# Development of *Lactobacillus agilis* La3 as a live vector for the delivery of vaccine antigens to the gastrointestinal tract of chickens

Benjamin Daniel Vezina
Bachelor of Science (Honours)

Supervisors: Rob Moore, Dena Lyras, Priscilla Johanesen

A thesis submitted for the degree of Doctor of Philosophy at
Monash University in 2018

Department of Microbiology
Monash University, Clayton, Victoria, Australia

# Copyright notice

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.

© Ben Vezina (2018).

# Abstract

The gastrointestinal tract of chickens can be colonised by a variety of pathogenic bacteria such as *Clostridium perfringens* that causes disease in chickens. Other bacteria, such as *Campylobacter jejuni* and *Salmonella*, are not pathogenic in chickens but are food-borne pathogens that can cause disease in humans. In the past, colonisation by these pathogens has been controlled by the use of in-feed antibiotics. However, the use of in-feed antibiotics is now reducing due to the increase of antibiotic resistant bacteria, and there is a need for alternative strategies to prevent or reduce colonisation by these bacterial pathogens. Vaccination, specifically live vaccination, is seen as the most suitable way of preventing colonisation of these organisms by inducing a mucosal immune response. *Lactobacillus agilis* La3 has been identified as a strong candidate for use as a live vector to deliver vaccine antigens to the gastrointestinal tract of chickens due to its strong colonisation and persistence within this niche, as well as its environment-modulating capacity that is mediated by lactic and short chain fatty acid production.

To characterise *L. agilis* La3, a combination of short and long read sequencing data was assembled into a complete circularised genome sequence. High throughput RNA sequencing was used to characterise the transcriptome of *L. agilis* La3 and genes that were constitutively expressed at high levels were identified. The presence of several highly expressed genes such as GAPDH, enolase and elongation factor Tu were identified. These could be involved in motility and low pH-specific adhesion, and these may play a role in the highly persistent characteristics of *L. agilis* La3 within the gastrointestinal tract. The predicted promoter regions from these genes were used to construct antigen expression cassettes on a plasmid system for integration into the chromosome of *L. agilis* La3. Construction of the pVEcXLa shuttle plasmid system allowed fast and selectable integration of genes into the *L. agilis* La3 chromosome *via* homologous recombination, which was completely stable within the chromosome over 100 generations. Promoters from the most abundant, constitutively transcribed genes were cloned and used to express the reporter gene CreiLOV from the pVEcXLa plasmid. A strong, positive relationship was found in *E. coli* between protein expression levels and transcript abundance and a moderately positive relationship between these factors was found in *L. agilis* La3. These promoters were then used to stably express *C. perfringens* toxin rNetB, a proven protective antigen, from the chromosome of appropriate recombinant *L. agilis* La3 strains without any loss of expression over 100 generations. Two recombinant strains, *L. agilis* La3::$P_{eft}$-r*netB* and *L. agilis* La3::$P_{cwah}$-r*net*B, show strong potential as live vaccines for use in chickens. Overall this work has made significant advances in the development of *L. agilis* La3 as a live vector suitable for the delivery of vaccine antigens to the gut of chickens.

# Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

**Signature:** _[redacted]___

**Print Name:** Benjamin Daniel Vezina

**Date:** 22/4/18

# Acknowledgements

This research was supported by an Australian Postgraduate Award (APA) scholarship (now Research Training Program) and the Poultry CRC, who provided a top-up stipend along with operating costs for the project. Without the Poultry CRC's support, this work could not have happened. I'd like to thank them for the funding, and I hope we've made something that they can benefit from in the future.

All the successes of this project haven't been solely because of me, and they wouldn't have happened if it weren't for the input and guidance of the network of scientists at CSIRO AAHL, Monash University, RMIT University and github. Thank you for taking the time out of your day for me. I am flattered, and here in this thesis I present the product of your time and company.

To my main supervisor Rob Moore, thank you for taking me on as a student, teaching me, explaining things and being patient. Your guidance and knowledge have been essential, not only in regards to the project but my development as a scientist. I don't really know if there is anyone else who has had such an impact on the way I think and go about analysing things. I've learned so much working for you and have been given so many opportunities I wouldn't otherwise have had. You definitely got the raw end of this deal. Thank you for high expectations and giving me the opportunity to learn and improve. It has been a pleasure working for you and being a part of the Moore Lab science squad, both at Lab 11 and the RMIT revival. It's quite special how you compile great people into excellent and productive labs and I really admire that. Being a part of your lab twice has been a privilege, and being there while you built up again after it was unjustly demolished was such a fantastic thing to be a part of. Moving to Geelong to work in your lab was the best decision I ever made.

To Dena Lyras for having me as a student despite not knowing anything about me. Thank you for supervising me, providing guidance and seeing avenues and perspectives not yet considered. Working in your lab was a great experience. Your contributions and perspectives have made this work better, and I thank you for that.

To Priscilla Johanesen for supervising me and guiding me through the most difficult times of my candidature. Thank you for the constant drop-in meetings and scientific discussion, especially during a period of time where nothing was working. Your contributions and insight have been invaluable to this work.

The foundations of not only the project, but my scientific development were built by the people of Lab 11 (CSIRO AAHL). Everyone taught by example and I look back fondly on all the expertise and company accumulated there.
Jake Lacey for his company and in-depth discussions about work. We worked through many things together, and it has made me better. There were many late nights spent at AAHL, shooting the shit and doing science.
Ben Wade for his company and mentoring. You're still my bearded angel of science.
Anthony Keyburn for being a funny and critical mentor who taught me a lot about the game.
Honglei Chen for the MiSeq sequencing run and assistance with library prep.
Karen Holden for her presence, knowledge, hijinks and being Lab Mum x 2.
Theo Allnutt for teaching me the basics of bioinformatics, which ended up being the one of the most important aspects of the project.
I am indebted to everyone in the Lyras and Rood labs at Monash. After starting and being established at CSIRO, I felt blind moving to a new lab mid-PhD, but everyone was so patient

and helpful, and were it not for the knowledge and community here, I would not have progressed as far as I have.
Plasmid Bay – Tom Watts, Tom Stent, Jacob Amy and Sarah Revitt-Mills for your late-night company, shenanigans and lack of groupthink which led to many interesting conversations, both scientific and non.
Vicki Adams, Jackie Cheung, Milena Awad, Pauline Howarth, Sherrie Young and all the others unmentioned, thank you for your expertise and assistance.

I'd also like to thank the Moore Lab RMIT edition. Maintaining work stability while packing up and moving again to Rob's new lab at RMIT was challenging, but everyone was incredibly accommodating, and the knowledge within the lab helped propel the work forward considerably.
Hao Van for her expertise, the sequencing run, dedication to the lab, and company. Hao also rescued the project by fixing the challenging *Lactobacillus* transformation I had attempted around 30 times, in a single go, allowing all future developments in this project to occur.
Mian Chee Gor, Emily Gan, Bron Campbell, Brydon Davidson, Khanh Chau and Tom McClean for their company, knowledge and troubleshooting discussions.

I'd also like to immortalize a few quotes which helped me throughout this work:

- "It's not gained through experience, it's lost during assumption" – Ben Wade on critical thinking, 2013
- "We're not repeating the experiment, we're account for biological variation" – Anthony Keyburn on how to justify redoing animal trials, 2014
- "If you do your job right, you'll put yourself out of work" – Someone at the Poultry CRC Ideas Exchange, 2015
- " 'Outcomes' are more frequent than 'results' " – Karen Holden on how to pitch milestones, 2016
- "Bring solutions, not problems" – Myself on remaining proactive about problem solving, also science pun, 2015
- "It's easier to design a house than it is to build one" – Tom McClean on why DNA work is simpler than protein work, 2017
- "The things that people overlook are often the biggest breakthroughs; otherwise you're just doing what everybody else is" – Rob Moore on following up on unexpected results, 2017

My time in research has been a clinic in failure, struggles and humblings, but because of that I've learned so much more than would have otherwise. It's been such a blessing.

Finally, to my future self:

1. Support staff are vital, be kind to them
2. This thesis is proof you need persistence more than intelligence
3. You can stop reading here, there's nothing more here for you

# Table of Contents

Copyright notice ..... ii
Abstract ..... iii
Acknowledgements ..... v
Table of Contents ..... vii
List of Tables ..... xi
List of Figures ..... xii
1. Chapter 1: Introduction ..... 13
1.1 The importance of poultry ..... 13
1.1.1 Commercial reality ..... 14
1.2 Colonisation of the broiler gastrointestinal tract ..... 15
1.2.1 First contact and colonisation ..... 16
1.2.2 *Salmonella* ..... 17
1.2.3 *Campylobacter jejuni* ..... 18
1.2.4 *Clostridium perfringens* ..... 20
1.3 Vaccination ..... 21
1.4 Live vaccine antigen delivery vectors ..... 22
1.4.1 Administration ..... 27
1.4.2 Therapeutic protein expression ..... 28
1.5 Stability of live vaccine and therapeutic protein expression ..... 30
1.5.1 Native, constitutive promoters ..... 30
1.5.2 Ribosomal binding sites ..... 31
1.5.3 Colonisation persistence of live vaccine ..... 31
1.6 *Lactobacillus agilis* ..... 33
1.6.1 Characteristics ..... 33
1.6.2 *L. agilis* as a chicken probiotic ..... 33
1.6.3 *L. agilis* La3 as live vector ..... 34
1.7 Aims ..... 35
2. Materials and methods ..... 36
2.1 Bacterial strains ..... 36
2.2 Growth and storage conditions ..... 37
2.2.1 *L. agilis* La3 ..... 37
2.2.2 *E. coli* JM109 ..... 37
2.3 Gel electrophoresis ..... 37
2.3.1 Agarose gel electrophoresis ..... 37
2.3.2 Formaldehyde gel electrophoresis ..... 38
2.4 RNA isolation from *L. agilis* La3 ..... 38

2.4.1 RNA for sequencing ..... 38
2.4.2 cDNA for quantitative polymerase chain reaction (qPCR) ..... 38
2.5 cDNA synthesis ..... 39
2.5.1 cDNA for RNA sequencing ..... 39
2.5.2 Single strand cDNA for qPCR ..... 39
2.6 Sequencing ..... 39
2.6.1 PacBio DNA sequencing ..... 39
2.6.2 MiSeq DNA sequencing ..... 40
2.6.3 MiSeq RNA sequencing ..... 40
2.7 Sequencing analysis ..... 40
2.7.1 DNA sequencing analysis ..... 40
2.7.2 RNA sequencing analysis ..... 42
2.7.3 Promoter identification ..... 42
2.7.4 Ribosomal binding sites ..... 42
2.8 DEPC-treated water ..... 43
2.9 Plasmid DNA isolation ..... 43
2.10 Polymerase chain reaction (PCR) ..... 46
2.10.1 Generation of PCR products for cloning and sequencing ..... 46
2.10.2 Diagnostic PCRs ..... 46
2.11 qPCR ..... 50
2.11.1 *rnetB* normalised transcription levels ..... 50
2.12 Restriction endonuclease digestions ..... 50
2.13 Phosphatase treatment ..... 51
2.14 Phosphorylation ..... 51
2.15 Purification of DNA from reaction mixtures ..... 51
2.15.1 PCR and enzyme reactions ..... 51
2.15.2 Gel purification ..... 51
2.16 Ligation ..... 52
2.17 Competent *E. coli* cells ..... 52
2.18 Transformation of *E. coli* ..... 52
2.19 Transformation of *L. agilis* La3 ..... 52
2.20 Chromosomal DNA isolation from *L. agilis* La3 ..... 53
2.21 Crude DNA isolation ..... 53
2.22 Induction and screening of chromosomal recombinants ..... 54
2.23 Stability assays ..... 54
2.23.1 Chromosomal insert stability ..... 54
2.23.2 Plasmid stability ..... 55
2.24 Whole cell lysate protein preparation ..... 56

2.25 Quantification of CreiLOV expression ..... 56
2.25.1 *E. coli* ..... 56
2.25.2 *L. agilis* La3 ..... 56
2.25.3 Statistical analysis ..... 57
2.26 Protein gel electrophoresis ..... 57
2.27 Western blot ..... 57
3. Chapter 3 Results: Characterisation of *Lactobacillus agilis* La3 ..... 59
3.1 Introduction ..... 59
3.2 Growth curve ..... 59
3.3 Genome sequencing and analysis of *L. agilis* La3 ..... 60
3.3.1 PacBio sequencing and assembly ..... 61
3.3.2 Annotation and composition of the *L. agilis* La3 genome ..... 62
3.3.3 SEED subsystem features ..... 63
3.3.4 Comparison to other *L. agilis* strains ..... 64
3.3.5 Bacteriocin mining ..... 66
3.3.6 Methylation ..... 66
3.3.7 Prophage gene identification ..... 67
3.3.8 Codon usage ..... 67
3.4 Transcriptome analysis of *L. agilis* La3 ..... 70
3.4.1 Promoter identification ..... 72
3.5 Chapter 3 Discussion: ..... 74
3.5.1 Genome composition ..... 74
3.5.2 Comparison to other *L. agilis* strains ..... 74
3.5.3 Methylation of the *L. agilis* La3 genome ..... 75
3.5.4 Prophages sequences within the genome ..... 75
3.5.5 Transcript abundance across the *L. agilis* La3 genome ..... 76
3.5.6 Potential probiotic effects ..... 77
4. Chapter 4 Results: Development of stable expression and integration systems for expression of recombinant proteins in *L. agilis* La3 ..... 79
4.1 Plasmid-based expression system ..... 79
4.1.1 Attempted transformation of pVEcLaSyn1 into *L. agilis* La3 ..... 80
4.2 Chromosomal insertion via double crossover homologous recombination ..... 81
4.2.1 Stability of chromosomal insert and plasmid ..... 84
4.3 Testing the predicted promoters ..... 91
4.3.1 EGFP as a reporter protein ..... 91
4.3.2 CreiLOV as reporter protein ..... 91
4.3.3 pVEcXLa expression system ..... 91
4.3.4 Strength of promoter compared with protein expression level ..... 93

4.4 Chapter 4 Discussion: ..... 95
4.4.1 Development of a stable expression system ..... 95
4.4.2 Identification of strong promoters by genome wide transcription analysis ..... 97
5. Chapter 5 Results: Therapeutic protein expression in *L. agilis* La3 ..... 99
5.1 Cloning of *rnetB* into the pVEcXLa-Promoter-*creiLOV* plasmids ..... 99
5.1.1 Codon-optimisation of *rnetB* ..... 99
5.2 rNetB expression in *E. coli* and *L. agilis* La3 ..... 100
5.2.1 *E. coli* ..... 100
5.2.2 *L. agilis* La3 ..... 100
5.3 Chapter 5 Discussion: ..... 108
5.3.1 rNetB expression ..... 108
6. Overall Discussion, Conclusion and Future Directions ..... 111
6.1 Comparison to other chicken live vaccines ..... 111
6.2 Improvements to pVEcXLa-$P_{eft}$-*rnetB* ..... 111
6.3 Conclusion and future directions ..... 113
7. Supplementary data ..... 115
7.1 Supplementary data 1 – RAST annotation ..... 115
7.2 Supplementary data 2 – pVEcXLa sequence ..... 423
7.3 Supplementary data 3 - pVEcXLa-$P_{cwah}$-*rnetB* ..... 426
7.4 Supplementary data 4 – pVEcXLa-$P_{cwah}$-*rnetB* ..... 430
8. References ..... 434

# List of Tables

Table 1: Summary of several live vaccines developed for animals. ..... 24
Table 2: List of bacterial strains used in this study ..... 36
Table 3: List of RBS sequences within *Lactobacillus* ..... 43
Table 4: List of plasmids and DNA used in this study ..... 44
Table 5: List of primers used in this study ..... 47
Table 6: Details and features of the three Southern blots used in this study ..... 55
Table 7: SEED subsystem results of RAST annotated genes ..... 62
Table 8: Comparison of *L. agilis* La3 to other sequenced *L. agilis* strains ..... 65
Table 9: Methylated motifs and their corresponding frequencies ..... 66
Table 10: List of Restriction modification genes found in *L. agilis* La3 ..... 67
Table 11: Prophage regions identified within *L. agilis* La3 ..... 67
Table 12: Codon usage within the *L. agilis* La3 chromosome. ..... 68
Table 13: Most frequently used codon for each amino acid ..... 70
Table 14: Predicted promoter sequences based on *Lactobacillus* consensus ..... 73
Table 15: Dunnett's multiple comparisons test between pVEcXLa plasmids. ..... 93
Table 16: Normalised qPCR results of *rnetB*. ..... 101
Table 17: Results of double crossover insert stability in *L. agilis* La3 chromosome ..... 104
Table 18: Summary of the pVEcXLa-Promoter-*rnetB* constructs. ..... 107

# List of Figures

Figure 1: Graph showing the yearly consumption of different meats in Australia.....14
Figure 2: Graphs showing global poultry production and consumption levels.....15
Figure 3: Graph showing incidence of salmonellosis per 100,000 population.....17
Figure 4: Graph showing incidence rate of campylobacteriosiss. .....19
Figure 5: Growth curve of *L. agilis* La3 under anaerobic conditions at 42°C. .....60
Figure 6: DNA gels showing the high molecular weight DNA .....61
Figure 7: Formaldehyde gel showing RNA extraction and purification.....71
Figure 8: Graph showing the TPM of the most abundant gene transcripts. .....72
Figure 9: Diagrams showing the chromosomal insertion site within *L. agilis* La3.....82
Figure 10: Graphical representation of double crossover in *L. agilis* La3 .....83
Figure 11: Gel electrophoresis image of double crossover diagnostic PCR .....84
Figure 12: Gel electrophoresis image of double crossover diagnostic PCR .....85
Figure 13: Graphical representation of probe binding site to *L. agilis* La3 chromosome....86
Figure 14: Gel and Southern blot of *Hin*dIII digested *L. agilis* La3 chromosome. .....87
Figure 15: Gel and Southern blot of *Hin*dIII/*Nco*I digested *L. agilis* La3 chromosome. .....88
Figure 16: Graphical representation of the probe binding site in pTRKH2-LCER. .....89
Figure 17: Image of gel and Southern blot of *Hin*dIII/*Nco*I digested *L. agilis* La3.....90
Figure 18: Graphical representation of pVEcXLa.....92
Figure 19: *E. coli* Relative fluorescent units vs transcripts per million reads.....94
Figure 20: *L. agilis* La3 relative fluorescent units vs transcripts per million reads.....94
Figure 21: Western blot of *E. coli* JM109 expressing rNetB.....100
Figure 22: Western blot of *L. agilis* La3 expressing rNetB .....101
Figure 23: Gel electrophoresis images of double crossover diagnostic PCR.....102
Figure 24: Western blot (B) of *L. agilis* La3 chromosomal recombinants. .....103
Figure 25: Gel electrophoresis images of double crossover diagnostic PCRs. .....104
Figure 26: Western blot of *L. agilis* La3::$P_{eft}$-*rnetB* colonies expressing rNetB .....105
Figure 27: Western blot of *L. agilis* La3::$P_{cwah}$-*rnetB* colonies expressing rNetB .....105
Figure 28: Western blot of *L. agilis* La3 colonies expressing rNetB .....106

# 1. Chapter 1: Introduction

The literature reviewed in this thesis covers the significance of poultry on a global scale. It then discusses the issues caused by the colonisation of the chicken gastrointestinal tract by several unwanted organisms including both chicken and human foodborne pathogens (Kaldhusdal *et al.* 2000; FSANZ 2005; EFSA 2010). The control of these organisms requires alternative treatments as they are currently controlled by in-feed antibiotics, which is undesirable due to increasing levels of antibiotic resistant bacteria and the worldwide imperative to reduce antibiotic use (Wagenaar *et al.* 2005). Live vaccination is an alternate method which can prevent colonisation of these unwanted organisms by priming the host immune system against them (de Zoete *et al.* 2007). Traditionally, live vaccines have been attenuated virulent bacteria, but there are several disadvantages with this approach, such as reduced colonisation of the vector within the host (De Cort *et al.* 2013), and the chance that the vector may reacquire virulence factors and become pathogenic (Salem *et al.* 2017). The alternative is to use live non-attenuated commensal bacteria to express antigens from these pathogenic organisms. This approach is advantageous as there is then little disease risk, the vaccine vector retains its colonisation capabilities, and it can also potentially provide probiotic effects depending on the strain (Lan *et al.* 2003). The strain *L. agilis* La3 has been identified as a rare chicken commensal isolate suitable for development as a live vaccine vector, due to its strong, lifetime colonisation persistence within chickens, across a variety of diets and microflora compositions (Stephenson *et al.* 2011).

## 1.1 The importance of poultry

Poultry, specifically broiler chickens, are an increasingly valuable commodity worldwide. With a rapidly expanding human population, projected to reach roughly 9.3 to 10.2 billion in 2050 (UN 2013; USCB 2013), economical and sustainable nutrient sources are vital. Poultry is considered to be an appropriate protein source to accommodate the rapidly expanding global population due to its many favourable characteristics including its efficient feed conversion ratio (FCR). FCR is an efficiency rating which indicates how many kilograms of food is required for one kilogram of weight gain. Broiler FCR ranges from ~1.3 to 2, with most studies and commercial settings converging on ~1.5 (Archer *et al.* 2014; Costa *et al.* 2014; Giannenas *et al.* 2014; Tanweer *et al.* 2014; Walk *et al.* 2014; Xu *et al.* 2014). In comparison, beef FCRs range from 4.5 to 7.5 (Shike 2013), while pigs range from 3.8 to 4.5 (DoAaF 2012). Chickens also have efficient water usage, with a worldwide weighted average of 4,325 $m^3$ / ton. Compared to other terrestrial protein sources such as goat (5,521 $m^3$ / ton), pork (5,988 $m^3$ / ton), lamb (10,412 $m^3$ / ton) and beef (15,415 $m^3$ / ton), poultry has a significantly smaller water footprint (Mekonnen *et al.* 2012). Chickens have a rapid turnover rate from 35 to 42 days to market weight under industrial conditions with the latest genetic stock that is currently used (Khosravinia 2009; Bibby 2014) which is unmatched by other terrestrial animals. Chickens have the capacity to tolerate high density farming and large-scale production, to an extent that is unattainable in other agricultural meat-protein producing industries such as beef, lamb, goat or pork. Mass production and production practicality on the scale of poultry is inappropriate to other agricultural protein producing industries such as beef or lamb (Mulder *et al.* 2013; Archer *et al.* 2014; Giannenas *et al.* 2014). Over the past 40 years, the global poultry industry has consequently expanded at a rapid pace, with poultry surpassing beef and veal consumption in 2007 in Australia (Figure 1 and Figure 2) (Mulder *et al.* 2013), and the global poultry trade also on a steady increase. This can be attributed to its low cost per kilogram due to the aforementioned characteristics

and advantages over other livestock protein sources. These factors have also been reflected in the Australian market (Figure 1), where poultry meat is now the most highly consumed agricultural protein source, with Australians eating an estimated 47 kg of chicken meat per capita in 2017 (ABARES 2013; ABARES 2016). Importantly, the growing poultry industry is experiencing challenges which, if not addressed will magnify with continued expansion. These challenges include colonisation of gastrointestinal tracts by unwanted organisms which are capable of causing disease in chickens and foodborne illnesses in humans. This will be the focus of the discussion below.

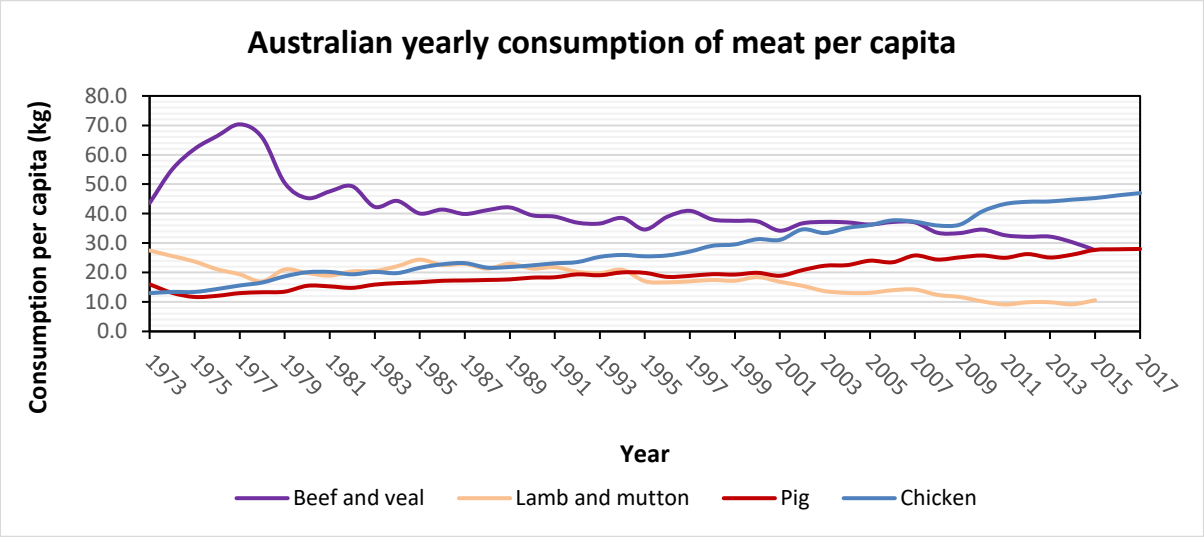


Figure 1: Graph showing the yearly consumption of different meats in Australia over the past 40 years (ABARES 2013; ABARES 2016). Poultry consumption has been steadily increasing, surpassing beef/veal in 2007.

### 1.1.1 Commercial reality

In a commercial setting, once broiler eggs are laid, they undergo quality control to remove perforated, cracked or unclean eggs. The remaining eggs, while still warm, are then fumigated with formaldehyde in a well-ventilated room (~35 ppm) to remove as many surface dwelling bacteria as possible. The outer egg surface is thought to be one of the first major contact points between the chick and microbes; therefore this rigorous selection and cleaning process is used to reduce the bacterial load of foodborne pathogens such as *Salmonella* as much as possible (Bibby 2014). If gas-producing bacteria pass through the shell and begin colonising inside the egg, they will be incubated along with the chick at 37°C, which is an optimal growth condition for many bacteria. The egg will then explode due to increased pressure, effectively killing the chick. This phenomenon is known as a 'banger' (Bibby 2014). This event can then cause a banger chain reaction through the whole flock of incubating eggs due to the highly colonised debris landing on other eggs. The poultry industry has found the most effective way of reducing this effect is to use the microbial control methods described above (Bibby 2014).

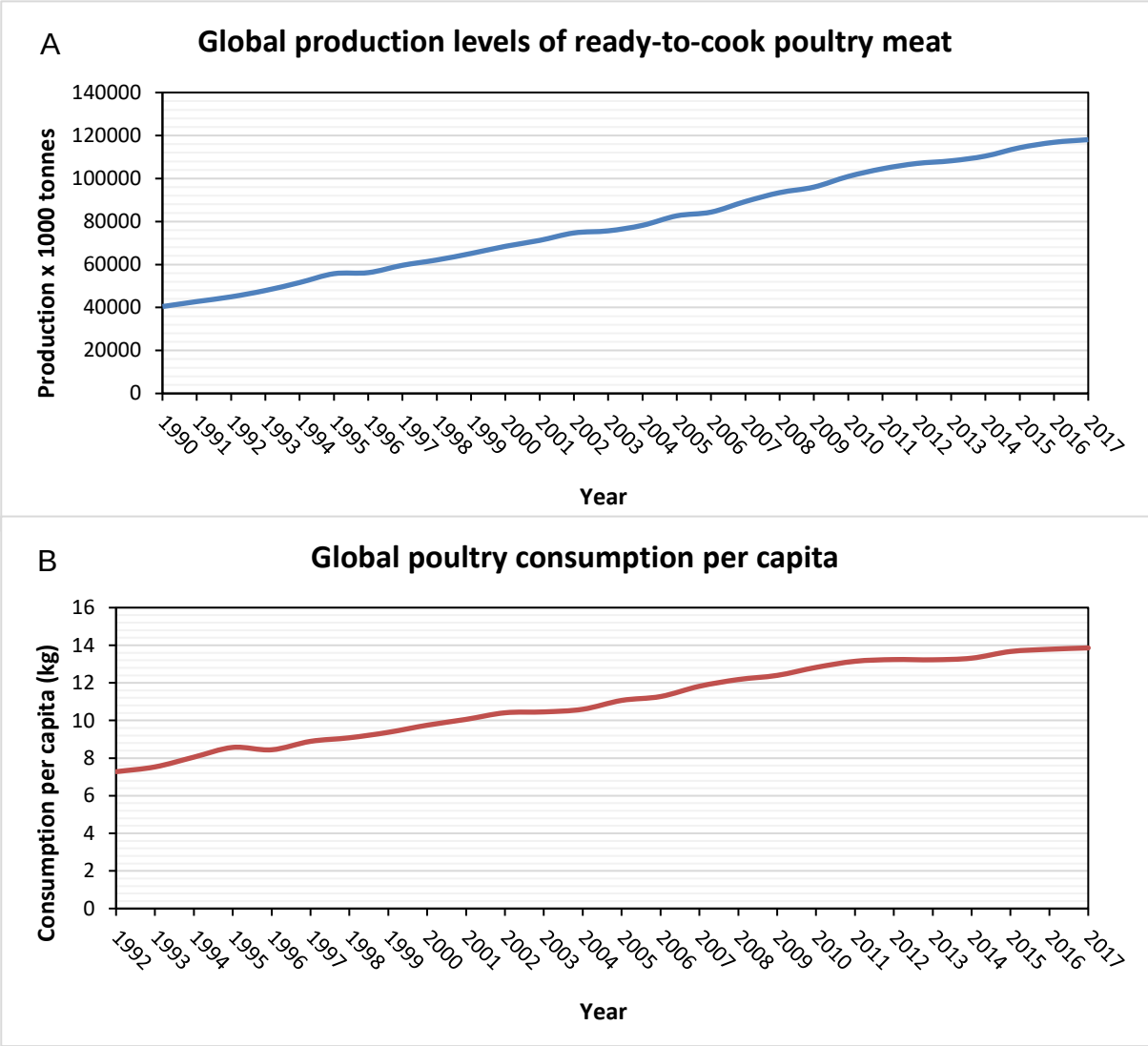


Figure 2: Graphs showing global poultry production and consumption levels over time (OECD/FAO 2017).

A: Global production of ready-to-cook poultry meat shows a steady incline over the past 17 years, currently around 120,000,000 tonnes in 2017.

B: Global consumption of poultry per capita has also steadily grown over the past decade, currently around 14 kg. These statistics indicate the increasing dietary intake of poultry within humans worldwide, along with the matching production levels.

## 1.2 Colonisation of the broiler gastrointestinal tract

In chickens colonisation of the gastrointestinal tract and the corresponding mucosal layer by both bacterial pathogens and healthy microbiota is contact and exposure based (Giovanardi *et al.* 2005; Crhanova *et al.* 2011). Therefore, microbes present in the environment of the newly hatched chicken have a strong chance of colonising the near-empty niches within the gastrointestinal tract.

In a commercial setting, hatched chicks have no contact with their mother or nesting environment, as would occur in the wild. Because of this, chicks cannot be influenced by horizontal transfer from older broiler chickens with well-established gastrointestinal tract microbiota. Eggs are hatched in batched flocks meaning that no direct contact between flocks occurs (Bibby 2014). Different batches of broiler chickens are housed in the same barn (which may or may not have been decontaminated since it held the previous flock) meaning microbial crossover between flocks can vary. (Bibby 2014).

### 1.2.1 First contact and colonisation

The uncolonized chicken gastrointestinal tract allows opportunistic, commensal or potentially pathogenic microbes to take up residence and to gain a foothold, potentially causing disease or posing other health risks (Giovanardi *et al.* 2005; Teirlynck *et al.* 2011; Stanley *et al.* 2013). A healthy microbial community within the gastrointestinal tract has been shown to be part of the first line of defence against disease-causing or undesirable colonisers via competitive exclusion, by occupying the same niche (Nurmi *et al.* 1973; Rantala *et al.* 1973; Methner *et al.* 1999; Crhanova *et al.* 2011), or by releasing antibacterial compounds (Shin *et al.* 2008; Messaoudi *et al.* 2012).

By using a variety of molecular techniques, microbial ecology within the chicken gastrointestinal tract has been shown to diversify in the first few weeks post-hatching, and then stabilises (Kimura *et al.* 1976; Crhanova *et al.* 2011). The composition of the microbiota also differs between sections of the gastrointestinal tract, specifically the crop, duodenum, ileum and each cecum, which contain different bacterial communities (van der Wielen *et al.* 2002; Yeoman *et al.* 2012). Disease and disease-causing agents, such as *Eimeria tenella* (Kimura *et al.* 1976) or *Clostridium perfringens* (Helmboldt *et al.* 1971; Wijewanta *et al.* 1971; Lacey *et al.* 2016), can also have an impact on the microbial composition of the gastrointestinal tract of broiler chickens. Although Kimura *et al.* used the dated technique of culturing (which is biased against non-culturable bacteria) they showed disease in the chicken can impact niches within the gastrointestinal tract, and therefore displace or indirectly kill native gastrointestinal tract strains, as well as allow opportunistic and favourable growth of certain genera such as the clostridia and enterobacteriaceae (Kimura *et al.* 1976).

The gastrointestinal tract of broiler chickens can be colonised by certain undesirable bacteria which are pathogenic to chickens (such as *Clostridium perfringens*) (Timbermont *et al.* 2011), or are foodborne pathogens carried by chickens that can infect and cause disease in humans, such as *Salmonella* and *Campylobacter jejuni* (Van Roekel 1964; Hood *et al.* 1988). Although this project does not directly involve *Salmonella* and *C. jejuni*, they are discussed below as there are no protective commercial vaccines currently available for these gut mucosal dwelling organisms. Thus, these organisms represent other important vaccine targets, which could be addressed by the live vaccine approach utilised in this work.

### 1.2.2 *Salmonella*

#### 1.2.2.1 Prevalence

Chickens are a natural reservoir of *Salmonella enterica* subspecies such as *S. enterica* Enteritidis (*S.* Enteritidis) which cause foodborne illness in humans (Van Roekel 1964; Kaufmann *et al.* 1968; Nurmi *et al.* 1973). *S. enterica* outbreaks have occurred in direct correlation with the mass production of poultry. The human incidence of salmonellosis is on the rise (Figure 3), and for many years poultry and egg consumption has been strongly linked to outbreaks (Van Roekel 1964; Kaufmann *et al.* 1968; FSANZ 2005). In Australia, from 1995 to 2002, it was reported that 498 *Salmonella* outbreaks occurred, and 18 were linked to poultry meat consumption (FSANZ 2005). The authors of this report also stated that under-reporting is common with foodborne illness. A study from 1995 to 2000 showed that poultry accounted for 25 out of 75 (33%) *Salmonella* outbreaks in Australia (Dalton *et al.* 2004).
*Salmonella* prevalence in poultry and poultry meat has been demonstrated consistently, where *S.* Enteritidis and related serovars have been detected in 5.26% of European meat samples and 3.6% of flocks (Catarame *et al.* 2006; EFSA 2010). In Australia, 48% of all *Salmonella* outbreaks were caused by *Salmonella* Typhimurium (OzFoodNetWorkingGroup 2015). These statistics clearly indicate that the *Salmonella* load within poultry is an issue of significant importance in human health.

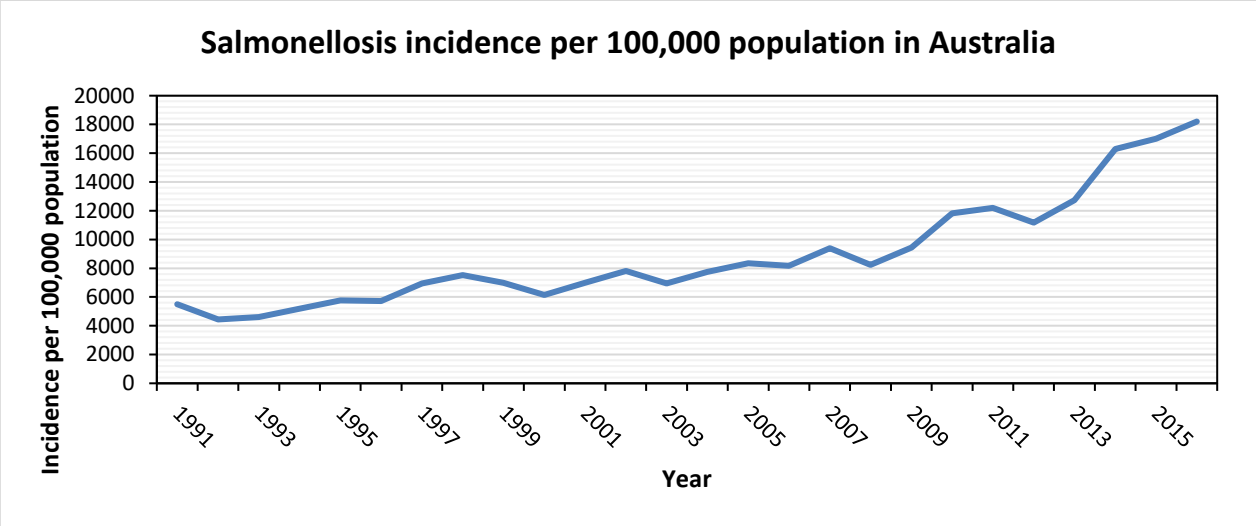


Figure 3: Graph showing incidence of salmonellosis per 100,000 population in Australia (DoH 2017). Incidence has been on the rise over the past 15 years.

#### 1.2.2.2 Host impact

*S. enterica* subspecies express many virulence factors such as toxins, flagella and adhesion factors, which are encoded in clusters on mobile genetic elements such as plasmids and across the chromosome in pathogenicity islands (Salem *et al.* 2017). *Salmonella* subspecies can have a wide variety of endotoxins (lipid A) and exotoxins such as cytotoxins and heat

labile *Salmonella* enterotoxin (van Asten *et al.* 2005). All of these contribute to human disease such as gastroenteritis and enteric fever.

*Salmonella* is capable of colonising the gastrointestinal tract of chickens and invades the liver and spleen, regardless of the composition of the gastrointestinal tract microbiota (Crhanova *et al.* 2011). The main factor in colonisation is age and, therefore, the maturation of the immune system of the chicken. This has an inverse relationship to *S.* Enteritidis loads in chickens. This colonisation begins around 10 days post-inoculation and then drops until market date (42 days), at which time it usually drops below detectable levels (Crhanova *et al.* 2011). The chicken immune system reacts to *S.* Enteritidis in a variety of ways, which is dependent on age. Day old chicks inoculated with *S.* Enteritidis express significantly higher levels of interleukins IL-1β, IL-8, IL-17, IL-18, IL-22, interferon-γ and inducible nitric oxide synthase (iNOS) within 3 days, compared to uninoculated controls (Crhanova *et al.* 2011). Chickens inoculated at 16 days express significantly higher IL-1β, IL-8, IL-17, IL-18, IL-22 IFN-γ, tumor necrosis factor-α (TNFα) and iNOS within 10 days of inoculation. IL-17 levels are increased by 50-fold, IL-22 by 47-fold and a significant reduction in *S.* Enteritidis loads occurs, indicating that the age of the bird at exposure time is an important aspect of colonisation (Crhanova *et al.* 2011). Chickens raise an immune response against *Salmonella*, however, this response is not sufficient to reduce the bacterial load below the infective dose for humans, as human incidence still occurs.

#### 1.2.2.3 Current control methods in development

There have been several methods experimentally evaluated for their effectiveness in the control of *S. enterica* colonisation in chickens, with varying success. Competitive exclusion or colonisation inhibition using attenuated *S.* Enteritidis mutants have been shown to significantly reduce the colonisation of challenged pathogenic *Salmonella* Enteritidis when orally administered on day one post-hatch, when compared to controls (Bohez *et al.* 2007; De Cort *et al.* 2013; De Cort *et al.* 2015). However, there are several key weaknesses to this approach, specifically, attenuated strains are required to be a similar serovar to inhibit the unwanted wild type *Salmonella* strain so as to effectively prevent host colonisation via competitive exclusion and antibody production (De Cort *et al.* 2013). For every new and existing serovar, new mutants would need to be generated and tested to provide protection. Attenuated *Salmonella* strains are transient (De Cort *et al.* 2013; De Cort *et al.* 2015), if flocks are colonised by a *Salmonella* serovar after the attenuated strain has passed through, their colonisation may be uninhibited. This approach therefore does not provide the broad protective effect required for an effective intervention for the poultry industry.

### 1.2.3 *Campylobacter jejuni*

#### 1.2.3.1 Prevalence

By 1988, it was known that *C. jejuni* was a major cause of enteritis in humans and was also highly prevalent among retail broiler chicken carcasses, and that infection with this bacterium represented a global issue (Hood *et al.* 1988). *C. jejuni* usually starts colonizing birds at between 1 to 3 weeks of age (Beery *et al.* 1988; Stern 1992; Newell *et al.* 2000; Sahin *et al.* 2003) but has also been found in 7 to 19 day old embryos (Hiett *et al.* 2014). It lives within the gastrointestinal tract and colonises the mucosa of the proximal small intestine at $10^2$ to $10^5$ CFU / g, the distal small intestine at $10^2$ to $10^5$ CFU / g, the cecum at $10^2$ to $10^7$

CFU / g, the large intestine and cloaca at $10^2$ to $10^5$ CFU / g, and the faeces at $10^2$ to $10^5$ CFU / g. The highest prevalence is found within the caecum (Beery *et al.* 1988). *C. jejuni* is shed from infected animals in high numbers, allowing successful colonisation of whole flocks and also of processed poultry meat via faecal contamination (Sahin *et al.* 2001). The design of slaughter houses makes cross contamination of chicken carcasses almost unavoidable (Corry *et al.* 2001). One study looked at the incidence of *C. jejuni* at age of slaughter and found that 7 out of 11 flocks of broiler chickens were positive for *C. jejuni*, with 79% prevalence detected within positive flocks (Beery *et al.* 1988). In a 2008 study, 10 chickens from each of 10,132 batches were sampled from 561 slaughterhouses in 26 European Union Member States (EFSA 2010). A total of 71.2% of the batches, and 75.8% of the total chicken carcasses, were found to be *Campylobacter* positive. *C. jejuni* was found in 60.8% of all *Campylobacter* positive flocks and *Campylobacter coli* was found in 41.5%. 40.8% of all carcasses sampled contained >100 CFU / g *C. jejuni*, and 5.8% contained more than 10,000 CFU / g *C. jejuni* (EFSA 2010). The minimum infective dose for *C. jejuni* in humans is known to be 500 to 800 individual cells (Robinson 1981; Black *et al.* 1988), and campylobacteriosis incidence rates have risen slowly over time in Australia (Figure 4). This indicates that *C. jejuni* is a prominent issue for human health that arises from the colonisation of poultry.

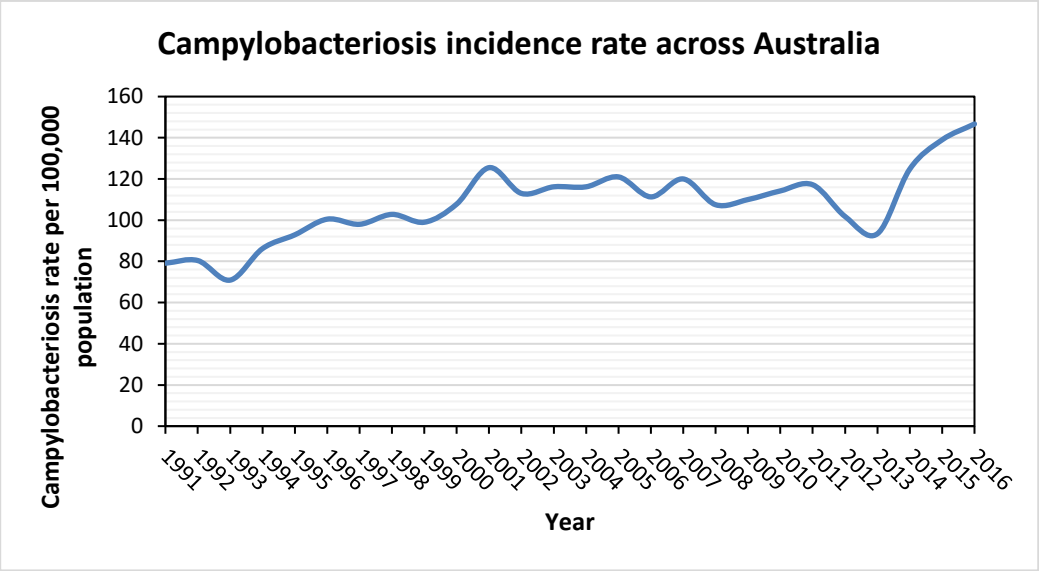


Figure 4: Graph showing incidence rate of campylobacteriosis over time (DoH 2017). Incidence has been trending upwards over the past 23 years.

#### 1.2.3.2 Host impact

There is evidence that *C. jejuni* is not a commensal organism and that colonisation by this organism negatively impacts chicken welfare and adversely impacts the quality of life of infected chickens (Humphrey *et al.* 2014). It has been shown that significant immune responses are mounted against *C. jejuni* and that inflammation occurs upon infection, leading to gastrointestinal tract mucosal damage and diarrhoea, which leads to wet litter and leg health issues (Humphrey *et al.* 2014). Chickens mount immune responses against *C. jejuni*, not only in the form of cell-mediated responses, but also in raising anti-*C. jejuni* immunoglobulin Y (IgY), immunoglobulin A (IgA) and immunoglobulin M (IgM) antibodies.

Vertical transfer of anti-*C. jejuni* maternal antibodies also occurs from mothers to eggs (Kowalczyk *et al.* 1985; Myszewski *et al.* 1990; Cawthraw *et al.* 1994; Sahin *et al.* 2001; Sahin *et al.* 2003; Humphrey *et al.* 2014). *C. jejuni* colonisation results in higher IgY and IgA levels at between three to four weeks of age. Towards the eight-week mark, IgA levels drop off. Mucosal IgA and IgY levels within colonised broiler chickens are significantly different between day one and nine weeks post hatch, as anti-*C. jejuni* antibody levels drop over time (Myszewski *et al.* 1990; Cawthraw *et al.* 1994; Rice *et al.* 1997). Interestingly, even though antibodies are raised, *C. jejuni* is persistent within the chicken even at 9 weeks old (Myszewski *et al.* 1990; Cawthraw *et al.* 1994; Rice *et al.* 1997). This is likely to be due to the combination of the ability of *C. jejuni* to evade the immune system, sub-optimal IgA production by the chicken, and antibodies being produced against non-optimal targets instead of vital colonisation factors (Myszewski *et al.* 1990; Cawthraw *et al.* 1994). Anti-flagellin antibodies (flagellin A, flagellin B) are prominent, indicating that the immune system partially targets *C. jejuni* surface antigens and virulence factors and, secondly, that flagellar proteins are relatively immunogenic (Cullender *et al.* 2013). Flagellar proteins are potentially suitable candidate antigens for vaccine development, although epitopes must be surface bound to have a greater chance at inducing an immune response (Power *et al.* 1994).

Mucosal immune responses, specifically IgA production, in unvaccinated birds are not strong enough for host chickens to clear the bacterial load, and the magnitude at which *C. jejuni* colonises the chicken gastrointestinal tract and replicates has been shown to increase from $4.5 \times 10^4$ CFU/g to $9 \times 10^8$ CFU/g after five days in cecal contents. After nine weeks it remains at approximately $7 \times 10^6$ CFU/g (Cawthraw *et al.* 1994), indicating how persistent and well-adapted *C. jejuni* isolates are to the chicken gastrointestinal tract. Of relevance to the work described in this thesis, the breakdown of maternal antibodies coincides with *C. jejuni* colonisation, as it is well documented that *C. jejuni* colonisation occurs within 1 to 3 weeks of age (Beery *et al.* 1988; Stern 1992; Newell *et al.* 2000; Sahin *et al.* 2003). This time period provides a good opportunity for immunologically unopposed opportunistic microorganisms to colonise the gastrointestinal tract.

#### 1.2.3.3 Current control methods

*C. jejuni* is currently controlled by the administration of in-feed antibiotics (Melo *et al.* 2013), although there are several alternative approaches currently being investigated such as phage (Wagenaar *et al.* 2005) and bacteriocin therapies (Messaoudi *et al.* 2012).

### 1.2.4 *Clostridium perfringens*

#### 1.2.4.1 Prevalence

*Clostridium perfringens* is a significant problem in the poultry industry, causing the gastrointestinal disease necrotic enteritis (NE) in broiler chickens, usually at two to four weeks of age (Williams 2005; Timbermont *et al.* 2011). Two forms of disease are recognised, defined as clinical and subclinical (Timbermont *et al.* 2011). The clinical form can kill 10 to 40% of a flock and cause a variety of conditions in birds such as paralysis, mucosal necrosis, liver damage and ballooned intestines within 1 to 14 days if left untreated (Helmboldt *et al.* 1971; Wijewanta *et al.* 1971; Riddell *et al.* 1992). The subclinical form of NE does not present with such obvious outward symptoms aside from wet litter, for which poultry managers will almost immediately start antibiotic therapy to resolve potential NE outbreaks (Williams

2005). The lesions and intestinal damage caused by *C. perfringens* result in internal discomfort and reduced digestion. This decreases the conversion of feed to muscle mass (Kaldhusdal *et al.* 2001; Mcdevitt *et al.* 2006). This loss of productivity is the greatest source of economic cost (Wade *et al.* 2016). NE is highly prevalent, with 1 to 40% of North American commercial flocks infected at any one time (Kaldhusdal *et al.* 2000), and is therefore a serious welfare and economic issue, estimated to cost the global poultry industry US$5 to 6 billion per annum (Wade *et al.* 2016).

#### 1.2.4.2 Host impact

NE is caused by the expression of an extracellular toxin, NetB, by infecting *C. perfringens* strains, which damages the gastrointestinal tract, significantly reducing absorption and digestion, and therefore productivity of the broiler chicken. The NetB toxin has been demonstrated to be an essential virulence factor in NE (Keyburn *et al.* 2008; Mot *et al.* 2014) and not all colonising *C. perfringens* cause NE (Keyburn *et al.* 2008; Lacey *et al.* 2016; Mishra *et al.* 2017).

Chickens produce humoral IgA and IgY immune responses against secreted proteins such as α-toxin from virulent *C. perfringens* strains (Kulkarni *et al.* 2006; Lee *et al.* 2012; Mishra *et al.* 2017). However, this response does not result in protection against disease. Chickens raised in a commercial setting that have NE or gangrenous dermatitis (also caused by *C. perfringens*) have lower antibody levels than non-afflicted chickens raised in the same environment, indicating that antibody function is sufficient to protect certain exposed chickens to virulent *C. perfringens* strains (Lee *et al.* 2012). The secreted α-toxin has been historically used as a proxy for specific anti-*C. perfringens* antibodies. It has been shown that the age of the broiler breeder hen is directly proportional to anti-α-toxin antibody titres of day old chicks, and that this results in a significantly lower mortality for these chicks at between 2 to 3 weeks of age (Heier *et al.* 2001), indicating that the older a breeder is, the better their progeny are protected. Older chickens tend to decrease egg-laying and increase IgY antibody production as they age, which leads to higher antibody titres in eggs and, therefore, greater protection against disease. Breakdown of maternal antibodies at between 2 to 3 weeks of age also coincides with NE outbreaks (Naqi *et al.* 1983; Heier *et al.* 2001).

#### 1.2.4.3 Current control methods

Traditional subcutaneous vaccination against *C. perfringens* has not been successful in reducing NE as no commercial vaccines are available. The use of antibiotics is discouraged as a long term solution and is outright banned in the European Union due to growing antibiotic resistance concerns (Mot *et al.* 2014). There are currently several approaches to vaccination against *C. perfringens* which have shown partial protection, such as the use of live attenuated vaccines (Kulkarni *et al.* 2010), live vaccines (Mishra *et al.* 2017), toxoid vaccines (Saleh *et al.* 2011; Keyburn *et al.* 2013) and maternal antibody subunit vaccines (Keyburn *et al.* 2013).

## 1.3 Vaccination

Vaccination induces an immune response against specific non-self-material. Modelling has predicted that a 100-fold reduction in *Campylobacter* load within broiler chickens may result

in a 30-fold reduction in the incidence of human cases of foodborne campylobacteriosis (Rosenquist *et al.* 2003). Antibiotics have been used to reduce the carriage of *C. jejuni* in birds but due to the high prevalence of antibiotic resistant strains, antibiotic use in animals is being, or has been, phased out as a control method (Wagenaar *et al.* 2005; Dhillon *et al.* 2006). Alternative solutions are therefore needed and are being investigated. Vaccines are a tool used to prevent or mitigate future diseases in organisms by priming the immune system against biological disease-causing agents. This is generally done by introducing antigens from the surface of the disease causing agent, or an attenuated version of the disease causing agent itself (De Cort *et al.* 2013), into a host to induce a protective response prior to infection. Vaccination primes the immune system for the disease before initial exposure, thus preventing colonisation or disease, or reducing the severity of symptoms (de Zoete *et al.* 2007). Vaccines have been especially helpful for circumventing the evolution of antibiotic resistance within bacteria through the reduction of antibiotic use, which could leave many disease-causing bacteria unchecked due to ineffective antibiotic treatments (Jacobs-Reitsma 1997; Giovanardi *et al.* 2005; Teirlynck *et al.* 2011; Millman *et al.* 2013).

Many vaccines have been used in livestock industries and have proved largely successful in reducing animal illnesses (Meeusen *et al.* 2007). However, it has been found in commercial broiler chickens that traditional systemic vaccinations against certain bacteria such as *C. jejuni*, *C. perfringens*, and *Salmonella* are not effective as these are gastrointestinal tract mucosal dwelling organisms. For vaccination to be effective, a mucosal immune response, involving IgA antibodies, must be elicited against the introduced antigens. Live vaccination has the greatest potential to induce an IgA-specific mucosal immune response rather than systemic IgY and IgM responses which traditional vaccines induce (de Zoete *et al.* 2007). Induction of serum IgY also plays a role in protection as intestinal permeability of young chicks is high within the first week post-hatching. This allows movement of maternal and self-IgY into the intestine, which partly confers mucosal protection within the gastrointestinal tract (Sahin *et al.* 2001).

Live vaccination is a suitable antigen delivery system capable of achieving a mucosal immune response (Bermúdez-Humarán 2009; Layton *et al.* 2011). This type of vaccination has the potential to be exploited for both the animal's wellbeing and health and for the prevention of downstream disease in humans. For example, human food poisoning is most commonly associated with bacterial loads within animal meats and the presence of *Campylobacter* species within undercooked and improperly prepared poultry has been shown to be the major cause of food poisoning in humans (Abelson *et al.* 2006). A reduction in poultry loads is likely to consequently reduce human disease. Protection of humans from food poisoning has not yet been demonstrated. Although there have been several approaches such as phage therapy (Wagenaar *et al.* 2005), competitive exclusion (Schoeni *et al.* 1994; Methner *et al.* 1999) and bacteriocin therapy (Shin *et al.* 2008; Messaoudi *et al.* 2012) that have been utilised to reduce colonising loads of these bacteria in broiler chickens it is thought that the most effective route to reduce pathogen carriage is likely to be mucosally delivered live vaccination.

## 1.4 Live vaccine antigen delivery vectors

Live delivery vectors are genetically modified bacteria used as antigen delivery systems (Cardenas *et al.* 1993; De Cort *et al.* 2013). They are a subtype of paratransgenic organisms (*para* meaning 'at or to one side of', and *transgenic* meaning 'genetically modified organism'). This means the target organism is not genetically modified, but an organism

within it (the live delivery vector) is. The live delivery vector comes into contact with the environment, such as the gut mucosal layer, and express antigens from target organisms, either in the cytoplasm, surface-anchored or excreted in a soluble form (Totte *et al.* 1999; Pawelec *et al.* 2000). The introduction of antigens into a host can prime the immune system, reducing unwanted organism load and improving resistance to pathogen colonisation (Cardenas *et al.* 1993; De Cort *et al.* 2013). Table 1 shows a selection of experimental live vaccines and their different properties and delivery strategies.

Table 1: Summary of several live vaccines developed for animals. The main characteristics of each live vaccine are detailed.

| Live vaccine | Strain origin | Host | Vaccine native to host (Y/N) | Attenuated (Y/N) | Dosage | Oral/nasal route/other | Host age when dosed | Sample size | Target organism | Disease caused | Antigen/therapeutic protein expressed | Codon Optimized | Promoter type | Excreted/membrane bound/cytoplasm | Chromosomal or plasmid-based | Outcome | Notes | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Salmonella enterica* serovar Typhimurium | Horse strain passaged through chicken (Curtiss III *et al.* 1996) | Broiler chickens | No | Yes | 100 µL (1.2 × $10^9$ CFUs) twice over 10 days | Oral gavage | 1 day | n = 15 | *C. perfringens* | Necrotic enteritis | Alpha toxoid and hypothetical protein | Yes | Constitutive via *trc* promoter | Periplasm and supernatant via β-lactamase type II signal sequence | Plasmid pYA3493 | Weak protection via IgA and IgY immune response | | (Kulkarni *et al.* 2010) |
| *Lactobacillus plantarum* | Yoghurt | Pigs | No | No | $10^9$ CFU / g per kg body weight | Oral in feed | 5 – 7 weeks | n = 15 | Classical Swine Fever Virus | Classical swine fever | E2 and Thymosin α-1 | No | Inducible via T7 | Cell wall anchored via M6 | Plasmid pYG301 | Strong protection via IgA and IgG antibody production | Thymosin α-1 is a human immune modifier protein | (Xu *et al.* 2015) |
| *Lactobacillus casei* | Dairy products | Mice | No | No | 100 µL ($10^9$ CFUs / mL) six times over 2 weeks | Subcutaneous immunization | 6 – 8 weeks | n = 5 | Human Papillomavirus type 16 | Cervical cancer | L1 | No | Lactose-inducible plac promoter | Cytoplasm | Plasmid pIAlac | IgG antibody production | | (Aires *et al.* 2006) |
| *L. casei* | Dairy products | Mice | No | No | 200 µL (1 x $10^9$ CFUs / mL) over 3 consecutive days on days 0, 16, and 37 | Gastric gavage | 6 – 8 weeks | n = 10 | *C. perfringens* | Enterotoxemia | β-toxoid and dendritic cell-targeting protein | No | Constitutive via P1 promoter | Cell wall anchored via usp45 secretion signal and *spaX* wall anchor | Plasmid pT1NX | Strong protection via IgG and IgA antibodies and IFNγ production | | (Alimolaeia *et al.* 2017) |
| *Lactobacillus acidophilus* | Human | Mice | No | No | 100 µL (1 x $10^9$ CFUs) six times from day 0 – 24. | Oral gavage | 6 weeks | n = 15 | *Escherichia coli* O157:H7 | Diarrhea and haemorrhagic colitis | *espA*-Tir-M | No | Constitutive | Supernatant via espA-Tir-M | Plasmid pMG36e | Strong protection via IgG and IgA antibodies, and IL-4, IL-10 and IFNγ production | | (Lin *et al.* 2017) |
| *E. coli* serogroup O78:K80 strain EC34195 (Δ*aroA*) | Chicken | Chicken | Yes | Yes | Coarse spray (3.2 x $10^6$ CFU/dose in 1 mL) on day 1 then in drinking water (1.5 x $10^6$ CFU/dose in 5 mL) on day 7 | Coarse aerosol spray then drinking water | 1 day | n = 30 | Avian pathogenic *E. coli* O78 | Avian colibacillosis | Δ*aroA* | N/A | N/A | N/A | N/A | Strong protection: 0% mortality, 36.7% cumulative colibacillosis compared to control: 28.1% mortality and 96.9% cumulative colibacillosis | Poulvac® *E. coli* vaccine (Zoetis) | (Ragione *et al.* 2013) |
| *E. coli* serogroup O78:K80 strain EC34195 (Δ*aroA*) | Chicken | Turkey | No | Yes | Coarse spray (10 mL of $10^7$ CFU / mL) (10 birds / 0.25 $m^2$) on day 1 then on day 7 in drinking water (1 x $10^7$ CFU / mL) | Coarse aerosol spray then drinking water | 1 day | n = 10 | Avian pathogenic *E. coli* O78 | Avian colibacillosis | Δ*aroA* | N/A | N/A | N/A | N/A | Significant reduction in colonization by E. coli serogroup O78:K80 in lungs, liver, spleen and kidney. At 42 days, IgY-soluble antigens were significantly higher than control. | Poulvac® *E. coli* vaccine (Zoetis) | (Sadeyen *et al.* 2015) |
| *Salmonella enterica* serovar Enteritidis phage type 13A S NCTC 12023/ATCC 14028 (Δ*aroA* Δ*htrA* Cj0113- CD154) | Chicken (heart and liver) | Chicken | Yes | Yes | 250 µL ($10^8$ CFU / mL) on day 1 | Oral gavage | 1 day | n = 55 | *C. jejuni* PHLCJ1 | Campylobacteriosis in humans | Cj0113 (Omp18/CjaD)-CD154 fusion. ΔaroA ΔhtrA | Yes | *lamB* native promoter | Cell wall anchored via *lamB* gene | Chromosomal integration | Significant increase in anti-Campylobacter IgY and Iga antibody production compared to saline control group. 4-log reduction of C. jejuni (below detectable levels) at 32 days. | CD154 (CD40 ligand) is part of the tumor necrosis factor (TNF) family and T cell immune modifier protein. Consistent across 3 trials. LamB is surface expressed. | (Layton *et al.* 2011) |
| *Salmonella enterica* serovar Enteritidis phage type 13A S NCTC 12023/ATCC 14028 (Δ*aroA* Δ*htrA* M2e-CD154) | Chicken (heart and liver) | Chicken | Yes | Yes | 250 µL ($10^6$ - $10^8$ CFU / mL) on day 1 then 250 µL ($10^6$ - $10^8$ CFU / mL) on day 21. | Oral gavage | 1 day | n = 10 | A/Turkey/Virginia/158512/2002 H7N2 LPAI | Highly pathogenic avian influenza | M2e-CD154 fusion. Δ*aroA* Δ*htrA* | Yes | *lamB* native promoter | Cell wall anchored via *lamB* gene | Chromosomal integration | No difference in mortality between vaccinated and saline control groups (80% dead in 3 days) in challenge model. Significantly reduced morbidity compared to saline control. Lower viral shedding. | CD154 (CD40 ligand) is part of the tumor necrosis factor (TNF) family and T cell immune modifier protein. LamB is surface expressed. | (Layton *et al.* 2009) |
| *C. perfringens* type A (Strain 11) non-virulent NetB positive strain (nvNetB+) | Chicken | Chicken | Yes | No | Oral gavage ($10^{12}$ to $10^{13}$ CFU / mL) on days 2 and 12 with adjuvant CT (List Biological) (0.025 mg CT/g body weight) then inoculated media in-feed at 3:4 (v:w) on days 2 and 12 | Oral gavage then feed | 2 days | n = 10 | *C. perfringens* | Necrotic enteritis | N/A | N/A | N/A | N/A | N/A | Significantly higher anti-*C. perfringens* secretome IgA and IgY antibodies post-challenge when compared to non-vaccinated control. Significantly lower NE disease scores. | *C. perfringens* type A (Strain 11) non-virulent NetB positive strain (nvNetB+) has low intestinal pathology in small percentage of birds | (Mishra *et al.* 2017) |
| *Salmonella* Enteritidis strain 76Sa88 Δ*hilAssrAfliG* | Chicken | Chicken | Yes | Yes | 1 x $10^8$ CFU / mL | Oral gavage | 1 day | n = 75 | *Salmonella* | Salmonellosis | Δ*hilAssrAfliG* | N/A | N/A | N/A | N/A | Significantly lower colonisation from *S.* Enteritidis challenge strain at days 14 to 28 | | (De Cort *et al.* 2013) |
| *Salmonella* Enteritidis strain 76Sa88 Δ*hilAssrAfliG* | Chicken | Chicken | Yes | Yes | 125 mL coarse spray (7.28×$10^8$ CFU / mL) | Coarse spray in 40 cm by 60 cm box | 1 day | n = 60 | *Salmonella* | Salmonellosis | Δ*hilAssrAfliG* | N/A | N/A | N/A | N/A | Significant protection from *S.* Enteritidis challenge strains until day 37 (last day tested) | | (De Cort *et al.* 2015) |

| Live vaccine | Strain origin | Host | Vaccine native to host | Attenuated | Dosage | Dosage route | Host age when dosed | Sample size | Target organism | Disease caused | Antigen/therapeutic protein expressed | Codon Optimized | Promoter type | Excreted/membrane bound/cytoplasm | Chromosomal or plasmid-based | Outcome | Notes | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Lactococcus lactis* MG1363 | Unknown | Mice | Unknown | Plasmid cured | 100 µL 2 x $10^9$ CFU / mL) for 3 consecutive days | Oral gavage | 6 weeks | n = 9 | Enterohemorrhagic *E. coli O157:H7* | Diarrhea and haemorrhagic colitis | EspB | Yes | $P_{nisA}$ (nisin inducible promoter) | Secretion signal protein SPUsp45 | Plasmid pNZ:SEC:EspB plasmid | Significant increases in IL-4, IL-10 and IFN-γ in peyers patches. Significant increases in fecal IgA and significant reduction of *E. coli* O157:H7 fecal shedding. | Plasmid-cured *Streptococcus lactis* NCDO 712 (Gasson 1983). | (Ahmed *et al.* 2014) |
| *Lactococcus lactis* MG1363 | Unknown | Mice | Unknown | Plasmid cured | 100 µL (1×$10^{10}$ CFU / mL) across 3 consecutive days, every 2 weeks x 3 | Intragastric gavage | 6 – 8 weeks | n = 8 | Influenza A (H1N1) 2009 HA1 | Influenza | HA1 and scFv protein linker ($(Gly_4Ser)_3$) | Unknown | CMV | Purified protein mixed with *L. lactis* and membrane bound via HA1/scFv protein linker/AcmA fusion protein | N/A | Significant survival rate increase and significant increase of intestinal, fecal and nasal fluid IgA levels | Plasmid-cured *Streptococcus lactis* NCDO 712 (Gasson 1983). Non-recombinant strain. scFv protein linker joins variable region of immunoglobulin light and heavy chains. | (Jee *et al.* 2017) |
| *L. plantarum* NCIMB8826 | Human saliva | C57 Bl/6 N Crl BR female mice | No | No | One group – 3 x 10 µL PBS ($10^9$ CFU / mL) intranasally 3 weeks apart. Another group – 1 x 100 µL 250mM $NaHCO_3$, 0.5% casein and 1% glucose ($10^9$ CFU / mL) intragastrically 3 weeks apart. | One group nasal gavage, other group intragastric gavage. | 8 weeks | n = 8 | *Clostridium tetani* | Tetanus | Tetanus toxin fragment C (TTFC) in | Unknown | Constitutive $P_{ldhL}$ | *Streptococcus pyogenes* M6 signal sequence and *Streptococcus pyogenes* M6 membrane anchor | pMEC127, pMEC160 and pMEC139 (pMEC4 base plasmid) | Cytoplasmic-expression levels highest compared to secreted and membrane-bound. Significant IgA and IgG increase. | No challenge. | (Reveneau *et al.* 2002) |
| *L. acidophilus* LA-5 | Dairy products | BALB/c male mice | No | Unknown | 0.2 mL / mouse ($10^{10}$ CFU / mL) for 3 days, and repeated twice two week later | Oral gavage | 8 weeks | n = 20 | Bovine viral diarrhea virus | Bovine viral diarrhea/mucosal disease | E0 | Unknown | P32 | Cytoplasmic | pMG36e-E0 plasmid | Significant increases to serum IgG, IL-2, fecal IgA, and IFN-γ and 90% protection rate compared to negative control (10%) | | (Zhao *et al.* 2015) |
| *S. enterica sv. Typhimurium* χ9718 | Unknown | Chicken | Unknown | Yes | 1 mL ($10^8$ CFU / mL) then repeated 2 weeks later | Oral gavage | 1 day | Unknown | *C. jejuni* | Campylobacteriosis | CjaA (cj0982c) | Unknown | $P_{trc}$ | Cytoplasmic | pUWM1161 | Significant increase to IgA and IgY at days 35 onwards and 1.4 log reduction of *C. jejuni* colonisation. | | (Łaniewski *et al.* 2014) |
| *L. lactis* subsp. *lactis* IL1403 | Cheese starter culture | Chicken | No | Plasmid cured | 100 µL ($10^9$ CFU / mL) repeated on day 8 and 17 | Oral gavage | 1 day | n = 19 | *C. jejuni* | Campylobacteriosis | CjaA, CjaD or CjaAD (fusion protein with CjaD) | Unknown | *L. lactis* IL1403 usp45 promoter | Usp45 signal sequence and cell wall anchored with YndF | Plasmids pUWM1392, pUWM1395 or pUWM1412 | Significant IgA and IgY antibody increases in CjaAD compared to negative control. No difference in *C. jejuni* loads. No difference between cell-wall anchored and cytoplasmic expression. | Cured of plasmids (Simon *et al.* 1985) | (Kobierecka *et al.* 2016) |

The earliest attenuated vaccinations were live vaccines, such as Pasteur's rabies vaccine in 1885 (Lackenbach 1912). The rabies virus was weakened then injected into the patient, essentially becoming an antigen delivery vehicle. This is an attenuated strain, rather than an engineered vector. Although these attenuated organisms were generally effective (Webste 1922) depending on the vaccine, there were many criticisms regarding safety, due to virulence reversion of the causative agent (Sadusk *et al.* 1946). This strain reversion phenotype endangered people who opted for vaccination and is a major drawback with the use of attenuated live vaccines.

It is hypothesised that a mucosally delivered live vaccine will be an effective means of inducing an appropriate IgA-based immune response as a control measure for the reduction of potential microbial targets such as *C. jejuni*, *C. perfringens*, and *S. enterica* loads in broiler chickens. There are a number of lines of evidence that support this hypothesis, as follows:

1) Broiler chicken maternal antibodies can partially protect young chicks from colonisation and disease (Kowalczyk *et al.* 1985; Myszewski *et al.* 1990; Cawthraw *et al.* 1994; Sahin *et al.* 2001; Sahin *et al.* 2003; Humphrey *et al.* 2014), therefore a vaccine-based approach which also utilises the immune system to protect against colonisation has merit.

2) There is likely to be a relationship between the infective dose of unwanted organisms and subsequent colonisation levels. This has been demonstrated in *C. jejuni* (Wassenaar *et al.* 1993). Dosage has been shown to be proportional with severity of disease using *C. perfringens* strains as the infective inoculum and resulting NE disease (Kaldhusdal *et al.* 1999).

3) Mucosal immune responses and production of IgA have been shown to be proportional to antigen dose (Cardenas *et al.* 1993; Curtiss III *et al.* 1996; Grangette *et al.* 2001; Heier *et al.* 2001; Reveneau *et al.* 2002; De Cort *et al.* 2013; Mishra *et al.* 2017).

4) An inverse relationship between antibody levels (IgA and IgY) and bacterial load has been established for *S.* Enteritidis (Crhanova *et al.* 2011; De Cort *et al.* 2013), *S. typhimurium* (Cardenas *et al.* 1993; Curtiss III *et al.* 1996), and *C. perfringens* (Mishra *et al.* 2017). This indicates mucosal IgA production is effective at removing or preventing colonisation of mucosal dwelling bacteria.

The ideal live vaccine vector must possess a number of properties to optimize the potential success of the vaccine, as follows:

1. Strong coloniser and highly persistent within the host
    a. The vector is the foundation of the live vaccine system, which acts as a vehicle for therapeutic proteins

2. Generally regarded as safe (GRAS)/non-disease causing to the host or anything above host in a food web
    a. Such as humans
3. Capable of genetic manipulation
    a. Transformation with plasmid DNA/mutagenesis to introduce foreign antigen expression cassettes
4. Capable of heterologous protein expression
5. Delivery of antigenic proteins must be spatially appropriate to the targeted organism
6. Expression level of recombinant proteins must be appropriate to solve the targeted issue
7. Heterologous protein expression does not harm the host

### 1.4.1 Administration

The aim of live vaccine vector delivery is to induce a mucosal immune response. It is hypothesised that this can be most effectively achieved by delivering antigens to the mucosal surface. There are a number of mucosal surfaces that could be targeted such as nasal, lung, and gut surfaces, and hence there are a number of potential administration routes that could be used to target these surfaces (Table 1). Administration routes include in-feed, in-water, coarse spray or fine spray, which can also be used to deliver material to lung mucosa (Grangette *et al.* 2001; Reveneau *et al.* 2002; De Cort *et al.* 2013; Ragione *et al.* 2013; De Cort *et al.* 2015; Kanojia *et al.* 2017). The inoculation route that is used can result in differing levels of immune responses and this seems vaccine and vector-dependant (Reveneau *et al.* 2002; Van Immerseel *et al.* 2005; De Cort *et al.* 2015). Once a live vaccine is developed, several administration methods should be investigated, primarily oral or nasal. An advantage of live vaccines is that they are rapid and inexpensive to dose large populations due to this oral or nasal entry into the animal. This reduces the labour-intensive process of subcutaneous vaccination, or the purchase of expensive automated machinery. The main requirement of live vaccines is selective bacterial media on which the strains are grown, which is considerably cheaper than expensive and laborious methods such as subunit vaccination of individual broiler chicks. If the live vaccine is a potent chicken coloniser, faecal shedding would also allow potential within-flock colonisation, enhancing the efficacy of the vaccine.

#### 1.4.1.1 Oral and in-feed

Live vaccines can be delivered in feed (Mishra *et al.* 2017) or water (Ragione *et al.* 2013; De Cort *et al.* 2015). Live vaccines can also be delivered intranasally via instilling culture into nostrils to allow subsequent colonisation of the gastrointestinal tract mucosa (Grangette *et al.* 2001; Reveneau *et al.* 2002) or intragastric delivery via gavage (Reveneau *et al.* 2002; Alimolaeia *et al.* 2017). However, intragastric gavage results in gastrointestinal tract injury, the disruption of the mucosal epithelium and reduced immune

responses (Reveneau *et al.* 2002; Kinder *et al.* 2014), along with being impractical on a large industry scale, and is therefore not a suitable method of live vaccine delivery.

#### 1.4.1.2 Coarse spray

Coarse spray is the process in which the live vaccine is sprayed onto the host, allowing the vaccine to make contact with mucosal surfaces. This is an alternative to in-feed delivery and has demonstrated superiority over other oral delivery methods (De Cort *et al.* 2015). It is hypothesised that sprayed delivery of live vaccines on chickens creates even and synchronised distribution due to droplet formation. These droplets are orally ingested during feather preening (Caldwell *et al.* 2001; De Cort *et al.* 2015). Strong protection has been achieved against *Salmonella* Enteritidis via course spray of the mutant *Salmonella* Enteritidis*Δhil*A*ssr*A*fliG* strain (De Cort *et al.* 2015). The administration route of the Poulvac (Zoetis) vaccine was compared between coarse spray inoculation and drinking water (Ragione *et al.* 2013). Both proved effective in reducing disease prevalence of colibacillosis and associated mortality from the avian pathogenic *E. coli* strain O78, however, a combination of coarse spray followed by drinking water inoculation of chickens was similar to coarse spray alone in reducing disease prevalence (Ragione *et al.* 2013). This indicates that either inoculation route may be effective, as both produced partial protection.

### 1.4.2 Therapeutic protein expression

#### 1.4.2.1 Antigens

The antigen expressed by the live vaccine is the most significant determinant of vaccine effectiveness. Live vaccines function by expressing antigens or epitopes from the target organism, allowing the host to mount a protective immune response (Łaniewski *et al.* 2014) (Table 1). The importance of selecting highly antigenic, surface-exposed proteins or epitopes that will be available to the host immune system is critical for a strong host immune response capable of clearing the target organism (Power *et al.* 1994; Layton *et al.* 2011; Łaniewski *et al.* 2014).

The NE-essential virulence factor NetB (Keyburn *et al.* 2008) has been used previously in several vaccine trials. It has demonstrated immunogenicity and partial protection against *C. perfringens*-mediated NE (Jang *et al.* 2012; Costa *et al.* 2013; Keyburn *et al.* 2013; Keyburn *et al.* 2013; Lillehoj *et al.* 2017). Recombinant NetB (rNetB) (S254L) is a reduced-activity mutated derivative of NetB (Yan *et al.* 2013). This has been used successfully in a maternal vaccine study where maternal IgY antibodies persisted in progeny up to 21 days in sufficient levels to provide partial protection when challenged (Keyburn *et al.* 2013).

#### 1.4.2.2 Co-expression of immune-modulating proteins

Traditionally, adjuvants such as aluminium hydroxide are commonly used in combination with vaccines to boost immune responses (Aguilar *et al.* 2007). More recently, immune-modulating proteins have been co-expressed by live vaccines as an adjuvant, alongside antigenic proteins (Table 1). This provides an interesting avenue to further improve efficacy of live vaccines. There are numerous examples of success using this approach. Significant Pig Swine Fever Virus protection was achieved by an *L. plantarum* live vaccine expressing the E2 protein and Thymosin α-1 (Tα1) when compared to the E2 protein alone (Xu *et al.* 2015). Tα1 is an immune-modifier protein which is involved maturation in lymphocytes, T-cell function and promoting antibody production (Xu *et al.* 2015). CD154, a tumour necrosis factor ligand has been successfully used in several chicken live vaccines as fusion proteins with antigen/epitope sequences (Layton *et al.* 2009; Layton *et al.* 2011) and has been shown to boost immune responses, both humoral and systemic. An approach using single chain variable fragments (scFv) protein linkers, which join the variable region of immunoglobulin light and heavy chains, has been shown to significantly improve immune responses by improving antigen display (Sletta *et al.* 2004; Yuvaraj *et al.* 2008; Jee *et al.* 2017).

Chicken interleukin-6 (chIL-6) is a proinflammatory cytokine implicated in the febrile response. Originally studied for its potential to increase egg yolk IgY antibodies (Nishimichi *et al.* 2005; Niederstadt *et al.* 2012), it could also be applied for use in a live vaccine. Co-expression of IL-6 by a live vector could improve humoral immune responses against the target. Co-expression or administration of chIL-6 with antigens such as *E. coli* K88 fimbrial protein FaeG, lipopolysaccharide (LPS) and *Clostridium botulinum* toxin BoNT A1 have previously induced significant increases in IgY, IgA and maternal antibody production in chickens and chicken cell lines (Cho *et al.* 2004; Nishimichi *et al.* 2005; Niederstadt *et al.* 2012).

#### 1.4.2.3 Membrane-bound/secreted/cytoplasmic expression

Another important consideration in therapeutic protein expression, whether of antigens or immune-stimulatory molecules, is protein localisation, specifically whether the protein has a cytoplasmic, membrane-anchored or secreted localisation. It has been demonstrated using *L. lactis* live vaccine systems that secretion or surface-binding of antigens can significantly improve IgA and IgG immune responses when compared to cytoplasmic expression (Ahmed *et al.* 2014; Kobierecka *et al.* 2016). Fusing the therapeutic protein-coding sequences with a surface-bound protein allows the immune system to come into contact with the desired sequences and mount a successful immune response. Maltose outer membrane porin (LamB) has been used successfully as a fusion protein to anchor both antigenic sequences as well as immune-modulating proteins in *S.* Enteritidis live vaccines (Layton *et al.* 2009; Layton *et al.* 2011). N-acetylmuramidase (AcmA), a cell-wall anchoring protein, has demonstrated stable fusion and surface expression in *L. lactis* (Raha *et al.* 2005). This work led to an *L. lactis* live vaccine against influenza A (H1N1) expressing antigen HA1 fused to AcmA, which resulted in significant IgA immune

responses and viral protection in mice (Jee *et al.* 2017). YndF has also been used to anchor *C. jejuni* antigens and improve immune responses of chickens (Kobierecka *et al.* 2016).

One study compared three *L. plantarum* live vectors expressing the model antigen tetanus toxin fragment C (TTFC) from *Clostridium tetani*, under the same promoter ($P_{ldhL}$), directed to three different locations, the cytoplasm, membrane-bound and secreted extracellularly (Reveneau *et al.* 2002). The effectiveness of each approach was compared via monitoring the immune response in a mouse model. Total amounts of TTFC produced were different, with cytoplasmic TTFC being the highest, followed by membrane-bound, then secreted. Additionally, membrane bound TTFC was only found to be expressed on <15% of *L. plantarum* cells, with large amounts of TTFC still present within *L. plantarum* cells (Reveneau *et al.* 2002). Interestingly, when the live vaccines were inoculated into mice, cytoplasmic expression resulted in the strongest serum IgG and IgA response at endpoint titres, followed by membrane-bound then secreted, indicating that the antigen level is proportional with the host antibody response. This was unexpected, as surface-exposed or secreted antigens are exposed to the immune system at higher frequencies when compared to cytoplasmically-expressed antigens, which rely on cellular lysis and protein leakage for exposure to the host (Reveneau *et al.* 2002; Sheedy 2006).
These results raise the question of optimal route of expression of the protein antigen, and several localisations should be attempted. The aim of live vaccination in this work is to induce a mucosal IgA and, to a lesser extent, serum IgG immune response. Therefore for this work it does not matter where the protein antigen is localised in the cell, as long as these immune responses are generated.

## 1.5 Stability of live vaccine and therapeutic protein expression

The key role of live bacterial delivery vectors is to effectively deliver antigens to the host immune system. Since antigen dose and host immune responses are proportional (Cardenas *et al.* 1993; Curtiss III *et al.* 1996; Heier *et al.* 2001; Reveneau *et al.* 2002; Mishra *et al.* 2017), vector persistence within the host is a vital aspect of continued, stable delivery of expressed antigen/s. Long-term stable expression of therapeutic proteins therefore provides the highest chance of successful vaccination.

### 1.5.1 Native, constitutive promoters

Promoters are short DNA sequences found upstream from genes that regulate gene expression by docking RNA polymerase, thereby allowing the transcription of genes (Browning *et al.* 2004). Promoters have a wide range of activity, which are sequence and species specific (Browning *et al.* 2004; Stephenson *et al.* 2011; Guiziou *et al.* 2016). They can be inducible or constitutive, and a range from both have been used in previously developed live vaccines (Aires *et al.* 2006; Kulkarni *et al.* 2010; Layton *et al.* 2011; Xu *et al.* 2015).

Strong, constitutive promoters allow high, continuous expression of antigens, and therefore increase the chance of successful vaccination. The *Lactobacillus acidophilus* lactate dehydrogenase (*ldhL*) and surface layer protein (*slpA*) gene constitutive promoters have been previously used with some success to express EGFP in *L. lactis* and *Lactobacillus reuteri* (Reveneau *et al.* 2002; Lizier *et al.* 2010). The *L. acidophilus slpA* promoter was not active in *L. reuteri*, indicating the species specificity of promoters. The erythromycin ribosomal methylase (*ermB*) promoter from pAMB1 (an *Enterococcus faecalis* plasmid) was found to be the strongest promoter tested in *L. lactis* and *L. reuteri*, followed by the *ldhL* promoter (Lizier et al., 2010). This group concluded that for the successful development of a live vaccine, a reliable and constitutive promoter that allows strong expression of therapeutic proteins is necessary (Lizier et al., 2010).

Several constitutive heterologous promoters, including the *ldhL* promoter from *L. casei*, the *slpA* promoter from *L. acidophilus*, a promoter-like sequence (P144) from *Lactobacillus paracasei*, the phosphoglyceromutase promoter ($P_{pgm}$) from *L. acidophilus* and promoter 23 ($P_{23}$) from *L. lactis* subsp. *Cremoris*, have been tested previously in *Lactobacillus agilis* La3, with expression compared by measuring EGFP (Stephenson et al., 2011). These promoters produced little to no EGFP, so the *L. agilis* La3 chromosome was mined for sequence homology to these promoter regions and the promoters $P_{ldhL}$, $P_{slpA}$, $P_{pgmr}$ were identified (Stephenson et al., 2011). These native promoters were then used for EGFP expression in *L. agilis* La3, and it was found that the $P_{ldhL}$ promoter produced the highest EGFP levels (Stephenson et al., 2011). This work showed that homologous promoters found within the live vector host work more effectively than heterologous promoters.

### 1.5.2 Ribosomal binding sites

Another aspect that influences protein expression is the ribosomal binding site (RBS) or Shine-Dalgarno sequences associated with genes. They allow ribosomes to bind to upstream regions of Open Reading Frames (ORFs), allowing transcription (Shine *et al.* 1974; Kim *et al.* 1991; Fitzsimons *et al.* 1994; Hols *et al.* 1994; An *et al.* 2006; Wegmann *et al.* 2013). They are short, sequence specific sites that have been shown to significantly impact on protein expression levels. For example, depending on the RBS used, expression levels can vary from 50-fold up to 600-fold, as previously demonstrated in *Bacillus subtilis* (Guiziou *et al.* 2016).

### 1.5.3 Colonisation persistence of live vaccine

Colonisation persistence of the live vaccine may be as important as strong, constitutive expression of therapeutic proteins. If the antigen-expressing bacterium is transient, the host will receive a lower antigen dose and the vaccine will not be as effective. Most live vaccines are attenuated disease-causing agents (Table 1). Attenuation removes

particular virulence genes which reduces the disease severity, however, there are risks associated with this approach. If virulence factors are on conjugative plasmids or in extracellular DNA (Bohez *et al.* 2007; Garcia *et al.* 2014; Kubasova *et al.* 2014; Salem *et al.* 2017) then the attenuated strain may reacquire virulence genes via horizontal gene transfer (Salem *et al.* 2017). Secondly, attenuated strains are less likely to colonise and persist when compared to the wild type strains because their regular properties have been altered in order to attenuate them (De Cort *et al.* 2013; De Cort *et al.* 2015). Live attenuated vaccinations can deliver therapeutic proteins to the mucosal layer of the gastrointestinal tract but colonisation effectiveness can be compromised, even if the attenuated strain was originally isolated from the target host (Bohez *et al.* 2007).

There have been partially protective attenuated live vaccines developed (Layton *et al.* 2009; Layton *et al.* 2011; De Cort *et al.* 2015) using *S.* Enteritidis, however, they appear to be transient, persisting within a maximum of 40% of the flock for approximately 11 days (Layton *et al.* 2011), and usually disappearing within ~16 days (De Cort *et al.* 2015). The lack of persistence of these attenuated live vectors may be a limiting factor regarding protection from colonisation or disease. In another example, an *Enterococcus faecalis* strain was inoculated via oral gavage into day old chickens at $10^6$ CFU. When administration was halted, the organism could not be reisolated after 5 days and was absent from the gastrointestinal tract. This demonstrates the potential value of identifying a vaccine vector bacterium that is able to reliably colonise and persist within the gastrointestinal tract of the target host, for the development of a live vaccine (Netherwood *et al.* 1999).

One live vaccine against *C. perfringens* used a chicken-colonising *C. perfringens* type A isolate (strain 11), which is a non-virulent NetB positive strain. This strain produced low intestinal pathology in experimentally challenged birds (Mishra *et al.* 2017). When chickens were dosed both via oral gavage and in-feed, significantly higher IgA and IgY antibody levels were induced when compared to the unvaccinated control, along with significant but not complete protection against NE. This strain was shown to persist until the time of challenge. An inverse, moderate correlation ($R^2$ = 0.52 for IgY and $R^2$ = 0.4 for IgA) between antibody levels and lesion scores indicated that some antibody-mediated partial protection against *C. perfringens* was achieved (Mishra *et al.* 2017).

Non-pathogenic, non-attenuated, live vaccines pose no known threat to the intended host (Aires *et al.* 2006; Xu *et al.* 2015) and retain their natural fitness within the host. This increases the chances of successful colonisation, persistence and effective antigen presentation to the host immune system. A good chicken live vector candidate will be GRAS, potentially a probiotic, be capable of genetic manipulation, capable of heterologous protein expression and be highly persistent within the chicken gastrointestinal tract. One such candidate, *Lactobacillus agilis*, is the focus of this thesis and is described below.

## 1.6 *Lactobacillus agilis*

### 1.6.1 Characteristics

*L. agilis* is a Gram positive, non-spore forming bacilli that has been isolated from sewage (Weiss *et al.* 1981), horses (Morita *et al.* 2007), pigeons (Baele *et al.* 2001) chickens (Mitsuoka 2002; Lan *et al.* 2003; Stephenson *et al.* 2010; Bujnakova *et al.* 2014; Chen *et al.* 2017), humans (Molin *et al.* 1993), pigs (Mitsuoka 2002; Murphya *et al.* 2013; Kinoshita *et al.* 2016), lemurs and tapirs (Kajikawa *et al.* 2016). It has gained interest over the years as a potential probiotic and food safety strain (Palop *et al.* 1995; Lan *et al.* 2003). A defining characteristic of *L. agilis* is motility via flagella (Weiss *et al.* 1981). The peritrichous flagella of the highly motile *L. agilis* JCM 1048 (BKN88 isolate) has been characterised and found to be encoded by the *fliC1* and *fliC2* genes (Kajikawa *et al.* 2016). *L. agilis* also ferments ribose, mannitol, lactose, maltose, sucrose, trehalose, cellobiose, melibiose and raffinose, and salicin (Weiss *et al.* 1981). Arabinose, xylose and glycerol are not fermented and *L. agilis* is non-nitrite producing. It is catalase negative and does not produce indol, $H_2S$, lipase or urease. The cell wall contains glucose and galactose but no teichoic acid. It is a facultative anaerobe that is obligately saccharoclastic, is homofementative and produces L-(+)-lactic acid and 10% D-(-)-lactic acid (Weiss *et al.* 1981). The GC content was estimated at 43.4% (Weiss *et al.* 1981). *L. agilis* has been found in the human small intestinal mucosa from both healthy and unidentified disease samples (Molin *et al.* 1993) indicating its presence within the human population and its status as GRAS*. L. agilis* R16 demonstrated an ability to detoxify glucosinolates from foods, which has potential value in food safety due to behaviours such as pH-lowering lactic acid production, sinigrin degradation capability and fast growth (Palop *et al.* 1995).

### 1.6.2 *L. agilis* as a chicken probiotic

*L. agilis* JCM 1048 (ATCC 43564) has been identified as a potential therapeutic chicken probiotic (Lan *et al.* 2003). Originally isolated from a chicken intestine, it was shown to have high tolerance of bile (40%) and 6.5% NaCl, along with being completely resistant to in-feed antibiotics. It inhibited growth of *Salmonella* spp. and *C. jejuni*, both making up most poultry-related foodborne illness incidences (Sahin *et al.* 2001; Catarame *et al.* 2006; EFSA 2010), when used as a probiotic mix with *Lactobacillus salivarius* ssp. *salicinius*. It was hypothesized that the inhibitory activity was due to acid production. This study showed a short-term significant increase of *Lactobacillus* isolates from chickens after 7 days when *L. agilis* JCM 1048 was introduced as a probiotic mix with *Lactobacillus salivarius* JCM 1230. After 40 days, chicken body weights were shown to be significantly heavier when fed the probiotic mix. The probiotic mix also had an impact on short chain fatty acid (SCFA) levels and bacterial populations within caecal contents. Streptococci and other aerobe counts were reduced, and anaerobes, bifidobacteria and lactobacilli counts increased. An increase in propionate and butyrate production was demonstrated, along with a reduction in acetate, while no lactate was present (Meimandipour *et al.* 2009). It was postulated that these changes were the result of cross-feeding, where butyrate producers within the caecal contents were using lactate produced by the probiotic mix as

a substrate to create butyrate (Meimandipour *et al.* 2009). These bacteria most likely were also consuming acetate, which led to its decrease. *L. agilis* BCRC 10436 and *L. reuteri* BCRC 17476 were isolated from the intestine of chickens and fed to one day old broiler chickens as probiotics (Chen *et al.* 2017). After 35 days, chickens fed the *L. agilis* and *L. reuteri* probiotic mix showed significant increases in cecal *Lactobacillus* counts within the chicken, but no difference in FCR. Contrasting with the Meimandipour (2009) study, by 21 days probiotic-treated chickens had significantly lower propionic acid, iso-butyric acid and ammonia, and significantly lower total SCFA than the control birds by 35 days. This again can be explained by cross-feeding. Chickens showed significantly lower amounts of low density lipoprotein cholesterol in the blood after 21 days. This can be explained by either the probiotic strains or environment-modulating impact of the probiotic strains selecting for bacteria which deconjugate bile salts, removing them as cholesterol synthesis precursors (Chen *et al.* 2017). These studies demonstrate the potential impact *L. agilis* can play within the gastrointestinal tract of the chicken.

The flagella genes *fliC1* and fliC2 from *L. agilis* JCM 1048 (BKN88 isolate) were recombinantly expressed and shown to not induce IL-8 production from Caco-2 human epithelial cells when compared to flagellin proteins from *S. Typhimurium* 92–35 and *L. monocytogenes* EGD (Kajikawa *et al.* 2016). Though the protein sequence between these organisms is highly conserved, they have different toll-like receptor 5 recognition sites (`LNR` instead of `LQR`), which may explain why there was little immune response generated. The flagella filaments were also shown to be stable in low pH (pH 1.3) and stable up to 47°C. The authors state that these features of the flagellin are characteristic of a commensal and symbiotic bacteria (Kajikawa *et al.* 2016). This observation is an important one because *L. agilis* is seen by the host as part of normal gastrointestinal tract microbiota.

### 1.6.3 *L. agilis* La3 as live vector

*L. agilis* strain La3 was identified in 2010 from the gastrointestinal tract of a broiler chicken, and has shown promise as a potential live vaccine antigen vector (Stephenson *et al.* 2010). *L. agilis* La3 is a rare isolate because it is highly persistent within the chicken gastrointestinal tract (from day 1 post-hatching to 42 days market date), even when chickens are fed different diets and have different underlying gut microbiota compositions. It has demonstrated an ability to successfully colonise chickens 18 days post-hatching after the gut microbial communities have been established (Vezina 2013). It is GRAS, culturable, can be genetically manipulated, and capable of heterologous gene expression (Stephenson *et al.* 2011). All these characteristics make *L. agilis* La3 suitable for development as a live vaccine vector for chickens.

## 1.7 Aims

When this project was undertaken, the aims of the project were as follows:

i) To further characterise *L. agilis* La3
   a) via genome sequencing and analysis, and
   b) to determine genome-wide transcription levels and identify new, strong, *L. agilis* La3 promoters

ii) To develop *L. agilis* La3 as a stable, live delivery platform
   a) by constructing and inserting expression cassettes into the *L. agilis* La3 chromosome
   b) by using a reporter gene to determine expression levels of bioinformatically-identified promoters

iii) To test a number of therapeutic proteins expressed by *L. agilis* La3
   a) to use *in vitro* assays to determine suitable therapeutic proteins
   b) to test suitable proteins *in vivo* within an animal model for their immunogenic properties

# 2. Materials and methods

The list of bacterial strains used in this study are in Table 2.

## 2.1 Bacterial strains

Table 2: List of bacterial strains used in this study

| Strain | Characteristics | Source |
|---|---|---|
| *L. agilis* La3 | Highly persistent broiler chicken strain, wild type | (Stephenson 2010) |
| *L. agilis* La3::$P_{ldhL}$-EGFP | Recombinant: Expresses EGFP under native $P_{ldhL}$ promoter in *L. agilis* La3 genome | This study |
| *L. agilis* La3::$P_{gap}$-*creiLOV* | Recombinant: *creiLOV* under native $P_{gap}$ promoter, in *L. agilis* La3 genome | This study |
| *L. agilis* La3::$P_{eno}$-*creiLOV* | Recombinant: *creiLOV* under native $P_{eno}$ promoter, in *L. agilis* La3 genome | This study |
| *L. agilis* La3::$P_{eft}$-*creiLOV* | Recombinant: *creiLOV* under native $P_{eft}$ promoter, in *L. agilis* La3 genome | This study |
| *L. agilis* La3::$P_{cwah}$-*creiLOV* | Recombinant: *creiLOV* under native $P_{cwah}$ promoter, in *L. agilis* La3 genome | This study |
| *L. agilis* La3::$P_{gap}$-*rnetB* | Recombinant: *rnetB* under native $P_{gap}$ promoter, in *L. agilis* La3 genome | This study |
| *L. agilis* La3::$P_{eno}$-*rnetB* | Recombinant: *rnetB* under native $P_{eno}$ promoter, in *L. agilis* La3 genome | This study |
| *L. agilis* La3::$P_{eft}$-*rnetB* | Recombinant: *rnetB* under native $P_{eft}$ promoter, in *L. agilis* La3 genome | This study |
| *L. agilis* La3::$P_{cwah}$-*rnetB* | Recombinant: *rnetB* under native $P_{cwah}$ promoter, in *L. agilis* La3 genome | This study |
| *E. coli* JM109 | *E. coli* strain used for cloning and expression, F´ *traD36 proA*$^+$*B*$^+$ *lacI*$^q$ *Δ(lacZ)M15/ Δ(lac-proAB) glnV44 e14*$^-$ *gyrA96 recA1 relA1 endA1 thi hsdR17* | Promega |
| *E. coli* MM294 | *E. coli* strain used for producing multimeric plasmid DNA, F$^-$ *endA1 hsdR17* ($r_K^- m_K^+$) *glnV44 thi-1 relA1 rfbD1 spoT1* | Microbiology department, Monash University, Victoria, Australia, (Meselson *et al.* 1968) |
| *E. coli* JM110 | *E. coli* strain used for producing *dam-/dcm-* methylated plasmid DNA, F´ *traD36 lacI*$^q$*Δ(lacZ)M15 proA*$^+$*B*$^+$*IrpsL* (Str$^r$) *thr leu thi lacY galK galT ara fhuA dam dcm glnV44 Δ(lac-proAB)* | ATCC 47013 |
| *E. coli* BL21 (DE3) | *E. coli* strain used for expression, B F$^-$ *ompT gal [E. coli* B is naturally *dcm and lon] hsdS*$_B$ with DE3, a λ prophage carrying the T7 RNA polymerase gene and lacI$^Q$ | New England Biolabs |

## 2.2 Growth and storage conditions

### 2.2.1 *L. agilis* La3

*L. agilis* La3 strains (Table 2) were grown at 37°C in De Man, Rogosa and Sharpe (MRS) liquid media (OXOID), or on MRS agar plates (1% Bacteriological Agar (OXOID)), anaerobically, using the AnaeroGen 2.5L Sachet (ThermoFisher Scientific). *L. agilis* La3 was grown with appropriate antibiotics (chloramphenicol 10 µg/mL, erythromycin 10 µg/mL, tetracycline 70 µg/mL), unless otherwise stated. *L. agilis* strains were stored at -80°C in 1:1 MRS media:glycerol.

#### 2.2.1.1 Growth curve

Growth curves were performed using *L. agilis* La3. Pure cultures of *L. agilis* La3 were obtained by streaking plates from glycerol stocks on MRS agar and incubating overnight at 37°C. An overnight liquid culture of *L. agilis* La3 was then grown and 100 µL subcultured into 100 mL pre-warmed MRS media at 42°C in a Whitley A35 Anaerobic workstation (gas mix: 10% $H_2$, 10% $CO_2$ and 80% $N_2$, at 49% humidity) with loosened lids (6 biological replicates). Hourly optical density readings were taken at 600 nm wavelength ($OD_{600}$) using an OD600 (IMPLEN) reader and macrocuvettes (Kartell). Serial dilutions were plated onto MRS agar and colony forming units per mL (CFUs/mL) were determined then plotted against optical density.

### 2.2.2 *E. coli* JM109

*E. coli* JM109 (Table 2) was grown at 37°C in Lysogeny Broth (LB) media (1% w/v tryptone, 1% w/v NaCl and 0.5% w/v yeast extract) at 250 RPM or on LB agar plates (1% Bacteriological Agar (OXOID)) with appropriate antibiotics (chloramphenicol 30 µg/mL, erythromycin 150 µg/mL, tetracycline 30 µg/mL) where appropriate. *E. coli* JM109 was stored at -80°C in 1:1 LB media:glycerol.

## 2.3 Gel electrophoresis

### 2.3.1 Agarose gel electrophoresis

Mini-, midi- and maxi-agarose gels were prepared by melting agarose (BIOLINE), at a final concentration of 0.8% in 1X TAE solution (40 mM Tris-acetate, 1 mM EDTA, pH 8.0) in a microwave. GelRed (Biotium) was added at a 1:50,000 ratio. Running buffer was 1X TAE and samples loaded using 1X Gel Loading Dye, Purple (NEB). Electrophoresis conditions were 100 V for 60 minutes unless otherwise stated. Gels were visualized under UV on a ChemiDoc™ XRS+ (BIO-RAD), and images were processed using ImageLab version 5.2.1 build 11 (BIO-RAD).

### 2.3.2 Formaldehyde gel electrophoresis

Mini-formaldehyde gels for RNA visualization were prepared by melting agarose (1% final concentration) in 1X 3-(N-morpholino)propanesulfonic acid (MOPS) buffer (20 mM MOPS (Sigma), 5 mM sodium acetate, 1 mM $Na_2EDTA$) with 6.5% formaldehyde and 1:50,000 dilution of GelRed. Gel tank and gel casts were cleaned with Pyroneg, then washed with 70% ethanol. RNA samples were prepared by mixing RNA sample buffer (64.5% deionized formamide, 15.5% formaldehyde, 20% 5X MOPS buffer) at a 2:1 ratio with RNA preparations. The mixed samples were heated at 60°C for 5 minutes then cooled on ice for 2 minutes. 1X RNA loading buffer was added (50% glycerol, 1 mM EDTA, 0.4% (w/v) bromophenol blue) to samples in total volume of 15 µL. Gels was pre-run for 10 minutes at 120 V, then samples were loaded and the gels were subjected to electrophoresis for 30 minutes at 120 V.

## 2.4 RNA isolation from *L. agilis* La3

### 2.4.1 RNA for sequencing

RNA was isolated from mid-logarithmic and early-stationary phases to identify transcript levels across the genome. All reagents for RNA isolation and purification were made using DEPC-treated water (method 2.8) to ensure RNase free reagents. *L. agilis* La3 was grown anaerobically at 42°C in a Whitley A35 Anaerobic workstation. 1 mL of culture was taken at the mid-logarithmic phase, ($OD_{600}$ 3.5) and early-stationary phase ($OD_{600}$ 5.7). 2 volumes of RNAlater (QIAGEN) were added to samples, then vortexed for five seconds. The tubes were then centrifuged for ten minutes at 5,000 x g and the supernatant discarded. Once the pellet was dry, 100 µL TEL (10 mM Tris-Cl, 1 mM EDTA, 15 mg/mL lysozyme, pH 8.0) with 1% 2-mercaptoethanol were added and vortexed for 10 seconds. Samples were incubated at room temperature for 1 hour. Buffer RLT (QIAGEN) was added (700 µL) and vortexed for 10 seconds, then the suspension was added to 2 mL screw cap tubes containing 250 $mm^3$ 0.1 mm zirconia/silica beads (Daintree Scientific). Samples were bead beaten at maximum speed for five minutes in a Precellys 24 Lysis & Homogenization instrument (Bertin Technologies) then samples were centrifuged at 14,000 x g for 10 seconds. The supernatant was transferred to a new tube and 440 µL 100% ethanol was added. An RNeasy Mini Kit (QIAGEN) was used to purify RNA as per the manufacturer's instructions with the following exception: The elution of RNA using 30 µL of *DEPC-treated water* was repeated to improve yield. RNA quantity and purity was determined using the Nanodrop ND1000 (ThermoFisher Scientific) then electrophoresis was performed using a formaldehyde gel (method 2.3.2) with 1X MOPS buffer to confirm RNA presence.

### 2.4.2 cDNA for quantitative polymerase chain reaction (qPCR)

The isolation method used to prepare cDNA for quantitative PCR analysis was similar to method 2.4.1 with the following exceptions: *L. agilis* La3 was grown anaerobically at 37°C to early-stationary phase ($OD_{600}$ 5 - 6). Buffer RLY (BIOLINE) was used instead of Buffer RLT; samples were bead beaten at 6.5 m/s for 60 seconds using a FastPrep-24 Classic instrument (MP Biomedicals); 350 µL 70% ethanol was added instead of 440 µL 100%

ethanol; and the ISOLATE II RNA Mini Kit (BIOLINE) was used instead of the RNeasy Mini Kit (QIAGEN).

## 2.5 cDNA synthesis

### 2.5.1 cDNA for RNA sequencing

cDNA synthesis was performed using 5 µg of isolated RNA (method 2.4.1) and 5.6 nmol random pentadecamer primers (Stangegaard *et al.* 2006) and SuperScript™ II (400 U) (Invitrogen) as per manufacturer's instructions. The reaction was incubated at 42°C for 50 minutes then heat inactivated at 70°C for 15 minutes (Invitrogen 2010). The solution was then stored at -20°C for 2 hours. NEBNext Second strand synthesis enzyme mix (NEB) was used to generate the second-strand as per the manufacturer's instructions. Double stranded complimentary DNA (dscDNA) was then purified using the MinElute Reaction Cleanup Kit (QIAGEN) as per the manufacturer's instructions to remove primers and small fragments.

### 2.5.2 Single strand cDNA for qPCR

Single strand cDNA was synthesised from isolated RNA (method 2.4.2) using Tetro Reverse Transcriptase (BIOLINE) as per the manufacturer's instructions with the following exceptions: 3.35 nmol random pentadecamer primers (Stangegaard *et al.* 2006) was used instead of random hexamers, and no RNase inhibitors were used.

## 2.6 Sequencing

### 2.6.1 PacBio DNA sequencing

*L. agilis* La3 was grown overnight in 10 mL MRS and 10 µL subcultured in 1.5 mL MRS broth for three hours until mid-log phase ($OD_{600}$ ~3.5). *L. agilis* La3 DNA was extracted using 5 independent preps of MagAttract HMW Kit (QIAGEN) as per manufacturer's instructions. DNA was concentrated and cleaned up using Agencourt AMPure XP beads (BECKMAN COULTER) as per manufacturer's instructions with the following exception: 0.5 µL instead of 1.8 µL beads per 1 µL of DNA was used (due to size selection preference of beads), and eluted into 84 µL ultra-pure water (Cascada™ Lab Water Purification System (PALL), 0.05 µm filters). High molecular weight DNA was verified using gel electrophoresis (method 2.3.1) and quantified with a Qubit 2.0 (ThermoFisher Scientific) using the High Sensitivity dsDNA setting as per manufacturer's instructions. DNA was size selected for >20 kb using BluePippin Size Selection (Sage Science) and DNA was prepared for sequencing using SMRTbell Template Prep Kit 1.0 (Pacific Biosciences). DNA was sequenced using a PacBio RS II (PacificBioSciences) instrument at the Doherty Institute, Melbourne.

### 2.6.2 MiSeq DNA sequencing

Short-read DNA sequencing was used to improve the PacBio assembly. Chromosomal DNA was isolated from *L. agilis* La3 using the ISOLATE Fecal DNA Kit (Bioline) as per manufacturer's instructions. The Nextera XT DNA Library Prep Kit (Illumina) was used as per manufacturer's instructions to prepare the DNA library for sequencing, with size selection at ~600 bp. DNA was sequenced on a MiSeq (Illumina) as per manufacturer's instructions. The 2 x 300 read length setting and v3 sequencing chemistry were used.

### 2.6.3 MiSeq RNA sequencing

RNA was isolated from *L. agilis* La3 at mid-logarithmic and early stationary growth phases (method 2.4.1). The Nextera XT DNA Sample Preparation Kit (Illumina) was used as per manufacturer's instructions with the following exceptions: The initial amount of cDNA (method 2.5.1) was halved, and during PCR clean-up, 25 µL AMPure XP beads were added to size select fragment sizes ~600 bp. Samples were loaded onto a High Sensitivity DNA Quick Kit (Agilent) and electrophoresed using a 2100 Bioanalyzer (Agilent) as per the manufacturer's instructions to confirm size library preparation and size selection. Libraries were sequenced on the MiSeq (Illumina) platform as per manufacturer's instructions. The 2 x 300 read length setting and v3 sequencing chemistry were used.

## 2.7 Sequencing analysis

### 2.7.1 DNA sequencing analysis

#### 2.7.1.1 Genome assembly

Sequencing data was analysed on Biolinux 8 (Field *et al.* 2006). Output PacBio bax.h5 files were concatenated into .bam file using bax2bam version 0.0.8 (pitchfork release, 10/6/16) (PacificBioSciences 2015), then converted to fastq format using the bamToFastq command from bedtools version 2.17.0 (pitchfork release, 10/6/16). All PacBio reads were incorporated for assembly using Canu Assembler version 1.3 with the following options: 'stopOnReadQuality=false', 'genomeSize=2.2m' and '-pacbio-raw' (Koren *et al.* 2016). Canu was used due to its speed and it's integration of the MinHash Alignment Process (MHAP) algorithm (Berlin *et al.* 2015).
50,000 bp from the beginning and end of the assembled *L. agilis* La3 molecule were aligned against each other using BLAST (Altschul *et al.* 1990; Wheeler *et al.* 2006; Boratyn *et al.* 2013) to determine if the molecule could be circularised. Seqret (EMBOSS version 6.6.0.0) was used to reformat the circularised fasta file to appropriate fasta format. Raw PacBio reads were aligned using blasr version 5.2.b99b47c (pitchfork release, 10/6/16) (PacificBioSciences 2015) and output in bam format. This was done to confirm correct circularization and to polish the genome. The aligned bam file was sorted and ordered using the 'sort' command in samtools version 1.3.1 (Li *et al.* 2009). The sorted alignment bam file was indexed using pbindex version 0.5.0 (pitchfork release, 10/6/16). Mapped reads were then visualised on Tablet (version 1.14.04.10) (Milne *et al.* 2013). Indexing of the *L. agilis* La3 genome file was done using the 'faidx' command from samtools version 1.3.1. The

genome file was then polished using the ‘arrow’ command from variantCaller version 2.0.0 (pitchfork release, 10/6/16) (Chin *et al.* 2013).
To improve the accuracy of the genome sequence file, MiSeq data was integrated into the polished genome for short-read error correction. An index of the genome was created using the ‘index’ command from bwa version 0.7.12 (Li 2013). The MiSeq output fastq files were concatenated using the ‘cat’ command from GNU coreutils version 8.21, then aligned to the genome using the ‘mem’ option from bwa version 0.7.12. The resulting sam file was converted to bam format, sorted and indexed using the ‘view’, ‘sort’ and ‘index’ commands from samtools version 1.3.1, respectively. The indexed mapped reads were used as input for Pilon version 1.18 (Walker *et al.* 2014) for short-read error correction of the *L. agilis* La3 genome. Reads were remapped to this final molecule and visualized on Tablet to confirm complete coverage across the whole circularised genome.

#### 2.7.1.2 Annotation

The genome sequence was submitted to Genbank for annotation by the Prokaryotic Genome Annotation Pipeline (Tatusova *et al.* 2016) and submitted to Genbank with the following details: BioProject: PRJNA336089, BioSample: SAMN05504077, Localid: *Lactobacillus*, Accession: CP016766, Organism: *Lactobacillus agilis* La3.
All coding sequences (CDS) from *L. agilis* La3 were isolated from the genbank file using genbank_to_fasta v1.1 (Rocap) and used in subsequent analyses where CDS format was required.
The genome was also submitted to RAST for annotation and SEED subsystem information (Aziz *et al.* 2008; Overbeek *et al.* 2014; Brettin *et al.* 2015).

#### 2.7.1.3 Bacteriocin mining

BAGEL3 BLASTN was used to identify potential bacteriocins present in the *L. agilis* La3 genome (Heel *et al.* 2013). Date accessed: 6/12/16.

#### 2.7.1.4 Prophage gene identification

Prophage regions were identified using PHAST (Zhou *et al.* 2011). Date accessed: 6/12/16.

#### 2.7.1.5 Comparison to other *L. agilis* strains

*L. agilis* strains were downloaded from NCBI Genbank and genome statistics were compared. BPGA (version 1.3) (Chaudhari *et al.* 2016) was used to compare the core and pan genomes at 70% similarity.

#### 2.7.1.6 *L. agilis* La3 codon usage

A codon usage table was generated using cusp (version: EMBOSS:6.6.0.0) after extracting the coding sequences (CDS) from the annotated *L. agilis* La3 genome.

#### 2.7.1.7 Methylation

To analyse the methylation patterns of the *L. agilis* La3 genome, ipdSummary version 2.3 from the kineticsTools package (PacificBiosciences) and the 'find' and 'reprocess' commands from Motifmaker version 0.2.1 (PacificBiosciences) were used to generate the individual base pair modifications and methylated motifs from the PacBio sequencing. The .gff file was visualised on Tablet (version 1.14.04.10) (Milne *et al.* 2013).

### 2.7.2 RNA sequencing analysis

Two million reads were taken as a subsample from output fastq files and quality trimmed using FastqMcf version 1.04.636 from ea-utils (Aronesty 2011) with the following options: "-l 50 -q 15 -w 5 -H -X --homopolymer-pct 50 --lowcomplex-pct 70". The build command from Bowtie2 version 2.2.3 was used to build a reference genome index file and then map the trimmed reads to the reference genome. Samtools version 0.1.19 was used to convert to bam, sort and index files. Artemis (Rutherford *et al.* 2000) was used to visualize the mapped reads against the genome and calculate Reads Per Kilobase of transcript per Million mapped reads (RPKM) values. 'Transcript per million reads' (TPM) values (Li *et al.* 2010; Wagner *et al.* 2012) were calculated using the following formula:
TPM = (RPKM of feature/Total RPKM) x $10^6$

### 2.7.3 Promoter identification

Promoters were identified from the most abundant transcripts via visual comparison to a *Lactobacillus* promoter consensus sequence (Pouwels *et al.* 1993; Stephenson 2010) by searching upstream of identified genes, then checked against Perform Promoter Prediction online (PPP) (http://bioinformatics.biol.rug.nl/websoftware/ppp/) to identify upstream putative (Lactococcal-like) SigmaA binding sites (Hijum *et al.*).

### 2.7.4 Ribosomal binding sites

Ribosome binding sites (RBSs), or Shine-Dalgarno sequences, are important for protein expression because they allow the ribosome to bind to the upstream region of the ORF, allowing translation to take place (Shine *et al.* 1974; Kim *et al.* 1991; Fitzsimons *et al.* 1994; Hols *et al.* 1994; An *et al.* 2006; Wegmann *et al.* 2013). RBSs upstream of Open Reading Frames (ORFs) were identified manually by viewing the upstream sequence and via BLAST comparison to consensus sequences (Table 3).

Table 3: List of RBS sequences within *Lactobacillus.* These sequences were aligned against upstream regions of ORFs to find identify potential RBS.

| RBS sequence | ORF | Distance upstream ORF | Species | Reference |
|---|---|---|---|---|
| **AAGGAGG** | *Lactobacillus plantarum* conserved sequence | N/A | *Lactobacillus plantarum* | (Hols *et al.* 1994; Chen *et al.* 2005) |
| **AGGAGGA** | *gap* | 10 | *Lactobacillus delbrueckii* sub species *bulgaricus* | (Branny *et al.* 1998) |
| **AGAAAGGA** | *ldh* | 10 | *Lactobacillus casei* | (Kim *et al.* 1991) |
| **AAAGGGGG** | α-amylase | 9 | *Lactobacillus amylovorus* | (Fitzsimons *et al.* 1994) |
| **AGGAG** | *gap* | 7 | *Zymomonas mobilis* | (Conway *et al.* 1987) |
| **AGAGGAGTGA** | *ldhL* | 6 | *L. plantarum* | (Hols *et al.* 1994) |
| **AGGAGG** | *cbh* | 10 | *L. plantarum* | (Hols *et al.* 1994) |
| **AGAATGGAG** | *pyrE* | 8 | *L. plantarum* | (Hols *et al.* 1994) |
| **GGAGG** | *repA* | 6 | *L. casei* | (An *et al.* 2006) |

## 2.8 DEPC-treated water

RNAse-free water was made by adding 0.1% (v/v) Diethyl pyrocarbonate (DEPC) to ultra-pure water and mixed for 1 hour at 37°C. Water was then autoclaved for 30 minutes at 121°C on slow exhaust to destroy DEPC.

## 2.9 Plasmid DNA isolation

Plasmid DNA was isolated from *E. coli* using an ISOLATE II Plasmid Mini Kit (BIOLINE) as per the manufacturer's instructions. DNA was eluted from the column with 50 µL ultra-pure water. The list of plasmids and DNA used are in Table 4. DNA concentration was measured using a Nanodrop ND1000 spectrophotometer (ThermoFisher Scientific) unless otherwise stated.

Table 4: List of plasmids and DNA used in this study. Construction details and notable features of each sequence are described. All plasmids confirmed via Sanger sequencing. pVEcXLa and pVEcXLa derivatives sequences supplied in fasta format (Supplementary data 1 – RAST annotationSupplementary data 2 – pVEcXLa sequence, Supplementary data 3 - pVEcXLa-$P_{eft}$-*rnetB*, Supplementary data 4 – pVEcXLa-$P_{cwah}$-*rnetB*). Sequences will be uploaded to Genbank.

| Plasmid/DNA | Size (bp) | Modification/s (in order of appearance) | Replication | Characteristics | Notes | Source |
|---|---|---|---|---|---|---|
| pLa3_4 | 2,676 | N/A | *L. agilis* La3 | Highly stable in *L. agilis* La3, small size | Native *L. agilis* La3 plasmid | (Stephenson 2010) |
| pVEcLaSyn1 | 6,315 | N/A | *E. coli* | *ermR* ($Em^R$), $P_{cwah}$ promoter, signal sequence, *egfp* and transcriptional terminator in pUC57 (*EcoRV*) | Synthesised by Genscript | This study |
| pTRKH2 | 6,727 | N/A | *E. coli*, *L. agilis* La3 | *ermR* ($Em^R$), MCS | | (O'Sullivan *et al.* 1993) |
| pJIR750 | 6,568 | N/A | *E. coli* | *catP* ($Cm^R$) | | (Bannam *et al.* 1992) |
| pTRKH2:$P_{ldhL}$:gfp:TslpA(hom) | 7,920 | N/A | *E. coli*, *L. agilis* La3 | *egfp* under native $P_{ldhL}$ promoter | | (Stephenson *et al.* 2011) |
| pTRKH2-LCER | 11,547 | BDV5* ligated to pTRKH2 - *Xba*I/*Sac*I, BDV6* ligated - *Sac*I/*Pst*I, BDV7* ligated - *Sal*I, BDV8* ligated - *BamH*I/*Pvu*I | *E. coli*, *L. agilis* La3 | *ermR* ($Em^R$), chromosomal recombination in *L. agilis* La3: *catP* ($Cm^R$), *egfp* under native $P_{ldhL}$ promoter | | This study |
| pVEcXLa | 11,042 | BDV13* self-ligated, BDV14* ligated - *Sal*I/*Eag*I | *E. coli*, *L. agilis* La3 | *ermR* ($Em^R$), chromosomal recombination in *L. agilis* La3: *catP* ($Cm^R$), *creiLOV* | No promoter upstream of *creiLOV* | This study |
| pVEcXLa-$P_{psp}$ | 11,193 | BDV15* ligated to pVEcXLa - *BamH*I/*Eag*I | *E. coli*, *L. agilis* La3 | *ermR* ($Em^R$), chromosomal recombination in *L. agilis* La3: *catP* ($Cm^R$), *creiLOV* under native $P_{psp}$ promoter | | This study |
| pVEcXLa-$P_{gap}$ | 11,364 | BDV16* ligated to pVEcXLa - *BamH*I/*Eag*I | *E. coli*, *L. agilis* La3 | *ermR* ($Em^R$), chromosomal recombination in *L. agilis* La3: *catP* ($Cm^R$), *creiLOV* under native $P_{gap}$ promoter | | This study |
| pVEcXLa-$P_{eno}$ | 11,124 | BDV17* ligated to pVEcXLa - *BamH*I/*Eag*I | *E. coli*, *L. agilis* La3 | *ermR* ($Em^R$), chromosomal recombination in *L. agilis* La3: *catP* ($Cm^R$), *creiLOV* under native $P_{eno}$ promoter | | This study |
| pVEcXLa-$P_{eft}$ | 11,222 | BDV18* ligated to pVEcXLa - *BamH*I/*Eag*I | *E. coli*, *L. agilis* La3 | *ermR* ($Em^R$), chromosomal recombination in *L. agilis* La3: *catP* ($Cm^R$), *creiLOV* under native $P_{eft}$ promoter | | This study |
| pVEcXLa-$P_{cwah}$ | 11,270 | BDV19* ligated to pVEcXLa - *BamH*I/*Eag*I | *E. coli*, *L. agilis* La3 | *ermR* ($Em^R$), chromosomal recombination in *L. agilis* La3: *catP* | | This study |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | (Cm$^R$), *creiLOV* under native P$_{cwah}$ promoter | | |
| pJIR1470 | Unknown | N/A | *E. coli* | *tet*(M) | | (Roberts *et al.* 2001) |
| pTRKH2-*tet(M)* | 8,932 | BDV20* ligated to pTRKH2 - *Sal*I | *E. coli*, *L. agilis* La3 | *ermR* (Em$^R$), *tet*(M) (Tet$^R$) | | This study |
| rnetB_La3 | 907 | N/A | N/A | *L. agilis* La3 codon-optimised *rnetB (rnetB_La3)* | gBlock synthesised by idtDNA | This study |
| pVEcXLa-P$_{psp}$-*rnetB* | 11,715 | BDV21* ligated to pVEcXLa-P$_{psp}$ - *Eag*I/*Sal*I | *E. coli*, *L. agilis* La3 | *ermR* (Em$^R$), chromosomal recombination in *L. agilis* La3: *catP* (Cm$^R$), *rnetB_La3* under native P$_{psp}$ promoter | | This study |
| pVEcXLa-P$_{gap}$-*rnetB* | 11,886 | BDV21* ligated to pVEcXLa-P$_{gap}$ - *Eag*I/*Sal*I | *E. coli*, *L. agilis* La3 | *ermR* (Em$^R$), chromosomal recombination in *L. agilis* La3: *catP* (Cm$^R$), *rnetB_La3* under native P$_{gap}$ promoter | | This study |
| pVEcXLa-P$_{eno}$-*rnetB* | 11,646 | BDV21* ligated to pVEcXLa-P$_{eno}$ - *Eag*I/*Sal*I | *E. coli*, *L. agilis* La3 | *ermR* (Em$^R$), chromosomal recombination in *L. agilis* La3: *catP* (Cm$^R$), *rnetB_La3* under native P$_{eno}$ promoter | | This study |
| pVEcXLa-P$_{eft}$-*rnetB* | 11,744 | BDV21* ligated to pVEcXLa-P$_{eft}$ - *Eag*I/*Sal*I | *E. coli*, *L. agilis* La3 | *ermR* (Em$^R$), chromosomal recombination in *L. agilis* La3: *catP* (Cm$^R$), *rnetB_La3* under native P$_{eft}$ promoter | | This study |
| pVEcXLa-P$_{cwah}$-*rnetB* | 11,792 | BDV21* ligated to pVEcXLa-P$_{cwah}$ - *Eag*I/*Sal*I | *E. coli*, *L. agilis* La3 | *ermR* (Em$^R$), chromosomal recombination in *L. agilis* La3: *catP* (Cm$^R$), *rnetB_La3* under native P$_{cwah}$ promoter | | This study |

Em$^R$ = erythromycin resistance; Cm$^R$ = chloramphenicol resistance; Tet$^R$ = tetracycline resistance; MCS = multiple cloning site; * Refer to Table 5 for PCR product information, bp = base pair. pVEcXLa: Vezina, *E. coli*, *L. agilis* La3

## 2.10 Polymerase chain reaction (PCR)

### 2.10.1 Generation of PCR products for cloning and sequencing

Q5 polymerase (NEB) was used to generate amplicons for cloning and sequencing due to its high-fidelity proof-reading activity. PCR reactions were performed as per the manufacturer's instructions. Briefly, 25 µL Q5 High-Fidelity 2X Master Mix was mixed with 2.5 µL 10 µM forward primer, 2.5 µL 10 µM reverse primer, ≤10 ng DNA and ultra-pure water up to 50 µL. Primers and annealing temperatures ($T_A$) can be found in Table 5.
Thermocycler settings were as follows unless otherwise stated:

- 98°C – 30 seconds
- 35 x
    - 98°C – 10 seconds
    - $T_A$°C – 10 seconds
    - 72°C – 20 seconds / kb
- 72°C – 5 minutes
- 12°C – hold indefinitely

PCR products were purified using method 2.15.

### 2.10.2 Diagnostic PCRs

MyTaq Red (BIOLINE) was used to generate amplicons used for diagnostic PCR screening as per the manufacturer's instructions except that 10 µM primers were used instead of 20 µM. Briefly, 10 µL 2X MyTaq Red was mixed with 0.5 µL 10 µM forward primer, 0.5 µL 10 µM reverse primer, ≤10 ng DNA and ultra-pure water up to 20 µL. Primers and annealing temperatures ($T_A$) can be found in Table 5.

Thermocycler settings were as follows unless otherwise stated:

- 95°C – 1 minute
- 35 x
    - 95°C – 15 seconds
    - $T_A$°C – 15 seconds
    - 72°C – 30 seconds / kb
- 72°C – 5 minutes
- 12°C – hold indefinitely

Table 5: List of primers used in this study

| Target | Primer Pair | Primer names | Primer sequences | Template DNA | Product size (bp) | Product name | $T_A$ (°C) | Reference |
|---|---|---|---|---|---|---|---|---|
| **Left Flank** | 1 | L Plas L Flank Clone<br><br>L Flank *catP* Clone | CCAACTTAATCGCCTTGCAGTCTAGAGGTACGGATGCCATATTTAACCC<br><br>ACCACAGGTTAGTACAAAGACCTTGTGGAGCTCGAGGCTTAAAATACGTCTCTCAACACT | *L. agilis* La3 chromosome | 1,143 | BDV5 | 55 | This study |
| ***catP*** | 2 | L Flank *catP* RC Clone<br><br>*catP* ExpCass Clone | AGTGTTGAGAGACGTATTTTAAGCCTCGAGCTCCACAAGGTCTTTGTACTAACCTGTGGT<br><br>ATTTAACTAAACATGTGCATCCATCGGCTGCAGTTCTTTTTCGGCAAGTGTTCAAGAAGT | pJIR750* | 950 | BDV6 | 52 | This study |
| ***egfp* under $P_{ldhL}$ promoter** | 3 | GFP F<br><br>GFP R | AGCTGCGTCGACTCTAGACCTCTTTAACTAACAG<br><br>ACTGCAGTCGACGGATCCAAAAAAAGACAGAG | pTRKH2:$P_{ldhL}$:gfp: T*slpA*(hom)* | 1,044 | BDV7 | 48 | This study |
| **Right Flank** | 4 | ExpCass R Flank Clone<br><br>R Flank R Plas Clone | GTTGGATGCATAGCTTGAGTATTCTATAGGGATCCCATTATCAAAGTGGCAACTAACATACTATAAGG<br><br>CCTCTTCGCTATTACGCCAGCGATCGGATTAAGGTGAAAATCATTGACCGA | *L. agilis* La3 chromosome | 2,174 | BDV8 | 53 | This study |
| **Chromosomal integration site** | 5 | Crossover F<br><br>R Flank R Plas Clone | TAAGCATCCTTACTAGGGCGTAG<br><br>CCTCTTCGCTATTACGCCAGCGATCGGATTAAGGTGAAAATCATTGACCGA | *L. agilis* La3 chromosome | 3,309, 5,075 – 5,414 [a] | BDV9 | 53 | This study |
| **Detect presence of pTRKH2-LCER and pVEcXLa derivatives** | 6 | pTRKH2pres1 F<br><br>Crossover S R | TATGTTGTGTGGAATTGTGAGC<br><br>TTGCAGGAATTGATAAATAGTTAACTTCAG | pTRKH2-LCER* | 1,322 | BDV10 | 51 | This study |
| ***egfp* for Southern blot** | 7 | GFP MCS F<br><br>GFP MCS R | CGACTGACTAGTATGAGTAAAGGAGAAGAACTTTTCAC<br><br>GCACGCTTAATTAATTATTTGTAGAGCTCATCCATGCC | pTRKH2-LCER* | 743 | BDV11 | 53 | This study |
| **pTRKH2-LCER for Southern blot** | 8 | R5 seq2<br><br>pTRKH2-LCER Seq1 | AGGCACCCCAGGCTTTACAC<br><br>ATAACCACCGTATTTATAACTACGG | pTRKH2-LCER* | 168 | BDV12 | 51 | This study |
| **Inverse PCR of pTRKH2-LCER** | 9 | Inv F R Flank Inc | AGCACGGGATCCCATTATCAAAGTGGCAACT<br><br>ACGCAGCGGCCGATGAGTAAAGGAGAAGAACTTTTCACTGG | pTRKH2-LCER* | 11,403 | BDV13 | 55 | This study |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Inv R R Flank Inc | | | | | | |
| ***creiLOV*** | 10 | CreiLOV *Eag*I F<br>CreiLOV R *Sal*I | AGCTCACGGCCGGCTGGTTTACGGCACACTTT<br>ATCGCAGTCGACTTAAGCTAAAGCCTTACCTTCAGTCTTAC | CreiLOV46A | 381 | BDV14 | 54 | This study |
| **Putative secreted protein promoter ($P_{psp}$)** | 11 | $P_{psp}$ F *Bam*HI<br>$P_{psp}$ R *Eag*I | AGCGCAGGATCCCAATTTACGGATATCTTTTTTTGGTA<br>ATCGCACGGCCGAACCATATCATTTCCCCCTA | *L. agilis* La3 chromosome | 186 | BDV15 | 49 | This study |
| **$_{Gap}$ operon promoters ($P_{gap}$)** | 12 | $P_{gap}$ F *Bam*HI<br>$P_{gap}$ R *Eag*I | AGCTCAGGATCCTTGAAATTCAAAGTTAGAGATTATTTG<br>AGCTCACGGCCGCTTATGGCCTCTTTCTCGCG | *L. agilis* La3 chromosome | 357 | BDV16 | 48 | This study |
| **Enolase promoter ($P_{eno}$)** | 13 | $P_{eno}$ F *Bam*HI<br>$P_{eno}$ R *Eag*I | ATCGTAGGATCCATTAATTGAAAGTCTAAAGGAAAACC<br>ATCGCACGGCCGAGACATTGATTTTACTCTCCTTTGAGTTTTG | *L. agilis* La3 chromosome | 117 | BDV17 | 50 | This study |
| **Elongation factor Tu promoters ($P_{eft}$)** | 14 | $P_{eft}$ F BamHI<br>$P_{eft}$ R EagI | ACTGACGGATCCTTCATTTAAGTATTTACAAAGTTAATTAGCCTC<br>AGCTCACGGCCGCATTTAAAACGAACCTCCTGATATTTTGC | *L. agilis* La3 chromosome | 215 | BDV18 | 51 | This study |
| **Cell wall associated hydrolase promoter ($P_{cwah}$)** | 15 | $P_{cwah}$ F<br>$P_{cwah}$ R | AGCCATGGATCCGTTGAGATAAGAATTACAAGTCAACTTACC<br>ATCCGACGGCCGATCTTTCACTCCCGGATATTAATT | *L. agilis* La3 chromosome | 263 | BDV19 | 52 | This study |
| ***tet*(M)** | 16 | SalI *tetM* F<br>SalI *tetM* R | AGCTACGTCGACTTGTAATCATGTACTCTTTTTG<br>ATGCATGTCGACCCTTTACACTTTAATTCAAATC | pJIR1470* | 2,223 | BDV20 | 45 | This study |
| ***rnetB_La3*** | 17 | rnetB_La3 EagI F<br>rnetB_La3 SalI R | ACTGCACGGCCGATGAGTGAATTAAACGATATTAACA<br>AGCGCAGTCGACTTAGTAGTATTCAATTTTATGATCTTGCC | *rnetB*_La3 | 903 | BDV21 | 51 | This study |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| ***dnaK* qPCR** [b] | 18 | dnak_qPCR1 F<br><br>dnak_qPCR1 R | CTTGACGCTTAGCAACTTCC<br><br>CTTAGGGACAACTAACTCAGC | *L. agilis* La3 chromosome | 146 | BDV22 | 51 | This study |
| ***rnetB* qPCR** [b] | 19 | rnetB_La3_qPCR2 F<br><br>rnetB_La3_qPCR2 R | GGTAGTAAGTATTATGGTAAGATGAAGTGG<br><br>GCAATCTTAATGTTGTTGTTAACATCAGC | pVEcXLa-$P_{psp}$-rnetB, pVEcXLa-$P_{gap}$-rnetB, pVEcXLa-$P_{eno}$-rnetB, pVEcXLa-$P_{eft}$-rnetB, pVEcXLa-$P_{cwah}$-rnetB | 89 | BDV23 | 56 | This study |
| **Putative secreted protein qPCR** [b] | 20 | Gpsp_qPCR1 F<br><br>Gpsp_qPCR1 R | GGAATACGGAAGTTATCTGTAGGAGTAGC<br><br>GTATCAGCTAATACCGCATGATTGC | *L. agilis* La3 chromosome | 89 | BDV24 | 55 | This study |
| **GAPDH qPCR** [b] | 21 | Gap_qPCR1 F<br><br>Gap_qPCR1 R | GGAAGTTGTTGCTATCAACGAC<br><br>GCATCGAATTGGCCATGAGC | *L. agilis* La3 chromosome | 87 | BDV25 | 54 | This study |
| **Enolase qPCR** [b] | 22 | Eno_qPCR1 F<br><br>Eno_qPCR1 R | CGAAGTCTTAGACTCACGTGG<br><br>GCACCAGATGGCACGATACC | *L. agilis* La3 chromosome | 93 | BDV26 | 54 | This study |
| **Elongation factor Tu qPCR** [b] | 23 | Eft_qPCR1 F<br><br>Eft_qPCR1 R | CCACGTTAACATTGGTACAATCGG<br><br>GTAGTCAGAAGCTTCTGCCAATCC | *L. agilis* La3 chromosome | 109 | BDV27 | 55 | This study |
| **Cell wall associated hydrolase qPCR** [b] | 24 | Cwah_qPCR1 F<br><br>Cwah_qPCR1 R | GCTTCCTCAGCTAGCGTTAAC<br><br>GCTCGATTGATTGAATTGAAACGC | *L. agilis* La3 chromosome | 106 | BDV28 | 54 | This study |

$T_A$ = annealing temperature; * Refer to Table 4 for PCR product information, bp = base pair, [a] = dependent on recombinant *L. agilis* La3 strain used (Table 2); 3,309 bp for wild type, 5,075 bp for *L. agilis* La3::$P_{ldhL}$-EGFP, up to 5,414 bp for pVEcXLa variants; [b] qPCR = quantitative polymerase chain reaction.

## 2.11 qPCR

Single strand cDNAs (method 2.5.2) were used as qPCR template using SensiFAST SYBR No-ROX 2X mix (BIOLINE) in 20 µL reactions. qPCR was done using twin.tec® PCR Plates (Eppendorf), covered with Microseal 'B' PCR Plate Sealing Film (BIO-RAD). Reactions were thermocycled using 3-step qPCR amplification in a CFX Connect Real-Time PCR Detection System. Primer details and annealing temperatures can be found in Table 5. Each sample was run in technical triplicate.

Thermocycler settings were as follows unless otherwise stated:

- 95°C – 2 minutes
- 40 x
    - 95°C – 5 seconds
    - $T_A$°C – 10 seconds
    - 72°C – 5 seconds
- 12°C – hold indefinitely

### 2.11.1 *rnetB* normalised transcription levels

Normalised *rnetB* transcript levels were generated by using an internal control gene, which could be used to directly compare between samples. These values were generated by dividing the Cq value from *rnetB* (Primer Pair 17) by the Cq value of the internal control, *dnaK* (Primer Pair 18). This ratio was used to compare between samples. *L. agilis* La3 wild type was used as a negative control.

## 2.12 Restriction endonuclease digestions

DNA was digested with the following restriction endonucleases (New England Biolabs (NEB)) as per the manufacturer's instructions:

- *Sac*I-HF
- *Pst*I-HF
- *Bam*HI-HF
- *Pvu*I-HF
- *Xba*I
- *Sal*I-HF
- *Eag*I-HF
- *Hin*dIII-HF
- *Nco*I-HF

Briefly, 1 µL of restriction endonuclease was mixed with 1 µg DNA, 2 µL CutSmart Buffer, and made up to a total volume of 20 µL with ultra-pure water, unless otherwise stated.

Reactions were incubated at 37°C for up to 4 hours and then heat inactivated at an appropriate temperature or purified using method 2.15.

## 2.13 Phosphatase treatment

After digestion (method 2.12) and subsequent purification (method 2.15) (if heat inactivation was not possible), the plasmid vector was dephosphorylated using Antarctic Phosphatase (NEB) to prevent self-religation, as per the manufacturer's instructions. Briefly, 1.5 µL Antarctic Phosphatase was mixed with 4 µL 10X Antarctic Phosphatase Buffer, the entire restriction endonuclease reaction and made up to a total volume of 40 µL with ultra-pure water. This was incubated at 37°C for 45 minutes then heat inactivated at 80°C for 2 minutes.

## 2.14 Phosphorylation

To phosphorylate undigested PCR products, T4 Polynucleotide Kinase (NEB) was used as per the manufacturer's instructions. Briefly, 1 µL of T4 Polynucleotide Kinase was mixed with 5 µL T4 Polynucleotide Kinase Reaction Buffer, 5 µL 10 mM Adenosine triphosphate (ATP), 1 µg DNA and made up to a total volume of 50 µL with ultra-pure water. This was incubated at 37°C for 45 minutes then heat inactivated at 65°C for 20 minutes.

## 2.15 Purification of DNA from reaction mixtures

### 2.15.1 PCR and enzyme reactions

PCR products and DNA from enzymatic digestion were recovered and purified using an ISOLATE II PCR and Gel Kit (BIOLINE) as per the manufacturer's instructions, if heat inactivation was not appropriate. DNA was eluted from the column by 2 x 15 µL ultra-pure water washes.

### 2.15.2 Gel purification

Entire restriction endonuclease reactions were loaded into agarose gels and electrophoresed (method 2.3.1). The desired bands were excised using a scalpel blade and purified using an ISOLATE II PCR and Gel Kit (BIOLINE) as per the manufacturer's instructions. DNA was eluted from the column by 2 x 15 µL ultra-pure water washes.

## 2.16 Ligation

DNA was ligated using T4 DNA ligase (NEB) as per the manufacturer's instructions. Briefly, 2 µL 10X T4 DNA ligase Buffer was mixed with 100 ng digested (method 2.12) and dephosphorylated (method 2.13) plasmid vector, 100 ng digested insert, and made up to a total volume of 19 µL with ultra-pure water. The mixtures were heated to 55°C for 5 minutes then snap cooled on ice for 1 minute. 1.5 µL T4 DNA ligase was added to the mixtures and incubated. For sticky-end cloning, ligations were incubated at 21°C for 3 hours, then 16°C for 12 hours. For blunt-end cloning, ligations were incubated at 16°C for 15 hours. Reactions were heat inactivated at 65°C for 10 minutes. Plasmid vector to insert ratio was a 1:1 w/w ratio unless otherwise stated.

## 2.17 Competent *E. coli* cells

Competent *E. coli* JM109 cells were prepared using a published method (Inoue *et al.* 1990). Briefly, 10 large *E. coli* JM109 colonies were inoculated into 200 mL SOB media (0.5% (w/v) yeast extract (ThermoFisher Scientific), 2% (w/v) tryptone (ThermoFisher Scientific), 10 mM NaCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, pH 7.0) in a 2 L flask and grown at 250 RPM to $OD_{600}$ 0.5 at 18°C. The culture was transferred to an ice bath for 10 minutes, then centrifuged at 2,000 x g for 10 minutes at 4°C. Supernatant was completely removed and the pellet was resuspended in 100 mL ice-cold (1 – 4 °C) TB solution (250 mM KCl, 15 mM $CaCl_2$, 10 mM PIPES, 55 mM $MnCl_2$, pH 6.7) and incubated on ice for 10 minutes. Cells were centrifuged again and the supernatant removed. Cells were resuspended in 20 mL ice-cold STB solution (TB plus 7% DMSO) via gentle swirling of tubes then incubated on ice for 10 minutes. 200 µL of cells were dispensed into liquid nitrogen-cooled 1.5 mL microtubes and stored at -80°C. All buffer solutions were sterilised through a 0.22 µm Millex-GP Syringe Filter Unit (Merck) before use.

## 2.18 Transformation of *E. coli*

Competent *E. coli* cells were transformed as per Inoue *et al*. (Inoue *et al.* 1990). Briefly, DNA was added to a maximum v:v ratio of 1:10 to cells and incubated on ice for 30 minutes. Cells were heat-shocked in a water bath at 42°C for 50 seconds without agitation, then transferred to an ice bath for 5 minutes. 1 mL of SOC media (0.5% (w/v) yeast extract, 2% (w/v) tryptone, 10 mM NaCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose, pH 7.0) was added to cells then incubated at 37°C for 1 hour at 250 RPM. Cells were then plated on LB 1% agar plates with appropriate antibiotics.

## 2.19 Transformation of *L. agilis* La3

*L. agilis* La3 transformation was adapted and modified using two previously published methods (Mason *et al.* 2005; Stephenson *et al.* 2011). Overnight cultures of *L. agilis* La3

were grown (method 2.2.1) and cultures inoculated at a 1/6 dilution into fresh MRS liquid media containing 2% (w/v) glycine. This was grown at 37°C for 90 minutes. Cells were then centrifuged at 4,000 x g for 2 minutes at 4°C and the supernatant removed via pipetting, unless otherwise stated. If the pellet was too loose, cells were centrifuged again for another 2 to 3 minutes. Cells were kept on ice from this point onwards. Cells were then resuspended in 40 mL ice-cold ultra-pure water then centrifuged and the supernatant removed and discarded. The cell pellets were resuspended by adding ~5 mL of each resuspension solution to cells, which was mixed slowly via gentle flicking of tubes with intermittent incubation on ice as resuspension via pipetting significantly drops transformation efficiency.

Cells were then resuspended in ice-cold 50 mM EDTA and incubated for 5 minutes in an ice bath. Cells were then centrifuged and the supernatant removed. Cells were then resuspended in 40 mL 0.3 M sucrose, centrifuged and the supernatant removed and resuspended in 100 µL 0.3 M sucrose containing 3 µg plasmid DNA. This was added to an ice-cold Gene Pulser 0.2 cm electroporation cuvettes (BIO-RAD). Cells were electroporated using a Gene Pulser and Pulse Controller (BIO-RAD) as per the manufacturer's instructions, with the following conditions: 1.5 kV, 200 Ω and 25 µF. Cells were immediately diluted 1:20 into 1.3 mL of prewarmed MRS liquid media in 1.5 mL microtubes and incubated for 3 hours at 37°C. 200 µL was spread gently and slowly onto MRS plates containing appropriate antibiotics and incubated anaerobically (method 2.2.1) at 37°C for 24 to 72 hours. All solutions were sterilised through a 0.22 µm Millex-GP Syringe Filter Unit (Merck) before use.

## 2.20 Chromosomal DNA isolation from *L. agilis* La3

Chromosomal DNA was isolated from *L. agilis* La3 using a GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich) as per the manufacturer's instructions.

## 2.21 Crude DNA isolation

Plasmid and chromosomal DNA was isolated from *E. coli* and *L. agilis* La3 by growing 5 mL cultures (method 2.2) then centrifuging at 4,500 x g for 10 minutes. Cells were then resuspended in 1 mL ultra-pure water. ~250 $mm^3$ 0.1 mm zirconia/silica beads (Daintree Scientific) were added to 2 mL screw cap microtubes (SARSTEDT). The resuspended cells were then added to the tube and bead beaten at 6.5 m/s for 60 seconds using a FastPrep-24 Classic instrument (MP Biomedicals). Lysed cells were then centrifuged at 21,000 x g for 5 minutes and supernatant used for DNA analysis.

## 2.22 Induction and screening of chromosomal recombinants

*L. agilis* La3 chromosomal recombinants were obtained by inoculating colonies into 10 mL MRS liquid media (OXOID) with 10 µg/mL chloramphenicol. Cultures were screened for chromosomal recombination via PCR using Primer Pair 5 (Table 5) and if positive, 100 µL of culture was spread onto MRS agar plates containing 10 µg/mL chloramphenicol. Colonies were then patched onto MRS agar plates containing 10 µg/mL chloramphenicol and the same colonies were identically patched onto MRS agar plates containing 10 µg/mL erythromycin. Colonies that remained chloramphenicol resistant while being erythromycin sensitive demonstrated the expected phenotype of successful chromosomal recombination. Individual colonies were also screened for chromosomal recombination by PCR using Primer Pair 5 (Table 5).

## 2.23 Stability assays

### 2.23.1 Chromosomal insert stability

Stability of chromosomal inserts was determined by passaging *L. agilis* La3 chromosomal recombinants in MRS liquid media under no selection for ~100 generations. 100 µL of the final culture was serially diluted then plated onto MRS agar plates. Passaged colonies were then patched onto MRS agar plates containing 10 µg/mL chloramphenicol and the same colonies were identically patched onto MRS agar plates containing 10 µg/mL erythromycin.

#### 2.23.1.1 Phenotyping

Antibiotic resistance phenotyping was used to determine chromosomal insertion of expression cassette. If the colony phenotype showed chloramphenicol resistance and erythromycin sensitivity, this meant the cassette was successfully retained within the chromosome, as the chloramphenicol resistance gene is within the expression cassette, and the erythromycin resistance gene is found on the plasmid. If the phenotype was both chloramphenicol and erythromycin sensitive, the expression cassette would not be present within the chromosome.

#### 2.23.1.2 Genotyping

DNA was isolated from individual *L. agilis* La3 colonies passaged using method 2.21 then used in PCR reactions (method 2.10) using Primer Pair 5 (Table 5). The amplicon size showed whether or not the chromosomal insert was present (5,075 bp for insert versus 3,309 bp for wild-type *L. agilis* La3).

### 2.23.2 Plasmid stability

Plasmid stability was determined using the same method as described in 2.23.2.

#### 2.23.2.1 Phenotyping

Individual colony antibiotic resistance phenotype indicating plasmid presence was both chloramphenicol and erythromycin resistance.

#### 2.23.2.2 Genotyping

Method 2.10 was used except Primer Pair 6 (Table 5) was used in PCRs. The presence of a 1,322 bp amplicon indicated plasmid presence, and the absence of this amplicon indicated loss of the plasmid.

#### 2.23.2.3 Southern blot

Three independent Southern blots were performed. Details of these blots are outline in Table 6. Briefly, genomic DNA was isolated from *L. agilis* La3 (method 2.20 and 2.21) then 1 µg of DNA was digested using *Hin*dIII-HF (NEB) or *Hin*dIII-HF/*Nco*I-HF (NEB) (method 2.12). This DNA was electrophoresed (method 2.3.1). The gel was then capillary-transferred using a weighted paper towel stack, onto a Hybond N+ nylon membrane (Amersham Biosciences). To make the probe, the DIG High Prime DNA Labelling and Detection Starter Kit I (Roche) was used as per manufacturer's instructions. Two probes were generated, PCR products BDV11 and BDV12 (Table 5), and 2 µg was used in DIG-labelling reaction. A probe sensitivity dot blot was performed, and probes were used at 30 ng/mL concentration during hybridisation. The probe hybridisation temperature was based off the following formula: (49.82+0.41(%G+C)-(600/l)) - 25.

Table 6: Details and features of the three Southern blots used in this study. Information about expected fragment size and target are included.

| Probe name | Probe size (bp) | Chromosomal digest | Target | Expected fragment size (bp) | Hybridisation temperature (°C) |
|---|---|---|---|---|---|
| BDV11* | 743 | *Hin*dIII | Chromosome | 2,856 | 40 |
| BDV11* | 743 | *Hin*dIII/*Nco*I | Chromosome and pTRKH2-LCER | 498 and 2,358 | 37.2 |
| BDV12* | 168 | *Hin*dIII/*Nco*I | pTRKH2-LCER | 2,829 | 38.9 |

* Refer to Table 5 for PCR product information

## 2.24 Whole cell lysate protein preparation

Whole cell lysate was prepared from *E. coli* and *L. agilis* La3 by growing 5 mL cultures (method 2.2) then centrifuging at 4,500 x g for 10 minutes and resuspending cells in 1 mL 1X PBS (prepared with ultra-pure water. 250 $mm^3$ 0.1 mm zirconia/silica beads (BioSpec) was added to 2 mL screw cap microtubes (SARSTEDT). The resuspended cells were then added to the tube and bead beaten at 6.5 m/s for 60 seconds using a FastPrep-24 Classic instrument (MP Biomedicals). Lysed cells were then centrifuged at 2,000 x g for 10 seconds and the supernatant used for subsequent protein analysis.

## 2.25 Quantification of CreiLOV expression

### 2.25.1 *E. coli*

To quantitate CreiLOV expression whole cell lysate was prepared using method 2.24 with the following exceptions: freshly transformed *E. coli* were grown to mid logarithmic phase ($OD_{600}$ 0.4) and then cell density was normalised to $OD_{600}$ 0.3 in 1X PBS. Cells were then bead beaten (CreiLOV is expressed in the cytoplasm) and 200 µL of each sample was transferred to a 96-Well Black Cellstar Plate (Greiner) and CreiLOV expression levels were quantified using relative fluorescent units (RFUs) in technical triplicates. Fluorescence intensity was determined using a POLARstar Omega Plate Reader Spectrophotometer. The 360-10 excitation and 490-10 emission filters were used. The gain was set on the sample with highest fluorescence, and 1X PBS was used as a blank. Wild-type *E. coli* JM109 was used as a negative control.

### 2.25.2 *L. agilis* La3

Whole cell lysate was prepared using method 2.24 with the following exceptions: freshly transformed *L. agilis* La3 were grown in a strict anaerobic environment by boiling MRS liquid media before inoculation and lids kept loose in an anaerobic environment. Cells were grown to mid logarithmic phase ($OD_{600}$ 3.5) in biological triplicate then cell density was normalised to $OD_{600}$ 0.1 in 1X PBS. Cells were then bead beaten (CreiLOV is expressed in the cytoplasm) and diluted 1 x $10^{-1}$ due to high background fluorescence of *L. agilis* La3. 200 µL of each sample was transferred to a 96-Well Black Cellstar Plate (Greiner) and CreiLOV expression levels were quantified using relative fluorescent units (RFUs) in technical triplicates. Fluorescence intensity was determined using a POLARstar Omega Plate Reader Spectrophotometer. The 360-10 excitation and 490-10 emission filters were used. The gain was set on the sample with highest fluorescence, and 1X PBS was used as a blank. Wild type *L. agilis* La3 was used as a negative control.

### 2.25.3 Statistical analysis

Data were analysed using GraphPad Prism (version 7.02). RFUs were analysed using One-way ANOVA and the Dunnett test was done to correct for multiple comparisons. Linear regression analysis was used to correlate RFUs to TPM (method 2.7.2).

## 2.26 Protein gel electrophoresis

Whole cell lysate samples were prepared using method 2.24 and 5 µL was mixed with 1X sample buffer (0.0625 M Tris (pH 6.8), 2.5% w/v SDS, 0.002% bromophenol blue, 10% glycerol, 5% 2-mercaptoethanol in ultra-pure water) then heated at 70°C for 10 minutes and loaded into 4 to 20% Novex Wedgewell Tris-Glycine precast protein gels (ThermoFisher Scientific) as per the manufacturer's instructions. Precision Plus Protein Dual Xtra Prestained Protein Standard (BIO-RAD) was used as a ladder of standards of known molecular weight. Gels were electrophoresed at 125V for 2.5 hours. Gels were stained using a Coomassie Blue staining solution (0.1% (w/v) Coomassie Blue R250 (BIO-RAD), 10% acetic acid, 50% ethanol in water) for 2 hours at 40 RPM. Gels were destained using destain solution (10% ethanol, 10% acetic acid in water) till appropriate, then digitally scanned.

## 2.27 Western blot

Western blots were performed identically to method 2.26 however no staining was undertaken. After electrophoresis, protein was transferred to a PVDF membrane using an iBlot and iBlot Transfer Stack (ThermoFisher Scientific) as per the manufacturer's instructions. The membrane was then transferred to a 50 mL conical tube. All washes/incubations were performed at room temperature for 1 hour at 20 RPM unless otherwise stated. The membrane was then washed with 40 mL blocking solution (50 mM Tris base, 150 mM NaCl (pH 7.5), 0.05% Tween-20, 5% w/v skim milk powder). The blocking solution was discarded. To preabsorb non-specific, non-anti-rNetB antibodies a primary antibody solution was prepared by adding 50 µL *L. agilis* La3 wild type and *E. coli* JM109 wild type whole cell lysates (method 2.24) to fresh antibody solution (1/500 dilution of rabbit polyclonal IgG antibody serum (anti-rNetB) in blocking solution). This was incubated at room temperature for 1 hour then centrifuged at 3,500 x g for 10 minutes and the supernatant kept. The membrane was then washed with the non-specific antibody depleted primary antibody solution. The depleted primary antibody solution was then poured off and stored at -20°C and reused up to 4 times. The membrane was then washed 3 times for 10 minutes in 40 mL TBST (50 mM Tris base, 150 mM NaCl (pH 7.5) 0.05% Tween-20). The membrane was then incubated with the secondary antibody solution (1/8,000 dilution of goat polyclonal antibody serum (anti-rabbit IgG)-AP (ThermoFisher Scientific) in blocking solution). The secondary antibody solution was then poured off and stored at -20°C and reused up to 4 times. The membrane was then washed 3 times for 10 minutes in 40 mL TBS (50 mM Tris base, 150 mM NaCl (pH 7.5)) and the blot developed with 1.5 mL Western Blue (Promega) while light protected. When the desired

level of development had been achieved, the blot was washed 3 times with deionized water. The membranes were then digitally scanned.

# 3. Chapter 3 Results: Characterisation of *Lactobacillus agilis* La3

## 3.1 Introduction

*Lactobacillus agilis* is a bacterium that is found as part of the normal flora of the gastrointestinal tract of chickens (Mitsuoka 2002; Lan *et al.* 2003; Stephenson *et al.* 2010; Bujnakova *et al.* 2014; Chen *et al.* 2017) as well as humans (Molin *et al.* 1993) and other animals (Baele *et al.* 2001; Morita *et al.* 2007; Kajikawa *et al.* 2016). It has also been identified as a potential probiotic in chickens (Lan *et al.* 2003; Meimandipour *et al.* 2009; Baldwin *et al.* 2018). Utilising *L. agilis* La3 as a live vaccine antigen vector in poultry means it will potentially become part of the public food system. To further characterise the specific isolate chosen for the live vaccine, *L. agilis* La3, and to confirm its status as generally regarded as safe (GRAS), the genome was sequenced and analysed. Sequencing the *L. agilis* La3 genome allowed further development of *L. agilis* La3 as a potential live vaccine, firstly through identification of genes that may contribute to the strong colonisation capabilities of *L. agilis* La3 and secondly, by providing insights into potential sites within the genome for chromosomal integration. Having a complete genome is beneficial when attempting to find sites within the genome to insert DNA sequences, as informed integration sites can be chosen which reduce the hazard of disrupting gene expression and cellular function.

In addition to determining the genome sequence to further characterize *L. agilis* La3 and to identify potential constitutive promoters that could be used for recombinant protein expression, RNA sequencing was performed to investigate the transcriptome and to identify strong promoters. The transcriptome is the total quantity of transcribed RNA within a cell (Wang *et al.* 2009). Analyzing the transcriptome is a useful tool for looking at comparative transcription levels under different growth conditions (Martin *et al.* 2010; Wagner *et al.* 2013), and can therefore be used to identify promoters that could be used for recombinant protein expression. Transcripts found in high abundance relative to the rest of the transcriptome may be indicative of highly transcribed genes with potentially strong promoters, although transcript stability also determines relative transcript levels and hence the abundance of transcripts is not necessarily an exact correlate of transcription level. Fine-tuning of therapeutic protein expression by using a range of promoters of different strengths is required for optimal protein expression in live vaccines (Stephenson 2010; Siegl *et al.* 2013).

## 3.2 Growth curve

To identify constitutive promoters that could be used for therapeutic protein expression in potential live vaccine derivatives of *L. agilis* La3, RNA needed to be isolated from *L. agilis* La3 at two different growth phases: the mid-logarithmic and early stationary phases. This allowed for the identification of genes that are highly expressed during vital phases of cellular growth. To determine growth dynamics of *L. agilis* La3 and to determine the

growth phases for RNA and DNA isolation, a growth curve was generated (Figure 5). *L. agilis* La3 was grown in MRS media under anaerobic conditions at 42°C and samples taken at 30-minute intervals for 10 hrs. The mid-logarithmic phase was determined to correspond to ~ $OD_{600}$ 3.5 while early-stationary phase corresponded to ~$OD_{600}$ 5.7.

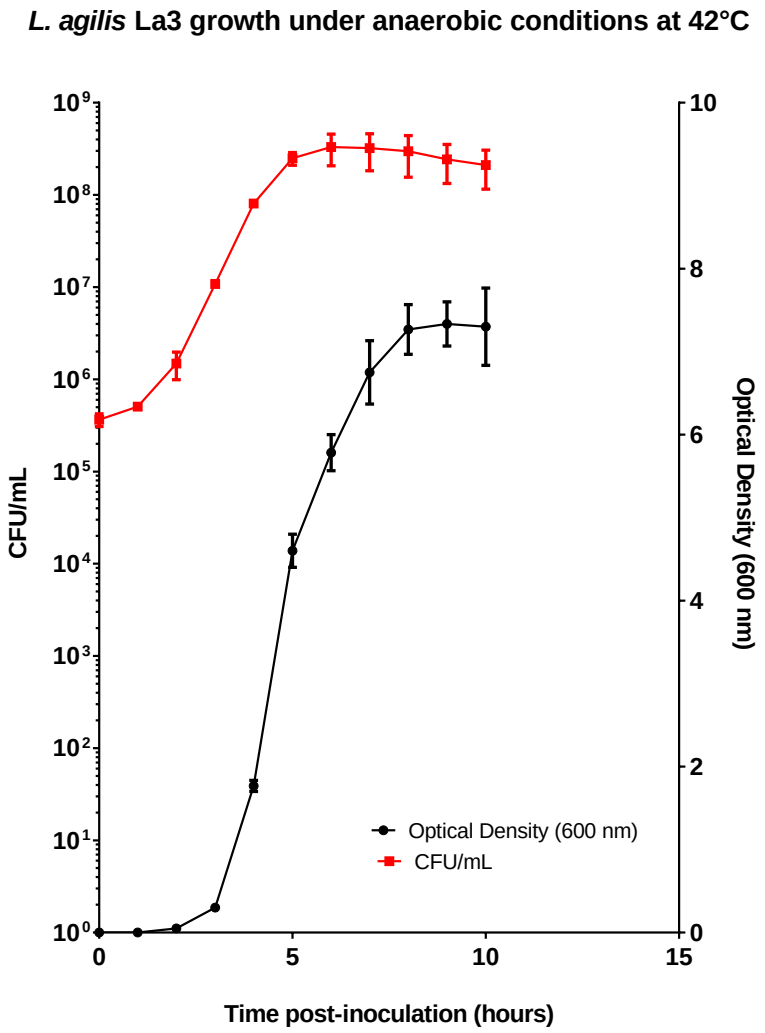


Figure 5: Growth curve of *L. agilis* La3 under anaerobic conditions at 42°C in 100 mL MRS media. The red line corresponds to CFU/mL, and the black line corresponds to Optical Density measured at 600 nm. The CFU/mL indicates bacterial growth in a closed system, with a lag, logarithmic, stationary and death phase clearly seen.

## 3.3 Genome sequencing and analysis of *L. agilis* La3

### 3.3.1 PacBio sequencing and assembly

To analyse the genome of *L. agilis* La3, high molecular weight DNA was successfully isolated from this strain at the mid-logarithmic growth phase (Figure 6) and sequenced by PacBio RS II (PacificBioSciences). A total of 3,274 reads (88,766,550 bp) were generated. The Canu assembler generated a read-corrected mean length of 26,880 bp and an N50 of 46,468. A 2,211,023 bp *L. agilis* La3 genome was assembled using 88,005,605 bp, indicating 39.8X coverage.

BLAST results of the first 50,000 bp and last 50,000 bp of the assembled DNA indicated that the terminal sections of the assembled molecule were highly homologous, with greater than 99% identity (24,045 out of 24,075 bp), indicating a large overlap. The few mismatches were most likely caused by erroneous 'deletions' called by the assembly software using the PacBio reads, due to the nature of SMRT sequencing which sometimes skips bases. 24,075 bp was therefore deleted from the end of the molecule, and the genome data was reconfigured to give a circular molecule. The MiSeq sequencing run produced 588,547,918 raw bases and 2,736,962 raw reads were generated. This was submitted to Genbank and has the accession number CP016766. The *L. agilis* La3 genome is ~2.2 Mb, indicating ~267-fold coverage. *L. agilis* La3 contains 4 native plasmids (pLa3_1, pLa3_2, pLa3_3 and pLa3_4) which were previously sequenced (Stephenson 2010). After correcting the sequence with PacBio reads and additional short-read error correction using the MiSeq data, the assembled *L. agilis* La3 genome produced a single, circularised 2,187,209 bp sequence with 41.3% GC content. The eight ribosomal RNA gene operons within the genome have a 51% GC content.

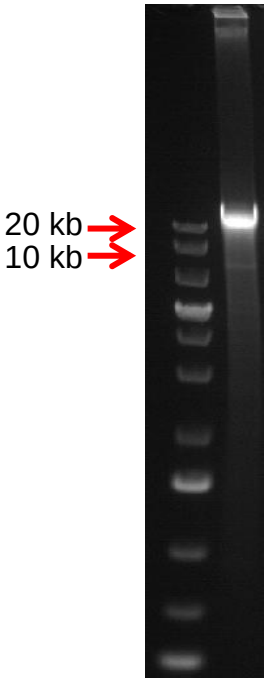


Figure 6: DNA gels showing the high molecular weight DNA extracted from *L. agilis* La3 after purification using the Agencourt AMPure XP kit. Ladder is Generuler 1kb Plus (ThermoFisher Scientific). Partial degradation of the high molecular weight DNA is present (demonstrated via the smearing of DNA at lower molecular weights).

### 3.3.2 Annotation and composition of the *L. agilis* La3 genome

The NCBI Prokaryotic Annotation Pipeline produced 2,055 CDS, compared to RAST, which produced 2,090 CDS. While both annotation pipelines will identify most genes, each can still miss potentially hundreds of genes due to differing parameters between the two (Chen *et al.* 2017). However, since they both produced a similar number of putative genes, the RAST annotation was used for analysis of gene function as it identified a higher number of potential genes (Supplementary data 1). Within the RAST SEED subsystem, 48% of CDS (990) could be assigned a broad gene function (Table 7). Broad functional groups are curated manually (Overbeek *et al.* 2005), which explains why many of these annotated genes are not classified in the SEED subsystem yet.

Table 7: SEED subsystem results of RAST annotated genes which clusters genes into broad functional groups. Several genes appear in multiple subsystems due to performing several functions.

| SEED Subsystem feature | Number of genes |
|---|---|
| Carbohydrates | 190 |
| Protein metabolism | 173 |
| Amino acids and derivatives | 149 |
| Cell wall and capsule | 125 |
| DNA metabolism | 107 |
| RNA metabolism | 102 |
| Nucleosides and nucleotides | 83 |
| Cofactors, vitamins, prosthetic groups, pigments | 67 |
| Fatty acids, lipids and isoprenoids | 65 |
| Cell division and cell cycle | 45 |
| Motility and chemotaxis | 43 |
| Membrane transport | 42 |
| Virulence, disease and defense | 40 |
| Phages, prophages, transposable elements, plasmids | 33 |
| Stress response | 32 |
| Miscellaneous | 28 |
| Phosphorus metabolism | 23 |
| Regulation and cell signalling | 21 |
| Respiration | 18 |
| Potassium metabolism | 5 |
| Sulphur metabolism | 5 |
| Dormancy and sporulation | 1 |
| Metabolism of aromatic compounds | 1 |
| Photosynthesis | 0 |
| Iron acquisition and metabolism | 0 |
| Secondary metabolism | 0 |
| Nitrogen metabolism | 0 |

### 3.3.3 SEED subsystem features

#### 3.3.3.1 Virulence disease and defense

Closer analysis of the SEED subsystem feature 'Virulence, disease and defense' showed no known virulence or superantigen genes, but genes mainly in the categories of resistance to antibiotics and toxic compounds (27 putative genes), invasion and intracellular resistance (11 putative genes), and adhesion (2 putative genes). The genome also possesses at least 31 predicted flagellar motility genes. Flagella play an important role in movement towards favourable environments and nutrient acquisition (Zhao *et al.* 2007). Flagella are an important characteristic of *L. agilis* (Weiss *et al.* 1981) and most likely improve the colonisation potential of *L. agilis* La3, probably allowing movement through the mucosal layer of the gastrointestinal tract of the chicken.

#### 3.3.3.2 Carbohydrates

The carbohydrate gene class responsible for sugar and energy metabolism had the highest proportion of genes, 190 genes in total. The presence of several notable genes includes L and D-lactate dehydrogenase genes, which indicates *L. agilis* La3 utilises anaerobic glycolysis pathways, resulting in lactic acid production (Dave *et al.* 2015). *L. agilis* La3 may have alterative fermentation pathways due to the presence of genes such as an alcohol dehydrogenase gene, indicating the capacity for ethanol and butanol metabolism.
A variety of sugar metabolism genes were identified, such as an maltose/maltodextrin ABC transporter gene, and genes encoding a permease protein MalF, a fructokinase and a trehalose phosphorylase indicate that *L. agilis* La3 is likely also capable of sucrose, maltose and maltodextrin utilisation, along with trehalose uptake and utilization. Beta-glucoside metabolism is also suggested, as well as mannitol, glycogen, mannose and fructose utilisation due to genes such as alpha-amylases, mannitol-1-phosphate 5-dehydrogenase, glycogen phosphorylase and mannose-6-phosphate isomerase.

#### 3.3.3.3 Amino acid metabolism and synthesis

149 genes were found to encode amino acid metabolism and synthesis proteins. Due to the presence of genes including tRNA-Cys-GCA, tRNA-Pro-CGG, tRNA-Gly-CCC, as well as genes such as glutamine synthetase type I, *L. agilis* La3 is predicted to be able to synthesise cysteine, proline, glycine, arginine, serine, leucine and tryptophan. The presence of tRNA aminoacylation genes for each amino acid indicate the capacity for utilisation of every amino acid.

#### 3.3.3.4 Nucleotide metabolism and synthesis

83 genes were found to encode nucleosides and nucleotides. The presence of genes such as deoxyuridine 5'-triphosphate nucleotidohydrolase indicate thymidine nucleotide synthesis capabilities. The presence of a uridine phosphorylase-encoding gene indicates a capacity for *de novo* biosynthesis of pryrimidines (cytosine, thymine and uracil). The presence of adenine phosphoribosyltransferase and hypoxanthine-guanine phosphoribosyltransferase genes indicates purine (adenine and guanine) synthesis is possible via nucleotide salvage pathways.

#### 3.3.3.5 Capsule production

125 genes were found to encode cell wall and capsule proteins. *L. agilis* La3 is most likely capable of capsule production due to presence of many capsule-related genes such as those encoding undecaprenyl-phosphate galactosephosphotransferase, dTDP-glucose 4,6-dehydratase, glutamate racemase, and N-acetylglucosamine-1-phosphate uridyltransferase.

### 3.3.4 Comparison to other *L. agilis* strains

The genome composition of *L. agilis* La3 is similar to that of two other *L. agilis* strains that were published by others (Drissi *et al.* 2015; Sun *et al.* 2015) with respect to genome size, number of genes and GC content (Table 8). The NCBI annotations were used to compare the strains as they were already annotated using the NCBI Prokaryotic Genome Annotation Pipeline. The large core genome is partially explained by the low number of strains used for this analysis and the 70% similarity cut-off. The unique genes of *L. agilis* La3 are mostly comprised of unknown function, phage genes, integrases and transposases. This demonstrates the high similarity between these strains.

#### 3.3.4.1 *L. agilis* La3 unique genes

A total of 303 genes were unique to *L. agilis L. agilis* La3 (Table 8). One unique gene of note, found in *L. agilis* La3 but not the other strains, encodes a mucin binding protein (BEN83_01135). The highest unique match on Genbank was 98% query cover and 46% identity from *L. salivarius* subsp. *salivarius* UCC118 (Accession number: WP_011475677.1) (date of search: 1/3/2017). *L. salivarius* UCC118 was isolated from a human gastrointestinal tract (Dunne *et al.* 1999). When analysed using HMMER (Finn *et al.* 2015), the protein encoded by this gene is predicted to contain a MucBP domain, LPXTG cell wall anchor motif and two small transmembrane regions. This may indicate specialisation of *L. agilis* La3 to chicken mucin.

Table 8: Comparison of *L. agilis* La3 to other sequenced *L. agilis* strains using 70% similarity as a cut-off. All strains share a similarly sized genome and similar number of genes. Despite being isolated from different sources, they have a large core genome, small number of accessory genes, and all carry 200 to 300 unique genes.

| Name | Origin | Length of genome (bp) | Genes | Core genome | Accessory genome | Unique genes | Absent genes | GC (%) | Plasmids | Reference | Accession number |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *L. agilis* La3 | Broiler chicken isolate (Stephenson *et al.* 2010) | 2,187,209 | 2,055 | 1,510 | 141 | 303 | 80 | 41.3 | 4 (Stephenson 2010) | This study | CP016766 |
| *L. agilis* DSM 20509 | Municipal sewage (Weiss *et al.* 1981) | 2,058,058 | 2,064 | 1,510 | 158 | 217 | 63 | 41.7 | None reported | (Sun *et al.* 2015) | GCA_001436215.1 |
| *L. agilis* Marseille | Stool sample of child with kwashiorkor, 2015 | 2,134,779 | 2,093 | 1,510 | 143 | 263 | 78 | 41.7 | None reported | (Drissi *et al.* 2015) | GCA_001243975.1 |

### 3.3.5 Bacteriocin mining

To determine if the reported probiotic role of *L. agilis* was due to bacteriocin production, the genome was mined for genes encoding bacteriocins using BAGEL3. No known bacteriocins or proteins homologous to those found within bacteriocin encoding operons (e.g. immunity or processing proteins) were present within the genome. *L. agilis* La3 could potentially be producing unusual bacteriocins or other antimicrobials, however these were not detected by genome mining which is limited by the completeness of the database used.

### 3.3.6 Methylation

The methylation profile of *L. agilis* La3 (Table 9) showed six different methylation motifs based on the DNA modifications indicated in the PacBio sequencing data. The known Restriction Modification (RM) systems present in *L. agilis* La3 are shown in Table 10, identified from genome mining. Three DNA methyltransferases and eight endonucleases were identified from the *L. agilis* La3 genome annotation data.

Table 9: Methylated motifs and their corresponding frequencies within the genome. Underlined bases represent the methylated base within the motif.

| Methylated motif | Motif frequency in genome | Motifs methylated in genome |
|---|---|---|
| CGATG | 3221 | 99.97% |
| ACRGA | 4910 | 99.90% |
| CNNTNNCNCNNTTCGNA | 12 | 91.67% |
| ANNGTNANANANCAG | 17 | 76.47% |
| CGATGT | 667 | 56.82% |
| ANTTKGTANYR | 255 | 33.73% |

Table 10: List of Restriction modification genes found in the *L. agilis* La3 genome based on the annotation.

| Name | RM gene type | Position in genome |
|---|---|---|
| DNA methyltransferase | DNA methyltransferase | 365,147 – 366,157 |
| DNA methyltransferase | DNA methyltransferase | 921,395 – 921,072 |
| DNA methylase | DNA methyltransferase | 1,148,209 – 1,147,535 |
| endonuclease | DNA endonuclease | 339,479 – 339,799 |
| HNH endonuclease | DNA endonuclease | 571,095 – 569,734 |
| restriction endonuclease | DNA endonuclease | 783,983 – 784,324 |
| endonuclease III | DNA endonuclease | 907,727 – 907,083 |
| endonuclease MutS2 | DNA endonuclease | 1,210,398 – 1,208,038 |
| HNH endonuclease | DNA endonuclease | 1,435,254 – 1,436,612 |
| restriction endonuclease | DNA endonuclease | 1,743,753 – 1,742,818 |
| HNH endonuclease | DNA endonuclease | 1,860,612 – 1,859,251 |

### 3.3.7 Prophage gene identification

Eight prophage-encoding regions within the *L. agilis* La3 genome were identified (Table 11). Only one complete prophage was present (prophage region 2). No known homologies to other prophages were found using PHAST. The GC content of these regions is very similar to the general GC content of *L. agilis* La3 (41.3%).

Table 11: Prophage regions identified within the *L. agilis* La3 chromosome using Phast

| Prophage region number | Size (bp) | GC content (%) | Predicted intact (Y/N) | Position in genome |
|---|---|---|---|---|
| 1 | 32,856 | 38.02 | N | 312,919 – 345,774 |
| 2 | 52,627 | 40.06 | Y | 630,214 – 682,840 |
| 3 | 11,416 | 41.13 | N | 694,449 – 705,864 |
| 4 | 14,112 | 37.9 | N | 958,729 – 972,840 |
| 5 | 34,887 | 34.96 | N | 974,163 – 1,009,049 |
| 6 | 11,189 | 43.54 | N | 1,461,618 – 1,472,806 |
| 7 | 9,003 | 42.33 | N | 1,831,912 – 1,840,914 |
| 8 | 10,453 | 41.22 | N | 2,091,183 – 2,101,635 |

### 3.3.8 Codon usage

Codon usage within the CDS's of *L. agilis* La3 was investigated (Table 12). Table 13 shows the *L. agilis* La3 codons most commonly used to encode each amino acid. Codon usage was determined to address gene expression in *L. agilis* La3 so *L. agilis* La3-specific codon-optimised genes could be utilised for expression of recombinant proteins. This will be explained in greater detail in Chapter 5.

Table 12: Codon usage within the *L. agilis* La3 chromosome. Calculated using cusp.

| **Amino acid** | **Codon** | **Codon preference (%)** | **Number of times used in CDS** |
|---|---|---|---|
| Alanine | GCA | 21.5 | 11,792 |
| Alanine | GCC | 28.8 | 15,772 |
| Alanine | GCG | 13.1 | 7,191 |
| Alanine | GCT | 36.6 | 20,069 |
| Cysteine | TGC | 36.7 | 1,504 |
| Cysteine | TGT | 63.3 | 2,598 |
| Aspartate | GAC | 39.1 | 13,310 |
| Aspartate | GAT | 60.9 | 20,764 |
| Glutamate | GAA | 76.1 | 29,079 |
| Glutamate | GAG | 23.9 | 9,133 |
| Phenylalanine | TTC | 26.8 | 7,085 |
| Phenylalanine | TTT | 73.2 | 19,385 |
| Glycine | GGA | 18.5 | 7,581 |
| Glycine | GGC | 23.4 | 9,584 |
| Glycine | GGG | 25.2 | 10,327 |
| Glycine | GGT | 33 | 13,517 |
| Histidine | CAC | 55.8 | 6,423 |
| Histidine | CAT | 44.2 | 5,095 |
| Isoleucine | ATA | 8 | 3,230 |
| Isoleucine | ATC | 28.5 | 11,542 |
| Isoleucine | ATT | 63.5 | 25,744 |
| Lysine | AAA | 46.3 | 21,731 |
| Lysine | AAG | 53.7 | 25,193 |
| Leucine | CTA | 14.4 | 9,731 |
| Leucine | CTC | 6.6 | 4,444 |
| Leucine | CTG | 7.1 | 4,785 |
| Leucine | CTT | 10.8 | 7,343 |
| Leucine | TTA | 41.7 | 28,232 |
| Leucine | TTG | 19.5 | 13,236 |
| Methionine | ATG | 100 | 14,604 |
| Asparagine | AAC | 47 | 14,066 |
| Asparagine | AAT | 53 | 15,879 |
| Proline | CCA | 36 | 7,654 |
| Proline | CCC | 16.2 | 3,441 |
| Proline | CCG | 17 | 3,611 |
| Proline | CCT | 30.9 | 6,576 |
| Glutamine | CAA | 72.8 | 21,563 |
| Glutamine | CAG | 27.2 | 8,068 |

| | | | |
|---|---|---|---|
| Arginine | AGA | 14.4 | 3,684 |
| Arginine | AGG | 9.3 | 2,388 |
| Arginine | CGA | 7.9 | 2,029 |
| Arginine | CGC | 20.4 | 5,229 |
| Arginine | CGG | 27.3 | 6,997 |
| Arginine | CGT | 20.6 | 5,273 |
| Serine | AGC | 18 | 6,703 |
| Serine | AGT | 29.5 | 10,990 |
| Serine | TCA | 18.6 | 6,924 |
| Serine | TCC | 10.3 | 3,839 |
| Serine | TCG | 7.5 | 2,804 |
| Serine | TCT | 16 | 5,971 |
| Threonine | ACA | 18.6 | 6,848 |
| Threonine | ACC | 30.8 | 11,378 |
| Threonine | ACG | 16.5 | 6,081 |
| Threonine | ACT | 34.1 | 12,605 |
| Valine | GTA | 21.3 | 9,672 |
| Valine | GTC | 18.4 | 8,369 |
| Valine | GTG | 19 | 8,630 |
| Valine | GTT | 41.2 | 18,695 |
| Tryptophan | TGG | 100 | 6,209 |
| Tyrosine | TAC | 47.2 | 11,887 |
| Tyrosine | TAT | 52.8 | 13,324 |
| Stop | TAA | 66.6 | 1,407 |
| Stop | TAG | 24.1 | 510 |
| Stop | TGA | 9.3 | 196 |

Table 13: Most frequently used codon for each amino acid based on *L. agilis* La3 codon usage.

| Amino acid | Optimal codon |
|---|---|
| Alanine | GCT |
| Cysteine | TGT |
| Aspartic acid | GAT |
| Glutamic acid | GAA |
| Phenylalanine | TTT |
| Glycine | GGT |
| Histidine | CAC |
| Isoleucine | ATT |
| Lysine | AAG |
| Leucine | TTA |
| Methionine | ATG |
| Asparagine | AAT |
| Proline | CCA |
| Glutamine | CAA |
| Arginine | CGG |
| Serine | AGT |
| Threonine | ACT |
| Valine | GTT |
| Tryptophan | TGG |
| Tyrosine | TAT |

## 3.4 Transcriptome analysis of *L. agilis* La3

RNA was isolated from *L. agilis* La3 from both the mid-logarithmic and early-stationary phases of growth (Figure 5, Figure 7). Isolating RNA from the mid-logarithmic and early-stationary phases was performed to allow for the identification of constitutive, highly transcribed genes that are expressed in two vital phases of growth. Constitutive, highly transcribed genes could be housekeeping genes or other highly expressed genes. Genes that also fit this expression pattern may also be under the control of inducible promoters, which may be induced by media components *in vitro*. It is therefore vital once these promoters are identified to test these promoters in a different environment, such as within an animal model, to determine if they behave in the same way under other conditions. After RNA sequencing, Transcripts per Million Reads (TPM) values (Wagner *et al.* 2012) were generated and the most abundant gene transcripts were identified (Figure 8). In general, the transcripts with the highest abundance at both stages of growth were within the glyceraldehyde-3-phosphate dehydrogenase (*gap*) operon. Each gene was highly transcribed. These genes and their functions are conserved amongst many close and distantly related species, from the bacterium *Staphylococcus aureus* (Purves *et al.* 2010) to the common capsicum *Capsicum annuum (Petersen et al. 2002)*.

The most abundant transcripts across both mid-logarithmic and early-stationary phases were from genes encoding glyceraldehyde-3-phosphate dehydrogenase (GAPDH), triosephosphate isomerase, elongation factor Tu, phosphoglycerate kinase, enolase and a putative secreted protein (Figure 8). Most of these genes are implicated in host binding, adhesion, and are associated with secretion and the cell wall surface of *Lactobacillus* species (Granato *et al.* 2004; Antikainen *et al.* 2007; Glenting *et al.* 2013; Kinoshita *et al.* 2016). The lactic-acid producing L-lactate dehydrogenase gene was the sixteenth most abundant transcript. Production of lactic acid is a defining characteristic of lactobacilli and has been implicated in the modulation of local environments via pH (Lan *et al.* 2003; Meimandipour *et al.* 2009; Chen *et al.* 2017).

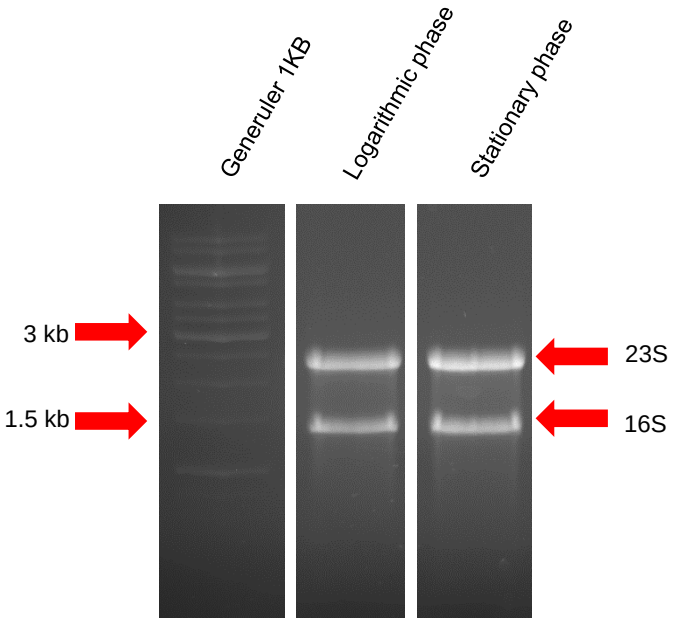


Figure 7: Formaldehyde gel showing RNA extraction and purification from *L. agilis* La3 at two different growth phases; logarithmic and stationary. Lane 1: Generuler 1 KB. Lane 2: mid-logarithmic growth phase. Lane 3: Early-stationary growth phase. Clear 23S and 16S bands indicate good quality RNA.

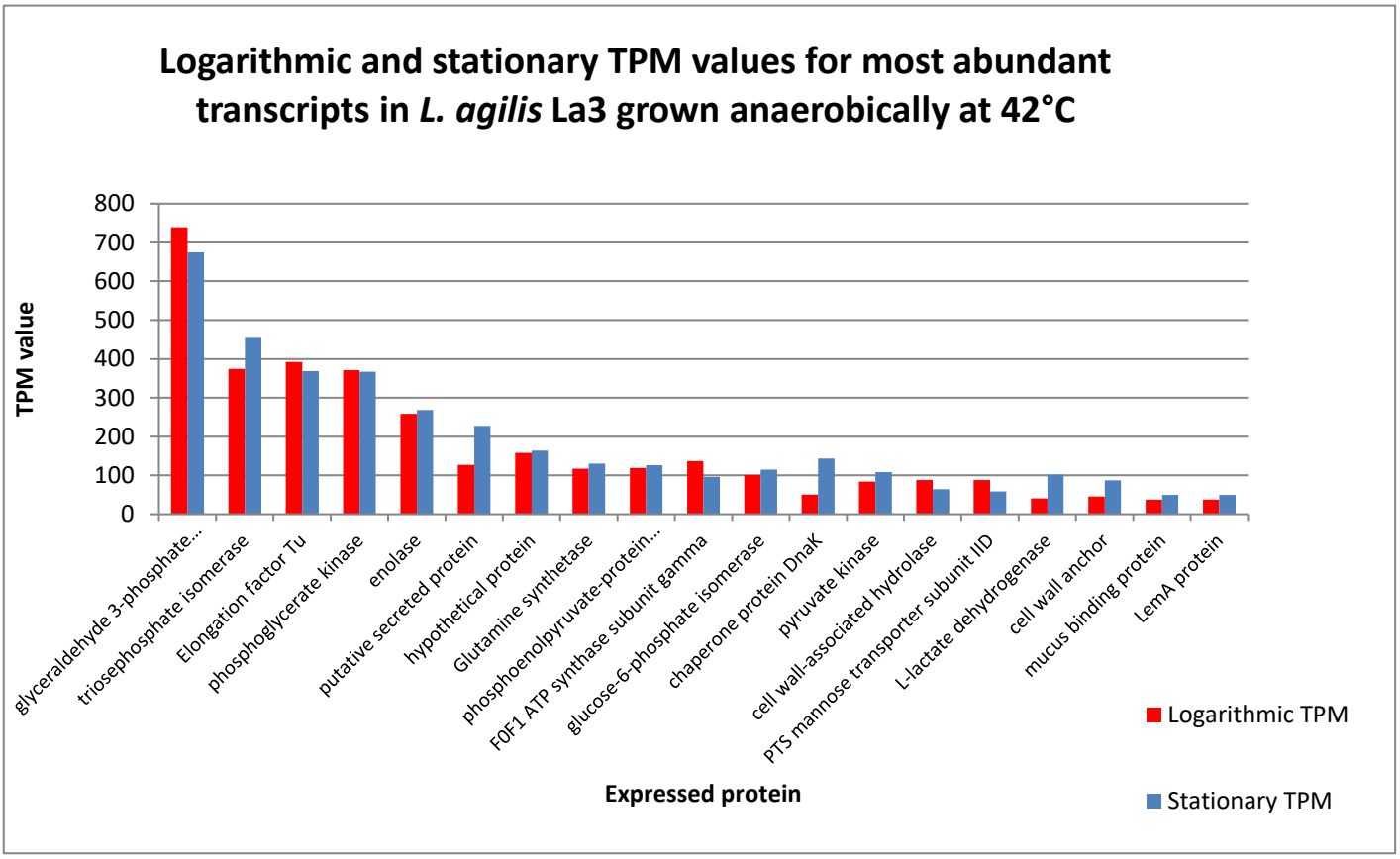


Figure 8: Graph showing the TPM of the most abundant gene transcripts present in *L. agilis* La3 when grown anaerobically at 42°C in MRS. Red bars represent mid-logarithmic transcript abundance and blue bars represent early-stationary transcript abundance.

### 3.4.1 Promoter identification

Putative promoters were identified upstream of the genes that had the most abundant transcripts (Table 14), where possible, based on comparison to consensus sequence motifs (Table 3) and the PPP toolbox software. Some of the most abundant transcripts fell within possible operons and it was predicted that they relied on the operon promoter for expression. Some promoter sequences upstream of non-operonic genes were unable to be identified as they did not match the *Lactobacillus* consensus sequence and were not found using PPP. A range of promoters from highly transcribed genes were chosen for future work (Chapter 5) to determine the relationship between transcript abundance and protein expression in *L. agilis* La3 (Table 14) through the use of a heterologous reporter gene.

Table 14: Predicted promoter sequences based on the *Lactobacillus* consensus sequence and PPP online portal. Promoters were identified from the most abundant transcripts, where possible.

| Operon/ORF | -35 region | TG region | -10 region | RBS | Start codon | Notes | Source |
|---|---|---|---|---|---|---|---|
| Consensus sequence | TTGACA | Variable (15 – 20 bp) | TATAAT | Variable (GA rich) | AUG/GUG | *Lactobacillus* consensus sequence | (Pouwels *et al.* 1993) |
| GAPDH | TTGAAA<br>TTGAAA<br>CTTGCA<br>TTAACT | TTCAAAGTTAGAGATTATTT<br>GGAAGCAAGTTTAAAG<br>AAATATTTCTTTTGTTGCTAT<br>GATAGGTAGCGATTGT | GTTAAT...286 bp...<br>TAAAAT...222 bp...<br>AATCTT...116 bp...<br>TATTAT...62 bp... | AAAGAGG...7 bp... | AUG | 4 potential upstream promoters, start of *gap* operon | This study |
| Elongation factor Tu | TTTACA<br>TTGGTT | AAGTTAATTAGCCTCTA<br>TTCTCTTAGGAAAGTAG | TATAAT...130 bp...<br>TATAAT...55 bp... | CAGGAGG...10 bp... | AUG | 2 possible upstream promoters | This study |
| Enolase | TTGAAA | GTCTAAAGGAAAACCTT | TAGAAT...37 bp... | AAGGAGA...9 bp... | AUG | Inside *gap* operon | This study |
| Putative secreted protein | TGGTAA<br>CTGACA | ATTGGGAGAAAATTATAAT<br>AGCCGGATCCAACGGTTA | TATAAT...307 bp...<br>AAGACA...88 bp... | AGAAGGA...12 bp... | GUG | 2 possible upstream promoters and inside operon | This study |
| Cell wall associated hydrolase | TTGTAA | CCTTTTTTGTACCCTAATGCAA | TATTAA...37 bp... | GGGAGTGAA...6 bp... | TTG | | This study |

## 3.5 Chapter 3 Discussion:

### 3.5.1 Genome composition

*L. agilis* La3 has demonstrated strong colonisation and persistence characteristics within the chicken gastrointestinal tract (Stephenson *et al.* 2010; Vezina 2013) despite the presence of well-established microflora (Lu *et al.* 2003; Saengkerdsub *et al.* 2007; Danzeisen *et al.* 2011). In this work the genome of *L. agilis* La3 was sequenced and genome analysis revealed a 2,187,209 bp genome with 2,055 genes. Analysis of the genome indicated a possible genetic basis that underpins the colonisation and persistence capabilities of *L. agilis* La3 due to the presence of genes encoding proteins likely to be involved in adhesion, invasion, capsule and motility. While many of these genes discussed here are common within *Lactobacillus* genus, genomic analysis is limited when explaining the superior colonisation characteristics of *L. agilis* La3. However, these results tentatively support the suitability of *L. agilis* La3 for development as a live vaccine vector for use in commercial broiler chickens. Food safety is a primary concern for a bacterial isolate that may be used as a live vaccine vector in commercial food production. The genome annotation did not identify any known virulence factors within the *L. agilis* La3 genome (Table 7). These results are consistent with previous literature which shows that *L. agilis* is a human coloniser (Molin *et al.* 1993; Drissi *et al.* 2015), and no evidence of *L. agilis* causing disease has been shown (Johansson *et al.* 1993; Molin *et al.* 1993; Drissi *et al.* 2015).
The ability of *L. agilis* La3 to utilise multiple carbon sources also possibly gives it a growth advantage in the highly competitive gastrointestinal tract, where it can most likely adapt to nutrient availability (Kamke *et al.* 2016), especially when combined with lactic acid production. Lactic acid reduces pH of the local environment, displacing particular bacteria and modifying environmental conditions to favour its own growth (Lan *et al.* 2003; Meimandipour *et al.* 2009; Chen *et al.* 2017).

### 3.5.2 Comparison to other *L. agilis* strains

The genome sequence of *L. agilis* La3 determined in this project was similar to the other sequenced *L. agilis* strains despite different isolation sources: *L. agilis* La3 from broiler chickens, *L. agilis* Marseille from humans and *L. agilis* DSM 20509 from sewage (most likely from the human gastrointestinal tract). Given that this bacterium is a gastrointestinal coloniser, this result is not unexpected. When comparing the results obtained here to the original GC% determination of *L. agilis* DSM 20509, the results were similar. In the earlier studies it was estimated to be 43.4% (Weiss *et al.* 1981) via the $^{14}C$ competition hybridisation method (Ley 1971), however, after NGS sequencing it was revealed to be 41.7% (Table 8), indicating the accuracy of the $^{14}C$ competition hybridisation method.
One unique gene found in *L. agilis* La3 when compared to the other *L. agilis* strains encoded a mucin binding protein. The highest unique match to the mucin binding protein was from *L. salivarius subsp. salivarius* UCC118, isolated from the human gastrointestinal tract (Dunne *et al.* 1999). The high coverage and low identity of the MucBP domains may indicate adaptation to chicken mucin, and therefore may be important in the effective colonisation capabilities of *L. agilis* La3 in chickens.

### 3.5.3 Methylation of the *L. agilis* La3 genome

Methylation is associated with several functions within prokaryotic systems including protection against self-Restriction Modification (RM) systems and defense against bacteriophages (Roberts *et al.* 1993). Methylation has also been implicated in DNA repair through methyl-directed mismatch repair (Cooper *et al.* 1993), and as an antibiotic resistance mechanism through 16S methylation (Moric *et al.* 2010). It is thought that there is strong positive selection for these RM systems; if methyltransferases are lost or mutated, host restriction enzymes will self-digest host DNA (Naito *et al.* 1995). The selfish gene hypothesis explains why so many diverse RM systems are present across almost all bacteria. Methyltransferases methylate specific bases due to the identification of specific sequence motifs (Sitaraman 2016). Analysis of the PacBio base-pair modification data revealed at least 6 different methylated motifs within the *L. agilis* La3 genome (Table 9). It has been demonstrated in prokaryotes that DNA methylation targets can be altered via phase variation of the base pair recognition domains of DNA methylation genes (Croix *et al.* 2017). This allows for a larger variety of methylation motifs than there are methylation genes, and could explain the presence of the six motifs, despite the three identified genes from the genome annotation (Table 10). It is possible that there are several unannotated or mis-annotated DNA methylation genes within the genome which could also explain this discrepancy. The relationship between the 6 methylation motifs and the eight RM genes (Table 10) also identified from the genome annotation is unclear.

It is not currently possible to determine the specific methylation genes responsible for each methylated motif and corresponding RM system without further detailed experimentation and analysis. Gene mutations or heterologous expression combined with bisulphate sequencing (Murray *et al.* 2012) could help determine the functionality of RM systems present in *L. agilis* La3.

### 3.5.4 Prophages sequences within the genome

The presence of both intact and non-intact prophages within the *L. agilis* La3 genome may provide an advantage to this bacterium due to the potential genomic incorporation of immunity genes, which have been shown previously to cause phage exclusion (Lu *et al.* 1989), however, no known phage immunity genes were identified in this analysis. It is thought that genes encoding immunity proteins have evolved within phage genomes to prevent multiple superinfections, reducing the chances of successful viral replication by foreign phage (Lu *et al.* 1989). This may give *L. agilis* La3 a competitive edge over other species within the chicken gastrointestinal tract due to the presence of such potential immunity genes. Alternatively, these genes could be non-functional remnants. However, due to incomplete phage databases and the phages and fragments identified within the genome being largely uncharacterised, the true function of these genes are unknown.

### 3.5.5 Transcript abundance across the *L. agilis* La3 genome

In this work, RNA was isolated from the mid-logarithmic and early-stationary growth phases of *L. agilis* La3 to determine highly transcribed, constitutive promoters. The RNA was sequenced and the transcriptomes were analysed by identifying the normalised transcript levels, or transcripts per million reads (TPM) (Figure 8). The promoters of the most abundant genes were then predicted (Table 14). These promoters were utilised in Chapters 4 and 5 for expression studies.

#### 3.5.5.1 Host-binding and colonisation characteristics

Transcript abundance across the genome further suggested the colonisation characteristics of *L. agilis* La3. GAPDH, elongation factor Tu and enolase are the first, third and fifth most abundant transcripts, respectively (Figure 8), and are likely to be involved in *Lactobacillus*-host adhesion via secretion of proteins or cell-wall anchoring (Granato *et al.* 2004; Antikainen *et al.* 2007; Glenting *et al.* 2013; Kinoshita *et al.* 2016). GAPDH and enolase are glycolytic enzymes found externally on the cell wall of many *Lactobacillus* strains via protein secretion and anchoring (Antikainen *et al.* 2007). It was shown in *Lactobacillus crispatus* that at low pH values, GAPDH and enolase are positively charged and will bind to negatively charged components of the cell wall. Binding to lipoteichoic acids (cell wall components) was tested and GAPDH was found to bind only at low pH (<5.6) and released at higher pH values. In *L. plantarum* 299v, it has also been shown these two proteins are important for adherence to human epithelial cells at a pH of 5, but not 7 (Glenting *et al.* 2013). These proteins bind to fibronectin, plasminogen, and mucin at a low pH. Two surface proteins from *L. agilis* strain MYU 227, a strain isolated from porcine mucus, enolase and phosphoglycerate kinase (PGK) were shown to be adhesive to porcine mucus (Kinoshita *et al.* 2016). *Lactobacillus johnsonii* NCC533's elongation factor Tu is also a surface protein that has been shown to bind strongly to several human intestinal cell lines and mucin at pH 5, but suffers a significant binding reduction at pH 7 (Granato *et al.* 2004). Another potential adhesion CDS in the *L. agilis* La3 genome is encoded by the *degV* gene, the expressed protein of which binds to short chain fatty acids (Nan *et al.* 2009).The homologs of the proteins that are found to mediate adherence in other *Lactobacillus* strains may be relevant in the gut environment from which *L. agilis* La3 was isolated.

The pH of commercial broiler chicken gastrointestinal tract range from 3.47±0.139 in the gizzard, 6.43±0.031 in the small intestine, 6.4±0.057 in the large intestine and 6.62±0.162 in the caecum (Mabelebele *et al.* 2014). However *Lactobacillus* strains, specifically *L. agilis*, have been shown to lower the pH of their environment by producing lactic acid (Weiss *et al.* 1981; Palop *et al.* 1995; Meimandipour *et al.* 2009). Through genome mining, *L. agilis* La3 was identified as likely to produce lactic acid due to presence of L and D-lactate dehydrogenase genes. Furthermore, L-lactate dehydrogenase was the sixteenth most abundant transcript detected, indicating that this gene may be constitutively expressed at a high level.

The presence of motility factors such as genes encoding proteins comprising a flagella apparatus suggests that *L. agilis* La3 has the ability to move freely through the mucosal layer of the gastrointestinal tract of the chicken. This probably improves the colonisation potential of *L. agilis* La3. Combined with the presence of genes encoding capsule proteins,

which have been demonstrated to be highly important for colonisation (Morona *et al.* 2006), and with the bind-and-release mechanism of highly abundant host cell and mucus-adhesion proteins such as GAPDH, enolase and elongation factor Tu, *L. agilis* La3 appears to be a highly specialised organism well suited to colonising and surviving in the chicken gastrointestinal tract.

### 3.5.6 Potential probiotic effects

Particular *L. agilis* strains such as *L. agilis* JCM 1048 have demonstrated probiotic effects (Lan *et al.* 2003; Meimandipour *et al.* 2009; Baldwin *et al.* 2018). The high transcript abundance of adherence genes such as GAPDH, enolase and elongation factor Tu (Figure 8) may play a role in the probiotic impact of such strains (Lan *et al.* 2003; Meimandipour *et al.* 2009), demonstrated by *L. agilis* JCM 1048. *L. agilis* JCM 1048 is a chicken gastrointestinal tract isolate, and considering all sequenced *L. agilis* genomes are highly similar (Table 8) despite their different isolation points, the genes encoded by strains such as *L. agilis* JCM 1048 may be involved with competitive binding to epithelial cells, and thus reduce the chance of infection by opportunistic pathogenic bacteria (Mangell *et al.* 2006). These genes are almost identical (100% query cover, 99% identity) across the three *L. agilis* isolates, indicating highly conserved function between them. This observation, combined with the high transcription levels for the lactic-acid producing L-lactate dehydrogenase gene (the sixteenth most abundant transcript in *L. agilis* La3 (Figure 8), which has been confirmed experimentally via a reporter gene (Stephenson *et al.* 2011)), suggests that *L. agilis* La3 alters its environment to increase acidity, therefore improving its adhesive abilities. The acidic environment produced by these strains may also reduce the ability of potential pathogens to gain colonisation footholds by changing (i.e reducing the pH) of their ecological niche, which is more suited to self and other lactic acid bacteria (Lan *et al.* 2003; Meimandipour *et al.* 2009; Chen *et al.* 2017). One study showed that a probiotic mix of *L. agilis* JCM 1048 (ATCC 43564) and *Lactobacillus salivarius* JCM 1230 inhibited the growth of *Salmonella* spp. and *Campylobacter jejuni*, both of which make up most poultry-related foodborne illness incidences across Australia (Kirk *et al.* 2014) and Europe (EFSA 2016; EFSA 2010). When the probiotic mix was given to chickens, significantly more *Lactobacillus* were isolated from the chickens after 7 days (Lan *et al.* 2003).

In these studies, the probiotic effect on bacterial populations within caecal contents resulted in a reduction in *Streptococcus* species and other aerobes, whereas an increase was seen in anaerobes, such as *Bifidobacterium* and *Lactobacillus* species. An increase in propionate and butyrate production was demonstrated in these studies, along with a reduction in acetate production, while no lactate was present (Meimandipour *et al.* 2009). These changes were most likely the result of cross-feeding, where butyrate producers within the caecal contents were using lactate produced by the probiotic mix as a substrate to create butyrate. These bacteria most likely were also consuming acetate, which led to acetate decrease when the probiotic mix was added. The propionate increase was directly due to the presence of the probiotic mix (Meimandipour *et al.* 2009).

Another study has shown that when *L. agilis* was given in a probiotic mix with *Lactobacillus ingluviei* and *L. reuteri* to day-old chicks, a significant difference in weight gain was seen when compared to the non-inoculated control group after 28 days (Baldwin *et al.* 2018). Despite *L. agilis* and *L. reuteri* not successfully colonising the birds, the overall microbial community was not different between the probiotic-treated group and control. There were, however, short-term significant changes in individual taxa including increases in

*Pseudoclavibacter*, *Nocardiopsis*, *Burkholderia* and *Brevundimonas*, and unclassified Planococcaceae and Clostridiales genera (Baldwin *et al.* 2018)
Studies such as these demonstrate the environment-modulating capabilities of *L. agilis* JCM 1048 and other lactic acid bacteria. These strains are potentially providing probiotic effects by impacting microbial composition and also possibly through subsequent competitive exclusion (Schoeni *et al.* 1994; Methner *et al.* 1999). Bacteriocin production has not yet been demonstrated by these strains to explain the potential probiotic effects that they confer (Klaenhammer 1993). Genome analysis in this study looked at whether *L. agilis* La3 possessed any bacteriocin genes that may explain probiotic effects, however, no bacteriocin genes were identified within the *L. agilis* La3 chromosome (Result 3.3.5).

These results suggest that *L. agilis* La3 modulates its environment to allow adhesion and successful colonisation and persistence within the chicken gastrointestinal tract (Stephenson *et al.* 2010; Vezina 2013), regardless of the current host microbial architecture within the host or diet. The potential ability of *L. agilis* to modulate the surrounding environment in the host could explain why it is a highly prevalent gastrointestinal coloniser across many different animals including chickens (Mitsuoka 2002; Lan *et al.* 2003; Stephenson *et al.* 2010; Bujnakova *et al.* 2014; Chen *et al.* 2017), horses (Morita *et al.* 2007), pigeons (Baele *et al.* 2001), humans (Molin *et al.* 1993), pigs (Mitsuoka 2002; Murphya *et al.* 2013; Kinoshita *et al.* 2016), lemurs and tapirs (Kajikawa *et al.* 2016).

# 4. Chapter 4 Results: Development of stable expression and integration systems for expression of recombinant proteins in *L. agilis* La3

Once the genome of *L. agilis* La3 had been assembled and the encoded genes characterised, the transcriptome of *L. agilis* La3 was analysed. Several predicted homologous, constitutive promoters were identified (Table 14). It was hypothesised that the development of *L. agilis* La3 as a live vaccine required stable, consistent expression of recombinant proteins. The aim of this chapter was to determine the transcriptional activity of these predicted constitutive promoters to confirm their characterisation and functional capacity, and to identify the most suitable one/s for protein expression. Before these experiments could be attempted, the construction of a genetic tool with which to test the putative promoters was required and several approaches were attempted to construct a suitable expression system. Previous live vaccines have used two main approaches to express therapeutic antigens in an expression cassette, either from plasmids (Ahmed *et al.* 2014; Zhao *et al.* 2015) or from a chromosomal location (Layton *et al.* 2009; Layton *et al.* 2011). Details of the plasmids used and constructed for this work are provided in Table 4 (Chapter 2).

The development of chromosomal integration systems is more time-consuming compared to utilising self-replicating plasmids due to the greater number of construction features required. However, chromosomally-based expression systems have significant advantages over plasmid-based expression systems because of their inherent stability in comparison to the potential instability of plasmids (Alpert *et al.* 2003). Chromosomal expression systems have been constructed previously within *Lactobacillus* species using site specific *attP-attB* recombination (Auvray *et al.* 1999) and homologous recombination systems (van Zyl *et al.* 2015).

The construction of a system in which recombinant proteins could be stably expressed from the chromosome of *L. agilis* La3 required the identification of a suitable recombination site. Ideally the chosen site would have minimal impact on regular cell function upon disruption and would thus avoid tempering the fitness of *L. agilis* La3 as a persistent coloniser of the chicken gastrointestinal tract. In this chapter, the genome of *L. agilis* La3 was analysed to identify potential integration sites, and an expression system was constructed.

## 4.1 Plasmid-based expression system

Initially, to construct a stable expression system for the development of *L. agilis* La3 as a live vaccine, a shuttle plasmid-based approach was attempted. *L. agilis* La3 has 4 native plasmids (pLa3_1, pLa3_2, pLa3_3, pLa3_4) (Stephenson 2010). The approach centred around utilising one of the native *L. agilis* La3 plasmids (pLa3_4) as the base for the expression plasmid, allowing replication within *L. agilis* La3, and fusing it with an *E. coli* plasmid, to allow simplified genetic manipulation. The plasmid pLa3_4 was seen as a desirable base for construction of a shuttle plasmid as it is the smallest native plasmid within *L. agilis* La3 (Stephenson 2010). The small size (2,686 bp) provided an advantage regarding transformation efficiency and cloning efficiency into *L. agilis* La3 and *E. coli*, respectively. It also contained the fewest restriction sites out of all the native plasmids, thereby improving cloning options. Since only the replicative regions of pLa3_4 were required for replication in

*L. agilis* La3, multiple shuttle plasmid constructs were generated to further reduce the size of the 2,686 bp native plasmid while maintaining replication within *L. agilis* La3 (Vezina 2013). A smaller, 2,145 bp region of pLa3_4 was demonstrated to allow stable replication within *L. agilis* La3 when used as a base in an expression plasmid. This new shuttle plasmid (pVEcLa4) demonstrated high stability within *L. agilis* La3 and significantly higher stability than another common *Lactobacillus* shuttle plasmid, pTRKH2. This result supported the advantageous approach of using native replication machinery over heterologous ones (Vezina 2013). This 2,145 bp replicating region of pLa3_4 was synthesised and cloned into *E. coli* plasmid pUC57 (Genscript), along with an erythromycin resistance expression cassette (*ermR*) and an expression cassette containing the native cell wall associated hydrolase promoter ($P_{cwah}$), signal sequence, and the region encoding the first two amino acids of the mature Cwah protein (to improve the chance of proper protein expression and cleavage), and the *cwah* transcriptional terminator sequence. The $P_{cwah}$ promoter was used as it had been previously shown to express and secrete a heterologous protein in *L. agilis La3* (unpublished data). The vector was named pVEcLaSyn1.

### 4.1.1 Attempted transformation of pVEcLaSyn1 into *L. agilis* La3

Once synthesised, pVEcLaSyn1 was used in transformation experiments with *L. agilis* La3 using a modified protocol (Method 2.19). Despite repeated attempts, using various conditions known to previously work for transformation of plasmids into *L. agilis* La3, the vector pVEcLaSyn1 was not successfully transformed into *L. agilis* La3, and no colonies were obtained (data not shown). Many strategies outlined below were employed to resolve the inefficient transformation, as the plasmid-base (pLa3_4) had previously demonstrated stable replication within *L. agilis* La3 (Vezina 2013).

#### 4.1.1.1 DNA source

To try and resolve the lack of transformants obtained for *L. agilis* La3 with the synthesised plasmid, pVEcLaSyn1 was transformed into several *E. coli* strains and DNA prepared for transformation. Firstly, plasmid DNA was transformed and isolated from *E. coli* JM109. This DNA was then used to attempt to transform *L. agilis* La3, unfortunately no transformants were obtained. Secondly, pVEcLaSyn1 was transformed into *E. coli* MM294 (Meselson *et al.* 1968), a strain capable of producing multimeric plasmid DNA (Krivonogov 1984). Some bacteria such as *Bacillus subtilis* show improved efficiency when transformed with multimeric plasmid DNA (Canosi *et al.* 1978; Vos *et al.* 1981). No colonies were obtained when transforming *L. agilis* La3 with multimeric pVEcLaSyn1 DNA isolated from *E. coli* MM294.

Thirdly, pVEcLaSyn1 was transformed into *E. coli* JM110. A previous study had demonstrated that DNA produced by *E. coli* with a *dam*-/*dcm*- methylation profile significantly improved *Lactobacillus plantarum* transformation efficiency (Spath *et al.* 2012). The methylation profile of *E. coli* JM110 is *dam*-/*dcm*-, compared to *E. coli* JM109 which is *dam*+/*dcm*+. However, transformation of *L. agilis* La3 with DNA isolated from *E. coli* JM110 resulted in no transformants.

#### 4.1.1.2 Conjugation-mediated transfer of pVEcLaSyn1 into *L. agilis* La3

To circumvent the inefficient transformation of pVEcLaSyn1 into *L. agilis* La3, conjugation-mediated transformation was attempted. The RP4 *oriT* locus from pJIR1456 (Lyras *et al.* 1998) was cloned into pVEcLaSyn1 (pVEcLaSyn1-*oriT*), and transformed into *E. coli* S17-1. *E. coli* S17-1 contains a chromosomally-based RP4 locus, allowing mobilisation of *oriT*-containing plasmids (Simon *et al.* 1983). Mating experiments were performed between *E. coli* S17-1 as a donor and *L. agilis* La3 as the recipient. These experiments did not produce any transconjugants when appropriate selection was used (MRS media containing erythromycin 10 µg/mL). pVEcLaSyn1-*oriT* demonstrated successful conjugation from *E. coli* S17-1 to *E. coli* LT101 (Palombo *et al.* 1989; Lyras *et al.* 1998), which was used as a positive control.

#### 4.1.1.3 Protoplast-mediated transformations

To try another approach of introducing DNA into *L. agilis* La3, protoplast and PEG-mediated protoplast transformation (Chang *et al.* 1979; Klebe *et al.* 1983; Simon *et al.* 1985) was attempted. Protoplast transformations involve degrading the cell wall of donor and recipient cells, then fusing the cell membranes together, allowing direct movement of selectable plasmids into the recipient cell (Chang *et al.* 1979). Protoplast-mediated transformations were undertaken by fusing donor *E. coli* JM109 (pVEcLaSyn1) with recipient *L. agilis* La3. However, no successful transformants were observed when appropriate selection was applied (MRS media containing erythromycin 10 µg/mL).

Since the second aim of the project was to develop *L. agilis* La3 as a live delivery platform, and the eventual goal was to move expression cassettes into the *L. agilis* La3 chromosome for maximal stability of recombinant expression, rather than continue to troubleshoot the plasmid-based expression system a chromosomal insertion system was constructed. The hypothesis on which these experiments were based were that once an expression cassette is inserted into the chromosome it would not require selective pressure for maintenance, nor independent replication, as would be required in a plasmid-based system.

## 4.2 Chromosomal insertion via double crossover homologous recombination

To construct a system where recombinant proteins could be stably expressed from the chromosome of *L. agilis* La3, selection of a suitable recombination site in the chromosome was required. Ideally the chosen site would have minimal impact on normal cell function when disrupted and this would avoid tempering strain fitness as a persistent coloniser of the chicken gastrointestinal tract. The genome of *L. agilis* La3 was analysed to identify potential integration sites. Several potential sites were identified and of these, a 21 bp site in between the non-coding regions of two converging transcribed genes, a hypothetical protein and a GMP synthase gene, was chosen. This site was chosen as it would not seemingly impact on any upstream regulation of the surrounding genes. Also, the RNA sequence analysis also revealed low transcriptional abundance across the site (Figure 9) as well as the flanking regions, making it an attractive site for chromosomal integration. Other potential integration

sites were ruled out due to high transcriptional activity, and presence of coding sequence, or regulatory elements.

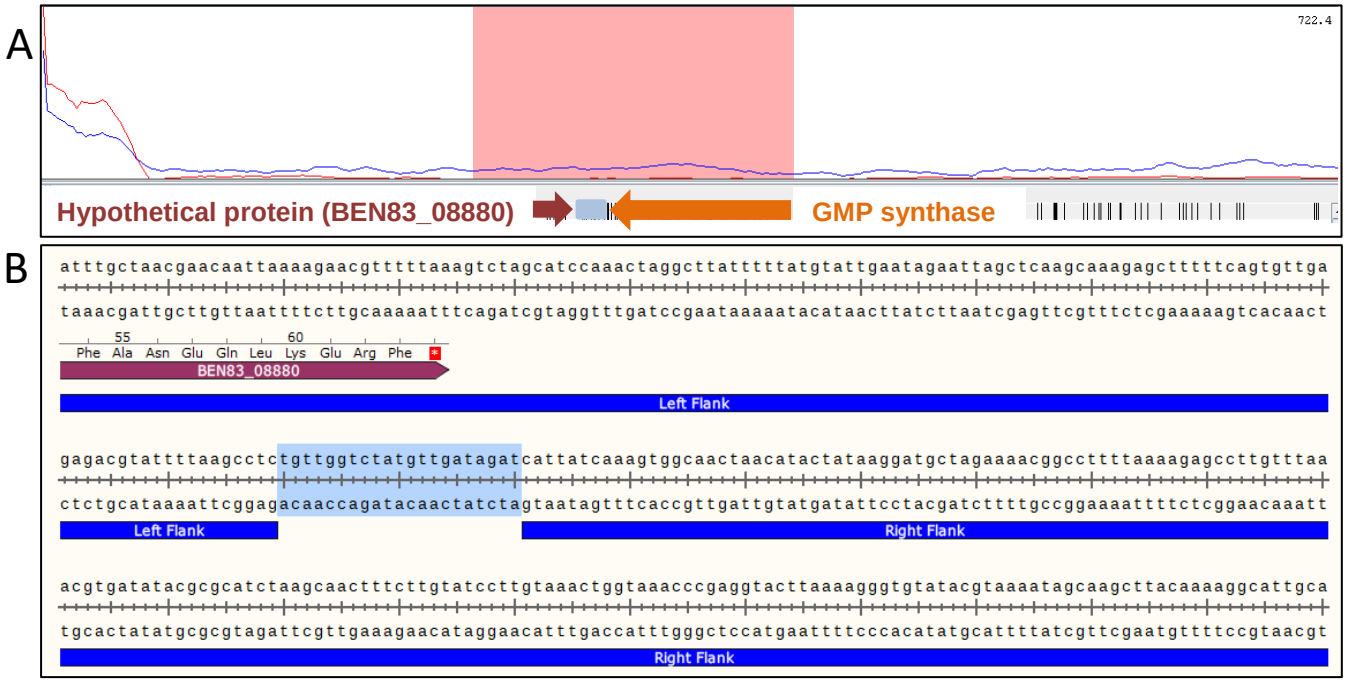


Figure 9: Diagrams showing the chromosomal insertion site within the *L. agilis* La3 chromosome.

A: Transcript abundance of the insertion site. The blue line represents transcripts from the early-stationary growth phase. The red line represents transcripts from the mid-logarithmic growth phase. Both show low transcript abundance in this area. The selected pink region shows the entire focal region of the chromosomal insertion site. The two converging genes are shown as coloured arrows. The hypothetical protein (BEN83_08880) is shown in purple, and the GMP synthase is shown in orange. The light blue box shows the region of insertion.

B: Sequence view of the 21 bp recombination site to be replaced. The selected bases (light blue) show the 21 bp insertion site. The dark blue bars show the planned regions of homology.

To construct a vector for homologous recombination of the expression cassette into the chromosome, the plasmid pTRKH2-LCER (Table 4) was generated (Figure 10). This plasmid was constructed by firstly cloning the Left Flank region of homology to the insertion site into the vector pTRKH2 using *Xba*I/*Sac*I. The chloramphenicol resistance cassette, *catP* was then inserted using *Sac*I/*Pst*I. Next, *egfp* under the control of the *IdhL* promoter, ($P_{IdhL}$) was cloned using *Sal*I, followed by the Right Flank region of homology using *BamH*I/*Pvu*I. The regions of homology were used in this plasmid to allow integration of the expression cassette into this 21 bp site within the *L. agilis* La3 chromosome via double crossover (Figure 10).

The integration/expression cassette used to develop and test the homologous recombination process therefore consisted of a reporter gene, *egfp,* under the control of the

*L. agilis* La3 native $P_{ldhL}$, and a *C. perfringens* chloramphenicol resistance gene (*catP*) expressed under its native *C. perfringens* promoter, that was used for the selection of chromosomal integrants. The expression/selection cassette was flanked by approximately 1 kb of homologous chromosomal sequence on the left flank and 2 kb on the right flank. These homologous regions provided the substrates for homologous recombination.

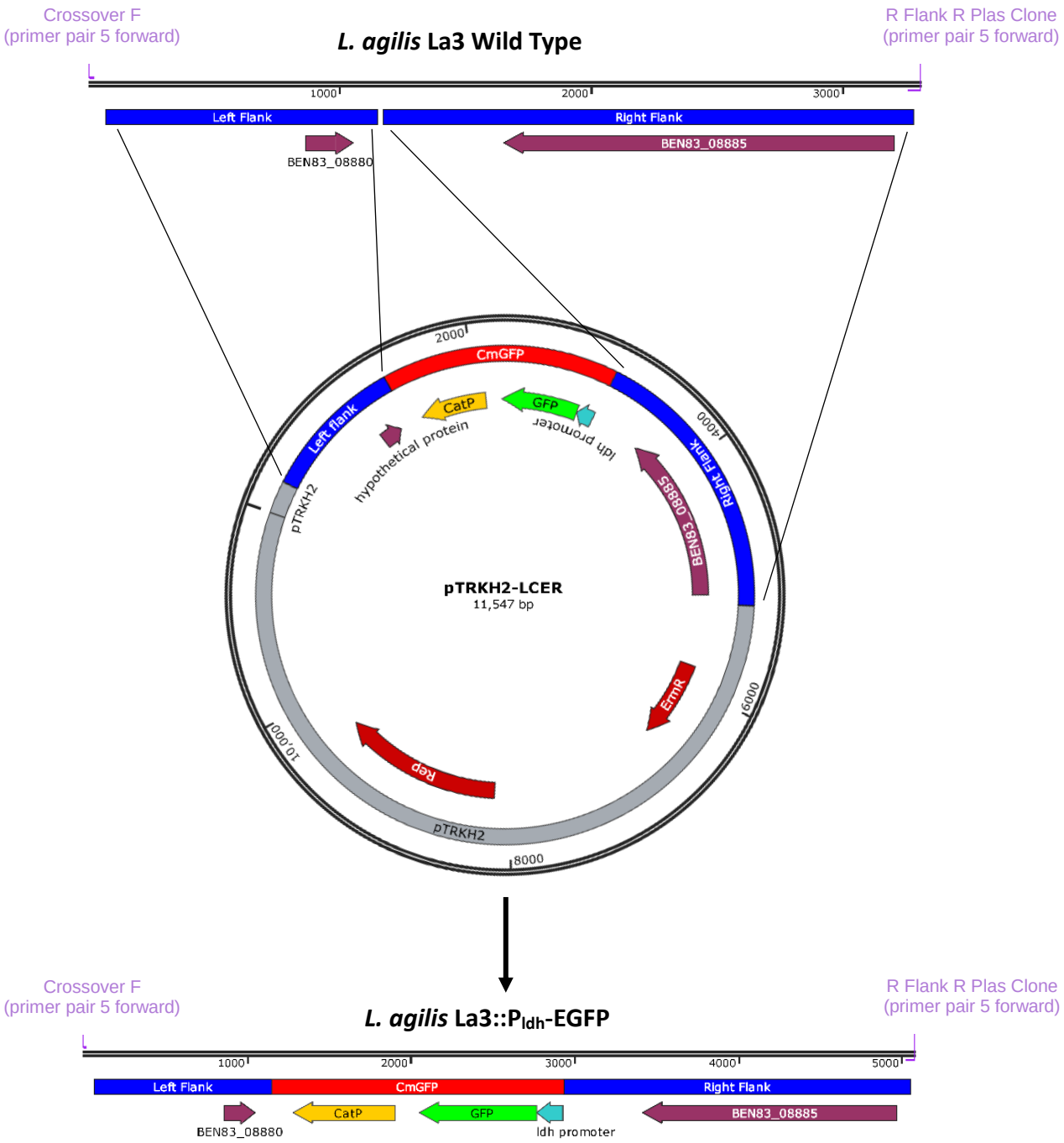


Figure 10: Graphical representation demonstrating the double crossover event in *L. agilis* La3 using pTRKH2-LCER to generate insertional mutant *L. agilis* La3::$P_{ldhL}$-EGFP via homologous recombination. The Left and Right Flanks (dark blue) allowed recombination of the catP and egfp genes into the chromosome, replacing 21 bp of non-coding DNA. The new mutant was CmR 10 µg / mL, as the erythromycin resistance gene (ErmR) was not integrated. The purple text shows the primer sites (spans integration site using primer pair 5 (Table 5) made up of Crossover F and R Flank R Plas Clone), which allowed identification of *L. agilis* La3 chromosomal recombinants via PCR.

pTRKH2-LCER was successfully transformed into *L. agilis* La3 using the modified *L. agilis* La3 transformation protocol (Method 2.19). To select for chromosomal integration and against plasmid maintenance, transformants were passaged using chloramphenicol selection (method 2.22). As pTRKH2-LCER is highly unstable in *L. agilis* La3, by passaging in the presence of chloramphenicol, highly stable chromosomal-based resistance will proportionally grow in number when compared to unstable plasmid-based resistance. After insertional mutants were generated, both resistance phenotyping (chloramphenicol and erythromycin sensitivity) and PCR-screening methods were used to screen colonies for those carrying insertion mutants. The bacterial population was usually dominated by bacteria in which chromosomal recombination had occurred within 1 – 3 subcultures. Chromosomal integration was shown to occur through homologous recombination via double crossover as opposed to single crossover by using PCR primer pair 5 (Figure 10). This PCR product spanned the entire insertion region and was used to screen potential recombinants. A single crossover would incorporate the entire plasmid, meaning the PCR would not produce either the expected wild type band size, nor the expected double crossover band size. This PCR result demonstrated that chromosomal recombinants via double crossover had been generated (Figure 11).

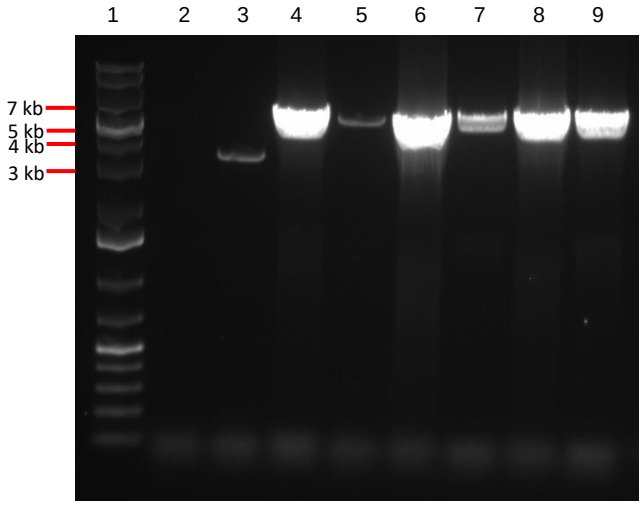


Figure 11: Gel electrophoresis image of double crossover diagnostic PCR (primer pair 5) on randomly selected *L. agilis* La3 colonies. Lane 1: Generuler 1 kb Plus. Lane 2: No DNA control. Lane 3: *L. agilis* La3 wild type. Lanes 4 – 9: Randomly selected *L. agilis* La3 colonies screened for double crossover recombination events. The ~5,075 bands seen in lanes 4 – 9 indicates successful chromosomal recombination. The *L. agilis* La3 wild type lane shows a ~3,309 bp band.

### 4.2.1 Stability of chromosomal insert and plasmid

Once individual *L. agilis* La3 chromosomal recombinants were isolated, the stability of the insert and plasmid were determined. After passaging in biological triplicate (n = 150), colonies were screened for chloramphenicol resistance and erythromycin sensitivity, and via PCR (Figure 12) and found to be 100% stable after ~100 generations using PCR primer pair

5 (Table 5) to detect chromosomal integration. Since these colonies were chloramphenicol resistant, erythromycin sensitive and primer pair 5 was generating the expected 5,075 bp amplicon, this indicated maintenance of the chromosomal insert within *L. agilis* La3.

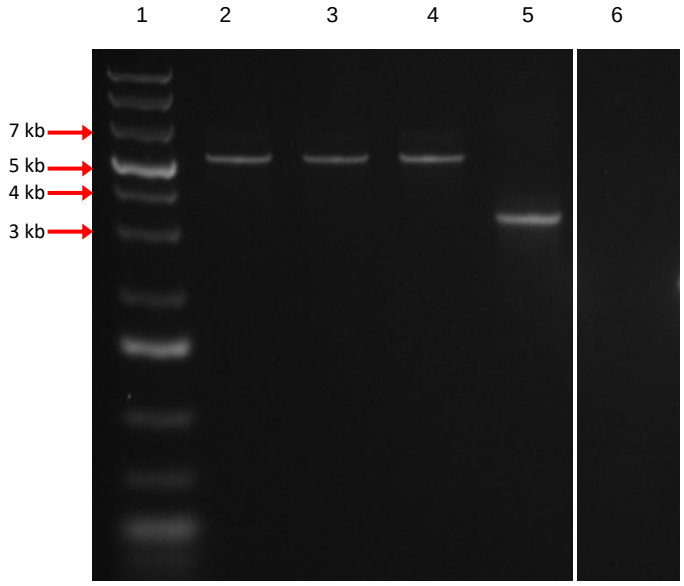


Figure 12: Gel electrophoresis image of double crossover diagnostic PCR (primer pair 5) on *L. agilis* La3::$P_{ldhL}$-EGFP MRS liquid media after ~100 generations passaging. Lane 1: Generuler 1 kb Plus. Lane 2: *L. agilis* La3::$P_{ldhL}$-EGFP colony #1. Lane 3: *L. agilis* La3::$P_{ldhL}$-EGFP colony #88. Lane 4: *L. agilis* La3::$P_{ldhL}$-EGFP colony #95. Lane 5: *L. agilis* La3 wild type. Lane 6: No DNA control. The ~5,075 bands seen in lanes 2 – 4 indicate successful chromosomal recombination, compared to the wild type PCR product (~3,309 bp).

#### 4.2.1.1 Southern blot

Southern blot hybridisation analysis was utilised (Figure 14, Figure 15 and Figure 17) to confirm recombination into the intended recombination site within the *L. agilis* La3 chromosome, and to show that no off-target integration of the plasmid into non-targeted sites had occurred. These experiments were performed using DNA isolated from *L. agilis* La3::$P_{ldhL}$-EGFP colonies before and after they were passaged for ~100 generations.

*Hin*dIII is a frequent cutter within the *L. agilis* La3 genome (1,184 times) and the chromosomal insertion region was predicted to be carried on a 2,856 bp fragment. BDV11 was used as a probe, which targeted the inserted *egfp* gene (Figure 13). The first Southern hybridisation (Figure 14) utilised this probe.

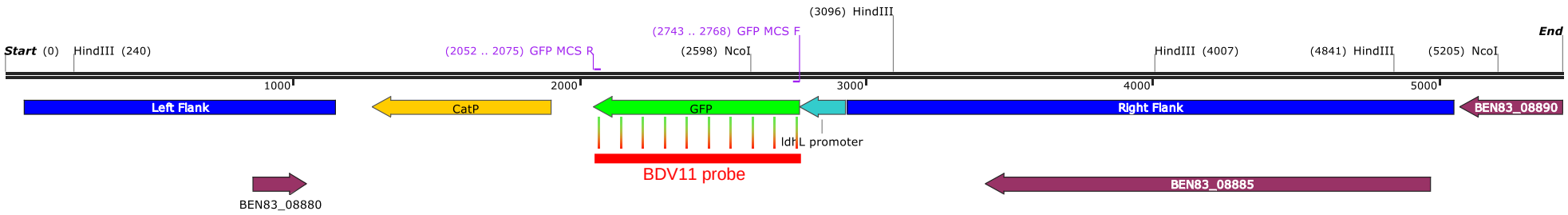


Figure 13: Graphical representation of the probe binding site to the recombinant *L. agilis* La3 chromosome. This probe was used for the first and second Southern blots. The probe is shown in red, and the binding site against *egfp* is shown. The purple text shows the primers used to generate the probe (743 bp). The *Hin*dIII and *Nco*I sites within the chromosome are also shown.

This Southern hybridisation was used to show that the chromosomal insert was localised to one site within the *L. agilis* La3::$P_{ldhL}$-EGFP genome. As shown in Figure 14, a single 2,856 bp band of expected size was observed. The probe did not hybridise to the negative *L. agilis* La3 wild type control, indicating that the insert is not present in the *L. agilis* La3 wild type strain but present in the double-cross over insertion mutants.

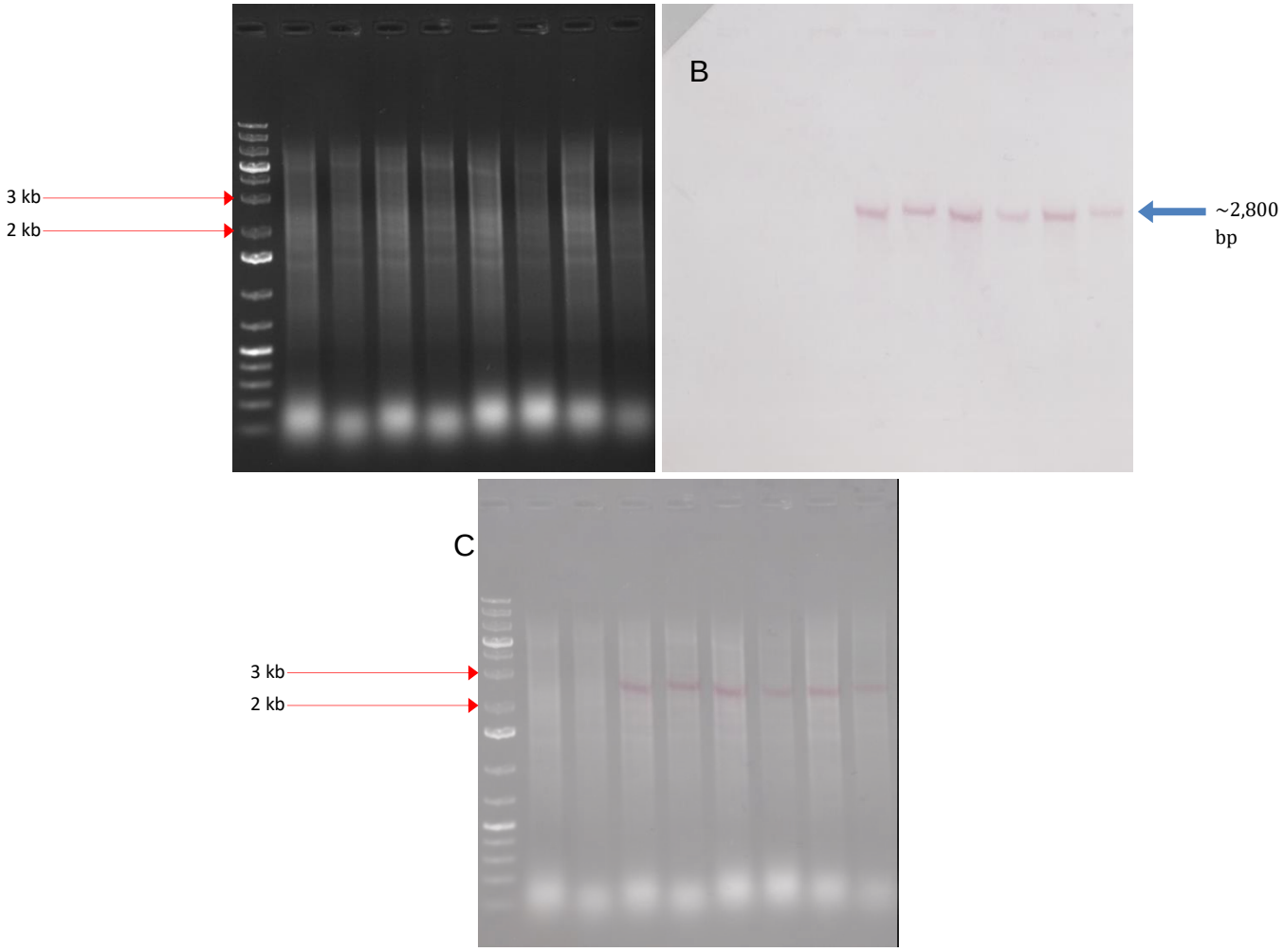


Figure 14: Image of gel and Southern blot of *Hin*dIII digested *L. agilis* La3 chromosome with BDV11 probe (*egfp* probe). This digest was intended to show only a single insertion into the chromosome had occurred.

A: Agarose gel electrophoresis image of the *L. agilis* La3 chromosomal *Hin*dIII digests. Lane 1: Generuler 1 kb Plus. Lane 2 and 3: *L. agilis* La3 WT. Lane 4: *L. agilis* La3::$P_{ldhL}$-EGFP colony #1. Lane 5: *L. agilis* La3::$P_{ldhL}$-EGFP colony #1 passaging ~100 generations. Lane 6: *L. agilis* La3::$P_{ldhL}$-EGFP colony #88. Lane 7: *L. agilis* La3::$P_{ldhL}$-EGFP colony #88 passaging ~100 generations. Lane 8: *L. agilis* La3::$P_{ldhL}$-EGFP colony #95. Lane 9: *L. agilis* La3::$P_{ldhL}$-EGFP colony #95 passaging ~100 generations.

B: Southern blot of the *L. agilis* La3 chromosomal *Hin*dIII digests using the BDV11 probe. Lanes identical to A.
C: Overlay image of Southern blot to agarose gel electrophoresis image of the *L. agilis* La3 chromosomal *Hin*dIII digests. This shows the probe annealing size to a band consistent with the expected size of ~2,856 bp in chromosome.

The second Southern hybridisation analysis (Figure 15), used *Hin*dIII/*Nco*I, which cut within the *egfp* target sequence, meaning the BDV11 probe would bind to two fragments. This was performed to confirm that the insert had only integrated into the chromosome at a single expected location. The two expected bands (2,358 and 498 bp) were observed in both the pTRKH2-LCER control and recombinant *L. agilis* La3::$P_{ldhL}$-EGFP strains, demonstrating the insert was present at one location.

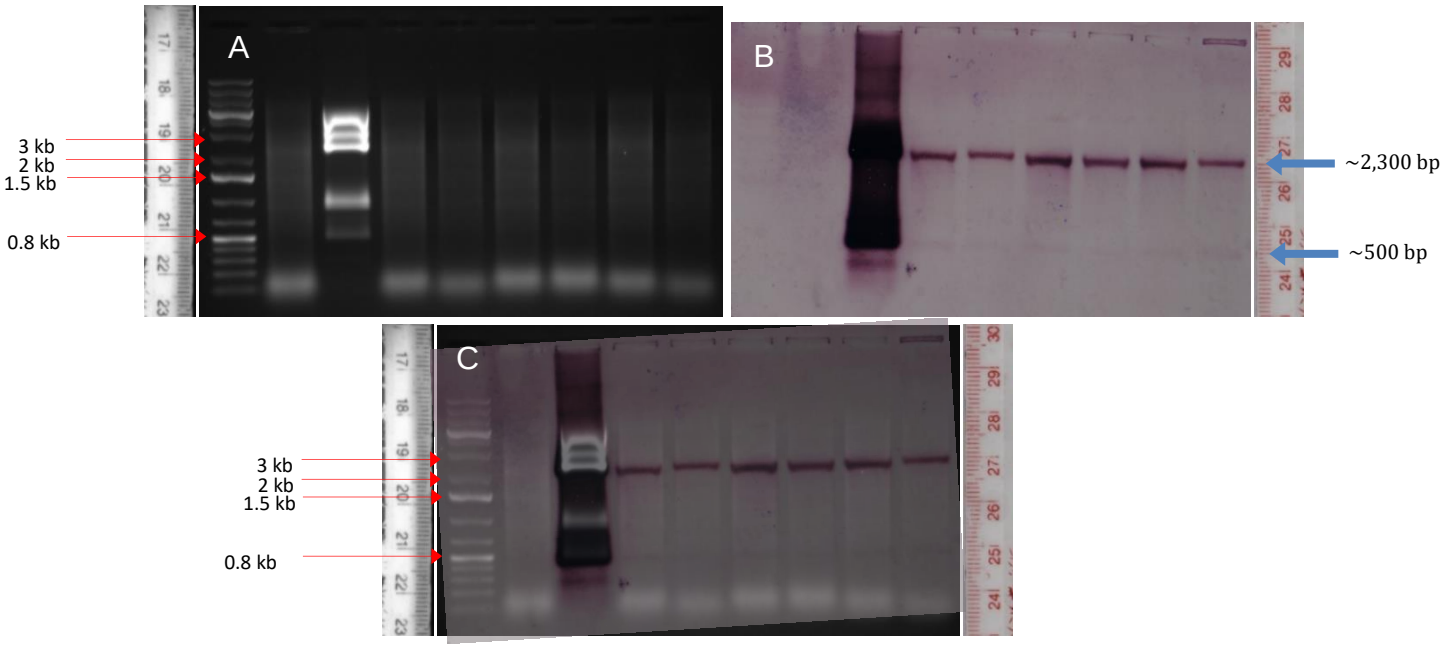


Figure 15: Image of gel and Southern blot of *Hin*dIII/*Nco*I digested pTRKH2-LCER and *L. agilis* La3 chromosome with BDV11 probe (*egfp* probe). This digest was intended to cut within the probed site to show only a single insertion within the chromosome had occurred.

A: Agarose gel electrophoresis image of the *L. agilis* La3 chromosomal *Hin*dIII/*Nco*I digests. Lane 1: Generuler 1 kb Plus. Lane 2: *L. agilis* La3 wild type. Lane 3: pTRKH2-LCER. Lane 4: *L. agilis* La3::$P_{ldhL}$-EGFP colony #1. Lane 5: *L. agilis* La3::$P_{ldhL}$-EGFP colony #1 passaging ~100 generations. Lane 6: *L. agilis* La3::$P_{ldhL}$-EGFP colony #88. Lane 7: *L. agilis* La3::$P_{ldhL}$-EGFP colony #88 passaging ~100 generations. Lane 8: *L. agilis* La3::$P_{ldhL}$-EGFP colony #95. Lane 9: *L. agilis* La3::$P_{ldhL}$-EGFP colony #95 passaging ~100 generations.

B: Southern blot of the *L. agilis* La3 chromosomal *HindIII/NcoI* digests using the BDV11 probe. Lanes identical to A.

C: Overlay image of Southern blot to agarose gel electrophoresis image of the *L. agilis* La3 chromosomal *Hin*dIII/*Nco*I digests. This shows the probe annealing size to bands consistent with the expected size of ~2,358 and ~498 bp in size in both pTRKH2-LCER and chromosome.

Southern blots 1 and 2 (Figure 14, Figure 15) showed that the chromosomal insert had only inserted once and at the expected location and was 100% stable over time within the tested cultures, as no altered bands are seen over 100 generations. In addition to the first two southern blots a third Southern hybridisation analysis (Figure 17) was performed using BDV12 as a probe (Figure 16) to show loss of the pTRKH2-LCER plasmid in all isolates.

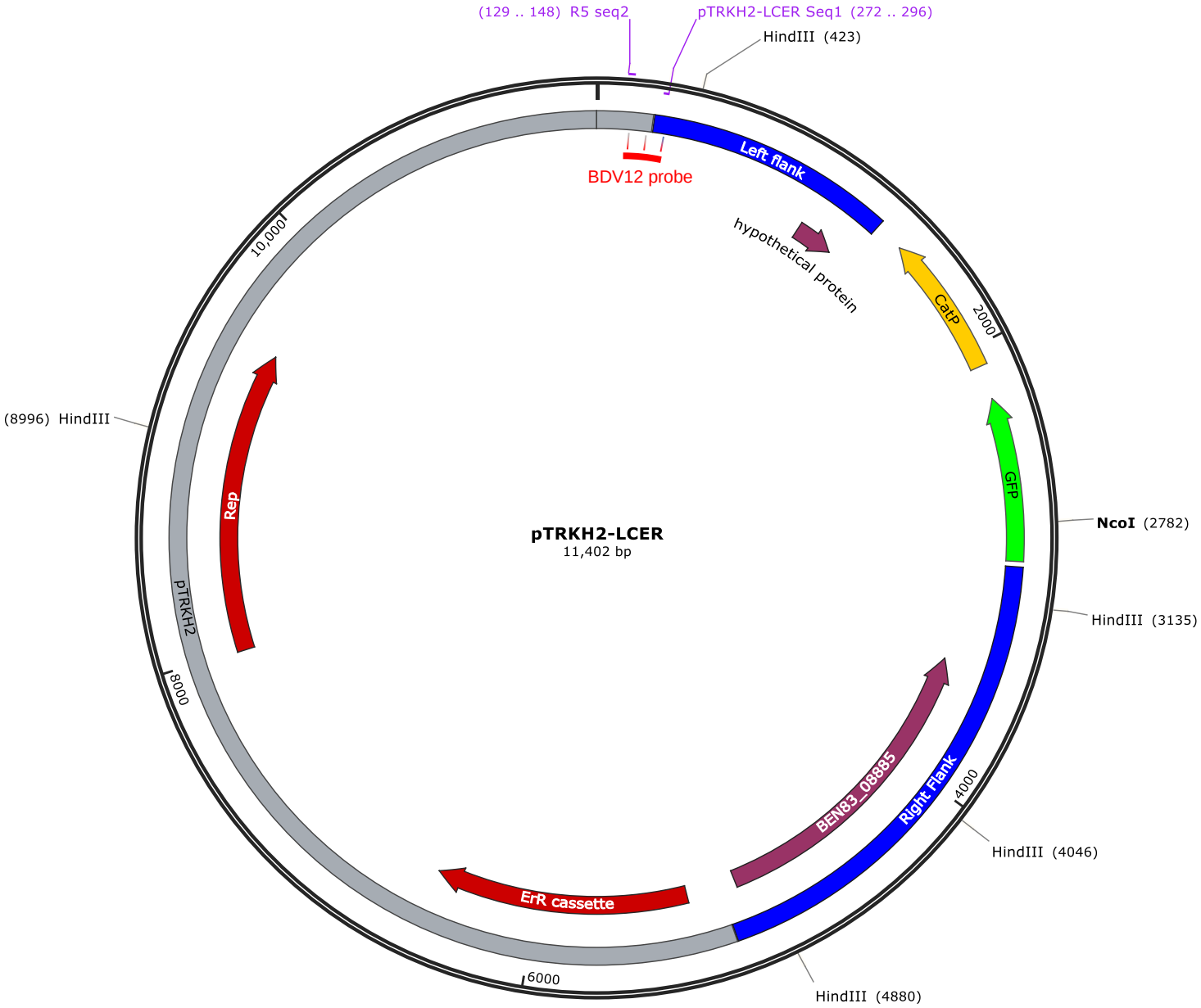


Figure 16: Graphical representation of the probe binding site in pTRKH2-LCER, demonstrating plasmid presence or absence for the third Southern blot. The probe is shown in red along with the binding site to pTRKH2-LCER. The purple text shows the primers used to generate the probe (168 bp). The *Hin*dIII and *Nco*I sites within the plasmid are also shown.

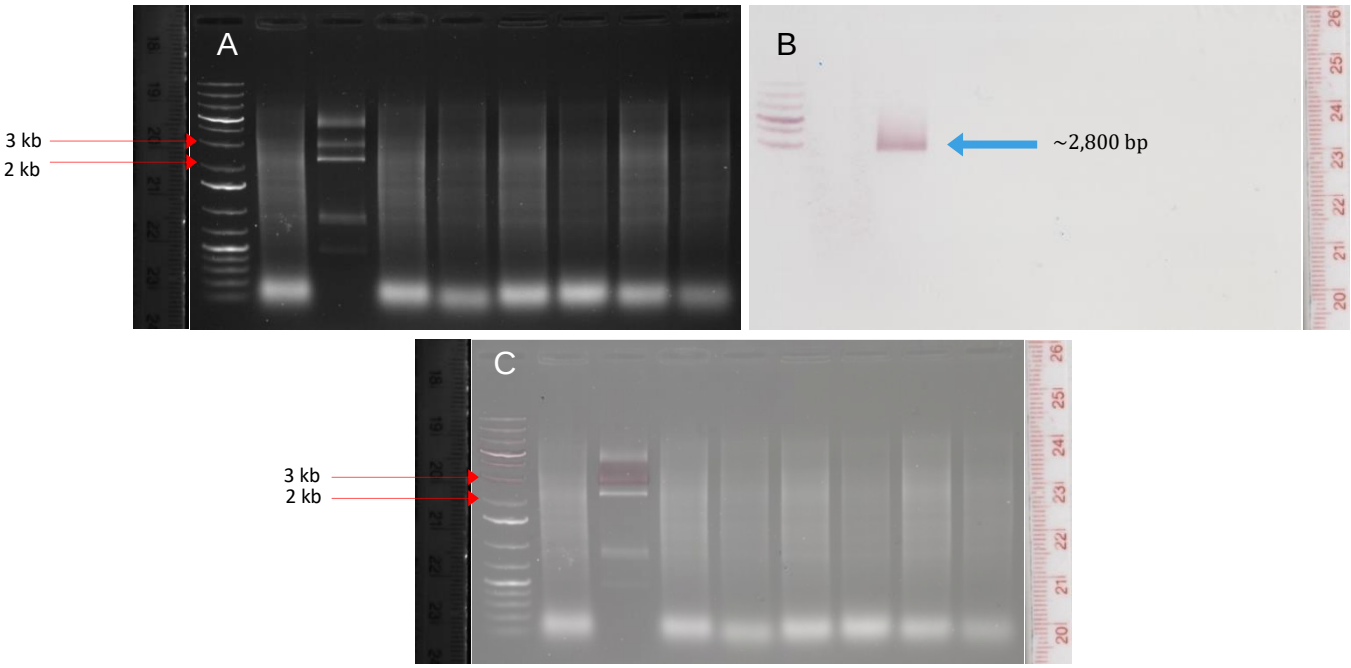


Figure 17: Image of gel and Southern blot of *Hin*dIII/*Nco*I digested pTRKH2-LCER and *L. agilis* La3 chromosome with BDV12 probe.

A: Agarose gel electrophoresis image of the *L. agilis* La3 chromosomal *Hin*dIII/*Nco*I digests. Lane 1: Generuler 1 kb Plus. Lane 2: *L. agilis* La3 wild type. Lane 3: pTRKH2-LCER. Lane 4: *L. agilis* La3::$P_{ldhL}$-EGFP colony #1. Lane 5: *L. agilis* La3::$P_{ldhL}$-EGFP colony #1 passaging ~100 generations. Lane 6: *L. agilis* La3::$P_{ldhL}$-EGFP colony #88. Lane 7: *L. agilis* La3::$P_{ldhL}$-EGFP colony #88 passaging ~100 generations. Lane 8: *L. agilis* La3::$P_{ldhL}$-EGFP colony #95. Lane 9: *L. agilis* La3::$P_{ldhL}$-EGFP colony #95 passaging ~100 generations.

B: Southern blot of the *L. agilis* La3 chromosomal *HindIII/NcoI* digests using the BDV12 probe. Lanes identical to A.

C: Overlay image of Southern blot to agarose gel electrophoresis image of the *L. agilis* La3 chromosomal *Hin*dIII/*Nco*I. This shows the probe annealing size to a band consistent with the expected size of ~2,829 bp in size, in the plasmid sample only, indicating absence of pTRKH2-LCER in chromosomal recombinants.

Southern hybridisation analysis using DNA from each double crossover *L. agilis* La3::$P_{ldhL}$-EGFP insertion mutants showed the absence of pTRKH2-LCER, while hybridising correctly to purified pTRKH2-LCER plasmid (Figure 17 - Lane 3). These Southern blots indicated 100% chromosomal insert stability over time.

The results presented here, including phenotype (antibiotic resistance profiles), PCR and Southern hybridisation analysis, when taken together, indicated that the expression cassette was successfully inserted in the desired location within the *L. agilis* La3 chromosome. These results also demonstrated 100% stability of the insert over 100 generations.

## 4.3 Testing the predicted promoters

### 4.3.1 EGFP as a reporter protein

Initially, EGFP was used as a reporter protein. Each of the five promoters were cloned independently upstream of *egfp* into pTRKH2-LCER, replacing $P_{ldhL}$. In each case, only non-fluorescent *E. coli* JM109 (cloning strain) colonies were observed. Growing these colonies and harvesting cells at early and late logarithmic phases did not produce fluorescence when analysed using the 360-10 excitation and 490-10 emission filters on the POLARstar Omega Plate Reader Spectrophotometer. Several different *E. coli* strains were used to try and troubleshoot and resolve the lack of EGFP expression such as utilising the strain DH5α and the recombinant protein expression strain BL21 DE3, however no fluorescence was observed. Upon sequencing these non-fluorescent clones, a large variety of mutations were discovered. These were a mix of nonsense mutations which truncated *egfp*, as well as mutations which disrupted the predicted promoter sequences. This indicated that EGFP when expressed using the five different *L. agilis* La3 promoters (Table 14) in *E. coli* on a high-copy number plasmid may be toxic to the cell and therefore mutations had occurred to prevent EGFP expression.

### 4.3.2 CreiLOV as reporter protein

As intact clones could not be generated using the five *L. agilis* La3 promoters to express EGFP, alternative reporter genes were considered. *L. agilis* La3 is a facultative anaerobe in a predominantly anaerobic environment (chicken gastrointestinal tract), therefore use of EGFP as a reporter protein outside of laboratory culturing is limited since this protein will not function correctly within the host. To circumvent this, an *L. agilis* La3 codon-optimised *creiLOV* was cloned in its place. CreiLOV is an oxygen-independent fluorescent blue protein, most active at 37°C and within a pH range of 6 – 9. (Mukherjee *et al.* 2015). It has an excitation wavelength of 450 nm (370 nm lower peak), and an emission wavelength of 495 nm (a maximum of 525 nm) It is also considerably smaller than EGFP (239 compared to 119 amino acids) These characteristics made it a suitable protein to use as a reporter.

### 4.3.3 pVEcXLa expression system

Following the success of the pTRKH2-LCER plasmid, a derivative plasmid, pVEcXLa (p – plasmid, V – Vezina, Ec – *E. coli*, X – double crossover, La – *L. agilis* La3) was constructed (Figure 18). pVEcXLa was constructed for the purpose of testing a range of promoters of differing strengths using a heterologous reporter protein, CreiLOV instead of EGFP. pTRKH2-LCER was modified to express an oxygen independent fluorescent reporter protein, CreiLOV, which was codon-optimised for expression in *L. agilis* La3. This sequence was flanked by suitable unique restriction sites which allow directional cloning of desired sequences. The plasmid was designed to be capable of replication in both *E. coli,* for cloning purposes, and *L. agilis* La3.

Once pVEcXLa was constructed, promoters (Table 14) were cloned upstream of the *creiLOV* reporter gene. Promoters were amplified from the *L. agilis* La3 chromosome (Table 5) and each was cloned independently into pVEcXLa using *Eag*I and *BamH*I sites resulting

in the construction of five plasmids: pVEcXLa-$P_{psp}$, pVEcXLa-$P_{gap}$, pVEcXLa-$P_{eno}$, pVEcXLa-$P_{eft}$ and pVEcXLa-$P_{cwah}$ (Table 4). Cloning these promoters upstream of *creiLOV* proved to be highly effective and mutation-free, compared to the EGFP constructs.
Each of these constructs were successfully transformed into *L. agilis* La3 (method 2.19) at similar efficiencies (data not shown) and passaged to allow chromosomal recombination to occur (method 2.22). Despite the only difference between these 5 plasmids being the promoter region, each construct took a different number of passages before bacteria with insertion of the expression cassette into the chromosome could be isolated. In the most extreme example, integration of the pVEcXLa-$P_{psp}$ cassette into the *L. agilis* La3 chromosome was not detected, even after extensive passaging and selection. By comparison, the integration cassettes carried by the other four plasmids recombined into the chromosome within one to ten subcultures. These insertions were confirmed via PCR using primer pairs 5 and 6 (Table 5).

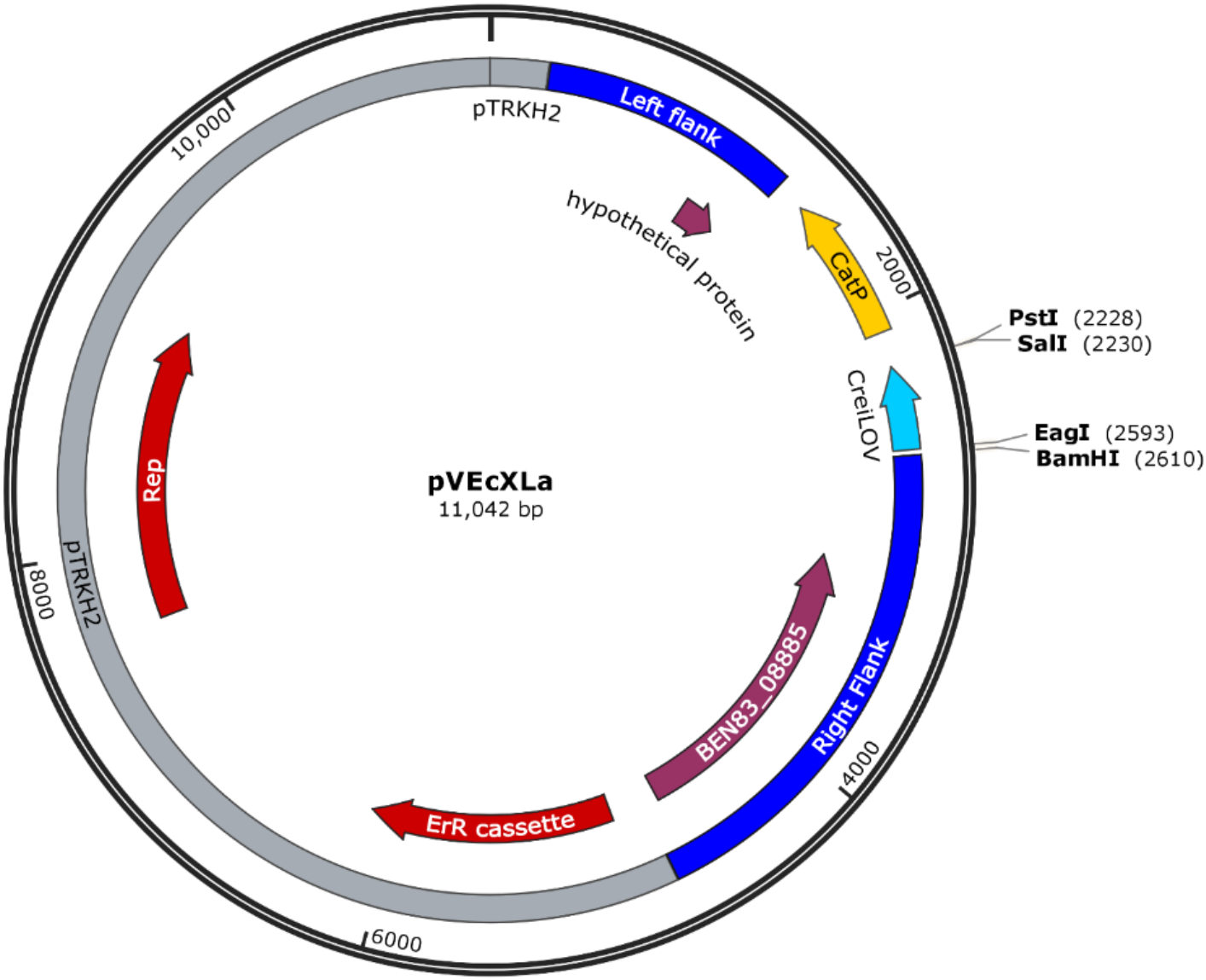


Figure 18: Graphical representation of pVEcXLa. pVEcXLa contains a Left and Right Flank allowing 100% stable, site-directed recombination into the *L. agilis* La3 chromosome. The chromosomal insertion cassette includes a codon-optimised *creiLOV* gene that is surrounded by unique restriction sites for cloning, and the *catP* gene to select for chromosomal recombinants on chloramphenicol.

#### 4.3.3.1 Attempted chromosomal integration of pVEcXLa-$P_{psp}$ using pTRKH2-*tetM*

To address the lack of integration of the pVEcXLa-$P_{psp}$ expression cassette into the chromosome, plasmid pTRKH2-*tetM* (Table 4) was constructed. This plasmid was constructed as an auxiliary tool to improve chromosomal recombination by exploiting plasmid incompatibility (Novick 1987). The mechanism works by having two plasmids that use identical replication machinery within the same cell (both have pTRKH2 rep genes and origins of replication), but under different selection pressure; tetracycline 70 µg / mL for pTRKH2-*tetM*, and chloramphenicol 10 µg / mL on the chromosomal insert of pVEcXLa plasmids). By selecting on both chloramphenicol and tetracycline, an attempt was made to force recombination and select for recombinants. However, transformation of the *L. agilis* La3 (pVEcXLa-$P_{psp}$) with pTRKH2-*tetM* did not alter chromosomal recombination frequency, and the lack of recombination from pVEcXLa-$P_{psp}$ remained unchanged.

### 4.3.4 Strength of promoter compared with protein expression level

To determine the difference between the 5 promoters and gene expression levels of CreiLOV, relative fluorescence units (RFUs) were determined (method 2.25). RFUs were measured in both *E. coli* JM109 and *L. agilis* La3. Of the promoters analysed a significant difference was only found between *E. coli* wild type and *E. coli* (pVEcXLa-$P_{gap}$) RFUs (Table 15) indicating that the $P_{gap}$ promoter allowed production of a statistically significant amount of protein (CreiLOV). No statistical difference was observed between *L. agilis* La3 wild type and *L. agilis* La3 containing any of the CreiLOV-expressing pVEcXLa plasmids (data not shown). When the effect size and relationship between protein expression levels or RFUs and relative transcript abundance or transcripts per million reads (TPM) was examined, there was a strong positive relationship in *E. coli* JM109 (Figure 19) and a moderate positive relationship in *L. agilis* La3 (Figure 20). High background fluorescence from *L. agilis* La3 combined with low fluorescence intensity of CreiLOV may have obscured the true relationship between promoter strength and protein expression levels, as seen in *E. coli*.

Table 15: Dunnett's multiple comparisons test between *E. coli* wild type and *E. coli* containing CreiLOV-expressing pVEcXLa plasmids under different *L. agilis* La3 promoters.

| Dunnett's multiple comparisons test | Mean Diff. | 95% CI of diff. | Adjusted P Value | Significant |
|---|---|---|---|---|
| No plasmid vs. pVEcXLa | -23,257 | -60,038 to 13,524 | 0.3108 | No |
| No plasmid vs. pVEcXLa-$P_{psp}$ | -19,199 | -55,980 to 17,582 | 0.4841 | No |
| No plasmid vs. pVEcXLa-$P_{gap}$ | -41,731 | -78,512 to -4,950 | 0.0240 | Yes |
| No plasmid vs. pVEcXLa-$P_{eno}$ | -21,415 | -58,196 to 15,366 | 0.3833 | No |
| No plasmid vs. pVEcXLa-$P_{eft}$ | -21,322 | -58,103 to 15,459 | 0.3874 | No |
| No plasmid vs. pVEcXLa-$P_{cwah}$ | -7,660 | -44,441 to 29,121 | 0.9743 | No |

## Transcripts per Million reads from *L. agilis* La3 vs Relative Fluorescence Units in *E. coli* JM109

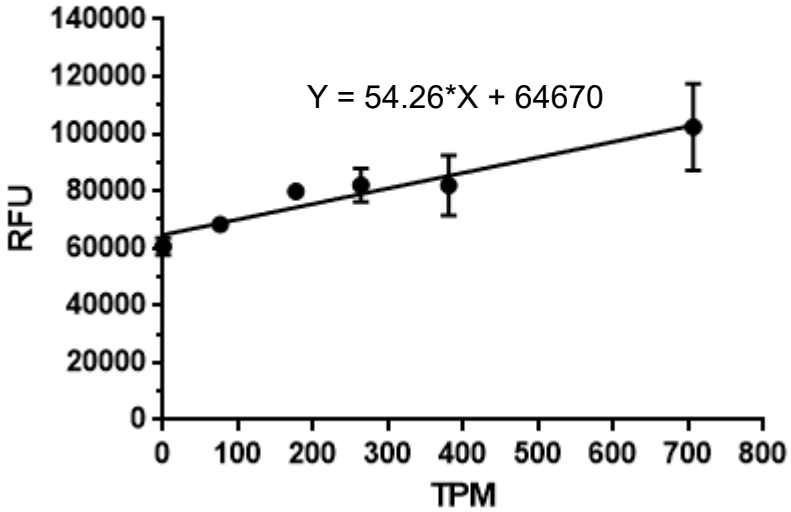


Figure 19: Graph showing relationship between relative fluorescent units (RFUs) from *E. coli* expressing CreiLOV under different *L. agilis* La3 promoters and the transcripts per million reads (TPM) of those promoters. Linear regression $r^2$ = 0.9325 indicating strong, positive relationship. n = 3. Error bars: ±SEM

## Transcripts per Million reads from *L. agilis* La3 vs Relative Fluorescence Units in *L. agilis* La3

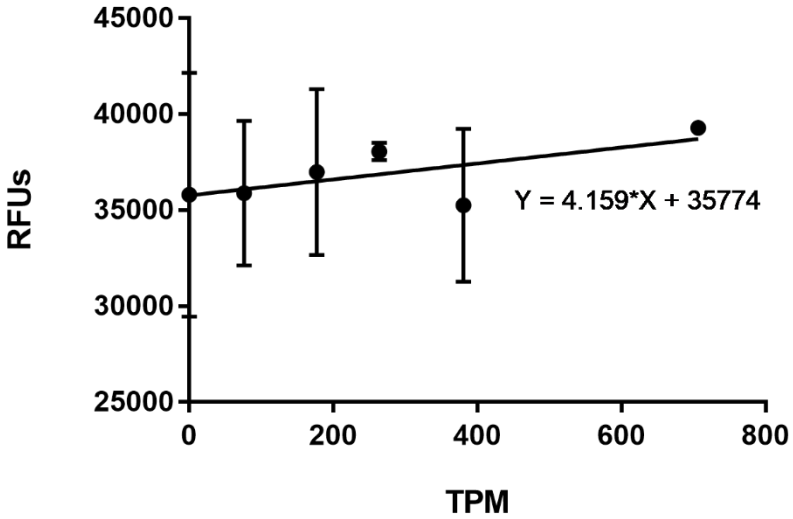


Figure 20: Graph showing relationship between relative fluorescent units (RFUs) from *L. agilis* La3 expressing CreiLOV under different *L. agilis* La3 promoters and the transcripts per million reads (TPM) of those promoters. Linear regression $r^2$ = 0.4637 indicating moderate, positive relationship. n = 3. Error bars: ±SEM

## 4.4 Chapter 4 Discussion:

### 4.4.1 Development of a stable expression system

#### 4.4.1.1 Plasmid-based system

The goal of this work was to construct a stable expression system for therapeutic antigens to develop *L. agilis* La3 as a live vaccine. Initially it was thought that a plasmid-based system would provide the simplest framework for stable expression, so pVEcLaSyn1 was designed and synthesised. When the pVEcLaSyn1 plasmid was used to transform *L. agilis* La3 no transformants could be recovered. This may have been due to the cell wall associated hydrolase (cwah) signal sequence and/or the promoter strength of the $P_{cwah}$ promoter in combination with EGFP expression having a toxic effect on the cell. Furthermore, when cloning the five *L. agilis* La3 predicted promoters upstream of EGFP in pTRKH2-LCER, there was a mixture of nonsense mutations in *egfp*, and disruptions to the predicted promoter sequences. EGFP has been previously demonstrated to not be compatible with particular signal sequences such as the maltose binding protein (Bassford *et al.* 1979), and may not have been compatible with the *cwah* signal sequence. It has also been shown on a number of occasions that strong promoters such as $P_{gap}$ will reduce plasmid copy number and stability in *E. coli* and yeast, as well as cause difficulties in obtaining clones (Stassi *et al.* 1982; Janes *et al.* 1990; Dillard *et al.* 1991; Silva *et al.* 1991). It is possible that the native plasmid backbone of pVEcLaSyn1 (pLa3_4) is a relatively high copy number plasmid due to its small size (2686 bp). Since both pLa3_4 and pVEcLaSyn1 contained the same replication protein and origin of replication, they would have been in direct competition for replication machinery, which most likely resulted in plasmid incompatibility (Novick 1987). While pVEcLaSyn1 was selected for, there is a 3-hour growth period before selective plating. This may have provided ample time for plasmid incompatibility to occur, and when combined with the low transformation frequency, resulted in no transformants. The metabolic load of a constitutively-expressed gene on a multi-copy plasmid may have selectively favoured mutations occurring that reduced pVEcLaSyn1 levels within cells, even under selection of chloramphenicol. If heterologous EGFP was expressed under the strong, native and constitutive $P_{ldhL}$ on this plasmid, it may have been too much of a burden for *L. agilis* La3 and was selected against, favouring instead the native pLa3_4 plasmid.

#### 4.4.1.2 Chromosomal-insertion system

##### 4.4.1.2.1 pTRH2-LCER

Attempts at constructing a plasmid-based system for expression of recombinant proteins in *L. agilis* La3 were not successful, so a chromosomal insertion system was constructed. The chromosomal insertion system, using pTRKH2-LCER, proved to be highly successful (Figure 14, Figure 15, and Figure 17). Chromosomal recombination has been demonstrated many times within *Lactobacillus* species through site specific *attP-attB* recombination (Auvray *et al.* 1999) and homologous recombination (van Zyl *et al.* 2015), and has been occasionally used in developing live vaccines (Layton *et al.* 2009; Layton *et al.* 2011). pTRKH2-LCER was constructed to have two regions of homology within the *L. agilis* La3 chromosome to allow double crossover via homologous recombination to occur, EGFP under the native $P_{cwah}$ promoter, and a selectable chloramphenicol resistance cassette. The

chosen chromosomal integration site had several characteristics which may have allowed fast and reliable insertion into the desired target location using pTRKH2-LCER. These characteristics include the low transcript abundance, indicating if any regular cell function was disrupted, it should theoretically have minimal impact, and being downstream of two converging genes, as to reduce the chances of impacting of gene transcription. 100% stability of the insert and expression of antibiotic resistance (chloramphenicol resistance) was achieved and demonstrated using phenotypic (Result 4.2.1), PCR (Figure 12) and Southern blot analysis (Figure 14, Figure 15 and Figure 17) The Southern blot (Figure 14) provided independent evidence of the accuracy of the recombination site utilised in the *L. agilis* La3 chromosome, as the chromosomal insertion site was digested into the exact fragment size predicted, indicating the correct *in silico* placement of that region of the genome.
The relative ease in which double crossovers occurred (briefly passaging under selection) indicated the chosen site was suitable for chromosomal recombination. If the integration of the expression cassette disrupted regular cellular function, there would be selection against integration.

#### 4.4.1.2.2 pVEcXLa and CreiLOV

The pVEcXLa expression system took the design of pTRKH2-LCER further by introducing more restriction enzyme sites and replacing *egfp* with *creiLOV*. Due to the mutations induced when cloning the promoters upstream of *egfp* in pTRKH2-LCER, it was thought that a different reporter gene would allow verification of the promoter strengths. pVEcXLa allowed successful chromosomal integration and formation of new, *L. agilis* La3 strains (Table 2), confirmed via antibiotic resistance profile and PCR. The only strain that was unable to be constructed using the pVEcXLa system was *L. agilis* La3::$P_{psp}$- *creiLOV* using pVEcXLa-$P_{psp}$, which, despite passaging for >200 generations, never resulted in chromosomal recombination. The mechanism at work here is unknown and further investigation would be required to elucidate the reasons for this outcome. However, since the only difference between the other pVEcXLa-Promoter plasmids (Table 4) and pVEcXLa-$P_{psp}$ is the 151 bp promoter region, the effect is most likely sequence specific. This may be caused due to sequence-specific secondary DNA structures forming, which have been previously shown to impact recombination, as well as gene regulation (Rawal *et al.* 2006; Pike *et al.* 2010; Bochman *et al.* 2012).

The pVEcXLa plasmid system, (Table 4) was demonstrated to be suitable for development of *L. agilis* La3 as a live vaccine due to several characteristics:

- High copy number in *E. coli* allowing efficient plasmid yields for *L. agilis* transformation, which requires 3 µg DNA
- Fairly small sized plasmids that are capable of efficient transformation into *L. agilis* La3
- Based on pTRKH2 (O'Sullivan *et al.* 1993) which is highly unstable without selection in *L. agilis* La3 over short generational time (<2% of colonies after ~100 generations) (Stephenson 2010; Vezina 2013), and reported here as 0% (Result 4.2.1, Table 17) and hence is efficiently lost from host cells.
- Simple induction of chromosomal insertion into desired site via homologous recombination (Figure 14, Figure 15)

    - Selective marker moves into chromosome
    - Long-term 100% stability of chromosomal insert (Result 4.2.1, Table 17).
- Unique restriction sites present within the plasmid, between the flanking chromosomal regions.
    - Allows ample flexibility for addition/replacement of promoters, signal proteins, cell wall anchoring proteins and protein coding sequences.

### 4.4.2 Identification of strong promoters by genome wide transcription analysis

In Chapter 3, five promoters were identified that had high transcriptional activity across different growth phases (Table 14). Determining transcripts via RNA sequencing in both logarithmic and stationary phases allowed the identification of constitutively-expressed genes. Genes were sorted by relative transcript abundance and the promoters of these genes predicted. In this chapter, these promoters were cloned upstream of *creiLOV*, a reporter gene, to determine the impact of these promoters on protein expression levels. The predicted promoters were successfully identified and functionality demonstrated by expressing the heterologous CreiLOV (Figure 20) in *L. agilis* La3.
The relationship between increasing antigen dose and immune response has been shown several times (Cardenas *et al.* 1993; Grangette *et al.* 2001; Reveneau *et al.* 2002; De Cort *et al.* 2013), so the success of the vaccine relies heavily on expression level of the therapeutic protein/s, ideally though a strong and constitutive promoter. This work demonstrated partial heterologous protein expression under the bioinformatically identified native promoters $P_{psp}$, $P_{gap}$, $P_{eno}$, $P_{eft}$, and $P_{cwah}$.
These results have identified several homologous *L. agilis* La3 promoters which all have higher TPM values than the L-lactate dehydrogenase promoter ($P_{ldhL}$) previously used in *L. agilis* La3 (Stephenson *et al.* 2011), which was found to be the 16$^{th}$ most abundant transcript (Figure 8). The GAPDH promoter ($P_{gap}$) produced the highest, constitutive transcript level, and has been used previously to express high levels of heterologous protein in the algae *Dunaliella salina* (Jia *et al.* 2012), which also produced the highest level of CreiLOV in both *E. coli* JM109 and *L. agilis* La3 (Figure 19 and Figure 20).

When these promoters were used to express CreiLOV in the pVEcXLa plasmid system within *L. agilis* La3, a moderate relationship was demonstrated between estimated promoter strength (TPM) and protein expression (RFU). Despite many of the hypothetical advantages of CreiLOV, it was found to not be a suitable reporter gene for *L. agilis* La3, however functioned well when used in *E. coli* (Figure 19) (Mukherjee *et al.* 2015). Determining the relative strength of promoters in *L. agilis* La3 using CreiLOV was confounded due to the high background autofluorescence of *L. agilis* La3 and low-level fluorescence of CreiLOV. The combination of these two issues made the relationship between relative transcript abundance (TPM) bestowed by the promoters and protein expression levels (RFU) less pronounced than the strong, positive relationship found in *E. coli* (Figure 19). Interestingly, the strongest predicted *L. agilis* La3 native promoters allowed high expression in *E. coli*, yet heterologous homolog promoters show significantly lower expression within *L. agilis* La3 (Stephenson *et al.* 2011) and other *Lactobacillus* strains (Kiatpapan *et al.* 2001). It is possible that a more accurate measurement of CreiLOV could be determined by using antibodies and ELISA, which is regularly used as a quantitation method when EGFP fluorescence is not appropriate (Jiménez *et al.* 2012). Regardless, a moderately positive

relationship ($r^2$ = 0.4637) was found between promoter strength and RFUs (Figure 20), indicating that a stronger promoter will result in higher protein expression.

Overall, this chapter demonstrated the development of a stable expression system (pVEcXLa) via double crossover homologous recombination of appropriate constructs into the *L. agilis* La3 chromosome. This expression system was used to test a range of highly expressing promoters, which could be used for expression of recombinant, therapeutic proteins. This system was utilised in Chapter 5 for the expression of antigens in *L. agilis* La3.

# 5. Chapter 5 Results: Therapeutic protein expression in *L. agilis* La3

Construction of the pVEcXLa plasmid system for the recombination of expression cassettes onto the *L. agilis* La3 chromosome allowed testing of the five predicted constitutive *L. agilis* La3 promoters (Chapter 4), which were demonstrated to be active (Figure 19, Figure 20). Although a moderate relationship was demonstrated between predicted promoter strength (based on relative transcript abundance) and protein levels (fluorescence), confounding factors such as the high autofluorescence of *L. agilis* La3 and weak fluorescence of CreiLOV made accurate determination of promoter strength difficult. To continue to address this aim, and also address the final aim of this thesis which was to test expression of potentially therapeutic proteins from *L. agilis* La3, all 5 promoters $P_{gap}$, $P_{psp}$, $P_{eno}$, $P_{eft}$, and $P_{cwah}$, were used to express a therapeutic protein, specifically antigen recombinant NetB (rNetB) (S254L), a ~32 – 34 kDa protein. rNetB has been used as an immunogenic vaccine antigen in previous studies, resulting in partial protection against necrotic enteritis (Jang *et al.* 2012; Costa *et al.* 2013; Keyburn *et al.* 2013; Keyburn *et al.* 2013; Lillehoj *et al.* 2017). The rNetB protein-encoding sequence was cloned into each of the five pVEcXLa-Promoter-*creiLOV* plasmids, replacing *creiLOV* (Table 4). By determining rNetB expression using specific antibodies together with Western blots, the fluorescence issues regarding the expression of CreiLOV from the promoters could be avoided. The strains constructed in this chapter demonstrate the development of the first potential *L. agilis* La3 live vaccine strain which was ready for efficacy testing in animal trials, although the latter experiments were not performed as part of these studies.

## 5.1 Cloning of *rnetB* into the pVEcXLa-Promoter-*creiLOV* plasmids

### 5.1.1 Codon-optimisation of *rnetB*

Originally, *rnetB* was amplified via PCR from a *C. perfringens* plasmid then cloned into the five promoter-carrying plasmids, replacing *creiLOV*. However, no rNetB was detected in either *E. coli* JM109, *E. coli* BL21 (DE3) or *L. agilis* La3 from these constructs (data not shown). Lack of expression was thought to be the result of rare codon usage present in the *C. perfringens* sequence, specifically the amino acid isoleucine which occurred 17 times within the protein. The majority of the isoleucines present in the sequence contained the rarely used ATA codon in both *E. coli* (used 7% of the time) and *L. agilis* La3 (used 8% of the time). A codon optimised *rnetB* (Table 4) was therefore designed, synthesised and cloned into each of the 5 pVEcXLa-Promoter plasmids, replacing *creiLOV* (Table 4).

## 5.2 rNetB expression in *E. coli* and *L. agilis* La3

### 5.2.1 *E. coli*

*E. coli* strain JM109 was transformed with the five rNetB-encoding pVEcXLa plasmids (Table 4) each containing the codon-optimised *rnetB* gene under the control of a different promoter ($P_{psp}$, $P_{gap}$, $P_{eno}$, $P_{eft}$ and $P_{cwah}$). Western blot analysis of cell lysates from *E. coli* containing the pVEcXLa-$P_{psp}$-*rnetB*, pVEcXLa-$P_{gap}$-*rnetB* or pVEcXLa-$P_{eft}$-*rnetB* plasmids showed positive rNetB expression (Figure 21). Determining relative expression rNetB was not possible due to variable protein loading. The lack of rNetB expression observed under the $P_{eft}$ and $P_{cwah}$ promoters may have also been due to the variable loading. Due to the polyclonal anti-rNetB antibodies used, a high level of background was seen, even after pre-absorbing the antibodies against *E. coli* JM109. This made screening large number of colonies, and clear conclusions about expression intensity difficult.

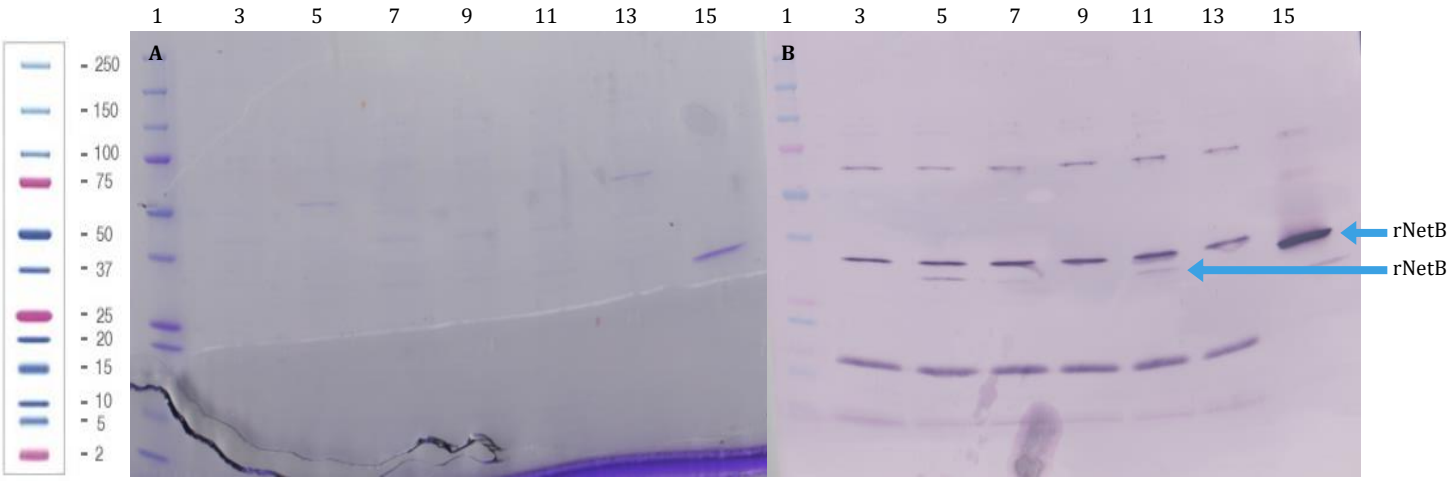


Figure 21: Coomassie stained gel (A) and accompanying Western blot (B) of *E. coli* JM109 transformed with pVEcXLa-Promoter-*rnetB* plasmids, harvested at log phase ($OD_{600}$ 0.4 – 0.6). Lane 1: Precision Plus Protein Dual Xtra (BIO-RAD). Lane 3: *E. coli* JM109 wild type. Lane 5: *E. coli* JM109 (pVEcXLa-$P_{psp}$-*rnetB*). Lane 7: *E. coli* JM109 (pVEcXLa-$P_{gap}$-*rnetB*). Lane 9: *E. coli* JM109 (pVEcXLa-$P_{eno}$-*rnetB*). Lane 11: *E. coli* JM109 (pVEcXLa-$P_{eft}$-*rnetB*). Lane 13: *E. coli* JM109 (pVEcXLa-$P_{cwah}$-*rnetB*). Lane 15: rNetB (purified). The pVEcXLa-$P_{psp}$-*rnetB*, pVEcXLa-$P_{gap}$-*rnetB* and pVEcXLa-$P_{eft}$-*rnetB* constructs show rNetB expression on the western blot. rNetB is ~32 – 34 kDa in size.

### 5.2.2 *L. agilis* La3

All five plasmids were then transformed into *L. agilis* La3 to construct *L. agilis* La3 (pVEcXLa-$P_{psp}$-*rnetB*), *L. agilis* La3 (pVEcXLa-$P_{gap}$-*rnetB*), *L. agilis* La3 (pVEcXLa-$P_{eno}$-*rnetB*), *L. agilis* La3 (pVEcXLa-$P_{eft}$-*rnetB*) and *L. agilis* La3 (pVEcXLa-$P_{cwah}$-*rnetB*). To compare between the non-integrating pVEcXLa-$P_{psp}$-*rnetB* isolates and the other four constructs, rNetB expression from cell lysates was examined via Western blot prior to the selection of chromosomal integrants (Figure 22). *L. agilis* La3 containing pVEcXLa-$P_{psp}$-*rnetB*, pVEcXLa-$P_{eft}$-*rnetB* and pVEcXLa-$P_{cwah}$-*rnetB* produced rNetB. None of the other colonies produced rNetB at a detectable level. Variable loading of the protein extracts may have contributed to the observed lack of expression and made it difficult to determine relative expression levels of rNetB (Figure 22). The small band at around ~32 – 34 kDa matched the positive control (purified rNetB) in lane 14. Colony-based screening of rNetB expression was

limited due to the non-specific binding of the anti-rNetB polyclonal antibodies, meaning higher-throughput methods such as ELISAs were not possible.

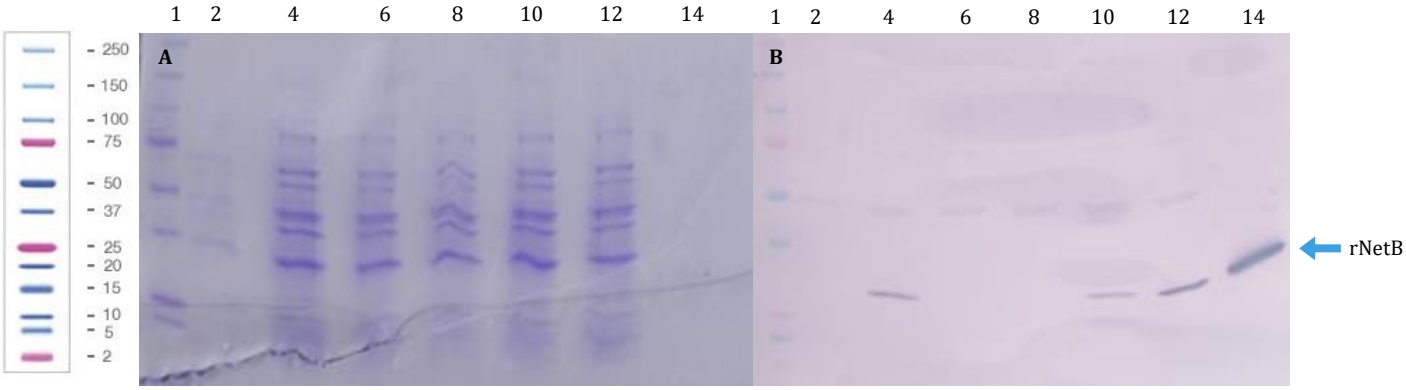


Figure 22: Coomassie stained gel (A) and accompanying Western blot (B) of *L. agilis* La3 transformed with pVEcXLa-Promoter-*rnetB* plasmids, harvested at mid-log phase ($OD_{600}$ 3.5 – 4.5). Lane 1: Precision Plus Protein Dual Xtra (BIO-RAD). Lane 2: *L. agilis* La3 wild type. Lane 4: *L. agilis La3* (pVEcXLa-$P_{psp}$-*rnetB*). Lane 6: *L. agilis La3* (pVEcXLa-$P_{gap}$-*rnetB*). Lane 8: *L. agilis La3* (pVEcXLa-$P_{eno}$-*rnetB*). Lane 10: *L. agilis La3* (pVEcXLa-$P_{eft}$-*rnetB*). Lane 12: *L. agilis La3* (pVEcXLa-$P_{cwah}$-*rnetB*). Lane 14: rNetB (purified). The pVEcXLa-$P_{psp}$-*rnetB*, pVEcXLa-$P_{eft}$-*rnetB* and pVEcXLa-$P_{cwah}$-*rnetB* constructs show rNetB expression on the Western blot. rNetB is ~32 – 34 kDa in size.

### 5.2.2.1 *rnetB* Quantitative PCRs

Following protein analysis, qPCRs were performed to determine RNA transcription levels from each of the expression cassettes (Table 16) to further investigate the differences in expression levels between the 5 promoters ($P_{psp}$, $P_{gap}$, $P_{eno}$, $P_{eft}$, $P_{cwah}$). This analysis was performed as $P_{gap}$ and $P_{eno}$ did not express rNetB in *L. agilis* La3 despite earlier demonstrating CreiLOV expression (Figure 20). RNA was isolated at the early-logarithmic growth phase (Figure 5) and cDNA was generated. qPCRs were performed using the resulting cDNA to determine transcription levels. The *rnetB* transcription levels were normalised between samples by using the housekeeping gene *dnaK* as an internal control. The results showed no statistical difference between promoters on *rnetB* transcription levels. Cq differences between all groups was ±0.08, indicating little variation.

Table 16: Normalised qPCR results of *rnetB* between different promoter constructs within *L. agilis* La3. Cq values were normalised by dividing Cq of internal control *dnaK* by Cq of *rnetB*. All were highly similar, indicating no difference between transcript abundance of *rnetB* when under different promoters in *L. agilis* La3.

| Sample | Normalised *rnetB* Cq |
|---|---|
| *L. agilis* La3 | 0 |
| *L. agilis* La3 (pVEcXLa-$P_{psp}$-*rnetB*) | 0.89 |
| *L. agilis* La3 (pVEcXLa-$P_{gap}$-*rnetB*) | 0.86 |
| *L. agilis* La3 (pVEcXLa-$P_{eno}$-*rnetB*) | 0.92 |
| *L. agilis* La3 (pVEcXLa-$P_{eft}$-*rnetB*) | 0.87 |
| *L. agilis* La3 (pVEcXLa-$P_{cwah}$-*rnetB*) | 0.84 |

### 5.2.2.2 *L. agilis* La3 chromosomal expression of rNetB

Once rNetB expression had been detected from the pVEcXLa plasmid in *L. agilis* La3, expression from a chromosomal location was investigated. Recombination into the *L. agilis* La3 chromosome was selected for by briefly passaging the strains under chloramphenicol selection, then individual colonies were screened via PCR to determination if an integration event had occurred (Figure 23).

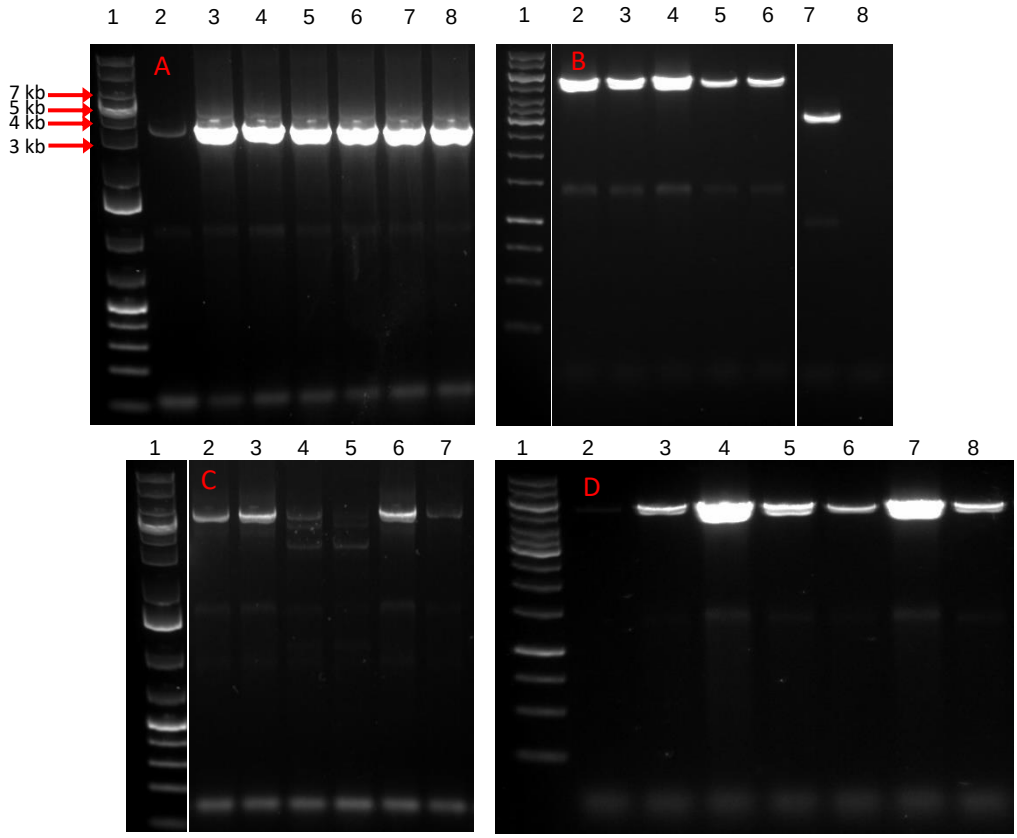


Figure 23: Gel electrophoresis images of double crossover diagnostic PCR (primer pair 5) on chloramphenicol resistant colonies after brief passaging to select for chromosomal recombinants.

A: *L. agilis* La3 (pVEcXLa-$P_{psp}$-rnetB) colonies. Lane 1: Generuler 1kb Plus. Lane 2: Colony #1. Lane 3: Colony #2. Lane 4: Colony #3. Lane 5: Colony #4. Lane 6: Colony #5. Lane 7: Colony #6. Lane 8: Colony #7. All colonies tested showed the wild type ~3,309 bp band, indicating no chromosomal recombination.

B: *L. agilis* La3::$P_{gap}$-*rnetB* colonies. Lane 1: Generuler 1kb Plus. Lane 2: Colony #1. Lane 3: Colony #2. Lane 4: Colony #3. Lane 5: Colony #4. Lane 6: Colony #5. Lane 7: Colony #6. Lane 8: No DNA control. 5/6 colonies tested showed the larger ~5,000 bp band, indicating integration of the expression cassette into the chromosome. One colony (Lane 7) did not contain the integrated cassette and had a band of ~3,309 bp.

C: *L. agilis* La3::$P_{eno}$-*rnetB* colonies. Lane 1: Generuler 1kb Plus. Lane 2: Colony #1. Lane 3: Colony #2. Lane 4: Colony #3. Lane 5: Colony #4. Lane 6: Colony #5. All colonies tested showed the larger ~5,000 bp band, indicating integration of the expression cassette into the chromosome. Colonies 3 and 4 contained a mix of the larger ~5,000 bp and wild type ~3,309 bp bands, indicating a mixed population within these colonies. This phenomenon was seen commonly, and further passaging under selection would eventually select for chromosomal recombinants (which produce a ~5,000 bp band).

D: *L. agilis* La3::$P_{eft}$-*rnetB* and *L. agilis* La3::$P_{cwah}$-*rnetB* colonies. Lane 1: Generuler 1kb Plus. Lane 2: *L. agilis* La3::$P_{eft}$-*rnetB* colony #1. Lane 3: *L. agilis* La3::$P_{eft}$-*rnetB* colony #2. Lane 4: *L. agilis* La3::$P_{eft}$-*rnetB* colony #3. Lane 5: *L. agilis* La3::$P_{eft}$-*rnetB* colony #4. Lane 6: *L. agilis* La3::$P_{cwah}$-*rnetB* colony #1. Lane 7: *L. agilis* La3::$P_{cwah}$-*rnetB* colony #2. Lane 8: *L. agilis* La3::$P_{cwah}$-*rnetB* colony #3. All colonies tested showed the larger ~5,000 bp band, indicating integration of the expression cassette into the chromosome.

Positive integrants were cultured to early-stationary phase ($OD_{600}$ 5 – 6) and cell lysates screened for rNetB expression using Coomassie gels and Western blots (Figure 24). Several insertional isolates of *L. agilis*, *L. agilis* La3::$P_{eft}$-*rnetB* and *L. agilis* La3::$P_{cwah}$-*rnetB* showed rNetB expression, however, *L. agilis* La3::$P_{eno}$-*rnetB* and *L. agilis* La3::$P_{gap}$-*rnetB* did not (Figure 24). In addition to this, a single *L. agilis* La3 (pVEcXLa-$P_{psp}$-*rnetB*) isolate showed *rnetB* expression yet did not demonstrated chromosomal recombination (Figure 23). Levels of rNetB expression yields within promoter groups look superficially similar, based on relative band intensity.

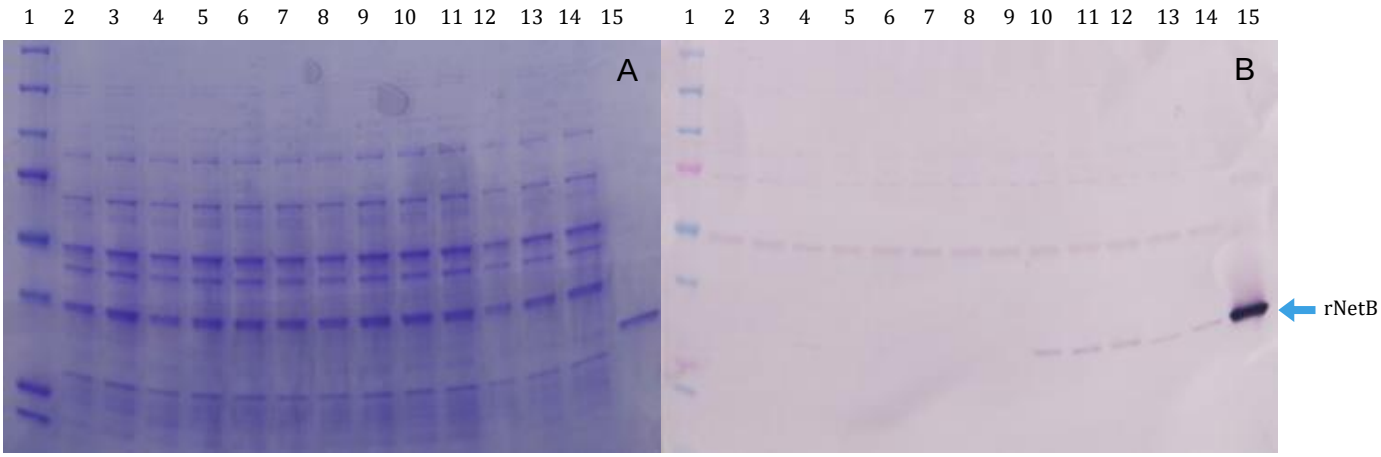


Figure 24: Coomassie stained gel (A) and accompanying Western blot (B) of *L. agilis* La3 chromosomal recombinants expressing rNetB, harvested at early stationary phase ($OD_{600}$ 5 – 6). Lane 1: Precision Plus Protein Dual Xtra (BIO-RAD). Lane 2: *L. agilis* La3 wild type #2 (negative control). Lane 3: *L. agilis* La3 WT #3 (negative control). Lane 4: *L. agilis* La3 (pVEcXLa-$P_{psp}$-*rnetB*) #1. Lane 5: *L. agilis* La3 (pVEcXLa-$P_{psp}$-*rnetB*) #2. Lane 6: *L. agilis* La3 (pVEcXLa-$P_{psp}$-*rnetB*) #3. Lane 7: *L. agilis* La3::$P_{gap}$-*rnetB* #2. Lane 8: *L. agilis* La3::$P_{eno}$-*rnetB* #1. Lane 9: *L. agilis* La3::$P_{eno}$-*rnetB* #3. Lane 10: *L. agilis* La3::$P_{eft}$-*rnetB* #1. Lane 11: *L. agilis* La3:: $P_{eft}$-*rnetB* #2. Lane 12: *L. agilis* La3::$P_{eft}$-*rnetB* #3. Lane 13: *L. agilis* La3::$P_{cwah}$-*rnetB* #1. Lane 14: *L. agilis* La3:: $P_{cwah}$-*rnetB* #2. Lane 15: 0.69 ng rNetB (purified). *L. agilis* La3 (pVEcXLa-$P_{psp}$-*rnetB*) #1, *L. agilis* La3::$P_{eft}$-*rnetB* #1 - 3, and *L. agilis* La3::$P_{cwah}$-*rnetB* #1 - 2 show rNetB expression. rNetB is ~32 – 34 kDa in size.

#### 5.2.2.3 Stability of double crossover inserts within the *L. agilis* La3 chromosome

To determine the stability of the rNetB-expressing inserts within the *L. agilis* La3 chromosome, a stability assay was performed (Table 17). 3 independent rNetB expressing isolates of *L. agilis* La3::$P_{eft}$-*rnetB* and *L. agilis* La3::$P_{cwah}$-*rnetB* were passaged for ~100 generations under no selection. *L. agilis* La3 (pVEcXLa-$P_{psp}$-*rnetB*) #1 was also included in this stability assay to determine its usefulness in continued development of *L. agilis* La3 as a live vaccine, even though it expressed rNetB from a plasmid location (due to the lack of chromosomal recombination). After ~100 generations, colonies were plated onto MRS, then patched onto MRS plates containing chloramphenicol to determine their resistance profiles and plasmid presence or absence. PCR using primer pair 5 was also used on these initial MRS colonies to determine if the chromosomal insert was present, which was confirmed in

each integrant, with the exception of the *L. agilis* La3 (pVEcXLa-$P_{psp}$-*rnetB*) #1 colony, which did not have a chromosomal integration (Figure 25). This assay once again demonstrated the highly stable insertion that occurs into the *L. agilis* La3 chromosome via homologous recombination using pVEcXLa (with the exception of pVEcXLa-$P_{psp}$-*rnetB*), as seen in Table 17. The plasmid was lost quickly over time, as demonstrated with the *L. agilis* La3 (pVEcXLa-$P_{psp}$-*rnetB*) #1 colony. This indicated pVEcXLa-$P_{psp}$-*rnetB* was not suitable for expression of rNetB, but the *L. agilis* La3::$P_{eft}$-*rnetB* and *L. agilis* La3::$P_{cwah}$-*rnetB* strains were suitable for continued live vaccine development.

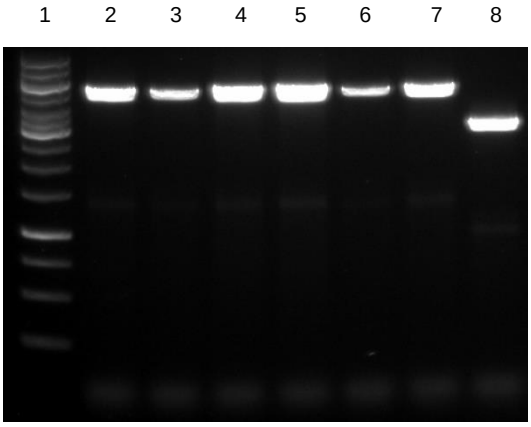


Figure 25: Gel electrophoresis images of double crossover diagnostic PCR (primer pair 5) on independent colonies passaged for ~100 generations under no selection. This was done to determine stability of the chromosomal inserts. Lane 1: Generuler 1kb Plus. Lane 2: *L. agilis* La3::$P_{eft}$-*rnetB* #1. Lane 3: *L. agilis* La3::$P_{eft}$-*rnetB* #2. Lane 4: *L. agilis* La3::$P_{eft}$-*rnetB* #3. Lane 5: *L. agilis* La3::$P_{cwah}$-*rnetB* #1. Lane 6: L. agilis La3::$P_{cwah}$-*rnetB* #2. Lane 7: L. agilis La3::$P_{cwah}$-*rnetB* #3. Lane 8: *L. agilis* La3 (pVEcXLa-$P_{psp}$-*rnetB*).

Table 17: Results of double crossover insert stability in *L. agilis* La3 chromosome of rNetB-expressing *L. agilis* La3 strains. Chromosomal inserts were stable 100% of the time. n = 51

| **Group** | **Stability (%)** |
|---|---|
| *L. agilis* La3 (pVEcXLa-$P_{psp}$-*rnetB*_La3) #1 | 0 |
| *L. agilis* La3::$P_{eft}$-*rnetB* #1 | 100 |
| *L. agilis* La3::$P_{eft}$-*rnetB* #2 | 100 |
| *L. agilis* La3::$P_{eft}$-*rnetB* #3 | 100 |
| *L. agilis* La3::$P_{cwah}$-*rnetB* #1 | 100 |
| *L. agilis* La3::$P_{cwah}$-*rnetB* #2 | 100 |
| *L. agilis* La3::$P_{cwah}$-*rnetB* #3 | 100 |

#### 5.2.2.4 rNetB expressed from the *L. agilis* La3 chromosome

Once the insert stability was determined to be 100%, the expression of rNetB from these strains was investigated. Screening rNetB expression was performed using Western blot analysis instead of quantitative methods such as ELISA due to the high-background of the anti-rNetB polyclonal antibodies. To examine rNetB expression, a small sample of four randomly selected colonies from each 100% stable independent isolate (Table 17) that had

been passaged under no selection for ~100 generations were grown to mid-stationary phase ($OD_{600}$ 5 – 6) and screened for rNetB expression.
All independent isolates from both *L. agilis* La3::$P_{eft}$-*rnetB* (Figure 26) and *L. agilis* La3::$P_{cwah}$-*rnetB* (Figure 27) demonstrated rNetB expression in all four randomly selected colonies from each strain. This indicated consistent expression within each isolate population. rNetB expression levels looked highly similar within strains.

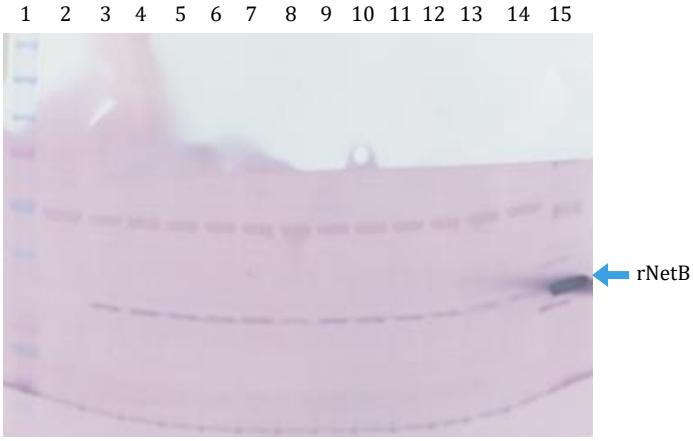


Figure 26: Western blot of *L. agilis* La3::$P_{eft}$-*rnetB* colonies expressing rNetB, harvested at early stationary phase ($OD_{600}$ 5 – 6) after being passaged for ~100 generations without selection. Lane 1: Precision Plus Protein Dual Xtra (BIO-RAD). Lane 2: *L. agilis* La3 wild type #1. Lane 3: *L. agilis* La3::$P_{eft}$-*rnetB* #1-1. Lane 4: *L. agilis* La3::$P_{eft}$-*rnetB* #1-2. Lane 5: *L. agilis* La3::$P_{eft}$-*rnetB* #1-3. Lane 6: *L. agilis* La3::$P_{eft}$-*rnetB* #1-4. Lane 7: *L. agilis* La3::$P_{eft}$-*rnetB* #2-1. Lane 8: *L. agilis* La3::$P_{eft}$-*rnetB* #2-2. Lane 9: *L. agilis* La3::$P_{eft}$-*rnetB* #2-3. Lane 10: *L. agilis* La3::$P_{eft}$-*rnetB* #2-4. Lane 11: *L. agilis* La3::$P_{eft}$-*rnetB* #3-1. Lane 12: *L. agilis* La3::$P_{eft}$-*rnetB* #3-2. Lane 13: *L. agilis* La3::$P_{eft}$-*rnetB* #3-3. Lane 14: *L. agilis* La3::$P_{eft}$-*rnetB* #3-4. Lane 15: 0.69 ng rNetB (purified). The 3 independent transformants (#1, #2, #3) demonstrated rNetB expression in all 4 of their randomly selected colonies.

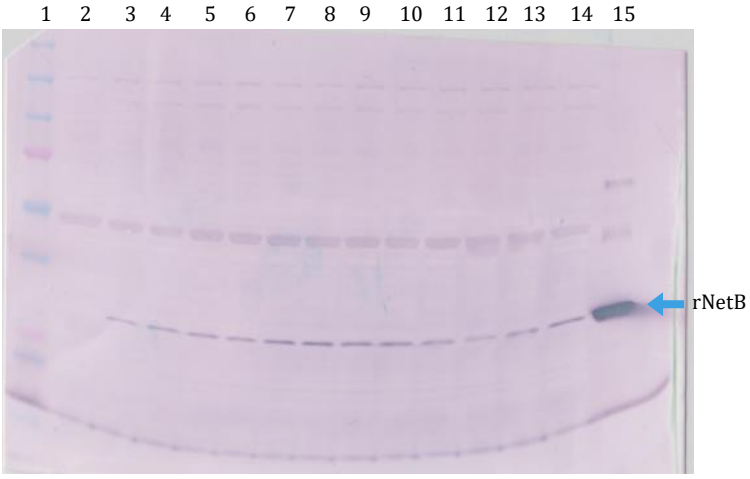


Figure 27: Western blot of *L. agilis* La3::$P_{cwah}$-*rnetB* colonies expressing rNetB, harvested at early stationary phase ($OD_{600}$ 5 – 6) after being passaged for ~100 generations without selection. Lane 1: Precision Plus Protein Dual Xtra (BIO-RAD). Lane 2: *L. agilis* La3 wild type #1. Lane 3: *L. agilis* La3::$P_{cwah}$-*rnetB* #1-1. Lane 4: *L. agilis* La3::$P_{cwah}$-*rnetB* #1-2. Lane 5: *L. agilis* La3::$P_{cwah}$-*rnetB* #1-3. Lane 6: *L. agilis* La3::$P_{cwah}$-*rnetB* #1-4. Lane 7: *L. agilis* La3::$P_{cwah}$-*rnetB* #2-1. Lane 8: *L. agilis* La3::$P_{cwah}$-*rnetB* #2-2. Lane 9: *L. agilis* La3::$P_{cwah}$-*rnetB* #2-3. Lane 10: *L. agilis* La3::$P_{cwah}$ -*rnetB* #2-4. Lane 11: *L. agilis* La3::$P_{cwah}$-*rnetB* #3-1. Lane 12: *L. agilis* La3::$P_{cwah}$-*rnetB* #3-2. Lane 13: *L. agilis* La3::$P_{cwah}$-*rnetB* #3-3. Lane 14: *L. agilis* La3::$P_{cwah}$-*rnetB* #3-4. Lane 15: 0.69 ng rNetB (purified). The 3 independent transformants (#1, #2, #3) demonstrated rNetB expression in all 4 of their randomly selected colonies.

#### 5.2.2.5 Semi-quantitative assay of rNetB expression levels

Next, to determine if rNetB expression had changed over ~100 generations under no selection, relative expression levels were determined by visual comparison to known-standards of purified rNetB (Figure 28). Both *L. agilis* La3::$P_{eft}$-*rnetB* #1 and *L. agilis* La3::$P_{cwah}$-*rnetB* #2 show similar levels of rNetB expression before and after passaging. The bands from *L. agilis* La3::$P_{eft}$-*rnetB* #1 and *L. agilis* La3::$P_{cwah}$-*rnetB* #2, which directly represented culture supernatants with no concentration or dilution, were estimated to contain ~150 pg/µL. After accounting for the 5 µL of a 5X concentration loaded on the gel, the yield was estimated to be ~ 6 pg / µL. Table 18 shows a summary of the developed pVEcXLa plasmids and their results across each of these experiments.

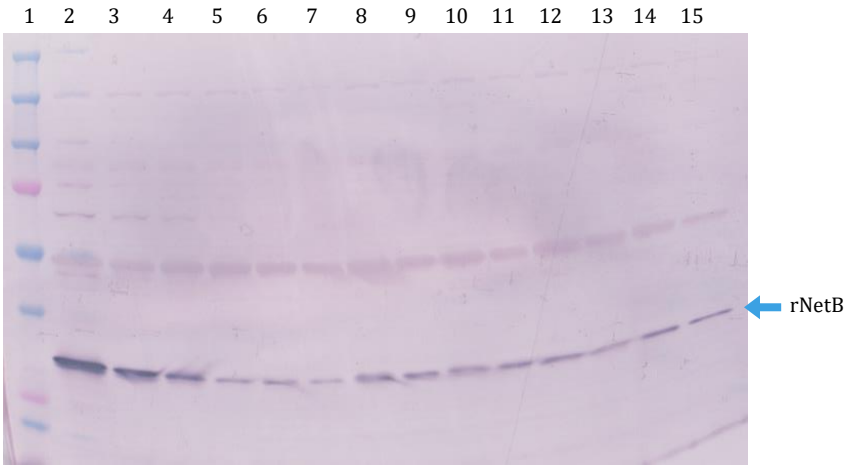


Figure 28: Western blot of *L. agilis* La3::$P_{eft}$-*rnetB* and *L. agilis* La3::$P_{cwah}$-*rnetB* colonies expressing rNetB, harvested at early stationary phase ($OD_{600}$ 5 – 6) either pre- or post-passaging for ~100 generations without selection. Lane 1: Precision Plus Protein Dual Xtra (BIO-RAD). Lane 2: rNetB 400 pg (purified). Lane 3: rNetB 300 pg (purified). Lane 4: rNetB 200 pg (purified). Lane 5: rNetB 100 pg (purified). Lane 6: rNetB 75 pg (purified). Lane 7: rNetB 50 pg (purified). Lane 8: *L. agilis* La3::$P_{eft}$-*rnetB* #1 pre-100 generations. Lane 9: *L. agilis* La3::$P_{eft}$-*rnetB* #1 pre-100 generations 1:2 dilution. Lane 10: *L. agilis* La3::$P_{eft}$-*rnetB* #1-2 post-100 generations. Lane 11: *L. agilis* La3::$P_{eft}$-*rnetB* #1-2 post-100 generations 1:2 dilution. Lane 12: *L. agilis* La3::$P_{cwah}$-*rnetB* #2 pre-100 generations. Lane 13: *L. agilis* La3::$P_{cwah}$-*rnetB* #2 pre-100 generations 1:2 dilution. Lane 14: *L. agilis* La3::$P_{cwah}$-*rnetB* #2-2 post-100 generations. Lane 15: *L. agilis* La3::$P_{cwah}$-*rnetB* #2-2 post-100 generations 1:2 dilution. Both *L. agilis* La3::$P_{eft}$-*rnetB* #1 and *L. agilis* La3::$P_{cwah}$-*rnetB* #2 show similar levels of rNetB expression before and after passaging.

Table 18: Summary of the pVEcXLa-Promoter-*rnetB* constructs and their phenotypic results in different strain backgrounds.

| ***rnetB* construct** | **Ranked transcript abundance (TPM)** | **Expression in *E. coli* JM109** | **Expression in *L. agilis* La3** | **Chromosomal recombination into *L. agilis* La3** | **Stable expression in *L. agilis* La3 after ~100 generations** | **Suitable as live vaccine*** |
|---|---|---|---|---|---|---|
| pVEcXLa-$P_{gap}$-*rnetB* | 1st | No | No | Yes | No | No |
| pVEcXLa-$P_{eft}$-*rnetB* | 3rd | Yes | Yes | Yes | Yes | Yes |
| pVEcXLa-$P_{eno}$-*rnetB* | 5th | No | No | Yes | No | No |
| pVEcXLa-$P_{psp}$-*rnetB* | 6th | Yes | Yes | No | No | No |
| pVEcXLa-$P_{cwah}$-*rnetB* | 14th | No | Yes | Yes | Yes | Yes |

*'Suitable as live vaccine' requires chromosomal integration into *L. agilis* La3 and stable expression after ~100 generations

## 5.3 Chapter 5 Discussion:

### 5.3.1 rNetB expression

The aims of this Chapter were to further test the five identified promoters from *L. agilis* La3 by using them to express a potentially therapeutic antigen. This also attempted to satisfy the next aim by expressing a therapeutic antigen, specifically the vaccine antigen rNetB, to further develop *L. agilis* La3 as a live vaccine. *L. agilis* La3-specific codon optimisation was essential for expression of rNetB (Result 5.1.1) and allowed expression in both *L. agilis* La3 and *E. coli*. Analysing rNetB expression across *E. coli* and *L. agilis* La3 demonstrated some perplexing inconsistencies (Table 18). pVEcXLa-$P_{gap}$-*rnetB* did not demonstrate rNetB expression in *E. coli* JM109 or *L. agilis* La3 (Figure 22, Figure 24), despite allowing CreiLOV expression in both *E. coli* and *L. agilis* La3 (when using pVEcXLa-$P_{gap}$-*creiLOV*). pVEcXLa-$P_{eno}$-*rnetB* had a similar relationship. pVEcXLa-$P_{cwah}$-*rnetB* allowed expression of rNetB in *L. agilis* La3, but not *E. coli* JM109, despite pVEcXLa-$P_{cwah}$-*creiLOV* allowing expression of CreiLOV in *E. coli* (Figure 19). There was no reasonable explanation for this, as the only changes between the CreiLOV-encoding and the rNetB-encoding plasmids was the coding sequences. When used in other plasmids the rNetB coding sequence (namely pVEcXLa-$P_{psp}$-*rnetB,* pVEcXLa-$P_{eft}$-*rnetB* and pVEcXLa-$P_{cwah}$-*rnetB*) was expressed in *L. agilis* La3 (Figure 24,Figure 26Figure 27).

The two plasmids that allowed expression of rNetB (pVEcXLa-$P_{eft}$-*rnetB* and pVEcXLa-$P_{cwah}$-*rnetB*) were shown to be highly stable over time, both in terms of the chromosomal insert (Figure 25), and rNetB expression levels (Figure 26, Figure 27).

#### 5.3.1.1 Expression levels of rNetB

The low expression levels of rNetB observed (~6 pg / µL) (Figure 28) were unexpected, as promoters and ribosomal binding sites were taken from the highest transcribed genes identified through RNA sequencing of *L. agilis* La3 (Figure 8). 6 pg/µL is considered a low yield in protein expression systems, which have been known to reach up to 394.2 ± 26.1 µg/mL for other recombinant proteins when using *E. coli* lab strains such as BL21 (DE3) (Lai *et al.* 2014). However, considering the highly stable, long term expression of rNetB, combined with *L. agilis* La3's colonisation and persistence capabilities within the chicken gastrointestinal tract, this may be enough rNetB expression to induce a local mucosal immune response without impacting colonisation fitness.

In Chapter 4, the CreiLOV expression level was only relatively quantified using RFUs, so absolute expression levels are not known. Therefore, from these results, it was not possible to compare the exact expression levels between CreiLOV and rNetB. However, in each pVEcXLa-Promoter plasmid (Table 4), CreiLOV was successfully expressed, yet only the pVEcXLa-$P_{psp}$-*rnetB*, pVEcXLa-$P_{eft}$-*rnetB* and pVEcXLa-$P_{cwah}$-*rnetB* plasmids resulted in rNetB expression within *L. agilis* La3 (Table 18). This indicated potential toxicity or selection against rNetB expression in the pVEcXLa-$P_{eno}$-*rnetB* and pVEcXLa-$P_{gap}$-*rnetB* constructs within *L. agilis* La3. The qPCR results showed similar levels of gene

transcription between constructs (Table 16), despite no detectable rNetB expression on the Western blot from either pVEcXLa-$P_{eno}$-*rnetB* or pVEcXLa-$P_{gap}$-*rnetB* (Figure 22 and Figure 24). This result was confusing, considering that all constructs and chromosomal recombinants were sequenced. A possible explanation may be that there were underlying issues with translation, as transcript levels were similar between the constructs (Table 16). This problem could be related to the size of the protein (rNetB: 293 in comparison to CreiLOV: 119 residues) (Onodera *et al.* 1996) or misfolding or aggregation in the cytoplasm (Schlegel *et al.* 2013). It has been shown in *E. coli* that by exporting heterologous protein out of the cytoplasm to the periplasm optimal protein expression resulted (Schlegel *et al.* 2013). However, as to why only two of the five constructs demonstrated rNetB expression despite all constructs inducing CreiLOV expression under control of the same promoters is not clear. Alternatively, the qPCR may have been contaminated, and repeating it further may resolve this issue.

Regarding secondary DNA structures, it is possible that the region of insertion into the *L. agilis* La3 chromosome also has an impact on rNetB expression when combined with promoter strength, as different sites within the genome can have an impact on expression levels (Rawal *et al.* 2006; Pike *et al.* 2010; Bochman *et al.* 2012; Bryant *et al.* 2014). The chromosomal insertion site was chosen partially based on its relatively low level of transcription. The working hypothesis was that by going into a very lowly transcribed region with no identified protein coding sequences, the chances of an insertion in the region disrupting any normal cellular function would be low. Avoiding any region that might express an uncharacterised regulatory RNA was also a consideration. However, considering the results obtained, it might be possible that this low transcriptional area may have structural barriers to high expression and this may explain the low levels of rNetB expression in *L. agilis* La3::$P_{eft}$-*rnetB* and *L. agilis* La3::$P_{cwah}$-*rnetB* (Figure 28). Variation in protein expression of up to ~300-fold has been demonstrated by inserting an identical expression cassette into multiple different regions of the *E. coli* chromosome (Bryant *et al.* 2014). This shows the importance of chromosomal location for therapeutic gene expression. It is therefore possible that the promoters identified in this study are not the strongest but rather have the highest levels of transcription due to the combination of their promoter structure and chromosomal location. However, the CreiLOV expression results indicated a positive relationship between transcript abundance and protein expression levels, so it is likely that they are the strongest promoters. The data presented here showed that promoters and ribosomal binding sites taken from highly transcribed genes (Table 14) may not equate to high expression levels when expressed in a different place within the chromosome. It is therefore likely to be of value to investigate expression from different insertion sites within the *L. agilis* La3 chromosome.

Alternatively, high levels of therapeutic protein expression may be sub-optimal for *L. agilis* La3 inside the highly competitive chicken gastrointestinal tract, as expressing large amounts of a non-essential protein may reduce genetic fitness and growth capacity. This hypothesis has been demonstrated via modelling which showed that the introduction of highly transcribed genes will have a global impact on other transcript levels via ribosome binding competition (Raveh *et al.* 2016). A colonisation and persistence chicken trial would a suitable way to determine this.

Strategies to circumvent this significant limitation in expression and increase expression may be achieved via secretion of recombinant proteins from the cytoplasm (Sletta *et al.* 2004; Mergulhão *et al.* 2005; Rosano *et al.* 2014). This approach has the added benefit of potentially improving host IgA and IgG immune responses by exposing the antigenic sequences to host immune systems (Ahmed *et al.* 2014; Kobierecka *et al.* 2016). Since there is a possibility chromosomal location has an impact on gene expression levels, fusion of therapeutic proteins with the highest transcribing genes identified during RNA sequencing should also be considered. For example; GAPDH is found on the surface of most *Lactobacillus* species (Antikainen *et al.* 2007). A fusion with GAPHD, the most abundantly transcribed gene in *L. agilis* La3 may allow far greater expression and remove the need for cloning signal sequences/membrane anchors as the fusion partner protein, GAPDH, is naturally found associated with the outer membrane. One study fused *C. jejuni* antigen genes and the CD154 gene to the chromosomally-based outer-membrane gene *lamB* of S. Enteritidis (Layton *et al.* 2011). Good expression was achieved, as the expression of the antigens and immune-enhancing CD154 ligand was under the native *lamB* promoter. Such a strategy could be employed for *L*. *agilis* La3

# 6. Overall Discussion, Conclusion and Future Directions

Taken together, the work described in results Chapters 3, 4 and 5 demonstrated the development of *L. agilis* La3 strains which could be used as potential live vaccines. Genomic characterization and transcriptomic analysis, development of the chromosomal recombination pVEcXLa expression system, and expression of therapeutic antigen rNetB was performed in this work. The developed potential live vaccine strains *L. agilis* La3::$P_{eft}$-*rnetB* and *L. agilis* La3::$P_{cwah}$-*rnetB* have several notable differences than previous live vaccines, as discussed below.

## 6.1 Comparison to other chicken live vaccines

Compared to other chicken live vaccines, *L. agilis* La3 is a highly persistent coloniser of the gastrointestinal tract. Most other strains are either attenuated or isolated from a different source and so do not perform ideally in the chicken gastrointestinal tract (Layton *et al.* 2009; Kulkarni *et al.* 2010; Layton *et al.* 2011; De Cort *et al.* 2013; Ragione *et al.* 2013; Łaniewski *et al.* 2014; De Cort *et al.* 2015; Kobierecka *et al.* 2016; Mishra *et al.* 2017). Partial to strong protection has been demonstrated in these studies, however, expression stability of the therapeutic protein is not usually discussed. Here in this study, long-term, selection-independent stable expression of a previously successfully used vaccine antigen rNetB (Costa *et al.* 2013; Keyburn *et al.* 2013; Keyburn *et al.* 2013) has been demonstrated.
Unlike other live vaccines, secretion signals or cell wall anchoring proteins (Layton *et al.* 2009; Layton *et al.* 2011; Alimolaeia *et al.* 2017; Lin *et al.* 2017) have not been added, nor has co-expression of immune-modifying peptides been introduced (Layton *et al.* 2009; Layton *et al.* 2011; Jee *et al.* 2017) via the pVEcXLa system. These are logical advancements in developing *L. agilis* La3 as a successful live vaccine. Although the two developed potential live vaccine strains *L. agilis* La3::$P_{eft}$-*rnetB* and *L. agilis* La3::$P_{cwah}$-*rnetB* have not been tested within a NE-challenge model, it is hypothesised that they will induce partial protection.

## 6.2 Improvements to pVEcXLa-$P_{eft}$-*rnetB*

The pVEcXLa plasmid system, specifically pVEcXLa-$P_{eft}$-*rnetB* and pVEcXLa-$P_{cwah}$-*rnetB* (Table 4), were demonstrated to be suitable for development of *L. agilis* La3 as a live vaccine due to several characteristics:

- High copy number in *E. coli* allowing efficient plasmid yields for *L. agilis* transformation, which requires 3 µg DNA
- Fairly small sized plasmid that is capable of efficient transformation into *L. agilis* La3
- Based on pTRKH2 (O'Sullivan *et al.* 1993) which is highly unstable without selection in *L. agilis* La3 over a short generational time (<2% of colonies after ~100

generations) (Stephenson 2010; Vezina 2013), and reported here as 0% (Result 4.2.1, Table 17).

- Simple induction of chromosomal insertion into the desired site via homologous recombination (Figure 14, Figure 15)
    - Selective marker moves into the chromosome
    - Long-term 100% stability of the chromosomal insert (Result 4.2.1, Table 17).
- Unique restriction sites present within the chromosomal insert region suitable for addition or replacement of promoters, signal proteins, cell wall anchoring proteins and protein coding sequences.
- Expression of codon-optimised rNetB under some of the strongest native promoters in *L. agilis* La3
    - Long-term, 100% stability of expression levels (Figure 26, Figure 27)

Because it has been demonstrated that transcript abundance has a strong, positive relationship with protein expression levels in *E. coli* (Figure 19), and a moderately positive relationship in *L. agilis* La3 (Figure 20), the future development of pVEcXLa-$P_{eft}$-*rnetB* should be strongly considered. The $P_{eft}$ promoter is the third strongest promoter within the *L. agilis* La3 chromosome, based on the TPM values, compared to the $P_{cwah}$ promoter which is the 14$^{th}$, and should allow greater expression levels (Figure 20). pVEcXLa-$P_{eft}$-*rnetB* also allows successful chromosomal recombination, and stable rNetB expression over time. Since antigen dose and immune responses are proportional in chickens (Cardenas *et al.* 1993; Curtiss III *et al.* 1996; Grangette *et al.* 2001; Heier *et al.* 2001; Reveneau *et al.* 2002; De Cort *et al.* 2013; Mishra *et al.* 2017), this makes pVEcXLa-$P_{eft}$-*rnetB* the most reasonable choice for continued development.

As *L. agilis* La3 will be potentially used in food, replacing the chloramphenicol resistance *catP* gene with food-grade selective markers such as uracil phosphoribosyltransferase (*upp*) or alanine racemase (*alr*) genes, previously used in lactic acid bacteria (LAB) is essential (Bron *et al.* 2002; Song *et al.* 2014) to reduce antibiotic resistance spreading within the chicken. The shortcoming of this approach is that *L. agilis* La3 Δ*upp* or Δ*alr* knockout mutants would be required. Codon-optimised cadmium resistance genes *cadA* and *cadC* from *L. lactis* (Wong *et al.* 2003) could be used to replace *catP*, provided *L. agilis* La3 has no innate cadmium resistance. Alternatively, a codon-optimised *Bifidobacterium longum* alpha-galactosidase gene (*aglL*) could be used to allow selection on melibiose (Sridhar *et al.* 2006), provided that the phenotype is not already present in *L. agilis* La3.

Rather than expressing a single antigen such as rNetB, co-expression or fusions with codon-optimised peptides or already secreted/anchored peptides could significantly improve immune responses and increase clearance of unwanted organisms. Immune modifiers such as IL-6 (Cho *et al.* 2004; Nishimichi *et al.* 2005; Niederstadt *et al.* 2012) or CD154 (Layton *et al.* 2009; Layton *et al.* 2011), as well as fusions with antigenic epitopes (Kobierecka *et al.* 2016), should be considered. Fusions with scFv protein linkers should

also be used to improve antibody-antigen interactions (Sletta *et al.* 2004; Yuvaraj *et al.* 2008; Jee *et al.* 2017).

## 6.3 Conclusion and future directions

Live vaccines are a practical and realistic solution to many infectious diseases, not only in the poultry industry but also to improve animal health in general. Looking towards the future, human health is likely to be partially managed by live vaccinations and paratransgenic delivery of therapeutic proteins. Live bacterial vaccine antigen vectors are also capable of expressing other therapeutic proteins such as bacteriocins (Shin *et al.* 2008; Messaoudi *et al.* 2012; Messaoudi *et al.* 2012; Lohans *et al.* 2014), host immune factors (Cho *et al.* 2004; Nishimichi *et al.* 2005; Niederstadt *et al.* 2012), digestive enzymes, and phage endolysins (Nariya *et al.* 2011; Gervasi *et al.* 2014). This project has advanced development of *L. agilis* La3 as a live delivery vehicle in several ways, by further characterising *L. agilis* La3, developing a stable expression system and testing expression of a therapeutic protein, rNetB.

Through single molecule real-time sequencing and high accuracy short-read sequencing the complete, circularised *L. agilis* La3 genome has been elucidated and the sequence has been annotated. This represents the first complete, circularised genome within the *L. agilis* species. Genetically it shows high similarity to *L. agilis* DSM 20509 and *L. agilis* Marseille despite different isolation sources, although they all share a similar niche within the gastrointestinal tract. RNA sequencing of *L. agilis* La3 has identified several highly expressed, constitutive genes that may be essential colonisation factors, indicating why *L. agilis* La3 is so effective at colonising and persisting within the chicken gastrointestinal tract, despite previously established microflora presence.

The pVEcXLa plasmid system was constructed in this study, allowing long-term, 100% stable integration of a selectable expression cassette into the *L. agilis* La3 chromosome via homologous recombination. The plasmid is highly customisable, small and chromosomal recombination is fast and simple to induce.

RNA sequencing was also used to determine the most abundant, constitutively transcribed genes and their promoters. These promoters were used to successfully express the heterologous proteins CreiLOV and rNetB within the pVEcXLa plasmid system. Two plasmids in particular, pVEcXLa-$P_{eft}$-*rnetB* and pVEcXLa-$P_{cwah}$-*rnetB*, demonstrated 100% stable, long-term selection-independent integration and expression of the therapeutic protein rNetB within the *L. agilis* La3 chromosome. Other genes such as other antigenic sequences or immune-modifying proteins can now be inserted up or downstream of the *rnetB* sequence. At the same time, signal sequence or cell wall anchor proteins can be cloned upstream, allowing proteins to be exported from the cell.

The continued development of *L. agilis* La3 as a live vaccine vector using pVEcXLa has great potential to reduce the incidence of broiler chicken-related diseases such as the *C. perfringens*-caused necrotic enteritis, as well protecting chickens against other disease-

causing pathogens such as *C. jejuni* and *Salmonella*. Good candidates for therapeutic protein expression should be considered to target these organisms.

When taken together, these results indicate two potential vaccine strains of *L. agilis* La3 have been constructed (*L. agilis* La3::$P_{eft}$-*rnetB* and *L. agilis* La3::$P_{cwah}$-*rnetB*) using the pVEcXLa plasmid system, and are ready for testing in animal trials. A necrotic enteritis challenge trial (Keyburn *et al.* 2013) should now be performed, to determine the efficacy of the two live vaccines developed in this work. Multiple dosage routes should be tested (coarse spray and in-feed), and these should be compared to a non-inoculated control group (MRS media only). Disease severity should be scored (Keyburn *et al.* 2006) to determine the protective effect, if any. NetB-producing *C. perfringens* and *L. agilis* La3 vaccine strains should be monitored over the course of the trial via direct isolation onto selective media from inoculated chickens. Blood and mucosal scrapings should also be taken to analyse specific anti-rNetB IgG and IgA antibodies, respectively.

# 7. Supplementary data

## 7.1 Supplementary data 1 – RAST annotation

>Glutamate N-acetyltransferase (EC 2.3.1.35) / N-acetylglutamate synthase (EC 2.3.1.1)
MLVKGAASKRDANVAAKAVVGSNLVKAALFGQDANWGRIISALGQTDVSIAVTKLDLAI
NQVAILKASQKQAYDEEGLQASLGQDQILIEVDLHTGQACGRAWGCDLTYKYVQINAS
YRS
>Acetylglutamate kinase (EC 2.7.2.8)
MQDLIVIKVGGNALDQLSSEFFEQLAAWRQAGKQILLVHGGGPMISKLCQKLNVPVTK
TDGVRVTDEQTLALTKLVLLGQTQPLLLQKLSDHHLNVQGLNAASNYLLKGRYLDKAK
YGQVGAITYVNQAALMAELKTKIGVLAPLALTESGEWLNVNADQAALAVASKLKAEKLY
LLTDVAGVLANGKLVAKLSENRAAKLVQKGIITSGMFPKVKAALTAQRSGISQVFITNSL
AKQGTQIESEE
>Acetylornithine aminotransferase (EC 2.6.1.11)
MDYVFPTYKRFPFEIVAGSDWELTDDQGKQYLDLTSGIGVCNLGYNQPKLNQAVAEQ
LTKVWHTSNLYESSLQDQVAQGIVNGKDYLVSFCNSGTEANEAALKLARKATGREKIL
AFNHSFHGRTYGSLSVTGNDQIKVGFTPMLAGVSFAEYNDDQALEQVSSDLAAVILEV
VQGEGGVYAADASWLKALAAKCQQVGALLIIDEVQTGMGRTGKLYAFENYGLKPDIFT
LAKGLANGIPVGAMVGKKSLGKYFGPGSHGSTFAGNPLAMAAAKEVLATMDSEFLEA
VAAKAEFAWHYLEDLTALECVTGVTGLGLMIGIHLSSEVKVAAVIEKLQAAGVLTLAAR
DNTLRLLPPLTIPASELLAGIEEIKAVLAEAILVGV
>Ornithine carbamoyltransferase (EC 2.1.3.3)
MENHFYGKSWLKELDYSAKELNFLIDFSLHLKDLKKKHIPHQYLAGKNIALLFEKTSTRT
RSAFTVTANDLGAHPEFLGQNDIQLGNKESLADTAKILGSMFDGIEFRGFKQDSVEQLA
KYSGVPVWNGLTDEWHPTQMIADFMTLKEEFGSLANLTLCYVGDGRNNMAHSLLVTA
AILGVNFHLAAPKDLQPKAEVLAIAQELAAESEAELLITTDPKQAVNGADAIYTDIWISMG
EKVDIHKRIEALLPYQVNQALLAQTGKQETIIMHCLPAFHDQETTLGKKLAAEYGLSELEI
TDEVFLGPNSRVFVEGENRLHSIKAIMAATLGELFIPEQYFN
>Galactoside O-acetyltransferase (EC 2.3.1.18)
MTVDINYVQNLTKAGQIYNDLDPKVVAARNHALKTVRAYNNKVYHQNDYDYQLLRPLF
KQLGSAPYIEPNFYCEFGFNLSIGDNFFANHDVVLMDCAPITIGNDVNIAPKVGIYTVNH
VEDPVARAHHQIYARPVVIEDGVWIGAGVNIVGGVTIGQNAIIGAGSVVTQDIPANVIAA
GNPAKVIRPLKTN
>DNA topoisomerase III (EC 5.99.1.2)
MKQLILAEKPSVARDLAQILGANQKNKHYFEGPKVVVTWALGHLMTLKMPEDLDKKW
QKWELETLPMIPKYIGIKPLPKTRPQLKAIGQLVKRKDIAEAVIATDAGREGELVARWILE
YVRFNKLVKRLWISSQTTKAVKEGFANLKPAKQYDNLYYSALARAKADWLVGLNVTRA
LTVKYQDNLSAGRVQTPTLALVGKKEAAISSFVPQKYYQVSLELNGVKAKLVQKNSQA
LKTSEQAQELVAKLRKTPGKVSEIKTKVKVEPAPLPYDLTELQREANSRYGYSAKKTLS
LVQSLYETHKVVSYPRTDSKYLSSDIKASLKERLQAVSEMFPKAKDYAKGQVKVVQAK
VFNDSKVTDHYALIPTEQRPNYAKFSNDEHRIYGLIVERFLGLFAPAHKVEQAKITVKFG
QALFNFKEERVIEAGWKKQGQTKQVTTNLKAGQAVAPNFKVEAALTTPPSPLTEGTLL
AQMEKYGLGTPATRAEIIEKLVKTELMERSGNHLRVTPKGKQLLKLVNPELVSPELTAK
WESQLEAIAKGKYASKQFLKEIEADTKRLVSEVKNSQAKYQDFSLTQKKCPECGSFLR
ERNGRDGKVLICTNPECSYRRRAEAKVSNHRCPQCHRKMLIVAGKNGDYFRCKYDGT
TEKMGGGKRGSKKMTKREERRLLQKVNQKDEDQESALAQALKAALGQE

>hypothetical protein
MVKSFVRFLWLVLAIVSFVGGIVGLLLPVIPQIPFFILCLFSLSKFSPRFHSWLQQTKLYQ
KYFLGFKRKLYSQADNLKEKEGSIKAQILAFFFQGEIEK
>Exodeoxyribonuclease V alpha chain (EC 3.1.11.5)
MDAKYQQARQLLAAGKVKAAQTIFREIFVLNPKYKSVCYFLGKCAYLQGKNKEVKHYL
KLHLDGRARKNIGPAFKYLCLLACKEKQYLKAQHYLAQAEKYGLKATSLAKLQKQLTQ
AASEPTTMAFTYQKLSKKGTPIASCKQAGFKGAKFFIHGLPPKLLARFQELKTLQLTVT
YPQKTDIPGYLTCRLTDGRTLYKDYLPYLPSLYYRSAFNDYLEDDKYFKFKWEKSRVF
SLLLVSGFSLEIVAKYLDANDLLDQAAAEHLQSCQGLVSLKEAGETLPVSAKEQLSDLD
LLKEFGFICAKECGKTIHLASKCRFERQLRRALKARSKVAFPYRKLQKQFRPSQEQARF
INWLKKNDQQVISLLGVGGSGKTYTLSKLLDPNAVLALAPTHKARLNLVDHGFSQNDT
VQHLLFKLEAEPELVLTGIKVILIDEISMVTTQMLAQLFDKLGPGYRYLLAGDDQQLPPV
SQDLDALNVCGDLMGLIKQEGSYYEFKTNLRCQNKQTAALIAAIRAKNLQLVTKKVGKF
TGAKRDMLNYKYFHPALEDCMILAHRNLTVGKINQQYYRVLTKDGRAKTPFFFKNDYG
HGGFFEGAQVVFYQNGDEKQRYGYTNSEFGIVTELHLEDKEPKIKVKTASNEYELPLY
VANRDLFLAYALTIHKAQGSGSRRVYMLEAANDSLLYTAVSRSRGELFFVDLDASQAV
EALSQPVAKKRNVY
>Cyclopropane-fatty-acyl-phospholipid synthase (EC 2.1.1.79)
MVLDNMLFKQLLKHSFNIPVEVTYPNGKVEQYGGLKPEVKIKINEKISLKALTENASIAL
GEAYMDGKIEVEGSLQKLITSAYANVDSFLNNNKFKKFLPKQSHTKSHSKEDVQSHYD
LGNDFYSLWLDKTMTYSCAYFENETDSLEQAQRNKVEHILKKLDPKPGKTLLDIGCGW
GTLMLTAAKKYNLKVVGITLSEEQAKFVRARIKDLGLEQQAEVILEDYREIQHEPFDYIT
SVGMFEHVGQENLGEYFQTIAKYLANDGVALIHGITRQQGGAYNGWIDKYIFPGGYIPG
LLENLGHIVDAGMQVADLETLRRHYQKTLEIWDQNFNQVLPAVRKMHDERFIRMWDL
YLQACAASFESGNIDCIQYLISKGPSGKNLPKTRAYMYQA
>Undecaprenyl-diphosphatase (EC 3.6.1.27)
MVELFKAFLFGIVEGITEWLPISSTGHMILLDEFVKLDVSKDFYSMFEVVIQLGAIMAVVV
IYWHRIWPFRLAKANGIGAPQGLWRFIDKETLIMWVKIFIACLPAGIIGVKFNDTFEKLFY
NPVCVALALIIFGVAFIWIETIHKGSRPKINSIAELSYYTVLMIGFFQLIAAVFPGTSRSGAT
ILGALLLGVSRTVAAEFTFFLAIPVMFGAAVFKLAKFGLVFSSMEMWIMVVGLITAFVSSI
VVIKFLMQYIKKHDFKVFGWYRIILGILVLAYFWLLK
>FIG00754943: hypothetical protein
MQFSFKRKNPRLYLSYSLAFWGLASLIYGLYLITGHQLVWYIDAANQHLPILESYRQTLL
DFIHHPSAGFPTWSPKMGLGADTFSVYSYYVIGDIFDYLTLLVPASKVVLAYQVLIILRM
YCAGLAFCFFASHFNFRNSTIVAGAAVYLFNAFLLYANVAHPFFIMPFIIFPLVIWALERV
LMQKSAWPLLLSFTWMLANNFYFAFMLGIGAIIYLLLRLAFTYKNKLNYLKTFGKLAWAS
LISLLLAAALLLPEFFAVTSSTRAGSAFANGLKVYPLYYYLTLPGQLINGANRDFYFWSA
LGFASISFYALVYVFSYRKKYPVLTASFSLGFVMFLIPAVGALFNGGMSPSNRWSLMLC
LPMALAVCFLLEALPKLDYRLLKRFVWATLIYSIYLGGYYFFDSDKRIFFPLLFLYASLGVI
SWFYFKPARFSQQLLSGLILLNVLVNALYFEAPYAGNYSNELLPKNAYQKVQANSYFDL
DQDLKDPLTYRVSTISNNYFFGSNFDLYNALNGRLALLNSYYSLQNNYLGNFMNDLQN
TQYEATRPIGQVADRTVLNNFFGVRYLFSQINQPNANKLPAGYQLDKLTPKVTDPNHN
SKNDTQARRYVTNNNFPLVYWQDQAFSLKQYRQFSPSQKERALASGVVVANSQLAG
LKSASTKQLSSQVLEVPYQLISSRGNLVSSQDLLKQDANESYQLVLNIQSADKQVQKKL
DQLYANSELHLEITDISYQPFSLKEQVKLEQRNNNNDLTRGIIGTNSALTKYKYFRYHILK
GSPDESFALNVTSNLGTEKISQPTQAATSFYKVVKNGTVNVGYFTKLPTSLTLTPSKLG

HYHFKLKVVALPLQGQYLKEVSRIQKHGLQNLTFTTTGFKGQITTKQTGILTSSIPYSRG
WSAKVDGHKAKVIRTNTAFLGLKLPKGHHQITFTYQIPGLALGKKLSLIGLILLGLSLAISL
IKWRIRHN
>Teichoic acid glycosylation protein
MIKELLLKYKAVIAYLFFGGLTTIINIVVFQVCSVNLGMYYQLANFLAWLLSVLFAFITNKL
WVFESKFTTVKSFLIEMWWFFFYRIISLGFDAGIMFIGISLLHANGLLTKLVDQVVIVLLN
YFFSKFLIFKKR
>Flavodoxin
MATAKVIFATITGNNEDIADIIAEELENKGLDVKVEEISQADPAELEAVDLCVICPYTYDE
GSLPDEGMDFYDELADVDLSGKVYGVAGSGDTFYGEYFCTAVDLFGEALAKAGASKG
ADNVKINLAPDEDEDIARLEEFADQLVKALGE
>Methionine aminopeptidase (EC 3.4.11.18)
MITIKSPREIEDMAKSGAILAGMHVALRDIIKPGISSWKIEEFCKNYIESHGAIPEQIGFEG
YKYATCVSVNDEICHGFPRKDLILKDGDLVKVDTVVNYKGAMSDSCWSYVVGKSTPEI
DKLMAVTKKALYLGIDQARVGNRIGDIGAAIQHYTEDEMGYGDVREFIGHGIGPTMHES
PNVPHYGEAGHGLRLREGMTITIEPMINMGTWEAEMDDPNGWTARTADGSLSCQFE
HTIAITKDGPKILTSQDPEMDAPYLLK
>UTP--glucose-1-phosphate uridylyltransferase (EC 2.7.7.9)
MKVRKAVIPAAGLGTRFLPATKALAKEMFPIIDKPTIQYIIEEAKKSGIEDILVVTGKGKRP
IEDHFDSVPELEQNLKEKGKKDLLQLVEQTTDINLYFIRQSHPRGLGDAVLMAKDFVGD
EPFVVMLGDDIMEDKEPLTRQLMDRYEQTHASTLAVMKVPHEEVNKYGVIDPAAQTM
PGLYDVKRFVEKPAVEKAPSDLAIIGRYLLTPEIFEMLETQKPGAGNEIQLTDAIDRLNKI
QRVFAHEFKGDRYDVGYKFGYLKTSIEFGLKHPEVKGELTSYIKDLAKKLK
>ABC transporter permease protein
MFKKTLWQDAFQAITHSLGRFVAIFLLMALSAFTLIGLKMTGPDMRQAVSDFYQRHQL
ADTSLTSNYGLDNSDEQILKADSDINSYELGYFQDAVLTKHQQAIRLFSKPTKLSTYALV
AGKLPHKDSELALDYLLAKKYHLGQEITVKQKGQLKYRHFRLVGFVRSSEYLDKNDFG
QTTVGSGSLAGYGIVAKSAFKANHYTIARVSYKQTAHLSPYSSQYADKVAFYQRRLKK
RLTNNGQEKYQVLKNQLTDSQTQLAQKQAQVQQLQMVNPTQAKLLQAQLEGAQARLI
NQKQALTKLGLPTYTLHERDYNPGYTLFRSNSERVDILANIFPLLLFAIAALVSLTTMMR
FVEEERINIGTLKALGYSNFDVSKKFLLYSLVASGLGVALGASLGFRLLPQLIFKAYTTNL
TLSQTKLLFSWTYLLVTLLVALSCTTVAALVALRKTLSEKPAQLLLPKPPKKGAKILLERF
KFLWKHMSFTQKVTARNLFRYKSRMFMTIFGVAGCTGLLVMGLGIRDSLRGITNIQFH
DLIKYDLVLSQKDYPTHSAKQKLAAKLDGSTVSKSAAVNYQLLTKKAGADQSVQSISMI
APHSKKLSSLVTLRNRTNQQQLTLRADSVIISEKLANLLNAKVGTTISLKDTSDQVRTFK
VTGITEMYTGHYLFVGKKAYQTAFKKAWTNNAYLVKLKDSGYQNVNHQARKFMETKA
LTTVVQNTNNRRTIDNVIAGLGNVIFILIGMATTLALVVIYNLTNINVSERIRELSTIKVLGF
YDRETTMYIYRETIILSLIGIGFGYLLGAYLHHLIITSLPPANAMFDPNLRLTNFLISALIPG
GITLVLAFIMHYKIKTVDMLAALKSVD
>ABC transporter ATP-binding protein YvcR
MAYIEVKNNTKTYQLANSTIHANKDISFSINKGELVVILGASGAGKSTLLNILGGMDTNTS
GQVIIDGQDISEYNRRQLTTYRRNDVGFVFQFYNLVPNLTAKENVELASEIVSDALDAS
QALKAVGLASRQDNFPAQLSGGEQQRVAIARALAKKPKLLLCDEPTGALDYHTGKQVL
QLLQDMSQQHGSTILIVTHNAALAPIANQVIHIHDGQVKKIKHNRQPAPIADIEW
>hypothetical protein
MAQKLSKEERKAMARKAMEKRAKQPKPAAGSGFAKLDADAKNLR

>cell surface hydrolase (putative)
MIKKFTYKLFLTLATSLLLLTSLCLTSPLTTRNHLHKQLVAQPTLYFHGWAGSANSTNHL
IAAAAALPQAKKVLTITVTESGQLAINGSWQDNVKYPLIQVLFANNQAEIPQQAHWIDKV
VTYLQTNYHIQSFNIVTHSMAGPATFYWAITSYRPTYPRFNKFVAIAGPFDGVIHIDDVP
NQNYFQADGLPAYQNAAFKNYYRQRQAFPKNTQVLNIYGNLADGSNSDSLVTNVSAR
SLGSLVKGQVASYQEITITGPSAQHSRLHNNNQVNHVVDNFLFK
>transposase
METQVIVKVKLNLANAEIASSFTNTMEQYRLACNYVSEYIFNYDFDMKQSRLNKSLYTN
LRSLFMLKSQMAQSVIRTVIARYKTVKTQMKRNPYKYQDINTGEWYRETRDLTWLKKP
ITFNRPQVDLQRNRDWSYLSSGQLSINTLNGRVKVNPVCHGFDQYFDGTWKFGLAKL
LKTGGKWYLHISATKTVANFDNQTVKHVVGIDRGLRFLATTYDEQGQTAFFDGKAVMR
KRAKYQKLRAKLQAKGTKSAKRRLKKLSGRENRWMTDVNHQLSKTLVQKFGANTLFF
LEDLTGVSFERTDLPKALRNQNKSWAFYQLAQFLTYKAHLNNSEVIEVSAQYTSQRCP
KCGTIKKSNRNYDLHEYHCTNCDYRSNDDRLGAMNIQLLGTQYISGQEQPKFELTTNA
>Histone acetyltransferase HPA2 and related acetyltransferases
MEIEIKATKDLASQIYQDALMIRKEVFVKEQGVDMALELDGELGPTHYVLYYKGKPAAT
ARTSFEQAGWHIQRVATLKSYRKLGLAKRLLSKIEEDARQSKVAYLTLGAQDSAQGFY
TKLGYQVVGAGFLDAGIAHHRMDKSLQIKEN
>LemA protein
MTLLIVIVVLVILALIYISMYNGLQKAKVNTEESWSQIDVQLKRRNDLIPNLVETTKGYAK
HEKETLAQVINLRNQLTELPADAHAQKMEVSNQLTDTLKSIFALAEAYPDLKANQEFTK
LMEELTNTENKIAYSRQLFNSSAANFNQRLLTFPSNMVAKIHHFTKVDYLEVPTEEKEA
PKVSF
>Heat shock protein HtpX (EC 3.4.24.-)
MLYQQIAKNKRMTVVVMVGFMLLVALIGSAVGYVFFNSSTAGLIMAIVIGLVYMLLMLS
QSTDVVMSMNNAHEITSPDQAPELWHVVEDMALVGQIPMPRVFIIDDPSPNAFATGPN
PEHAAVAATSGILTRLNREELEGVMAHEVSHIRNYDIRLQTTALALAAAISMLVNFASHS
FFWGGGRRDDDREGSSALDLIMMVVSILAIILGPLAASMAQMALSRNREYLADASAVE
LTRNPLGLISALQKISSSEPMQAADPSSASMYIANPFKDGSWTHLFDTHPPIEKRIERLR
NM
>UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-alanyl-D-alanine ligase (EC 6.3.2.10)
MKMQLAEIARALGAEVKAGYEDICITSVSFDSRTLEAGALFIPLVAGNDGHNFIQSAIEH
GAVATLWQADHDLGQVDLPQLVVADTLEALQTLSRYYLDKINPKVVAITGSNGKTTTK
DMVAAILGSQFNVQKTHANFNNQIGVPMTILGMEPNTEVLVVEMGMDHFGELDHLSRL
AAPDVAVITMIGEAHIEFFGSRDKIADAKMEITHGLKEDGTLVYNGDEPLLRTRAEDLKQ
LTVTFGRNDNNGLQATNIVSDDTQTGFRVVSWPEENFTIPMIGAYNVNNALAALAVGN
LFHVPVEAMRKALAAFDLTANRTEWVKGAKGERILSDVYNSNPTAVKQVLQAFTKVET
DGKRIAVLGDMLELGDQAEIMHTGLAAEIQPEEIAEVYLCGPIMAKLAQVLDEKYPADK
VHHYDTDAKAQLVADLKASIQPEDIVVIKGSHGIHLEEVLAALQK
>DEAD-box ATP-dependent RNA helicase CshA (EC 3.6.4.13)
MKFSELGLEKDLLTAIERSGFEEATPIQEATIPLVLAGVDVIGQAQTGTGKTAAFGLPVL
QHINPKKNIIQAVIISPTRELAIQTQEELYRLGKDKKAKVLAVYGGADIRRQIHQLKHNTP
QILVGTPGRLLDHINRKTVDLSLVNSLILDEADEMLDMGFLDDIESIIESIPPNRQTLLFSA
TMPKAILNIGEKFMMEPQVVKIKAKELTTDLVDQYFVRAKDFEKFDIMTRIMDVQAPELT
IVFGRTKRRVDELAKGLEARGYNAAGIHGDLTQQRRMNILRKFKAGQLDILVATDVAAR

GLDISGVTHVYNYDIPQDPESYVHRIGRTGRAGHHGTSVTFVTPNEMDYLRVIEKLTKK
RMEPMKPPTEAEAFEGQISSALVDIKELVAKTTTAKYEAQAKELLEQYDAVDLVAALLN
NMTKDDASQVPVKITPERPLPRRKKGNHTGNFRRSNKNWRPKNKRYDERRRSERDR
DRDQFKHRTKKRKFTIRQK
>Holo-[acyl-carrier protein] synthase (EC 2.7.8.7)
MIYGQGIDIASLKRIKAAYQKNPKFAAKVLTEGELAAFETLSSKRQIEFLAGRFSTKEAY
SKAVGSGIGSQVGFSDLEILNNKQGRPVIVKHPLQNKLRAHLSITHDNEVVVTSVILEKL
GGFERLKSYFK
>Alanine racemase (EC 5.1.1.1)
MVVGKHRNTKVIVDLGAIKENVAQEKRQLEANQKLFAVVKANAYGHGLVEVAQAASQ
AGADGFCVAVIDEGVALRQAGFEQEPILILGVNPASEATYMASYNLAVAVSSLDFLQAA
QPELLKAKLKLRVHLALDTGMGRIGFKDVATLKAAVAYLNANPQAFVFEGIFTHFAQAD
SKDTSYYHQQLTKFKELTGSLEQLPPYVHVANSAAALWHHDCGGNVIRYGISMYGLNP
SGRELELPRPLKPALSLVSELVHVKQVAKGESIGYGSTYTASEDEWIGTVPLGYADGW
LRRMQGFAVLIAGQYCPIVGRVCMDQFMVKLPANLVLGTKVTLIGKDGSEEITAQDVA
DYAHTIHYEITCCLGERLPRIYRE
>hypothetical protein
MMEQVQSRRLFDDAKLSEYCEFYGISAQSLDDKALKYIESRYEKLDLLITGYGEMAKLN
QEICTEFLSCDCDELDELVR
>Programmed cell death toxin YdcE
MAAEVKRGDIFFADLSPVVGSEQGGMRPVLIIQNNIGNHFSPTVIVSAITSRISKPKMPT
HVGLKAQKNGLEKDSVILLEQVRTIDKQRLREKIAHTDQETMKKVNEALRISFGLIAIHK
>hypothetical protein
MQLVTLYLTWSNWWSLLPIVANIATTIGEYTYNSQKIRLVGLAINSPLKISYDFLVGSWA
GIFDELVTLASVIISIWRYGWKNLDRVEEIG
>Branched-chain amino acid transport system carrier protein
MLNKLKEKKLTKKDYFVVSSLLFGLFFGAGNLIFPLHLGQLAGHNWGPAVLGFLVTGVL
LPLLSVLAISITRSEGVYDIGRPLGPAFALAFMVLIHATIGPFFGTPRTATVSFAVGLAPLL
PAGSTKVALFVFSACFFLVAYLFSRKENKIVSNVGKLLNPVFLVLLALVFFVAFLHPLGS
LSTATPTKEYLGQTSFIHGFLEGYNTMDALAGLAFGVTVVTAIRSMGVKRETKVAAVTA
KAGVFGMGAVALIYLLLTLLGVQSLGKFKLSADGGVAFSQIVNQYAGVAGQAILAALITV
TCLTTAIGLVAAFAQDFHKHFPKVSYHTWLVLSCASSFFFANFGLETILAWSLPFLMFLY
PLAMVLILLSVFSPLFKRDSYVYFFVVLFTVVPAFFDMVATFPAVISQSNFGKAVYAFQH
HYLPLADYSLSWLVPALVGLALGLAVHFYKEVFVKTQASEMD
>FIG00742398: hypothetical protein
MLLKTLVKPKQVLTTVREDVTLAEALETLETSGYRCVPVLDTTGKIFRGNIYKMHIYRHK
SRGGDMNLPVTHLLKNATKFITVNSAFFNVFFSIRDLPFIAVLDENNYFYGILTHNRLLE
MLSQSWNFNVGSYVLTVVSAGERGDLVDMAKVITKYTSIASCITLDVQQGELVHRTLFT
LPADVSHEKLERIVASLERKNFRVSEIEDLQA
>L-lactate dehydrogenase (EC 1.1.1.27)
MTVSKTQNHQKVILVGDGAVGSSYAYAMALQGIAEEFGIVDVVKDRTEGDALDLADAT
ALTYPKKIYSAEYSDCKDADLVVITAGAPQKPGETRLDLVNKNLKILSSIVKPIVESGFDG
IFLVAANPVDILTYATWKLSGFPKNKVIGSGTSLDTSRLRVALGELTGIRDPRSVHAYIM
GEHGDSEFAAYSAATIGGLPILDWAEKNGVTKEQLAKLEDDVRNKAYTIINKKGATFYG
VATALARISRAILRDEDAVLPVSAYMEGQYGLNDIYIGTPAVVNGSGLDSILEVPLNEEE
HAKMTASAKTLKEVLTNGLKTLEEGK

>Peptidyl-tRNA hydrolase (EC 3.1.1.29)
MKMIVGLGNIGRKYDETRHNVGFMVLDQFAKEHQVSFDKEDFNALIASVFINGEKVLLV
KPTTYMNESGRAVGPLATYYNIASEDIAVIHDDMDLEVGKLRLRQKGSAGGHNGIKSLI
HHLRTQEFKRVRVGIGHPQKMAVVDWVLSKFTAEQRQAIDVTANEVVAALDYWLASD
DFMKTMNKYNHK
>Transcription-repair coupling factor
MDLSEILLEETKVTDWLKQLNENKRRNLVTGLVGSAKTLFLQTLFQELKRPLLIVTSDLY
RMQELEDDLKNQVEPGTLYTFPVEEVLAAEMATSSPEFKSQRVQALNALASGQAKLIL
TSVSGLRRKLPAVDYWRSLQLKLTVGSQLDLDALGQKLTAMGYRRQKLVERPGDFAI
RGEIVDVYPVAGLEAVRIDFFDTEIDSLRTFDVASQRSIANLEEVSLAPASDFLVAKDDL
KAAGVKLVADFERLGPELSDAKQRQALQATVAQKATAWQAGALTETDLPWGAYLYPE
AVTLLDYLSSAGLVIVDDYSRIVDKNQKLINDDQNWQVTKLASGELLPNLQLGFDSAEL
LKNNQRTGSIFLAPFAKGMGGLRFSNKLAIKSRAVQQFFGQMPLLKTESERWRKQKQ
TVVILVTSNERLHKVEQTLTDFDIQVLLSEPSQLQKGNLQLVLGGFRSGFELTEANLIVL
TEHELFAKVNRHQARRQTMSNEQRLKSYTDLKVGDYVVHVNHGIGKYLGMETIEVSG
RHQDYMTLLYKDDAKLFIPVSQIDRIQKYVSSEAKQPQIASLTSDKWRKTKSRVARKIE
DIADELVELYAKRESEQGYAFSKDDSYQKEFEDAFPYTETPDQLRSSQEIKRDMEMTK
PMDRLLIGDVGFGKTEVALRAAFKAIQDGKQVAFLAPTTVLVQQHYETILNRFDGFPVS
VGILSRFSSPTEVKQTLQDLRLGLIDIIVGTHRLLSKDVKFADLGLLIIDEEQRFGVKHKE
RLKELKASVDVLTLTATPIPRTLNLSMLGVRDLSVIETAPLNRYPIQTYVLEQNFGVIIDGI
KRELARGGQVFYLHNRVEDIESVANEIRALVADARVAYIHGQMTESQLERVLMDFIAGE
YDVLVTTTIIETGVDIPNANTLFVENADKMGLAQLYQLRGRVGRSNRIAYAYFMYQPDK
VLTETGEKRLEAIKDFTELGSGFKIAMRDLAIRGAGNILGKQQHGFIDSVGYDLYTQML
NTAIAKKQGKKVQAKTDAEINLEIEAYLPTEYIEDARQKIEIYKRIGQFENLEQYREVQTD
LIDRFGDYPQEVANLLEIGLLKMYADLALIEKIKQTGNKVTVLLAEKFVKQVPAQALLKVL
AQTKLVTKLVPKKEQVQIEWVLQPTMDTAQELSQLKELAQLLGKELQAKKG
>Low temperature requirement B protein
MQTKNATKTLLKGALILSLASLIAKILSAFYRVPFENIVGNTGFYVYQQVYPLYGIGMTFA
LSGFPVFISKLIAEQEGQLAKVNLARRLFVILSFFAGGIFIFLQLTSGLIANSMGDGQLELL
IKSVSFMFLFMPLLAVGRGYYQGTFNMVPTAISQVSEQLVRVAVILLAAYLFTKRGWTY
YEMGAWAMAGATIGALCASLSFGKFYLGSFLSQPFATSKLSMQTLIRRLLSEGLIICLFA
SLMVLLQLVDSFTVKQALEVKGFSSELAKEIKGTYDRSQPLVQLGIVVAVSFSSTLLPSL
TEALQKRHLNEFRKTAKALLRISVGLGAAATVGLIALMPEINTVLFGDASLSKTISIYVVS
VFMMTLISNYNSLLQSLNQYQMTVIALLAAIVVKALINSYLIERLGIAGASIATVLALIVAQA
LIAGSLPALLKGLLSGDNYLLKLGLVTGIMGGVAAVVAKVGTFLLPGGRLALLLVALISV
GIGVLVFIGLALWLRLFSIREWLAIPGGKKLLRLTNKFTNRKGLS
>S4-domain-containing heat shock protein
MRLDKFLKVSRVIKRRTVAKEIADQGRILINDKVAKSSSTVSVGDVLTIKFGNKTSKIKVL
ELKETTKKNEAAGMYELVSEEYKEDYRL
>Cell division protein DivIC (FtsB), stabilizes FtsL against RasP cleavage
MEPKRKQIRVLNNEYFQAKQKEARLQAKRQKRQQQRSLLHLICMATMGLLFTYGVIRV
AQAYLQIKALKQETVVAKRELKRTQAKNDRLKREVEQLKDTDYVQKIIRSQYFYSKDGE
TIYILPDDAKD
>RNA binding protein, contains ribosomal protein S1 domain

MSVEVGTKVTGKVSGITNFGAFVDLGDKQTGLVHISEVSDGFVKDIHDVLAVGDEVQV
KILSVGDDGKIALSIRKAMAKKPSVKEHKPYKKHEGDNNGRFHHEHNRNRDDKRHSFN
NSKRNNDGKKKDFDSLLAGFLKESEDRLTSLKRNTEGKRGGRGGRRS
>tRNA(Ile)-lysidine synthetase (EC 6.3.4.19)
MLRKFQAQLAKLGLTGKKVLVAVSTGCDSMVLLDLLLKSQRRFDLQIYVAYVDHQLRP
ASKVETAFIKDYCHKSQLPLRMYTWPLAAHPAKGVEAAARKMRYDFFKRVMATEQLT
TLLTAHHADDQAETFLMKLIRGGQLAQLTAIAPQRDFGVGKELVRPLLGFTKQELADYA
QANNLTYFEDETNASDDYQRNRIRHHLLPLLKAENPQVLEHIAAYRTQLSDLLALGDEL
AAANIAKMKAADGSYNLRVFQTYSPKLQVLCLQHLLMEQGVELRTQLPEELRAGLLDD
RKPNLKIKLNKEYYLLKSYERFGFESVCPKSQARAVSYNLEPNEWIDLGAEGKVGLVSF
EDYQARSGDDCLELSELPAGGLVLRHRLSGDKLVTTHGRQKVKKIMIDNKVPLAQRKT
SWLVADQAQNVYWVLDLKKSDLSRGAVNAKIQYIVVHRS
>Hypoxanthine-guanine phosphoribosyltransferase (EC 2.4.2.8)
MNNDIKEVLYSQEDIKAVAEKLGQEITKDYEGKNPIVICVLKGAILFMADIAREIKTYCEL
DFMDVSSYGDATVSSGEVKIIKDLDTSVKDRHVLIIEDIIDTGRTLERLVDLLEHRNAASV
KICTLLDKPERRVKGVSPDYVGFEVPNEFVVGYGLDYMGKYRNLPYVGILKPAVYENN
D
>Cell division protein FtsH (EC 3.4.24.-)
MNNKKNGLLRNSLFYIVIFLGFMGMLYYFFGSKESTNSQQIQSSQFVTELRKDNVKDFT
IQPSGSTYKITGTYRKAKKAKAPAGLAGLSAPSAKVTTFTTNVLTNDTVIKTIENYAAKN
NVKNNTKEEESSSVWIQLLISLLPLVFIVIFFYIMMGQAGQGGGNGRVMNFGKSKAKPV
DKKNNKVRFSDVAGAEEEKQELVEVVEFLRDPRKFLALGARIPSGVLLEGPPGTGKTL
LAKAVAGEAGVPFFSISGSDFVEMFVGVGASRVRDLFENAKKNAPSIIFIDEIDAVGRRR
GNGMGGGHDEREQTLNQMLVEMDGFEGDEGVIVMAATNRSDVLDPALLRPGRFDRK
ILVGRPDVKGREAILKVHARNKPLAGDVDLKMIAKQTPGFVGADLENLLNEAALLAARR
DKKKIDASDIDEAEDRVIAGPAKRDRVISKQERETVAYHEAGHTIVGLVLNDARVVHKV
TIVPRGRAGGYAIMLPKEDQMLLSKKNMEDQIAGLMGGRAAEEIIFRQKSSGASNDFQ
QATQLARAMVTQYGMSSKLGPVQYEGSANIMPGEMGGAAPYSGQTSNLIDQEVKRIS
EEGMQKAKEIIEAHRKEHKIIAQALLEYETLDEKQILSLYKTGKMPVKASDEFPSEKAAT
FEESKEALERKEALKKQALADKDTKDTAELYPSEKAAEDDSLKTTVYYPKSKTSESDSE
DKQK
>Chaperonin (heat shock protein 33)
MTDYMVKSLVYKGALRAYAVDATEVVREAQKRHDTWSAASAALGRSLVGTLLLAAAS
MKGDEKMTVKINGNGPVGGIVIDANSKGEVKGYLQHPHVHLPLNPNHKIDVKGAVGTE
GFLAVTKDLGLKEPFTGQVPLISGELGEDFTYYLAKSEQIPSAVGLSVFVNDDNSIETAG
GFMIQVLPGASDEILTKLEQRLAKLPLVSEMMRQGKTPEAILETIFPDTEVEFLERQPVA
FKCDCSKERFAEALASLPQADIEDMIAKDHGAQAVCHFCGAKYDYSEADLKQIMKEAR
A
>tRNA dihydrouridine synthase B (EC 1.-.-.-)
MKMEWKIGDIVIPNQVVVAPMAGVTNVAFRLICKEFGAGLVVCEMISDKGIQFRNKKTL
GMLFVDPTEHPVSVQIFGGSKETLVEAAKFVAENTSTDIIDINMGCPVPKVTKTDAGAH
WLLDPDKVYEMVKAVTSSIDKPVTVKMRTGWDEDHILAVENALAAQEGGAAALAMHG
RTRKQLYSGHADWDLLHKVANSLTEIPFIGNGDVTTPEEAKKMLDEVGADAVMIGRAA
LGNPWIVKQTVNYLENGEIIAEPSPAEKITVAKEHLHRLVKVKGEKMGPREFRSQAAYY
LKGIPRSARTKAALNSAETEAEMVEIFDNFLADTLARQAR
>Lysyl-tRNA synthetase (class II) (EC 6.1.1.6)

MYENFESEVKHSMANQQSMNDQLRVRREKMQELRELGIDPFGHRFERDHTNAQLHE
KFGNESKDELNDQHIEVTIAGRMVAKRGKGKVGFADLQDRTGRMQLYIRKDEVGEEM
YHIFKRSDLGDHLGITGEIIKTDMGELTVKVSKLTFLSKALRPLPDKFHGLQNVEQIYRQ
RYLDLITNRESYQRFVNRSKIVSAVRSYLDGQGFLEVETPVLHNQAGGAAARPFITHHN
ALDIQLYLRIALELHLKRLIVGGMERVYEIGRVFRNEGIDTKHNPEFTMLETYAAYFDFK
DVMDETEGIFKAAAKVVTTDGKINYQGTEVDFNKEFKRVHMVDLIKEKTGVDFWQEMT
LEEARKLADEHKVHYESYWQVGHIINEFFEQFCEDTIVDPTFVYGHPVEISPLAKKNAD
DPRFTDRFELFILGNEYANAFTELNDPIDQRQRFEAQVAERQAGNDEAEGIDEDYIEAM
EYGMPPTGGLGIGIDRLVMLLTDAPSIRDVLLFPTLRPE
>Small Subunit Ribosomal RNA; ssuRNA; SSU rRNA

>tRNA-Ala-TGC

>Large Subunit Ribosomal RNA; lsuRNA; LSU rRNA

>5S RNA

>tRNA-Asn-GTT

>tRNA-Glu-TTC

>tRNA-Thr-CGT

>transposase
MVLTSGGIALTLGSLEPTFYPFASLCCVFTEQKCYNVSKESEVNLMEIQITIKVKLNLAN
VEIASSFTNTMEQYRLACNYVSEYIFNHDFDMKQSFVY
>transposase
MYTNLRSLFMLKSQMAQSVIRTVIARYKTVKTQMKRNPYKYQDIKTNYL
>transposase
MQRNRDWFYLSSGQLSINTLNGRVKVNPVCHGFDQYFDYKTGANL
>Mobile element protein
MVKKKGNKFGYTRQYHISLSYQQIDLYEKAIKTQNELYQFALKYLYKTYGRKHIGRPLP
FGQGINYLNNKIKAMFINEKYELKRWNVKKLGLSSHAADEFLKIIFTNFSQYRKRLEQAG
QMSAVEKYNLRMNITKDKDGKHKNPKHRSWYRKGSMNFLRNNNSFRTITSQKLPKGN
TKLVSPHHLNVADFGEIMVFENLKHLNFDEIALTKIKRLPDNTFRLQITFTREKKRSSQN
KVVGADWNMFNNEVFRTSENKTIAIPKEIVKKANELEAAKDKLKSLRDSEYNKRGKTAL
WQRYQRKQAKLSTKRANLLTETYRKLVHELVDEFDTIIIEKIDAFEMRKRSWSLNKAQN
TGKNKRLALIKPYELAKIVESLVNKQNKTLIKVDPYKTSQVEYGTEYEEKHELRATNKDG
KRIYVSAYTGKEVDRDYNAALNIKEWGLHPEKHVKLRDYPKLNASNLVEII
>Oxidoreductase, aldo/keto reductase family
MIINQVQLSLKHTPLIDQGIFVNRFEDEASDRTGGLLEYAQQNKITLQAWSPYQYGMFK
GVFIDNPKFPELNQVLQTVAAKYGVTKNAIATAWITRLPQDIQVILGSMNPVHLKESLAG
ADIRLTRQEWYDLYLAAGHQLP
>Di/tripeptide permease DtpT
MNDNTKEATFLGHPRGLSTLFFTEMWERFSYYGMRAILLYYMYYEVSKGGLGIPQATA
ASIMSIYGSLVYLSSTIGGFISDRILGSRRTVFIGGVLIMFGHIALATPFGKTALFISIALIVI

GTGLLKPNVSEMVGGLYSEEDPKRDAGFSIFVFGINLGAFISPIIVGYLGQNVNFHLGFS
LAAIGMFFGLIQYTLDGKKYLSKDSLYPTDPLEPHEFKSLAKKVTIGVVAIALILFIMSLFG
MLTILNVINLITILAVLIPIYYFVAMLSSRKITKKERSRVLAYIPLFIASILFWSIEEQGSVVLA
LFAEQQTRLHILSWSFPASIFQSMNPLFIMLYVPFFAFLWTKLGSKQPSSAAKFAYGLF
FAGISFLWMMLPGMLFGTGAKVSPFWLIISWALVIVGEMLISPIGLSVTTKLAPKAFQSQ
MMSMWFISDAVAQAFNAQLVRYYSTKTEVAYFGSVGVATIVFGLILLAFVPKIKKLMEG
VS
>tRNA-Tyr-GTA

>tRNA-Gln-TTG

>Pyrroline-5-carboxylate reductase (EC 1.5.1.2)
MELGFIGLGNMAKAIIAGWLKADAGLATKIKVHSAHKSNYETYAKKWGLTACASNKEVV
TNSDFIFLAVKPFLLAEVLKEIKPALTENKVLVSMAAGVSLADIAAVVGPSVKLIRIMPNV
NAEISMGMTALVANEKVSQTEYQKVEALMASEGEALAVAEKDFSIFSALAGCAPAYVY
YFIDALSRAGVKHGLKKSEATKIAAQTVLGSAQKVLSASLSPMDMVDQVCSPGGTTVA
GLLAMEEAGFMTAVVKGIDAAVAKDQGK
>N-acetylglucosamine-6-phosphate deacetylase (EC 3.5.1.25)
MATVLKHAKIYTGTEVIADGYIRFDKEIQALGPVSDYQVASGDQVVDLAGKVIVPGFIDV
HCHGGYSFDTMDGDAAGLDEMVNDMVKDGVTTIFPTTMTQSPQNIKQALEAVAKVAK
QNPVIQGIHLEGPYISPVFKGAQPEEYIKAPDVKQMAEWNQVADHLIRLVTYAPEHETA
REFEEYCLANKIVPSVGHSNATREQMHASKASHVTHLYNAQREFKHREPGVTGHAML
ENDMYCELIADGFHVCPDMIKLAYELKGADKLELITDSMRAKGMPEGISELGGQKVIVK
DKQARLESGNLAGSVLTYDDAFVNIMKFTGCSVQEAIKMSSLNQAKEFGLSQKGSLEV
GKDADLNVFDFDWNLKATYSYGRLFQK
>Predicted transcriptional regulator of N-Acetylglucosamine utilization, GntR family
MGSPRYIKIHNEIKQMIEAGKWRVGDRIPSERVLALEFDVSRMTLRQAVQTLVDEGILE
RHVGSGTYVSRQRVQEKMSSGVTSFTELTEAQGKVPSSKTVTYKVATPSLSELEKLHL
KEGAKVLRMERIRYADGRPICFEVATVPYDLVANYGREEITASLYRTLEANDLQIGKAQ
QTISAMLASERIAEYLELKRGAAILRLRQVTFLTDGRPFEYVRTQYAANRFEFYLEK
>N-acetylmannosaminyltransferase (EC 2.4.1.187)
MAKFQTVNVLGVDFINTTNHDFLTQLKTDSQANKNRFVVTANPEIVLAAKQNPQLAAAI
KSSDYTTADGIGIIKGAKILKTPLPERITGYDTMCALLDWANANQKRVYFLGAKPSVLAD
LKKKLAQNYPDLVIAGSHDGYFQDEGPIVNDIKQAKPDFVFVALGFPKQELFIANHRQV
VSAIWMGVGGSFDVLAGHVKRAPAFWVNHHIEWLYRLLKEPSRFVRMLALPKYLLLVY
RYALIKK
>NAD synthetase (EC 6.3.1.5)
MRPLQAEIIKALRVQAEIDPQAEIRRSVDFMKAYLKKHSFLKALVLGISGGQDSTLAGKL
SQMAISELRQETGDSDYKFIAVRLPYGQQADEQDALDAIEFMQADLTVRVNIKEAADQ
MVAALEANGLTISDFNKGNIKARQRMIAQYGIAGAYQGAVVGTDHAAEAITGFYTKFGD
GGADITPLWRLDKRQGRAMLELLGAPTHLYQKVPTADLEEDRPALPDEVALGVTYADI
DDYLEGKDVAVDAAEKIEAWYQKTEHKRHLPITVYDEFWK
>ATP-dependent DNA helicase RecQ
MTPNETLKHYFGYDHFRPGQAEIINQVLAKKNTLAIMPTGGGKSLCYQIPALILPGITLVI
SPLIALMKDQVDALNDNGIPAAFINSSLDYASIQNNLSLALTDQLKLLYIAPERLENPHFL
NELASLNISLVAIDEAHCISQWGHDFRPSYLNLQASLATLGQDFPTLALTATATDQVAQ

DICQQMQIDPQQIIKTGFSRPNLIFKVAKGQDANTFCLNYLRKNPHQAGIIYAATRKKVE
QLGSFLSKHGQKVAIYHAGLSEETRHQAQEDFIFDRVNVIIATNAFGMGINKSNVRFVIH
AQAPGSLEAYYQEAGRAGRDGLLSEAILLFHEQDMMTQRFFIDNSDAQADYQTIQYHK
LQAVSHYANTEECLQRFIVRYFGEETEPCQKCGNCTDERASEDITLASQKILSCVVRM
RGNFGKKIVAQVLFGSTASKITQFNLQTLSTFGIMHDLRLKEIEALIDYLLASGYLKTAAG
KYPTVAITALGVQVLKGKQGVKRKVTKAKVIKSNQAVDRALFERLRNLRLDLALKQEVP
PYVIFPDKTLEELCLLKPKQLDDLLAISGIGQAKLAKYGQLFLAEIRAYQNET
>Transcription accessory protein (S1 RNA-binding domain)
METAVLNETIKALNYRPRQVRASLELLTAGKTVPFIARYRKEMTGSLDEVALREIQATY
HQLEKLSQRKQEVLTKIAAQEKLTPKLAAKIKQATSLQAIEDLYLPYKQKRRTKAQIAKE
QGLAPLAAAITRFESDQALAKLSHTLLTEAVADEAAAFLGAHEILAEAFGEDARLRAWV
RQKLSRQGQVVASLKDESQDEQGVYQDYYDYEQPLKKLASHQVLALDRGERQGVLK
VKLVLDEAYVLNYFHSQFVAGKPSGKAASEIQAAYTDAYKRFIKPAIEREVRKQLTEQA
ASEAIEVFGANLYHLLMQAPLKGKVILGLDPGFRTGCKLAVVDATGKFLAKAVIYPHEK
AKGAQANPKLRAEAKQIVTDLITRYQVEMVAIGNGTASRESEQFIAEVLGKLTQTVYYVI
VNEAGASVYSASANARAEFPDLNVEERSAISIARRLQDPLAELIKIDPKSIGVGQYQHDL
PEKQLDEKLAAVIETGVNQVGVNLNTASPDLLTHISGLNKTIAQNIVAYRNEVGAFKSRA
ELKKVARLGPRAFEQAAGFLRVLAGTNPFDNTDIHPESYPLAKQLLAKVGCSTADLGS
ELLKAKLAQLDLKEVASQFNAGYETLKDVALSLEKPGRDLRDTMPAPLLRQDVLKIEDL
KSGMKLEGTVRNVVDFGAFVDIGVKQDGLVHISRMSKRYVSKAAEVVAVGDIVTVWIE
SVDLKRQRIQLSMLPVASDD
>Putative metallopeptidase (Zinc) SprT family
MTDEELQAWVQALSQTAFSMDFSHRAYFNARLRTTGGRYRLADHNIEINPKMLGYGK
EVVAGIIKHELCHYHLHLAGLGYRHKDRDFKRLLAKVGGSRYAPAAPKQRKYHYQCTN
CQQDYFRQRLVNTDKYVCSKCHGRLALVGLAQKN
>tRNA-Leu-AAG

>3-oxoacyl-[acyl-carrier-protein] synthase, KASIII (EC 2.3.1.180)
MSKKIKLLATAGYLPATVVTNDDLSQIMETSDEWIVAHTGIKTRRYALGENTSDLATKV
GKELLAKANLAATDIDLIIVATITPDALTPSTAAIVQANLKAQNAFAYDLSAACAGFIFALA
TAEKFLQAGTYRRALVIAAETNSKMMDFTDRTSAVFFGDGAGGALLEASDAGSSVYLA
EKLVTQGNKEVIHSGRIAPLSAIAATNYPKTDAFYQDGRAVYQFVTTTVIAHIKDFLAQN
NLKPSDLAMVITHQANLRLIEAIATALDLPLSKFAQDVVEAGNTSSAGIPLALAKAHAAG
LHGKVLLCGFGAGLAYGSILLDL
>hypothetical protein
MIIWTIQPYSVYQELVAKGRFYCDPNKSCNLKEPDFQAAYAWMIAQLKAKVGIPPNDQI
QTPIWAWYRARAFSHTRPDFREHRDYADQVCLELAIDDDQVLLSDFDSWHFVLGDSY
FAPASNEQEWEALNAWFDGLTKKDQQAVKEASWQQIFDITPSHDPWTLNGAYVQAC
FWQLELSQVRKVWRMKKGKKVRELDMGSV
>Abortive phage resistance protein
MNLDDYKLLKTKEIVAVLDGDEDFGDFNSRQKMSMPYLSGSAICDVGKRFGLALEYEG
MKSRWSYMYDLIDLGIGSSNISDVLSYLFDESNFSKYILSSLTNEEFKEYYKRSVESAIG
RINSLLHYSGKKLVIKNNKAVLVSRNKEEIELSVKNITTINRDYIKRVAENAIDNIERSNYD
SALTQARTILEETFCFVLEQRNIQANNKGNIKNLYKEVKNEYRMHIDEKADRRVKMLLS
GLENIVSAVAEMRNKSSDAHGVGENRLTIDKHHARLAVNSAMTMADFILSVEKKANGK
LDY

>Mobile element protein
MVKKKGNKFGYTRQYHISLSYQQIDLYEKAIKTQNELYQFALKYLYKTYGRKHIGRPLP
FGQGVNYLNNKIKAMFIKEKYDLKRWNVKKLGLSSHAADEFLKIIFTNFSQYRKRLEQA
GQMSAVEKYNLRMNITKDKDGKHKNPKHRSWYRKGSMNFLRNNNSFRTITSQKLPKG
NTKLVSPHHLNVADFGEIMVFENLKHLNFDEIALTKIKRLPDNTFRLQITFTREKKRSSQ
NKVVGADWNMFNNEVFRTSENKTIAIPKDIVKKANELEAAKDKLKSLRDREYNKRGKT
TLWQRYQRKQAKLSAKRANILTETYRKLVHELVDEFDTIIIEKIDAFEMRKRSWSLNKAQ
NTGKNKRLALIKPYELSKIVESLVNKQNKTLIKVDPYKTSQVEYGTEYEEKHELRETNKD
GKRIYVSIYTGKEVDRDYNAALNIREWGLHPEKHVKLRDYPKLKAGNLVEII
>hypothetical protein
MQLLTLYATWANWWSTLPVVANIAATIGGYTHNPRKIRMVGLTVNSPLWIIYDMAVGS
WAGIFDLKRMEVGLSTRQRKVIT
>hypothetical protein
MDGFLAIVFYDKKFHTIKVYSKGGGEYHSHLAAYVLKYYEDTANRDTTLLYEIVEPVMD
EHIVWYDEAWTTLLAAVKNDLRGSLMLVSNIMGIIHNEDELKTFLSLYTNIIEEGYVFYDA
NDKMVKVKTNYFRIARELRKGMALGKPLNSWKNTDDKIKKIAYEWYKKAQSMGMRSF
SPKVAYDVYWEFKSKNIDEFLPTPLTLEGLTYKK
>hypothetical protein
MVFVFLGLAGIAGTIFSIAFIVWIVALIINKKDLIKRAWKFTWISFAVTLVCSLGYGFTLWG
VFDSSNDTEQVSQSETKEDKEDDKYAEQQENEREQEEAVYTKNAPKKVTFEKDDYD
GAGVLTKYNMRNGYKYNYYKDSDGDYFYYSINDKDTFPPDRRYLHDELATIATNEKYA
KTHKKDMVWDDDPNGLSKLKESGVYTMNPPQRDKVNGGILSKNFNRQKGYQVNYYY
GTFKTPNNVKDPISSDSKQYADGYYYYDSDGYSTSVNGDYIKTAEDKLKNNPDDLLQQ
>Ribosyl nicotinamide transporter, PnuC-like
MISWWHDQLFSNWKKFEIGYISLLLALQVGVYLLAPDSWIGMFSGIFGTLCLIYGMKGR
KVSFIFGFFQCLAMTYVAWISHAYGSFAMDIFYVISQPIGWFLWEKSSLTKSLPKKQRQ
LLFGFALLAWGIGWWLLSLLHGQLPYFDSINFVVSFIAQILYICKFQENWSLWIVVNLANI
IYWGILAYQLLVGATNIGTFGAALSQVALQAALLFNSLYATKVWASGEADYEGGTSKQN
>Quaternary ammonium compound-resistance protein SugE
MAWISLLIAGLFEVVWATTMKLSNGFSHPLWTLATIIGMTISFGGLVLATKHLPSGLPHL
DRDWGRWLRYRRRPLLPRPA
>hypothetical protein
MNNAKLKQLLSEIRACKQQLERMEADEFFKTQTAPLKKELAELIATYQQQTKRDPLELL
ARQDEKRRRNFLANWPQLKELRFSVGGYPGDYATGLAVILPKKVVLLQQPHSLIEGTP
CLVNQLERKQFLTAVQACHLEDWQREYFNPQILDGTQWSLTCYYQGLKQTFTAEGSN
DYPASYERVKNLLLAKDEAAKEVALNLMDQEAVTDFLTTF
>carboxymethylenebutenolidase-related protein
MTPTPKRKRYLWGGLLTLAILIGLAGIILHVKTYQPTASANQASQAATVSKNVTTFKAKN
SKLTVVFYPGGLVEPASYSNWASQLAKAGYTVKIVHFPLNLAVLAPNQAKKVVGPHEQ
YVIGGHSLGGAMAARYAAQADKKNLKGVFLLAAYADQKGRLDHSKLPVLSVTASRDG
VLNWSHYEANKKYLPRNTTFKTISGGNHGGFGSYGHQQGDKAAHISNATQQQQVAHL
LIKWLKRIN
>2-oxoglutarate-malate translocator
MKRLEFVNYRGFVLPLIVGLILWGLTSIRPAGIPVAGWHMLALFVATIIACITRPLPIGGVS
IIAFALTVLTNTVPMDKAILGFGNAAIWMIAMAFFIARGFIKTGLGRRIALNFVKAFGHKTL
GLAYALIGVDLILAPATPSNTARAGGVLYPIIKSLSQSFGSDPKDGTASKIGSFLIFAEYH

GDMITAAMFMTAMAGNPLAQTLAKSNGVTISWMDWFLGALVPGLVSLIVVPWLIYKLY
PPEIKATPNAKAWAKDQLAQMGPISLPEKIMAVIFVCALILWITGSFLGIDATLAAFLALAL
LLITGVLSWQDVLAESGAWNTLTWFSVLVMMAGSLSSLGFIPWLSKVIGNSLSGLNWF
LALVLLIGALFYSHYLFASSTAHITAMYGAFLAVALSLKVPPMLAALLLAYFGNLMSSTTH
YASGPAPVLFGSGYVPQNKWWSLNFILGIFYLVVWIGLGSVWLKVLGMW
>hypothetical protein
MTIEEYAARQSAYVREEKENQTIKGLVKLNLSQEQIIEFLVQNFKLDKQAAVKAYERAM
ATV
>FIG00754245: hypothetical protein
MDRETKISQSVSSFRLPRYEQLPEVELYLDQTTAYISERLALLGDVKLTNSMISNYVKH
KLIRRPVKKKYAADQIAELFFIAVAKNVMQLNDLKEAIELQRQSYSTAVAYNYFAAELENI
LPYVFGLKEQLADVGHDNSEYQRMLRNIIMAVCYKAYIDKYYADLQAELAQK
>DegV family protein
MFAPVLTGIKHQSVALSLLGQHATFLDEAGFSQAAFAEFIAGQTGPIKTSCPNVDSYLA
AMQGADEVYVFTISSALSGSYNAAQTAKQMLLEEDGGKKIAVFDTKAAGPTERMAAIK
ASQLIAAGLAFEEVVATVQAYIDRAKIFFGLQSVINLANNGRINKTVAKLAQVLKINVLGW
ANEAGEIEQLGKARGAKKARHNLLALLEQHHFDGQHLVIDHADNLAGAKAFKDLVLAK
YPACQVEIGNCNALCSFYAEKGGLMIGCLVKG
>transposase
MVLTTGGIALTLGSLEPTFYPFASLCCVFIEQKCYNVSEESEVNLMEIQITIKVKLNLANA
EIASSFTNTMEQYRLACNYVSEYIFNHDFDMQQSRLNKVLYTNLRSLFMLKSQMAQSV
IRTVIARYKTVKTQMKRNPYKYQDINTGEWYREIRDLTWLQKPITFNRPQVDLQRNRD
WSYLSNGQLSINTLNGRVKVNPICHGFDQYFDGTWKFGLAKLLKTGGKWYLHISATKT
VANFDNQTVKHVVGIDRGLRFLATTYDEQGQTAFFDGKAVMRKRAKYQKLRAKLQAK
GTKSAKRRLKKLSGRENRWMTDVNHRLSKTLVQKFASNTLFVLEDLTGVSFERTDLPK
ALRNQNKSWAFYQLAQFLTYKAHLNNSEVIEVSAQYTSQRCPKCGTIKKSNRNHDLHE
YHCSNCDYRSNDDRLGAMNIQLLGTQYISGQEQPKFELTTNA
>hypothetical protein
MQEQVDDTQAQILEIESGDKAAKSKWRLQLKLVMVTAICVALGLHLLLHLFS
>MutT/nudix family protein
MKHKKIPTTLTNMVMIYDQDGNVLVEDPLNPTWPGITFPGGHVEADEPLVDAAVREVW
EETGLKVSQLELCGIKDWLEADGSRYLVLLYKTQTFSGQVKSSREGNIFWTKLADLKQ
LKLASGMETMLEVFLSPKYSEHYLDTTNGRHDNLLK
>COG2110, Macro domain, possibly ADP-ribose binding module
MEKLALPVGTFYLEADGQRHDFTVKELTEETNEYLAKSERPPIEKCYLLRPNLPAEFSY
RQLLLRAEWADLDVSFDDNLSDEFFCGASWFVASKVVGLVSRVDNGELDDHVASGQ
GSLIPLAYVDVDERFRADLNFLIVYQGQSAYQADELSLFFAADQVCYDVAAPVSNASEV
LGPTSYYSYCQVCFHPGQKAYHYLSADESLQEGDLVIVPVGANNNLQVGQVVACANY
PATKVPYPLLQTKQVLRRANEAEQLDWWERQEVTPEGKQPENEAEFYLKLGAQAYSR
AEYGLAKEYYQQAVALGSSQAACNLGYIYAYGRLGQPDYEQAFYCFTQAALATNPNG
SYKLGDAYFHGQFVAANMEIAFKYYRQAEIFLARDGSDDDLKADIYYRLARAYHQGLG
VEANELAALDYINQAQLYSYVDEQAGRFKGAKLRQEITQLQVEVLAKLNQVQTKIALTL
GDITKQGDVEVIVNAANKSLLGGGGVDGAIHKAAGPQLLAECRQLNGCQTGQAKITGA
YNLPNKYVIHTVGPIWNGGAHGEADLLAACYENSLRLAQKTGLKTIAFPAISTGIYRYPL
EEASQIALATVQAFVKANPGAFELIKFVLFDQAALDCYYRKLGGLQVEL
>Methionine ABC transporter ATP-binding protein

MGEFTQRIQLSGVKKSFGPKEVLSAVDLTVNTGEIVGLIGPSGAGKSTAIKVMLGMEKA
DAGQALVLGQVMPKRKLLGQIGYMAQTDALYENLTALENLRFFGKLKGLGNQQLAKEI
EHLATVVELSGALKQRVASFSGGMKRRLSLALALLGQPDLLILDEPTVGIDPALRQKIW
AELAKLKAAETSILITTHVMSEAELTDRVALLFEGKIIATDTPDRLEAKYQVETIEEVFLKA
EGEDK
>ABC transporter permease protein
MRSIVVAKRVLLELLRDKRTLALLFFAPVLILFLLNLMFAASTDVKVKLGTSGLEPALSEK
LAKVKGVSVKNYQTTTLARADLKKYRLDGVISYQADKYTVTYANLDASKTQLTKQALK
QALLSQQLAGLQANLALGRGEKSLQAKMPVVVNHYQYGDSQTNFFTKLIPVLMGFFTF
FFVFLISGLALLKERTSGTLARLLATPIKRSELVYGYTLGYGLVACLQASLIALVSVYLLKL
EVVGSLVAVVVVSVCLGFVALAFGILLSTFASSEFQMMQFIPLVVVPQIFFSGIIPLASLA
PWAQVIGRILPLTYAGDALSKIIMQGATLRQVSGDLLVLVGFLVCLLGLNILGLKRYRKV
>Transcriptional regulator, TetR family
MTVKTILATYQEVVARKKLTPGQQKALLAAIELFAKQGYDATSTAQIASWAHISEGTIFK
YYPSKEQLLLAVVELLTDDLLPRLQAEFLAKVVASDKVALPALVRVIVTDRLAFMKANQ
QLLRILLQEGLTKPKVRTLVSQGFIQSLKANGEVFTRFLTLVNHYHPDYEPAVIIRLILGP
LVAYFSQRFILAPDLAYNEKRDIDLMVAQIVAGLS
>hypothetical protein
MAKSLETVYLKDGQVDVAHYHDLALQKQKEHAKFLATLKKKKPKHLDQLVQKVHNEVF
EQIDCTKCANCCKSLGPLFTEADIERIAKAMRMKLGKFEETYLRVDEDGDKVFNCMPC
PFLGADNLCSIYALRPKACREFPHTDRKKIYQINHLTIKNTLYCPAAYLFVEKLRQELG
>N-acetylmannosamine-6-phosphate 2-epimerase (EC 5.1.3.9)
MQVLFSYTEGNSSLGGIFMTNNFLKAVAGGLIVSCQALENEPLHGSHIMAKMALAAKQ
AGAVGIRANSVSDICAIKQAVSLPLIGIIKKDYPDSKVYITPTLAEMRQVAHTKVEVVACQ
VTNQPRPHNEDLQASLRAFRQEFPDTLLMADTGDLADVKLASKLGFDIIGTTMHGYTP
TTQNANIANNDFAYLKEVLATTELPVIAEGKIDTPAKAKRVLELGAHAVVVGGAITRPQEI
AQRFVDRIKASH
>Uridine phosphorylase (EC 2.4.2.3)
MLIKNEWPILEYDTSQQAVLAPDHEQLDLKLPKLAVFAFLGSQIEAYAAKHQAKRVAEF
VSATKTYPIFVIERAGQQICLVQAPVGAPAATQLLDWLIAYGVSRVISTGSCGCLVALPE
NTWLLPKKALRAEGTSYHYLAPSRFVELNLKALAALKRALTRQGTDYQVVTTWTTDGF
YRETKELVAYRKAEGCQVVEMECAALAACAQMRGVIWGMLLYTADSLADVSHHDSRS
WGELAAESALELSLAAVLELGDENEI
>putative arylalkylamine n-acetyltransferase
MKFRQVQKSDLEAVVAIEQAGFSPAEAASASTFAARIENMAETFLVAVADQGQVCGFV
CGLTTTTRYVEDWMYEAQAKPISDGKYLNILSIAVAPKKRGQKLGSALLTALEEVAKAK
EISSISLTCLADRIGFYERNGYQNQGQSASTHANEVWYDMEKIISKKF
>Methionine synthase II (cobalamin-independent)
MTNLAHYDIVGSFLRPQELKQARADFEAGRIDAAALKAVEDQAIKDLVAKEEASGLLVV
TDGEFRRSYWHLDTFWGFAGIEHIVRKHGYFFHDEETRPDSARVVGKIAFAGTHPDLE
AFKFLKQLTAGTNLTARQSIPAPAQLYAELVRGPENVAALRKVYESDDELFTDISNAYR
DLILALYEAGCRDVKLDDCTWGMVVDDRYWEQRDETEFDRDALQEIYLRLNNEALRD
LPADLRVSTHICRGNYHSTWASKGGYQVVADHVFAKENVSDFYLEFDDERSGDFRPL
AALPAGKRVVLGLVTSKRPELEDKATLKKRIEEASQYVSLENLALSTQCGFASTEEGNI
LTEEEQWQKIKLVIETAKEVWGN
>Methionine synthase II (cobalamin-independent)

MTNLAHYDIVGSFLRPAELKQARAEFKEGKLSKNDLRALEDKLITDLVAKQKEVGLQAV
TDGEFRRGFWHLDFLENLNGVEGYHPQAYNQQFQGTNAPAYNVRVVDKLSFNFNHP
FLKDFSFLKEVAGPEVLAKATIPAPTMLYRQEFLANDGTSKLSEVYPDLNDFYHDLAKT
YSDAIQAFYALGCRYLQFDDTNWAFLVDEHKRADLKAKGIDPVEATKISRQIINAALANK
PSDLTVTTHICRGNHASSWLFSGGYDQIAPDLFATDYDGYFLEFDDERAGGFAPLKYV
PAGKRVVLGLVTSKRPQLEAPTELKARIKEASQYVDLADLALSTQCGFASTETGNKLTE
ADQWAKLKLVVSTAKEVWG
>Homoserine O-succinyltransferase (EC 2.3.1.46)
MMTVTSKNGLVKTKQLWQENSAAGLELLIVNLMPTKLETELQFLKLLAKTKINCQVTFA
YPDSHHFRHGDGPLVKQAYLPLSQALKQNFAGVIVTGAPVEHLAFEQVDYWHDFCQL
VAWAQATNTYALFECWASQAALKVQYGIEKTARKAKLFGIYQAKIQAETNHFLKGFGS
GGLLRMPQSRHTDLSLPPKLPKGLQLLASSNEVEAFLLATSDLKNLFITGHPEYATETL
AGEYQRDVRLGKKISLPSNYYRQRQIVNSWQPTSVRLYTNWLQVIQKIG
>O-acetylhomoserine sulfhydrylase (EC 2.5.1.49) / O-succinylhomoserine sulfhydrylase (EC 2.5.1.48)
MTEKQVTTTNDHFETLQVHAGQGLDETGARAVPIYQTTSYVFKDAKQAAERFALKDP
GNIYTRLTNPTTEVVDKRVAVLENGSAGITLASGSAAITYAILNVAGSGDEIVAASTLYG
GTYNLFSVTLKRLGITTHFVDPDDPANFEAKINEKTKAIYVESIGNPGINLVDFEALGEIA
HKHGLLFIVDNTFGTPYLVRPLDHGADVVVHSATKFIGGHGTTMGGVVVEKGDFDYQA
SGRYPGFTEPNEHYEGLKFADLKGAAFTTKIRAEVARDTGAAISPFNSFLLLQGLETLS
LRVERHVENTRKLIAFLQSRPEVAWINYPELADNKYNALAQKYFPKGVGSIFTLGLTGG
KQAGQELIDHLRLFSLLANVADAKSLIIHPASTTHAQLSEEELKAAGVTPDLIRISVGIEN
ADDLIADFAQALDKLSD
>Protein of unknown function DUF419
MEQKELLAYIEQKYQVQAEYLWQKFPTYAVFRQQANQKWFALVMTVPAAKLGLSGQ
GNEELLNVKLDPELVGLLKQMPGFLPAYHMNKATWISIRLKQIAAEQIKDLLAQSYDLT
QK
>hypothetical protein
MVTIPKKFKVAVGQLFIARLLENSGGKILFTPIPSKWKLLRHADKHPK
>hypothetical protein
MIVKTRKVGNSTVLTVPKDFNIKVAKEYKPRLLADGSILFAPKSKKRLGTVRPEN
>hypothetical protein
MNQFDNQEPNWGNINITDKEMFNQVMQHDDLCLRFLQTVFPSAKLTKLTPVSESERQ
EALGADSDLLDLLVYDQNQYLFHISLLDQPVSPDDIADLVVATSKALVPFHQKFKSTAS
QASGLYNIIVYPYHYLDDDEVLHRFHGLFVLKSLQVVNVKTVLITTSDLPA
>FIG00742402: hypothetical protein
MIGTLTNTCSILVGTSIGCLLRSGLKEKYRQVLFLAMGLAAFGIGLENLTNAMPKSHYHV
LFIVSLAIGALLGTWWQIDKKFNALVTKYSQSQVGQGLTTGVLLYCFGALSIVGPVMAA
VKADYTMLFTNATLDFVTSIIFGASFGWGMLLGAPVLFCWQGSIYLIAKFLSASFFSGPL
VTELTIIGGFLITASGISLLKLREIKTLNFLPALFVPVIFFLVKGWLGW
>Transcriptional regulator, MerR family
MPTTTQPKTYTIGQVAAMFNLSTYTLRYYDKEGLIPNLKKDAAGNRAFSEENLATINIIE
CIKKAGMPIKDIKQFIAWCNQGDASLKQRLAMFQQLRQDVLGQLAQLEETLATIEYKCD
YYQQAVAAGTETYVKAKAHNTENVHD
>Ethidium bromide-methyl viologen resistance protein EmrE

MAYCYLLLAICLELIATNFLKLSAGFSRLGPTIITVSAYCACFYFFSLSLRGIKLSIAYATW
SGVGIIFASLLSYFLFKESLTGWQLLGLGLIVCGTIIANLTGGH
>hypothetical protein
MVKLFLGILTLLFLVALVMILVNRSHTEYDEMQVALQAKGYRYAFWSVLLFNFISLGGQL
VGMISARWALISPVISIWLGLVVYTCYAISKGAYLPFKTKVSLGEIGLVWFGPGLISFGLA
CFDWLSAKQVAVVLDVYLLLAGLFMLLVAFALTVSLKKKRVSD
>Transcriptional regulator, Cro/CI family
MKNLKLKAARAGKDLSQQELAQLVGVSRQTISAIEKGNYNPSLNLCIKICQTLDKGLDD
LFWPND
>ABC transporter, ATP-binding protein
MSIIQINNVTFAYPNGSENVFTDLNLQLDSSWKLGLVGRNGRGKTTFLKLLLGKLDYRG
NITSSLKFEYFPYPVVDEDQFVGELLQEIAPTAQEWEVVRELSYLGMEADSLWLPFSTL
SKGQQTKVLLAGLFLKPDCFLLIDEPTNHLDVAGRKLVANYLKKKTGFILVSHDRNFLD
ACIDHVLSLNRADISLTKGDFSTWLQNFENDQANKLAQNDRLKKEVKRLKTAARQSAS
WSDKVERSKNGTTNSGSKLDKGYVGHKSAKMMQRSKSLQKRQAKELAAKEGLLNNL
ETISELKLSPLTFHEQRLAWAENLQITYPNGWKSKAVSFELKQGQRLALTGANGSGKS
SLLKLLLGQAIPYQGNLTVKPQLKISYLPQDASFLAGPIKTYAEQLGIDETLLFTLLRKLD
FERDLFSKDMTNFSEGQKKKVLIAASLCQSAYLYIWDEPLNYIDVYARMQLEALLAKVK
PTMILVEHDEYFLQGLATDYLEL
>Phosphoglycerate mutase family
MTKTLYLMRHGQTLFNQMHKIQGVCDSPLTDLGINQAKLARAYFNDRQIKLTAAFCST
SERASDTLEIVTNNSLPYTRLKGLKEWDFGRFEGQDEALNPPLPYEDFFKTFGGEGQS
EVGLRMEATILDLMATAAGENVLVVSHGAACACFYRQVTGQSQLPVRGIKNCSIFKYE
YKNGKLTYQDLWQPDFSQLDNK
>Phosphoglycerate mutase family
MTKKLYLMRHGQTLFNQMHKIQGWCDSPLTEAGIKQAQAAGRYFKEAGIHFDAAFCS
TAERTSDTLEIATNNSLPYTRLKGLREWSFGRFEGQDEALNPPAPYGDFFKYYGGEG
QSQVEARLVKTLTDIMKQTSGQNILAVSHGGALINFYWGITNKRSEFLPITNCAICVYDF
TDGHFTFKDLIQPTLLYQEEKIS
>Cof-like hydrolase
MEQKLIFLDIDGTILGYDGTLPASAVTAIKTARQNGHRVYLNSGRSKAELPKELLAIGFD
GIIGGNGNYVENDGQVIFHQALPLETEQAIVTWLTERDLPFYLESNAGLFGSPDFEIKAK
PAFQAYIHGKNPQADTSSLKVEDELHGLILGADLTRTDVNKISFALKSYQDFLDAKAAFP
DLEVGTWGGKDELALFGDLRPKNITKSNAIDHLLTFLAATKEQTVAFGDAKIDIPMFTYC
NLGVAMGNAGPETKAAADLITTDLMDDGLFNGFKEIGVLSND
>hypothetical protein
MAEGARLELCHKADLPQVATERLGLAHLAFNLGSKEAVVTDLEGNLLELTTKKRPGKF
PTSF
>Aminopeptidase C (EC 3.4.22.40)
MTKEISPEQLADFQANLNTQPAHAVIERAVTKNGILKTAADWQAEVDMAPVFSIDLDTG
KVANQKQSGRCWMFAALNTMRHDMKNKLKIEGDFELSQSYTFFWDKLEKANYFYQN
VLATAQLPTSDRKVAWLMTTPQQDGGQWDMFMAIIEKYGVVPKSVMPESFSSSASRE
FNATLNLKLRKDAVELRKLVAANASEAEIAKRKEEMLAEVYRMASYSFGQPPVKFDFE
YRDSDKNYHIDRDLTPKAFFAKYIGWNLSDYVSIINAPTADKPYNHMYTVEMLGNVVG
GRQVRHLNVDMATFKQVAIDQLKAGESVWFGCDVGQESDRQKGIMDTNLYHKDELF

NVDMQMTKADRLDYGESLMTHAMVLTGVDLVDDKPTKWKVENSWGEKVGTKGFFV
MSDAWMDEYCYQVVVNKKFLPAELQAVLTEEPKVLAPWDPMGALA
>Ribose 5-phosphate isomerase A (EC 5.3.1.6)
MNQTELKKLVGQKSVDWIKDGMIVGLGTGSTVKFMVDELGRKVKEEGLNVVGVTTSN
VTAKQATELGIPLKSVDEVDHIDLTIDGADEISADFQGIKGGGAALLFEKIVATNSKKNL
WIVDESKMVETLGAFPLPVEVIPYGSQRIFDRFEEKGLNPSFRKKADGGLLLTDSDNYII
DLHLGRIEDPHALADYLIKQVGVVEHGLFLDRVNTVIVGKQTGPEVLEAR
>acetyltransferase (putative)
MYTWTDNSMDYANEEGRHLANIGFSMIDANQTYVVEHTWVSEDARGMGLAGKITVNF
LEHARAEGKKVLPLCPYTKKFFEKHPEYADLLKA
>Deoxyuridine 5'-triphosphate nucleotidohydrolase (EC 3.6.1.23)
MKRGFKVVSKYKEQEIKLPTRATKNAAGYDFYAREDFVLPSIWKLGFLKVLHALRKQK
EISDEELLLGQKILKPYLVPTGIKAYMQADEFLLLANRSSSPLKRSLILPNGVGIVDADYY
DNEANEGEIFFQLLNFGLTDQVIKKGERVGQGIFMPYLLADGEQSPSQKRTGGFGSSG
R
>DNA repair protein RadA
MAKTKTKYVCQECGYSSPRYLGRCPNCGAWNSLVEEIEEKTSPKAAANARVTISGTKV
SPQLIEQVSIEKTPRIKTNLKELNRVLGGGIVPGSLVLIGGDPGIGKSTLLLQVSGQLADL
KSRVLYVSGEESASQIKMRADRLGVGGSEFYLYPETDMGAVRQSISDLKPDYVIIDSVQ
TMQEPEMTSAVGSVAQIREVTADLMQLAKTNGITIFIVGHVTKGGAIAGPKILEHMVDTV
LYFEGDKQRTFRILRAVKNRFGSTNEIGIFEMREGGLYEVANPSEIFLEERLQGATGSA
VVASLEGTRPILVELQALVTPTAFGNAKRTTTGLDHNRVSLIMAVLEKRAGLLLQNQDA
YLKAAGGVKLDEPAIDLAIAVSLVSSYRDKETRASDCFIGELGLTGEVRTVTRIEQRVAE
AAKLGFKRVIVPKNNLDGWDFPAGIQVLGVTTLNEALHIAFD
>Membrane-associated protein containing RNA-binding TRAM domain and ribonuclease PIN-domain, YacL B.subtilis ortholog
MKGIVNNMRKRLIKIVFAILGIGIGYSFFPWAWRLANFNQPWLVNPIMNMFLGLVIMSFL
GLLLSNPLEALIKKIESSLLHQKPEVLITEGISMIVGLVIAALISTLFSHSQLFFFNTILPVVL
MIIFGYLGFIVGKMRFSDLPKLFALKKKAESEVVKATPEEEVLERKVGENFYRYKLLDTN
ILIDGRIYDLAKTGFLEGTLIVPNFVLYELQYIADSGDSIKRVRGRRGLDILNKLRTEKLVP
VEMYEGDFEDIQEVDTKLIKLAKLLDAEIITNDYNLNKVSEFQNVRVLNVNALAKSLKPR
VIPGEKLNVMVIKNGTERQQGVAYLDDGTMIVVEDGRYFMNKTIEVEVTSAIQTDAGR
MIFAKPAHSRREIKEKN
>Methyl-accepting chemotaxis protein
MRARFTVGKNVKTKLLFPSVALIFVICIIFSLSSLNSLRGAITDLGISQGKLAAKAVTDAIN
PADISKVTISHQNKQAINRIRASLDSVSANSDIKSAYTIIKKNDQFVYQVDGAKNNPATLG
TAFTSNHALLSQAFAGQVATSSKIQKTNQGYLITAYRPVFNKNDQVIAVVGVDFRADKA
ITMLKKKALILIGISLFCLISSILVLNACLSILQKFIKSIDQNIYDLAHTDGDLTKRLDIKTNDE
FEVMANHINELLDNLTKVMTKINDGSTAVTTASEKMSVNITGVQENVADVSASMEEMS
ASMEETTASMNEINDSIQENKTTTQDIASRANDGFNDSKKTIKKAKQIYQRSSKEQDSA
KNDVLNITKSLNAKITGAQAVEQVNDLTNEIIAISRQTNLLSLNAAIEAARSGEAGKGFAV
VAQEIGKLASQSSDVANEISTVSAQVIQIVTELSGEAKKVGDFLNETVTNGYSQLLAIAK
DYQNQSETLGVLLQDFAQSSHGLNQKLDVIADSAQAVNTATDESAKGITQVTESAMDI
TNKIDELKQAAKQNEQTATTLQKTVAYFKLQ

>PTS system, sucrose-specific IIB component (EC 2.7.1.69) / PTS system, sucrose-specific IIC component (EC 2.7.1.69) / PTS system, sucrose-specific IIA component (EC 2.7.1.69)
MFTLFKKKAEPFVVSAPISGKLIDLADVNDEVFSEKMMGDGFAVIPAKSKSIVSAPVSG
VIEMVADTMHAVGIKNDDGVEILVHIGLDTVTLNGQGFKSFIKAGKRISRGDKMIEFDPE
FMKEKGIDMTVMTIFTSGYDKDIVLNVAYNTQVTPNSVLLD
>Beta-glucoside bgl operon antiterminator, BglG family
MRAIKKINNNVAVCIDGNGDELVAFGKGIGFKKMPYEIADMSLISMTFYRVDMRSMQLL
KEIPENVMQISAEIVVRAQALLPDKSFSPNIIFSLADHLNFATIRLKDYKGMKMPLDYDIK
QMYPVEYQIGRYACQLFKERVQVVFPDSEAVAVAMHLINAENEVVYTLNDESFDDMA
RMISNYIEDFFQIKIDRDEFSYNRFLTHLRYYFKRIQDDTSNEGEINPTMLSAFKREEPK
VYQCSVEVVNQIDHRFHTTTSDDEMFYLMVYINRIVKKSQIKNGG
>PTS system, beta-glucoside-specific IIB component (EC 2.7.1.69) / PTS system, beta-glucoside-specific IIC component (EC 2.7.1.69) / PTS system, beta-glucoside-specific IIA component (EC 2.7.1.69)
MAKKDYSQLANSILENVGGSANVASVTHCMTRLRFNLKDMSVAKDDTIKGLSSVLGVV
RSGGQYQVVIGQTVNEVYDAVVSAGNLTTAAPVNENLDDKPKEKMTWKSVGNTILNK
VAGSLTPLIPMLVAASMFKMLAAVTGPSMLNWFSANSDTYKLFTFVGDAGFYFFPIIIGY
TSAKQFNTNRILAMFLGAILIDPNLINIVTAGKAFHVFGINMSLVNYSSTLVPIMLSVWVM
SYIERFFNKHIVASLRTILSPTLTITVMLPLALCVLGPLGGFIGDGICKGILAFGNMGGIWA
ILGVAIIGALWEFLVLTGMHLVMISAMTLLFAQGGHDNFVTLGAVAASMAVAGMCLGAS
FRLKGEERSLALGYTVASIIGGVTEPGLYGVAIRYRRPFVGMMIGGFCGGLYAGIAHCV
GYVMVPVANFLALSTYIGGNSANIINGIISGAVAFIVAAIATALLGIESKSKVKTEEAETTH
ADVSDVAHA
>Cellobiose phosphotransferase system YdjC-like protein
MTSRILVRSDDLGYSKGVNYGIYEAVKNGFINNVGFMVNMPTSQHGLDLVKDIEGLNL
GLHTVICAGKPLTDPKLIPSITTEDGYFKASKTYRSAKEDFIDFDEVVTEIEAQYQKFVEL
VGRKPDYFEGHAVVSDNFVKGLRTVAKRHDVPFLDFAFDSEAHVLFKSVAFKSVMES
MGNPNYNPVEVVKKAAEFAKTHANVIPMVIFHCGYLDDYILGHSSLTLPRTKEVEASCD
KDLAQWLKDNDVQLMRYSEV
>Malate permease
MNPNIIAIIIGLILFIFSIHLPVVVNQFLGYVSDAFTPLSMIVIGSNLSNMELKDIKLPGYLWL
LVLLRNLIYPIMTGFVLRACGLIRVPLMATIILSACPFAGGVSWFSLQTHGDAKAATVAM
SVSTLLSLVTITVVFALLQNI
>methyl-accepting chemotaxis protein
MHIRFTFGKSIKTKLLVPFVTLVFIICVIFSLSSLNSLKSTVTNLGISQSKLAAKTAADSLNP
ADVKKVSVTHKNNSSYRKIRADLLRITTGSDIKYAYTIIKNNGQYVFQVDGDKNDPTALG
SAFTSDKELVTKAFSGQVTSNSKIQKTKDGYLITVYRPIMDKNGKVVAVIGVDFGADEA
MAMLKQKALFLVAISLICLATSTIFLNFSLSVLQKFVRNIDQNIYDLAHTDGDLTKRLDIKS
NDEFETMANHVNDLLDNLQEVMKKITAGATSVDSASKQMVTNITGVQAGVADVSASM
EEMSASMEETTASMSEINDSIQEIKTTTQDIASRANSGFKDSQATITKAENIYQTSSQQK
AEANDEVAAITAKLNSKIAGAQAVEQVKQLTSQIIAIAQQTNLLSLNAAIEAARSGEAGK
GFSVVAQEIGQLASQSSDVANKISTVSAQVIQAVTELSGEAKHVGTFLNETVTTGYNQL
LDIAKDYQGQAETLGSLLKDFAQSSQGLDQMLDEIAQSTISVNTATEESAKGITQVTES
AMDIANKVTDLKEAATSNEKVAADLEGTVAHFKL
>hypothetical protein

MSKLTLILATLVALEFFYIMYLETFRTTSAKTAKVFKLTPADLSNPTVQTLFKNQGIYNGL
LAVLILLAIFVFASKVATVILMSYIILVAVYGSLTSNPKIILMQGGLAILTLLSTLF
>Rrf2 family transcriptional regulator, group III
MLAAKFAEDLLMKFSYKLSDAIHILAYLEIFKDGDRSSRQIAASIEANPSVVRSLMSNLRT
AGLIVTRQGAAGANLAKSPAEINLYDVYQAIDMQRLLHIDPKTNPNCPVGANIQTVLDQ
VYTAAEQAAFAKLKDTSLADVIAAIKVHF
>Oxidoreductase
MKYAVTAATGKFGQAAVKALLEKLPSQELVIIARNLTKAKQLFPDLEIRPADYDSEVELT
KALTGIDKILFISSQPGGAVSRDIQHQNVVNALKKAGVTYVAYTSYPHANSAQAGLAAD
HKLTEQLLQASGIKHSFLRNNWYFENELDLISGLAQTQTGAYWADNKTGWALESEYA
QAAANVLLMAQPKEVYEFAGTPLSYEELAQAVGTATNQTVTITKLSHEDYVNSLVQTG
LDHETAALYASFQIPNSLGDLAQPSDDLTTVLGKQPLEITAAIANLLQN
>Maltose operon transcriptional repressor MalR, LacI family
MTTIRDVAKLAGVSPSTASRVLHDSDMISEATKKKVRQAMKELDYSPNYLAQNLANKS
NNMIGIVLPVRHDQPALSNNPFFMQIIQGIVTVCNQRNYMVSLATGQTDEELIKNLTTLS
KQGRINKFIFVYSKENDPVFDYVQKRTEAETIVVGSPYAKVTKRARYVNNDNVQAGKD
LTNYLLKRGYQEIAYVYSNLSEVVQEDRYMGYKDALKAKAEHLLCLESSDQTELVANFL
ETNPEVDAFVACDDMTAIFLQQILTGLGLAANAYGLAGFNNSLIGQLVHPSLTSVEIHPY
ELGMEAAKLLTQATAKNYLLVEHTIIKRDSTK
>Maltose phosphorylase (EC 2.4.1.8) / Trehalose phosphorylase (EC 2.4.1.64)
MKRIFEIDPWKIISHDLNPADKRLQESMTSIGNEYMGMRGMFEEKYSADSHKGIYLGG
VWFPDKTRVGWWKNGYPEYFGKVINAVDFVSVDIKLNGESIDLAQDKYSDFEVSLDM
QKGILTRTYTVHKGEAKLKVTSTRFLSIVLKELYANRLEFENVGSVPVRVEVASKIDADV
YNEDANYDEQFWNVLGKEVSAKEGRLVAQTIPNNFGTPQFMTEMAMFNSGDLEAVDL
GQTAEKEVVNTLATELAVGAKANFEKRVIVTTSRDYETKADLDQATSKLLDLQAEKTFD
QLLADQKEAWAKRWQKADVTISGNDEAQQGIRFNLFQLFSTYYGDDARLNIGPKGFT
GEKYGGATYWDTEAFAVPLYLSTADQSVTRNLLEYRYKQLPGAYHNAKQQGLKGALY
PMVTFNGIECHNEWEITFEEIHRNGTIAYAIYNYTRYTGDQEYVTHDGFDVLVGIARFW
ADRVHYSQRNDKYMIHGVTGPNEYENNVNNNWYTNMLAKWCLTYTDEIAAKVGEDKL
AELGLNQAERDKWREIVAKMYLPEDEELGIFVQHDGFLDKDLRPVSAIPADQLPLNQN
WSWDHILRSPYIKQADVLQGIYYFIDKFSPAQKKANFDFYEPFTVHESSLSPSIHAVLAA
DLHEEDKAVEMYSRTARLDLDNYNNDTNDGLHITSMTGSWLAIVQGFAGMRVIDDQLV
FNPFVPKDWDGYSFHINFRGRLLSVNVDKVGTKIDLLAGDDLTIKLNGTEETLTA
>Beta-phosphoglucomutase (EC 5.4.2.6)
MKGVAFDLDGVITDTAKFHFEAWSKLALEKFDLELPAEFESQLKGISRIESLERILAFGK
LSGKYTPAQVAELADEKNTYYVAAISKLTQADILPGISQLLADLKAKQVKLSLASASKNA
PMILEKLGLLTYFDAIVDPSKLKAGKPAPDIFIEAAKAVGLAPSECVGLEDASAGVAAIKA
AGMVAVAIGDKEELAQADTVVSSTADLNYDLLVTSYQQAQGK
>Maltose operon transcriptional repressor MalR, LacI family
MVTLADVAKEAHVSKMTVSRVINHPEKVTKELQKIVFEAMAKLDYQPNLAARALASHKT
RIIKLFILEEMTVVEPYYMTLLVGISKQLEKHGYTLQVTPKLELNPGLCDGYIITGLRELDY
SKLAAFHKPLILFGDNRHGYDFVDNDNYEGLRLATHYALKCGYEHLVYVGTDIEEPFEL
KRRQGYLDVLAESKLQSELYFCANHSHVAANLIKAKVAQWQANTCFICASDRIALGVS
RALKASGKAVGEDFGVIGHDGVFLDQVGDLTTIKQPILEMGQALVEILLAKINGQDSSN
YQRIYKSKLVVRTSTRKETDAYATNHLSDLS
>Oligo-1,6-glucosidase (EC 3.2.1.10)

MLMQPIIYQIYPKSFQDSNHDGIGDLKGIEQRLDYLKDLGVTMLWLNPIYKSPQVDNGY
DVADYYQIDPSFGTLADFKALVAAAHERGLKIIMDFVMNHTSDEHPWFKEALKGANNP
YHDYYLWQTAPSGQLPNNWASFFGGSVWEPVGHNEYYFHLFDKRMPDLNWHNRKV
RQEMVKIAQY
>Oligo-1,6-glucosidase (EC 3.2.1.10)
MDAFIHIVKADFKQNVLTPKKGPQLAEEYYANLPLVQEYLEEFISELKQLKPDLFIIGEAA
SAEVSLGFDYTNPKRDLCDTVVTFRYFKEKPSPDLPALARQTLDLVHLKTELSKWQQG
LGLGALPTLYWNNHDLPRILSRFEVPSEAKEKLAKSLAVAMYLQRGIPCIYYGEELGLE
SLELKTVEAYQEKLELKQALAAGYSQNQIEAKLPLKHKMAARTPMPWDDTKVHQGFS
DATPWLETKEDAPVTVTAQLHRADSCLSAYKQLLALKKTPLFTAGQEIILPSAEDLVVFE
RQLNNKRALVLANLSAKPQVFKGEFAPQTCAFKNGDYQVVKQEIRLAAYSSCVFINDE
EEA
>Neopullulanase (EC 3.2.1.135)
MNLAALYHRPDSEMAYLKTPTVFQIRLRTAKADVSKVSLYYGDPYGNKVNEKDESQW
EYQEIAMTKLGTSELYDYWQAEVSLPLKRLQYLFAIDFKDGQSYFYDDRHVWPATKKH
LLASTCFRMPYFHEVDRKKAPAWVKDTVWYQIFPERFANGDKSNDPKGTLAWGSVPP
SRTNYFGGDLQGVIDHLDYLKELGIGGIYFCPLFTAGSNHKYDTIDYFNIDPAFGDKAKF
KELVDEAHKRGIKIMLDAVFNHMGDFAMQWQDVVKYGQNSRFSDWFHINSFPVSYEE
DGLEHAKNISYDVFANTPHMPKLNTANPEVKAYLLEIATYWIKQFDIDAWRLDVANEVD
HHFWKAFYEATHALKPDFYVVGEVWHSAQEWLAQEEFDAIMNYPFTEAILDGVILKKT
KPKEMLAALLDNLMKYRDQTNQVMLNSLDTHDTARILTLAKGKQDLVKQALALMFVLP
GSPCLYYGTEVGMAGGNDPACRACMVWDQTKQDLEMLAFTKELIALRKRYASYLNQ
VPSWSEQAGLVTLNLGQKLQASFNLSPKAVAYPKVSQPLLANGLSDKQLKAGGFVISL
AD
>ISSth1, transposase (orf1), IS3 family
MSRKSKYSVEQKLNILNEAVHSSFKRVAKKYGVNESTILTWYRLYKYQGIDGLRSIRSN
RSYSKEFKLFLVEQYRNSDDSLELFAIKHGLKSKSQLLNWIILYNESKLKAYTPRKRDSI
MPGRKTTFEERLTIIEDLIKHDVNYNWAVNKYHVSYQQVYGWYQKYRKSGNDPQSLR
DRRGKAKPKAEWTELDKLRAENRLLKAEMKRNEMEVAFAKKLIEIRNREAKKSSNTKS
>Mobile element protein
MLEAILKLEEKHKWTLGYLGMTTQLAFENKLSFKAGLKRVTNCMRNHGIKANIRKKKHN
RIKRHKEYINDNLLNGQFDRQNKNEVWVTDTTEVLYGSEQVRKVRVHVVMDLYGRYIL
SYNISATETATSAIEAFKRAFKVEPGVQPMIHTDRGAAYCSRAFNDYLAEQNCIHSMSH
PGHPWENSPIERWWNDFKLVWLAKRARSKSLSELEQSVQQAIKYFNTERAYASKNGL
TAEKFRAQAA
>Maltose/maltodextrin ABC transporter, substrate binding periplasmic protein MalE
MGIWKKVGLATLAAGLTVGLAACGSKSSESNQKTLTVTVDKGYKDYINDIKGDFEKKN
GVKVVVKTKDALATLDSLKLDGPAGKAADVMMAPFDRVAVLGKQGQLATVKLADDGR
YNDADKKNVTLKGKTYGEPVTIETLVMYYNKDLIKKAPTSFKELETLSKDSKFAYENDK
TKNVAFLTQWTNFYNAYGLIKGYGGYVFGDNGTNPKEIGLNNKGSVEGLTYISDWFKN
VWPKGMMSTSSNENFVTDQFTSGKTAVVIDGPWMAAKYKKAKVNYGVAALPTLDNG
NKYQAFGGGKAWVISNYSKQKGLGQKWLDYVTNKKNQTKFFKDTNEVPASEQARGA
VAKGDNDLAKAVIAQYKDADPMPNIPEMAEVWTGAENLIVDSAAGKKTPQQAADAAVK
QIKQTIQQKYKD
>Maltose/maltodextrin ABC transporter, permease protein MalF

MKSKNPQKAMLLSLIPGLGQISNGQKAKGGIFLGIFALFIIEMLTFGASAVDGIVSLGHNP
GVDNSMFLLVRGTLQLLILVLFVMFYVINMVDAKRVAQILNEGGQVGTSFADIKESLATH
GFPYLLTFPAYILMIFTIILPVLVTLFMAFTNYDFNHIPPAALISWTGLENLKAIFTLSSYKA
TFGAVFGWTLIWTLCASTLQIVLGIFTAVVAHQDFIKGKKIFGIIFLLPWAVPAFVTILSFS
NIFNDSAGAINVQVIPFLNHLFPFLHIPQISWMTDPFWTKVALIMIQGWLGFPYVYVLVT
GILQSISDDLYEAATIDGASAWQKFRNITLPAIFAVAAPTFVTQYTFNFNNFSIIYLFNNG
GPGDVGGGAGNTDILISWIYKLTTGTTPQYGLASAITIIISLIVIGISLAVFKKTKAFDMEG
>Maltose/maltodextrin ABC transporter, permease protein MalG
MAAKKKSRSSIKRQRFIKHFFTYLYMVVLAVIIIYPLLITMLTAFKNSDVVAFTLSFKGPW
TLANFTSLFTDTLYGTWYWNTLAISIMTTIVQVAVVTLAGYTYSRYNFIGKKWSLTFFLV
VQMVPTMAALTAFYVLGMLLNGLDHYWFLTLIYIGGGIPMNAWLMKGYFDTVPYDFDE
SAKLDGAGNFTIFWKIVMPLVKPMIAVQALWAFMGPFGDYMLARYLLREENHLTVAVG
LQEFISDPKDQQVALFAAGAILVALPICALFFYLQRYFVTGLTAGGAKG
>Maltodextrose utilization protein MalA
MSRRPFPFNYFTNFASAKRLFLERGSYSWGQLILVTLFLLALVVLPVPFFYQGQTSTNL
ALYMPRVNKLLANPKLEQVAKTTPYTKSGYAFKGKQVLLTNRSEIVGVNLDKQDIASHK
AALSLNNNSFVLKDKGVSYHLFYTRQFAPNQGKLKQQLSQTWFLRNKAAISFSMMTLV
LNLFLVTNLLFSFGAAFFLWLARKSPALTINSYKEAVGLMVGILGPATLLTMIIEFVKFDIS
LAISIQMIAAVCLMLVVYAKTHFNEHYLNDERIGNR
>Oligo-1,6-glucosidase (EC 3.2.1.10)
MKWWQKAVVYQIYPRSFQDSNGDGIGDLPGIISRLDYLAKLGIGAIWLSPVYQSPNVD
NGYDISDYQAIAEEYGSMEDMDLLIKEARKRQIRIVMDLVVNHTSDQHPWFIEAKRSKT
NPYHDFYVWSDQANDLGSTFGGSAWEYDEHLKQYYLHLFTKQQPDLNWANENLRKA
IYQMMNFWIDKGIGGFRMDVIDLIGKQPAQKITANGPKLHTYLQEMNQATFGNSDLMT
VGETWGASPKEAKLYSDPKRQELAMVFQFEHMTLDQGQSKWDVKSLDVAKLKQVLS
KWQQELSQTGWNSLFWNNHDLPRIISRWSDEGQYRLKSAKLFAILLHLMKGTPYIYQG
EEIGMTNRKVTSIDQVDDIESRNAYQELLAQGLSQKEALSAINKVGRDNARTPMQWDA
SPNAGFSQGQPWLAVNPNYQTINVAQALADPNSIFYTYQKLIELRKNNPLVVLGDYQLL
KTSPEVFAYYRTYQEQKWLVVANLSHKENSIEITDEISQVILDNGAKPLVKGRNLLEPYD
ALVVKLA
>Multiple sugar ABC transporter, ATP-binding protein
MVELQLKHIYKHYDNAEPGVYNVTDFNLDIADREFIVLVGPSGCGKSTTLRMIAGLEDIS
KGEFLIDGRVVNDEAPKDRDIAMVFQNYALYPHMSVYDNMAFGLKLRKYDKAEIDKRV
HHAADILGLEKFLQRKPADLSGGQRQRVALGRAIVRDAPIFLMDEPLSNLDAKLRVSM
RAEIAKLHRELKTTTIYVTHDQTEAMTMADRIVVMRDGKIQQVGTPQEIYNNPTNVFVG
GFIGSPAMNFFKVTLKGNKIFDIKHDLELEVPEGCLKVLKDKGYEGKEVIFGIRPEDIHSE
RAFLETFPNSVVKAEVVVSELLGAESQLYSKVGETEFVSKVNARDFMQPGESLELGFD
VNKAHFFDPESELAVY
>DNA-binding response regulator
MKKILIFEDEPAIIEAIKQELTKWQYQVEAISDWEQVLEQVQAFKPELILMDISLPAFDGF
YWTQKLREFTKVPIIFISSADLDQNAVRALMIGADDYLVKPFALDLLVAKMQVLFKRLDQ
QPDEHLLTFASYQLNQLTNELSQGEAVVKLTPTESLILKLLVVNQGEIVTKLKLIQALWQ
GNNFIDENALNVNISRLRTKLEPVGLKDKLVTVRGKGYLLVL
>hypothetical protein
MDKWPLILAYFASILAVTLLVYLKSLSFELVLDYLRLTWLFALGGLYLVYCLKQGSWCY
>two-component sensor histidine kinase

MGSSYLEQAYEAYLADQEAQLKETERQLRTIYEERADYLKIWSHEIKTPLTALTLLAQTK
EEIPATEVNHQVDVASYQLNLLLNYERMADFNHDLEFRPVDLLALIQAVAQKNMNFFIA
KNIGLQIEVKPVKVVTDPKWLSFIVEQILINAAKYSGANKTVKISYQAGKLEVSDQGIGIS
ASDLPRIFEAGFTGANGRKYGAATGMGLYIVKNMAQKLNIDLKVSSILDQGTVVSLNFN
>ABC transporter ATP-binding protein YvcR
MAILTLKHVTKSFNKQTKVLKDINLEVEKGEYVAIMGESGAGKSTLLNIIATLDKATSGEV
SLTGFDLNKLKESEAAKFRRQHLGFVFQHFNLLDTLSNRDNIYLPMVLNKEKHQEMAQ
RLEPLASQLEIDALLDRYPYEISGGQAQRIAVARALITKPELLLADEPTGALDSQNTKRLL
DLFTKVNQAGQTILMVTHSALAASYSKRTLFIKDGLIYHELYRGDSSQQEYLQKISAGLT
ALTSEGGLN
>ABC transporter permease protein
MLELKLGWQSLRNNKQLYLPFTLATSLLIGVFYIFQAIINNDSLAKIPTASAINSIMQVALV
FTSLIIVVFMFYINNIVAKQRNKELGLYSMLGMTKANLFFLVLGVDVCLFGFSLVLGLLIG
LSFIKFVLLAFVKLLDLKLVHLTVFSGQALLLTVAFFLAIFMALLLVDLAKLSQVRPVELW
NATKKGERLPKNSWFTKLLGLAGVATLIGGYYLSVITKPNASVITTFFLAIILVVIGTYALF
TSASIFFLNLLKQNKNYYYKRNHFITVSGMLYRMKQNGVGLASICLLCTASIVALVATSS
LMAGKDNQIKQSAPLDVSITSKKITPATYKKTQELAAKYKVTLINKKSQRISHLALANIEA
GQLKTSQDANTANYEVLTLPLADYNRVEGTNYHLKNNELLVYATDKKVKLKQLQIKGK
NYRVRMLSEYKYRLGDVATLTGLYLIAPDAKSAQSLLNIPLNETYSFDLKGTDQAKENY
SQAASQQLKGRNSENFVTTKVEVSEIMNQFYGGMLFVGLMLSVTMLLTTGMIIYYKQV
SEGYADRARFKTMQQVGLSLAETKQAINSQVLMVFMFPIVVAAIHLCFALPALASVLKL
FSMYDLKLLLIVAVTMLAILIIIYLLIYSLTTKVYRRIVNAD
>Universal stress protein family
MNFNKILVGVDDSADALKAFDYAINFAKQNSAELDIISVLEEHDINVYESLDKDFIHGKYR
DLTEHLIKYQRQAQEAGVSQVKTLTAKGNPGETIVKQIIPQIHPDLLIIGSLAKKGLARHL
GSQAAYMAKYAPISVWVIR
>Manganese transport protein MntH
MEANSSMSRKHHLIEYANGPSLEEINGTVEVPKGKGFLKTLFMYSGPGALVAVGYMDP
GNWSTSITGGQDFSYLLMSVILMSSLIAMLLQYMAAKLGIVSQMDLAQAIRARTSKSLGI
VLWILTELAIMATDIAEVIGAAIALYLLFNIPLVFAVFLTVFDVLLLLLLTKVGFRKIEALVVC
LILVILFVFCYQVALSNPNWSQVVAGLIPSAQTFSDHPNVDGLTPIKGALGIIGATVMPH
NLYLHSAISQTRKINHHDEEDVARTVRFTTWDSNIQLTFAFIVNSLLLIMGVAVFKTGSV
QDPSFFGLYHALSDTSTLSNGILVEVARSGILSTLFAVALLASGQNSTITGTLTGQVIME
GFVHMKMPLWARRLITRLISVIPVLICVAMTNHLSVIKQHEAINTLMNNSQVFLAFALPFS
MLPLLLMTNSEVEMGKRFKNSPVVKFFGWLSVISLTVLNMKGLPDSIADFFGDNPSAS
QLAWAHGIAYFLIVAILALLVWTVVELHQGDKKLQAKLAAQSQTKGSVLDEL
>Alcohol dehydrogenase (EC 1.1.1.1)
MKKAIFEEAGKMGVVEAPVPVIENEDDVIIKVVRACVCGSDLWAYSAGDAKEAHSENS
GHEALGIVEAVGPAINTVKKGDFVIVPFTHGCGHCPACLAGFDGVCQAHSNNFSMGF
QAQYLKFAYGDWALVKVPGQPSDYSEGMLKSLLTLSDVMATGYHAAVTARVSKGDKV
VVVGDGAVGQCAVIAAKMLGAAQIIIMSRHEDRQQMALESGATAFVAERGQEGIDKVL
ELTGGSGADAVLECVGAQAATEQAFGVVRPGGRIGRVGLPHGEKMAPTATFYRNITL
AGGPASVTTYDKKVLLKAVLDGEINPGKVFTHSYKLDQINQAYQDMAARKTIKSYIVFD
D
>Substrate-specific component BL0695 of predicted ECF transporter

MKVKDIITTGIYTVLYFILMSLGTLIAVLLTHNPNMKFAPAITALLAGTVYMLLVVKVQKFG
AVTLMAAVLGLFFLVSGHSMLAFIPSLLCGLLADVVAKLGRYQSKFLNLLSYCLFSLGNL
APIIMMWTMKQAYISQLLAKGKDMAYVHKVLINFDLGNVSVLVGAILICAIVGGLFGQHL
VSKHFKAAGLV

>Transmembrane component BL0694 of energizing module of predicted ECF transporter
MNRGYDVRTKLLFLLLANYYALITPGIGVIWVLILLFSLDFIWRHEYVRLASYLIFFGLATA
CLSLNIDEASHPYLGFFLLGGLTIMPCFMVGLSILRTPVHEFIQTLRKWRVPESILLVLAV
AMRFLPTVTESYLNIRAALKLRGEFVSWYQYLLYPLAYFEKILVPLLLNATRLGQDLTIAT
LTKAVGSHHKTSYRNYVFTGKDYLGWFGLGGLVLFTLGGFL

>Duplicated ATPase component BL0693 of energizing module of predicted ECF transporter
MIEVENLSFTYAGQKKPALKIPHLEIAAGRCVVLTGISGSGKSTFLRVLNGLIPEYFPGNL
QGKVTLNQQDLTQQSVEEIAAQVGSVFQNPRRQFFFAKVSDELAFPAENQGQAPERI
WQQLRVTEQELGLAGLLSAKLAELSGGQKQRVALAAAQMQRPQILVLDEPTANLDKQ
GVALLQANLAKLKAQGLTIIVAEHRLDYLREVADDYYYFTAGEIKRSWTKAEWLKLTDR
QRLAYGLRPLQVETFSKARLQVTGQNTLALKGFTLKWHQQELGWLKEFSLTSGIIYGL
TGANGLGKTSLAKVLVGLKKARSHLTWNGRALTQRALLAKSSLLMQDVSQQLFLPTVQ
AELNASGATEQEVTAMAAKLDLSELLASHPFSLSGGQQQRVALGRALLAQKELFILDE
PTSGLDEQNMRRVADLLLELKASRRIVMVISHDQELINLTCDEVLNLSDFLTKK

>Glutamyl-tRNA synthetase (EC 6.1.1.17) @ Glutamyl-tRNA(Gln) synthetase (EC 6.1.1.24)
MAKNKIRVRYAPSPTGHLHIGNARTALFNYLFARRNKGTFVIRIEDTDTKRNIADGERS
QLDNLTWLGMDWDEGPDKPGNYGPYRQSERKEIYAPLIQELLDKDLAYESYMTEEEL
AAQREEQKANGEAPRYVYEYEGMTKEEIAAKQAEAKAKGLKPVVRIRVPHNKVYEWD
DMVKGHISFKSDTIGGDFVIQKRDGMPTYNFAVVVDDHLMEISHVLRGDDHVANTPKQ
LVVYEAFGWQAPEFGHMSLIINTETGKKLSKRDESVLQFIEQYRSLGYLPEAMFNFITLL
GWSPVGEDEIYSHQEFIKMFDPKRLSKSPAAFDGKKLEWINNQYVKAASEDEIMHSAL
LQLIKAGKVEADPEPLKLEWARKLISLYKHQMSYTGQIVDMAEVFFEEPPALSQAAKEE
LADDSAKVVLEAFAKKAKALPLFDGPAIMAAIKEIQKETKVKGRKLYMPIRIATTRETHG
PELGPSIELLGRKKALAHLEQTLAEMN

>Cysteinyl-tRNA synthetase (EC 6.1.1.16)
MGKLGGLEMIQLYNTLTKQKEEFKPQVPGQVLMYVCGPTVYNYIHIGNARSAIAFDTVR
RYFEYRGYQVKYVSNFTDVDDKIIKASQEMNLSVEEVTAKFIAAFYEDTAALNVKKASL
NPRVMENMADIINFVADLIAKGYAYEVQGDVYYRARKFKHYGQLSGQSIDALEQGASE
RLSVEDQAKKEDPLDFALWKAAKPGEISWQSPWGLGRPGWHIECSVMATKYLGKTIDI
HAGGQDLEFPHHENEIAQSEAETGQRFANYWLHNGFVTIGEDDEKMSKSLGNFVTVH
DLVKQVDPQVIRFFMATTQYRRPIRYSQANLTEAANNLAKLKTAYDNLTFRLKDATEAE
LEAEVKQEFAGYEQQFVAAMDDDFNAQNGIAAVYEMAKQLNIYAEKEKVVKETITYLLA
NFTKVLAVFGIEFSENDVKPDAEIEALIAERDAARAQKNFARSDEIRDQLKEQGIILEDTP
QGTRWKRS

>COG1939: Ribonuclease III family protein
MIETDYQQMNGIALAYMGDAIYEVYVRQHLLAKGMTKPTKLHHRATHYVSAKAQAFLIE
KMLEQELLEKEELDYFKRGRNSKSYTSAKNTSVVTYRISTGFEALFGYLYLSGQKERLD
YLAQWCIETVEREGK

>23S rRNA (guanosine-2'-O-) -methyltransferase rlmB (EC 2.1.1.-)

MVDKKTEDFVIGRHPAKAALMSGQTINKVFIQAGLKADALNEIIKLAKKQNLIISQVPKSK
LDLLADRQNHQGVVLAVAAYEYAELDDLFAAAKAKDEAPFFLILDGIEDPHNLGSIMRT
ADAAGVHGIIIPKRRAVQLTATVAKTSTGAIEYVPVVRVTNLAQTIAELKERGLWIFGTD
MAGQDYRTWDANGPTALVIGNEGKGISPLIKKACDGMLTIPMVGHVQSLNASVAASLLI
YQGFSSRGK
>Hypothetical protein DUF901, similar to C-terminal domain of ribosome protection-type Tc-resistance proteins
MKKQTLIVDAYNMIGNWPQLAKLKKADQLEAARDELLRILSNYHKQSNSDIIVVFDAMF
VPGLSKSYKQFNLEVVWTEEGQTADSYIEALAGKMQSTLNQVTVATSDQAEQWTVFS
QGALRKPAWELYQDIKRANKEVHATARRYQDENTVRHSPFDEQQLEQLVALRDQLSQ
HD
>LSU ribosomal protein L33p @ LSU ribosomal protein L33p, zinc-dependent
MAKKKVALACSECGSRNYTLTENPNRQERLEVKKFCKYCNKHTIHRETK
>Preprotein translocase subunit SecE (TC 3.A.5.1.1)
MKFFKSVVQTMKETTWLNAKETRRDTTTVVIMSILFVAFFAVIDYLIQLALNLL
>6-phospho-beta-glucosidase (EC 3.2.1.86)
MTFNEINASTWGFTGTGALDDELSDQEAMQLRYQALHHQFLASAQAVKIGHQIDPEN
QIGCMIARMQTYPKTCNPADVRAAQLQDQLNLFFTDIQVRGEYPNYMNRYFAENNIKL
QMAENDQALLKEGCVDCLGFSYYMSTVTSAASQVEKASGNFALGEANPYLEASDWG
WQIDPVGLRVTLNEYWDRYQVPLFIVENGLGALDKLEADGSIHDSYRIDYLRKHIIQMQ
EAIKDGVQLIGYTMWGPIDLISASTSEMSKRYGFIYVDQDDDGNGSLKRYKKDSFYWY
KKVIASNGQDLSD
>transposase
MEIQITIKVKLNLANAEIASSFTNTMEQYRLACNYVSEYIFNHDFDMQQSRLNKVLYTNL
RSLFMLKSQMAQSVIRTVIARYKTVKTQMKRTPYKYQDINTGEWYREIRDLTWLQKPIT
FNRPQVDLQRNRDWSYLSSGQLSINTLNGRVKVNPICHGFDQYFEGTWKFGLAKLLK
TGGKWYLHISATKTVADFDNQTVKHVVGIDRGLRFLATTYDEQDQTAFFDGKTVMRKR
AKYQKLRAKLQAKGTKSAKRRLKKLSGRENRWMTDVNHRLSKTLVQKFGSNTLFVLE
DLTGVSFERTDLPKALRNQNKSWAFYQLAQFLTYKAHLNNSEVIEVSAQYTSQRCPKC
GTIKKSHRNHDLHEYHCSNCGYRSNDDRLGAMNIQLLGTQYISGQEQPKFELTTNA
>hypothetical protein
MAYEQLTKDILAGVGGKENINSVVHCTTRLRFKLKGQLPLTKVRGL
>Choloylglycine hydrolase (EC 3.5.1.24)
MCTSIFVSATDKSYLLGRTMDWYDLYVSPIFVPRNYQWQSIYNHHPYQNKYAVVGTGF
QDQNLIDLSDGINEWGLSVQKLTFANGTSLTTEKNDANIQLAAFEFVLYTLGNFKSITDL
EKQLPTIQLMALKSQLKHGGDELHFAVSDASGRSIVIEPSHSPLKVIENPLGIVTNMPSF
EKQLAKMDSYLDFTPEFLAGNSPLGAFHVTTGKQAGKKQPSGGYTPSQRFVKAAWLK
EMVDVPANGLEAADLAFKLLDAVSVPKSKAHRPTYTVYQTVTSAQSLSYYFKPVGQAQ
IFGVSLKSNLLSRPTVKVFQLPSTWQATMLF
>Transcription antitermination protein NusG
MEKIKGEETMVESNEKKWYVLHTYAGYENKVKENLESRAQSMGMEENIFRVVVPEEE
KVEKAKNGKEKVEMLKTFPGYVLVEMVMTDQSWYVVRNTPGVTGFVGSHGAGSKPA
PLLDEEIQHILHELGMSSRHADLDLEVGESVKIVAGPFTGLVGKITEIDTEKMKLTVDVD
MFGRETSTELDYDQAEKLA
>Cell surface hydrolase, membrane-bound

MLRRKNLLVGGLLALGALFLRKFEAWAKPDIRPQVRISNQVPTLFIHGYNGNRFSFGPL
IERFGKYGWGKKTAVIVVTPDGSMKISGDPSLPGGLVQVLFVKNRFPLTKQSDWLYRL
TYVLKQNYHLKSLNIVAHSQGGVTVLNYLSEHANDIMLTQVVKVVTLGAPFNDFETGKT
TKVLERNAPEILTPLFKKLKKSRWRYQNQIEFLNIAGDKDDGSFSDGQVALSSVLTLAN
LIDDSVRNYQEYVVKGKKAGHNYLLRNQQVDQVIGNFLFDLK
>Cell division topological determinant MinJ
MLFLKLVATYLVEPLFLLGLVASYILYLTRIKRERHFFRIAVNRDFYELRHFLKHGLLLGIS
GSIVTLGLGLTLGWQVVVIYQALACLALVAGIQANLAFLALLLTGLLAGGVAYLNPSWW
SFDKIVPSQLTGVMLLVVLYLAFKLVLLRAQKFDWFSPQIKAGKRGRRIATYFWREVSV
VPLVVLVPGSWLKVSGAWPVLTIGHQSYSLFIMPFLIAVSVKLAKQVPTVALGLYRKQT
WLLLGLASGLTVVSYFYPSASLVGFGSLALLWLYQAYRRKQADLRADFWYVETNQGV
RVVAVKPKTPAAKMQLVPGDIIIECNQQLVNSEAELYQALQANSAYCKLRVRGFNGEL
RLAESAIYAGAPHEIGLILFH
>Alkaline phosphatase synthesis transcriptional regulatory protein PhoP
MVDDEPAIVTLLKYNLENEGYQVTTCDNGQAAFELALANSFDFILLDLMLPGLDGIEVTK
RLRQEKVQTPIIMITAKNSEYDKIIGLELGADDYLTKPFSPREVIARLKAISRRLPAQEAG
QLKVEKERYHFADFTVDLANVAVVVHGRKVKLTPKEFELLAYFVKRPNRVLSRDKLLS
GVWGYDYAGQTRLVDMHVSHLREKIEPNPKQPQYIQTIRGFGYRFTGEVNET
>Phosphate regulon sensor protein PhoR (SphS) (EC 2.7.13.3)
MKLKNYVLLYCLTCLGLVVSVIGYDWYEGSNFDSQWHSLGLYSLVVSFVLTSLYFGYD
LWQQKRQAAEIKLLNDKLVDLLKEGAYHHVILKPSDPYYQLAQSINAIQSKQKRFVKNF
VRQQRGYFSLLEYLTSGVMLIDQDNLIYTANHALNEIFAKDLNQKQSPYYYVLDNFELIQ
LIEQAIRTKSDQAKEIKHHQDGEKLLAVQVRYIPLSKHHFLIMCLVDDVTQKRKLEQQQK
DFLANLSHELKTPITSISGFSETLLAGALSDQTVSRQFVEIIHRESKELAELIEDGLSLARI
EDDRALNLTSFNLYAFCNASLDSFTKLIFEKQIKVKLEVEPALEIVCDQRKLRHILNNLLQ
NGLRYNRPEGKLTIKANLLANEWQLIVSDTGIGIEKAEQKRIFERFYRVDQSRTKAIAGT
GLGLAVVKEYVELLSGDISLVSKPNEGSTFTVTLPRFEL
>hypothetical protein
MAKKLYRSHNRLIAGVCAGLAEYFNIKVKTVRLSFVVATLICGILPHLSVVPVGIYIILALV
MPLNLQQEQFSYFDLFKLFTGQTAKPNQTNSEYTKSGRKIIKDARERDIH
>membrane protein
MGFWQRAAVNALTFIALAGFFPEIFHVRSIWMAFVAAVVLGVLNMFVKPFLVLISLPITF
LTLGFFYWVINALMLEMTSTFVGNAFEFANFSSALIVALILSVVNLIITNYLAR
>HPr kinase/phosphorylase (EC 2.7.1.-) (EC 2.7.4.-)
MNKVSVAELVDANGLIVYSGGMYLKERFITTSDISRPGLELTGYFDYYPEERVQLMGM
TEISYAHKMEPAARKEIMERMCGENTPCFVISRYLSVPEEMEQAAKEHQIPILRSKLPTT
RLSSKLTDYLESKLAERKSLHGVLVDIYGLGVLITGDSGVGKSETALDLIKRGHRLIADD
RVEVYQQDEQTIIGEAPKILQHLLEIRGVGIIDVMNLFGAGAVRTSTDISLIVHLENWSKD
KQFDRLGSGEQSVQIFDVAIPKITIPVKVGRNMAIIIEAAAMNYRAKTMGYDATKAFERN
LNSLIQENSQD
>Prolipoprotein diacylglyceryl transferase (EC 2.4.99.-)
MLSAINPIALKLGPLSVHWYGVIIASAVIIAVTLAVIEGKNLGLAEDDFYDFLLWALPVALIT
ARSYYVIFQWPYYASHPSEIYRIWDGGIAIYGALIGALIVLIIFCRQKQISAWLFLDVIAPTV
ILAQGIGRWGNFINQEAHGGVTSLSFLHQLHLPAFIINQMYINGNYYQPTFLYESLWDVL
GFLILISLRHRQGLFKQGEVFLTYVLWYSFGRFFIEGMRTDSLMLGGLRVSQWLSVVLF
IVALGIWAYRRYYNPLNAWYVEGSKQDTLQKTK

>Thioredoxin reductase (EC 1.8.1.9)
MAKEYDVIVIGAGPGGMTAALYASRANLKVLMLDRGIYGGQMNNTAEVENYPGYKSIL
GPDLAEKMYESTIQFGAEYAYGNVESVELIADNLKLVKTDSDEYQAPVVIIATGSQHRK
LGVPGEEEYGGRGVSYCAVCDGAFFRDKKVAVVGGGDSAIEEGIYLTQLAQNVNVIHR
RDQLRAQAIIQQRAFANDKMTFTWDTVVKEIKGDGQKVTGLLVHNKKTGEDSEIEASG
VFIYVGIEPMTEAFQNLGITDEAGWIETNDRMETKIPGIYALGDVRKKNLRQITTAVGDG
GIAGQEAFNYLQTIK
>Phosphoglucosamine mutase (EC 5.4.2.10) / Phosphomannomutase (EC 5.4.2.8)
MSWKETYNTWKDYADLDAELKDELANLAGNDAALEDAFYTPMEFGTAGMRGLLGAGI
NRMNIYTVRQATAGLAAFMLTLDEATRNRGVAISYDSRHHSQDFAFEAARVLGAQGIP
TFVFESLRPTPELSFTVRHLHTYAGIMITASHNPKEYNGYKIYGEDGAQMPPKESDLITK
YIREVDDLFAIKVASKEELIDAGVLKVIGAEVDQDYLDQVKTVTINKELVAEEGKTMKLIF
SPLHGTGAMLGEKALRQAGFENFTMVPEQSKPDPDFTTVKKPNPEDPAAFNLAIKLGK
KEGADLLVAVDPDADRLGAAVRQPNGDYQLLTGNQIAAVMLNYILTAHKEAGSLPANA
AAVKSIVSSEFATKVAASYGVEMINVLTGFKFIAEQIKHFEDTNEHTFMMGFEESYGYL
VRPFVRDKDAIQSLVMLAEVAAYYKKQGKNLYDGLQELFEKHGYYAEKTISLTFDGVE
GAKEIADLMTKFRNDAPSEFAGYEVAALEDFAKSEKRLADGTVTKIDLPTSNVLKYILTD
GTWIAVRPSGTEPKIKFYIGTQADSLAKAEEKRAAFEDAITAFVKG
>ABC transporter ATP-binding protein
MATIELVNISKTFGTGLGQVTALKNVNFKADKGQLILILGPSGSGKSTFLTILGNLQTPSS
GQVKLNGLDLAQLSSRQKEKIRLEKIGFVLQAYNLVPYLKVKDQFNFVSRIKKSNNLSS
AKLNALLAKLEIADLVDKYPANLSGGQNQRVAIARALYTNPEIILADEPTAALDSHRVEE
VGLMLHDLAQKENKAVIVVTHDLRLRKHADSIYQITDGILTKEA
>ABC transporter permease component
MFLALKEFRYEKLRYSLIVVMFLLISYLIFILTGLAFGLANQNTNAINNWHFKQIVLNQDA
NVNLNQSLLTKEMVATLPQNAHTALIGQASIVVKKPHHANVGASFIGLKPQQFIYRNLKL
TAGKMPTNSHQLVVDENFKNSGYHLNDTVTLNSLKDKFKIVGFVKNHQYNVAPVAYG
QLTAWQNLKNVGPNFTASALLSKEKLHFKATNLKTYSVTQLINKLPGYSAQNLTFALMI
GFLLVISLIIIAVFLYIITIQKLPNYAVLRAQGIPAKTLVATTIIQAMLLALSGTILGGLLTWLS
ALFIPKAAPIFFDPVALGLVAAGLLFASLLGSLLPVRTILKVDPVTVINS
>Transcriptional regulator, TetR family
MVKSTFLNLPAEKQACITQALLHEFSRVPLATAQVAPIIKQAQIARGAFYKYFTDLTDAY
QYLYQLALTDIHQDLNFSKALTAKDYILLITNFLSGTKNSPYYDFIRLSVTQNDYFLRLHS
PMKQLASKDWAVATLCHEAIFACFFEPQHQELYLARLEEALTTFLKGV
>Excinuclease ABC subunit B
MIERQADRKFELVSDYQPTGDQPEAIEELTAGLKAGKKAQILLGATGTGKTFTISNVIKN
VNKPTLVLAHNKTLAGQLYGEFKEFFPNNAVEYFVSYYDYYQPEAYVPSSDTYIEKDS
SINDEIDKLRHSATSSLLERNDVIVVASVSSIFGLGDPTEYREHTLSLRVGQELSRDKLL
RSLVDIQFERNDIDFQRGRFRVRGDVVEIFPASRDDKALRVEFFGDEIDRIREVDPLTG
EVYGQREHVAIFPATHFMTNEDRLKVAIDGIESELTERVAELTAEEKLLEAQRLKQRTT
YDIEMLREMGYCSGIENYSRHMDGRRPGQPPYTLLDFFPKDFLIVVDESHVTMPQVR
GMYNGDRSRKQMLVDYGFRLPSALDNRPLQLSEFEKHVNQIIYMSATPGPYEMEQTA
TVVQQIIRPTGLLDPTIDVRPSMGQMDDLVAEIQKRIEADERVFVTTLTKKMAEDLTDYF
KELSIKVRYLHSDIKTLERTRLIRDLRLGKFDVLVGINLLREGIDVPEVSLIAILDADKEGFL
RNERSLIQTIGRAARNAHGHVIMYADNVTESMQKAIAETKRRRSIQEAYNEAHHIVPKTI

KKEIRANISASEKVAAKSKEAELDLDFVDMNKKDQQALIASLEAQMKQAAKQLDFEEAA
TLRDTILELKAQLD
>Excinuclease ABC subunit A
MAQNKIIVHGARAHNLKNIDVEIPKDKLVVVTGLSGSGKSSLAFDTLYAEGQRRYVESL
SAYARQFLGQMDKPDVDSIDGLSPAISIDQKTTSKNPRSTVGTVTEINDYLRLLWARVG
KPVCPNDGHEITSQSPEQMVNKVLALPERTKLMILAPIIRAKKGQHKKVLEKIKAQGFVR
VIIDGQMHDVEEQLELDKNQQHTIDVVVDRIVIKDGIRSRLFDSFEAALHLSDGYASVQII
DGELLQFSENFACAYCGFTIGELEPRLFSFNAPFGACPDCDGLGMKLEVDEDLVVPDR
TKTLREGAMAPWNPISSKYYPSLLEQACQAFGIDMDTPFEKLAKKDQQLVLYGDDSRV
FHFTFESDFGGLRDLDTKFEGVIPNVERRYRETNSDFTRDVMRQYMTSLTCQTCHGY
RLNKQALAVKIAGQHIGQVSDYAIGDSLKFFSKLQFSAQDEQIAKPILKEICDRLTFLENV
GLDYLTLSRASRTLSGGEAQRIRLATQIGSNLSGVMYILDEPSIGLHQRDNDRLIASLKK
MCELGNTLIVVEHDEDTMRAADYLIDIGPGAGANGGQIMAAGTPKQVARSKKSLTGQY
LAGKKFIPIPDRRRSGNGKFVKIYGAQENNLKKLDVSFPLGEFIAVTGVSGSGKSTLVN
MILERALAQKLNRSSEKPGKFTKITGTEALEKVINIDQSPIGRTPRSNPATYTGVFDDIRT
LFAQTNQAKLRGYGKGRFSFNVKGGRCEACKGDGIIKIEMNFLPDVYVPCEVCHGSRY
NSETLEVEYKGKNIAQILEMTVEEALEFFEAIPKIRRKLQTIVDVGLGYVALGQPATTLSG
GEAQRMKLASELHRKSNGKTFYILDEPTTGLHSDDIKRLLEVLQRLVSKGNTVLVIEHN
LDVIKTADYLIDLGPEGGAGGGTIVACGSPEEIVKVSASYTGKYLKPILEKGHYFTPTDE
Q
>S-ribosylhomocysteine lyase (EC 4.4.1.21) / Autoinducer-2 production protein LuxS
MAEVESFTLDHTAVKAPYVRLITVEQGKHGDEISNFDLRLVQPNENAIPTAGLHTIEHML
AGYLRDHMEGVIDCSPFGCRTGFHLIMWGKHDTVEVAKALKASLNDILNASWEDVQG
TDIKSCGNYRDHSLFSAQEWCKKILAEGISSDPFERKVI
>Hypothetical ATP-binding protein UPF0042, contains P-loop
MTETQSLVIVTGMSGAGKTVAMQSFEDLGYFCVDNMPPSLLPKFWDLVKESGKVKKV
ALVIDLRSRAFYDEIGEMLSEMDETDFVHKKILFLDSSDEELVARYKETRRAHPLAMEG
RILDGVHLERELLATIKAQAQLVIDTSKLSPRQLREEIFHNFETNDVKSFHVEVMSFGFK
YGLPIDADIVMDVRFLPNPYYVAELRNKTGLDQAVSEYVMNSDTTEEFYQSFIKVLETV
LPGYKREGKSSVTIAIGCTGGQHRSVALAERIGKALAANYPVRISHRDMNKRKESVNR
S
>FIG002813: LPPG:FO 2-phospho-L-lactate transferase like, CofD-like
MTKFTSKRKRPKIVVIGGGTGLPVILNSLHKENADITAVVTVADDGGSSGIIRDYINVVPP
GDIRNVLVSLSNLPQVYKDIFQYRFESKDQFFSGHAIGNLVIAALSEMKQGDIFSSVQEL
SEMMKVDGHVYPAANIPLVLNARFSDGAVISGEAEITAAGKTITDVWVTTGDDSLPEAR
PEVIQAIMEADQIVLGPGSLFTSILPNLMISNVGQAVCETQAEVVYICNIMTQKGETERF
SDADHVRVLNQHLNKKFINTVLVNTEPVPTDYMDRQKYDEYLYQVHYDFKGLRDQGC
RVISDNFLLLRDHGVFHDGDKVVEELMRLIGQLPV
>FIG001886: Cytoplasmic hypothetical protein
MSYASEVKKELTMLEVHFGNAKAELMALIRMNGSLGIYNHKFVLNVQTENPAIARRIYK
LLKQFYDIDSELLVRRKMKLKKNNLYIVRLKTGSDYVLKDLDILDGFQIKETVPLKFLDDD
AKVRSYLRGAFLAAGSVNNPETSRYHLEIYSLYEDHNNTICEMMNRYGLNARKTERRS
GYITYLKEAEKIADFLSLIGATNSMLKFEDIRIVRDMRNSVNRIVNCETANMNKVADAAR
KQIENIDYIEQTVGLSKLPDKLQAVAQARKAHPEVSLKELGELVAGGPISKSGVNHRLR
KINEFAQKLQVN
>ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92)

MNLVPTVIEQSSQGERAYDIYSRLLKDRIIMVSGEVNDDMANAIIAQLLFLDAQDPEKDI
YMYINSPGGSVSAGLAIYDTMNFVNADVQTIVMGMAASMASVLATAGTKGKRFALPNS
EIMIHQPLGGAQGQSTEIQIAAEHILKTRKLINQILADGSGQDIETINKDTERDNFMTAQQ
AVDYGLIDGIMKNKKKA
>hypothetical protein
MLAFFINLDIIPYKTLNKKLGLKKTRTKLARVFFNLKLKLFSCFS
>tRNA-Arg-CCG

>RNA polymerase sigma-54 factor RpoN
MALEQGFSQEQRQVQKLAMTQRMQQSIQVLRYNAEDLQAYLKQQELDNPFISVNLAT
NYNSSQAGAESKDDWQTYTSVRKQQSLFEYLLDQIHLTMRATPIRKWVIYLLEHLDSN
GYLRLDLKTVSEEQGVSEITLLDALTLLQRLDPPGVGARNLQECLLLQIENDAHAPKLA
QQVIKNEFEAFADRKWEQIAKKLGVRLAEIQTVFDYVRTLSPAPGAAYDQDQVGYIIPD
LILIETNQQLELKTTKYTQPQLEFKQSYYDRLAKQDDQEVKKYLAQKQKEYEHLYNDLL
QRGQTLLRVGQEIVDYQRDFFIKDEHPLKPLLLRDVAHKLQLHESTISRAVNGKYMQTS
FGTFELKHFFSQAVNYRTETGEAVSSDDVHKYISELVANEDKRKPLSDQKLSELLKQK
ELTVSRRTVAKYREQLNIPSSSKRKRFD
>hypothetical protein
MKGSKFKVKFGKIGKCLITLSIYQFFYEIKAVFGLQNISFVAIILIVIEVW
>Central glycolytic genes regulator
MHQELEWIELIAPDMVDVMSQRFIILRNISWAQPVGRRSLAQTLGMSERVLRTETDFLK
KQELISVTKSGMILTAKGQETIYGLTGLMDQLLGLQQMERRLSKALHIKRCLIVSGDSDK
QPQVIDNMGKLVNDALKMLLPEGTNTIAVMGGTTMAQVANCLSPDLSNQRQLTFVPA
RGGIGERVEIQANNVCAQMALKTGGKHRTLYVPEQVSEEAYRPLLKEPAIQEVVSLIRS
SNVAIHSIGEAIHMAKRRAMSPEIIAMLTKQKAVGEAFGYFFDENGKIIYRIPRIGLQLED
LASMDCVFAVAGGSSKAKAIRAYMQLAPSNTCLITDEGAAKLILKE
>NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
MTTKVGINGFGRIGRLAFRRIAEVSNDLEVVAINDLTSPAMLAHLLKYDTAHGQFDAEIS
ATENALVVNGKEIPVYAEADARNIPWVKNNGVDLVLESTGFYTSEEKASAHLEAGAKK
VLISAPAGKVKTIVYGVNDDTLTADDKIVSAASCTTNCLAPLAKAVNDAFGIKAGTMTTI
HAYTATQKLQDGPDRGGNVRNARAAAQNTIPHSTGAAKAIGLVVPELNGKLDGHAQR
VPVITGSLTELVATLDKTVTAEEVNEAVKKVTEGNASFGYNADGIVSSDIIGTEFGSVFD
PTQTQVVGEGEGQVVKTVAWYDNESGFTAQMIRTLEKFASML
>Phosphoglycerate kinase (EC 2.7.2.3)
MAKLIVSDLDVKDKKVLIRVDFNVPIKDGVIGNDNRIVAALPTIKYVIENGGKAILFSHLGR
VKTEDDKKKLSLRPVAERLADLLGKPVTFVPVTRGAQLEEAISELNDGDVLLVENTRFE
DLDGKKESGNDPELGKYWASLGDVFVNDAFGTAHRSHASNVGIASNMTQAAAGFLLE
KEIKFLGEAVDEPKRPFVAILGGAKVSDKIGVINHLLGKADKVIVGGGMTYTFYAAKGIKI
GNSLVEEDKIELAKELIEKAGDKLVLPIDNVVAEKFDNDVPNQVVEGDIPDGYMALDVG
PKSVALFEDVLKDAKTVVWNGPMGVFEMSNYAKGTLEVGKFLGELSDATTIVGGGDS
TAAVQQLGVADKLTHISTGGGASLEYLEGKTLPGIAAVSDK
>Triosephosphate isomerase (EC 5.3.1.1)
MRTPIIAGNWKMNMNPVQTAEFVKAVKDQLPASSEVESVIAAPAVDLPALLEAAKGSE
LKVAAENCYFEDEGAFTGETSPKVLKEMGVDYVVIGHSERRDYFHETDEDINKKAHAIF
KNGLTPIICCGESLETREAGKAEEWVEAQVTAALKDLTADQVASLVIAYEPIWAIGTGKT

ATADQAQEICAVVRQTVAKLYNQETADNVRIQYGGSVKPANVKELMAKEDIDGGLVGG
ASLVPDSYLALVNYKN
>Enolase (EC 4.2.1.11)
MSAITEIYAREVLDSRGNPTVEAEVYTEAGAFGRGIVPSGASTGEHEAVELRDGDKSR
YQGKGVLKAVANVNDVIAKEIVGYEVTDQVAIDKAMIALDGTPNKGKLGANAILAVSIAV
ARAAADELEMPLYNYLGGFNAHVLPTPMMNVINGGAHSDNKVDFQEFMIMPVGAKSV
SEAIRMGSETFHNLKSLLGADGKVTSVGDEGGFAPDFANNEEPLQYLIKAIEAAGYKPG
KDIAIAIDVASSELWDTETGKYKLRWSTGEEFTTEEFINYLEGLTNKYPIISIEDPIDENN
WDDWKAITDKLGKKVQLVGDDFFVTNTQYLQKGIKMGCANSILIKVNQIGTLTETMEAI
EMAKEAGYTAIVSHRSGETEDTTIADLVVATNAGQIKTGSMSRTDRLAKYNQLMRIEDQ
LTSTVAEYKGIHSFYNISEEARQNVINK
>Voltage-gated chloride channel family protein
MMEETHKKLLFDVTRLRFILKGIVVGILVGIVVSLFRLAIEKLILVSKWGFAQIHNQPWLV
LVWLLVAVLISFIVGKLIQGDIDIMGSGIPQVEGQLAGELEMNWWSVLWRKWLGGVLMI
GSGLFLGREGPSIQLGSAVGQGFAEKTCHEGTGRRVLIAGGAAAGLSAAFNAPIAGTL
FILEEVYHNFSPLVWTTSLAAALTANAVSMAFFGLTPVLHMTYRTSLPLKYYPCLLLMG
VVLGLLGYLYQYLTLNVNKVYAKLRWLPRQFDSLIALALLLPIGLLWPNTLGGGNSLIISL
GHAKLSLLVLFGLFVLRLSFSTISYGSGVPGGIFLPILTLGAVLGALIGKAFCLLGIFPDYLI
VNFIIYAMAGYFACISKAPFTAILLITEMVGSLSHLMPLAVVSIIAYTVVDVLRGQPIYEAM
LANLRPNPDQIHGTYEDRLEFPVLAGSKLQDRQVRDIKWPKDCLLISIRRGENEQIPHG
DTVIRSGDTLISLTNYEQRAKIRKQINFITQNLSAK
>Preprotein translocase subunit SecG (TC 3.A.5.1.1)
MYDTLLTLLIIVSVLIIIAVMMQPAKTDNASSALSGGAADLFSKQKPRGFEAFMQKVTAV
LGVIFFVLALALVYISSN
>3'-to-5' exoribonuclease RNase R
MENEKLKQLLMKRFEEEPDNEFTVKELVEYLHMTSTGQFKFVVKALAELEHDKLLTLN
QSGAFTLARSKQLPTYQGIFHSNDRGFGFVTVEELEDDFFINPTQTMAALNGDEVEVQ
MLTPGDAAKGMRPNGKVVQILGHALTQVVGAFNSAPDQNLPKGILGTIHLKDKKLANY
TFLLEENGLHPVDGEVVLCDITSYPDAHQPGILKGVAVKVIGNVNDPGMDILEIVYQHNI
PTEFPEEVMEQIAQIPDYVTPEEKVGRRDLTDQQLVTIDSIESKDLDDAVTAWKLPNGN
FHLGVHIADVSHYVRPKTPLDKEAFERGTSVYLTDRVIPMLPPKLSNGICSLNPKVERL
AMSCEMEITPEGQVVKHEIFPSVIKTTERMTYVAINKILESHDEKMMARYQDLVPMFET
MHELHKILVKMRKRRGAIDFEYPEAKIIVDEAGHPTDIQIRERGISERMVESFMLAANET
VAAHYNKLHVPFVYRIHENPEEDKIKNFFEMLSGLGIEVTGKADAVSPKMLQNILKKVA
GKPEEQMVQTMMLRSMQQAKYAPTPVGHFGLGAEDYTHFTSPIRRYPDLMVHRLIRY
YEDHGKNNDYQEKLPEITTHSSEMERRAIDAERDTDAMKKAEYMADHVGEEFDAVVS
SVTKFGMFIALPNTVEGLIHISELQGDFYVFLEKQLALVGRRTKRTFRIGQSVRVKLTNV
NVEQKEIDFSLLDADKAPQTNLLDGVELPAEPKRRFNRNRDDKRYNGNRKNDRGNKR
YGNNRNANRSNNGGKNKKPFYKGNKKANKFHR
>tmRNA-binding protein SmpB
MAKKNSPKDKPLAQNKKASHDYHILDTIEAGIVLTGTEIKSIRAARINLKDGFAQIKNGEA
WLMNVHISLYDQGNQFNHDPLRNRKLLLHKKQINQLAGEISKKGVTLVPLKVYIKNGFA
KVLLGLAQGKHDYDKRESIKKRDQEREIRRTLKNYR
>Acetyltransferase (GNAT family) SAS0976
MLIRPATSDDAPQIRALFRIILKQMELKKLRNLNQQRLNLALELSFKTKECRLSKGQIIVA
EVDGEVAGMAFGYPSQADGDIDELLCHYFPKVGLPAYESLYDNDDSQLTKTEWYLDS

LAVKPKFRGLGIATALLKRLPKIAARCGKTKVSLDVDLTNPKAKKLYRKLGFKQTDTLQI
GSHRYHHLMRAI
>FIG00742364: hypothetical protein
MRNLYTKRENLTGTGTTTVKDERGLACYFLTGRWGMHPAVLSVYASNGDLKAEIKQR
SLGFWPKFELFSDRKLVGSLRRYYGLGTEMLFVKKLNWLILGNSLTFNYKIYHGRRCVL
SLHEVELASGNYLELTIADINDEALCLCIVAILDFWAKGGLKGGSKENDPLLNLSYIH
>hydrolase, haloacid dehalogenase-like family
MIQIIASDMDGTLLNDEMVISDFNAAAIKEAQAKGVHFIVSTGRAYAEVKPLLDKVDLTC
PLITINGAQTFTAAGKLLNTVALPDQLARKICLKLKKKGLYFELVTSTGVYSDNKASRIAN
FAELLTRVNPDTPYKLAVTLASARLELMNINYIDDYFTLIDDPHMPILKIVAFSQEGRPVL
DPIRAELSANPALKVTSSSDSNIEINHINAQKGVALQAFADSLNIPMDNVMAIGDNNNDV
SMLEVAGVSYAMGNGSTEVKRIAKRLAKKNTDNGVGHAIKEVLAEN
>Uracil-DNA glycosylase, family 1
MKTFIHNDWQEVLGPQFESAYYAHLHNFLKSEYQHYTIHPDMYHIFEAFEWTPFSKVK
VVILGQDPYHGPNQAHGLSFSVLPGVKVPPSLQNIYKELQADLGYPPVNHGYLEKWAK
QGVLLLNSVLTVRDGQAFSHRNQGWEQLTDAAIAALSARPKPVVFILWGKAAQAKQAL
IDQSQNIIIKSAHPSPLSANRGFFGSRPFSKTNDALMAMGEKPIDWQLPAEVTAK
>Phosphate acetyltransferase (EC 2.3.1.8)
MDLFDSLKEKIRGKNISLVFPEGDDTRIVTAAVRLAKEKLVHPVVLGNKDEVKKLVESN
GYTLTDDVEILDPTTYPEADFEAMVQAFVERRKGKNTAEQARQMLRDVNYFGTMLVY
MKKVDGLVSGAKHPTGDTVRPALQIIKTHPGVSRTSGAFIMSRNDGKERYLFADCAINI
NPNAQELAEIAVESAKTARLFDIEPRVAMLSFSTKGSAKSDEVTKVVEATKLAHELAPN
EEIDGELQFDAAFVPVVAAQKAPHSEVAGKANVFVFPELQSGNIGYKMAQRFGGFDA
VGPILQGLNAPVSDLSRGCNSQEVFDVALLTAAQATV
>TsaE protein, required for threonylcarbamoyladenosine t(6)A37 formation in tRNA
MRIVATNPDQTQAVGKKLAALLQSGDLILLAGDLGAGKTTFTKGLALGLGIKRNIKSPTF
TLIREYKEGRLPLYHMDVYRLEDGGAEDLGLEEYFNGDGVSVVEWASFIPEELPAEYL
EIRLNKDETDDNKRYIDFNAQGKHYEKLVAKLAE
>L-O-lysylphosphatidylglycerol synthase (EC 2.3.2.3)
MKNFFVKAKLALEKRLTLLKIVFVLSVLVFVINAISRILKGTDWQLVSQQLMGRSLGQLA
LLLFGGLLAVSPMLIYDFSIVKFLPGKYSRAYIIKSGWITNTITNIAGFGGVLGATLRASFY
SKGASKKQIFYALSKIALFLLAGLSLLSWLALLLIYGFDLGEHLANYGLWLIGGALYFPLL
FIVTKFKNNGFFADLSLRQELTLTVGSTLEWLFVAGFFLLVGRILGVNHDLTAIFCLYIIAS
ILGIVSMLPGGLGSFDVFMLIGMATFGINRELALTWLLYFRIFYYIVPVMVGLAFLVRSM
GGRINDYFDGIPQTFLATLAHRILTIFLYFSGIFMLMESAVPSLTVSNKVLVRFYPYTFFF
LHQLTNICYAFILLGLARGIQAKVKKAYWPMLAVLLVGIINTLFQDWSLSLTIYLGIVLAITL
FSKKALYRQKFQYSLGSLAFDGVIFIGSFLLYVVIGVINAPAYSNKHKIPSVLFFPGQQL
WLSGLIGIGIAVIILYFIMQYLTTGRDPFADRDFESERVKQVIATFGGNETSHLAFLRDKN
FYFYQVDGVDQLFFMYRRKNDKLIIMGEPVGNQAYQEEALQSFLHQADLYGYELVFYE
VDKDYTMLLHEYGFDFLKTGEDGLVELDKFTLVGKKQRAQRALMNKFEREGYVFSLA
PNPLSDELLAEMKRVSDEWLNGQVEKGFSLGFFDEYYLNQAPVALVRDAEGTLVAFA
SMMPTGGKEILTIDLMRHSKDAPSGIMDEIFISLFKYGQEEGYTYFDLGMAPLSNVGQS
KFSFIEERVAHLIYQYGYKLYGFQGLRAYKNKYATGWHPKYTAYRKKSSLLSTMWQVI
MVVNQRADYRTANGGKKQLPFIPKFLQP
>hypothetical protein

MKKNWLQILNQGSILFVCLVIVDCLRGILLIKGNGNIETLLGIKLSLIRTASQSSFNLSGTW
TIVIYYGLCLGLVWLVTIIKSPKRG
>transcriptional regulator, MerR family
MADLTITQVAAKYNINPDTVRYYERVGLIPKVPRKENGNRYYPDSLQHWLEMLVCLRH
SGVSVEVLHDYTELILQGDVTLETRKELLQEQVTVLKQRQADIQRSIERLEHKISLCQLP
LTKVRGL
>transposase
MFNFTTLINLGLSLTLGSLEPTFYPFASLYCVFTEQKCYNISKESEVNLMEMQVIVKVKL
NLANAEIASSFTNTMEQYRLACNYVSEYIFNHDFDMKQSQLNKVLYTNLRSLFMLKSQ
MAQSVIRTVIARYKTVKTQMIRNPYKYQDINTGEWYRETRDLTWLQKPIAFNRPQVDL
QRNRDWSYLSSGQLSINTLDGRVKVNPICHGFDQYFDGTWSFGLAKLLKTGGKWYLH
ISATKTVANFDNQTVKHVVGIDRGLRFLATTYDEQGQTAFFDGKAVMRKRAKYQKLRA
KLQAKGTKSAKRRLKKLSGQENRWMTDVNHQLSKTLVQKFGSNTLFVLEDLTGVSFE
RTDLPKALRNQNKSWAFYQLEQFLTYKAHLNNSEVIEVSAKYTSQRCPKCGTIKKSNR
NHDLHEYHCSNCGYRSNDDRLGAMNIQLLGTQYISGQEQPKFELTTNA
>ABC transporter, ATP-binding protein
MQILEVDKLSKAFKSKPVLKDVSFKVQAGEVVALVGKNGAGKSTLLKIILGLLTADSGQ
VTFLGKKRDFKEQIGVMLQEVPTIRRIKVKEIIQLWRSYYAKPLDYDELLKLADLATKEN
SYLTELSGGQKKRLYFALSLAGNPQILFLDEPTVAMDSLSRADFWKQVDSLRKEGKTFI
ITSHYPEELERIANRFLILAKHRLVFNGSLGELQSKQKACLVSFASDLKRAEIANWSGVV
RISEQNHHYFLQVIDLPAFLTDLLPKIRQVTELTIKPNGLAELLDNLGRNYDEEI
>ABC transporter permease protein
MKKFSYQLVLNFKRVVLRNPGFLVGAVLLPIVFYLVFTKALNQNAASEWNLNYLASMIV
YGILLGSVMTTTSVMIGDAQGGYNKLLRLTPISKWSFYLNMLVTFEVLNLLCILGISSVG
VIVNQVQLQLKAWLGLILIVPLASLPLILIGFIISRGCDYSLGTTVSNLIIFPLAIVSGLWLPI
GLLPKWMQAVGKVMPTYYLRQLTFNLINGQSLGSKNCLWIGGWLLVLLGLLVILTKFKT
KEV
>sensor histidine kinase
MGYPYLFIFTGFFIGSERLEKKSRLSLVGGYYLLAVGLYIGAMLTLKGGNLYEALINGLFII
LSLPVAYYSSERWRRHIQIVQDNERLTYIVKKAERERIAQELHDNLGQSFSVLALKAELA
KKLLVKDPILASQQLDDIAKTARADLDLVREIVAGLRKDLLVQTLAQESINLKLVRIGLKT
EGEKIALNWPEDIQELSAQVIKEVTTNCIRHSGARYLTIIFSQDDANYYLKASDDGQGFK
AKPFKETYGLSGIQKRVASFGGEVSFQNQSGAVIELSLPKKG
>DNA-binding response regulator, LuxR family
MIKLYLAEDQEMLNSALVMLLNLEDEFQVVGSCRDGRQALEEITRLQPQVVILDIEMPS
LSGLEIAKKLRAANQPLKIIILTTFARKKYFQEALAAKVNGYLLKDSPSDSLISAIYAILEGN
TVFDSKLVSGILNEVPNPLTKRELEVLRALKTTATTKELAASLYLSEGTVRNYISAILSKT
GTKSRLDAVNLAQAKGWL
>Ribonucleotide reductase of class Ib (aerobic), beta subunit (EC 1.17.4.1)
MENYTAINWNKLEDEIDKATWEKLTEQFWLDTRIPLSNDLADWRTLTEADKDVVAKVF
GGLTLLDTLQSQDGMASLKKDVRTQHEEAVLNNIEFMESVHAKSYSSIFSTLNTPAEID
EIFAWTNSNELLQYKAQKINDLYQHGNPLEKKVASVFLETFLFYSGFYTPLYWLGHNKL
PNVAEIIKLILRDESVHGTYIGYKFQLGFNELTEEQQEELKAWIYDLLYDLYENEEKYTAS
LYDQVGWTDDVLVFLRYNANKALMNLGLDPLFPDTAEDVNPVVMNGISTSTSNHDFFS
QVGNGYLLGNVEAMEDEDYLFRFDDED
>Ribonucleotide reductase of class Ib (aerobic), alpha subunit (EC 1.17.4.1)

MSLKTIDPNQVTYYDLNNEINLPLNGQIQLGKDKEALAAFLKENVIPNTMKFNSLKERFD
YLIANDYLESEFLAKYDFAFIEKLYAKLTAANFTFKTFMAAYKFYAQYALKTNDGAYYLE
SFIDRVAINALYFADGNEELALKLADELVNQRYQPATPSFLNAGRKRRGELVSCFLIQM
TDDMNSIGRGINSALQLSRIGGGVGITLSNLRGAGAPIKNIQGAASGVVPVMKLLEDSF
SYSNQLGQRQGAGAVYLNVFHPDIEAFLSTKKENADEKIRVKTLSLGLIVPDKYYELVA
NNEDMYLFDPWGVERVYGKLFSYVDITKEYDKMVANPAIPKKKIKARELETEISKLQQE
SGYPYVVNIDTANEASPIAGKIIMSNLCSEILQVQKPSIVNDRQEFDELGTDISCNLGSTN
IVNLMAAPDFGQSVETMVRALTFVTDHSDIQVVPSVQKGNAQAHTIGLGAMGLHAFLA
KNQIEYGSKEALDFTSVYFMLLNYWTLVASNKIARERHTTFVDFAKSKYADGTYFDKYL
ANDFGPKTAKVKALFEHVFIPTAADWAQLKEAVMTDGLYHQNRLAVAPNGSISYINDTT
ASLHPIINRIEERQERKIGKIYYPAPYLSNETIPYYKSAYDTSMLKVIDTYAAAQQHVDQG
MSLTLFMRSEIPAGLYPWKKAGNNKMTTRDLSILRNYAHSKGIKSIYYIRTFTDNNGEIG
VNECESCTI
>Ribonucleotide reduction protein NrdI
MKIAYFSVTGQTRRFIKKLALPADEVYEIDPTGAFKEFDEPFVLITPTYEDNITEPVRDFL
DTGANATYLKGIVGTGNRNFAQLFIFTARDLARDYQVPIIYAFEFNGTPKDVTNFKKAVK
KLESQND
>Glutaredoxin-like protein NrdH, required for reduction of Ribonucleotide reductase class Ib
MSLELYTKNNCIQCKMTKNFLMQHDIAFEERNVNENPEYIDYLKNKGFQSVPVLEQDN
NPIIRGFRPDLLKKIVAQ
>Ribosomal RNA small subunit methyltransferase C (EC 2.1.1.52)
MTNYYYTKNPDVEHNEQNFNFTLLGNELNFTTDNGVFSKRRVDFGSCVLLAALDQPA
FQPRKLLDVGCGYGPLGLAVAKKFPQAQVDMVDVNELALSLATKNAAANQIKNVNIWS
SNQYQTVTAKDYDAIITNPPIRAGKEVVHGILSQAKDHLVVGGTLTAVLQKKQGAPSAK
KKLEEVFGNCQVLKKDKGYFILQSIKEV
>tRNA-specific adenosine-34 deaminase (EC 3.5.4.-)
MDQEQKEYFMNEALKEAKLAFDAGEIPIGCVLVYQGKIIGRGHNLRQASQQAIDHAEIL
AISQANQQLASWRLAGCQLFVTLEPCPMCAGAILNARIDEVYYGAADLKAGVAGTLMN
LLNDGRFNHQAYIEKGILEDKCLSILQTFFKRVRKEK
>DNA polymerase III subunits gamma and tau (EC 2.7.7.7)
MSYQALYRVWRPQKFSELVGQEMVTKTLRHALMTHQTTHAYLFTGPRGTGKTSAAKI
FAKAINCHQQTDGEPCNECELCQAITAGRLNDVIEIDAASNNGVEEIRDIRDKAKYAPTV
ADYKVYIIDEVHMLSTGAFNALLKTLEEPPANVVFILATTEPHKIPATIISRTQRFDFRRIT
PQDILGRMEYILKQKEVSYEDSALKVIAKAAEGGMRDALSILDQVISFGEDGQVTFDNA
LLVTGTVTRQKLTAYLKQVFAQDTSAALETIHTIVNEGKDAKRLIEDLIGYCRDLLLYKQA
PKLVEESELGLLDADFKEFAQKIDSQVLYQAINILNEQQENMRYTSHETVYLEVLTIKLS
QSVSQGQAKAPVATENSADHEEVVRLAKQVAQLQEQLQTLSQASPKQVSVPAKRPKP
SQASSGQFKVNLNQLLPILNKAKKENLLQLREVWGDLLNMLNPAQRAMLRSFEPAAAS
DDGLILKGAYPYLCQRAQSSQELQDAIGNGLSRLLPQAPSLHFVPNENWQEIRQEFYR
QNQPQTEEKAPKKEEVAPIVLTAQKLFGKENVEIKAD
>FIG000557: hypothetical protein co-occurring with RecR
MMRGMGNMQNMVKQMQKLQKQMQKDQEELHSTEFTGQAPDEAVVVKFTGDRKMV
DIAIKEEAVDPDDVEMLQDLVIAAVNDAMDQIEAQTQKTMGKYSRNIPGL
>Recombination protein RecR

MQYPEPIAKLIDSYMKLPGIGAKSATRMAFYTIDMPADDVTDFAKALISAKRDLHYCSIC
GNITENDPCVICADKTRDQSKLIVVEQPKDVMSLEKMREYHGLYHVLHGVLSPIEGKGP
EDLNITSLVKRLQDNEAIQEVILATNATPEGEATAMYISRLLKPAGLKTTRLAHGLAVGS
DIEYADEMTLFKAIEGRTEI
>FIG015094: hypothetical protein
MFGRNKQKLRRTYDDLLLADVEQAKVDWDNAKLTQKSVYDADDELEAETKLAKAKYQ
LLFREARLRRIKGHLQATMIKVNEFNN
>Thymidylate kinase (EC 2.7.4.9)
MTGFLVTFEGTDGSGKTSVINGLTEQLEQLGYQDRYIVTREPGGNKISEAIRHILLDEEY
EGMDPKTEVLLYAAARRQHLVQTVLPALKAGKIVLCDRYVESSLVYQGVGRNIGIEPVL
AINNFATDELKANLTFVLDIDPEVGLKRIANNRLDEVNRMDKESLEFYQAIRRAYLKLVA
KEPHRLVAIDANQELAQVQSDVWRVFESKILATNN
>protein from nitrogen regulatory protein P-II (GLNB) family, ortholog YAAQ B. subtilis
MKLIIAIIQDKDSSKLSNALVEADIRATKLSTTGGFLRAGNTTFLIGVEEDKIETVLAIIKES
SQKREQYMAPPINLDSPMEATHTYPIEVQVGGATVFVLPVDSFYHF
>DNA polymerase III delta prime subunit (EC 2.7.7.7)
MEVEVRKTQTQLVEHFTKLVAKQQLVHAYLFSGPAGVGKMQLALYVAQAVFCQQKEA
GKPCLKCPECRRIAENNHPDVVRIIPEGLSIKVEQIRYLKTEFSKSGVEGKQKIFIIQDAD
KMTIGAANSLLKFLEEPSGQVTAFLLTSHLNQILPTIISRCQLVELSALTVKQLTAKLVEQ
GLPPDKAQLLAHLTGDLKLAKQLAEDDAFNQVIQSVWEWYRLILKNDLQAFVLVQTKL
MTYAETRAEQELLLQLLVLIVRDAMLLHYQRDDEVAFSMYKTELSQQILPLSATKLVAA
MELVLKSWPEMALNVAFQNILEELTLNLCNCYHQR
>DNA replication intiation control protein YabA
MDKREIYDSFEKMEQQTRTVSEMLSKIKEELTQVLEKNAELEIENQHLRERLQEIELPQ
HKEEPGLSKSRQNLEKLYKEGFHVCNDMYGSRRVNDEPCIFCQDVIYGERH
>rRNA small subunit methyltransferase I
MLIERTSSFKGVTTSGRLYLVPTPIGNLGDMSPRAVTTLQEVDLIAAEDTRNTAKLLNYF
EIKTPQISFHEHNTMERIPLLVDKLLAGQNLAQVSDAGMPSISDPGQELVKACIEKQIPV
VPLPGPNAALTALIASGITPQPFTFYGFLPRKNKDLKAELATLANERYTLIFYEAPHRLKK
TLKALLEVFGGDRQVVLCRELTKQYEEFLRGDLQAAYDFSSTEEVRGEFVLIVAGNSTL
NQAEQVEQDTNSSLKAQVEALVATGLKPNAAIKEVAKRNNLKKQVVYNDYHEID
>Acyl-ACP thioesterase
MAQEFSEQHRVVYYDTEVTGEMGIGKLVDMMMLASEDHSDSVGVTTQKVLGLGLGW
VITQHLLEVKRLPKTNEVVTVSTKAASYNPYFCYRDFWVEAADGELLAKMHTVFVLMD
QAKRKITRILPELVAPFESEYTKKIERNPLPMAPKQPTASHDYGVRFMDIDSNHHVNNV
HYFDWMLDVLGADFLTSHRLKKMNIQYKQEVRYGQVATSEVEIADLTSYHQIKVGQDL
SCVAQCEWKLR
>Internalin-like protein (LPXTG motif) Lmo2821 homolog
MFRDKSNKDLKFAIRKKKNGGGAASFIIGSVVLGGLMVSGITNASADTTTTPDSTSAST
TSAASSSATDKNEVTLNSTATPATASSEETSAAETTTAAPASEATPNSSVATTATDTSA
SSATATSEEAPVEEGKVVVKYVDESGNQVGMPKLLAGNVGESYDTTAVELPAGYALA
GSTSIENATGTYTKDPITVTYTVKKISNTEAATQALDRAAAVSANQLATNKVSISNVHFN
KSVVKESQGLDIDYAYDWTGKGLVKGDTLVSEMPAAFTSITKLVETPFYSGNDEIGVLV
LDYTNHKIYTRFTGEMDPNKIYNGSINIATFVDRNHFKEVENDEMVQLQMPDGTTVER
PLHVTFDAAQISPQLTLVSVYANESKDNPDKSMDVKWATTINKAGIDLNEAVVYLSPDT
VLGINPEFTKLGYDREGNLIGDPSSSYKAINPEYDGNALYKINKDSIEVYEANIHDSMGY

EVTKKLVAGTDYKLVEDSSSPNAYVIEFIGDYVKTNKSFVITYGGRVADSSSDRVAINQ
AVTDSLLAYYEGPYNKYTDKYAGPFNASWSEAAITVNNSSVEGGAQDITGSVTVVHID
ASTGKVLKTENYAKTTDGQVAHDAKQGTGYVTAPEEFAGYKYTTMGYQSAPAEGIVK
QGIQRVVYLYVPADKKGSVDVVHKTTDGQILADVKPVVTDVNVGTAYDTEKGNFNGY
HFVGMAEDSAAADGVVSEGIKHVIYLYEKDVTPEVKKGNVDVTYVAEDGTILEATSDVV
KDGEIGSNYETTQKNFDGYHFVRMGEFSADATGNVEEGTKHVVYAYAKNPETPVEKK
GSVDVKYITKDGEVLEDVASVKDNAPVGEDYTTEEKSFDGYHFIGMDKTSDPATGLVA
EGTKHVIYVYEKDVTPEVKKGNVDVTYVAEDGKVLEATSDVVKDGEIGSNYATEEKSF
EGYHFVRMGEFSADATGNVEEGTKHVVYVYAKNPETPVEKKGSVDVKYITKDGEVLE
DVASVKDNAPVGEDYTTEEKSFDGYHFIGMDKTSDPATGLVAEGTKHVIYVYEKDVTP
EVKKGNVDVTYVAEDGTILEATSDVVKDGEIGSNYETTQKSFDGYHFVRMGEFSADAS
GQVEEGTKHVVYVYAKNPEKPVEKKGSVDVKYITKDGEVLEDVSTVKDNAPVGEDYT
TEEKSFDGYHFVGMDKSSDPATGLVAEGTKHVIYVYEKDVTPEVKKGNVDVTYVAED
GTVLENTSDVVKNGEVGSTYATEEKSFDGYHFVRMGEFSADASGQVAEGTKHVVYVY
AKNPETPEVKKGSVDVKYITKDGEVLEDVTSVKDNAPVGEDYTTEEKSFDGYHFVGM
DKTSDPATGLVAEGTKHVIYVYEKDVTPEVKKGNVDVTYVAEDGKVLEATSDVVKDGE
IGSNYETTQKSFDGYHFVRMGEFSADATGNVAEGTKHVVYVYAKDPEKPEVKKGSVD
VKYITKDGEVLEDVASVKDNAPVGEDYTTEEKRFDGYHFVGMDKTSDPATGVVAEGT
KHVIYVYEKDVTPEVKKGSVDVKYITKDGEVLEDVASVKDNAPVGEDYATEEKSFDGY
HFVGMDKTSDPATGVVAEGTKHVIYVYEKDVTPEVKKGNVDVTYVAEDGTVLEATSDV
VKNGEVGSNYETTQKNFVGYHFVRMGELSAEAAGKVEEGTKHVVYVYAKNPATPEVK
KGNVDVTYVAEDGTVLENTSDVVKNGEVGSNYETTQKNFDGYHFVRMGEFSADKTG
QVAEGTKHVVYVYAKNPATPEVEKGSVDVVYVDEQGNVLPGGELASIKKDVPVGESY
TTDQKTFEGYTFNRMGTGSADAEGKVTEGVQHVVYVYTKNPEVTRGTVTVTYVDKNG
NPLPGGEKTTVKNDVPVGEDYTTEEKTFDGYHFVGMDKNSDPFEGKVTEGTKNVIYV
YEKDPVKKGSVDVKYVTVDGEVLSDTTPVVTDAEVGTDYSTLQLDFNGYKFVKMAEG
SAQANGQVVEGNQHVVYVYTRLINTLPSKDEETPTPTEAKKGSVDVKYVTTDGTVLQD
VTSVKENAEVGTNYETTALEFAGYKLVGVAKDSDSTSGQVTEGAKHVVYVYEKVTPVL
TDKHEDKETPSTSTQTEQKPVKEAPATPVAAKLVTKKQVASKQADTNKQALPQTGET
QSETNATVLGALALGLAGIFGLGAKKKEDRD
>cell wall surface anchor family protein
MFRDESEKQLKYAIRKKKTGGGAASFVIGSVIFGGMMLGTMTANADTVATNDTSNDAV
TNQTGTDSTADATTTLKATTTPAATASSEGNLESTAAAATNENTTSTKPVNTATTSEAT
SKVETASSTSEVASSQADSTSTASEEASSAVANTNDKLQTSETSTPQASGTDLTRRAI
AAENTNINVNNKVNITNVTFNKDEVVESEGTDINISFDWEAQGLHKGDTMVVPMFEGF
DSVTKEVKYNFSSDSLSNMGTLVLDYNEHKIYVEYTGDIDPNKIYRGTMNIGTFVSRNY
FKNEKNEAQINIPLPDGTSISKTLKVNFDQAKEWDDAYAAAINETLSEDKTRADVSWVV
EVNPKKKAMKSTAVYITPDNIDEVKPFYVDGKLQSASIYYTSAQNPYTLDPDSIMVYEA
NITPSLGYTKGKKLVLGTDYEILRGAYDANTRSQKDNVWIVRFKGKYDVINGDQFVIEY
NTHYDNLDPTLESTTGDRAVAIGTDVEANNLISYFDPNNSVMNFKGGIIYAKVRLIDSSM
TSVEGENITGSVIVAHINGKTGEYLKPEAYVINGGTTLKDVKQGTPYTTSPETFDGFTFT
RMGYDSAPATGTVNNSVQRVVYVYMPDELKGSVDVKYVDRATGEVLPFADAALTTVK
DNAPAGETYDTDKKDFAGYTFVGMTEDSAAADGNVEANKTLHVIYAYDKKPEPVVEK
GSVDVVYVDEQGNVLPGGELTDVKKDDPVGEKYTTEQKTFNGYMFKRMGTGSAAAN
GEVTKGVQHVIYVYTKNPEAPKTGSVDVKFVDKTTGEMITGTGVETVKDGAPVGEGYF
TTPKNLTKQGYQYVGIREGSDDPAGLVAEGTKHVVYEYEKIPEPIVKKGSVDVKYVDR

ETGEVLPFDEAALTTVKDNAPEGEGYNTTKKNFAGYTFTGLTEDSAAANGSVVADKTL
HVIYAYDKNPEPIVKKGSVDVKYVDRATGEVLPFDEAALTTVKDNAPAGEKYDTDKKD
FVGYTFVGMTEDSAAADGNVVADKTLHVIYAYDKNPEPVVEKKGSVDVKFVDKTTGE
MITGTGVETVKDGVPVGEGYFTTPKDLTKQGYQYVGIREGSDDPAGLVAEGTKHVVY
EYEKIPEPIVKKGSVDVKYVDRATGEVLPFDEAALTTVKDNAPEGEGYNTAKKNFAGYT
FVGLTEDSAAANGSVVADKTLHVIYAYDKNPEPVVEKKGSVDVKFVDKTTGEMITGTS
VETVKDGVPVGEGYFTTPKDLTKQGYQYVGTREGSDDPAGLVAEGTKHVVYEYEKIP
EPIVKKGSVDVKYVDRATGKVLPFDEAALTTVKDNAPEGEGYNTTKKNFAGYTFVGLT
EDSAAANGSVVADKTLHVIYAYDKNPEPVVEKKGSVDVKFVDKTTGEMITGTSVETVK
DGVPVGEGYFTTPKDLTKQGYQYVGTREGSDDPAGLVAEGTKHVVYEYEKIPEPIVKK
GSVDVKYVDRATGKVLPFDEAALTTVKDNAPEGEGYNTTKKNFAGYTFTGLTEDSAAA
NGSVVADKTLHVIYAYDKNPEPIVKKGSVDVVYVDEQGNVLPGGELTDVKKDAPVGEA
YTTEQKNFDGYTFSRMGTGSAGANGKVTEGVQHVIYVYTKNPEPVVEQKGSVDVKYV
TTDGKVLEDVSTVKDNAPVGEEYTTDEKSFDGYHFVGMDKTSDSANGKVTEGTKHVI
YVYEKDPETPAEVEKKGSVTVTYVDKDGNPLPGGEEKTVKDNAPVGEEYTTEQKDFP
GYHFVGMDKHSDPVDGKVTEGTKRVVYVYEKNPTPTVATDKEEEQKGSVTVTYVDKD
GNPLPGGEEKTVKHDVPVGEDYTTEEKDFPGYHFVGMDKTSDPANGKVTEGTKHVIY
VYEKDPTPTVATDKSEEKKGSVDVKYVDTEGNVLEGPNSVVKDGLVGSDYNTSGKTF
AGYKLIGMELGSASANGKVVEGTLHVIYLYQKVTADVPTAPTATSETPTVNKVQPSVTA
SVTPSAAQPKSEKQASTQVALPQTGETKQNSSMLGVIALGLAGMLGLGFKKEEKENK
>hypothetical protein
MTKILKREANYPITIPAKVRFDLRLTANEKLLYGELYTFGGMAGTFTVSNQRLAALYHVT
VRSISAWLNKLAKCGYIELANSLDKNERIISFVPGVDPLDRDGYLYDEEISHEKVVEAVA
EPSDKYDYTVDTKVDPEVIEAEMTERFGAIPDHVTPEQREQLLMVRESLYYKLHSKNR
QPYNQIQNVELRRSMFIDSEVKKEFNKLKNQGVKLTMDEEYVLINKVTARATEQFNKE
HGINCEKETSMEENFYPLEKKSSRVEANF
>hypothetical protein
MGKNTSRVENNFQSKLDNINSIVDKIIVGFADIIKNINKLSDLRTTKNLISNQSA
>hypothetical protein
MQAAEKWAQATITDNIILGWIFSEHPELLAKLLKICLPGFKVRQNSEG
>hypothetical protein
MRVNREEVTKKDLVYYGPRFDIYTEDEFGTIYDLEMQNSDKGDLAERMSMHQASLER
NALLAGEDFRKRRRTVVIFMYSRKDWLLITSVLRYSTV
>hypothetical protein
MNLALTIMATIPKLTAANKIHEEIITTLVAKYNIDQAQAESYYEQVMALRA
>Cation transport ATPase
MKKIYQVDIKQLATNYEQENFRAGLTSAQAHSKLTADGPNVLVSKQTPKWKIFLRQFN
NIVIYILLLSALLTLMIGHYTDAVVILAVVIINSLIGYFQETSAANALAKIKELLAQNATVYRD
HTRQDVAASDLVKGDVVFLEAGDNVPADLRIVEADNLRIEEASLTGETTAVTKTPLALS
ANNVPLAERTNMAFASTSVITGSGLGIVVATANETEIGQISQEVAQIKSSKTPLIKEIDGV
GKVVSYLSIIISLVVFIIGLMLKIYSLPALALAVVAMLVGAIPEGLPATTSVILAFGVSKLAK
EQKTIIKSMPAVETLGSVNVIATDKTGTLTKNEMTVTDLWVDDNQYQVTGSGYAPEGQ
LLQAGTKAQLTPNLELLLAAGYQANDTILKQTESGWTINGEATDGAFLTLYHKLKSSAP
TYHSLDLLPFDSDYRYIARLVENSQKEQLIFVKGSPDKLIPMAQKGNPNFDQALWTKRA
NAWSQKGKRVIALGYKSVTKQKQVTHEQLAAGITWLGLAALQDPPRPEVVTALKKMN
QAGVSVKMITGDHPKTAQAIGQQLGLAPGKIKALTGAEWDCLSQEEKQQAALDHQVF

ARTTPQNKLEIISALQEQGQVTAMTGDGINDAPALKKADIGIAMGIKGTDVAKDAADMIL
ADDNFATMATVIKEGRKIYTNIKKSILFLLPTSFAEGLIVAFSILTGQEVPLQPAQLLWINL
VSAITIQFAFVFEQAESGLMNQPPRSRTQKMFSKADLWQIAFVSSLMAIFALIGYEWFL
RLGVSELSASTMVVNTIVISKICYFFSIRTPKPAFSKASLENKPAWAIIALMLAFQLFLTYV
PFMQTAFKVTALSLLEWSVIGGFSALILLLAELKKFLKSHR
>L-asparaginase (EC 3.5.1.1)
MKKILVLHTGGTIAMSEDEHGAVAPGAENPLNKFSEAFTDQLDLVVENPFNLPSPHMS
PKEMLSLAQRINDAANNNFFGVVITHGTDTLEETAYFLDLTVTSDIPVVVTGAMRSANEI
GSDGLHNFQTAILTAASDEAVGKGVLVVMNDEIHTARYVTKTHTTNVATFRTPTFGPIG
LVFKHNVRFFEQLLRHTALPIDHVCEQVFLVKAYAGMEPDFLDFLASQATTKGVVIEAL
GAGNLPPKVLPSLQKLLGQNIPVVLVSRCFNGVAQDVYDYEGGGVQLKKLGVTFCHG
LNGQKARIKLLVGLSAGLAGSALKHFVSDAVS
>TsaB protein, required for threonylcarbamoyladenosine (t(6)A) formation in tRNA
MKILAIDTSNKPLSVAVLEDQELLAEITTNAQRNHSITLMPLIKEVLQRANVKVADLDRIA
VAQGPGSYTGLRIGVTTAKTLAFTLNKELVGVSSLKTLAAAYPNTEALLVPVFDARREN
VFAAAYQWQAGHLVEVIAEGHYAFANLCQRLADFNQKVIFIGRDSQNFASNLKASDLD
YQLAGPNFAYPRAMEVALLGQVATPVAVQTFVPNYLRLTQAEMQWLEKHPEGKAHDE
SYVEKV
>Ribosomal-protein-S18p-alanine acetyltransferase (EC 2.3.1.-)
MSPMLKKFKRLVTKVLYGESEQVVEFNSRTDEVEGIHYLICRAIVPDIPEMLAIEKSVYE
GNTPWDFETFKAELSKRKTTLYLVARYEDELVGFAGVDFKRANGEAHITNVAVTPEHQ
NRGLGSKLLRVLIESARKQECQNIVLEVKASNDNAQKVYQDLGFKILNRKKNYYTEEQE
DALTMRLDLVPAMGRG
>TsaD/Kae1/Qri7 protein, required for threonylcarbamoyladenosine t(6)A37 formation in tRNA
MAERNLILAFESSCDETSVAVIENGSKILANIVATQINSHKRFGGVVPEVASRHHIEAITL
CVADALEQAQVSYEDLDAVAVTYGPGLVGALLVGVAAAKAVAFAHNLPLIPVNHMAGH
IYAARFVSPLKFPALALLVSGGHTELVYLAGPTDFKIIGETRDDAAGEAYDKVGRVLGLS
YPSGKEIDELAHQGHDSFHFPRAMENEDNFDFSFSGLKSAFINTVHHADQVGESLDKA
DLAASFQASVVDVLAHKTVKALRKFDVKQLILAGGVAANKGLRARLDQELTAFPEVELI
KAPLKLCGDNAAMIGAAGYEAYQAGIRADLDLNAEPSLEFEWLD
>Galactoside O-acetyltransferase (EC 2.3.1.18)
MISKCINDIKNIDYNIELGENFYANHNLVILDGAKVTFGDNVFVAPNCGFYTAGHPLDYE
RRNAGLEYAYPITVGDNVWIGGGVQVMPGVTIGSNVVIGGGSVVTKDIPDNCVAVGNP
CKVIRPITEVDKTRSYDRL
>First ORF in transposon ISC1904
MKAKEVLETLQISRQTLQRWREQGILKARLLPNGRYDWDPESVYAVLNKGEQRGTYL
YARVSTPKQKADLKHQVENLQTFAMKNGYQINGVFQDVASGISFEKRKEFFKLLDLVL
AHRVQRVIITYKDRLSRVGFDLFKYLFAKYHVEIVVMSEITDKKTDQQEIFAEITSLLHAF
SMRMYPGRRKKIKEALSDNRESDVETHGTKNEKE
>FIG00744915: hypothetical protein
MWRRMAQKTKKNKLLRYQRIRQKQTEQQYEYLTERHVVRPSHPKYSVLDRFSHLANN
VYNQGLYRARQALFKGSWLSYNKLDASLKQSNRQKDSMIYSSMNSVHLVQQQLKLVA
RNMNSWNKARKAYAKNPGKFTGRPKLPKYHTKGGLTTILVDNQTAKLRPNGYVEIPC
MDKFKVKLAHPETTAIQEVRIVPQNRRFIVEIVYKTNQTIQYKPDNGRYLGIDPGVNNAF
TLATNVPGITPVIVNGRPIKAVNQFYNKQRAQLTRIHDLLGQNRSSHRLASLDFYRNQK

MNSFAHEASKRIVDFALSHGLNTIIIGKNKGQKQRSRMNKQNNQNFIGIPHQIMINLIKYK
ANLAGIVVIQHSEAYTSQTSFLDKEEPIRENGDKARKRKGISPAKRRLKRGLYKADTGQ
LINADVNGALQIIKKVVPKAFADGIEGIGLCPVKWSLRF
>hypothetical protein
MKTFWYNVFKLKRKKATPMSEKEKMLNQQLYDANYAPELIKERTFIKELCYD
>hypothetical protein
MERKKFRDSKIYQLFVKLNSKDFGNLLDQVKIQRHEQEELRKKGGLRFRFEKEYIDAFN
GIRERRKRD
>Hypothetical membrane spanning protein
MSSPSRSGKIKAQAVSEVKIKMKKLAERFEQLTQMLAQFFSGLSVIVIALLGLILLVRIGL
ELKVLVSLSLKLDAANNFYLIMDKIVAFFILFEFFAMIISALKNKGHVSITLLMGLGLTALLR
NLLIIHDDYLDIILNVVGILLLVVGMAIYRHYVHQDSDQ
>Ribonuclease BN (EC 3.1.-.-)
MTRINSFFKLFFERFVDSNINNAAIVIAYNALLAIFPTVILIGNLLPLLNLKADTIMSYLQTA
VPPTIYKTLNSLVRTFLTSGSGGIASVSALVAIWAASRGINALKRAINDAYGVDSQSAFM
TRIVSIAIIFVFAIFLVALFIFFSFGQLALEKLSPILNLPLSWLDTFRSVKWPTTILGIFAILCLI
YRFLPNAKVHWRCIFPGAVLGTAGWLLITQGFSIYVYYFAQRVLSYGALGTFIVLLFWL
NYSGWTILLGAVVNASLEYCLYGKVTPRRPTFRRILDNVIPKKDNS
>Regulatory protein recX
MPKITLIQAQKKAGRYNIYLDGTYAFPVAESVLIKYQLAKGMEIDKQLLSQITSADLQAKA
YNKALDYLSHQLRTQKELSDHLKELGIDEATIKEVLTKLLDLRLLDDQVYANSFVRTQAN
LSDKGPTVIRQKLRQKGIGENLIDRALLEFPLEAQVANAKKLGQKLAKRHQASPLKIRQ
NKIITGIMTKGFSASVANSALEQLDLQVDPNQELVALTTQAKKAWRRYQRFDLKTRQLK
TKQALYRKGFNLDEIDKVLAELAD
>RNA methyltransferase, TrmA family
MEKKQVVVKTGQKVMLTIKRLGINGEGIGYYKRKITFVKGALPGEVIVAKVTKVHPKYLE
ASLVKIEKPSQDRVEPVDKYDVGGIELEHLAYPAQLAFKEDVIRQALEKFKPKGYQDYE
LRPTLGMDYPYAYRNKAAFQVRKTKSGQVIAGLYQEGSHRLVDLPTFSSQMPLTMKIM
RTLVDLLAKWQVPLYQERQHTGILKTIIVRESVAKQNAQVVLVTRTSKFPQLNNLVADIK
SQLPEVVSIMQNINPQKTSLLWGEQTLKIYGQDYLEERLGEVEFRLSARAFFQLNPSQT
KVLYDQVKQALDLQADESLIDAYCGVGTIGLYVAPAAKAVYGVDIIPEAIEDARQNAALS
NRSNVNYEVGAAEKVIPAWVKQGVKPDALVVDPPRTGLDAQLIKTILQVKPAKFVYVS
CNPSTLARDLVALSKVYQVDYIQSVDMFPQTARVEAVVKLSKR
>Cysteinyl-tRNA synthetase related protein
MVLKTSENALIGLNDHTLVTEADGRRWVTREPALVYFHKHYWFNIVTMIRDNGISYYCN
LASPAYLDREALKYIDYDLDVKVFPNGEKRLLDVDEYEEHGAKWHYSTKTDKILKHNVR
TLVDWIDHGKGPFSKEYVEIWYNRYLELSRRQ
>UPF0118 membrane protein YrrI
MMFDKKKKARLMYWSLELLLVAGLIFMCSRLDFLFKPIGTFISTLFAPIIISGFLFYLLNPL
VNLLMKVRFKRFKVGRTLAVAVVFLLLIALIVLLSASVMPKIIGQFQELIAKVPAYIKEAQK
LILSLNQQENLPTWLKNLDLDSYMKQLEGSLSEMIKKFAFSLTSGIGSVIGAVANVAVTL
VTVPFIVFYMLKDGPKLLPNIKKVFSPRHADAIENLLGQMSATIAKYISGQAIECLAVGTM
SVIGYGLIGLPYALIIGIFAGITNIIPYLGPYIGLIPALFIASTHSLKLLILVVIVCVVVQQIDGN
LVYPNVIGKSLEIHPLTIILLLLVAGNLAGLLGMILGVPVYAVAKVIIKYFYDIYRLRTSEDT
KSQNMD
>Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94)

MTEKIAVLGAGSWGSVLAKVLVENGHEVALWSNSQAQVDELNQQHTNAKYLPDLIYP
EALKATTDLKAAVKDAGTVLFVVPTKVIRLVAQQLIEVLEELGTKPLIVHASKGLELGSH
KRISEVIAEEIPTKYRSGMVVLSGPSHAEEVARQDITLITAASEDLANAKKVQALFMNDY
LRIYTNSDVVGVETGAAFKNVIAIGAGALHGLGYGDDAKAALMTRGLAEISRLGVSFGA
DPLTFIGLSGVGDLIVTCTSVHSRNWRAGDQLGQGKSLEDVIANMGMVIEGINTCKAAY
ELAQQKGIEMPITQAIYNVLYKQADIKSEIANLMQREGKTE
>Glycerol-3-phosphate cytidylyltransferase (EC 2.7.7.39)
MKKVLTYGTFDLLHRGHVRLLKRAAALGDYLIVGLSTDEFNAIKGKKAYTSYADRKYILE
AIRYVDEVIPERDWEQKVKDVQANQVDIFVMGDDWQGKFDFLKDYCQVVYLPRTQGI
STTKIKEDLEKKGRG
>Signal peptidase I (EC 3.4.21.89)
MIKVKEIVATLIPIIIGGLIALVIKTFLVTAISVDGSSMAPNLYANETVVLLRHARVKRNRIVV
FKAAGVDKNNHYLKPNTLYIKRVIGMPGDRIRYSKQGNLYINGKYQAQSYISPLQRKLG
TTLLNPGVKPAKGICLGSNRTFKVPAKTYFVLGDNRTISNDSRYYGFVPAEKIQGVAKA
LPWSRNYQLVNSYK
>ABC transporter ATP binding subunit
MAKLEIKYITKSAPMLTDSKDNKGAPVKFWELRGIRMDISAGEAVGLIGMNTASKSLLL
QIIAGNVKQTTGFITTKHPINLASLSEIDNSLSGLENIRKVIKQKEVDELKGNHLTNAVINFI
DFGQWLYQPASSYSLGMKARLALGLALFIEPKVVLIDEVLGLLDRPFFNKVAQKIQDLK
DVGCSFIIADTKQLYMEQFCERTMWLQFGQVQDFGPTKEVLEQFTFAQEWFNSLSLP
EQNEYLADKQFEQVQFDVNKVYEEFKVEQFKHGYTRKDEPRMRKAFFVERGLDPVVV
EQDQKDKQVQSKQPKTSSNFTAWALGALVILILAGGTTYWWHAHQSTSTSQVAKSKQ
SSSKSSKLTSSQLAASRKAASSVKAAEEASKQKASASQAQAAAKKSSEDAAKAASSAA
AASSSSLAAKRANAQTIEVVDGDTLESLAEKYATTVSEIQSLNDLGASNTIKTGDTLYVP
K
>Magnesium and cobalt efflux protein CorC
MISDFARIALALFAIIWSFLAHATVKASQLLDAHNQDSLRVLADKFPKKQAKLARLADNL
ADYSLRLFVAKGFLYLLGVVLVASLAASYVGFNVWWVLAITLVETYYFVWQDGRAQALI
SERFKNYVGFAYYLGLLFAPLTWLVELLLPKIKVPVESKTNSWKEIVHQIELGHNSGELD
NDEFEMIDGVISMHEKMAKEVMVPRIDAFMIDITNDNERSIDNIIEMNYSRVPVYHEDKD
NIVGVIHIKKLLKAARRYGFDHITIRQVMQPAFFVPETIMIDELLFQMKKSQNQLAILLDEY
GGVVGLVTLEDLLEEIVGEIEDETDEPTETVKQLSETEFLIDGKMPVDDFNDEFAAHLET
TTADTIAGYVIEQLGFFPEEGQVLNVRTPEGIILETHQVIDGTRIAEIKLRLPQQLAKLYDK
RILEQAKLDEKLGN
>Peptide chain release factor 3
MDTQTLKQAVEKRRTFAIISHPDAGKTTITEQLLLFGGVVRQAGTVKAKKSGNFAKSD
WMEIEKKRGISVTSSVMQFDYAGKRVNILDTPGHEDFSEDTYRTLMAVDSAVMVIDSA
KGIEPQTKKLFKICKMRGIPIFTFINKLDRDGRDPMDLVAELEEVLGIEGYAMNWPIGMG
KGLKGLYDIYNNRIELYRTEDKGESFLPLDEEGKLAVANPLEEESIYQQVLEEVELIKGA
GSEFDADKIAHGELTPVFFGSALTNFGVKTFLDAYLQFAPAPSAHKTEAGTEVAPTSEN
FSGFIFKIQANMNPNHRDRIAFVRICSGEFERGMDVVLNRTGKKMRLSNSTQFMADSR
ETVETAVAGDIIGLYDTGNFQIGDTIYTGKKALQYEKLPQFTPELFMRVTAKNVMKQKS
FHKGIQQLVQEGAVQLYQSYSTGDYILGAVGQLQFEVFQFRMANEYNSEVLMTPMGH
KTARWINPDQLDEKMSSSRNLLVKDRYGAPLFLFENDFAERWFKDKYPDVELTSKL
>Pyruvate formate-lyase (EC 2.3.1.54)

MLKVDERLDNKAWDGFKGGDWQTEIDVRDFIQNNFTQYNGDQSFLEGPTEATTKLND
QLVDLKLKERAAGGVLEADTKIVSTITSHKPGYLDKSLEKIVGVQTDKPMKKALMPFGGI
RMAKDALEAYGLKLDPEVEYIFTELSKTHNQGVFDAYNADIRRARHNKIITGLPDAYGR
GRIVADFPRVALYGIDRLIAAKIEDLNNYGDGEMTDDVIRMREEISDQIKALKDMKKMAE
SYGYDISQPAQNAQEAIQWLYFGYLAAIKQQNGAAMSIGRIDAFIDIYIERDLERGVLTE
KEAQELIDQFTMKLRMVTFARTPEYNSLFSGNPIWATLSLAGMGLDGRHHVTKTAFRF
LNTLHNMGAAPEPNITLLWSENLPETFKKYAASVSIDSSTIQYENDDLMVKTWGTDYY
GIACCVSAQPIKDGVQFFGARANLAKTILYAINGGKDELTKQQVGPKYEPITSEYIDYDE
FMAKFDDMLEWVADVYVNALNAIHYMHDKYYYESAQLALKDTRLDRTFATGISGLSVA
ADSISAIKYGHVKVIRDEDGMAVDFVAENDFPRYGNNDDRVDDIAKFLVKDLYTKMNK
HHLYRGAKLSTSVLTITSNVVYGKNTGSTPNGRKAGEPFAPGANPAYGADTSGALASL
LSVAKIPYRYATDGISYTFAVTPQTLGHDEAAKENTLVNMISGYMKNDGHHLNINVFNR
DTLLDAQAHPEKYPTLTIRVSGYCVYFADLTKEQQDDVIARTFHSSL
>Pyruvate formate-lyase activating enzyme (EC 1.97.1.4)
MREKNPLLKPVPMRDGQVEGYVHSIETFGAVDGPGIRFVVFMQGCHMRCKFCHNPD
TWKTNVGTKMTSDEVLEQALPYKAFWGDQGGITLSGGEILLQIDFALELFTKCKELGIS
TCLDTCGQPFTRRQPWFDKFQKLMDVTDILLVDIKHINSDEHKRLTGFPNENILDLTRY
LSDIGKPVWIRHVLIPGITDVDKYLMELDAYIKTLHNVEKVEILPYHTMGVHKYHEMGIKY
RLEGVEPPTEDRVKNAQELLHCDDYRGYLSWRPGIKTFNP
>hypothetical protein
MKYLFLSLRLLAYLLTGCLAFLAILIYRRELDFTDLSYHFRLSQLYFGFNLLLLPWFRLLH
PLVFISQPENFKSYKTRALDFYHQDLAAHSKRKKWSSDGVNVNPYEYLSPYTMSYDN
ALNWAYNLCKALLMSILFLLAGPCFWLYHYSKRKLISPRH
>Predicted metal-dependent membrane protease
MVENRWSFGRQLYWGLIFFLGFAFLSLFFFDESERFKNLFELSLGLLAVLVACYSKRGL
KILKRRPSFGRQTRWLQWLGLGCLLYMMLVGHVFSSIYALFFDRHFLDDTLIALGAATL
EEGICRGFLLSGFLGLAVYYGFKRKLVWASFAAAVTFSGFHLINLLGTPVTVVSQQLFY
SLAMGVFLTASYFATNNLLVPLGIHFLLDWGPSYNYQPQPSSWLTLLAVFLPLLLVSSW
YLVVLDRQVSDYF
>protein of unknown function DUF1440
MQKSLKNTLVAGLAAGLISGLVKLGWENILPPRTPQRDAVNPPQTFLQQLGIPAKVTQL
TYSYSGHQMPWASFLVHFGFSVTFATTYALALEKKVTWLTKKQGLPFGVAMWIAFHLV
IMPAMKTIPTAKRQPLAEHVSEILGHMFWMYTNHEVAKEILDRLDRK
>Phosphate ABC transporter, periplasmic phosphate-binding protein PstS (TC 3.A.1.7.1)
MKKIVQITALSLFLGLFLSACGNENKIKQVTIVGSTALQPMVEVAAEEYQTHNEGISITVQ
GGGSGTGLSQVQAGAVTIGNSDIYAAEQPGIKADKLVDHKVAVVGIGPIVHKGVGVKN
VSMEQLAGIFTGKYTNWKQLGGKDQEIIVINRAQGSGTRAAFEGKVLPKQKALNAQEQ
ESSGTVAKMVAATPGAISYVAFSYFKDNVTALAVNGVKPTDQNVMTDTWKIWAYEHM
YTKGKPDLETKKFLDYMRSEEIQTKVVKKLGYISINEMKVKVD
>Phosphate transport system permease protein PstC (TC 3.A.1.7.1)
MDKIKEKITSNSKETRQELRGKILTYSAIGLIILVVLAIFIFVASKGLATFIQDKASLKTFLTT
TTWNPNVLGADGKPLVGALPMIAGSFGVTILAALLATPFAVGTALYMTELAGKRGRDFL
QVVIELLVGIPSVVYGFIGLNVVVPFVRNHFGGSGYGILAGAVVLFVMILPTVTSMTVDA
LKAVPKHYKSASLALGATQWQTIYKVLLRASIPGILTAVIFGMARAFGEALAVQMVIGNA
SLLPTSLLNPASTLTSVLTMGIGNTVMGTTANNALWSLALVLLLMSLLFNVLVRWVVRK
RKM

>Phosphate transport system permease protein PstA (TC 3.A.1.7.1)
MNAKLQNKIALTIIKLIAGIVVLILAFLIGKILLEGLPHISWQFLTSPAKAFEAGGGVGIQLF
NSFYLLVLTLLISFPISLGAGIYLSEYAPKNAVTATVRMAIEVLSSLPSVVVGLFGFLLFVV
QFKFGYSILSGAIALTFFNLPLLTRNIEESLQAVPNLQREAGLALGLSKWRTVNKIILPEA
LPGIVTGIVLSAGRIFGEAAALIFTAGQSATNVDFTNWNPFASNSPLNLMRPAETLAVHI
WKINSEGIMPDADAVSAGASAVLIIAVLVFNLGARFLGNYLYKKLTAAASK
>Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)
MNQYDLKQNYILPLPEPEHEIALETKNLEVFYGAKQAFFEGNLKFERYKITSLIGASGCG
KSTYLRSLNRMNDGIGTVKGQILYRDLDINGPDINVYELRKHIGMVFQRPNPFSKSIKEN
ITFALKENGLKDKEKLAEIVEESLKKAALWDEVKDDLNKSALALSGGQQQRLCIARAIAM
SPDILLMDEPCSALDPISTLKVEETMLELKKNYTIIIVTHNMQQASRVSDYTAFFHMGHV
LEFDKTKKIFTNPTIQATDDYVSGNFG
>Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)
MTELITTQDVHLYYGKKEALKGIDLNFPDKGIHALIGPSGCGKSTYLRCLNRMNDLIENV
KITGSFKLAGQDIYGPKTDLVELRKKVGMVFQQPNPFPFSIYENVTYGLRLAGIKDKQLL
DERVETSLKQAAVWDEVKDNLNKSALALSGGQQQRVCIARVLAVKPQVILLDEPTSAL
DPVSSNLIEDMLLTIKDDYTIIIVTHNMQQASRISDTTSFFLNGELIETGKTKHVFLNPKKQ
ATNDYLSGRFG
>Phosphate transport system regulatory protein PhoU
MRRMLDQQIDNLKKEFTTMSLNVEAAIKNSLLAFAEQDLELVQEVIANDETINAQEIHLE
KQCTQVIALQQPVASDLRLIIAMLKASSDLERLGDHAISIAKATNHFENEQRDEAIDATVI
EMGQAIQAMLAKIIRAYENLDTKQAIEIANQDNEINDYFKRIRKLSLTEIHNDTDFAESGID
YLNVANHLERMGDYIINLAEWIFYATEGKITELGRNDL
>Fructose-bisphosphate aldolase class II (EC 4.1.2.13)
MSFVNGNEIFTAARQGHYAVGAFNTNNLEWTRAILKAAQEKNTPVLIQVSMGAAKYMG
DYKLVKTLVEEEMRVMGITVPVVMHLDHGNYEAAKACIEAGYTSVMFDGHDLPFEENL
AKTKEIVELAHAKGISVEAEVGSIGGEEDGIIGAGELADVEEAKQIAALGVDYLACGIGNI
HGKYPENWQGLNFDRLKELADAIDTPLVLHGGSGIPQEQVVKAIKMGVSKVNINTECQ
LAFADATRKYVEAGKDLEGKGYDPRKLLAPGTQAIVDTVKEIIDWLGTPAVK
>hypothetical protein
MDTVMLANWLKNFQAPNLPTWEEFPDFELYMDQLVSLGNRYLDNLTDSPITSSMVNS
YVKKGLMAHPIKKKYQPEHIAELIVISLLKNVYSLEIIKKSIREVIAQSTVKEAYNHFASVF
NSTLIQVRQQEEKILQVDTSDSVINFTVQFAVRAVIYKLISSRLIELQHQQHQHE
>Mobile element protein
MLCVQKVRLYPNQTMKQVLDDLCDYRRYCWNQGLALWNDMYDASLILDDKKLRPSA
HKVRDELVANKEDWQYQLSARCLQLAISDLGKAWQNFFKKSLPDWGKPKFKSKKTAR
QGFKTDRAQIVNGKLRLDKPQGIKAWADISFKGADDLNGDLKVVSIYRENGKYWASLP
FEVKVTKKTKTGQKTAVDANVGHFDYTEGQVKTLPNNLKALYKRIKHYQRLLARKRVA
NGKKATQANNYVKTKAKLQRDYRKVANIQHDIVQKFTTMLVDNYDRIAIEDLAVKQMQ
MSHVASKGLHQSMFGYFKQVLKYKCKWYGKELILANRYYPSTQRCSACGHVKTGKD
KVGLDGNKKHKTKHNEYICYDCGAVMDRDENAVRNLLALL
>COG1272: Predicted membrane protein hemolysin III homolog
MDRKSQIINEVLSAVSHGIGVGLSIAGLVFLIIKGVASNSGQALFAYLVYGISLLGLYSFS
MLFHSLYFTRARRLFQIFDHCGIFLLIAGTYTPYCLLAIKGSLGLGLLIIIWLLAIFGIIYHILA
KKRLQWIETLTFVIMGWLCLVGARPLAQALGSHGLTLLVFGGIVFTLGALVYSIKGVKYA
HVYWHFFVMLGSLLMFFSIYYYI

>Transcriptional regulator, ArsR family
MPVVKLTAKQETVLDRFVNEINVLNALADVKRQQIIVILCKHEKEGLTVTDITNQMEITQP
AVSHHLKILREAGLVSFRKNGLQSVYYLTVSEPLKNLEEVVHALRTEIVEA
>Ribonuclease HI (EC 3.1.26.4)
MGKFYAVKRGRQTGIFTSWPDCQKQVSGFPNARFKGFTDKTAALNWLNAESSTTQP
STSLKKAAIQVYTDGGSRNHGNKLGQHVKADDKAAWAYLINTPTGSYEESGGEYGST
NNRMEIMAFLQALKKLLALGLNHEPLEFVLDSRYVLDSLTKGWLTSWQRRNWRKADG
NAVLNKELWQEVATLLPNFSTPYYTWTKGHANNPGNNRVDYLLNQTMDQM
>ABC transporter, ATP-binding protein
MTALLELNNISTIVNKGTSEERQILKNLSLQVNQGDFITLLGTNGAGKSTLLNVINGALSP
DQGSISLKGQDLTKLNEVKRAQYIAQVFQDPKMGTAPRMTVAENLLLAYKRGQKRTLR
LRHLKQHRQEFESLTAKLPNGLSEHLNTATGSLSGGQRQTLSFLMATLNRPELLLLDE
HTAALDPNTSKQLLALTDEVVQAEKLTCIMITHQLKDALKYGNRTIILNNGQVVLDVAGA
ARAQLSEQEVLSYFTDL
>ABC transporter permease protein
MTLIVSAIGQGLIWGILGLGLFLTFRILNFPDMTVEGTFPFGASVCVAALVHGASPLTATL
LSLCAGLLTGLVTGLLYTKGKIPILLAGILVMTGIYSINLRILGKATVGLLNKPTILTTSFFN
SLPEGFNSVFLGIIVVLIVISLLALFLNTDLGQAFIATGDNEKMARSLGISTDAMKILGLMIS
NAMISLAGGLIAQNNGYADVNMGIGTIVIGLAAIIIGEVAFGELSLTARLIAVVIGSIFYRFV
LLIVLQLGFSTNDFKILSALVLALCLMLPLIERKLNLTGLISKGVKKP
>ABC transporter substrate-binding protein
MKRLITTIIALVAFLTFAFFHEQKIATQATKTPTVGILQLMTHPALDSIHKGIIAGLKENGYV
VGKNLKIDYQNAQGDQSNLKSMSDRFANEKVNVMVGIATPSAQSLAHVTTKTPIVMGA
ISDPVGSGLVKSLTNPSNNGNITGTQHIEPVAQQVKLIKKLMPNLKTIGGIYTSSDDSST
AEFKEFKQLAQKAGIEVKTYTITSTNDINQVAATMAHEVQAVYVPTDNTVASGIQALLKN
TNAAKIPVFASVATMVKSGAVATYSVSQYDIGRLTGKLVAQILKGKKPNQMPIQHVTQG
DYAFNLSQAQRFGINIPQDLIKQAEKKGVIYK
>Beta-glucoside bgl operon antiterminator, BglG family
MDFIKNFNNNAALVADQAGNEWIVLGKGVGFGQKLGQPIDEAKIERRFKAAGSDDTTL
ATIKSVSPLTLEATSAAIKLIEAESSIRFDNFQYLALADHIDFAIIRSEGGIDMEDRALRWE
VKRLFKQEYSLAKRVVKLINGLTGASLPASEEVLMTYHLVNAESDGAKVQDTMKITKLIA
GIVDIVQYQYGLTLDVESFNYTRFIGHLRAFMVQRLSNARPYGAELDGELLVLMEQKYP
QAAVTVTRIDNFLQTKMGWTLNPDDRVYLILHVWRVTHRQEQ
>ErfK/YbiS/YcfS/YnhG family protein, putative
MEIINKKVKVIGASIVALVVLALGGLTWYQHTHFNSQVTINGTKVGGLTADAALKKLRSV
SLTNAIYLDGKLFYQGQKTAAGFTASDLGAVKKVLKKQVTLFPTKAAKDYEIAPSKISNY
RKVTLKQELKAKLESENLKRTKAQDAYAILKDGKVSVVPAKKGNQYDVAKILKEYDRIS
YSSDLNLRVQYLQPLSAKSQEVKTEQTKLAELAGRKVDYQVQDKHYELKANEMLTSA
RYVDGKYQFDTAALKNKVNEINQAQATLGKTFTFTTSTGKTIQVPGKTYGWALRDSDAI
ASITKAYETGKPSLNAVNDIYGTGYLTYGTGYDTTLNGGLGNTYAEVSIVDQHVWLYKN
GQQVASIDVVTGKKSTGEDTPTGVWYIMYKQSPSVLRGSSAGSGSYEVKVNYWAQFT
NSGCGFHDASWRKNWAKDAYISDGSGGCVNVKPSEMPNIYNNLSQKEAVIIY
>D-serine/D-alanine/glycine transporter
MSKDKKVANHSGLNRSLSSGQMEMIALGGTIGVGLFMGSTSTIKWTGPSVLLAYAVVG
VVLYAVMRALGEMIYISPGTGSFADYATHYIHPLAGYLTKWSNIFQYVVVGISEVIAVTQ
YLNYWWPNLPDWVSGLVVVVTLTLANLASAKAYGTLEFYFAMIKVVTIVLMIILGLLVIFL

GLGNGWHPVGLANLWQHGGFFTGGFKGFMFSLSIIVGSYQGIELLGITAGEAQNPKHA
IVASVKSIVWRILIFYIGAIFVIVTIYPWNQLAAVGSPFVETFTKVGIAGAAGIINFVVLTAA
MSGANSGIYSSSRMLYKLALDKEAPGVFAKLSKHIVPNVAILTISGGILIGFVLNIVLSLLS
KSSENLFVIVYSSSVLPGLVPWFVILLSELRFRHQNQASLVDHPFKMPLYPCSNYLAIIF
LLVIVLFMFINPDTRVSVSVGAVFLLGLSLLYYVKHGRQK
>Deblocking aminopeptidase (EC 3.4.11.-)
MEKAAQIQMLTDFSNANAVPGFEHEFVKMFVDRIKPYADIEIDGMLNVYASKKENQGK
RPVIQLDAHSDAVGFMTQAVRPNGLLKFVPLGGWVKYNIPGTKVKVRNKEGVYIPGVV
ATKPPHFMAAGEGDKVPEIAEMQIDVGSSSRKETLEVFKIDTGCPIFVDVDCQYNEQTE
VFFGKDFDDRFGAAVMVNLLAELQGKQTNLDVVCALSSQEEVGLRGAYVTARKIKPDV
AIVLESCPADDTFEPDWLSQTGMKRGPMLRDMDKTFIPNPVFQQYVCELADQRKIPYT
RSVRTGGGQDGAAIYYENGAPTIVIGIPVRYEHSPYCFASYQDYQASLDLALAIVEDLTV
EKLASFKQPRL
>DNA-binding response regulator, OmpR family
MKLLMVEDNKSVSEMMSMFFKKEQWDVEYAYDGEEAVEKFKQDPDSWDMITLDLNL
PKKDGMQVNAEIRAISKTVPIIILTARDSESDQVLGLEMGADDYVVKPFSPITLIARIKALH
RRAEIPSVKVANVLNDTEGYDVVTDHFKLNTKTREAFLFDKQINDLTPKEFDLLKTLAQ
KPRQVFSREQLLQLVWDYEYYGDERTVDAHIKKLRQKLEKVGPQVIQTVWGVGYKFD
DSEVKLK
>Two-component system histidine kinase
MKLIYQQMLAFFVVIITLLVILGASFFHLMRENVYQSTWQQMETYAYSLKSTSLKLVDTS
KGEQIALNVSNLEISELLLQNQKVHFTIYTGRDKVLYPLDGFQSVISKTDWQKLQAGQIL
RRRKDIGDRKKVNQKQLSTAQKKLSSAQKQLPNWGKQVITTDVFVPCFNSQGKLVAVI
SVGAMVSNIESSIKEIQKSLIYALLVSALLGILISYILAHYSTKRISRLRKASHQIAGGNFDV
QIENNGRDELDYLAADFNKMASSLKESQEEIKRQEERRRQFMADAAHEMRTPLTTING
LLEGLAYDAIPEESKGQSIELMRNETKRLIRLVNENLDYEKIRSGAISLSKSKFDAVTVIH
NLVEQLQQKAQESNDRFELALPKQLEVYADYDRFVQILFNVMQNAVQFTDDGVINISA
ERGFNEVIIRIADTGIGMSQEQLRNIWERYYKADPSRNNTKYGESGLGLAIVHQLMQLH
QGKIEVTSQEGKGTTFSLTFPSPK
>Lipoteichoic acid synthase LtaS Type IIb
MRLKKISEFLNTRKGFFTFLVGLFWLKTIIVYCVDFKLGVATPYQWLVLLINPLATTMLLF
GLALYVHRTLLSYLTMVLIYIANSALIIFNVIYYREFTDFMTFNVIFGYSSVSNGLSTSSLA
LLEPADLILILDLVLILGALITKKIKLDPQPVARKTAVAISSLGCLMFFFNLTLAEISRPQLLG
RTFDRVYLVKYLGLDAYTVYDGLKTVKTNQARSDAESADLNKVLTFTKEHYAKPNPDL
FGIARGKNVIVIHLESFQQFLINYKLDGKEVTPFLNSLYNGNETYSFANFFNQVGQGKT
SDAENMLETSTYGLPQGSLFTTLGSDNTFQAAPAILNQEGNYSSAVFHGNVGSFWNR
DNTYKNMGYQYFFDASYYNTNSNNLTEYGLKDKLLFHDSVKYLERLQQPFYAKFITVS
NHFPFELDKKNTSFKAANTGDSTVDNYFVTAHYLDQAVNEFFNYLKESGLYNNSIIVIY
GDHYGISDSRNLKLAPLLGKSATTWSDYDDAQMQRVPFMIHIPGITTGGIQEQYGGEID
VLPTLLHLLGIDSSKYVQFGTDLFSPNHDQVVAFRSGTFVTPKYTVVGSTIYQNGSGTV
ITHPTNKLKQELKADKQKVRTELALSDSLNNKNLLRFYVPAGFTPVNPQNYNYKKQLN
QLEKLQQILGNNSTSLWSQKGNQTTANLYNSDAPELKTTDQEATDASASTASSQTTTQ
>glucose uptake protein
MLIALIPALAWGSVGLISGKLGGNANQQTLGMTWGALVFSILMTLFSWQHFISSNSLKL
WLVGIVSGLFWSLGQNQQFHAMKAVGISKAVPISTGAQLVSNALAGALLFHEWTTGR
QLSLGTIALVVLVLGATLTALKDKSVATDNTVKENWALGSRALVLSTLGYLGYTVVVTW

ASVDTQAMVLPQALGMVAGASVFALGKDAFEVATFKNIITGLVWGTGNFFMFMATAAV
GLAISFSFSQMGIVISTFGSILLLGEHKTKKEWLYVCLGSILVIVGGIILGQLK
>Glucose-6-phosphate isomerase (EC 5.3.1.9)
MAHISFDSSNLTKFVNPNELGEMQALVTAVDKELREGTGAGNDFRGFLNLPVDYDKEE
FARIKEAAKKIQSDSDVLVVIGIGGSYLGARAAIEFLHHSFYNYLPAEKRNAPQIFFAGNS
ISSSYVHDLLDLIGDRDFSVNVISKSGTTTEPSIAFRVFKEKLVEKYGQEGARERIYATT
DRARGALKTEADAEGYQTFVIPDDVGGRFSVLTPVGLLPIAVSGADIDKLMAGAAQAR
EDYADADLEKNEAYQYAALRNILYRKGYTTEILENYEPSLQYFSEWWKQLMGESEGKD
QKGIYPSSANFSTDLHSLGQYIQEGRRNLMETVVKVGKPNHDVAIPKEEENLDGLGYL
EGKSMDFVNTKAFQGVVLAHTDGDVPNMIVNIPDQTAFTLGYTIYFFEIAVAISGYLNGI
NPFNQPGVEAYKKNMFALLGRPGYEELTAELTKRLEK
>putative oxidoreductase
MAQISYGILGASQPATEFVKAIEQVEEAKVVAIAGRSLADAQKQAERLGLERAYGSYEE
LCEVSDVDVIYIPKYNRAHYACAKYALEHGKHVLMAQPFTRHKVGAGELFQLAANKNL
LLMEEQSVLFLPLMQRIKAEVAKGTIGKLRFIEIKRYLKLANNNDWLQDLAAGGGALNV
GGSNPLELAAYLTDGEVSEWAGFAENNIGEADRSASLTFKVADLLVNAFVTTELTGED
KLTLVGTDGKITVPNYTEAKVATVETTSGTNQLVDQNYQSGLAYLLAHFNECLAKGLVA
SPLVTPAQTIKALTVIEDLYQRWYGDPLN
>Glycerol uptake facilitator protein
MHFALTARLAAEFFGTAIMVILGNGSVANVELNGTKGSRSGWMTIAVGYGFAVMIPAL
MFGSVSGNHINPAFTLGLAFAGDFPWSEVLPYIAVQFLGAMVGQLVIVACFKPHYDLTE
NPEAILGTFSTIDASKSRLNGFVNEYFGTFTLVFAALAITHSPIFAHGNQGAGFIGIGLLV
MVLVTSFGGPTGPALNPARDLGPRIVHALLPLKHKGSSQWDYSWVPVLAPICAGISAA
LLYKAVFGL
>Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205)
MIAKAVEEMFSQRKGGYLIPAEIVANVSETNNLNHALLVLTKVKYAKIPVLDKEQRVTGL
LSLAMITETMLGLNGIDPSRLNRMTVKEIMQTDFPKIQAPYDIEEILHLLVDQPFLTVVDQ
AGVFLGIITRREIMKEVNYFAHDLEKQYLIEPKAGYLEETKNTGTKVNLVVPKRRSSSAK
VRR
>2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-acetyltransferase (EC 2.3.1.89)
MPKLDANEIIRTIAESKKQTPVKVFIKGNLNELEVPASIDAYLNDKAGVLFGDWQEVEAF
LKANQAKITNYRVENDGRNTGVPMIDKKQFNARIEPGAIIRDQVEIGDNAVIMMGAVINI
GAEVGEGSMIDMGAILGGRALVGKHCHIGAGTVLAGVVEPPSAQPVTIEDDVLIGANAV
VLEGVRVGKGAIVAAGAVVTKDVEPYTVVAGTPARKIKDVKDVAGDKKEIVDDLRKL
>N-acetyl-L,L-diaminopimelate deacetylase (EC 3.5.1.47)
MMAFTAEELIKIRRQLHQIPEIGMEEFETSAKLMAIIKSLPQTFLEIKTWKTAILVHVQGK
NPSKTIGYRTDIDGLPVTEQTGLPYASKHAGRMHACGHDIHMTVALGILSHFAEVQPEC
NLTFLFQPSEEILSGGKQVYESGVMDEWMPDEIYAFHDNPQLPAGTIGCRLGTLFAGT
CEIHATFTGKSGHAAYPHMANDMVVAGAQFVNQIQTIVSRNVDPIQAGVVTLGHFEAG
TTGNVIAGKCQIDGTIRALTQVNNRLIQKRVEAIAKGVAQAFDCQLDLKLEQGGYYPVE
NNDQTTKAFIDFMQANPEVDFVETQPAMTGEDFGFLLSKIPGTMFWLGVDSPYSLHSE
YLAPKEEAIFAGVKAMVAYIEHRQTEVFAAKY
>Potassium efflux system KefA protein / Small-conductance mechanosensitive channel
MTRNIFHYTILFFYLYAILSILGVPVGTLLAGAGILSVALGLGAQGFVSDVITGAFILLEQQ
LAVGDSVKIGTIEGTVHSVGLRTTQVLGYDGTLNFIPNRSITIISNLSRNDMRALIEIGLDP

KKEIKEMIALVTKVNEELKPNYPDITNGPNILGTFDSGNGRMVLRIAIYTKNGAQFAIQSA
FLERYLEVLRSNGFVIPTSPLNLAGK
>hypothetical protein
MENIAGQRIDRFDTIASQVLDKTGVSKTSFNNAFANEVKQDEE
>Pyridoxamine 5'-phosphate oxidase (EC 1.4.3.5)
MRKVGKSMVLTFGQLAGLIAAGAFLILVIALCFVLKGLIKTLEETTKSIVAVTEDINGISKE
LEQLVNKTNTLMDDVNQKSAKLDPLFETMAELSESVSDLNAASRNVAERVSGSAKAAT
QASMAFRLGKQALKFYNKHKKNPTD
>hypothetical protein
MMKKTFLLGLGAVAGASFLLSKKVTEEQREKIALKLDEALLDGREKALKYDRYLRDFLQ
ANSLDFAPLKDKVLAKTNALKENEKMTDAITSLKQATTELREHLRVAKEELADEDLAED
MQDDIIIDGRSAFGEAKDVADFESEHPTEVFYPKKSTEVTD
>FMN reductase (EC 1.5.1.29)
MKIVALAGSVVGKKTLTVMNYVAQKMATDFSEIDFELIELAKKDIQFSDGRNYLEYSGD
TLEVTTKIMAADALIIGSPIFQAAIPGSVKNIFDLLPEKALRDKVVSLVITAGSPKHYLVAE
MQLKPILAYMKAQVLPEIVFVEGQDLYRNEIINHDIIFRLDKLVEDTVLMVETFKSLRQKQ
EDALF
>transposase
METQVTVKVKLNLANADIASSFTNTMEQYRLACNYVSDYIFNHDFDMKQSRLNKVLYT
NLRSLFMLKSQMAQSVIRTVIARYKTVKTQMKRNPYKYQDINTGEWYREIRDLTWLQK
PITFNRPQVDLQRNRDWSYLSSGQLSINTLDGRVKVNPICHGFDQYFDGTWSFGLAKL
LKTGGKWYLHISATKTVADFDNQTVKHVVGIDRGLRFLATTYDEQGQTAFFDGKTVMR
KRAKYQKLRAKLQAKGTKSAKRRLKKLSGRENRWMTDVNHRLSKTLVQKFGANTLFV
LEDLTGVSFERTDLPKVLRNQNKSWAFYQLAQFLTYKAHLNNSEVIEVSAKYTSQRCP
KCGTIKKTNRNHDLHEYHCSNCGYRSNDDRLGAMNIQLLGTQYISGQEQPKFELSN
>Luciferase-like monooxygenase (EC 1.14.-.-)
MSMTDYRINPKTGLEFGLYSLGDHMPDALTGERVSAKERVNQIIQEAKLAEDAGLDFF
SIGESHQQYFVGQAHAVLLGAVAQATSKIKLGSAATIVSTSDPVRIYENFATIDLLSNGR
AEIIGGRASRLGLFELLGYNVSDYEELFEEKFELLLKINQEKRITWQGNFRAPLHDAEILP
RPLNDKLPIWRAVGGPAESAIKAGMQGVPMAIAMLGGPVAIFKRSVDAYRQSLRLSGY
QPTDFPVATTGLFYAAADTKTALREFYPYVDKGITLANGTGFNKRLFANGAKIQDVMAI
GSPQAIIEKILYQYEQFGHQRYMAQIDFGGLPFEKVMKNIEIIGKDILPAIKKYTRK
>Proline dipeptidase (EC 3.4.13.9)
MSHLTKVQQWLAENNYDIAYVSDYKNIQYFTGFGSDPIERVLALFIFPDKDPFIFAPALE
VEAVKEVGWKYPVYGYLDHENPFALIKDHILQTAGNPVKWAIEKGNLTVERYEAFKQQ
FPSADFTGNISPLIEQEKLYKTPADIELLKVAGKWADFAFEVGFNALAQDKTEQDIVAEI
EYALKKKGILHMSFDTIVQSGANAADPHGAPKADKIRPNELCLFDLGTVYKGFVSDASR
TVAFGQIDAKSKEIYDVCLEAQLTAQAAAKPGITAAQLDKIARDVIDKAGYGQYFIHRLG
HGMGQSDHEFPSIMEGNNMELEPGMCFSIEPGIYIPGVAGVRIEDCVYVTDTGVEPFT
HHPKDLITLPVK
>Catabolite control protein A
MKEVMKNKFWRKANMEKQTITIYDVAHEAGVSMATVSRVVNGNPNVKPATRKKVLEVI
DRLDYRPNAVARGLASKKTTTVGVIIPDVTNVYFSSLARGIDDVAAMYKYNIILANSDGN
PKKEIQVLNTLLAKQVDGIIFMGNNITDDLRAQFERSKTPIVLAGSVDPKEEVESVHIDYV
AAVEEAVSDLINHGNKRVAFISGPLSDPINGQYRLKGYKNVLAKHGIEYDPELVFETDY

SYRSGEILWPAMVAAKATAAFVGDDELAAGVINGAQDAGVKVPDEFEVITSNNSKLTEI
IRPQMSSITQPLYDIGAVAMRFLTKMMNKETVEEKTIELPYGFIKRSSTK
>Predicted L-lactate dehydrogenase, Fe-S oxidoreductase subunit YkgE
MKKPKVCIFSTCLVDLLFPKVGIAMVEVLERLGCETVLPEKQICCGQPTYNSGYSKASIK
TFKNEIDALLSIDADYIVGPSGSCVFMIKEYKDIFKDDPEYGPKAAEIDSKIYEFTQFIYRV
LGILDCGAQLDAKVTVHRSCHMTRLIGERTAPFVLLDHVKDLKRIPLKNVQLCCGFGGT
FSAKRPELSVEMVDEKVKNIMDTDAEILIGTEESCLMNIAGRMNRLGHQIKVMHIAEVLN
HNVDPSRIKYISDTDHVVVPANGEGVF
>Predicted L-lactate dehydrogenase, Iron-sulfur cluster-binding subunit YkgF
MGISTSNKPYLERIADAKSDKFMQAAVAKAQDAQFIKRTKARREMGHWNEWRDLAEQ
IRQHVLKYLPDYLEEFAGNVEKQGGHVFFAQTEKEARDFIKELVIKKQAHNVVKSKSMV
TTEIDLDKTLQDIEGVNVLESDLAEFILQEDNWDEPTHIVFPTLHKNRDQVQAEFKKLGY
TGDNDPQHEARFVRGYLRDYFMKADFGITGCNFAVADKGIINLDTNEGNADLTMALPK
TQVVVMGMERIVPSMKEAEVLDNMLARSAVGQKLTTYCTFSGKQVAGEVDGPEDFW
VVILDNGRSNALGTEFEPILQCIRCGACMNVCPVYRHIGGKGYGSIYPGPLGKVLSPILG
GYDQFEDLPYACSLCAACSETCPVKIPLHNLIRKHRIVEMDEKHLDHSMTNLIMKVIGT
GTGSPFLFQSALKMEHFGLGMFGKAQDEKDVSGTIYENGKIKHLPKVAPKLVKGWSDI
RDIPVPPKNKDNFRNWYKKHQPVEVAIRVKQKGDESDD
>Predicted L-lactate dehydrogenase, hypothetical protein subunit YkgG
MTNLLSQANSKENREAFLDRIAKKAGRPRHQVKAFTPLNDLPEQVLAGKSGVELLTIAK
ENAEEVNAKVVVKKRAELEEFLKDYAQEIGAKRLLLPGVKPAKWQEYQLDHWAKATGI
EQTDCWSTELSRMENEQLANQADLAVGFADYLIANTGTITVVVSPEQGRGFNFLPEHY
LALVPKSGLVASTRQAVTKYEQRLAAGSLKTSAINFISGPSNSGDIEMELVVGVHGPLD
CTYLVVEDL
>Transcriptional regulator, Cro/CI family
MTLRTIKARLAEIEENLMVEDFPKVKQLLKKIDVSQIKTKEEVLRYYLLEGMTNVLTNQP
LGRVTASFNPILKALDTKNLSLYVYLAYLGEGIFYMRHDASRQAHFFFHKVERFFNKLP
KKLGDELTYAYLTMAYFLAEYHSLMADFEESNHLIKLALAKCASSHRTYYLPRLKLLEAE
NNLNLKKDSQVILEALEDARAFARLNNNNAVQLQVAALLRSYQRLVLD
>integral membrane protein
MTWIPVLVLCLLVIVSFNRIYFFKKWTSWQKITLYSALIYSITVVWLTFYPAGMAFFAGSY
KPIKYFGNVAYNLVALKYFDGGFVENVLMTVPAGVYLRLAKPKLSLVQLLGWALLPGLA
IESCQFISDMAVNLQRVVDIDDLCANTLGVLLGFVLLQLLERTSLKKLIIKLSLAN
>FIG000859: hypothetical protein YebC
MSGHSKWHNIQGRKNAQDAKRGKIFQKISREIYMAAKSGGPDPDSNPQLRLMMDKAR
AANMPKDNIKRALDKATGAGGENYEEITYEGYGPGGVAILVHTLTDNRNRTASAVRTD
FNRNGGALGETGSVSFMFDRKGYIAIAREELDVDEDQMFEDILEAGAEDMQASDDVFE
IYTDPKSFAEVRDALEKKYDLATAELTMIPQNTVPVPEDKVEQLQRLIDRLEDEDDVSE
VFTAAEFPED
>Ribokinase (EC 2.7.1.15)
MNKVTVLGSLNVDTILKINRLPKPGETMQMNSKQNAGGGKGANQAIAAARCGAKTSFI
GKIGNDANGAMMLELLKEAEIDTDAITYSASGTGQAFILLQESGENSILVYGGANQELSL
EDIDQAQAKIAASDFLITQFETPLAGGKYAFELAKKAGVKTILNPAPAKQVPPEMLSLTD
LITPNETEAQALTGIEVKDEVSCQQAAQKLQALGVKNVIITLGSKGAFYQTKTTHGFVPA
FKVQAVDTTAAGDTFLGSLSSQLDTDFTNIEAAIRFANKASSLAVQKLGAIPSIPTYAEIS
ANN

>Inositol transport system sugar-binding protein
MLKKFFSLCLILTCASFFLAACGHDQADGNSKKIAIMTYTKNDTFSKKLIASMKESAQAN
GIEVDIYNANDSSDTQIAQVKKVAKMNYAAILIRPVEANNSQEITAIAGKTPIIFYNLQPEK
SVLKPNQNIYVGPNEELAGKLQANYVLKQTSDKDTINVAILRGPSGLAAEKRTSAFQDT
LRKAGKNVNYVFLDNGKWDFKTAQHLTELFLKTNQPCDALVANNDDMALGAIAAFKQ
AKRKLPIVCGVDLSAAGKKSIQNNQLAFSAYQPTDSQGKTIIKAALLLANKKKVTDLKGA
SDDDLYVYTSFEGVTAENVANY
>Methyl-accepting chemotaxis protein
MMKNKNFSIKAKLLLAIIPITCAMIVALVIIAYNVSAMMIKQNATQLLETSVAKQGTSIENW
MNENLSSFNAAKRAIEQAHPKDVELQKIVNGYYNFNTNSPDGLYIADQNGTVIKATKST
KNTTNAVNSPWFKEGLSHVNMTFGDPYVAGGHKVISATGLIDDGSNNLRILGADFTLD
RISTIVNSNISMEGAQAFLVNKDTMKVMADRNSTFVAQKVGSSDQPAIYKAAAKQIDRG
NYNTQTLGADFTAFHTIAGTNWVLVSYIPTKVVLANLSKLRLIMIAISIVALIIMCIIVERVT
NRTIKPVKKMTKVITAMADGDFTVDVDTKGNDEIAEMGKSVDKFIVYMRQMIKAISQIST
SLQTQATGSEKVAHEMNDASETQSKSMDELNMTVDQLSVSVNDIAKNASQLAEVVTD
TKDNSDLVKVKMEDTVSVSEHSREDMKQVVVALDNIQNSVTNLQESINKVGTASGEIV
NIVKIIGEIAHQTNLLSLNASIEAARAGEAGKGFAVVATEINNLAKSSANSVSQISDLIKEV
NDLVGDTVKQAGDSVNDINSSAVLINTAVDSFDTIFQNIHETSNMISQVVERIDGIEEVA
TNAAAISEEQAASSDEILATSETTLAQAKNISGNSKVVAKEAQQLAQTAKDLANQVKHF
KV
>Late competence protein ComGA, access of DNA to ComEA
MQSELRTILAKAIEQRASDIYFLPKDDYFSVGFSLAGKYCFQSKLPLVVGQQCLNYLKF
KADMSLSEQRRPQLGAWVYKWQGEVIFCRLSTVGDFLNRESLVIRLIYDGVKQQQGY
FFSDQWELVRTACQKRGMIIFAGPMGSGKTSAMYYFARQLQDKQVLCIEDPIEIYEPDF
LQLQVNDKALMGYADLLKVALRHHPDVFIIGEVRDEGTAQAVVRAALSGHLILTTVHAQ
SALGVVERLENLGLSRADLNQTVQLISYQRLIPTCDGSVKVLFDQVDNQTLLTAKPTQR
MSTEWGQRLEKCYQNKWISLETKKKYQAG
>Late competence protein ComGB, access of DNA to ComEA
MADLLAAGFSLKQAMSNLVILNNKQGSALTEILAKLEEGQQLSLALKDYVSQATYYQLLI
AERHGQLDESIAQLGKLLKRRIAQRAKLQSLLIYPLLLLGLLVLIIVAVHVWLRPALAQFG
MMPNPAGQINWVKLSGYCLAVILVIVGVYLGKLFWWWKKQPTLTKHHWYSQLPLIGKF
YRQYSYYYLSFNLGLLLQSGMDFQQICYFLGDFDKQSLLYQLGLQLKLWLEKGDELRG
FIALYPFIPPELVFFFSKGQTQAELSKEILLYSQLAYQKLLRQLDNLINLVQPLLFLVIAAIII
GTYLAILLPLYRNMGGLYR
>Phage integrase (Site-specific recombinase)
MIRVATYTRVSTLEQAEEGYSINEQQDKLEKYCELKDWTITQRYSDPGFSGSNIKRPGI
SELVAAAKQGDFDLVLVYKLDRLSRSQKDTLYLIEDVFQANQVEFVSLSENFDTSTPFG
KAMIGILSVFAQLEREQIKERMTMGKIGRAKSGKAMGWSKIPFGYNLVDDVYEVDEFQ
ATVVKRIFSDYLSGESPATIAKVLNGEGHLGKGKPWSWKTVKDVLMNVIYTGYISFKGE
LYPGMHKPIIDMETYKKTQTQIEIRRVNADNPRPFRAKYMLSGLLKCHYCGATLRIAVS
VNQRTKARSYRYNCPSSHPRRASTYKTKAVCPFKLVYKDEMEKRVIAEIEKLPKNFKQ
EKESVVDLDAIKKQIKLIKAKQSKLMDLYLVDGIDMEELNRRNGEFNKQIKGLEKELSST
PSKLDVSEMLNRAKNISELTYDEQKKLVKLLINEIEVANNSIKIHWRF
>hypothetical protein

MPFTQTEEPETEAASSSSSSSSEAELESEASAASDSKQKSSSKKESDDTNTRNELDTA
VNFINNQNDAVKASIGQNNKSIVLTVNDSVAEVAKNAANDQTNQDNWIKFTGEINKLTE
SMLTGKYNVKVPFIIQAPDGTKLMESKHGYTVYDIIDTE
>hypothetical protein
MCMIINYRGIDVSTSFYNDACNQVLEFKIFLAKKYTVPLSLISWEHVYEYMKELVKNQQL
IVQTTNATIKFSGLTKFIDNRFLIIVNTSPKINDSRLRFTLTHELMHIILHHSELDLGFKIKLQ
PRFVNITEDERELQADFAANDLLMSSWLIKDRMSRGYSFERLKDSFQVSWSAMQTRL
HNYLVFEIGIDFDDARKLTNQYRYYNDKTICSYVTARDEFIHYFVDFCEIPCPPSIEDIES
ETMNLLPVGGYLNHLGIELIKNDIENVLLNQLTDLAV
>Phage transcriptional regulator, Cro/CI family
MNLRDKIKELAAKRKISVAELERILGFGNGSISKWNKQSPSTEKLKQVADYFNVSLDYL
VGRDVKNEIDSNAYFRMDSDNLSDDEIEELKKQLKFAEALALQNIRNKKE
>hypothetical protein
MKIYDTVKRLSSEKNISVYRLEKDLGLANGAISKWNKQIPNALRLQEVADYLGVTSSYIL
NEARKDDNYDE
>antirepressor, putative
MMNNLVIMHDQQAVTTSLILAEAFDKQHKHVIEAIEKKISTAENSALLKTMFVEDCYTAS
NGKQNKMYYLNRDGFTFIAMGFTGRKADEFKLKYIDAFNKMETYIKDNSAKLSTDKRL
EIMEENAKTRKANLLYKIAMATKSEFAKESLLANAGEVLTGQKAIPVMRKEEFSATTIGN
LLGISANKVGRIANELGLKAEQPGQNEYGRWGIDKASNGKEVPNWVYFKPAIAEIEKYL
GR
>hypothetical protein
MNRIKEVREMHGLTQKQLAEEVGVTKATISNWEINFNPNVQMVKKVSQVLSVSPLYLM
GLSDSLYDNFELLKHLLTKETTMITSDEDSMIQIKRILSEN
>hypothetical protein
MNNLAQASYSRLYRAITMKEFYVSDLKPTYVYWTMDEDYGLHKQKKKEDPEVWPVLT
EHEFEELRYLLSEAYAYSSLLSHVFEDLLDETDGVSFTERGFDKVANILIDGEFYVEVRH
DVEGD
>hypothetical protein
MDELKPYKSHFKKLERGVNNKPIIPMETYKGFLYYGQANVMVKTKVDVCVNQVESNA
DKLAMVDDRYFIDKLYSHQITLETKTFKNIFQVIPKGLKTNDYLKITLNPDSIIMKPHESLR
ELMEVKYNVDLGIEEPFMFAVNPVYVRDVLMLFSSLKIKNFKLKHTGSNLRPVMFEADD
IQVVVAPVRLPNM
>FIG00529064: hypothetical protein
MGLGKTFVGSEKLHRLGASQNVLICQRSLIPMWLNHFKTYYDYQVYDGSQPKQLEAV
LENKLDNFILVINYDLVFRRPKLKRAFKDYTLMLDESSMIQNDRSKRSRFILGMKPTNVI
LLSGTPTSGKYENLWSQAKLLGWDITKTAYNKTYINWMDIYLGPKKLKVVDRSNPYKN
VQRLKRKFREHGAVFMKTEEVFELPEQNFIPIKVKPSKIYQHFLEDKWVAVNATTELLG
DTTLTYRLGLRQLSGVYSKEKLQAMKDLIESTDDRLIVFYNFTSELNELKKLTKKPTSIV
NGEVKDLKSYEQEENSVTFIQYQAGSMGLNLQKANKIVYFTPPEKSDLYEQSKKRIHRI
GQERACFYYNLIVEGSVEERIYKALKERRDYTDELFIDEYKR
>hypothetical protein
MTGTEAEKLLKEYVKQKKLFNKKKAKSSKITAPLVNGISIENNWHIKIRRTW
>hypothetical protein
MNELTEFEQANLVLMTKLAQSEQLVREVKEMQESYKRELKDAMEKYNIKSIDNDLVKV
TVVPESESTTVDLKTFREKEPDEYDGLLADYPKVVKRSSYVRIKVK

>hypothetical protein
MAGEKRFEKQVERYLESKHIYHNTTRKQDKVLPERGWFFKVWGGGYQSAGVPDIIAN
VNGYFLAIELKDVNGKPSALQKRNVELINQTNGVAVILYPQGFEQFKEIIEVMLACNSAI
PGYLPMRGIRLSTNLGILRD
>FIG00527559: hypothetical protein
MSYPLVTDKHINEAIKLELLSKKALAVLPSTGKYEVTLSNQDYIGFIDMLVKNEDGTFDI
WDFKHSNSIENYMKSAQLHLYKYYLEQTGRKVKHLRYLFIPKVRLKQKKTESEYQFRQ
RLKRAIDKDYQVTIKEVEYNPDKVIEWLASVKHALEAKEFPLNEDDPFWKYSPYRDYVE
KGLTYNMVTLPENKRRKLTEVTRRKLWIYGAPFTGKTTFANEFPDPLMLNTDGNTSYV
DAPVVAIKDEVTTTGRRTNRKFAWQVFKEYIDELEKKDNDFKTIVVDLVEDLYESCRLY
MYDKLGIEHESDDPFKAWDEVRTEFLSNMRRVTNLDYENIVLISHEDSTKDILSRAGAK
LTTFKPNIQEKVANKIAGMVDLVCRVVVEDNEHILSFKTTTTQFGGGRLSDLSVDSINLN
YDELVDVYNTSLADSNKPKKRKQTKLEEPENEEPEQETPEKEEPTEETESVEEVEEKP
RRRRKRRESTESVETAEDVETESAETSEPVEDEVEEEKPRRTRKRRARRTTEED
>FIG01253079: hypothetical protein
MTNWSKFDKEFDTKGLQEEIKEAAENDRGGYEDVPDGDYEVSVEKLELKESRKGDPM
LSVWFNILDGKFTNSKIFMNQVLTKGFQIHNANELLRSLGTDLDIKFDSFAQYAELIDDV
FEQIDADGLEYLLDYGTNKKGYNTFKIKEVYSD
>DNA polymerase( EC:2.7.7.7 )
MIFYDYEVFKEDWLVVIKDTATQSETVIVNDSDKLKKFNEAHEDDLWVGYNNNHYDQ
WIHKSILCDLDPYETSKAIIEQNIPGWKISNLFRRIKMFNYDVMIRGDGGLKSLEGFMGS
NIKESDVDFNIQRKLTQVEIDETIKYCRHDVEETMEVFLHRKSDFEAQMALAKLPHDHL
ELKYLSKTKAQMAAIVLEAQKQSYSDEFDLQFPDNLEINKYTEVLEFYRNPDNRNYSQK
LTTMVAGVPHVYGWGGVHGAKEKYIGDGYFINMDVTSLYPSLMIQYGLLSRSIKDPRK
FKEIYDTRVKYKHEGNPLQAPLKIVINSTYGAMKDVSNPLYDPRQANLVCLYGQLFLTD
LMEKLEPVCEIIQSNTDGVLVKLRNEDDFDLIDDIAWDWEQRTHLSLEFSEFKRVYQKD
VNNYAMVGTDGHVKTKGAWLKKLNPLDNDLPIINEALKRFLVDKVPPQQTINDCDDLE
QFQQVCKISSKYKTLVHNGEVLNERCVRAYASLDADDGGLFKQKYSTNRLEKFANSPK
NCFIYNGKVKGAKTPRKLDRRYYIDMVNDRIRAFGMIPPGEWYQGDLWEHTGGLMM
>phage / plasmid primase, P4 family
MIYRGYVKSNKKGAIEPYKDVEKLRTLEEVKEFDSYGGVLADDTILIDVDSIEDSEKLLD
LIEGEDLKCIVRYTNHGAHFLFKNHTPDLRNGTKLMLPIGIEVDVKFGYNPSFEPLKIKG
KEREIGWDYPDYQEVPSYLLPLSKSTSRRVKTRFNEMRAGDGRNQSLFEYILTLQGEG
LAKDDIKLTINMINKYILEQPLSQQEIDTITRDEAFKKQVFFKNKQLQVDKFSEYLKSEYHI
KLINGQLGIYENGAYVLGRKLIEAKMVQKIPTIKRSMRQEILSYIELTTLENHTSQDANFIA
FKNGLYDIKNRKLVHYSPEHIITNILPWDYNPNAKSKLVVDTLMKIACNDKKIYDLLFEIIG
YCLFRRNELGKFFILTGSGSNGKSTYLDIIRTALGLLNISSLDMSELNERFKTAEISGKLA
NIGDDISDGYINDTSIVKKLVTGEALTVERKGTDPFMFENYSKLLFSANSIPRLGKGSDT
KALNRRMMIVPFNATFSKDDPDYQPYIKYDLRQQDAMEFVLYQGVQYLHKILEEKQFT
VPEKAKQELAKYERENNPILNFFDDLEETDYLRQPVKEVYKKYTEYCYRNGLNPVSNIS
FSKQITTNFELDSKSVRIKGKVTRIYTKVEES
>hypothetical protein
MTKNKMEDLNNLLFEQLERLNDESLDLEQELKRAKAISDVSDKLIQSADLSFKVMKLRA
DMTGNVETPEMLEVKKIETKND
>Bacteriophage phi 1.45 protein-like protein

MKQKTIEWLRETVPGKPWKEVYKLYQETFPDEPRKFSYIKTVCYKNGITNGLNGQFKK
NRVPLNKGMKGWCAPGSERTWFKKGQSSPNERTIGSEYDSAGYLLVKIKNEGRRTDK
WRPKHVLVWEKANGPVPEGHIVVFLDGNHKNFNLDNLACIPRGINAILNRSRWRSDNA
DITKVRITQAELKRAIKEATRHD
>hypothetical protein
MNIDISMPFPLAYLDWCKEKGIKTLKDTENNALIMEFGKLYTQKGGKIETYEILIARGDKII
>Site-specific recombinases, DNA invertase Pin homologs
MIERAGYRVDKWLADEGVSGTIDWTKREINVAIEEASEGDTIIVAELSRLGRSLKQILEIV
ELCQKKKVNIIMIREGIQTNNDSPVNKLLLSILGSLAEMERNLISQRTKDALALKKKNGVV
LGRPVGRTTPVEKMKLYPKSEQILEWHKDGVSYTEIGRRLGVHRVTVSNFISLLESGE
N
>hypothetical protein
MKTYYVLGTKDSKFYIEKVNQRGKLTLTDNLGEAVKAMSFSKKYDADKEAEKTGLIVYE
LTVKEVE
>hypothetical protein
MMKNLNDFHTALNSLWTDVVLTRKNFNDSYLPLSASDRVKFNDDFRELAAKLELMMM
KTEEIKNEEESK
>hypothetical protein
MTTGVYLITDYKGGILNVDGTLQGAYFKAKRLVEKANFPVNHIRILAMGSGQTFKYQDL
MNREWH
>hypothetical protein
MKVKAKGSPSEMLGILAKSIISIAKQFKLTDDEILEIVKVTLKDED
>Phage DNA-binding protein
MKDLSEFLLKEPTVGVPFYGFVKTETSLQSLAIVPNRKSNNYDLLNLKKGEVVKNGTLE
EISDYPGIRKKFSRLINYDEIDKFADLIAEEKAKQIMSQYVEYVFSKEVSTSYSVLSRKIIN
RFTDIGIVKPDNTKTQFMKVTEELGELAEGINKDKPEQVKDSLGDVLVTLILLAEDLNLN
LLDCLNSAWNEIKNRKGEVKDGSFVKEEDLNV
>hypothetical protein
MCRTRIKIIRGFIDSEGIKELEDTVNMFIADPVNKVTKVNSVEFEVYYDACLLAIINYELGE
>hypothetical protein
MTKYNDIGLTREDLIFELKQKYGSSWELDSPFEDETLQMFRNFPRLSDGPNNFILRNSR
EIYIRLYKQGLSVNEIADRLNRKPNTVRSVLIGSGVLPKFSKWHNIKVLDKKGKLVQMG
SIRRISQDFKIHPTALKRKIKDGTADKNGYRYIEVE
>hypothetical protein
MGRRKQCYIAYNKADEFLALGTVDEVASVLGITVNAVKSKMSHWKRGDIKNPRIKIYKV
EEDESC
>hypothetical protein
MLKIKLNDETQVINVDNLNYQLQELRNNKKPTWVTLGYYPSLKSALKAYVLDHRNQNIE
SVNDYLTMIDDKMAEVKEVLGE
>hypothetical protein
MKQSIFQASEAVIKTADRQARKHHLNTLDILERSFVSAMLGLNVTPEEFALVINLILENYK
KEFKQIEGIKR
>hypothetical protein
MKAQSKDYKTHVMNSVKKFLELKLDEFGISKTELVRQLNAQGYPISYATVNGYITNRNL
ITGSNLLMLADFFETSTDEILGAYDL
>hypothetical protein

MSDYQEVYDYFKHYKTIHEAITRLEERLETIKSKYELQAQKYSAEPSGKGGKPASYETM
LIQTERLYKQLEKLYKEADAIKATFDKQSILVDDQLAMRVIEMYFFDGMSHQEVAINLHL
APSRVSQLKRKGLECINLN
>hypothetical protein
MNEVAERVVLMREEEFDEILQGLDALPFDIKHTTASNGKFVSSVITISKSKVEASLKAMD
YNQLQDKDVTYRAYVEWVD
>Phage-associated homing endonuclease
MDKYLKFYHSKEWKLARTQALARDHYLCQVCLKRGLVKPANTVHHIIPIKDDFKKRLDL
NNLETICLECHNQEHQERGSNEPERYKKLKARKRKAYVFKANPDYKL
>Phage Terminase Small Subunit
MPMPAKSAMMQLTQKNPNRRTKAELEKRVRNEKKLQMADDQLQPPAWLEPGAKKVF
NRVVEVMTPTQILNNADLEMLAVYADTYYDYISYKRKIRKTGLVADNGKPNPFIAQKRN
AAASLSRLANSLGLTPAARASLAIHMEDEDDDDDF
>Phage terminase, large subunit
MTMTSNLLKASQTEIERWWDDYRQQRLGWAYLKEPSPVVLTTMYAQMVVEGSIPVSK
PNMLACKRHLHDLEIQGTDDFPWVFDEEKAHRPIRFIEQKCSPSKGARSQLVMQPWQ
HFVIGSLFGWVHRDTGIRRYREGLVFVGRKNGKTTLQSGLADYMAGFDGERGANIYFL
ANSQSQANILFEESKSMIQSSEFLSKRFKVMSRQIIHTKSKSKMVAMSSEKKDKDGEN
VHFAVFDEIHEYQDTDLISLMKQARGTRTQPLIMFITTAGKVLDGPLMDYIEQGEEVLG
DYDSHLDDRFFYYLARLDNPDEADNPEMWVKANPNMCLMEIPNLISDYKSDKRIPAEY
ANWLTKQFNIFSSTDELSFVTPEIINKNTRTIPSASILGRSCVGGYDLSETEDFTSAGLEF
PLDDGSVFWLNKSWIPESRFKLDKNQDRLREWEKQGYLEIIPGDYVQYEHVYDWFVE
QSKKYKIQQINYDPNKALFLNQALKDKGFKTEVVKQGFVTLGGPMQNIKELLLDGKLIT
NNNRMFRWYLNNVKLVTDRNNNWMPTKQSRNRKIDGFAAFLTAHATVVKNLNTNPTP
GRVKFISFKDL
>Phage portal protein
MNLAMGNQFPISLGDSQLQTNETVFSVITQLANAMASMPLKLYRNYEEVTDSDLAMM
MKYHPNPSMTSFSFIQKIETDRNEYGNAYVLIERDQYWQPINLWPIPFNNVTVMQNQD
DNSIWYQVTGIDGNMLVSEANMLHFKHVSGSSRLLGISPLKTLKNALEFDLAVQKFSLS
EMSKTDSFKVTYATNVDDEQRDNIIANFRMYKQENGGVLFQEPGVEIDEIQRDFVSSD
LLNTEKITDIRIANAFNVPLAFLNQVNITNNEDLMSQFVQRTLIPIVRQYEQELTNKLLTEE
QLRQGLYFKFNVNSLMRGNVQARTAYYQALRRAGILTTNDIRALEDLPLVKDEFADKLF
VSGDLYPIDMDPSQRKGVSANGAKTEESSQVLGNEQERRQRW
>Prophage Clp protease-like protein
MSKSDDNVGEISIYGEIVSEKWFDEETSATSFKSDLDELGDVDTINLHINSPGGSVFEGI
AIHNMLKMHKATVNVYVDGLAASIASVIAMSGDTIFMPENSMLMIHNPWTFAMGNAKE
LRKQADDLDRIAQASIKTYLSKSNGKIDEETLVKLLDEETWLSAQEAVDYGLADEVLEA
NQAVASLPGEFLERYKRVPSQLRKKSAPASSINRQKLITQANHKIAYVNNTLGGINKWV
>Phage major capsid protein #Fam0068
MGVTLYEKKQNLGTLGQQLKKINEEIAMKAGDPTVADKDLMQLQEQSQSLETRYNML
KEQVEREETAQRAKFTETQAPVMTAEEKLIHAKAEFYRGQKLSADYKQVLGDDDSTTN
GSKLLPVTIANDIIAEPTENNPLRDDELVTAVTNLERPRIDVTIDDDSFVNDQEVAKEIKL
KGDSVKFGRNKTKLKVAISEAILNGTDTNLVEHVNSQLQAGLAKKEKTVAFAETPKSGE
EEMSFYSTQNNIKKVSGDTLFDAITQAAGDIADAFQSDIKVYMTRPDYLKMIKELSNGA
VSLFGKAPEEILGYPVRFTELAKKPVVGNFKYAQLNYEISSALYEQWKDYDKGINNFQL
TAWFDHKILLASAFRIADVTSK

>Phage related protein, possible DNA packing
MDLIHLKQAIAIDTDVFDDVLTVSMKAAERYIQNAVGTEYLEFYQDNDLYDLAVIQLTDH
YFKTRSATTEKGTEVTLYGLSSIILQLKPLYRIYAQKQETSEES
>Phage head-tail joining protein
MIAETGDLTEVIKIIEPGVEDVDEYGDPVSTEKVIYPKLFAMLRSKNANDVEKNLSTLST
SAQFVIRHRLSNEPEITTDMQVVHRNEIYKIDDFNIDTQFKKFDVIICHKSP
>hypothetical protein
MTESFKIDENITVKLQELGNKANRITRKAVRESAPGFAEILKANTPYESIADRSWKAQRR
MDRLTGKKSKFKHMRDDVQFSGVTQTGQVNVGFGKDTYWRVHFVELGTVNQKANPF
IERAIEAGAGPYTESVSNVIRSELGL
>Phage head-tail joining protein
MRISAIELGNLVKTMQDFTNQNTFILREVPQTILEKKVLPFAQITFLGNEPYDYASNVKQ
GELSESQVDFYVKDNKSAENLICNFEKALKNTGFEVFFCDIDTNFEYDFQVVHMRVRR
YQRI
>Phage major tail protein
MTTTNFTIGTIGFSRILFGIYADNAETVTDVVTIDETNGGAIELKTSGFQGQLNTTYASNI
AYYISDAGTGTGKVEITALELSKDVATKVLGDKEENGILSTYSNVTQPYVSVIAEAEDLQ
GKKMWIGIAKAKFGTPDGDELKTSEDKGKTPNHITLTGSAITRRSDKLVKCKASETSGA
TFSAFATKMFPGFSDISKFDRTTAVSVGGAG
>hypothetical protein
MISIKLYNPETGETNYYEQKKVPFGKIKKIIEFNKAQATKETELDLLQRKLEENGSLSKTE
ETKYLNLQTEVEFGMVDPMVELVVDLFNHPKVTKEAIYDGLDLSDGIQGLSDILTQAMG
GKTDSAKK
>Phage tail length tape-measure protein
MAVGKPLGSMIVTLGLDAAKFTDGLKSAQNQMKLANSEMKANLAAIAATGTEYDKASA
KVDSLSKVMEANQRRIDALREAYDKQVKTQGEYSNAAVRTASRINDAVRLQASYERQL
QSARVALNEATRGTNDLRQSLDALERTTRATVSGLQAQNKTNQANLAQYRGLKQALQ
QYNQLLEAEEAKLKDLINLKGKDAQETRTQSVQVAELRSKMQQSQAQFDSLSDKYGK
MSTAQAIARDRGAKFKSEMTDLGDKVSKTGEKLKGFSTIAGFGLAAGLKVSIDSLAELK
NTLNEIKNLAVTGGESVSEATKNVAKIQKDALEYSNKYGVSINKITDGYEDLIKRGYTTE
QALGAMKSELQASLASGDDFNDVVTVSSQTLEAFGMKVNNTAGMLKNTKIVTNELAYA
ADATSTGFSELGIGMSYVGSTAHTAHIALSETASAMGVLSNNGLEAEKAGTGLRNVINN
LVNGVKKIDDKDSVLNKLGVTRDELVDSEGHLRNLSDVFEVLNSKMKGMQPVERAAIF
KSLFGTTGQQAGIILADNVKELDELNKKVKEAPKNNYIGALSEKNLKSAQNQLKIFKESA
SNAGMALAEVLLPSISRVAAQLSKMMQWVAGLPAPMRTLISVTTLFAVAIGPLTIALGAL
IKSIGLIGAALSALSLGALLNPITLTIAAIAALGVGFTLAYKNIKPFRDTVNGVIKTIKNLWTV
LTQGKDSKSKAQANLDLAQIFPQGALIKINQLGNNLQSVAKKIRTAIQAIKDTFTIFTKGD
NSKTKANAYTNLLNILPQSKAQTLINTANNIKKAFQGMFAVIKGQNKQGSNLLADVFPK
PLVTFITTTIKILKADFKSLGNTISSVFGAIGKTLGTVFGSIGSSFLKLLKSLFSWFGKYGP
DLLAAFRNIFGLIAGVVAVKFTAILTTATVIMKVLGTVIKGALDVIRNVFSAVWGSIENIVG
GVLKAIGGIFEVFAGAFTGNWKLVWQGVKDIFSGIWNSFKGIVGGVINVIIGIVNSGIDGI
NWILSKFNAKKVGHISPVKWATGTTRYYPSGLPETQLAMVNDGGKREAIVYPNGQIGM
FRDMNVTTILPKGSHVINGDDTERLGLGANPVTHYYAKGTVDFSKIWNGIKSGASKLW
DDLTDGLKFAKNIIAHPIKALQGAFESHLNLGGQTSFVIDTVKGIGSFIVNAIKNGVVSEIK
KWIGENEASSDATSPKPTGDHKHWMKQAGIPESWFSLVNWIVTRESGWNPKATNPS
SGAYGLPQSLPGNKMASAGKDWRTNPITQLKWMYSYIKGRYGNAANAQRFWQAHH

WYANGGFVTQEQIAHIAEGNRPEAIIPLTNRTRAMQILTQVRDKYGLSAGNVVLSGGE
QDNTDLSSFERKFDTVISLLGQIAGLGAEQVNALKVMKPSQNFDKNKFYQDMFRDQTI
SNYMNM
>Phage tail fiber
MEKLYLKIGNQDEFDICEKVQGLHFLGDDSIPVTTNQFLEISGTDGSQFQYATFGKYQV
VANFFLEFKSWEDYKLAKHQINRIFSTKKLIRIRTNVESAIVRYVYPNFPEIKPISDGANS
STFSVNFDNPSGYRYSIDRSDKLTNVQYGMYLLDDIYPEYHFTNKSFRVYNPSDIAIDP
YLGKHDLKIISKFSGSSLKITNTTNGTSWSYNKSSNGNETVLLDGIVTTVNGNPATVNTD
YGHIVLNTGWNDIVVSGTNSNDITFSFPFIYI
>Secretory antigen SsaA-like protein transposon-related
MFQGKILVQGVNRAEKEPLNLFDPKSVQIQWGVNQTWSLQLTAYNDGSLAYQMLESE
ASIFLDNQEYIIKQVADDSSSGLDSIQVTATHVYFEVQKIRKYKEYVDPEDKDKQTDVKV
LKDDTDSAKSDDGDNAKTDTSEKTEGNTTTKVTTKTTDETQQDNQNQVIYSIQDVLDH
WLKDNKLGFTYEVIGSFEKKELEELQDGTGADMLSKISDTWDNAIIYPDNRKIRVYSTD
KFNLNRGNRIDYLNNASEIKFSTDSTSLTNMAYCIGGKYSVETTTETTTETTTETSSNSN
TQEYYYFAPFIYRDEESIKKYGEHPAEPIEDGRFKDKNAMIEYVKTKLQPNPSLSIDVTT
TTGVRPIAGDVVHVMIKSQDVSTNFTLTGFTWYPYSYPVDNPTSITLNSNVQNILDYQN
SRQRQFSKAISELKSSTNEAINNSNSFNEFGGNQQLKTWLNDFVGG
>Phage lysin, glycosyl hydrolase, family 25
MNIWEWIDKLTKALQKLDSRITVIESVLFDDKTNPDTPKPQHIGKIIDVSEWQGAIDWNK
VIADDVTLSIIRVQDGSTHQDLKYMENIQKCISAGGKYAVYAYFRGASTADAQQEARDF
YNRTQRVVAGKQQPVFYALDIESVEMGGAASQMRAGVEAYMSQLNALGVPDNKIVLY
IANHLYDKFNLNVARPGAIWIPSYGQNDGTLANSLKPTHPYDLHQFTSKGSVKGISGNV
DMSAEPSEKFKELIFSA
>hypothetical protein
MNKLSSLLNESDYQEYDNFGNVVTVSYEAYLDDKLELDKDNFNKYYQQLQIILDKLVKF
KHDNVSEQYLKGGEN
>hypothetical protein
MAVANNQYINFDLLRYQNEVLDITNKFKGRVGDTQDYIKLFVTSNSYPVDLRGMKLLFG
GVDPKQQAHRHYLDFRADQKTDNLEQGRCTVYFDENTFNYEGVWEQAYFKFIDAEG
NTVSTVDMVLKVLDDRFYAAVGQSENVVIAEFKKLVEKLTDKEQEAEKLFNSLSTDAKA
KFQAVYDDYKKTIEEAYDAIFNAETGLKVNYTRLQEMAQHIQETLRQAQFHDRPFQFD
TVAIMKKCLELQDGDLVITSGWDSKDDGHGNMWQVRAKKHDETPDEVNVIALQSGYV
AERNLSMISADSLEDIMYGYSIKIVHNQKDYPKPTVFYYENALGTETGGLGSGSFDETL
TKLVPCEAEYTDNNSIVVRIPRNFYMDAKPYYKYGDWYLGSGNKTIKISLGNIDDGAAK
TGDGKGSSYLSHTTGYFNYPTAPSDLRAVYVNDTTERLIWE
>Phage major tail protein
MKFYIYRGIGSSGELTKIAEVTDAKTYTDTGLQPKTIYRYAVSAYNGLRESAKSNIITVTT
AEIPVATITLAIDKTSLEVGGTAKITVTVTPPNQTSGTPTLASTNTKVATVDNSGNVRAV
GVGTTTITASLAGKTSNMLTLNVYEALVNVSNLTSSNVTPNSVTLSWT
>hypothetical protein
MTYRVYNGTQLLGDVNTKSTGINGLLPATSYQLGVTNYNGNRESSKATINVKTRGIRLII
PTSLTVNSTITLNYQEYSLGLVPIGTEPKGMFGGGNKRNIQAKVISVDNGRSTVELLEPL
ESSSTVSVGGTNLLRKASNEWKTWTGTGWANNTVDAGPINLDAGTYTFRAEIDNTGN
TQRGVHTEVYVDSEKKFLDQIVKSYTQTFVNAGSKGYVTTTFTIDKPLTDVRIHPHASV

NSPIPSTVTIKWRREKLEKGSVATPWTPAFEDTQYIVDNTQMNRLQDGSFAAFNGYRA
IYFR
>hypothetical protein
MLIYCFDDAKRFAYTDVIADDVPVPVNATTVPPVNTDGSGMYAPTWDGTKWIPMTEEE
FEKEFAQPQRPKGVSEDLIPTEAEQLQAQQVMMIASLTKQVQDLQGAVKTLVLQNAVN
KEEK
>hypothetical protein
MYTYSFVNSFYKMGLFTKDDVKLLLEVKQFSQDDYNKMFPEDSSLTA
>hypothetical protein
MAVISVLFSSVYWLIRHGAKALNNAITIGTYPLQQQFKELTSTIKQLNNNFEEQRKSLKK
LEDEVEEHDKTILLHDEKIKRLEERK
>prophage Lp2 protein 58
MKKVFLDKDGKLNRKTVTSLVLLLIVLVQQLCAIFNLKFTGDVGQIMDLVNTLLTIGGILG
LVDGVTVDTETVKSIDKTANEALKLASQHQGKGGKHD
>Membrane-bound lytic murein transglycosylase D precursor (EC 3.2.1.-)
MTKLNRKWLLSLTILVGLFLMPVTVNAARYYGVDGSRYQGNTLKKVTPEDSFAISQIGG
YYNGTFIPQTTYQSQVASGIAMGLRMHTYIYMETGSNQAQTKQMLDYYLSKVQTPKQ
SIVALDYESGASADREANTDNVLYGLRRVKEAGYTPVLYSYKPYILSHLNRQRITAEFP
NCLWVAAYRDYSVMTRPDYNYFPSMDGINMWQFTSTAIAGGYDYNVDLLGITLNGYK
NGNVEHPRSETKAIEQGQKADNTSKSAIQVGNTVRVKFGVKHWANGVGMPSWVQSN
TYKVQEVSGSKLLLGGIMSWINASDVEIISVANSNPSVTGTIYYVVKSGDTLGGIASRYG
TTWQRLQALNGLSNPNRIYPGQRLKVTGNVYAQRTYTVRYGDTLSSIASRYGVNMYTL
AHKNRISNINLIYPGQKLNI
>hypothetical protein
MVSEAQKKANEKWKAANKEKQKIYGYRSQAKKFISEFADKEDLKTLEELIRLRYEKIRG
EGNV
>prophage ps3 protein 01
MYDVFKIVNWLRVKNNADMKRNENVEELTQMKAMKLLYYIQAASLVVTGKRMFNENL
VAWKYGPVVERVHEKYRGQRAIVGEITDADLDDYSELEQDTNTADILNSIYDIYGYSSA
YDLMRQTHKEKPWQETKQSEIISDEAIKNYYSGVFEVEA
>hypothetical protein
MDQLKALVGRKVEAKAGEGEIIGVDEFSIRAEVKIGDEIKHWPLPTVLDHLI
>Late competence protein ComGC, access of DNA to ComEA, FIG007487
MAVVLFIISLLLLIVIPNVSNQKKHAGSISDEALKTELTTQRQLYLSDNPEATSVSLEELQ
AANYLTANQVKQIREHKLDEG
>hypothetical protein
MKVKYNKQAFTLIEMIIVLALVSVLLLIGFFPVKTMIASYQEQQFWRTFNQAWSKTAITA
ERKGCVAEVAFHSKEIEFSDYAENTRGQKQHSQIITLPPSLQVLTKFAYADFYKNGGTK
FMKVVFYSKLHQVKYTLIPQVGYGGRYRLEVTK
>hypothetical protein
MEGGIGLKLRSKQGFILADSLVGLLIGIVGMSIFIGVQTQFEPLRKDAVREAINYQSLADK
SEQQYLAGKQGGLDKIVVNGVTVQKEA
>hypothetical protein
METIIALVVTSIGLSLFMGFMLNLKSNLSSVRQDDLAQVEQLVSVLKTRGLSLEYKGTSF
GNVVFYSPPRQANYQLKYERNKIWLNQDGQGYMPLLFDVEAFKTKWHDENYTLDLEV
KLHGKVIKREVVFAKAKPPKEKQTPKNEG

>Adenine-specific methyltransferase (EC 2.1.1.72)
MGIEKIEQTFTSLKEAVALLQKELQVSYLDALIETGDNLLAKEVHVEDGLPTAEIKTRLEE
LYANAQLAGLTTEQRRQVMQLLLLQAYKQEKIQANHQMTPDTIGFLVEYLLEKVLPKKQ
GLALLDLTVGTGNLLTAILSKLNQNGWPKIQAYGVDNDDTLITLASMATQIEQISVDLFH
QDALDSLLLPKVDVVVSDLPVGYYPLDERVTQFKTRAKKGHSYVHHLLIEQAVNQLKP
AGFGIFVVPKGLFESPEAKGLLDYIQANAYLQGMLNLPQDLFASKRGQKAILLLQKHGS
GAKQAKQILLGDFPSFKHKEQFQAFLGQIDYWERNSLMK
>D-serine/D-alanine/glycine transporter
MAQEEYARKLKSRHIQLIALGGTIGTGLFLGSGKSIHLAGPSIVLAYLLTGGICFLLMRAM
GELLLSDLKTHSFIDFIAKYLGKEVGFVAGWTYWICWITIAMADVTATGMYIQYWFPDIP
RWLPGLIVIAILLGLNLITVGLFGETEFWFALVKILAIVALIMAGIVMVIVGFKTPLGHASLG
NLVNYGGFFPKGIKGFLMSFQMVTFGFIGIEMIGVTASEAENPTKVIPKAINEIPTRIIIFYV
GSLLALMSIYPWTKISTTDSPFVQVFGDIGISAAAGVINFVVLTAAASACNSSLFSTGRM
LFSLNYGKEGKFARKMSSLSRTQVPANGLIFSAVVIALAVIAEFIVPNASVVFTFISSVAT
VCFLFIWGAIVLAHLRYRKSLQKQGQKNLVSFKLPLFPVANYLILLFLAFVGVIMCFDQT
TFYALLAALIWFGGLYMALRLQRTK
>Hypothetical DUF1027 domain protein
MNSEKKVRREAQLARKAETIKAFASEIRMQNETEMTIDGHPYVLLLNYRQGFQIEKLQE
RFSQILTKYDYIVGDWGYDQLRLHGFYEVGSKKGSPSQNIDRLEDYLYEYCNFGCAYF
VLQNLDVQKPEPIPASAKKKPANRKARERRYKDKAYLEERVYPLKKKRPRKFKPARAN
NVKKLGKKRHFTIRQKG
>Hypothetical NagD-like phosphatase
MRDYKGYLIDLDGTIYKGKIKIPAAKRFVEQLQARKKDFLFVTNNSTKSPEEVATNLATN
FDIHVSPENIYTTALATADYVADLDEKKRTVYPIGELGLKQALLDKGFRFEEYQPDYVIV
GLDYDVTYHKFELATLAIKRGAKFIGTNADTNLPNERGLVPGAGSVIALVECSTQQKAT
YIGKPEPIIMEKALKKIGLSKDEVIMVGDNYMTDIKAGINFGIDTLLVYTGVSTKEQVAKM
AVKPTYELDSLDEWSFDE
>COG4478, integral membrane protein
MNKLKLTEGLLLVVFYLAIFSACIAITINFRGLYYYFVWHQDLGSKVGLSNHQLITTYQRL
LAYLNYPWVQQLDLGLPTSFNAMHHFTEVKALFFKNYLILAASFPVSCWWLKGLWQK
RRLFILITPCYYLLSLGVVVLTLMVTNFNKFFVTFHRLLFANDDWLFDPKLDPIINALPAS
YFLAAFSLFILLVFISLVGIIGIARYQLKHL
>DedA protein
MNIIQLFIDFVLHIDTHLVTIVNTFGSWTYLIVFAIIFIETGAVIMPFLPGDSLLFAACALSAN
PIYNLHTGIFVVLFWVAAVAGDSLNFFIGHTIGKSITSHPFFGRFIKEENITKTEKFFEKHG
GISITLARFMPIIRTFAPFVAGASGYKYSQFIRYNLVGATAWVALCCGGGHFFGNFPLVK
EHFSMIVLGIILVSLLPAVITFIKSRLGQKA
>Thioredoxin reductase (EC 1.8.1.9)
MTETAYYDITIIGGGPVGLFAATYARMRLAKTQIIESLDQLGGQVSTLFPAKKIYDIPGFY
EISGEDLIKNFEKQTLAFKPDIFLGETVLGFEKTPEGFLIKTSKRTTYSRSIIIAAGIGAFEP
RRLTLDNASQLEGKQLRYFVTDPTTFAGKDLAIAGGGDSAIDWALTLYPQAKSVHLIHR
RDKFRALESNLAKLKETKTNFETPYTIKELSPAATDKLTITLQKVRSQETKTITVDHLLVN
YGFVSDNKIISSWGLEMQNRAIKVNQAQETNIPGVYAIGDSAYHEGKIDLIATGLGEAPN
AVNNALLYVYPEKRQPAHSTQLIKAFNQAQKK
>Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8)

MYPQLNLAAVEGPVATFKTNKGDFAIKLFPKQAPKTVENFLGLVAKNYYDGIIFHRVIKD
FMIQGGDPTGTGMGGQSLWGAPFEDEFSNELFNIRGALSMANAGPATNGSQFFIVQN
QNMPAQMEEQMVAAGYPKEIIEAYRQGGTPWLDHRHTVFGQVVSGMEIVDEIANVET
GIQDKPVEDVVIKTIEISE
>General stress protein 13
MDYKIGMILEGKVTGIQAYGAFVSLDAQTQGLIHISECHHGYVDNIEHFLQVGQKVRVIII
DIDEYTGKISLSVRCLEKPFDFSKANKRVHYQHKKFWTNKHAQEGFAPIAERLDGWVK
QALVDIDAGK
>Small Subunit Ribosomal RNA; ssuRNA; SSU rRNA

>tRNA-Ala-TGC

>Large Subunit Ribosomal RNA; lsuRNA; LSU rRNA

>5S RNA

>tRNA-Asn-GTT

>tRNA-Glu-TTC

>tRNA-Val-TAC

>tRNA-Lys-TTT

>tRNA-Thr-TGT

>tRNA-Gly-GCC

>tRNA-Leu-TAA

>tRNA-Arg-ACG

>tRNA-Pro-TGG

>tRNA-Met-CAT

>tRNA-Met-CAT

>tRNA-Pseudo-TGA

>tRNA-Met-CAT

>tRNA-Asp-GTC

>tRNA-Gly-TCC

>tRNA-Ile-GAT

>tRNA-Ser-GCT

>tRNA-Glu-TTC

>tRNA-Met-CAT

>tRNA-Asp-GTC

>S-adenosylmethionine synthetase (EC 2.5.1.6)
MSERRLFTSESVSEGHPDKIADQISDAILDALLAKDADARVAVETTVTTGLVIVVGEVST
SAYVDIQKVVRQTIKDIGYTNGKYGFDGDNCAVLVSIDEQSPDIAQGVDASLETKSGQI
DPLDQIGAGDQGMMFGYATDETPEMIPLAVALSHKLMRKIAALRKSNELPYLGPDAKA
QVTIEYGEDNKPERVDTVVLSTQHDEDTKLDQIRTDVIEKVIKTVIPAELLDEKTKYFVNP
TGRFVIGGPQGDAGLTGRKIIVDTYGGAAHHGGGAFSGKDATKVDRSASYAARYIAKN
IVAAGLAKEVEVQIAYAIGVAQPVSIAVNTFNTSDIANDKLVEAIRKTFDLRPAGIIKMLDL
KRPIYKQTATYGHFGRTDIDLPWEKTDKVAKLQELLK
>Membrane-bound lytic murein transglycosylase D precursor (EC 3.2.1.-)
MKTRKDRILAAKNEQLAKNLKSIKKATTYIGTTVIMGTAGLALSKTKAAADTVVQADASS
ANTTSTTTTMSSSSAANASSTSSSVITTTATTNNEATSSQASSSSNVTTENSSSQASSA
VSQAPVVAAPRVLASNYSGYSSKVANFLNNVAPSAQAIAAQRGLYASMMIAQAALESG
WGTSSLSTNAYNLFGVKWNGYGAYINMPTQEFYGGAYHTVYAKFQRYSSYAESLNAY
ANLIVSHFPNSTKANASSYAVAANNLRNGVYGTYATAPNYASSLINLIQQYDLTKYDTP
GSTTTTSDSNSGNISSTPNNVVATSSYTVKAGDTLYAIASRYGLSVASLKALNNLSSNLI
YVGQTLKVSGTAATNQSNQTSNQTSSQTSNNNGATSSSNQSGSYTVKAGDSLWGIS
HKYGMSIDSLKSLNNLSSNFIYPGQVLKVGQSNNQSNTTSKPSASQSINSGSYTVKAG
DSLWGISHKYGMSIDSLKSLNNLSSNFIYPGQVLKVGQSNNQSNTTSKPSASQSINSG
SYTVKAGDSLWGISHKYGMSIDSLKSLNNLSSNFIYPGQVLKVGQSNNQSNTVSKPSA
SQSTNSSNQSGSYTVKAGDSLWGISHKYGMSIDGLKSLNNLSSNFIYPGQVLKVGQSS
NQSNTASRPSASQSTNSSNQSGSYTVKAGDSLWGISHKYGMSIDSLKSLNNLSSNFIY
PGQVLKVGQSGSQSNTVSKPSTSQSTNSSSQSGSYTVKAGDSLWKIATEHGLTIAQLK
SLNNLTSNTIYVNQKLTVTMKKSVNDVIPTTKAPTQGYTVKAGDSLWQIASAHGTTVNR
LKALNKLNSDLIYVGQHLRLN
>Leucyl-tRNA synthetase (EC 6.1.1.4)
MSYNHKEIEKKWQRYWEEHKTFKTTEEEAKPNYYALDMFPYPSGQGLHVGHPEGYT
ATDIISRMKRMQGYNVLHPMGWDAFGLPAEQYALNTGNSPREFTKKNINNFRRQIKSL
GLSYDWDREVNTTDPAYYKWTQWIFEQLYKRGLAYEAEIPVNWSPDLGTVVANEEVI
DGKTERGGFPVIRKPMRQWVLKITAYADRLIDDLDDLDWPEAIKEQQRNWIGRSVGAS
IKFTLENDAETKIEVFSTRPDTIFGASCLVLAPEHDLVAKLTTSEQKAAVDAYIDEVAHKS
DLERTDLAKDKTGVFTGSYVINPVNGQKVPVWIADYVLASYGTGAIMVVPAHDERDYE
FAKKFDLPIIPVIEGGNIEQEAYTGEGKHINSDFLNGLNKEDAIAKAIAWLEEKGIGAKKV
NYRLRDWTFSRQRYWGEPIPVIHWEDGETTLVPEEELPLYLPATNNIKPSGTGESPLA
NLEDWVNVVDENGRKGKRETNTMPQWAGSSWYFLRYVDPHNHEAIADKAKLEKWLP
VDLYVGGAEHAVLHLLYARFWHKFLYDLGVVPTKEPFQKLVNQGMILGNNHEKMSKS
KGNVVNPDDIVEQYGADTLRLYEMFMGPLDASIAWSEDGLNGAYKFVNRIWNLIIDDE

DKLRDRITTINDGKLDKVYNQTVKKVTEDYQALHFNTAISQLMVFVNDCYKADSLPLEY
VEGFVKLIAPIVPHIAEELWTKLGHVGSIAYAKWPTYDESKLVENVVEIVVQVNGKVRQ
HLQVAKDASREELQTLALNDDKIKQELTGKEVKKVIAVPGKLVSIVVAK
>hypothetical protein
MENNLTFQNQTLIISPRGISKLLAFKSKLAIPFNHIVGATIDPGILKDFKGIRSPGTAVPGL
YYAGTFYNQQEKLFFNIRTSSTPVVIQLRNEEYDRIVIGVEDPRKLVEYLNNLIYSN
>D-serine/D-alanine/glycine transporter
MKNELSRALTHRHVQMIAIGGAIGTGLFLGSGSAIQKAGPALILAYLIAGIFCFFMMRAIG
ELVLSDTTKTSFIEFVRQYLGSRFEFIIGWTYWLCWISLAMADLTASGIYIRYWFPQVPQ
WLTALIIILVLLAFNLFSVGLFGELETWFSSIKVLAIMALIVTGLILLFTSADIGGHHVTIMNL
VNHGGFFPKGTLGLLAAFPMVIFAFTGIEMVGLTSGETANPEYDLPRAINNLPIRIGLFYI
GSMFVLMCIYPWNEIVTSSSPFVQVFEGIGIKFAAAIINVVVLTAALSACNSAIFSTSRTL
FVLANGKKAPKALGVTNNRSVPVKALLFSSSVLFLIVILNYVLPATVFEIISGVATVSFVF
VWIALVWCHYKYRKANPKKKTVFPMPLFPITNVLTIIFFIGVLVLLAFNTSTLISLLFSLVW
FLALAVAYSILKKKLN
>Membrane protein involved in the export of O-antigen, teichoic acid lipoteichoic acids
MKNKQTLSAKEKMLRGSAWMTAASIFSRILGAIYIIPWYAWFGEHRLQANALYAKGYTV
YAVFLMLSTAGIPSAVAKQVAHYNSLNEYEIGRRLYRKSFVIMLLLGIIMGGILWGAAPV
LAAGDMRTVPVFRSLATALAIIPLMSLTRGFFQGYQDMAPSALSQLVEQVIRVFYMLVT
AFIIMKVINGSYQTGVVQSTFAAFVGAVGGILTLAYCYLKKKGEFDYLRANSSNSLEVSD
GQLIKELLAQSLPFIFIGIATNLYNLIDQYTFPWVMSKVSTYSSNYIDEMYALFAANANKLI
MIVVSLATALAATAIPLLIEFVTKGQKRDAKKQLGDSLELFFFIMLPSALGMAAVARPLYI
VFYGYSAYGIAVLKIAAYTAITQGLLIVLASLLQGLYENKRAVAYAICGLIVKLVVQFPLTA
KLEAFGPLVASGIGMSCSIYLMLRYLRRGYGMNFKRVQHTFNALMILSLAMFVVVVSLN
FLWTILFDSESRIVAILELIVTAGSGGYVYVYLTLKTHMADKILGPRVARLRQLLRIR
>Ribosomal small subunit pseudouridine synthase A (EC 4.2.1.70)
MRLDKFLAASGVGSRSEVKQLLKQKRITVAGKVVTSPKIQVEADSQVELDHQPVNYQT
NYYLLMNKPKGVITATTDKSQRTVIDLLGAGEQKLGLFPVGRLDKDTTGLLLLTTDGEL
AHRLLAPKKHVSKVYQAKIEGIVTDQTVARFAKPLTLKNGEVTKPAKLEVISGQKDANIS
RVIIEITEGKYHQVKRMFAAVGMHVLELHRLKMGQLELPADLKPGQYRPLTKEELASLT
K
>NAD(P)HX dehydratase
MEQIKQTILNKVIRPRPLHSYKGNYGRIITIGGNYNLGGAIIMSASACVYAGAGLVTCATD
KHNLNALHTRLPEAMFVDYSDFSDLLENLTKADVIVLGPGLGLDKLAQNICQLVLSNLT
PDQLIIIDGSAITLLAANPGLLTYLKKGQAIFTPHEMEWQRLSQIKIADQIPLKNKQAQLKL
GQTVILKKYHTEIYHPECDTPHQLTIGGPYMATGGMGDTLTGILAAFLGQFSSAPLTDVI
DAAVYTHSAIASKLGVNNYVTLPSAIIAELPHFMNCTIKD
>Acetylornithine deacetylase/Succinyl-diaminopimelate desuccinylase and related deacylases
MAIDWQAEVAKRKDDLLNDLKTMLEIESVRDESIATADAPLGPGPKAALDKFLEIGKRD
GFETMELDGLAGHIAYGQGDETLAILGHVDVMPAGKGWDTDPFKPVIKDGRLYARGA
SDDKGPSMAAYYGLKIVKELGLPVNKKVRFILGTDEESQWRGMSHYFEKMPAPDFGF
SPDAFFPIINGEKGNVSFFINFAGSNGDAEVKLVHFEAGLRENMVPRDAKAVLAGIKLE
DVEAAFADFVAANPVSGQAYEEAGNVVLELIGKAAHAQEPRNGENAGTYLATFLANYN
LAGDAKNFIEFTAKYLHKDSRMNALGATFSDEIMGDLTMNSGIFAFDENKGGKITLNFR
FPKGMDENDIKASLTKATAAMNVTIDQGRVQVPHYVSPEDPLVKTLLGVYERQTGQEA

HGMVVGGGTYGRLMKRGVAFGAMFPGVPDTMHQANEFIPVDDVLRAAAIYAEAIAEL
VK
>hypothetical protein
MKKGIESCIMKMKAKKGVAIMAESKSTFIKKLLDNVIASDGLSKTYAVPSKREQLALKID
NRVIEKLLATNHLEKMVRTLIRKRSKSTQKEIVNITLKNYRIFL
>tRNA-Leu-CAG

>FIG00743414: hypothetical protein
MFLTTGGINRTYAIKGIVNATIKKTLLAEEIDKVDQYDELYEEIKRRLAKRATALNGDAVIE
VRFIPEIVRVSVGPKYMLLHGYGTVISFPRKK
>hypothetical protein
MAKDFLHYYVQRAKIYRDEAQRAITCTTLDEYERAEIIKKTLLRSVTAELANLSTEISAYY
ELLEAIQTYSQKQPELVLELTYIRTACQKFIDSYA
>Universal stress protein family
MLRQEYKTILVPVDGSAIAEKALVTGTEIAKRNNAHLDILNVIDLKQFSVSFGGMIDANG
DVVYQSFEDVEKYLEGLKKHVEEQGVTDVAIHVRFGSPRNVIARDFPNDHETDLIVMG
PTGLNPVERVLVGAVSDYVIRTAACDVMIAR
>hypothetical protein
MLGEVIKWIIYVLLVVLLSYNIYLSVEMTKFKRRSLVRSVEELSQMETQMIAELSKTKRK
WSLLEQTLFLIAIFMAFKGKQIEVAFFIDLYTVASIVTNKLTYQIFRTLALKD
>hypothetical protein
MLSLKDYGIALAKGLLPLVVAIALVNLINASNWQYLLNYYVICNYLTFYPYNLNQAKPLE
NEQPAEQPLRRSKSKPAPEGLRRSHDDRTAGDKASEILSSWVLAFGKHLLLLLVAPGL
FIYGLLFSNEVGEF
>Lysine decarboxylase family
MDNICVFCGSNLGKFTEFEQQLKQLAAYLVEHKKTLVYGGGRLGLMGALYQDVVTAG
GKVIGVIPKVLEKHAVVADKNTSLIHVKDMSERKQKFVELADVFVVFPGGLGTMEEFFQ
VYSWAQLGLVTKPIVLVNIGGYYDKLLAFLKDGVKQGFMPQANLDALIVVEEVTAGLAA
AKNFEYHQANKWQDVSQR
>Amino acid transporter
MIEMKNIWKRMLAKEDPSVYKAKDAHLLQVLTTKDFLALGVGTIVSTAIFTLPGVVAAKY
AGPAVALSFLAAAVVAALVAFAYAEMASTMPFAGSAYSWINVVFGEFFGWVAGWALL
AEYFIAVAFVGSGLSANFRGLIEPLGFKLPAALSNPFGSNGGVIDIIAVIVLLLIAFLISRGV
SEAARVQNALVILKVLAILLFIVVGATAIKFSNYTPFIPAHHTNPDGTPFGGWQGIYAGVS
SIFLAYIGFDSIAANAAEAKNPSKTMPRGILGSLLIAVLLFVSVSLVLVGMYKYTNFADNA
EPVGWALRNAGHPVVATVVSSIAVLGMFTALIGMMLAGSRLLYSFGRDGMLPSWLGK
LNKNQLPNNALLTLTVVTVILGAVFPFETLSQLVSAGTLIAFMFVSLGIYAIRPREGKDLQ
MPGFKMPFYPVLPAVAFLGALAVFWGLDIEAKEYAGYWFILGVIIYFSYGMSHSILAKKK
>hypothetical protein
MKYQEQLEQLLINNSEINEIFAILRKFHLGQAYLCAGSVRTLVWNSLMQRETNLKLGNL
DIFYNAQFESAEEHQVLTTRLNQSYSKYLWSLNNLSQINPHAGHHQNGQALPDVIAAF
PETASAVGVNLDFQGKVSIIAPYGLNDLFEFKLVASPNFLTDSTKLAQFHKRVERKNWL
TIWPQLTLS
>Acetate kinase (EC 2.7.2.1)
MKKILAINSGSSSLKFKLFEMPAEKVLAAGLIERIGFKDGGVTLKYGTGQVYTETRKVSD
HRVAVDLLMDLLIEFKIIEDYGEITGIGHRIVAGGEYFKESTIIDDEAEAKIEQLIPLAPLHN

PAHLVGIRAFKKKLPQALEVGVFDTAFHQTMPEENFIYSLPYEWYQKYNARRYGAHGT
SHRYVAQEAAKLMGKPLADLKLITCHLGAGASLTAIKDGKSFDTSMGFTPLSGVTMAT
RSGDVDPSLLSYLIEKDPEIADMSEMIDRLNHEAGLLGISGVSSDMRDVLAAAPSNHRA
QLAVDIFVKNIVKYIGQYIAEMQGVDGIVFTAGIGENSQPVRSLVMKRLAYLGIKEDVVA
NATFGKNTFVSTPDSKVKVMRIMTDEELVIARDVQALAR
>Universal stress protein family
MKKYQVILVGVDGSKNAKEAVNKALELAKYSDAKLVLASVIKPYVSVGFGGPGLEVDS
NSDDQMTNEERKHREEMLQKYAKEASQAGVTAETVVVIGSPKEELAKNLAQQYQADL
IIVGATGLNRVEQVMIGSTASYVVRASKRDTLVVTSQA
>Lipopolysaccharide cholinephosphotransferase LicD1 (EC 2.7.8.-)
MEIAAVHEVELALLAEFKRICDQNKLNYFLIGGTLLGAVRHRGFIPWDDDVDIGMLRSD
YDKLLKILPQQLSKDYYFLQTPWSDPHYALSYAKLLDERFIIQERFNYNNARSGLFIDIFP
FDTIPMGKVERSLQMVNFKVANDQLLLKLNYSLIPMPERMVNSAKEKVISKLDEIEQLK
AKREELMRQYNVDDLLPEVKNLASQYSYEKEVFAKAELRKLVLLDFEGLKLKAPQSYA
AILTQMYGNYQKLPPLEQRTSKHLSNIKER
>Deoxyadenosine kinase (EC 2.7.1.76) / Deoxyguanosine kinase (EC 2.7.1.113)
MITLAGIIGSGKSSLTELLAQELGSRAFFEPVSDNPVLPLFYKGNELAAKKRQAGQKDA
TNPYAYLLQTFFLNRRFKMMKEARASHNAILDRSIYEDAMFMKMNTDMGNATAIEYEIY
QELLANMLAELEQVSQEEASDLTVFIKVSYETMIKRIEKRGREYEQISSDPSLVDYYQRL
LKYYAKWEQDFKGSPLLIIDGDKYDFVENQSDRNYVLNLIETKLVELGQLSEPEFARLK
AARA
>Mobile element protein
MLRVQKVRLYPNETMKQVLDDLCAYRRYCWNQGLALWNDMYDASLVLGDKKLRPSE
RKVRDELVANKEDWQYQLSARCLQLAISDLGKAWQNFFKKSLPDWGKPKFKSKKTAR
QGFKTDRAQIINGKLRLDKPRGVKAWADISFKGADDLNGDLKVVSIYRENGKYWASLP
FEVKVTKKTKTGQKTAVDVNVGHFDYPEGQVKTLPNNLKALYKRIKHYQRLLARKRVA
NGKKATQANNYVKTRAKLQRDYRKVASIQHDIVQKFTTMLVNNYDRIAIEDLAVKRMQ
MSHVASKGLHRSMFGYFKQVLKYKCEWYGKELILADRYYPSTQRCSQCGYVKTGAD
KVGLDGNKKHGTKHNEYICYECGVVMERDENAVRNLLALL
>Phosphoribosylaminoimidazole carboxylase catalytic subunit (EC 4.1.1.21)
MSAQIAIVMGSSSDWATMKEAAKILDQFNVSYDKQVISAHRMPKEMFAFAQGAVAAG
VKVIIAGAGGAAHLPGMVAANTPLPVIGVPVQSRALNGLDSLLSIVQMPAGIPVATTAIG
VAGAKNAALLALQILGISDSSVQAQINDYRQEMHDQAQESGDQLV
>Phosphoribosylaminoimidazole carboxylase ATPase subunit (EC 4.1.1.21)
MTNLSNYIEQGQTIGIIGGGQLGQMLAFDAKQAGMKVIVLDPDETCPTGQVCDEQIVAA
YDDIAAIEKLAELADVLTYEFENVDLNALEDVADKVSIPQGTELLRITKDRLREKTFLKEH
GLKVAPFRSVATETDLKQALAAIGYPSVLKTCQGGYDGKAQVVLRTEADYAQAQELLQ
KGQCILEGWVAFEKECSVMVGRNLLGQVTVFSVSENIHRNQILHESIVPARISTTLQEQ
AQKMAVTIAKGLNLCGILGVELFVGQDGQLYVNELAPRPHNSGHYSIEACDFSQFAIHN
RAICNWPLPQVKLLKPVVMVNVLGQHLEPSRQLIKQKPAWHFHDYGKAESRVDRKMG
HITILTDDLEATLAEIKATKIWDKE
>Adenylosuccinate lyase (EC 4.3.2.2)
MISRYTRPEMAALWTDEKKYECWLEVELAADEAWSKLGHIPAEDVEKLRKNAKFDINR
IAEIEAVTKHDVIAFTRTVSESLGDEKKWVHYGLTSTDVVDTAYGYQLKQVNDHLRKDL
AEFKEIIANQARKYKNTVCMGRTHGVHAEPTTFGLKLARWYSEIKRDIERFEHAAAGVE
AGKISGAVGTFANIDPFIEEYVCKKLGIRPQEISSQVLPRDLHAEYLAALALIATSLEEFAT

EIRGLQKSETREVEEHFAKGQKGSSAMPHKRNPIGSENICGLARVVRGHMVTAYEDV
VLWHERDISHSSAERIILPDTTILIDYMLHRFGNIVKNLTVFPENMKRNMDRTFGLIYSGR
VLLKLIDTGMSREAAYDLIQPKTAQAWDEQRPFRPLLEADPEITGRLSKEDLDDAFDYH
WHLKHVDDIFKRVGLE
>Endo-1,4-beta-xylanase B precursor
MYYHEINLRPDNPQAGMKLYLLDEEITHKKHRNRPLMIVCPGGGYLSLAHKESEPVVA
RFLGLGFNVALVNYSVYFKQRPDESQPAIVNEDYRVKDSLLDLVASLNLVEQHAAEWFI
DTSRVYLLGFSAGSHLVTSLAERYLDKQWRREAQLKITFKIRGLLLSYPMLSANAVHVK
EQAHFDELMAKESLAQLDLVTQVNAQMPRTFIWQGNEDLAVSPLETSKFVVKLQEAG
VACEYHLFATGGHGLALADRTSAVFASDIQPNLAAWVELARVWLAQDEPATIHLYPDN
RE
>Potassium uptake protein, integral membrane component, KtrB
MYSRRKTKQKHIFSRVQVITLSFLVAILIGTGLLMLPVMSVTGKATPFVDAMFTATTSVC
VTGLVTVTTASHWSALGHLVILLLIQLGGLGTITLMSLAFFATGRRISLQNRLLMADAFNL
DNNQGIVRFVIHVIRDSLIVEAIGALAYACYFVPHFGLGRGSWYAIFHSISAFCNAGIDLL
APNSLINYAGNWWLNFVTMFLIICGGLGFTVWNELGHGLVNQYRYRKKGRGKRLKTTL
SQHSKVVLLTTAILILGGAVLYYLFEHNNPATLGKVGTAQAMLEALFQSVTTRTAGFVTI
SQADLTIPSVMLTVFLMFIGGSPTGTAGGVKTTTLAVIMLEVKAIIKGEEDVICFRRRIGR
QTIRKAIAITVISVIVSVMAVLLMYFIQPGDMTDNIFEVYSALGTVGISRDLTPDLTTAAKY
LIIACMYLGRVGPASILLALTKREHQADIKYAAGNIRVG
>Trk system potassium uptake protein TrkA
MDRSYAVFGLGEFGRSVCHELMELGADVLAVDRNEAQTTAIADDVTMAVAIDCLNEM
AYEKLGLNNMDGVVVSMSGNLGASIMAIMAAKDAGVPLVIAKASDDTQRTIFQKVGAN
RVVIPERDGAVRAARNLLAKNFLDYIELSDKISIVEINVKREWLNQRLADLNLRSRYGLN
VIAVRRDGELLTDIGPDATFIMGDTILVVTDKQDLGLERGK
>FIG00749078: hypothetical protein
MVVIKDRTVIQDLGTDFFRDIDQNAKGLTIFRGAVYNTSLSFIKKYLLPRFEKIELVIGLSD
TKANTALENYLYASNEQFKLLTSLDSELRQRFEDGSLELKFSKEALVHSKVYILTNPNG
DDYRAYLGSMNLSEQALKRNKEVLLCDHGKTEDLAYQEIYQAVYQDIYQTAAEFLNRKI
VTGFFGQENTEKQKIYLFDQTVENLAAAKNKPTFASEDITKIKQETRQVIDVAAIKREQK
AKEQVVNLIFNKSGKVKEKLNLLTDEVKRKQIFKVIYHEQESKANFDFLNRQAHEFYPK
PLYVYDEDQTSLLKAPSFGSSLQESVIAQKPTKEEITDILEIVRFYRDRKQSDESQAAFS
FLMYVLQSANIWKIRKIISEHKGNVDQIPVVAGLIGQGNTGKTTLLKIASNLTIGNTNNIVD
GTDAMFRIKKKVKDALAAGKSPEELGKTKPLMESKKDINRNVWSFTQEYMETKSSVTP
MCIDDPDISLFGASAEATLKYLANRYHGSVHPVVLFGLNDKDGGKVSISSQVGKRAYAI
GQENQFKEFTVQDEEFLDFLEGIKSRQGRMVNNNVFLYLSQAIDRQLDNLSLEDYQML
TKDFLSPVKANLRRLIEDYELDVSDLEAYFGEDNYDINFDKGRRNWRTLLSDKQVIEQI
SVTAGDDQCLIPVGVFPNRNDIKRFFDYLPPKLEICASYVSVGLTVNIDNMDKWLGVNL
LREHYDEETGVKKRREELEKIERQSEFQAMYTVKYQAEQKQKEKLERRQKNPWFRLV
DKFIK
>Phosphoribosylaminoimidazole-succinocarboxamide synthase (EC 6.3.2.6)
MSELIYTGKAKQMWTTEDEDVLRVVYMDQATALNGKKKDQIKGKGQVNNQISTLIFNY
LKAKGIPTHFIKKLSETEELVKKVTIVPLEFVTRNIAAGHFASRFGVKEGTPFKQPVEETY
YKSDELDDPFMNESQALALGIATKAELDYMWELSRKVDQLLIPLFAKANLQLVDFKLEF
GRLADGTIVLADEFSPDNCRLWDLTTKAHMDKDVYRRDLGDLTPVYEEVLTRLEKVLA
KEAN

>Phosphoribosylformylglycinamidine synthase, PurS subunit (EC 6.3.5.3)
MVEVRIYVTYKQSVFDSQGETITDAIHALGHDEVQKVKVGKFFDVTLDPSKKDIDQAVK
EISEQLLVNFNLETYRYEVMEEA
>Phosphoribosylformylglycinamidine synthase, glutamine amidotransferase subunit (EC 6.3.5.3)
MKFAVIVFPGSNCDIDLYEALKTVCHAEVEYVSHKETSLAGFDAVMVPGGFSYGDYLR
CGAIARFANIMPALIEFAKAGKPVFGTCNGFQILTEAGLLPGALKQNDSLKFVCKTVALE
VNNQTMFTSEYNDNEVIKLPIAHADGSYYADEETLAKLEANNQIAFRYHGENPNGSLH
DIAGIVNEQGNVLGMMPHPERAVEALLGNEDGLRLFKSILKHGHQTEEVVNG
>Phosphoribosylformylglycinamidine synthase, synthetase subunit (EC 6.3.5.3)
MAKELTPQEIKDQKPYLEWGLTEKEYDYIATELLGRLPNYTETGLFAVMWSEHCSYKK
SKPVLRLFPNKNERVLQGPGEGAGIVDIGDGQAVVFKAESHNHPSAVEPYEGAATGV
GGILRDIFSMGARPIASLDSLHFGEIDRPRTKYLINEVVAGIGGYGNCMGIPTVAGEMTF
DECYTGNPLVNAMSVGLMDQKDMQKGRATGVGNAVMYVGAKTGRDGIHGATFASA
DFADDKETQRSAVQVGNPFMEKLLMEACLELTSQHADWLVGIQDMGAAGIVSSSCEM
ASKANSGLELNLDYVPQREANMSAYEIMLSESQERMLLCVKKGHEEEVQAIFKNYDLE
AVVIGRVTAGHDYVLKHQGQVVTNIPVSTLTDDVLEEPSQERVPARLANADTEKWQPE
ISDASETLKKLLAQPTIASKEMFTQTYDSQVRTSTVVGPGSDAGVIRIRGTKKALAMTTD
GNGRFVYLDPKVGGKMAVVEAATNVISSGALPLAMTDCLNYGDPNDPEIYYELHQSV
QGMADACLALDTPVISGNVSLYNENNGEAIYPTPMVGMVGLIKDVNYVIPSFVQTAGDL
VYLVGQTGDDYAGSELQKLLTGKISGSLADLNLTEINAYLKKLLQAMETGLVKSSHDLS
EGGLGVALAETAFKTDLGLDLDLTLTKAQLFSETAGRLVVTVAKEKQADFEELLGQAAS
LVGQVTAEHKLVAKLSDGQLQLDLAAGQKIWEEAIPCLMKSKA
>Amidophosphoribosyltransferase (EC 2.4.2.14)
MSYEIKGLNEECGIFGVFGTPNASALTYFGLHSLQHRGQEGAGIVSSDDQRLHQYRNR
GLLAEVFADQKDLDRLVGQSAIGHVRYGTSGNNSVANIQPFLFRFHDGDIALAHNGNL
TNAKSLKYQLEDQGAVFQSTSDTEILIHLIRQKRELGFIDALKASLNEVKGGFAFLLLKK
DCLVAALDPNGFRPLSIGQLENGAYVVVSETCALDVIGAKFVRDVLPGELIIIDKDGLKID
HYTKDTQLAICSMEYIYFARPDSIIHGVTVHNARKKMGRLLAKEQPTVQADMVIGVPNS
SLSAASGYAEEAGLPYEMGLIKNQYIARTFIQPTQELRERGVKMKLSAVRGVVEGKSV
VVVDDSIVRGTTSKQLVKLIRDAGAKKVHMRIASPPLKFPCFYGIDMSTRAELMAANHS
VEEMRQIIGADSLGFLSIPSLIKAIDLPDAGDAPNGGLTVAYFDGNYPTPLYDYEEDYLK
SLRQG
>Phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1)
MSRYQEAGVDVNAGYELVRKIKADVKSTARLGVMGTVGSFGGMFDLSSLNVKEPVLV
SGTDGVGTKLMIAQMLAKHDTIGIDCVAMCVNDILAQGAEPLYFLDYIATGHNDPEKMA
AIVAGVAEGCRQANCALIGGETAEMPDMYAEDEYDLAGYSTGVAEKSQLLTATLPQAG
DILVGLPSSGLHSNGYSLVRQIIFKDAGLALTVKPAELSGKTLGEVLLEPTRIYVKAILPLV
KAGLVKGISHITGGGLIENLPRMYGDDLKAKVDATSWPKLPIFTYLQKLGQLSEQDCYE
TFNMGIGLVLAVSQTDYPKVSQALSQAGEAHYQIGQLVERQADEDKISIEKSR
>Phosphoribosylglycinamide formyltransferase (EC 2.1.2.2)
MRVAIFASGNGSNFEVLAQKFQQGQMAGELVLLFCDHPDAYVVKRAEKFQVPVETFT
VKACGGKQAYEERILQVLKDYQVDFIALAGYMRVIGPTILADYEGRIINLHPAYLPEYPG
LHSIERAFADKKKQTGVTVHYIDSGLDSGPIIAQEHVAILPTDSLASLEERIHACEHRLYP
QVLQTILANQANN

>IMP cyclohydrolase (EC 3.5.4.10) / Phosphoribosylaminoimidazolecarboxamide formyltransferase (EC 2.1.2.3)
MKRALVSVSDKSNLVDFVKTLVANDYEIISTGGTKKVLDEAGLKTISVEDVTGFPEILDG
RVKTLNPYIHGGLLGRRDLPAHLEKMAELKIKPIDLVCVNLYPFKETIQKPGVTLADAIEN
IDIGGPSMVRSAAKNYQAVTIVVDAADYKQVAEEIAANGDTTLETRQKLAAKAFRHTAA
YDALIADYLTKQAGISAPEKLTLTYDLKEEMRYGENSHQKAYFYEDALPEAYSIAQAKQ
LHGKQLSYNNIKDADAAIRIAREFSQPTVVALKHMNPCGIGQADNLEQAWDLAYEADPI
SIFGGVIVLNRQVDLATAEKMHKLFLEIIIAPGFDQDAFEVLAKKKNLRLLQLDFDQKDTA
QKPEVVSVLGGLLVQDQDVLAEDVKDWKVVTKVAPTEEQLASLMFAWKAVKHTKSNA
IVVANGERTLGIGAGQPNRIDSTKIAIKHAGAALDEHAVLASDAFFPMDDCVEYAAKHGI
KAIVQPGGSIRDQDSIEMADKYGIAMVTTGVRHFRH
>Phosphoribosylamine--glycine ligase (EC 6.3.4.13)
MTKKLRLLVVGTGGREHAIAKSLAQSSLVEKVFCAPGNVGMASDKIEVVNIGELEFDAL
KGLVAEEQIDWTFVGPENALVAGIVDSFEADGLKIFGPNKEAAQLEGSKDYAMRFMDQ
YQIPTAKFATYQSSEAAIAGLKDFSEPIVIKADGLAAGKGVVIAPSKDEAEAEIRLMFKKG
QKQIVLEEFLAGDEYSLFVLIGKSGYRILPMAQDHKRAFDQDKGPNTGGMGAYSPVPQ
LAKVDYQRMLIEVVEPTIAGLKQANLNYQGVLYIGMILTAQGPKVIEYNVRLGDPETQVV
LPRLATDFAELINAFVTEANLPEVKVKPNACLGVVLAATGYPQAPLKGQAISLQLADESL
DIDYANVTGSLDDLKGAGGRILTVLASAPSLAEAQLKVYECLNSQVQAETFYRKDIGFR
ATKASF
>oxidoreductase of aldo/keto reductase family, subgroup 1
MTKIPTKTLNDGNQIPVLGFGTYKLNGAQGVNAINSALSNGYLLLDSAFNYENEGAVGE
AIRRSSVKREDIFVTSKLPGRHQTYDQAIPTIQESLYRAGLDYYDLYLIHWPNPKEDHYV
EAWQALIAAQKFGLVRSIGVCNFLPEHLERLVAETGVTPAVNQIELHPYWSQASQRAY
DQAHGIFTQAWSPLGRASSVLQEPVINAIADKHHKSAAQVILRWEFQLGVGIIPKANSP
KHQLANLDIFDFELTPEELATITELDKPNGRLKNQDPAVYQEF
>Transcriptional regulator
MADFTYQVFSEVAKQKTLVAAAKQLQVTPSAVSHSLAGFEKQLGFPLFVRNRNGVKLT
PDGAKILPIVQELLNTEAKLAEEAARINGLNQGQIRIGAFSSVCINWLPDIINSFKQTYPEI
EISVRQGTFNEVVEQVRLGTLDLGFSALPIKENLEVIALHQDPIYCIAPVGFKPRNGKTIT
QEDVVNEEFILQQIDYDRDTKRALDAYNVSVNSIQYSIDDASIIAMVESGLGLGILPDLAL
QNLAGKVQTYPFSKEFYRTICLVAHTKSQQTPSTKAFIKTVQEYIKQRYA
>hypothetical protein
METKIEVGDKVKCKKFGSLKHDFIGSVEKKYENSAVVAILEHDNEDNVAVTDFHNRAVV
RFDCMKKISA
>tRNA-Gly-CCC

>Arginyl-tRNA synthetase (EC 6.1.1.19)
MQDYKLQVANAISAAVDGQLSQDEILSKIEIPKALNLGDYAFPAFVLSKVLRKAPQAIAA
ELVEKIDQTGFEKIEAVGPYVNFFLDKQAYAKAILAEVLAAKEDFGNQDLGAGRQVPID
MSSPNIAKPMSMGHLRSTVIGNSIANLMAKVNYKPVKINHLGDWGTQFGKLMCAYEM
WGSEAEIKADPIATLLKYYVRFHEEAEAKPELNDQARHWFKRLEDGDAEATRLWKWF
REESLKEFMEVYDELGISFDSYNGEAFYNDKMDEIVELLEEKHLLKESQGAQIVDLDKY
NLNPALIKRSDGATLYITRDLAAALYRKRTYDFAQSLYVVGNEQTNHFKQLKAVLKEMG
YDWSDDIHHIPFGLITSGGKKLSTRAGRVILLQEVLNDAVNLATKQIEEKNPTLANKAEV
AHQVGVGAVVFHDLKNERLNNFDFDLEEVVRFEGETGPYVQYSRARAMSILRKAGLD

AASVTGSYDLNDPEAWETLRLLGDFPATVKRAVTEYEPSVVAKYAIHLAKAFNKYYAH
SRILADDEQKQARLAMVAAVATVLKEALRLLGVGAPDEM
>Predicted cell-wall-anchored protein SasA (LPXTG motif)
MKRRKYNKLYKETDRKTRVKMYKAGKHWVQALLSRIGLLRIASDDVADKSVQVNNIEL
KRAATKDSIDKENDAKKKALRTAVTAGTVLGGAVVSTQVAHADTTQATEQNVTAQSTS
DSQASTSVSQGSTTSTISEKSVVTSTSELNTESDSSSQVNSTSVSTSVSASQSSSVSA
SQSNSASASQSNSEAHSTSETTSSASSETTSGSEASTVASEANTVTESESATNASSVQ
PSSEVSATTSTTNKDKLESLLNQSDEITATDGYKQADQSYQAMYQAVIQHYREVLANP
TQNQSDEYYGYGLAQLEELQGLFTNSTVVQPQIFAQSDMNRDSNWGNVEDDNTKAG
DTKAYWFDKEGSGGFKYIQGADRLWSNVDVHTSLNAKDFTSKAVDSNLQISTKQLDD
GKTKWTITFFPKKGLDSTGNKSENDVYGLSNGQMGFYLTKDYKIIGDINIHVSAKSGTS
YKINGLNGFPNFYTVDGWGKMSVPSEASFDFNGSQADGRDYWSIHNKDFAGSVKTT
NADRDWKLTQKQGFATNQYIMHLRYADRSIMDDSTFNQVVFTNKGSDKLDESFATAN
DFDKLTVRDKPPTFNFVNNTVPFDTSFNKDTIGTAVYMQSVGMFNDGIRASYTMTFTT
EHSNAEQAKLASGSNGGAFSGIYAIANSYQQGNENAGGGRLIGQEVYNLNKNGKIDST
IPSQIESTSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASE
SASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESAS
TSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSA
SESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASES
ASTSASESASTSASESASTSTSESASTSASESASTSASESASTSASESASTSASESAST
SASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSAS
ESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESA
STSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTS
ASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASE
SASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESAS
TSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSVSESASTSA
SESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASES
ASTSASESASTSASESASTSASESASTIASESASMSVSESASTSASESASTSASESAST
SASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSAS
ESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESA
STSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTS
ASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASE
SASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESAS
TSASESASTSASESVSTSASESASTSASESASTSASESASTSASESASTSASESASTSA
SESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASESASTSASES
ASTSASESASTSTSESASTSASESASTSASESASTSASESASTSASESASTSASESAST
SASESASTSASESASTSASESASIGSATSMSDKPGTLPSQSEESSDSASTSASESAST
SASESASTSASQSASTSASESASTSASQSASTSASESASISASTSMSGKPSTLPSRSE
ESSDSASTSASESASTSESESASTSASESASTSASESASISASTSMSGKPSTLPSRSEE
SSDSASTSASESAFTSESASASTSAIPSGSNLSSNSSSDGVVPLANHRTVTKIDLAENS
KVKNDKVGRQALPQTGQATNSYAEIGLASATLSALLLAAARKKGKKNEE
>Mobile element protein
MKDIANHNNVSPSTVVNVLNKSVTLSSFTLHDSLPEHLCFDEVRTNDNQLSFICIDAVK
HTLVALLPGKSSYDIKKFFNGFSLAQRARVKTVTTDLNAYYSSVAERLFPNTEIIIDRFHII
KMINQAFNHERVKKMKLFTKLSVEYNLFKCHWHLFLLDPAKLDNEHPHYRRQLKCSMT
DAQIVSEALELSEELLLSHNVIHKLHRAIICNDVLTLARCLRAFKQSFKQVSVQKAQ

>Mobile element protein
MNFENYYNTLSSHLDEINQINHLYLNITSKGLIMDEMPHVEPYKNFDRALFVPAKLKFPQ
RPTCIFCGSKAVEYHGYTPSTLKHIAADPSFPVFIELKNHRAKCFNCGHCFQAEDADLA
PKYAQKTN
>Ribokinase (EC 2.7.1.15)
MANKIAVLGSINIDTTYRVNEFPKPGQTIQALSKSSTGGGKGNNQAVAAARSEAQAYFI
GMVGTDDEGKFMLENLAEEGINTDLISQNKEIGTGTALVTLNSAGQNDIMVYGGANQA
MTTEVLTGSEVVLAQVDFLLAQNEIPQEVTLVAFKLAKSLGVTTIFNPAPAGKIMPALWE
YTDIIIPNETECQALTGSEITDEASMQAAANAFKEMGVGTLLITLGSKGVYYSTPSSQGL
VPAFKVQALDTTAAGDTFIGALASQLSKDLSNIVPALTYAQKASSLAVQKMGAASSIPS
AEEIKTALTR
>Inosine-uridine preferring nucleoside hydrolase (EC 3.2.2.1)
MTKIIFDCDPGIDDGVALAVALTDPALELQLISTVAGNVTVDKTSKNALKLVEFFNKEVP
VAAGAKKPLIKQYEDASHIHGQSGMDGYDFPEPQRTVVAKDAVVAMHELLQASPEPLT
ILATGAYTNLALLLAQYPEDKAKIKQVILMGGTLSRGNMSSVAEFNVYSDPQAAEIVFKS
GLDLVMVGLDVTLKGRITEASQQAIAKCGPAGQMLDGIISADGDKTEAGTAIHDLQTVF
YLQHPEAFVTKDYWVDVVTEGPALGATVADVRAAYHQETNVKVCLDIDTDAFNEWFV
NKVKDNMQ
>FIG00748805: hypothetical protein
MYKAFPGVNQAAVQTLSTIPNFGIVLGLVISPFLVKFLGEKVTILTGLLLTLISGVFPMFA
SAYTPILISRFLIGLGIGTFNSLAVSLIPEFYSQDEEEMATMIGYQNVMSSIGAALSSFLVS
YLVTISWHAAFAIYFLVIPALILFTLFVPLKKQPVVTSQHEARPKQSINSKVILIAGLMFAIF
AFYMPMAYELPSLIVSQQMGTASTAALIAGLSTLIGIPIGASFGFFFKRLHDKIFPLGFILV
TLGFLITALSPNIAILFCGVIVLGFGFGLGVPYMYNWLDWSAPSNSVNLATTIVLVLVNV
GCALSPSLLALIATPLGLNHPRGIMLISTCAFLVIAIYATIHYLGVHKANVAKETVHSKEPH
NAH
>Esterase/lipase
MVSQETLTAILEMRTLFKQGDDKRDAGLDPNPQGVIRYDNLAYGPDPKWHLLDLYLPQ
NASGKVPLIINIHGGGWCYGTKETYQFYGLNLAQHGFAFLNPNYRLAPAVSFPGQLDD
VNRYIHWISEHAAEYNLDTNNVFLVGDSAGGQMAEQYITLLTNPKFRKLFGYELPQLKF
RAAALNSAAVFVTDPGMMTGAVTGYFTAEVLATKAELLATEKYITKDFLPVYLSTANED
FIHDQTIKFAGFLEAKGVEYVCKSFGDKKHPEPHVFLMNQRDELAKQCNEEEVAFFKA
HLVN
>Sugar kinase and transcription regulator (EC 2.7.1.-)
MTNHKYLSIDIGGTNIKYGLINQAGQISFKAKLATPAEKEPFLAAIDQLIADYAPKVSGIAF
CAPGKIEKSTIRFGGALPFLDGIDFAQRYRHLALPLTVINDGKAAALAESWLGNLKDIKN
GAAIILGTGVGGGIIINGQLLAGSHFQAGELSYMIMDPTQTKSTEDSPMGTGFTGGTLS
AVGLIKKINELTKHEDLLDGLVAFKELEAGTNQQVTALFEEYCHHVACLIMNIQSVVDLE
KFVIGGGISAQPLVIKGINQAYDDLYHTNEAVALTLCRPQITVAKFNNDANLYGALYQLL
LTTATEKFN
>Transcriptional regulator, AraC family
MSNDLQHETVLVTPPLPVWLFYHHNNSANHIAAHWHQAIEFSFTLAGSIDRFKIGNQVY
QTKPGQILVVNSQVIHSIDVINHSDNQALSLSVPFNYLESFYPKLSESYFELNQPKKFTA
QQQSAYRKLQSLFYQLAWRYQDQDELKFIDLQLLINQILKLLLTHFTKPLASKTNLSRQT
SYTLSRLHEITKFVNEHYQADIGLVEIAAHCNISRVYLARFFKQHMDLTVSQYLNNVRA
QHAYKELLTTKKTLSQLALTNGFSGIRTMNRAFTNLYGQTASKLYRERKTAND

>Methionine ABC transporter ATP-binding protein
MASIVLKNIDVTFKNKAQTVNAVKNVSLEVEQGDIYGIIGYSGAGKSTLVRVINLLQRPT
NGTVIVENQELTKLKTQELRQVRKKIGMIFQHFNLMKSRTVAGNIAYPLLSTGLSKAEIN
AKVTRLLELVDLSEYRDAYPSQLSGGQKQRVAIARALANDPQVLISDEATSALDPKTTK
SILELLKKVNEQLKITIVLITHEMQVVKEICHNVAVMEQGQIIERGPVREVFINPKTALAKD
FIETSVNIKKALERIEAKNSEWNSEKRLFLLQFAGEETEDGAISRISKQTGVDISILFADM
DELDGITIGYTIIKISGKKAAIEAAYNSFKAAKIKVETLK
>Methionine ABC transporter permease protein
MKTEFIDAIYQTIYMTVVSALFAGILGLLTGVALVITQEGGLAEDKLLYNILDKVVNLLRSV
PFIILLAVISPLTIFLVGTSVGTTAALVPLVFGIFPFYARQIQAALLTVDAGVIEAAESMGSS
NLEIIFRVYLREGLTEIIRTSTITVISLIGLTTMAGAIGSGGLGDVAISIGYARYENDITLAAT
LVILIMVFVVQVIGDLLVKLTERG
>Methionine ABC transporter substrate-binding protein
MKKSYQIILGVIVVIVALVLGIHALNSSKKSAASSTITVGSMNTDYQVWKHIAESKEAKKL
GLKIKVKSITDGVQLNNATAQGSVDVNAFQSWSYLETYNKQHKKTKLAALGTTYYEPM
GLYSNKYKKVAEIPDGATIAIANNPSQAARGLLLLQKAGLITLKKNFDSATGTTSDINANP
KNLKFKEIDDTTGPRVLKSVDAVIISNSVAFDGGLNVLKDSIYHETISKNTKDNINVLATA
AKNKDNKAYQKLVKLFHSKAIQAYVNKKFSGTKVEVKKPVSYLED
>S-ribosylhomocysteine lyase (EC 4.4.1.21) / Autoinducer-2 production protein LuxS
MAEVESFTLDHTKVKAPYVRFITVEEGKNGDKIANYDLRLVQPNENAIPSAGLHTIEHLL
ADLLRDRLDGVIDCSPFGCRTGFHLITWGEHTTTEVAKALKESLQEIAENIEWDDVQGT
DIKSCGNYRDHSLFSAKEWSKKILAEGISDEPFERHVI
>2,3-butanediol dehydrogenase, S-alcohol forming, (R)-acetoin-specific (EC 1.1.1.4) / Acetoin (diacetyl) reductase (EC 1.1.1.5)
MGKLAIVTGAGQGIGEAIAYRLAKDGYAVAVADINRETADKVTATINKTGGQAKAYYLD
VARRNEVFDLVAKATADLGELAVFVNNAGVAFIETFVDSDPQKVERLLDVNLKGTYWG
IQAAAKHFIKQGKGGSIINAASLAGVEASALQSAYSASKFGIRGLTQSAAKELAQYKITV
NAYNPGVVRTPLRDSIDETTSKIKGFTVPQQQANVLTEIALGREAQADDVAKLVSWLAS
SEAEYITGQSIMVDGGMRYL
>hypothetical protein
MINYATDNFKLKIVPSAVLIYWRLGLFFVKKLSGSKVNPYYTI
>hypothetical protein
MLIDTALALLMAANASSKLTQLVVEVPQGTSALLKLAQVSEQPLELSAGKHVFKL
>Dipeptidase (EC 3.4.-.-)
MKVAECTTILVGKKASLDGSTLIARSEDGGREILPESFKVVLPEDQPRHYKSVLTGCEV
ELPGEPLRYTSVPDAHGKYGVWAAAGINEENVAMTATETITTNARIQGIDPLLVDDGLG
EEDFVTLTLPYIHSAREGVLRVGELLAKYGTYEMNGMAFSDKDEIWYLETIGGHHWAA
RRIPDDAYVVAPNRFNIDEFDFESPDFMASADLQAIIADYKLNPDFEGYNLRHIFGSATV
KDTHYNNPRAWYVQHYFDPNSQTTPLDQDQPFIVHADRKISVDDIMFLMASHYQNTP
YDVYGDQGTSAEKKRFRPIGINRNFETHILQIRNDVPAGLEGVQWLGFGPNTFNAMVP
FYTNIVDTPASFRDTGSKFNLQSIFWLNKLMSQLGDTNFKLYGELESAFEQKTVAACY
QVQHQTDKLAQELQGAQLQTCLTKANQEMADITFANTVELLGKMVEEGHAHMQLKYD
LLD
>hypothetical protein

MNQREELIRLWFQMWLQKKDLGILKLFTADAVYVESWGPKYQGNAEIKHWFDEWNTR
GTVLSWEIGQFFHKQNQTVVEWQFKCQMKNSEVTEFDGISLLKWTKDNKVQFLKEYG
CKSDNYNPYQVAADSPKFRDEKVDWL
>hypothetical protein
MAKKKLSNKKLLHQKKANERKVTGIKKTDSAITPMNFISQEQFISEATKLLQKVKQSDN
PTNIATFPAQLYPERKALARLLQGYILADANLKATDQLTNLTDYQNQNILVRGLVVDFIP
GASDEEGRLVIQAPTIIAHNKDGQTSWNYLTEPLALAPSYQLALKDILSASDNVRVSIGD
YIILDTLVTAKETLSQIAIVDSGLIITTKHGFNTLDSISSAFPRKKDFLFEFTDLPNDPGQW
GLVLNNEGLTYWSPLASRSEQVDPKHLAEFNN
>NrdR-regulated deoxyribonucleotide transporter, PnuC-like
MFARYYDVLTNFSSYSKNIFRKGYFAWLWQQCQGWGRFSYGLLGFNFILQVISLGQS
FRQTPLLAVIAFIGGNLSVACVIGISNRSGIQGWAGAISALAIATTGFIAGNYATAFEQLG
YLIFLDLFCILDPKWNDDIQVEKFESGLEWIKYGLFFLVTWLITYLLFSLTSDPRVFLDSL
NLAMAITGSLLELNRKREQFFVWTLASLFTIALWAQTMLQGDGNFALIFSYSVFFLNDM
YALFSRKGWFRLAE
>Purine trans deoxyribosylase (Nucleoside deoxyribosyltransferase-I) (EC 2.4.2.6)
MKNTDPVANTKIYLATSFFNEEQRARIPQALAQLEANPTVGVVHQPFDFQYKDARVDS
DPAGVFGSLEWQIATYNNDLNAVGTSDVCVALYDMDQIDEGICMEIGMFVALHKPIVLL
PFTKKDKSAYEANLMLARGVTTWLEPNDFSPLKDFNFNHPMAQPFPPFKVF
>NADH dehydrogenase (EC 1.6.99.3)
MKITVAGGTGLVGQSILTELRHHPEYELHSLSRTGRRPTSIPGVTYHAVNLLADTSWQ
QLIETSDWVIDCVGILFENRRKQITYQTSSVLPAKRLIDTLANTSGNFLYVSANYAPFFLK
DYLATKKEVELYAYSRLGQRAVCVFPGIVYSKKRLSNYYPARIIANLIRTRCFNWLTKIR
PISSQAFAHEILNIITGLPSPLRKRIP
>MSM (multiple sugar metabolism) operon regulatory protein
MNNEYYTLSQNGIESNILFFGQEKCQPGYAFRGNNIRDNYVLHYILAGQGDFSVAGHK
TVSLKAGDFFLLPKGVACFYQASFTNPWHYAWIGLTGIKIANIFAASKLAGNYYLKATSQ
SNFAKSFQRLFEVLHQKTSLANNLLVENLTYEMFYHLVTEFPHTNTDPSQNRPAQQFE
LAVRYLENNYTYNCTIEDLCVSLSLSRSYIYNLFQKFAHTSPQKLLTQMRMEDAKSKLT
SSNQSIQEIADLVGYHDTFTFSKAFKRYSGYSPSTYRKHFS
>Lactose and galactose permease, GPH translocator family
MSTTNKKMSMGQRLAYAFGAFGNDAFYGLLSGYLIVFITSHLFNTGNKALDNHMISLVT
LIIMILRIVELFIDPFIGNAIDRTKTKWGHFRPWVVIGGTISAVLLLLLFTSLGGLYAKNAFV
YLVVFAIMYITMDIFYSFKDVGFWSMLPSLTTDSREREKTATFARFGSTIGGGLVGVLV
MPAVIYFSATKTTSGDNQGWFVFAAIICTIALVSAWCVGLFTREVNSEIRENKEDTVGV
FGVLKAVAQNDQLLWIAIAYLFYGVGNNILCSLEVYYFTYIMGKPTSFSILSTINIFLGMM
AAALFPILSKRFSRRAVFTGSLTIMLIAIGLFALSGSNLLMVLASAVLFAFPQQIVFLVVLM
IITDSVEYGQWKLGHRDESLSLSIRPLIDKFGGAVSNGVVGQIAIMAGMTTGASASTITA
AGRTNFKLMMFAVPALMMVIAIVIFWKKIILTEAKHAEIVAELEKTWGKDLKETDVEATS
QVTSQTLKLASPIAGQLMPLAEVNDSTFAMGGLGQGFAIKPSDGKVYAPFDATVRQVF
TTRHAVGLVGDNGVVMLIHVGLGTVALKGTGFISYVDTNQRVKKGDQLLEFWQPTIQK
AGLDDTVIVTVTNSQSFTDFELTLAPASSVEALTTVMELKRGEG
>Alpha-galactosidase (EC 3.2.1.22)
MTKLIEVDEQQQIFHLHNEQISYLLTVEAGQTLSHLYFGKRVRGYHGQLTYPRVDRGF
SGNLPGSDDRTFSKDTLPREYSSAGEMDYHLPATVVRQADGANALNLQYVGYTVLEG
KPKLAGLPAAYVEDDSEAETLVITLRDEVTGLDYQLSYTIYRALPVITRSVKVVNNGTQA

VQLEKVASMQLDFVDRDFEVISLPGAHVNERHPERRPVSFGVQSFESRRGTSSHQMS
PFIALVDKQTDEFSGDAYGFNFVYSGNHKFSLEKDQLAQLRLLVGINDYNFTWKLEPA
QSFQTPEVLMAYSDQGLNKMSQTFHQLIGKRVVRSKYRDLKRPIVVNNWEATYFDFD
EEKLKPIVAEAKQLGIEMFVLDDGWFGHRDDDYSSLGDWQVFAKKFPQGLGHFADYV
HNQGLKFGLWFEPEMISYDSDLYRQHPEYLMQVPNRAPSPSRSQYILDLCQAEVRDN
VVEQVATILSQGQVDYVKWDMNRHLSDIYSLSLPADRQGEVYHRYVLGYYDLMERLT
AKFPDVLFEGCSGGGGRFDAGVAYYMPQIWTSDNTDALARLTIQMGTSLAFPPSMMT
AHVSAVPNHQTGRVTPFETRGAAAMSAVFGYELDLTKLTSAEKAAVKEQIAAYNQMR
ALVLEGKFYRLKSPLDSNQAAWNFVSADKKQVFVAVFNVLAQAQPYLTKTKLAGLDKD
KCYRNLATGQVYGGDELMNLGFYDPVVRYDFASQTYLFEAVTEE
>Sucrose phosphorylase (EC 2.4.1.7)
MALENKVMLITYPDSLGKNIKELTEVLTDDFGEAIGGVHLLPFFPSTGDRGFAPVDYQQ
VAPEFGDWADIKKLGERYYLMFDFMINHISRQSRYYKDFQAKSSASAYAELFLSWDKF
WPEGRPTKEDVDLIYKRKDRAPYQEITFANGQREKLWNTFGEEQIDLDVTSEVTQKFIK
DTLRFLCAQGADIIRLDAFAYAVKKLDSNDFFVEPEIWDLLAKVKADLAGHGEILPEIHE
HYSLPFKIAEHGYFVYDFALPMVTLYSLYSGKTAQLAKWLKMSPMKQFTTLDTHDGIG
VVDARDILSPAEIEYTSQELYQVGANVKKKYSSAEYHNLDIYQINTTYYSALGNDDQKY
FLARLLQVFAPGIPQVYYVGLLAGENDIELLEKTREGRNINRHYYSRAEVKAEIERPVVA
ALLKLLSFRNREEAFDLDGTINVTCPSQNEIVIKRQNKDQTRTAVLTANLKTNAYQVKV
NDQLVTFE
>DNA alkylation repair enzyme
MGQFADLKALLEANADKETARGQAAYMRNQFKFYGIKTPQRRSLTQNLIKAAKRAAKV
EWDFLDQCWLAEQREYQYFVCDYLKAVKGQLVYADLKRLELYVQTKQWWDTIDALDT
VIGQIGLSDKRVDQLMLDWSQASDFWLRRIAIDHQRGRKEKSKQELLAKIIDNNLNREE
FFINKAIGWALRDYSKTNPPWVKEFLASRKNGLAKLSLREASKYLR
>tRNA-Gln-CTG

>/ Glutamine transport system permease protein GlnP (TC 3.A.1.3.2)
MKKFFIWLSFLLGITFITTLVAKPEQVLAAEKTYTIATDVTFPPFVYANADNKYVGIDMDLI
KAVAKEEGFKVEIKPIGFNAAIQAVQAGQVDGMIAGMTITNERKKTFDFSTPYYSSGIV
MAVKPGSGVTKLTDLKGKKVAIKTGTSGADYANSIKKKYGFKTVTFDDSDNMYDDVKT
GNAAACFEDSAVLEYGVKTGLGLKIVTKPTNTGDYGFAVKKGANQELLAKFNHGLAAL
KKNGTYDKIIAKYTGKKALADKKNAASGHSEYTIMGLLKQNKTTLWHGVLETLWLTIVGI
VFATIFGVIIGLLGVVPNKLAQGISTTIIYIFRGLPLIVLALFIYNGVPSLTGTKVPAFIAGVIT
LTLNEGAYIGAFVKGGIESVDHGQMEAARSLGLPFGKAMKKVVLPQGIKLMVPSFINQF
IITLKDTSILSVIGILELTQTGKIIIARNLQGFRMWAIIAVIYLLMITILTLISKWIERRMKN
>Glutamate transport ATP-binding protein
MTEYKVDVTNLQKSYGSNHVLKGIDFKVANNEVVVLIGPSGSGKSTLLRCLNKLEEPTS
GSIVIDGQDISLPSTNIDKARENIGMVFQHFNLFNNMTVGENIMLAPVELGKMDRQTAE
TEARQLLETVGLSDKFDAMPASLSGGQKQRVAIARALAMNPDVMLFDEPTSALDPEM
VGDVLEVMKELAKNGMTMVVVTHEMGFAKEVADRVVFMADGYVVEEGKPEAVFDHP
QHDRTKAFLDKVINV
>hypothetical protein
MIKVILVMHDQNGDYYKMNKTFFESMPKVGEYIYNTDGLAYVVEEVAQFAGYVSSKGA
IAILVVHQADENHPVSNLYGLDIERDLDD
>Multidrug-efflux transporter, major facilitator superfamily (MFS) (TC 2.A.1)

MVTEGEKMPSTWRKTFIALWFGCFMTGLGASMTMPFLPLFISSIGNFSRWQLNIYSGIA
FAGTFLTQALISPYWGSLADRKGRKLMCLRASGVMAITIFLTGMAVNVWMIIALRLTQG
LFSGYINNATAFIAGETPRSKSGSVMTMIMTAGVTGNLLGPLFGGALADTFGYRIPFYIT
GVCMALAFLMTLFNTTEHFKPISKKEMQPMRQMFADLEKPGLIIIMFLTTLIVQSSLMSIS
PIISLLVKELMHNHGNVSFVSGIVAATPGFGTLLVASWLGHKMDVVGPLKVLMVGLLAA
TLLFIPMFMTDSPWLLAFWRFLLGVANAALLPAVQTALTINVPAKSFGRIFSYNQSFQA
AGGVIGPLMGSLVSSALDYQYVFLVTASLVFVNLLLVVGAKVQQIKTK
>contains DUF161 domain
MFLVALEIIALSINFFYAPIDIAAGGTTGIAILVEAVWQIDRSLTVLVCNLAMIVLAGVFLER
KVVKNIAVGSLLLPLLMKITPSFKVTDNDLLAVIIGGVLMGFAVALLYRVNASSGGTTVP
PLILKKYFYLDTAFSLMLIDLTVILMNVFVAGLNAFFLAAFSQVVTALVMRYTEAGLDLKY
QVQILSNDLLPQIKEALLAENSSLTVYDVVGGYSQTPRQQLLLIVDRQDYGRLITSIHEID
PQAFIVTSQVTKVHGGKWGM
>unknown
MASNSDSIFYVLSYLRRHPEAFYFAKNKYDNVVQIFIKDDLQIADADIYFPQNRLMVNRL
QDDFLAQHGNLLDHFWEQSGHRPSGYHEVWATSSHLIDSNVFLIELSYE
>Phosphoglycerate mutase family
MAKRLGQELKQIKFEAVYTSDLKRTQKTAQYIIDENEAYYPPVYQDPDLREISFGKFEE
AKNTKMIPKALKALGLRRILQALLNEEHVAELTELFRSLDGNEVIEDADHLTLRMTRALT
NIGNQYQDKGGNILIVTHGLILSNFIESLEGEVPLFLLDNSRASRVDYENGKFKVIYVNR
MKLDDKGMAKDAIS
>Beta-lactamase class C and other penicillin binding proteins
MLYPKTTALIHELVTKQVVPGVSYAYLHQGQLQTGVFGLAQLKPTKEKLRPGMYYDLA
SLTKVIGTTTVILKLIASGRLGVDDAVTKYLPAFKDSRVTIRHLLTHTSAISGYIKNRDELS
PAALLAALHQLKVGSGLGKKVVYTDIGLIFLGEIIESFYQQPVQRVIEKEVLARLGLSQST
FNPPLSLAVPTELTAKRGLIRGQVHDPKAYTLKEHCGSAGLFMTLHDMVKFSQWLLQP
DSTPGPLDSKMVSALFKDWTPSGHDGRSLGWDLRYTWQNQACLYHTGYTGTMLLID
QANQDALILLTNRVHPAGNNQAFLAQRELLLGQYLREKYFYEKDGSK
>Mannose-6-phosphate isomerase (EC 5.3.1.8)
MTAPLFLKPYFQEKMWGGKRLKTEFGYEIPSDTTGECWAISTHPHGPAVIENGPYQGL
NLAELWQQHREVFGNAKGDVFPLLTKILDARDDLSVQVHPDDAYALKHEGELGKTEC
WYVLAADEGAEMIYGHHAQTKAELKQLIEAGDWQHLLRRVPVKAGDFLYVPSGTVHAI
GKGIMVLETQQSSDTTYRLYDFDRVDKKTGQKRELHIQQSIDVINVPHVDPKLNREVK
QVAGIEVTTFVDTDFFKVQRWQLAGQAGSLIRNSDEPYTLVSVIAGSGKLTVAGIDYEL
QKGQHFILPNDVKEWQLEGTLTLIASVPGKKA
>Inosine-uridine preferring nucleoside hydrolase (EC 3.2.2.1)
MKKVYFSHDGGVDDIISLFLLLQMESVDLIGVGVMGADSYLQPASEASRKVIDRFAPTK
KLAVANSNARAVHPFPKEWRLDAYSENALPILNESGKIVTPVAAKPAHLDLIDKLEASSE
KVTLLFTGPLTDLAMALELKPTIADKIEELRWMGGSFLKEGNVAEPDCDGTQEWNAFW
DPEAVKTVFDSKIPLTMVALESTNNVPLTNDIRLSWAKNRRYIGLDFIGNSYAFVPELRD
FPTNSTYYLWDVLTTCSLYKPDLVKEKEIKCDVMTQAPSDGRTYLTENGREAKLVYDV
DHDSFFETIMELAKRAQ
>MutT/NudX family protein (putative)
MENRLIAHSLIKAHGKYLVIRRNKIKRGQKNVYPAHWDIPGGRVEAGELPRMAAIRECR
EEVGLEIELENILHEDSNLDGQLVFTRLVYAAHLLVDSKTDVNLDFEEHDQYRWITGLG
DLAGENVVPYLQSLLK

>Putative amidotransferase similar to cobyric acid synthase
MPYSLHVAHLYGDLMNTYGDIGNILVLNYYAKKINVTLTSEVVSLDQTFDPTKFDLALFG
GGQDYEQVIVSKDIQSKKAALTSFIEDDGPVLAICGGYQLLGQYYIGANGQRIEGISALP
HYTKSQDHNRFIGDIDIKNEETGEVYVGFENHNGVTYLGAGERPLGKVLTGHGNNGQ
DQTEGAIYKNTFCSYFHGPILARNHNLAKRILLLALKRKYPQADFSQAEATLAQID
>proposed amino acid ligase found clustered with an amidotransferase
MSIKSSLAISAGKSSYWFLHKFLHGGTALPGKITLQLDPDVLKSLAKDYEIVIITGTNGKT
LTTALTVKVLRQKYPNLLTNPSGSNMKQGIVTAFLTHKKAKSGKNLAVLETDEANVPLIS
QYIKPSAFVLTNLFRDQMDRYGEIYTTYEKILTGVRMHPDATIIANGDAPIFNSVELPNP
VIYYGFDHRVDHDQKAPANTDGLLCPKCQRILAYKNLTYSNLGKYYCPNCGFKRPELT
YKVTKLISQSPQASKFEVNGQPLTLHIGGTYNIYNALAAFAVGAHLGVSPDLIQAALDDP
DEKVFGRQEVIKLADKELTLILVKNPVGLDQVLEMIKTDPEPFSLAVLLNANYADGIDTS
WIWDGRFEDLPFERIPAVLTGGERYKDITFRMRVAGVPEANFDQEPKLEAVLAKVKQM
PTKHVYVLATYTAVLQLRKLLASEGYIKGGME
>Thymidine kinase (EC 2.7.1.21)
MKKVAQLFFRYGAMNSGKTIEILKVAHNYEEQDKQVIILTSGLDDRDEVGYVSSRIGMK
RKALPVFSQTNIYDEVVKHNADADCILIDEAQFLQKHHVLELTKIVDELGIPVMTFGLKN
DFRNELFEGSKYLLLYADKIEEMKTICWFCQKKATMNLRVHNNQPVYAGEQIQIGGNE
AYYPVCRRHYFNPDLEMMSKGE
>Peptide chain release factor 1
MDKLFDRLQMLEDRYEELGELLSDPEVIADTKRFTELSKEMADLRETVEKYNQYKEVS
QRISDDEEMLNEGLDDDEMTEMVKAELADSKDEKQALEEEIKILLIPKDPNDDKNIIMEI
RGAAGGDEASLFAGDLFNMYSKYAERQGWSIEVIDRNMTEVGGFKEIAMLINGKNVYS
KLKYESGAHRVQRVPVTESAGRVHTSTATVVVMPEAEDVEIELDPKDIRVDVYRSSGA
GGQHINKTSSAVRMTHLPTGIVVAMQDQRSQQQNREKAMKILKARVYDYYASQEQDA
YDENRKSAVGTGDRSERIRTYNYPQNRVTDHRIGLSLNKLDRIMNGELEDIIDALILFDQ
TKALEKLQDA
>Protein-N(5)-glutamine methyltransferase PrmC, methylates polypeptide chain release factors RF1 and RF2
MPKMTYFKAQQWAFSFSAEHGYLKEDVDFLLLGQLNWTLAQLLSNYQREMPAAAWE
KFQTNVKKMCQGYPPQYLLGSTSFYGLKLQVTPATLIPRPDTEELVDFILTDLQTKKDL
RVADIGTGTGAIGLALKSNKPSWDVSLTDISEAALAVARNNAHALGLDVELKQGDLLAP
LVGRYDVIVSNPPYIPKREADLMDESVKRYEPALALFAAKDGLAIYERLAKEIRPYLQPK
ARLYLEIGFSQAKQVQALFLQEFPRAQVTVKQDLTGHDRMIKVEF
>TsaC protein (YrdC domain) required for threonylcarbamoyladenosine t(6)A37 modification in tRNA
MLITKIYQPSEVKQAAKELTLGELIAFPTETVYGLGADATNEQAVKRVYLAKGRPSDNP
LIVTVASREMVKQYTGELPDAAVKLMDKFWPGPLTMVFPIKDGSLPKAVTGGLKTAAF
RCPANQLTRTLIAEAGVPLVGPSANSSGKPSPTTAQHVYHDLEGRIAGVLDDGPTSVG
VESTILDMSSKQPTILRPGAVSPQEIAAVLGQEVISDKHHVGQNEVPKAPGMKYKHYAP
AAQVLIVAQDKWDAALAWAKDHANSVGIMASPQVISRAPGLNTYDLGADVVTASQHLF
AGLRYFDDQAEVKWILAPAYPEEGLGEAYMNRLKKAAGNQFFEA
>Uracil phosphoribosyltransferase (EC 2.4.2.9)
MGKFEVLDHPLIQHKLTIIRDKNCGTREFRQCVNEIAELMAYEVSRDMALEDVEIETPM
GKTVQKRLAGKKVVIVPILRAGLGMVDGILELIPAAKVGHIGMYRDEETMKPHEYFVKM

PDDLENREMFIVDPMLATGGSAIMAVDALKKRGARSIKFVCLVAAPEGVEALRAAHPDI
DIFAAALDEKLNENSYIIPGLGDAGDRLFGTK
>ATP synthase F0 sector subunit a
MDEKTQIVTFLGIPFNLSNIISGTVVAFLVFCLVFWLSRNIQMKPSGRQNLLEWIIDFTNG
IVKSAIPDEAGKSYGLLAFTLFVFIFVANQLGLAIEVSYNDYNFMKSPTSMPMTTMTLAFI
VFLLAHYMGVKKFGFKGYFVNTFMSPMKALLPINILEQFTNFITLALRLYGNIFAGEILLT
LIYKLAKINAFVFIPASVLELIWQGFSVFIGSIQAFVFTTLTMVYISQKIEKE
>ATP synthase F0 sector subunit c
MGSLTLIGAGLAAAGAAIGAGIGNALVISKMLEGMARQPELSGQLRTNMFIGVGLIEAVP
IMAIVISFLLVFQ
>ATP synthase F0 sector subunit b
MNSTFLLGAAEEGFAWGDFLFYLVLFIILMALIKKFAWGPVTEMMEKRAAKIATDIDAAE
ESRKKAEDLAQKREEALQASRTEASEIVNRAKQNGEQQKATIVSAAHDEVVTMKANAK
KDIQQERQEALASVKNDVAELSIEIASKIIQKELSAADQKALVDSYIEGLGKQNGIR
>ATP synthase delta chain (EC 3.6.3.14)
MALDNPTVARRYSQALFEVAQEKGQLKDLSLELNELKKGLLAQPQFMTFMASPAIKPE
AKLAMINAMTEEASALTKNLVMMAYDYGRIANLVAIIDAFNNLYDEYEKVVRVKVTTAIE
LDKEQEDKLSDSFAKLVGAKKVIVNPQIDESIIGGVVLQSKSYIYDGSIKTKIERIKRLLLK
>ATP synthase alpha chain (EC 3.6.3.14)
MSIKAEEISALIKQQLAGYQDQLEVAEVGTVTYVGDGIARANGLDNALSGELLEFENGT
YGMAQNLEANDVGIVILGAYKGIREGDTVKRTGRIMEVPVGEELIGRVVNPLGQPVDGL
GEIKTTKTRPVERKAPGVMERQSVAEPLQTGLKAIDALVPIGRGQRELVIGDRKTGKTS
VAIDTILNQKDQDMICIYVAIGQKESTVRTQVETLRRYGAMEYTIVVSAGPSSPAPMLWL
APYAGAAMGEEFMYNGKHVLIVYDDLSKQADAYRELSLILRRPPGREAYPGDVFYLHS
RLLERAAKLSDELGGGSMTALPFIETKAGDVSAYIPTNVISITDGQIFLNSDSFYSGIRPA
IDAGTSVSRVGGAAQIKAMKKVAGTLRLDLASYRELESFAQFGSDLDEATQGKLNRGR
RTVEVLKQPLHKPLPVEKQVVILYALTHGFLDTVAVDDILPFQDQLFDFFDANHQDLLA
EIKTSGQLPDTAKLDAAIKDFAQTFQASVK
>ATP synthase gamma chain (EC 3.6.3.14)
MPASLQEVKRRIESTKKTGQITGAMQMVSTAKLSQIQKHAVSYQVYAQKIRSVITHLAK
SHLLDSATTGESETSQDGKINTLGMLKQRPVKKTGVVVITSDRGLVGSYNSNVIKQTM
SLVNSEERNKDNVVILAIGGTGADFFKKRGYEVGYEYRGVKDIPTFREVRPIVNAIVTM
YSNGQYDELYVCYNHFVNTLTSAFRAEKMLPISAENMGSDDFNTPDENGLMADYETE
PSQDAILEVVLPQYAESLVYGAILDAKTSEHASSSTAMKSATDNAKDLIASLELQYNRA
RQSAITTEITEITGAQAALE
>ATP synthase beta chain (EC 3.6.3.14)
MSLGKVVQIIGPVVDVQFPLNEALPDINNALHVKREDGSVLVIEVTLELGDGVMRTIAM
DSTDGLQRGAEVVDTGAPISVPVGEDTLGRVFNVLGETIDNGEEFAEDHRRDPIHRDA
PAYDQLNTSTEILETGIKVIDLLAPYIKGGKIGLFGGAGVGKTVLIQELIHNIAQEHNGISV
FTGVGERTREGNDLYFEMKESGVLEKTAMVFGQMNEPPGARMRVALTGLTIAEYFRD
VEGQDVLLFIDNIFRFTQAGSEVSALLGRIPSAVGYQPTLATEMGQLQERITSTKKGSV
TSIQAVYVPADDYTDPAPATTFAHLDATTNLERALTQQGIYPAVDPLASTSSALSPEIVG
EEHYQVATEVQHILQRYRELQDIISILGMDELSDEEKVVVARARRIQFFLSQSFNVAEQF
TGLPGSYVPVEQTVKGFKEILEGKYDDLPEEAFRLVGPIEDVVEKAKKLANQ
>ATP synthase epsilon chain (EC 3.6.3.14)

MDAKSVITVNIVTPDGSVYDATTDLVVCKTTVGEIGIMPNHIPTLAALAIDRVRVRKGEE
KFDEIAVSGGFMEFSDNTLSIVASAAEKSEDIDTNRAQRAKERAEKRIAQAKESHNVDE
LKRAEVSLRRAVNRISISSHR
>Phage protein
MLLTLKKWGNSSAIRLPKSLLEEVGIQDDASFEVEVKDGSIVLKEKKEVESLEELFSDFD
TESYFRNKDTLEYDWGKPQGKEIF
>UDP-N-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7)
MEKIIVRGGRRLEGEVEIEGAKNAVLPILAAAILPTKGRTVLTNVPILSDVYTMNNVIRFL
NVEVGMDEASKTIELDATGKLSYEAPFKYVSKMRASIVVLGPLLARLNHAKVAMPGGC
AIGSRPIDLHLKGLEALGAKIKRHDGYIEASTDGLVGANIYLDFPSVGATQNIMMAATLA
KGTTVIENVAREPEIVDLANILNKMGANVKGAGTETIKIVGVDELHGTEHAIIQDRIEAGT
FMVAAALTKGNVLIKDAIAEHNIPLISKLEEMGVTVKEEARGIRIIGPDKLKPVVVKTLPHP
GFPTDMQPQVSILQLAASGTSLLTETVFENRFMHLEELRRMNADFKVEGNTVILYGPT
DFNGAEVAATDLRAAAALILAGLVARNITQVTHLQYLDRGYYHFHQKLAALGAEIKRIND
VTGDVEYRLQRDLIK
>Rod shape-determining protein MreB
MAKDIGIDLGTANVLINVKGKGIVLNEPSVVAVDTMSNRVLAVGSEAYKMVGRTPSNIR
AIRPLKDGVIADFDMTEQMLSYFIKKLNVKGFMSKPNIMICCPTNTTEVEKKAIRQAAEK
SGGGNVYLEEEPKVAAVGAGLDIFKPQGNMVIDIGGGTSDIAVLSLGDIVVSRSLRVAG
DKLNQDIVTYLRKKRGLLIGEHTAEKIKIEIGTVLKGHRDKEMEVRGRDVVKGMPESITV
TSAEIEEALHDTMMSVIIAAHEVLSQTPPELAADIIDRGIVLTGGGALLDGIDELFSNQLQ
VPVVVADSPLDDVAVGTGYLLDKVEQTGK
>hypothetical protein
MGFSRILRCNPFVPGGIDYVPDKFSLRRNPDKTYREIKSDKE
>hypothetical protein
MSKKDKNIQITEEEKMVNGQLVTELTAKKSKLGQVIADQDKFIAVLPSGERFNVKTENE
ALDLLIRDFHLHRVKS
>Rod shape-determining protein RodA
MLKKQFLSMKADDDARIDWGIIFCVLVLAVIGMVSIYVAVKHDTSTASITRVMLSQAIWY
TLGAIIVIFLMQFDAEQLWKVAPVAYGIGIFLLIMVLIFYSRSYAALTGAKSWFALGPFTF
QPSEIMKPAFILMMGRVITQHNSEYYEHTMRTDWLLLGKMIAWLIPVAVLLKLQNDFGT
MLVFFAIFGGMVIVSGLTWRIIAPVIAIATVLGSTAIFLVVYNRAILEKIGFKAYQFSRIDSW
LDPSASTNGTSYQLWQSMKAIGSGQLLGKGFNVSKVYVPVRESDMIFSVIGENFGFIG
SCLLIFIYLLLIFQMIQVTFDTKNEFYAYISTGVIMMILFHVFENIGMSIGLLPLTGIPLPFVS
QGGSALLGNMMGIGFIMSMRYHYKSYMFSNQDESDFK
>Histidine kinase of the competence regulon ComD
MGNYVLGMISTLLIAIITVTYVSKIKFDLAHILCFTLLMLGCIPIKISNLIMILGGVLITHHQD
QGKNQWLRYTLFVMYAFVTTWILQVILNFYGSQFASWINYNNYLGTFVCQSVVLLIQLG
VIYLFRDSLEILALSPSNKKKGAEIVAIVILAVYLVYGSTFDNDYNNWWISAYVVGLILFW
SYVRALALLETKRELAELRNFELKNLENYANEMEEAYQKLRRFRHDYLNILLSLDQAIKS
GNLGLIRQTYDKVLAPSKAKINMDYYDFGKLNHIKTSAIKSILYNKFSYATHEGIKLEIEIID
DVKIQRTKLLDAVRIISILLDNAIEAALESNTPKLLVAYIKDANTERVIIENSIALERVELAPI
FKEGYTSKSGDRGQGLATVKRLISHYPNVVLNTSSHDYKFTQELILSKEE
>Response regulator FasA or ComE or BlpR
MRVFILEDDSLQSLRLQQILASITDEVIVQSFKKPEQLLAHVSQPAKDNVYLLDIEIGKDL
YSGFETSLLLRSIDNLGTIIFVTTHSEFMRKTYEYRVAALDFIDKFQTDMEIRQRLLSDFR

LIKERQALSVQTKMIDLESENDFLRLPISDIYFFETIDDFTKRKLSLYTKSQHLIVKGNLKL
VEAKSERFMRVHRSFVANIDNVIKLDRKKRQLYFNDNTYCLVSRSKLAKLREILQTKFIE
D
>Conserved T+S rich domain protein
MKIKGKVVIGLLLLLGFAISFKEQRVLAADNDRYLAYVGQSQDSFGMPIYVENIKTGKVD
VAYCINESKKLPYSAIYNRFDATSERLMALDKPLVSRELLPGIIKRIIYNGYPLNATNLQG
SLSDGAFRRLTQLAIWHYSDSFNVDESELTPQERWVYHQLIQGTNLVPVNLNLNIYVR
DDSSYQDLLGTSFGKLPEHPQFSSSSMASSSTSSSSTSSSESSNSSSSITKAESSSST
PTVSSSSMASSSTSSSSSSSSESSNSSSSITKAESSSSTPTVSSSSMASSSTNSSSTS
SSESSSSSSSITKAESSSSTPTVSSSSMASSSTSSSSTSSSESSSSSSLITKAESSSSTP
TVSSGSTTSSSTSSNSSSSSESESSSSSSSVSTTNSSTENLVNHSEDMSSSGNKSSSV
KLAKSDAKQVVTQKSLPTNSQVISKPRQDRKSAQKQVKLPQTGESSEHFLWLSLAGCL
LLMGALFLFERKYR
>Transcriptional regulator, RpiR family
MNFNTDIIDKINSHLPSMSESDQKIANVILNNPQAIVNYTITELANKAAVSTASVSRFCKN
LDLVGFHQLKILLAQTDSKRDDYFKEVSLDDFSASLNKISANKVAEIRHTLLPLDSKKILPI
LELIKKARLIQISAEGATFPVAEDAAYKFNQIGFLAIATASYETALAQTINLSEADLLIAISN
SGESRELLSQITIAHQKHVPVLALTNRQDSPIALDADFHIKTAVRQRVLESQYYFSRLAT
LSVIEVLFLLLIADDPKRLAKIEFHEQLVAQKKI
>Ribose 5-phosphate isomerase A (EC 5.3.1.6)
MQTIVDLALQKLRPQMSVSLGGGRNVLALTEELAKQPSKNLKLYSPSELTRAYGQKLG
LEILASEQGRHFDLAFDGCDSIDYQFNCLKSKGGINYFEKRAAQKSDEYFILTPKTRLNE
TLDAQLALSLEVKLERCFEILAAAQQLGLRAEIRLGQAVASYARTPLGNVLIDCFATDW
QQIKDYNEQLLAKPGVVASSYFEGLVTGIIASEENGQGIIFTR
>oxidoreductase, Gfo/Idh/MocA family
MAKKYNWGLIGTGWIAHEMANALQAVNGEIYAASSPNQAELEKFVKEKQVSKAYLKAD
DLINDDQVDIVYVATPHTYHYEYIKKALLAGKHVLAEKAMTVNLKQFNEVYELARKQNLI
LMEGFTLYHMPIYRQVADLIAQGKLGEIKLVQVNFGSLKDYDPSNRFFNKELAGGALLD
IGGYATAFARMFLKQQPNSLLTTVKFFETGVDEQAGIILKNADEQLAVMALSMRAKQPK
RGVISGTKGYVEISNYPRATKAEITYTKDAHQEEYQTLTAGKSEEALNYEVFAMQKFIE
QGHDEGQLELSRDVAEILNSIRTNWGLKYPFD
>Arsenate reductase (EC 1.20.4.1)
MKKFYCYSRCSTCKKAKKWLDEHGVVYEQQDLVELPPTKEELLTWISNSDQNLRYFF
NTSGQHYRQLGLKDKVAQMSVAEAAELLASDGKLIKRPLMVEGTKVTCGFKEDVYQA
TWLN
>hypothetical protein
MAQLKVTPKVNWGRKILLVALGLALLGNYLRGTQLPVDVSGLSPLILFRISQLGVGLATI
KILFIDANSWRFKAGFLAMSLFLYYVGSNAHSLDIFFYFVILVAAKDEDYRWLLRCYLIEV
AVIMLSLALLTFGGILNDMTMVRPGRTLRMSFGSVTPSDFAAHVFHWLVAYFCYRNFK
LSWIEYLVSLVATWGVFRFTGTRLDLILMAVTLVLAVSYPKLKKLINYLDGWSLGGIMLV
YLGFNLAINYLFTPKWTWFRRVDELLSHRLTYGQLALQKPLTLFGRYFRENGNGKFNP
HQPYFFIDSSFVRLLAMQGILLFTCFVIMLIFLAWRFKAERTYSLLIGLGLVLLSAAIDHHL
WELSYNFIFLASLTDNQHFIVKRK
>FIG00630584: hypothetical protein
MSMNKEVLKTKSIGKWILFFLIVVGLVPLFTMLAASNIITKNVLIERNNLSKVSAVDMIQE
ERTHLQKSTEKMLIATAKYGEFTSGNFDKPKLKTEMQKLVSANSSLLGATVADQAGEF

ASTINYPQGYQVTEQNWYRQAKQNLNRVVWTLPYKDPKSGVFVITAAYGYKAANGKF
FVISSNVSFNDIERPLGQLKIGKTGRVTLISDKGIALASKGAIDDAHYSRGNDLSSDKMY
QAIAASSKRKGYVHVKGNSKVMDIYFNKGAVGSKYWAYAYVMKDDLKAELHSVLISAS
IIALVVAVLVLILAFIINGIFKEIVGHILSYFRAAEKGKMNPIEVKPATSRNFKARVAAKLYT
PDEHGSEFNQISYGYNKMIAAVGGLVKKVKGQANNVAEKSDSLLELSTQTDKAIDEVA
QTITGIAEVTSSQAQETQESVTRLHNLSEIIERLHANAQTMNDEADKSASLNQANMETM
DSVNANWNQELVEMKDLSQSVQDMNTSIQDITKIIGVINDISRQTNLLALNASIEAASAG
EAGKGFSVVATEIRKLAEQSAASTKEIEAIINQIKDKSHEMVTKTNNSIEGGQKQSELIQA
AIKSTAEVYKSNQTMTEQVAQVANASAEIEVVQGKVLSRLENISASTEENAAGTQEVSA
NAEEVLATMEEFTSHVADLHDIAAKLKAETDKLTVN
>Para-aminobenzoate synthase, amidotransferase component (EC 2.6.1.85)
MKPLIGIIGNSFDNPIVEHANLPINYTPKGYTNAVHAAGGLPIIIPLVNEDNITEIIDRFDGL
VFAGGQDVSPIFYGQEPHPKLGSLNYQRDLVEIALIKAAFAKGLPILGICRGLQVINVAL
GGSLYQDLSEYPNWSVKHQQFGTAWAMPTHSIKVEPANPLHDIFGDTGLVNSLHHQA
IKDLAPGLQALATSADGLVEAVTNKQQNLLAVQWHPEAQFEDFPEALAIFAWLIQASQA
KD
>hypothetical protein
MGLSKLLPIIPINGFMKNHSKKKSCSLIIEEEAKKYH
>Iron-sulfur cluster assembly ATPase protein SufC
MSTLEVKNLHVTIDGKKILKGVNLTLKTGEIHAIMGPNGTGKSTLSEAIMGNPAYEVTEG
EVLLDGQNLLELPVDERARAGLFLAMQYPAEIPGVTNAEFIRGAINARRSEDNPISIRKF
LKKLDENMDFLDMSEEMADRYLNEGFSGGEKKRNEILQLMMIEPKFAILDEIDSGLDID
ALKVVSKGVNKMRGPEFGSLIITHYQRLLNYIIPDVVHVMMDGRIVAQGGPELAKRLED
EGYAGLRDELGLDIKLTDEDA
>Iron-sulfur cluster assembly protein SufD
MDLTNYPELATYAKQVAEPAWFTEERLLALQASETLKMPSFQKIKYRHWPINLTKQLAY
EESKLEVPEEIDQAKATAVQVGQTTKLLNLSPALKEQGVIICDWQTALKEHADLIQAYF
MKKAIKSTENRLTAEHVAKLTSGLLIYVPKNVVLTEPLVTYFVQDATTGADYVHHVLLIA
DSNSEVSYLENMLTVGQSESTANIIVEVIALDGSHVKFSSIDRLGQKTTAYLNRRGYLK
RDAKIDWSLGMMSDGNIVGDYDSELVGTGSHAETKVIAISTGKQVQGIDTRVTNYGKN
SIGHILQHGVILQNSNLIFNGIGHVMKGATGADAQQESRVLMLSRHAQGDANPILLIDDN
NVTAGHAASVGRVDEEQLYYLMSRGLPKPIAERLVIRGFLGPVLAAIPSRAIQIELENMI
ERKLVDGQRDE
>Cysteine desulfurase (EC 2.8.1.7), SufS subfamily
MDKEMSKYRADFPILAQTVNDEDLVYLDNAATSQKPQAVIDAIVNYYRNDNANVHRGV
HTLAERATAQFEAVRQKVARFINAPTSEEIIYTKGTTEALNWVARSYGERFVKTGDEIVI
SYAEHHSNLVPWQELAKRVGATLKYLELTADGAVDLAKAEATITAKTAIVAVNQASNVL
GVANPLAQLAQLAHQNDAILVVDGAQGAPHMVTDVQKMGADFYAFSGHKLLAPTGIGI
LWGKSELLAQMAPLEFGGEMIDWVELTQSSYKKAPLKFEGGTQNIEGVIGLGAALDYL
EQVGMDKIAAYEQELVAYVLPKLQAIPGLKLYGPKDPKKHTGVISFNLDQIHPHDVATA
LDMEGIAIRAGHHCAQPLMRYLDVAATARASFYFYNTKQDADRLVEALLLTKEFFGNG
TI
>Putative iron-sulfur cluster assembly scaffold protein for SUF system, SufE2
MGLSKLNNLYRQIILEHSQHPHHHGHLPRADKSLELRNPTCGDVLTVELALKDGKVSE
VAFSGTGCTISQASASLMTDEVIGKTPKEIEKQVLLFSKMVTGEVTDEDQLDEVLGDAS
ILAGVAQFPARIKCATLAWKAVYQVLESEGK

>Iron-sulfur cluster assembly protein SufB
MMTEEMPKLGGEYQYGFKDDITPVFTTGEGLTEEIVRAISAAKNEPEWMLEFRLNAYKI
YRKLEMPKFGPDLSKLDFEHINYFRRDTDKVARDWDDVPEDIKKTFDRLGVPEAERKY
LAGSSAQYESEVVYHNIKEDFEKMGIIFMSTDDAVQKYPELVKEYFGKLIPPEDNKMAA
LNSAVWSGGTFIYVPKGVKTDVPIQSYFRINEGNSGQFERTLIIVDEGASVNYVEGCTA
PNYSKDSLHAAVVEVNVLKDAYCRYTTIQNWSDNVYSLETKRAQALENATMEWVDGN
LGSKVTMKYPSVYLNGEHARGTMLSIAFAGATSEGEGVDSDTGARMIHNAKHTSSSIV
SKSLCKDGGRVDYRGQVRFERDSDGSFAHVECDTIIMDDKSSSDTIPFNEILNGNVSM
EHEAKVSKISEEQLYYLMSRGISEEKATEMIIMGFVEPFTKELPMEYAVELNRLIEYQME
GSVG
>hypothetical protein
MHFESRSHYSSNLGREMEFNVYGHGGKIFLVFPSSGGSQNEYADFGMVSACQTFIDQ
GLVTFYTPNSFDNESWLAKDKSPHEMAAAHDAYDRYIIEELVPLIKYERNYYGALGVTG
CSMGAFHAVNFALRHPDVFDTVIAQSGVYDARFFTGDYYGDELVYHNSPVDYLWNLD
DTWFLDQYRQNDYIICIGQGAWEEVADTRKLEEAFNAKQIPAWFDYWGFDVDHDWP
WWRKQMPYFLTELRAAGKL
>Carbamoylphosphate synthase large subunit (split gene in MJ)
MTAPLNFIVVSPHFPPNFEPFTLRLHETGFRTLGIGDCPYDQLSPSLKTALTEYYRVDD
MEDYNQMYKAVAYFAFKYGKIDRLESHNEYWLLQDAKLRTDFNIPGYKNADVDVIKYK
SKMKEVFVKNGIPAAKGQLVKDKEQTEQFVSQVGYPVIIKPDSGVGASDTYKVNNQAE
LESFFTKKRPGVAYFLEEYLEGEIVTFDGLTDQDGKVVFYSSMVYGQPALETVLDACN
MYFYEAREIPADLKELGLKSVAAFKITERFFHFEFFRTKDKKLYALELNCRPPGGPSIDL
MNHAHQIDLFKAYAVLVKENHFIYGNDCPYNSVYVSRRKGPEYLHSVADVKTKYAAEL
KDYQEVPAGFADIMGDIGFIFNTKTVEQLKEISAYVHASKQ
>Predicted glutamine amidotransferase
MRPRIAVPADTIFDATNVINMRNAAFAPRPIIEAIVKSGGVPVIFPSVAPKYVKDYLDLFD
GIIFAGGVDVDPTFYQQEPHQNLGATYYKRDVFEIELCKQALAAGKAIMGICRGMQLIN
VALGGSLYQDLSEDKKAYLKHSQIAPGNMPSHHVSTKENSRIRGLVGERPYVNSRHH
QVLNQVAPGLVVSAWADDQVPEAVESLASDQILAVQWHPENMYKHYDYSQALFNDLI
ARSQRAH
>Glutamate--cysteine ligase (EC 6.3.2.2)
MFSKVGKVIFDNEVVAKTSDFNMGLEIEQHRVDEQGYLSQEPYPATIGDGAKNPWLTK
DFLEAMSESVTPAAATALDALHYLYRINNSLRSALAPGELLWPLSMPPKLPADKSTIQL
AKTTPEKDAYLKEWAKRRNFSSGTPCGVHINLSLNPRVVDIVYNNLRGQFANRMQAQ
TYLYTIIAQGFVRYRWFLTYLFGASPVAEENFFEKNQGPTKPVRSLRQSHYGFGTHFS
GDYSSVQAYVDRIEQGAKEGKLISDYEFHGSVRFKGGSSLKKMPAEGVDYIELRMLDL
DPSSSVGVRSDTLRFVRLLASYFVMTPALKPADVNEVVARADKMNEEVSLEEPEAVSK
YQVLARAFMKRLEIFANKLQLGPEYQEVLQDLEDRIENPSTTPSARLLKHLKDGSLVPY
ALERAQRYQDAALQSLKVFAGFDSEQILSATELSQQLFEPDAKATLAKTTLAKTK
>Amino acid permease
MDSNETHGKYISWPILAMMAFVTVIGFEDIMYNFKNQGMGVITSWVLMLFLYVIPYALM
VGQLGSVFNHEGGGLSSWIRGTKGEFLGYFTAWTYWAASVPYVVDSANSVVVGFGW
AFTGSGKFQDTMSNSTFMLLSALIFVIFIFVQRHLAHSMELLSTIGGFAMFGMTVLFVLL
AFAGLIKNGGHMATQPMNIHTIVPKFDLKYLTTIGMLIYAVNGSELVAPFITRMKRPQKD
FPKAMITLSIMTAFLTICGSFALGIFFDANHLPNDMKMNGDYYAFQALGKLYGLGNFFM
YLYAWTSVIYMCALLAVLLDAMTRMLIADTGSQYMPAFLRKTNKEGLPINGYFLTCFLS

AFIMLLGAFLPEMNDIFNWLLNLNGIISPGVTCWIFYAFMKIRANSAKYPSSYVFIKNDKL
AWLCGFLLLAVTAIATVFGIFPQDTSQFSSTWWYELTINIVAIVVLIGLGAILPSIRAREVK
YGLAFDKWQRNSLIILCLGSITLGVWLGGSQLALRGLWIFIQALVALGLMYLIGRRQPKQ
H
>putative alpha-dextrin endo-1, 6-alpha-glucosidase
MMARKANNYYLELKTHWLDLPYYHEKRRVRVLLPKDYHKETWATYPVLYLHDGQNIF
YSREAFAGYSWKLIPTIKQNIDLPKMIIVGIDNHGEKRLDEYGPWQSDLPTEGQAGGDG
MAYADWVVNTVKPFIDHKYRTKPEHQSTLLAGSSMGGLITAYMGAAYPDVFGTLGVFS
LCSWFSQRDFLRFIASHPLNKQTRLYIQTGTKEGDEADAAFISDMNQAYIDCALNYYQR
TLEAGLPLTQIWLRILANETHSEYYWAKHFKEFLSFAFVDQKFWQ
>Permease of the drug/metabolite transporter (DMT) superfamily
MSEKKKGFSLAIIGCMFWGLSGNVAQDLFEQAHIGPMWLVGVRLLGAGLLLVIWSHLT
LPVAKVKALKTRAYLPRLVCFALLGMIPSQLTYFFAIKHGNASSATILQFLGPLFIIIYLALK
QLQWPRRVDIISIVVALVGTTLLVTNGNLNQLALEPLGLFWGIMAGVSQASYTLLPSKLL
ASFDAKLVVGWAMIIGGLCFIPTLITTKVAALSWLNLAEIGFIIVFGTMLSYLFYLKSTSYIA
PATTGMLSAFEPLTASLVGLGLGTVHFTLIGILGACLVLSTAFLQSLASK
>FIG00749076: hypothetical protein
MLVDFLNEYYQAKLANIHDVAGNQHLVGYSQLYRCDLFVKIFKAEAMFYAEQHVNQVY
CPEIYLDSVIFEDNYIVVLKDRQLTDIDSNEVKVTEKKAEEYGRLLANFHQRLTDNVSVY
HDQRKLSQRIKVYVDALQASRYQDDINQAYALLLADLPTADREYELLPKVVLHGDYSLR
NIKGYQGKEILIDFERAHIGVAYEDFIKIFYNEVHNVALRHTFIQGYKQVMNFEIPRSELQ
RALLFLIALEICEFHLSHHKEKFGDMAHKMLATIERGDAVLKV
>D-alanine--D-alanine ligase (EC 6.3.2.4)
MVAKKLHVALLFGGNSSEHDVSKRSAHNIYDALDKDKYEVSVFMFTKDGFLLGNKDS
MRIFDGENEDEVVAQAIKDVDFSNPLANVQNLSEVKDVDVFYPVIHGNMGEDGTVQGL
FRLLNKPWIGSGVASSAVSFDKDLTKQLLTLHGIRNTRYVVVTPENQAEYPYEKVAADL
GETLFVKPARQGSSVGIYKVRNKEEYDKAIQDGFHYDFKVMVEEAIKNPREVECSVLG
NRKADASQLGAINIPEGDDFYDYNNKFVDASGVVFQLPVNLPKKLTKEIQQMSLDAFR
VLDNRGLARMDFLVDENDVPYFGEVNNLPGFTNISLYPQLWALSGISYAELIDKLIQLAI
DEFNDNAKLHYDFVELGQEKVGKGLIKKDGSKTE
>hypothetical protein
MKLHTVKAYEPLALNLAAAKDNIYRDLAAKRKVGKAGLLDWYLTSQKIAGREYHLLVQP
EAKLYIPFTPQLGDYFEDILRFILQRLAVTGEQVEKMHAGLFNGRGVEVILNASEPELGF
ISANEAKLKRLLPGELAEFKSRDAQIGYCALLGYWLSHFATEKEALGEITINQILGNYYPI
KMKKPNKKFSYVTLTAKYKDFTTWEADLGLSSDQFGFERVKKEIMANNKKAINQFIALN
ASHLKDWGGKKVVRSYLEGYLNDQLFLGNQLRTVLTNAADFPLYVYKQGEYFLNSNQ
PAVMIGVVSCLYQSVLPFFDFLMQSGGLAKKDNERLQASFADALYWLKNLLLADKDKA
KRAKEVFAKIPATEIERLPVEIKAWYQELVKALK
>Thiol peroxidase, Tpx-type (EC 1.11.1.15)
MQVTFHGEALDTSGTPVAVGSDFPQFKLTTKTGQTLTKDELAGKVTLISVVPDINTRVC
SVSTKKFNQEVDKFKQVAFYTVSTNTVAQQAGWCAAEGVAKMEVVSDEAKDFGKAT
GLYIEAKGIDARSVWVLDQAGKVVYRELITEQTNEPNYQAVLDFLTKLA
>Lysyl aminopeptidase (EC 3.4.11.15)
MGQVAHLLDVFVPKNYQLYFDINRQTKKISGVTKISGRALKTDLALHQKFLNIKTVKVNG
QAVAFTSSDDSETVNFTSPTTGPVELEVSYDTALTDTMMGIYPSYYEVDGVKKQLIGT
QFETTSARQAFPSVDEPAAKAIFEVAIKFDEQAGETVLSNMPEVQVEDGVHYFAPTLK

MSTYLVAFAFGDLQAKKTETKSGVEIGVFATKAHQAKELDFALDIAKRSIEFFEDYYQTP
YPLPHSWQLALPDFSAGAMENWGLVTYREAYLLLDPDNTPLATKRVVATVIAHELAHQ
WFGDLVTMNWWEDLWLNESFANMMEYVAIDALEPNWKIWEAFQTSEAPLALQRDAT
DGVQSVFVEIKDPAEIDSLFDGAIVYAKGSRLLVMVRALIGDQALRTGLKNYFADHQYA
NSVGADLWDALGQASGLNIGAIMKSWLEQPGYPVLDLKVVDGDLIVSQKQFFIGQGQD
QNRLWQVPLHANFDAPQLLTTKELNLGSYQELRAKAGHALRLNLGNTSHLVVDYSDEL
LADILTEVDTLSAIDQLQLLQDLRLLAESGRKSHAELLPLLLKLRDSKSAIVTETLSQVIAK
LKFFVNPDSQGEKDLKALVASLSQAQLKRLGLEAKASDSNDDLLVRPLVLDAALYAEN
AELIAAAHDLFLANQAKLVALPVAVRALVLKNEVLNFNSDQLFDNLLVSYQKTSDSSFK
QDLLAALTKTKKQPQLEKLVSHFEDADVIKPQDLRAWYVNVLANPRGEQLAWDWLRN
EWSWLEAKVGGDMEFPTYLTVTARVLHTAERLAEFKAFFEPKLTNPEISREITMDTKVI
ASRVDLIQADQANFLAALAKVTA
>Mobile element protein
MLCVQKVRLYPNQTMKRVLDDLCAYRRYCWNQGLALWNDMYDASLILEDKKLRPSA
RKVRDELVANKEDWQYQLSARCLQLAISDLGKAWSNFFKKSLPDWGKPKFKSKKTAR
QGFKTDRAQIVNGKLRLDKPHGVKSWADISFKGAKSLTGDLKVVSIYRENGKYWASLP
FEVKQTKKVKTDRQTAIDVNIGHFDYTEGQVKTLPNNLKALYKRIKHYQRLLARKRVAN
GKKATQANNYVKTRAKLQRDYRKVANIQHDIVQKFTTMLVNDYDRIAIEDLAVKQMQM
SHVASKGLHRSMFGYFKQVLKYKCEWYGKELILANCYYPSTQRCPACGHVKTGENKV
GLDGNKKHKTKHNEYICYDCGAVMDRDENAVRNLLALL
>Shikimate kinase I (EC 2.7.1.71) # AroE I
MLIGFMGSGKTTISQLLGKILAKPVVDLDQAIVKQAGSEITEIFNRHGEQYFRQLEHDSL
KRIITESGILATGGGTPLRADNAEILATSGAKVILLEASGETILERLKTDCSRPLAKQMAK
QQLLALKQARQSHYQQCADLTIQTDKLTPKEVVAKILASLETVV
>Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)
MAEFRVNQRLFGVALGIVLVASQLFTYSVLAVGVALIFSPLLVGAVNFRQKRQVNYISW
LITMLALGIIGLPFLGLSTYFDLTTHLRLGWHWLDFFTNNRRLMVPVLGLGCGLLTLASLI
WVKRLGEFSPQTLAKRYLTGALLVVSGQLILSFLLIKLGQVENSLAYLTLLASAAWVLHL
GLVYLFYSFLLTGHLVKLRLGKLGLLGIVCALSLVGWQTRTYLSQPLARGVKVIAHRGV
NQDNALGNSTLALQKTKQEVKPNLVEMDVQATKDRKFIVSHDEDLKQLTGKALVVDQS
NLKQLTGLRERQAGKQAKLVSFTTYLSKAKRLDQPLLVELKVPNKNYRQTNSGQEFG
QTFANELAARDQVHSMSLRALWQLKAVSPRTSVGYVLPINMWGLVFGPLDFYSVEQN
TLTRSLVWQLHHRHQQVYVWTVNRPADIKRVLAKGADGIISDDSSQVKREIKRLQNDK
KSLRLAKFRVILNEFGYGLRI
>DNA-3-methyladenine glycosylase (EC 3.2.2.20)
MLKCPWAYLSPEMEAYHDKEWGKPLYDERALFELLCLEQLQAGLSWQTVLKKRPAFR
KALANFEVAKLANFDESDLAQLLANPALIRNRRKLEAILANARLIYEQQLSLAKLCWAYV
DEKPLINHWQSSEQVPAQTDLASQMAKDLKQVGFKFVGPTTIYSFMQASGMVNDHLE
SCPCKEA
>Abortive infection bacteriophage resistance protein
MNIWQLSYEQQLALLKLRGIAEINLNYVELSKLRLNKTRYLSYQRQLRSIESIGYYVIKEY
ANPFYNGKRYNGIKFDDIVNRYFYDKHLKQHVLQAIETIEVALNAKIAHILGNRCGAFGY
LDFNNWCQLSGINGYLGKKKKVNKYLLQDEELRFKKKLANAVRKSSNPDIIDFIGTTQNI
YPSVWLMVGMLDFGGSVRILKLMDEELRKEMAKYFKLSEYDLIKWLETINLIRNICCHN
RNFIDFSLRTKPPINDRIKPFLYSHSEGKKIVYTNKVALPILMIVYLVKAINPRYDIGKLMK
SVRKFATSDSEAQKIGFENKNSIGIKRVNLSI

>Universal stress protein family
MLQRYKNILVPVDGSYEAELAFKKAIDVAKRNGNDTAIHLVHVVDTRAFQNISSFDTAM
VEQVTDTAKKTMAGYVEDAKKDGLTNIDYAVEYGAPKVVIAKDLPKEKDIDLIMIGATGL
NAVERLLIGSVTEYVTRVADCDVLVVRTDLENKPALSKAQRKAQQKKD
>hypothetical protein
MILKILAIILLIIVSSILLFMGGYLLGHQRKPFLIFHPESTPTLGGVLKICGLILLALGALSLFIT
IIGQILWALLIVLCAAFFVSILSFIMLTYL
>Uncharacterized ATPase (AAA family) associated with cysteine desulfurase
MQPLAYRMRPQKIEDVVGQADLVGPGKIIRRMVEAKLLSSMILYGPPGTGKTSIASAIA
GSTKYAFRTLNAATDSKKELEQVVLEAKMSGTVILLLDEIHRLTKPKQDFLLPHLESGRII
LVGATTENPYISINPAIRSRTQIFEVKPLSPADVKIAINRALTDKKNGLGKLKINLTSEALE
HLSQVTHGDLRSALNGLELAAKSTPPAADGLITIDLGVIEECVQKKALVHDKDGDAHYD
VISAFQKSIRGSDADAALHYMGRLIEAGDLPSISRRLLTIAYEDIGLANPQAAARTVQAV
TAAEKIGFPEARIPLAVAVIDLCLSPKSNSAITAIDAALADLAKGNYGEIPSHLRDSHYQG
AAKLGRGVDYKYPHDYLHDWVKQDYLPSKLKQKHYYQPKDNGRYEQALKEQYLKLC
QAQGRPFS
>Glutamate--cysteine ligase (EC 6.3.2.2)
MQVDAIGNQIQIDNLTDVLENISYQLTKNVEVLEADGQWLKADSLKENQYLKQNHHSV
SLTTNPNLNLAELKSELSLLTETLSKQLTRGQKVFPFSQASKKHHGRQPGYSLKVVLPE
NLFKLLYTSKFKETKIDYLSFRNQVYLKLAQQLVAKRAYMTYLFGATPFAWQEGVSEEL
SSPKRSVTNSLNQVELKEANALDYLNLESYLASKSSASLTSDNVNLNLIEIDGKQTIEAL
QITGLDFNPVSETLIEGPALEFLNVCLGYFLMTEGIGAGDLKASLSQSRVLNQTVASEN
PFAPSVVAGELRKFLEEVNHFASAYYYPAWQPAYTKLRKRLADPKASLSASLLRAQGE
ADSLYSLALSGGFKTETGKQTLSYELQTMLTAAIMANHKFRILNQELSLVQIDETILQAG
LKTKENSALLEAMWANKQVSKQLVSAGGFETLKAWQVKSLQELETLAPKLADKALAIK
SVSDHAAKASRLFRLPPSSRQLKASVQAVLKEQKQALLEQVAPGSTYRALIIKGKLVSL
VERIPANVVGNGRASLKELTASKHLILGPVERETLASQGIGLNDVIARGIQVLLRYDATE
NTGASQVESLADLDESYRTAIEKIAQILGMSEGQIDLIIPNIYQAYQQEPGQLYFLGAHR
QINLALHLQVLMAANKELPATILDKLLKAN
>FIG00743012: hypothetical protein
MSDQLEDRLTQSLAGTPHLKPDEQRRYLGTFRERVSLTLTYDQAQSDFYLARLKDEM
LAHPDYQLAIRSDLNETPLNKLLSFASQHELKTVLKADEASQNPSPYALVLASPSQALH
VEVIDVAEKYPQEENSHPKPSHPSFWQKLFD
>SSU ribosomal protein S4p (S9e)
MSRYTGPSWKVSRRLGISLSGTGKELARRPYAPGQHGQNNRRKLSEYGLQLREKQK
LRMTYGVSERQFANLFKKAGKIREGKHGDNFMILLERRLDNVVYRLGLATTRRQARQL
VNHGHITVDGKRVDIPSYEVTPGQVISLRERSKDLTVVKEALEAVVGRAPFVSFDENTM
SGSLVRLPERDELDANIDEALVVEYYNKL
>Free methionine-(R)-sulfoxide reductase, contains GAF domain
MTQSLLAQQIDALLTGETNQMANLANASALLNDSLTDINWVGFYLYQADTAELILGPFQ
GKVACMHIKVGQGVCGTAFAKKEVLRVANVHEFAGHIACDSASNSEIVLPLHNAAVT
>Septation ring formation regulator EzrA
MVAALVAIVVLAVVIYLGALYFQRLNNARISAIEDITAKIEANSQEDELAAISKLNLTGESL
AAFEELDKSYRALINRELPFLVERLKDVKEANQHFSFPTVKSDLETLEAKGTAANMAFT
KIIKQIKQIKQQTDQHQQSIDSLQTKYQEIRKKLLSKNFFYGPAITRLEAALADLEKDFKE
YENIAENGDYVKADKALAKLKDDTKTLEVALKQIPPIYQKLHNVYPDQLTDLKDALAKM

QEEKYLFKEDMTAKLADLDQQSQALTTALTKLDLETAKLLEKELGHQIEATYTVFEEEY
TAREKVEAKTKFYADFIEHAEKQERRLSLELDRLGQNYTLNNDEKARERQLAEQLKEIH
AAYTSYQAALMTDGVIYSQIWQQMEQAFNDLKTIEEEQGRINDGVANLWKEEKQAWE
AVNQFNLDLAKYRRDLEKLNLPGLQTEYLDYYYRVKDEVEAVYQSLDHVKIDMDAITKS
LIDTQADLDALEEKTKDIVDSSSLTEEMLQYSNRYRNRYPEVAQAYQQAKKLFSEDYD
YVAALDTIAQAVDKVEPGAYDKVSARYHKRQQG
>Arsenate reductase (EC 1.20.4.1)
MTKPKVAFICVHNSCRSQIAEALGKLLASEVFESYSAGTVKTALNPDAVRLMKAKYQLD
LTQSQTSKTLAALPAVDIVITMGCNVSCPALPCRYREDWGISDPTGKDDATFEAVIATIE
AKVRQLARRITAGEFENDV
>Cysteine desulfurase (EC 2.8.1.7)
MIYFDNSATTKIAPAVLETYTKVSQQIWGNPSSLHKMGEQAYLLLEQARKQIADLLGVM
PGEIFFTSGGTEGDNWVIKGTAFEKRHFGNHLITSQVEHPAVLNTMAQLEKLGYEVTYL
PVDKNGQVNPADLKAALRKDTILVSIMAVNNEVGAIQPLREIAAILDDYPRVHFHVDAVQ
ALGKGLGELLSLDRIDFLTFSGHKFHAPRGIGFIYAKAGRKLAALMAGGGQEQNQRSG
TENLPAIAAMAKAIRLLLTDEDAKVAKQQAVRKMIYDHLAKFDKVTIFSQLTPNFAPHILC
FTIDGVRGETIVHAFEEFGIYISTTSACSSKKGTISGTLSAMKIPEKIATSAVRISLDEANN
DKEAAEFIKAFDILYKRFAKINS
>tRNA S(4)U 4-thiouridine synthase (former ThiI)
MQYTEIMVRYGELSTKGKNRRNFIDRLATNVRQSFHDFPNLVIKAHRDRMHMHLNGE
DADQVMARLRQVFGIQNFSPSIKVDRDMEEVKKVAAQIMQEKFFDGMTFKINTRRSDH
RFMYDTNEINDMLGGYILEHVPGIKVKMKDPDITLRVEVRLNGIFLSYETIQGAGGLPVG
TAGKGMLMLSGGIDSPVAGYLALKRGVDIEMVHFFSPPYTSEQALNKAKELTAKLTPY
GGHIKFIQVPFTEIQETVKHDVPSGYLMTVQRRFMLRLTAALAQKRHGLAIFNGEALGQ
VASQTLESMAAINDVTTMPVLRPVVSMDKNEIIKIAQDIDTFDLSIMPFEDCCTIFAPPSP
KTKPKIDKARAYEAKLDVDGLIERALAGVEITEIKIGDTFLNQEEESFAQLL
>Valyl-tRNA synthetase (EC 6.1.1.9)
MAKEMSTKYDPSQVEPGRYQKWVEAGLFKPNADKNAHPYSVVIPPPNVTGKLHLGHA
WDTTLQDLIIRQKRMQGFDTLWLPGMDHAGIATQAKVEARLAEDGISRYDLGREKFVD
KVWEWKGEYADIIHNQWAKLGLSLDYDRERFTLDEGLSKAVRKVFVSLYKKGLIYRGE
YIINWDPKARTALSDIEVIHKDDKGAFYHVKYPFADGTTFNGKDYIEIATTRPETMMGDV
AVAVHPDDERYKDIVGKTLVLPLQGRHIPIIADQYVDPEFGTGMVKITPAHDPNDFEVG
NRHNLERINTMNEDATMNANAGKYEGLDRFEARKAIVKDLEDQGYMLRIEPIVHSVGH
SERTGVQVEARLSTQWFVKMKPLAEAALANQKTDDRVDFVPGRFENTFTQWMENVH
DWVISRQLWWGHQIPAWYNKQTGETYVGEEAPSDIENWEQDTDVLDTWFSSALWPF
STMGWPDEKAPDFERYFPTNTLVTGYDIIFFWVSRMIFQSLEFTGRRPFKHVLIHGLIR
DEQGRKMSKSLGNGIDPMDVIDKYGTDALRWFLTNGSTPGQDIRFSYTKMDSAWNFI
NKIWNASRFVIMNLADDTKPVLPPKSEWQLADKWILNRLNQTVAEVTRLFDNFEFGEA
DRALYNFIWNDLCDWYIEMAKENLYGDNAKLKANTQNILCYVLDQTLRLMHPIMPFVTE
EIWLSMPHDGQSLVTAAYPVFKEELVDEDADRQMSVLIELIKAVRNTRAEVNAPMSSAI
DLLIKTKDEEVKAMLEANKDYLDRFCHPKELQIGADVTAPKLAMTSVITGAEVYIPLGDLI
DLKEEIARLQKEETKLVAEVTRAQKKLGNERFVANAPAEVVNAEKEKLADYQAKLDAT
KERIATLAKEA
>hypothetical protein
MAQSKFEQEMYDEAYKVASYLRQTGQDSIFLRLVKEKGSRFFLNATALSLTLAGMSMK
LALRLTTLVLRYGLRVVTAILDILDGKAKGIKVTIETNK

>Substrate-specific component BioY of biotin ECF transporter
MTRSTNITAIRLTKIALITAVVVVLGFLPPIPLGFIPVPIVLQNLGVMLAGVILGAKDGTISIL
LLFLIGIILPAFSGANTQGVLVGPTAGYVIAWLFVPALIALCLAKLPKQSATFQFVVIWLFG
VLFIDAFGAIYLAVYSHMPIFKSLLANLAFIPGDTLKAIITLIVANRLTKLK
>hypothetical protein
MIKIYGALTDQAGRCRHYHSQLDIVALKCATCQKYYACYRCHDEVTDHHFVASSSSEP
YPVLCGNCSCLLSKSQYLLGHCSYCQAAFNPNCRFHHQIYFKA
>Exonuclease SbcD
MRVLHTADWHIGKKLYEYDLYAEQVAAFEQIRQIARQKKVDAVVIAGDLYDRSVPSEEA
VSVLKQMLVKLNIEDKLPLLAISGNHDSARRLGVGDEWFEHNQYFLKTELGIDAKGRRT
LEPVELGNCQFFLVPYFNPKASRTYFEDDSISTMAQAMAKVVAECESYFKPGLTHILVG
HCFVAGSSHEDSEVSSTVGGLDEVPVSLFKNFDYVALGHLHNHQALLADNARYAGSP
VKFSLSELTNQKGVYIYDTETNDREFVPLKDPTPMVSLTGSLDELCQPAKYNEVADQA
FVFVNLTDRQVIANLVEKLRACYPRLVHIERQNGLEGIVQTPSVELLKEANSPLKMTAAF
FEQIQGLTLSANQKDYLAQLLAEIEEEQ
>Exonuclease SbcC
MKPLKLTLKNFGPYLDEEIDFTRFEASPLFLINGKTGSGKTTIFDAMTYALYGTTSAAGK
TGRDAKEMRSKFADAKAQTKVSLTFSHQGKVYQVERSMKEKKTATTLDSKVELSLIEP
KPAKLLAAKTREVERQVTALIGLDAKQFSQVIVLPQGEFDNFLKADANQKEALLERFFD
TAIYKRLEEKLLAQSKELEAQLESSNQLINNQLAAYPEYEVKTAPEWLAKVTETLAALEE
SSQQEASVITGLEDKQTQLNQQLGFQEQLHAQLEAKETLVKRQTELAAKQAEFEANKT
RLDQLTWAKDQLSNWQLFQQNQRELATLKERHQKEELTQGDLQAQLAALDQQAPAV
AKNTADLQALRQKVSYLKEIEPHYEAYVNKQSQQLALKQNLTSKQVHYNDLQADLKAA
QTTEQALRQQISNLPDLKQLQEEFYRTERQLANLTQERTSLTKLYQAYLDAQAKLTRTK
EALADKKAKLAVAKEKFTTLDQEYTSLQIVRLTQKLQPGQPCPVCGALDHPLVNQGQV
ADLAKVKQVEGDLEQAQANLTSLQSDYNQLEGRLKVQKQTHLAAQTDYQAELKQVLA
NHGLQSGGELKQLANQFQVQKQHLESSLKAASAREAELKAKQADSQQKIDQLNLKLE
QAQAAVTQAQQDFQAVNDELIGLKAKLAPDYQTKDQVLSERQKLELRCAQLDQELTN
FERNKSQVKEALLTTKTNLSNLEENLVKTQTQIANYQTSLKAAIAQVNFVSNFDDLGYLL
SHAAESEALKAQVTAYQNEVKLVTEQLAAIEQQLAGVTVLDLEELRAKLGAVKQDLQV
KRTALEEVTRKKERLHDSYVQVKREVSKNEALLAKLDRLNQLTAPISGKGVGKIDLQRF
VLQQYFQKILGNANVILQTLSQDRYLFQMTTTQAKSGTKAGLVLDVVDKISGERRRTQT
LSGGESFVASLSLALGLAETIQAENGGVRIEALFIDEGFGSLDTDYLARALDYLQTKEGD
SRMVGIISHVRELQESIPDRLEINNRNGKSTVSYHL
>ATP-dependent nuclease, subunit B
MALGFILGGAEIDHEQAILEQLLAWQKQAPSENYFVIVPDHIKFETEVKVLEFLRQHYSS
AKGLYAQTKTQVFSFSRLAWYFIKDQQSVQGRHLSQMGLTMLVQSILSRLDSLHVFAA
EKNQVGFAQKLAQQLLELKQGGISPDDLDNLVANLPKDQAKSMDLRAKLADMALVYR
QFELELGDKYVESSKVLQSLNDYLLEADLSQTSFLISGFDQLSALELQLLKTLVKRAKEV
RISLVLTKQGLEQASGDDNLFNRPKRLFTALTNWADKNHLPLAELSWPTVKRLSPDLE
RLGKYWQTSVSLKANKPAPLSSKENVSVVALANRIEEVRYVASQIRQEVAKGERRYRD
YMVLAPDLADYENLIAPIFNQYQLPFYLDLSKGMGNHPLVEFLNALFTLAKHPNLYSYH
DLMRLLKTELFIPQLYVSDANGQLVLKQDLATFRDELDWLENYLLRTGISSRKRFEADD
FWLVERLPLKPNQSQAEHEAAQRKLLHNQNANLLKAQVAKLLASFSRALAKVKTNRDF
ARVLYQFLIKAGVDQSLATSQARAQKRFEEGKLNLVGINAQDATRPEQVWNTLCALLD
EYVLALGDQDFSLSNFSEILNAGFKAAKYKQVPTTLDQVTVSKANLSQRRDRRVLFFL

GASDQALPARFENNSLISEEERQLLASGDWLARDQFLAPGAEMAMADEPYQAYLSFL
TAKDRLVFTYPQNDADGQLLACSPYVKRIMAAFGLKETDVRDRLTAFLGTPRTLLTDLL
TQSRAAFQNQADLDSQWQAILTYELAVNQTLTLKLLNSLDYHNEILPNTKPSADGHKRL
IKELLKELYRRGKNAALFGSVSRLESFYANPYEFFLKYGLQLEKRKEFMLSPADTGTYF
HNLLDEFMRLIHADKTYLKRLTSREFDHYMKRAVTNVQQQAKEQEGYLFDASARLDFL
RQQLESRAYQVGQAIWKQRQDQAIYTLETEATFGFQANAQLDGLVFPAGKRFAQDET
LNLQGRIDRLDLVISELNRHYLSVVDYKSGNTDLKLKDFLTKALNGLSLQLLTYLAVLQK
DSNQAKLKQLLLEMDLPPAEQLEVGATTYLNLFSGRLAFKDYQPLLEKKKTFEDFFMQ
QYQYKGLFRLEEDADFLLALDQKLKAKPAKSLPYNISHQSSDKFKANGNSQLLTFEQLE
QLVQLNLQKIEQARQEIFGGIIDLKPYKLDNQTDGLLYSDYLPIMLFDPMVGNEYRKIKT
LSTEEVWAQIKHELQKG
>ATP-dependent nuclease, subunit A
MESKFEGYATKNQIQAVKSEAQNILVSASAGSGKTAVLVERILYQVIGHYLADDQADDE
SLLGEVANLLVVTFTNKAAKQMKDRLEKALRERLSQVQAASLPAHKKQRASQHILKQL
RTLNLADISTMDAFCSRLLKRYYYALDLDPNFRILADQAESQMIKADVWERVREQAYAK
MGQELAQKQPSAFFAVYQTFSNGKGDEALTDVVMSIDTYANASANPEAWLANLAKLY
QFDAQADLTQSDLWQNYLALDLNKRLAKVSSEFTELLTELTAFLEAAKHFWQTDVYGR
YYEGKGAAKEKKFSADLAKREQKFADLLAAYQATSQALTAYQAQLTTANYDQLRSGLD
LNFSFKPILRQAGKTAGDFPESLAAKWNQVKAKKDELKTEVAKIKASYFIYSQADFKAIL
TKSQLIIGELARLTTEFRSEFAKEKTRRHLLDYSDMEHLALELLTGSQPEQVRIQAGIKA
HYQEVMVDEYQDTNELQEALLTAIAASPNSSYFMVGDVKQSIYRFRQADPSLFINKYQ
TYQAEAAADELIILPQNYRSAANVTGFTNLIFSQIMDASLGEIDYDQAAQLIAANQAYSK
ETSQVPTELLLYETKQDESASPIAIDFELDSKEQAQVNLIAHKIQELVGQQEFYNRDSG
RTETLTYKDIAILASTHAQGTLIAEEFKRYGIPVEVSDTDNYFKTTEIQIMLSLLRIIDNPH
QDIYLVATLRSPLYGFDENELAYLRIVTRGSDFYGTLVGFKQTYENAKDQLFRFKLDND
FAFEDQLAKKVNHFLTELEAFRKLSRSSELAPLIWTIYEQTGFLEYVSGMPGGPQRVAN
LHALYERASQYEEISYKGIFQFLRFIEQMQKNDQDLAELEVKADDNAVSFMTIHKSKGL
EFPVVFVMDLGRSFNLGSVQRNNYILDDKLGLGLALVDEKQNEQADLPESLNVTSPTP
VYELIKTRQLKLALAEEMRKLYVALTRAEQRLYLVGHVKDAQDALGKWQLASQSDGKL
LAEAQRLKGRTYLDWLGASLMRMPNLADRISFAGKREEQLDLPTPSPVLADFKEVNFK
LEFWNEAELSQILVASDQPKVDQELSLNQWLAKQTAVASQAPASVKQVADFTYPHEAL
TQTTAYQAVTDIKRLFDDPDNSLMENIEITDVTATRLNYRQAEQGQASLNLYGQANFK
QPEFMQTEVKVSPAAVGTATHLVFQELDLNQPVTLDTVAATIDKLVLAKLITTEVAALIN
QKAIGRFYESDLGQAVIANRASLERELPFSLLYPAGHLFEGVDQASADPILIHGIIDGCFE
TPDGQLVIFDYKTDAINDFQTPEKVVARYRGQLNLYSLALESIKHKKVTHRYLYLVASG
QVVEI
>Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA (TC 3.A.1.5.1)
MKRKKLWQSGLLLAAGVVGLASVNVTSVSAASKSVLNWSELSELQTLDTAASLDTTSS
EALRNSYEGLYRLGNNGSIHKGLVKTAKLSKDGLTYTFTLQKNAHWSNGDKVTAKDFV
YGWRRAVNPETKAANSQLYDGIKNASEIVKGQKDPKQLGIKAVGNYTLKVTLDRKIPYF
KSLVAAPVFYPLDQKAVAKYGKEYGTAANKAVYNGPYTVSGWTGSNQKWTLVKNKH
YWDKQHVKLAKVNFQVAKSTSTSYNMFQANQLDETNLSTEQARQMKTTKEFVARKQ
ARLNYLQFNLKNKELANANIRKAVSYAIDRKQLVKRVLGDGTVAAHSFVPKELMVTNG
KDFADYAAGKVGTSYNKKQAQAYLKKGLQELGLKQLNLDLLGDDDDTSKKINEFIQSQ
LEDNLGSKKVTVNVTNINKKSRIARMQSGDFGIVQTGWGADYQDATSFLGLFSADSVY

NFGKWNNPTYDKLLAQAKTQSGTKRLATLKQANDTLLADQAVVPLCQPALATLLKDKV
TGVAYNTIGQYNFKEAQVK
>hypothetical protein
MYYHALLANNMLAPGEDYQNLRETYVIFLCGFDPFGLDWAVTDYQMTANHGKHILGD
GTHTLILNAKGDWREVNDELKGIFQLVLKQRATYGHLGLEIARNIKNFKESKAGGRKYM
ALAGDFWQNVWEDAREEGREEGRKEIVLLFVQKLAKDGQSQPEIMTYLETELGLSED
EAQEYYQLAMTKA
>hypothetical protein
MLELGKGYQVAAGQDLVEGYEQTSETYLTANVGVKKLVEVCQHFIAIHPEPLFFILELP
VTADREKEVAPGVLTESHRDVYYLDGCSQEVCLWLLEQYGDLLINDGMSKFGFGCHQ
SQDEIMLGQYNVVTIYSQRLENYADFFAAHEIFQVNQLVTAWDTFSAQTPGKAERVDY
QGKSVYDLPAELKEVGMYLGETRAE
>COG0553: Superfamily II DNA/RNA helicases, SNF2 family
MYPQFIDNERKTLADTLRDIAKDYDVLRIATGYWDLAGTLKIINEIKDYKKIQLLIGQEPLA
DYLQKKYNISDIDSGLFPDTYVQSDLEGYGSSEEINALRETAKELVRLIEEGVLEVKVFR
KPRLHAKVYIFGDIGDGNSVGIIGSSNFTGAGLSSSKELNFLTNDYKIVEFVPSSDKQEN
GHLTWFKNLWEDENAEDWTGDFTEIISKSPVGNMTYGPYDVYIKTLMEVFPDELIDISP
FDTKTEKILYEFQNQNALSLRRKLEKMGVAMLSDSVGLGKTITAAAIINQYINDGKTNIVII
PPASLKQQWVDELESDRWNLVEHRDFEIYNQQDKDKIAELIKKSKQRKNTRNEVDLFV
VDEAHNLRNVNSVRHQQILELFQENPHAKVLLLTATPINNSLMDFANIIQLGSKGDLVSV
NVPYQMQKSDLEYIDFFKALRNIQSEVTRAEKRGERINWDLYKNTLTSGIRHYIVRSTR
QGVEKRNAMKPLKEGRHLFPETKVEQISYSYGKENAEYISNLIVKGIDSYLEGLDPRKL
NLDFVSELTQRTEHPIDLYKKIKKMQDGGKYEQVAEDNNVSEKYARMNLFDDTIAKTVI
PLLYQLTNFIGFAPYKPDAYRNDIYGKTIPEIRAKEVGGKIRTQLAIHNMLHVTWLKRLES
SMSALEKSVKNYERRLELFEKWLDRGYIISLSDISTLENDYGEDINRAFDDYEKYLEELA
TVESGNEELIKKKGIEKRVADESEFRINKLKKDLDRDKRIVDELLKILDLLIESNHDEKLN
VFANNLVQLTKDKKYGKKVLVFSFFSDTIDYLKEALPPLLEDKIPDFSKKAVFVSGNSKE
VELNAKRFSPKSKKYKLENGETELDYLFATDVLSEGQNLQDAGILVNYDLHWNPVRMI
QRNGRINRLGSPYDEVLIANEIPNDDLEAYLKLVRRLERKIDTINFTVGNDQSVLGEEAN
PIEFNDMLDFFSSNSKKASEAAKKLENQDDPLTWTDDYSLELRQFLEDHKDDGEIERIK
SIPKGKWNYLPKRKNEVSDPSEVIGLYTALGKYTGTGEKIRDVGFVKIKFSGGRGPFKA
ISAEYIDEQEALKMVKTDSSDNRGSMDTINLNRKTYIHSGEEEIGVHFKKSDSLFKIEPK
QEDALKRISTHFTKDVLGIVRKGIKRSNEKRIFDRLAREINREVKEGGAPFPSTVRKFEIF
INRLLEKEAQEQKLNKIEGVLFYAHK
>putative type IIS restriction/modification enzyme
MPSHTEDAKWMSLAIALDLADELLKVMMPENGKWLNQLNRDTQHTGALPEALPLDTF
KNKKQVTWFKEHPNSKSARSAFKLFYDEDCALESNFFWFSESATKQKISAATEYSAN
WEDSDLTRRPEYKVGIDFFLTPDTNKLLLVLSNHYKLRVLELHEHLSNTQKQIIKNYLNG
AAAYSGIKDGQKEDLEPQRTIHTTLWNALQLKEVNKQFYGYIAGHFTELVSKLEEQGKS
SDDAKQFSSRLLGRLLFVWFLRKMEIINESMGYFDVNDLSATDYYNQKLKVLFFKTLNT
EVDSRETGDLVTPYLNGGLFEAKENDFADELLEFPEGFFVRLYNHFNEFNFTTDESSA
DFELVAVDPEMLGQVFESLLASQTDEADSNERNNTGSFYTPREIVGYMVKETLRQYLY
TKIDKVAHKGIDELLDSSDYQWLARKSTSKVDVWGVNSKRVISKIKTALDNFKVIDPAA
GSGAFPMGMMQQLLKTYERIGTDKDSYNLKLQIIENNIYGVDLQPMAVEISRLRAWLSV
IVDETNKEDVKPLPNLDFKFVAANSLVKLADGQTSLFGDPDLDIKLETLRDKFFNARKA
HTKKEYQEKYYKLTFGQLNMFDDERTHQLRSFDPFKNRTHADFFDPHVMFGVDEFDA

VIGNPPYIHFEKMDVEFRNFYKKRAKNLGFETYAARGDIYTLFYEKGISLLKKDGILAYIT
SNKWMRAKYGEKLRNFFVDKTNPIRLVDMGAGVFESATVDTNILIVQNSSNKHELQAL
TYTENQTENMSDYIRQHAVAIDYKLDESWIILSEIEQSIKKKIEAVGTPLKDWDISINYGIK
TGLNEAFIISKEKRDELIKADPKSAEIIRPILRGRDIKKNSYEFANIYVITAYKGINKIIKKEY
PAIYLHLKKYEKKLRQRGQVEGKPGKPGSNQHHWTELDNNVSLEKLDDFSKPKILYSEI
VTRPQFYFDSSGRFMPEATAFLITGSNIGTLIKYLNSKTISWIFKNYYAGGGLGNGFRYK
KAFMEQLCVPSEILTEVVDEKIEDEILQIYKFNNDEKRYIKSLI
>hypothetical protein
MGGGIEKEFKVNNLMNLPVPLPGEKIELTNCEENYIRDLM
>putative type IIS restriction /modification enzyme, C-terminal half
MTKYLPFYFDEEQYLINQKAFIVTGESIGTLTAFCNSKLFKFVYLDNFPELQGGTRELSK
VFFEKLFVKKASISDELEYLKLIKEIQKCKLLNKETTFLERKLNEKIFDFYKLTYDEILAINQ
VID
>putative type IIS restriction/modification enzyme
MYKHIIILLYMLLSSLMFYYFLFSQIWAKAIYLQLICTIFSKSHLRFNFMSDYIRQHAVAIDY
KLDESWIILSEIEQSIKKKIEAVGTPLKDWDIKINRGILTGLNEAFIISKEKRDELIKADPKS
AEIIRPILRGRDIKRYSINFANLYLIYVPWHFPLQNNPSIVGASKEAESLFKKGYPAIYNHL
YQYKDKLSKRNRAETGKRYEWYALQRYGAKYMDDFLLPKIVWSDISTKPQFALIDKHIL
INNTAYMITRIHPGAIHILNSRLMKWYFTKISSDLGVGSRYFKQFVELLPLPLFLENKNFY
DTLSEDNTEEKLSEFYHLNQAEIRIINQ
>Site-specific recombinase, resolvase family
MIAVIYGYIRVSTNKQTVENQQFEINNFCKNRNIEIDKWISETVSGTKKINHRKFGKLMK
KIKKDDTVIVSELSRLGRNLMQIMGILNECMEKDVKVIAVKEGYELGNNINSKVLAFAFS
LSAEIERSLISQRTKEALARKKSEGVILGRPKGESRFSSLKLYRKRRIIQEMLQEGYTYS
SIARKLKVHRVTVANFVKRMRDEDNLQP
>hypothetical protein
MTIISNKKNSTLSQKLANQTKNASDKLYIITNYSTLTKEQFDNEIQMGISDFDRNNILTPE
EVQTELLKRRKA
>HNH endonuclease
MQHRLDFTIKEIKRLYTFLPITNLVIEVTPFDNQKLVNPNIKSWEYTHGKMHDFKTIKDYL
LARDNYRDALDGKQYPASQLRVHHLVQRKDGGTNQPDNLVLLSDINHSQANHNNGSL
AKLRENRQKTIDYRGAYFMSILATRLSDYFDNYTTTQDYLTANLRQKYQIEKSHLNDAFI
IVGGTDTTLRTNNVYSRQKLRNNNRMLQKFYDAKYIDSRDGKQKPGKELSSGRTRRS
RELNYDNLRQFRKEKIKKGRISIRRNHYQLRPHDVVLNTRTNKIETVTGVQNSGTVIKF
QTGKTCSIKTVVSLYHVNGMLAKKIGDI
>Paclitaxel/taxanoid biosynthesis susceptibility protein TS1
MMSEIKKVDYIYVVNSNGTPLMPTSRLGMVRRWLKTGQARWFGNSRHTIQFVRPVTT
NTQELTLGVDAGFHLDLAVVGNQCEYYVQESLRKSEKDRITSRRELRRIRRNRLRHRQ
ARFDNRKRKEG
>transposase
MEMQVTVKVKLNLANAEIASSFTNTMEQYRLACNYVSEYIFNHDFDTKQSRLNKVLYT
DLRSLFMLKSQMAQSVIRTVIARYKTVKTQMIRNPYKYQDINTGEWYRETRDLTWLQK
PIAFNRPQVDLQRNRDWSYLSSGQLSINTLDGRVKVNPICHGFDQYFGGTWSFGLAKL
LKTGGKWYLHISATKTVADFDNQTVKHVVGIDRGLRFLATTYDEQGQTAFFDGKIVMR
KRAKYQKLRAKLQAKGTKSAKRRLKKLSGRENRWMTDVNHRLSKTLVQKFGSNTLFV

LEDLTGVSFERTDLPKVLRNQNKSWAFYQLAQFLTYKAHLNNSEVIEVSAQYTSQRCP
KCGTIKKSNRNHDLHEYHCSNCGYRSNDDRLGAMNIQLLGTQYISGQEQPKFELTTNA
>hypothetical protein
MNLQTTLDATIRYDHNLTANEKLLYSELTARMDENALCALSNQEVANLYDVSLCTASH
WFKALIDYGLVKVIKAIPRLLKILPATPHQKAPTPEVSPSAKEAKKVTANSVSSKRDYIRN
PLPADEFTALLLAGFKDNPNLVSQAKPAPTQANISTPLKVNANKEFANSTTTNKADYRN
ESRVVKTNFKTEPNNIKTSFTSSPSTSFQLSKLAVKKLGVTKLLYPLLNKFIVEMINIIINLF
NDLLLKGADLTASFSEA
>hypothetical protein
MIKQHTITQVNFDVRLNLTKVLQSLEPSCLVVLLLLGLLFITIIKAGKTLTPTPMTPRLFEL
SWLTTLVLSLVWVSLLGLLTYLR
>hypothetical protein
MKILDLKFLTPNHTVRHLKFENVNTNLDAATLKDCMEQIIALNIFQTTDADGIASPLYGTP
VSASYITKDEEIVYSAAQGLVDQPAA
>hypothetical protein
MEKQWTKKGLRLTTINAEAIDGHRSFSSIGKDVDDQKVIDFSKLVEKLTGDDVKFIAVTE
VNQLAE
>hypothetical protein
MEEEFITILPAQVRYNRDLKPNEKVLYSELLIKSYPLGYCEINNMELATSFHVTRQSISG
WLSDLVRAGYIEIKVKITDPKLGRIERKIYPLVSAETRQALSKLYAEG
>Replicative DNA helicase (EC 3.6.1.-)
MFSEQIEERIVGIALKNPNLADDPDFKSRYFMGERVRAIAEALCVLPTEKRSLANIWNEL
DNGYNQHNYPISYDDLVALEELIAGQASFYRDLQVVKRNYLENALQQAMVNYSQNRS
QDHLDKVTSIIEALNTINTCDSGELTVASQEFSDQLAGKGPKAIKTLPKLDVLFNGGLSG
AKLIALGARPGLGKTAFGLNLVQAILANNAEVHVDYFSLEMSKLEIFKRLISLQTQIPHSY
LINPTNRKLDRIKLEQAQEQILASNLQVYDSLRTLGSILAVIQAHASQYRQQYVVFIDYLG
LITLDQQYQLDRYLLVSEITRQLKMCTLDYQVPIIMLSQLNRAVESRLNKAPQLADLRES
GSIEQDANLVAFLYRPEKDLHKVNLIIRKNREGCLGELTFKFKGETMTFE
>hypothetical protein
MNDETIAKYRMKLPIQLDIPEDQLNAREFIRLIKKHKLQVTKAVETHLASACRYRNLINGL
NKESYIKNSNGFMAKKLVELDKLRVISISNALSVAFYEKELDCRLVEDADDFLKLTKVKH
QGDLDKVNEYEQCILKLSQNYCESYRVMVTEDYIGGKLPKFE
>hypothetical protein
MDNLYSMGQNILYFTGSLLGTQIIMWLAGKLVKDDEKRNKFKSSGYNIIFSVLFLWFNT
SLGIILLVFNVLVVLFGIRFTLLKDKKISDWGTGIILVLTTVLNIISGKYDKQWEFAVLGIAW
LLFTVVFQNEKIWKFTRMGLFTLLSMVSGVVIYLVGYNYQFETAHILTAEYTTTMDGMS
ADVSGYATANAKVRTYLDGEEVKPAEESDEDGYFSFEANKPGKWVVKVTKNGKTDSD
YTIVKPSKKYKDKKAKKVKEKFANAFAQMADSLEEIGVDEYTRWSDADASGTGGSVD
DLIAQVQKEHGTDSTVAMLGKYNLNDYLKQLKNLESTDYDTYSDYYSDMLSYYRTVM
NPPAGYKDNFGDKFDDLRDKITNDFDEMGR
>RecD-like DNA helicase YrrC
MGEMKKVYNQGNQLLENKQYQLASELFMKVRKRDRKYLSVNHKLARCFLGMKEYQK
SIYYERNHLRVNHRSNQRQANLLLYRAYLKLKMPLKAAFYLETAHELGLADEMYNTKK
QRVEKELAKRADDKETKRELYQRDIDYLTKVKFRYKDLATSGSPIAVSNMPQDQGLKF
FINGMNDAELTKFSQSKGWQHLSVSKPEITKIPFYRSVKMSDERTLYKYYLPKMDDIAQ
IKAFNDYLEDNPFFKKKWADSKVFSLLLIEGFTPEVIEHLLELNSLLTKQEFTDLQTCDDL

VFLKEEGELLPLDRKNEISNRKLLTDFNQLCFENCGNTLHLRSKCVAERQLLRRLKQVD
GVGLPLVKENKTFKPTAEQQDFIDWLQHQRHNVISLLGVGGSGKTYTMGQMLDTKKV
YALAPTHKARLNLGKFFKETATVQRVIADIERGRSMYYRDSDVILIDEVSMLPLEMLVKI
LNYFGGKKVILIGDDKQLPPVSKDTDAFEVVGNVMELLQEGQVYYFRDNIRCKDEETG
QFIKASRNKDTYTLNKLNYRHESLEEMLAYKAKYPSIEDCMILAYRNQTVGKINQSMYE
KLSVNDPHRVEFNYEKNRGLGGFFVGAQVVFYENDRYNNGYTNSEFGIVEEIYTKLDE
DRRERKWIRVRTNQDCYDLPLGIAYRDLRLAYAITIHKAQGSGAKKVYVLESEDYGLAY
TAVSRAKEQLVFVNLEEEEFLESLQKPAPKKKNVQLLENNLLRKRYLVVLNQEQWQNA
SQDIICNYVYNLDKDALIYLYLSSRFEQRIVGIARYVDEFYEGGITRKIKLRLVKKLPQAT
DLSKDRLINEGILLTNGKLLNSSQIEYLQKITDEIVVVS
>hypothetical protein
MKEIRITDKFSFAGMAILIVSEDIPKGIWMGDKVKIAGETFIITGMPISDNRTFVIEDTAKV
RVNQVVEFYKG
>hypothetical protein
MTKVEVIDVMKLDEKFLVTLKGSNELYPGTKLISATNVYQVEARETQSIYAYRNKIVGVL
LAVEPTETVLELG
>conserved hypothetical protein NMA0318
MTKAKIIMIDGEKYIHCPVCNRLVQLFDVCECNWKNTGETNIDGGPNKLTLKEAQAAYA
KGQKIY
>Quinolone resistance protein norA
MPKQKTFFALMICFSLASSYAHPVTPTYFKLLGFNDSMFGLAFAFMALGMFLASPFWG
KVSALINSKLTMALGSAGYGLGQIFFAYCTSAGWILVGRLIAGLACGAFFVGALTYITNK
STNASRGQNLAINATIQTVAGALGYFIGGLVGAWSLPLILNLQIIQLFLTACGLYWLLETD
KTTDQPLALRQLVKEANPFRAFLASRQFMNWAYVIIFLAATLVFSAYTAVDQAFNYYLK
DVYHLNSSYNGFFKGAVGLISLLANTTVGLYLMKKTNVKKSLVVINFLASLSLLLALVLP
GIWAYVGLIIIFYGFYAIVTPLLQDVIAMQATKANRSLVLGFYQAINSLGMIIGAFLAGGLY
DLNTQLPFIFGLGAFLLASLLSLYYQGLLTKSA
>FIG011178: rRNA methylase
MEFTKVIKATGNQQIKDLKKLHTTKGRKKAGRYLLEGWHLVKEALASQADITTILVTADF
KHLAELEGLSGIDLIEISPEVAQSLSDAKSAQGIFAVLSLAKTAPVEPKTARGAWLLLDGI
QDPGNIGTMVRTADAAGFSGVVFGTGSADLYQPKVLRAMQGSQFHIKLAQGELLPWI
EAFQAKHYPVFGSELNEAARPYNQVGKHQDFALIMGNEGNGMQRQHLELTDLNLYIPI
KGQAESLNVAVAAGILMFELLA
>FIG005986: HD family hydrolase
MLKETWRNDQEYMSYVGELLAKPEVQKLANYTQHHFSTRLQHVITVSYLSYRLAKKFN
LNAKATARAGILHDLFYYDWRVTKFDRGTHAWIHPRIAVRNAEKLTTLSPLEKDIIIHHM
FGATICPPKSLEGLIVSMVDKYSATQEYSEHLIQEFEGKMLKKFSFAKVSK
>Transcriptional regulator, HxlR family
MSTEVVTAPNFKLCPKFEKTFNILGKKWNGLIIDVLLSEGPQRFKDLASKVDKCSDRVL
VERLKELEADGLLERVTQPNSCRVLYQLTDQGQALRPVMESIHTWAETWYSLADCQ
>Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)
MDLWQKLEELKKTSLSDIQKADNEKKVNDIRVSLLGKKGALTEILRGMKDLSPEDRPKV
GSFANTIRDELTKEIAKQAEQIEVAMMAKKLEKEKIDVTLPGVPVKEGQPHILMQVIDQL
EKLFLGMGYKIVDGPEVEEDHYNFEMLNLPKDHPARDMQDTFYITKQILMRTQTSPVQ
ARTMEKHDFSKGPLKMISPGKVYRRDTDDATHSHQFHQVEGLVIDKNITMADLKGTLE

LVAKTLFGADREIRMRPSYFPFTEPSAEVDVSCFKCGGKGCAVCKQTGWIEVLGAGM
VHPNVLEMAGVDSSVYGGFAFGLGPDRFAMLKYGVDDIREFYLNDVRFLNQFSQKG
>Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)
MKVSTKWLNEYVPVSDLDRSALAEKIARTAVEIAADYRLDEGLKKVVVGHVLSMEAHP
DSDHLNICQVDVGLDEPYQIVCGAPNVFAGAKVIVALPNSWIGNHTKIKKSKMRGVVS
MGMICSLQELGFSDSVVPKKYADGIYILPADAPVGQDIFSYLGMDEDIVDVDITPNRADL
LSMHGVAHEVAAIYDRKVSFPEVTVAEDPNDSVANYLTVTAPSELAQTYMMRVIKDVK
VKESPIWLQSRLWNAGIRPLNNVVDVTNYILLDYGQPLHAFDYAQLNSKQVVVRLAKE
GEKLVTLDGQERTLKATDMVITNGDVPVALAGVMGGLDSEVSDQTTTVALEAAVFDSS
KIRKTARRENLHSEASTRFERGVNLADVQTALDKAASLIAELGAGQVVSGIAAVNDLDL
ADKQIEISAPRINKVLGTDLSLETISNIFARLGFGVEANGADLTVAVPVRRWDIEIYADLIE
EVARLYGYDNLPTTLPTGETTPGHYTYEQKMIRQARALLEGAGLSQAISYGLTTEAKAK
RFALEPAKLTKLDFPMSSEHTTVRMNLISGLLDDLAYNAARKNEDVALYEQGRVFFRD
EGSERPREVEHIAGAITGLFHQASWHDGKKAADFYLTKGIVDYLLTSFGLENVSYQATA
AHDEMHPGRTADIFVGDEFVGFIGEVHPTILKEYKLKRTYVFELDLAKLISLDKKELVYQ
PVSKYPTITRDIAMAVAKEVSNAAIVETIEANGGPYLKEVKLFDLYEGEHIDFGFKSLAYT
LVFNNPEKTLVEEEVNQAFDKVVAELKAKFDIVVR
>FIG004453: protein YceG like
MGYNRKDEFFRKGAKKSVADDKNKQAPIDGAPKKRAKGIVIGVVVVICALVIAICGFGY
HYVQDSLKPLNPKSHKLIQVKVPVGSTNKQIGSLLEEKHVIKNGMVFDYYVKTKKFNGF
KAGYYVLQPSMTLSQIAKKLEKGGSSTPLINTGKVLVKEGVTADEIGTVVEENTRFKKK
DFLALLNDQTFVNQLAKQYPELLSSAVNAKDTRYKLEGYLYPATYVVTKSVSLKDLVTQ
MVARANQALSPYYSQIAAKKYSVQEVMTLASLVEREGVTEADRKEIAGVFLNRLDKDM
ALQTDISVLYALNVKKANVTYKDLKVDSPYNLYLNKGYGPGPFNNPSVDSIEAVLNPSD
RDKNYLYFVANMKTGKVYYAATYAEHQANIKKIETENAN
>Uridine kinase (EC 2.7.1.48) [C1]
MSQQASNKRPIIIGVTGGSGSGKTSVSKAIFEQLHGHSLLMLQEDSYYKAQDEKPFEE
RTKVNYDHPDAFDTDLLIKQIHELLNWKSIEVPVYDYAAHTRSKEVIKKDPKEVIIVEGIL
VLNDPRLRDLMDIKIFVDTDDDIRIIRRIQRDIEERGRSLQSVIDQYLATVKPMYHQFIEPT
KRYADIIVPEGGENQVAIDLLVTKVRDIISKTN
>Transcription elongation factor GreA
MSEKTFPMTLEGKEKLEKELEYLKTEKRAEIIKRIQIARSFGDLSENSEYEAAKDEQAFV
EGRVKEVENMLNHAEIIDNGAIDADEVAVGKTVTFQEDDDEPETYQIVGAAEADAFAG
KISNESPIAAGLIGHKLGETVAIETPGGVMEVKIIDVK
>Type I restriction-modification system, specificity subunit S (EC 3.1.21.3)
MDKEQIGIPKLRFPGFTGAWEQRRLGEVADSVEYGINAPAKEYDGINKYLRITDIDDTT
HEFLQSDLTSPDTDLTLAEKFKLLEGDILFARTGASVGKSYIYSREDGSVYYAGFLIRMR
FNKSISPNFIFQNTLTTAYKQYVEVTSQRSGQPGVNAKEYSNYSLYIPSKEEQQKIGSF
FQQLDTLIALHQRKLEHLKQQKNGLLQKMFPKNGESVPEVRFPGFTAPWEQRRLGEV
ADSVEYGINAPAKEYDGINKYLRITDIDDTTHEFLQSDLTSPDTDLTLAEKFKLLEGDILF
ARTGASVGKSYIYSREDGSVYYAGFLIRMRFNKSISPNFIFQNTLTTAYKRYVEVTSQR
SGQPGVNAKEYSNYSLYIPSKEEQQKIGSFFQQLDTLIALQQRRIEHLELLKKGLLQQM
FV
>Oxygen-insensitive NADPH nitroreductase (EC 1.-.-.-)
MELNETLKKQLDHRTIRAFKEQILSQEQLVTLLEVARHTSTSMFMQGFSVMHITDPKKR
AAIRNISTQPYVGANGDLFIFMVDLYRNYQIRKQTGHDAGRVHTTDIFLQAYEDTVLAV

QNVLNAAESMGLGGVILGSINNDPAKLVEVLDLPKLTYPVLGLQVGVPDQAPQLKPRL
PLEFIAFGNDYPKDFKASDLKDYDEVVQTYYDLRDANRRIDSFTNQVIGPKLSVASKKR
DEQVKVLKAQGLCLDFN
>hypothetical protein
MVSKQKIRQKLARADFTDDVMFGLVMSQLEICRKFLKAVLPQHDFTKLKIVDQSHLKTD
YLSKAGRLDIHCTDAYGNQFDIELQMTNEHNVAQRSKYYHALMANHMLTAGEEYQAL
KETYVIFLCPASFTDADQAITHYEMVSRQGNHLNDGTHTLILNASGNWAGVSDELKGIF
QLLLRQPASFGKLGPQIMDKIDEIKHSKTGGHKYMELTGAFWQDAREAGIKIGERRGK
KAGIKEGRKEGRKEGRKQGRREGTDEAIIKLIPDLVALTNDRQAVINILMKKMDLSQAQ
AESYYKRAMEREKRD
>Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-) / Penicillin-binding protein 1A/1B (PBP1)
MEKKPQRQAKRRPQPKKRPVKKRRPKKRKRKITWKGRIRRLIASLVFLAIVFICFVLGLV
IIYARQAPALDNNYLISSGTTTIYDRQGQKIFSLGLQKRQPVKANEIPSQLADAVTSIEDR
HFYKHFGLDPVRILGAAVADLKHPDRALQGGSTLDQQLIKLSYFSTKKSDQTFKRKAQ
EAWLALQLDHQYSKADILKFYVNKVFMGNGIYGMQTAAKYYYGKDLADLSLAQTATIA
GIPNAPSTYNPYSDLQKATKRRNTVLKAMLANKKISQPAYLAAINEDISSGLVPKASHA
QSQTNQKAILADPYLKQVLSSAKQAGYDPYRSDLKIYTNLDMQVQTKLYQLANSNELA
YPNDQLQVGASVVDPKTGAILGMIGARKVDNVTFGFNRAVQTSRTNASTAKPLFDYAP
AIEYLEWPTYHELDDGPYTYAGTDKELHDFDNRHLGKMTMQDALIQSRNIPAIRALEAV
GITRAQNFISQLGFTYPNNLELQNGIGLPSSSLQNAAAYAAFANGGTYYEPSYITKIVLP
AGETRKFTSKKKQVMQASTAYMITAMLKDVIYSGAGTGRAASIPNFYQAGKTGTNAYP
ADVTTNFPSDALMDAWFNGYTKNFSVSVWMGYDNPYQPNHYLTPAQSQLPIEFYRHL
IEAISQGLPNQDWARPSDVATKVINGKTYYYLTADDDDTSDSFSVPASSSSSLGGFAST
TTSSATSSLSSASASSATSSLNSSSSQISSQTSQTSVTSSSQASSSQASLTPSSSSSEE
TLTP
>Flagellar motor rotation protein MotA
MDFMVVVGILMGVIAVLGGMIVKGANLMVFVNPEALLIIFVGMFAAVINSYPLSDIKRIPG
MFKVLFTNRDYNPVATIESIVEMSNIARREGLLALEEPVNQIENEFLKRGLEMVVDGIDA
EQIEEIMYTEIEATEERHHAGAGIFKTAGTTSPTLGVLGAVIGLIGALGDLSDVDKLGESI
SSAFVATLLGIFFGYVIMIPFSSRLARKSAEEIQVMNIALEGVLYIQAGQSPKTIKQKLYS
KLEPKLRPEDEE
>Flagellar motor rotation protein MotB
MKRRKKKREEEEVGEGWLLPYSDMMTLLLSLFIVLFSIAKVDSHKLQQVSSVFSQILSS
TGSTTIATSQETAIDNVIDLGDTPQKGKGGTPNVRGEANVKSQANVNTISQAEKAEARL
QAEMQQKQTFNQVTQEVKDELAQAGIQGNSTVKQEKDGIHINLDSNILFDSGSAELSD
QVQAALKKLVPTLQKVHTYPVIIAGHTDNVPMAHTWRYASNWELSSARAISVMHFFVN
NGAFNEDGVSVEAFADTKPIASNDTAAGRAKNRRVEIIIQKAKDQS
>hypothetical protein
MNTKGTKSVGRIIIAILMAVSLLSIVTLTVGAYKSTKDLLIERNQLSEKSAVTLVNGEATNL
RLSTFKRVKDIRKNGFTILSTDGFNISALFV
>transposase
MNNSIKTILGIKDPYLKLDEKNFDNPIEDQPNQIIVHLIQTYPMHCPKCGHLMCKNGYKT
VNCLGPELHFKPTIWSIKKQKYICKASSSCPEVVTKLAAVEDIHYRNHISLAIKQRAMML
LTKNESQSDLAKELNVSDWTIRRVITNLDQFFKPNYHWLPRHIAFDDFKSGRFAPSGM
SMILMNIENKRTLDIILSRKNSYLRKYFLRYDRSARLAVQTVTVDLYTPYRHLIHELFPHA

LIIADHFHIVAQAYRAFNKIRIQVMNRAGAGTHKWRALKHFWKLLLTPANELKYDNYWS
RRNFSYAQLTDVEVIHRLLSFDNELKRAYEYYQNLILVIAHRSKKELKNLLAIKWTQLPQ
ALQKVQRTLRRHKQEIYNSFKYDTYTNGPVEGTNNKIKVIKRTAYGFRNFFNFRIRILLA
LPNTYIAINWRNKQTAHAKVQAQAA
>FIG00630584: hypothetical protein
MDSNQTGNDGIVGVHFSNDQGKAIYGLREMPKNMDTRTRPWYKGAVAQEGTVYLSK
PYKDIASGQYVSTVALAFRNKANQLGVIGVDVSFSAINHTIKHLSIGRTGSVTMITRQGD
VMSSEGKGKKYKDGTNIANSEVYQAIKNSPKLSGFVQLKDGNKVSAVYFNKIRANSTL
WIFATVSKNDLAVETRALIRNSAILVVVMLIVVFAISFFLTKQIKGILDKFGSFFDKASKGQ
LAKIPSIKAGHKASLVDKFTNPDANGYEFNRLAVSYNQVVDGISTMVHSLQSESQRVD
EMAASLLELSQQTSKATEEVSQTITGIAEVTSTQAAETAHGVDQLHKLAETINDMRQSI
VQLHDQTNDSSALNQENMDTMDKVHANWNNELAEMKQLMESVNDMSANVQNINQII
NVINDISRQTNLLALNASIEAASAGEAGKGFAVVAAEIRKLAEQSKESTAEIEKIIETIRQK
SDEMVAQTNNSIEGGQKQTDLILSAIKSTTAVYQGNNQMTAEVNTLDTASASIEKAQEK
VLSNLENISASTQENAAGTEEVSANSEEVLATIEEFTGHVADLANIAKELRKETNKFKILD
>Aggregation promoting factor
MQVSASQLATITGLFPNSALKKDLSSQQVAFSQQLNQAQVTVKAGDTLSQIAADHNLT
TEQLVKVNHITNPDLIYPGQVINLSSSQASTPVQANQSQASTVTTVANATTSPANSTVP
SLASQAEEEARAWIAAHESSGSYTAKNGKYYGKYQLDLAYLKGDLSPANQEKTANNY
VHNRYGSWVKAKEHWLSHNWY
>similar to Golgin subfamily A member 2 (Cis-Golgi matrix protein GM130)
MQTKAYLGGAGLLLLAPPTTVLASPLKQTISQNLSAHQQRLSFLTSFQTRLSQTQGLLD
PGVAYPSLQAAYQEWLRTSQAITNLTRQIGVLKTQLAAKQTDLDQTLVNQAQLNDHLA
AVKLQLQDLSANKAELTQELADLKEQLANASAQPDELTAGIAELTTKLAQLTDQLAELT
KTQADLEANLTELAAKLARLKEELATMTAQLNLLTSQLNQLEAKQAQLVARFHGTLASL
ISQEQTTVTNLTNQLKTSQQHLTSQPLTLRTAEQKTSVSFPPSPKQISFSQPSYQPNYP
VQANPTTPLPVAKQTKAVQTKKSSASQAKGDGKLASYTAPQAKFAKKEKHHSATNWL
YSLLAAAFALFSAGFFYQRTRRD
>hypothetical protein
MKSRIIRINPKQKLKLISNYINGNDNLIIVNKKIGFALRFLL
>Flagellar basal-body rod protein FlgB
MVASESTYSLLRQAMAVSNQRASLISSNIANVNTANYKAKQIVFESQLADATQGLALNR
THKAHMGYLGASQGAYVTASNTGAVKENGNNVDLDSEMVSQAMNTIYYEALEKQISG
RLSMRNYVINN
>Flagellar basal-body rod protein FlgC
MSIFNGLQINASGLALERLKLDTVSTNIANLNTNQTRNSAAYHAKSVVFSENMKTSLNN
GLTSSGVKVLGIDENLDRNLKYDPTNEAADANGYVEQSNVNLSDEMVNMIQAMRTYE
ANVSAQEANKSILTKALAISKN
>hypothetical protein
MNLGAIDPSLAAYQANLGQSLGLTKEQATSAKQTFSDYLDKSIASLNQSMAVVNDGTK
DLVSGKAEDLGQIMVNMTEAQLNLQTAVQVRNKVITAYNDLKEMQF
>Flagellar M-ring protein FliF
MDKVKDSLANLGQWWRDLSRLARIAIILGLVSLLTVGGIIYYLQTKVDYGVLFSNLSQAD
AGKITEQLEKENVAYKLADNGSTILVDKTKIDQERINLAMNNDLPNTSKGFDLFDSTSV
MTTDQDRKILYQRALQGELQNAIQSLDAVSAAKVILVMPNNSSVFSDSSDNNKAKASIT
LTLKNGTISNQAVRGIVSLTTGAVNGLSPSNVKVVDSNGNVLNSDTASGDDNTGGSSK

YIAMQNKYASQLEKKIRSLLQPTLGDTSFTVSVNASLDFNAVENKTTTYANPQVRSEAT
AAGGDAATVSSAQAANEASNVANVTNQQNGNNASSSASYSHTQNNELDTSVTKTIQA
PGTVKRLTTSVLVNRRLSRKNRADITRAVQAAIGFDAGRGDAIVVQGMTLNNKRKTTA
TPKATTDNKSFPWLYLGIGGVILALVIGGITFFMIRRRRAAEEEYYDDDYDVDVSDEETP
TATEQADEVETPAPSPVVEEPKVSPEVQKQLDRDEQARNYAKENPEVAADLIKAWMK
DETNRRGGGK
>Flagellar motor switch protein FliG
MAEMDGFKKAAILMISLGTETSSQVMKHLPADVIQKVSYEIANTDMVSPEEKAQVLNDF
VNTAQARQYVLDGGIDYAKNVLKQALGPQKAKEVIDLLTQIQYRERPFNIARKADEQQL
VNLLLDEHPQTIALVLCYIQPDKAARVLSEFPKELQVDVAERIGTISSTSPKIIHRIEKVME
SKFSTLIENDSEKVGGVKALVDILNSASRSTEKSILGDLEKTQPTLAGEVKASLFTFEDIV
TLDKMDVQKVLRDVANDVLALALKGSNDDIKNFIYENLSNRAVDNLKEELEFMGPTRLS
AVEEAQQEIVGVIRKLDEAGEIYIGRGEEDAVV
>Flagellar assembly protein FliH
MYNNVVKQGALLNGKKAPIVTVQPKKHHQQADLLADGNLTPEELELKRQSKRHEAKL
NQLRSEIISQAQAEAQQIKDQAHQAGFESGQKAGYEAGYQEGLTEGNKVATEIVNQAQ
ANLEAANKEVKDFVQTEKQAIINYGIKMAEVLLKSELKADSSKILALLEPILFELEKPDQVI
IVKAHSRYHQDLNQRLEIMKQNSNVRYAILDDNSLGVDKVVVESDEALLTLDIHEELEN
YFKALIEKEQNEN
>Flagellum-specific ATP synthase FliI
MKIDFDELTNEARLHRFYSKYGKVAEVVGLIIKATGLEVFVGEICEVQVKNSHKVALTEV
VGFVKDTVLLMPLDDLEGIGPGCIVRATKKTLQVKVSDGLLGHTLDGLGRPFDLSDVDT
SSLVDYPVMQDSPDPFRRQPIANIMSTGIRAIDGLLTVGEGQRIGIFAGSGVGKSTLLG
MIAKYCQADVIVVGLIGERGREVLEFIQKDLGEGYQKSVVVCATSDMPPLVRLKGAHTA
TAIAEYFRDQGKKVVLMMDSVTRFAMAQREIGLSIGEPPTTKGYTPSVFAELPKLLERS
GMSDKGSITAFYTVLVDGDDMNEPIADAVRGILDGHIILSRKIAAQNHYPAIDVQTSLSR
LMKGLVSEQQSKDASALRENLAAYADAKDLIDIGAYKNGSNAVTDFAIKVHHPIDSFLK
QEVSEYSSYEETVTKLHQLFDQTSVL
>Flagellar protein FliJ
MEKFQFSLEKVLTYRQQLEQEAKLEFAAVQRQVNEREAKLDQIAKERAQMLTVPEQTL
GRMQVQHRYLAQLEQVSSELKLEVDQLQGKLDQALDGYVKAQQQRKVIEKLRAKKEA
EYQQAFRQAEQKQLDDLANRKQLIS
>hypothetical protein
MLKNIPLNVANPTLGTTGSDDQAAPLVTGEQVLDFSKLLAQVTTETKPVSSPLEETTNQ
QQPLKQQPLDQAQVKDALATKEVEVVPLASLTNAKTSAQLELEELGQKQAAVEKQLSA
VTTNDQVKPVKVAVTIASNQADLQPADLGLKPTTQLRKQPVAVQLEAEPVQPDLKLKP
NLTAVKVKQSPTKISGQADEQVDLGQLALLTSAAFKENLAVANGQTNLKTNLPSQVQEI
SLTTGQTKELDSLANKIVAGLKSENLNDHKTLTLQIKPEKLGKMEVKFKATAENFTLEFK
VETKETKAILEAVSDKLQKIIDKQLSPTFSQALNLQAPSEVLKVSPSQTASEMTLGQFDF
DQGQSQQRQFGQTHQAKKASPKQAVTLVEDSPATPEVTGKVSILV
>Flagellar basal-body rod modification protein FlgD
MDVANLSDAVASLATNPNLKSASQQALAKSNSKLDMEAFLKIMAASMSNPSFSGDSSS
GGASGATDYVSQLAQFSVLQELTDLETTIAQSAAIQQQQQAFALMGKKVMLNDGQGS
TIEGTVSGVQFSKGYAKLMVNGQAYSMGNLIQVGDAK
>Flagellar hook protein FlgE

MLRSLFSGVSGMKNLQTDLDIVANNIANVNTTGFKSSRVHFADTFSQTVASASAPTRG
NGGTNAKQVGLGDQVASIDTNMTSGSTQSTGIATDFYITGSGFFMAKDASSGQTYFTR
DGAFQRDSAGNLVNSNGYLIQGVNTVANPILYANYNENSLTTPTTVNATQGITIPPTLTY
QGQNYSFNSFSVDSNGVVTGEYKNTTNTDADAQKFVLGKFSLATFNNAAGLEKQGNN
LYSASNNSGTPIVGDVGQNGAGTVQSGALEMSNVDLSTEFTNMIIANRAYQANARIITT
SDDILQELVNLKK
>Flagellar protein FlbD
MIRLTSFKGKTFYLNPDLLYRIDETPDTVLTLTDGKNLVVKETASEVVALIIAYRQKIYRQ
PAENLVEKVGD
>Flagellar biosynthesis protein FliL
MAKNAPKTEEQPEATPKKKGNKLVLILAFILLAAAFGFGGSYLANKYILPKKAEAAKNDK
TTISKDEVLVPISKFVVNLSGNSNNNYQYIRLKVSCLVADKDASENLTKSMPLVRDSVIS
VLNSKTANDLLDANQGINNLKTAIKDKINQEYGQTIVKQVYVTDLVIQ
>hypothetical protein
MEMALAVVRLIIGLAVVIFLINVLMRYLNKYTTQNNDSLRLVLRVALSKNLSLGIVEILGAY
YVVSMSEQKLEIIRQLDAEEIEQLEANLAQKQAANEQQVEKFKEIFTKAKLTVRKDHKD
DED
>Flagellar biosynthesis protein FliP
MMKIKARPSKVYLLALGLSLGLLLMGLFPEAIHAVTTNELSSELNKALGGTSDTKTTINT
FVMLTLLSFLPIVLVMTTSFTRIIMVLSFTRSALGTQTNPPNQVLMGLALFLTLFTMKTTY
TNIYQDAYKPYTENKISQTTALSRAEGEMKKFMYKQTRSKDIKLFLDLAKDKHRYRKAQ
DLPLVSLVPAFVISEMKTAFSIGFLIYIPFLIIDMVVSSILMSMGMMMLSPVMISLPFKLLLF
VLVDGWYLLVQSLVAGFH
>Flagellar biosynthesis protein FliQ
MTISLVLTLIREASLRILVVSGPALIVSMIVGLIISIFQATTQIQEQTLTFVPKLFAIFITIILAG
NFMITVLLEFTRYVFRMISQL
>Flagellar biosynthesis protein FliR
MDFVQILGLILCRISAFIVLAPLFSERGFPRMAKLIVAVCLTVAAFPEVTTSYRAVSGDLY
LLLALKESLFGMVMGYISKLVFDAVTMSGVFIDFQTGLTMAQAYDPTFQVQNSTYGKIY
NWMAIMVFFALNLHLLMIKGVMYSFEMIPLGTANILTGSVIEGIVKLFVKAFAMAISLAAP
LVVAVLVVDVVLGVIARTIPQINVLLLGMSIKTIVALILFLLALPNVVDFLQANLPAAYTNIV
DLVKALKGR
>Flagellar biosynthesis protein FlhB
MNNDKSDKTEKPTPKRIREARKEGQVAKSQDLGAAAVLLGLAGFFIPAWQFVIQNYLP
YMVASLQNLTIYQAQMNDLPKVAVQAIWLVVLGSLPFMAISLGIGLFINLVQTRLNFSVK
AIKPDFKKLNPLTGVKNMFSMQALMNLAKTIAKLAVIVYLCVKKYQAAISTIAQLSQMEM
GQILYFVLHFAQELMLQIALVFLVLGGLDYIYQRHSLMKQLKMSKQEIKEEVKQNEGDA
NIKAQRKAKYQQLLQNAIANVKDATVVITNPTHFAVAIQYEKDGLGIPQVIAKGADEIAQ
RMKAEAKKLDVPMVENVPLARGLYAKVKVGDVIPDEFFGPVAEVLALVYQLEEEAKHKI
>Flagellar biosynthesis protein FlhA
MAGQSKKKAKGNKLLAADIIISFVVVGVIGLIIIPLPAAVLDVLIVTNMAIAINILLITLFAKSV
LEFATFPTLLLITTMFKLGLNMASTRLILTDGSAGKVIAAFADVVAGSNYIVGIILFIIIMVV
QLVVVTNGSGRVAEVSARFTLDAMPGKQMAIDADLNSGLITEEQAKNRRKDLQREAD
FYGSMDGASKFVKGDAIAGILIVLVNLIGGSIIFSANGSMSIVDALGQFGKLSIGDGLVSSI
PSLLISIASGIIVTRSGNDGSFGKDIAADAIRNPSLFRIASAILIILGLVPGFPTIPFMVLGVG
LLVASTMITSNEEKKQAAFRREQQKLALEKRKSSQEEDEDVASFQVEPIAVEIGYGLVS

LVDKSQENSLMKQVVAIRKQTAHELGILVAPVRIRDNLYLDANNYSIKIRGNEVAFGDVY
PDKYMIISNGQDEFPFQGIPTKEPAFKLDALWIDEADKEQADLQGYTIIEPLTVIATHLKE
VIYGHASELLGRQEVKTLLEGIKAANNVVIDELIPDILRLGEVEKVLQNLLDEKIPINDLTTI
LETLADYGTVTKDTEVLTEYVRQALKRTIANKYSEDGRTLNVVTVHPKTEEMIAQSIQKT
ATGSYPVLNPDAVNQVLQAVDKVQAQLATKEQSYVILTSPKIRLAFRKLIAFNFPDLAVV
SLNEVPNEITIEAVGMIEV
>RNA polymerase sigma factor for flagellar operon
MYEVDLESQVLKYLPLVERIASRISIKNTNYEYEDLYNIGVIGLIDSLKKYDPSKKVPFGG
YASIRIKGAIIDEVRRHSKLPRYKMNLINEYYRAKEELESQAQHTASDEEIAAKMGIDLN
QLGEVYDSMHYLASVSLEGTMFSKQEDQIHLADLLEDQTATDAEENLLVAEQKQILAK
AIKRLTEREKLILSLYYQEEATLKEIAAVLDISIARVSQLHGKAIAKLRAYIQKELEND
>Flagellar hook protein FlgE
MIRGLDTLRADFNVLQARQQNLSSNVANVNTSGYQQKHLFQAALADVRLHNYENGPQ
NDQYNGLFGGISWHNQISGAKIDTTMGALKHTERPTDFAINGDGYFVVQGPNNQQLYT
KNGNFRLNQQNQLVTQEGYPVLGANGQAVTGIANPNLMVVTFNDNNNIQNLGATYYT
TTGGVSQVQNLQVQAGYLVDSNVSMADAMTEMMQTVREFQSSQQVLSATNQTLDKA
VNNLGKV
>Flagellar basal-body rod protein FlgG
MGINNLFATNVTGVTGLQRQLDIIANNVANSTTTGYKARQASFQELLNQDIKAGVDPAS
QPLGLNAGLNVTNQAPDFKQGNLEETGRQTDLALNGPGFFGLRDANGNLYLSRDGAF
SFDSDQNLVSAQGLKVELNESLPKAQWPKGTLQIGRDGRVYIKEQNQTTAVGQLVLY
QAQNNEDLLSVGNNLYQVAPGRTLQSNLTGANVGAVIQGSLESANVDLAAEMSNLIIT
QRSYSMNTQMLKANDELWEAINAFSR
>Cell division transporter, ATP-binding protein FtsE (TC 3.A.5.1.1)
MIRFKQVSKTYGENLALRQVNFEIKQGEFVYLIGPSGAGKSTIIKLLTREELATSGQVIIG
NTDLSKLKDQHVYKLRRQLGIVNQQDVFLENRTVKENYQAVLGATGVAETEWLAKIKR
TLDLVGLGTYLDAYPRQLSVGQKKRVAIGRAMLNRPKILLADEPTANLDPKTAIDIMRLF
FRINRQQTTVLMATHDSTIVNTFMRRVLELEEGKLIRDQAAGNYNVFGNPNDIYIL
>methyl-accepting chemotaxis protein
MQNVFSRFKKINGKSFIPVGITLLILVCCYFFVISGNADLAASFWIFSLVVLILQALGICLA
KMYFTGATDQSIANLEDKIAKVSQGDLRDLESMVTDDKESIAGRVQTSLLGIAKAFKGM
AIGMKEETKEIEQMVTKLGNLTDNAKASISDIRDSVADIADKSNRQNSETDQTVEYINQL
ADDIDSINADISQMVARADQSRESNERNFNLMHDVATSWQEDHQNQTQIVKEMDNMN
KDIQSIGNIVNLINDISEQTNLLALNASIEAARAGEAGRGFAVVADEIRSLAEQSSKSTKDI
REIIESIRDKSEHMTNDINRSYKKGQEQSQDIEKAIGATREITDLVNSFNTSLDSIKGRMT
DVIEQKDVVTEASQQVMDSMTSISASTQEVNTNVDDFYKMVQELEKDVAHLKDASEIK
KLNAAAFITD
>Positive regulator of CheA protein activity (CheW)
MEQYIVFTSHDQLFCLPVTVVRRIVEAENFVPIPDVPAYILGVYEYEEQMLPIIDLKAKLF
NDPTVKNTENKVILCQWQEQTIGIYVESIRGIMMLEDTDYERELTKAKLKRGYIGKFLKL
DDEVVISLELEYLFNNEQAQELMDSVDQAKEDQILNQPAKDTEGDDQ
>Chemotaxis protein CheD
MINSSLRVGIADYKVAKAPNSLITVGLGSCVGTFIYDERSGVGGLSHIMLPDSSLYTGSL
ALNPAKYADTALELMVKELKRKAPYGRFKAKIAGGANMFDFNKTLKNGNVGKRNVEAV
EAELSRLHIPLIAENVGGNSGRTMRVNLETFEATISIVHHDSIIL
>Chemotaxis response regulator protein-glutamate methylesterase CheB (EC 3.1.1.61)

MIRILVVDDSAFMRKVITDLIAKMPGTELAGVARNGKAAVDFVKADSELDLVLMDVEMP
LLNGLEALKQIKQFSSVPVVMLSALTNQEVTIEALESGATDFVEKPTNIMAIKDDWIREF
FAKIKAASRQQRRPLSRRREVAPTATTSGQVKSSLELRRNLKAVVIGSSTGGPRALLHL
ITALPARLAVPVIIVQHMPKGFTKTFAERMDTEAKPKVSEASDGMLLKNEVYLAPGDYH
MIIKDGRLHLNQEPKMHGTRPAVDPLFDSAAAYYGQNLVGIILTGMGKDGAEGMGTILA
KGGYTIAQDKDSCVVFGMPRSAIEKGVVNEVLSLDQIRVKLEELVR
>Chemotaxis protein methyltransferase CheR (EC 2.1.1.80)
MSELNFEFFNDWVEKKLGIQLNSYKQKQMQRRIGNIMHTTGAKTLEEYAQILTRDKAA
RDAFVEHLTINVTEFYRNKEIFESFEETVKKIILPKYKMPKIWSAACSIGAEPYTLAMILD
QNQIKGKIVATDIDKVILDKAKAGTYKLAELKNLAPKVLERYFTKQDNLYEIKPQIKKYVT
FKRHDLLKDSYEPNCHIIVCRNVTIYFNSDARDEVYQKFSDVLMPGGILFTGATETISNC
EQFGLKKVDSFIYQKI
>Signal transduction histidine kinase CheA (EC 2.7.3.-)
MADDNSVYRDLFFEESDDNLQALNDNILELEQEPDNIDLVNNIFRVAHTLKGMSATMG
YDVMTKLTHKMEDIFDLFKSGKLKVTSDHVSLIFKCLDKLQELVDDLRDDKELSEEQITD
LWNELDQVEKGVEGKTPAAQAPKEEQASDGPKKLTTDFEKVEEADLDVIKKAKAEGY
NAFAIAVRIDAGSLLKGPRVFLIMEKLDQEGDVIHSEPSTDILEAGDFETDFKLIYLTHDS
KEQVEENINSNSEIDQVIVADFDPEAKPAAAEAEEKEVAKKEAPSPAKEAKPKAAPKKV
AKKDAHHNNRNQSIRVDLTRLDNFLDLVSELVVYRNQLEDAGKRNNADEVRDSMEQV
SRLTSELQDLVLKIRMQQVSVVFSRFPRMIRDLANELKKDMELVIIGEETELDKTVVSEL
SEPMIHLLRNCADHGIEQPEVREKLGKPRKGKITLAAYQEGNKVIITLEDDGKGLDPEAI
KASAERKGINTEGMSEREIQMLVFHPGFSTAKKITDISGRGVGLDAVQNKISELGGSLE
MESQVNVGTKVTIKLPLTLSIIQALMISISNEDFAVPLDVVERVVMIKEEDILQTQQQEVY
RFQDQLIPIVRTNRLLGMPEIESKKKFAIIVKINKQYYGILADGLIGQQEIVIKKIDQMLMNI
NKYQGATILGNGSIALILDVNAICNAQKVANS
>Chemotaxis protein CheC -- inhibitor of MCP methylation
MAYSEIELDGLKEIVNVGGGNAATSISQMINSRVDMDVPEVEVMAYDELYQKIIADDVE
MHAVLSKIVGDIDGALLFVIADESGQKIAKMMLGSDDNPSNEIIASAVTELTNILFNSFLR
AIGDMLQIQLIASLPISRYDFFGAIISSAYMAFDHYDEQILVIHNEFTYNNESLDASLFLIP
GEGVLDKIFKALGI
>Chemotaxis regulator - transmits chemoreceptor signals to flagelllar motor components CheY
MAKSVLIVDDAVFMRMKLKDILEKNGYEVAGEAQNGQEAFEKYQATKPDLVTMDITMP
EVDGLEALKMIKASDANAKVVMCSAMGQQGMVMDAIRAGAVDFIVKPFDTERVIQAID
KALS
>Positive regulator of CheA protein activity (CheW)
MQMILFKMQGQNYLISADDVEEVVDTIQITKVPLASNWIKGLINLRGTVVTVVGLAELLGI
PSSDNNLNILIMKNGEERKGLLIEEVIEVLDIDPTEIQLGSDQDEKHIKGVIAFPDKVANIV
DVTDAIF
>Flagellar motor switch protein FliM
MDQVLSQQEIDALLNAMSSGTLDEQKIEEEAQEAAKVKTYDFRRPTRLTKEYINTLHMV
FDDFAKLSGTLLASQLRTSVNLKVAAIEQVSYDEFIHSIARRTLIGLFHSDPMKGSQMIEI
NPQLCNFLIELLCGGQEIGSDEANAKRDIEIRENFTDIELSILEEVVTDLMGVYQQVWKD
IVEFNPVLDTIDTNPQLLQTMSPNEPVVLVTFTVDLLDRSTFINLCIPYIFFEGILDKLSIRN
WFDSSKDNDEDDSAELKRNLNGVEMTIEAILGGAKMDLADFLNLEIGDVIKLDKKTSEP
VETYIGEQPFYTTKLGKLSNGQFAVELIDKMEGEED

>Flagellar motor switch protein FliN
MSDTLSQAEIDALMNGGATAEPEKPAETKEEPASSEADQVQASEASAEASSLDEQAK
SDLIGEVGNISMSQAATTLSSILNHRVSITTPRVTRVRFKDLIDGIKAPKVATTVEFKEGL
TGVNLMLLEVHDANVIANLMMGGDGTPENEEFTELELSAVAEAMNQMIGSAATSMAT
MISTKVDILPPKVNLWDEPGNTKYDFIVDDEVIYKISFSLNVDGLIQSEIMQIFTDQMVDD
IAQAMLADKATVVDNRETATSQSEESKSQAPAASQAPSQAPSQASSPAPAPTPKPTPK
PKHQEPVEVQSPEFQNFEPTQPAAESNDNLDLIMDIPLNLSVVLGRSKKTVRDILSFNT
GSVIELDKLTDEPLEILLNGKLIATGEVVVINENFGVRITDILSPQQRLDYMR
>hypothetical protein
MNIRGDYNNYRNLYVNKAKPAYSTAASQEPVMGSKAANVELSETAKKLGQQETKQSG
TFDQAKVDRLKAAIADGSYELSAEKIADSMLQAMKGK
>hypothetical protein
MEEKEFYRQLSQFKRLLLKEKQALIKNEGQKVADLLPQKEKYLPSLQTYTGPLGEQTK
SLLREIKELQETNLLLTKQALSFGENLMKSIRTNLKSDQANVYQDQKRKNPYQQRPTN
SYIIDSQA
>Flagellar hook-associated protein FlgK
MVGLFGTLGTASSGLNASDIGLQTAANNISNAGTEGYSRQRVNLHTKLPYTLAGVGSI
GTGVKVDGVKRDVSDYIRQQIRVATSKQASATQRSTTLGNLQDILNEPSDNGVINQLA
NLSKAWTSVANDPEMTTTKSVVLEHASSFVDAIHNMASSITDLKNTTVNGVAQAALDF
NKKLETLNELNTQIYNSYAIGETPNTLLDSRDQVLKELAGLADISTVTDEYNRVSVYLGE
NPQAENVVLDSATKQVNNLSVVSDSQAGKSSVIVDGDITGKQVELAGDYPAGTLLILKQ
NASGAYSISKEAQVVEGDLGGYQDALAEINSAMTELNNFTNTIVKTVNLVYNSNGTTSP
GKAGLYDGFFTLATDSQNPAVDFEVNKSLLANPALLKTGTTTASGDSSIATAIVAAFAA
PIPTDGSDASVLAAYDPANLRFKTNTSGSTYATRFNNIVTTNGIAKQKADNTQAAQDTL
LNSLNQQDQSVSGVSINEEMADVIRYQNAFQANARMMQVVADCLDTLINKTGV
>Flagellar hook-associated protein FlgL
MRISSSMTYTNFLNNYQVNANTMQGTLDQLASGKVVNKASDNPLLVSKIMDSKVALAN
NAAYANTIADAKSWTENQYSVLNQVTKSMNRIITLITTAASSTNGIDELTAHKHEINQDL
QNIVDSLNTSYGGKYIFAGKNTDTKPFEMQKDANGVLTGISYGGTKDGLPREISAGVTV
DLFADGSKFMQVTDKEGQTTNLTDFVANLMNALGNNSSQSIDQAKLGTELLSQAQAF
YDNFTTVQTTVSAKETRLQAAASRNETEKTNLSTQLSDEQDVDYAEAYMHYQNQMVA
YKATLAVGTKILQTTVLDYMR
>hypothetical protein
MNNYNIDRVERAERMVNEIAAVEKIQKIADSQNNQVENQLRDEAVGEQRLSQPPIKKT
KTQKRLKHSVTSTDIDNVDLDDEFLLEIIRRLRDRFKLADLDLDLRLERTLPRQVALLVK
SKQELIKILSVAELVALYRRESFSKQGISLDEKG
>Flagellar hook-associated protein FliD
MATSSISATTSSAFSAATSNLNKSTTDTDKTTSKSSLTKATDASIKNNLGKYTGVTLDVI
DQLIEADPLSTRKSSDEDLIKNLNEQTTAWKNISASLSDLKTALNYLKDDDQYYTKATSS
SNQNVATISASTESGIGSYSIKVSQLATAMKISGGIPAVLGTTDADSSQKLNLTGSLNLT
SSVGTATLNINDSQNINDIATAINKLTVTVKDDKGKDVVKSLGITATVVANHLILRSVATG
KDNTISVSSTDGTADSLGLGSQAHSVAGQDAVYNIDGYDLTSASNTITNAINGATITLTG
LSAKEADGTLADTTLNVTHDTTYLEAAVENFVNKYNATMKYLNDYTKIDTSTDNLDEKQ
NGVLEDDSAASRLKSELARLAVSLGVDTDSKTGQKIDTGMTSSKLGISVDRDGIMSIDA
EKLTKAIDSGNGIETIKNFFYNTTGRISKEEHGYVTGFLETIKAYNEPDYVKNTKGIINDK

LEMYSKQIKDLNSDVTDLETRLATKRESLINQYSAVDAAMEKAEAMLKQIDDWFKTDD
NNK
>hypothetical protein
MDEQQAVKLILASLAAFDGSAQMALEIVNQLEPLFAQLQTRPPQLSEAELRQLEAGYN
HLIAVLAAERRNLSKQLSELSGPKQKRIVNAYSKHMKSSMKMTF
>Flagellar biosynthesis protein FliS
MVYGYGPKAGKADYLKTQVLSASPAKLTSMLFDGAIKSIKLAQIAIEQHEIEKAHNNIMN
AENIIYELANSVDTKVEGELPGQLLSLYDFMNSQLQEANIKKDKQKLDRVLDMLTTLSN
SWKKIAE
>Glycosyl transferase, group 2 family protein
MCKMDLAICILTKDEDDKLRRCLDSVKDVGAELIVIDSGQNESTKALCQELAAEYHPYD
WADNFALARNYALEQTLSNNIYFLDSDEWWAKEDVGFNTSKLLKYFNRPDLAEAIGTA
TIINHRQKDDGSEEEVGQLKQRAFNKSYFRYDKEVRVDEVLSSTLGFDSLAFNLVNVT
KVHHDGFLTAESRTKKSAQKLGLALKDLEAHPGDNKYLLTLADCYQEGGQTEAAKLIY
NQVVSATTGRDGAELYQTVIENEKMLVNSPDYHFMAGEIYGMHHEDAKAQAHFAKILA
LQDEGIDSTYVPVAHYFIGKLAEAKGDFETAKKHYRLSGEYPDAIDALAAIL
>CDP-glycerol:poly(glycerophosphate) glycerophosphotransferase (EC 2.7.8.12)
MNTRDSKEILNLITTLEEANEYLASSEEIGLELLKQVITNCLTASENISQTLQTFGEDISYQ
AELTTYQRLLQNFLANISQAGELAYPAQTIGECLAKLTAKLSAKHKKQVVFLPYKASMW
DSLATIYEAALADPNCEAYVVPIPYFTKDEHGKLKDLHYEGNLFPKEVPITSWQDYPLA
KKRPDIIYIHNPYDDDNLITTVLPAFYTDKLKQCTDLLVYVPYFVSLHERKHYEYALNKGI
ENADVVFIQDEVVKAGYQGAWKKFTQENPTYPKESLEQILAKMVALGNPKYDALNKTN
PSNYPLPSEWKQKIYRPDGSKKIVVFYNTTIIDLNASGRQMLAKYADVFNFFLKHQADY
TLLWRPHPLLIDAIRSQHPDLLADYLQLASQFKESQMGIYDDSPDFYLAYTYADAFYGD
MSSMAELFRKLNRPVIKQIPTYPTEQVPAEFDLNLVKEHLAKQKCLAEGKLPLAKLPAC
LAANTHSQKATDYHFGLDIHRYVMNLANS
>Integrase
MAKIKSYKSKGKTRYMFKVYLGTDPLTGKRIDTTRRGFSSAREAKQALTQLKAEFDAG
KFKTKNQDITLGELFEKWWAVYEPSVKAATSGHRYRTYSNHFKKYDDLKLSKLNALWA
QDVVNELAKEFKSYKTLLSVLKLVTNYGIRLELIEKDPFALVVYPKPLENSTKVHNTKNN
FYSKDELKLFLDRAKASSNPVFYPFFRLLAYSGIRSGEISALNWSDIDFNEHTISITKTVS
RDSRNNLNFIAEPKTEKSTRTIFIDDETLKALKAWKTTQAKFLLARGVNALSDSQLMFPT
HDNRIAGTTVANSRINTFYNQNPDLRRISIHGFRHTHATLLLEAGLNIKDVQERLGHKNV
QITLNIYSHVTQQKKQETVEKFANYLES
>Lj965 prophage protein
MAQAEFLKNNYINWLKEKLVFTDVENGYVSISTPFIDSNYDNIELFARFISKDKIEITDFG
ETLFNLEESGLKINKRSKTIYRLFEGILNDFGIMYNKENQSLSITTDVQRFPTAKNRLLQA
VMRVNDLLYLNKENVSSTFNDIMTDFLIKNKILFSPNIEIVGQNGISSHFDFSIPLPEGKE
RLIKTAARPNDINQAKIFNFDVRETANSRDAKYVLLLNDSNNPIGGNMNTNALVGLDKH
LAEVKGFEEIQKKPNLLAS
>hypothetical protein
MNGEADVLYFSGFNNHSKIEMRANRRAVKLLLEYCNCHDLSLEPVQFLENFGIPATFE
DLVKEESARYYLG
>Transcriptional regulator, XRE family
MLGNNIKELRKSKRMTQRDLAVAMRVSQQTVGAWETGRAIPGSDTLGELADYFNVTT
DYLLGRPEKKKEKQNVELTDDDIIMTYQGKELSDEDREIIKRLMNGK

>hypothetical protein
MNWKKFLLGNFDYTKTTKNGKYDVKINIQGGILPVIIIIALIVWLIIK
>hypothetical protein
MATRLKVLRAEHDLTQAELAEALGTTQKVISAWETGRVNPRPAMMQKVEDFFQVPKE
EIFFTAFNYSN
>Phage antirepressor protein
MNELIRVIKDNNGKSVVSGRDLHEFLGVKDNYTDWFKRMVAYGFTKNVDFIGLSEKSD
KLGGRPRTDHALTLDMAKEISMIQRTEKGKQARQYFIEVEKAYKEQQATPFKLPQNYH
EALLQLAEQVEINEANQPKVDYYDEFLSNKGLITTTLIAKQYGMSAVELNKFLHAKGVIY
KESGKKAWVLYANYAGLGLADYEEFAPNERTIRKTLKWTAKGERFIRDLLAEEGIKTNT
QLAEEVALSEPEIEYDGPYFSASEIAWQLRLPNEWVKIIGELANREHLKPIFNDQNIFCR
KTFDEYGRLRWEYTKYGARKIEAAINEELSKNVG
>hypothetical protein
MGIFNDGFLKELLLELMKRIDKLVEWKMKEAERQYMSVKEAAKYCGVSRTAFETTFAE
AGLRPIVIPGFARKLYSKKALDEFMMAYQR
>hypothetical protein
MKEKRVSVRTNVGFGPARLAVQEEQLIGNLDRMFNFCVTEADGQRHERGNIWEQIMF
PRDVAEAANLHLTDAVSWFKKIEEVRSK
>Recombinational DNA repair protein RecT (prophage associated)
MAQQTQQNQVALLQKDITDQVSKRITELEQTGLALPANFNAQNALKAAFFALDKTKDR
SGRPALEVCTKPSIANALLDMVTQGLSPAKTQCYFIVYGNELQMQRSYFGTVTALKRL
SDVEDIYAEVIHEGDEFEIDADEMGRTRVAKFKPSFANLDKPLVGAFAVIKKSDDSLVYT
IMTKTEIDKSWSQSRNKNNHVQANFSQEMAKRTVINRAAKMFINTSDDSDLLTGAINNT
TENEYDEDRNEPRKVVAKEKQPEQKSISDLFGQVEAGKPEKEEKPVIIEAEEKEPAESE
EAEVEEVKDIITDYEKNGEGEIDMFRQQQELSKEAGAK
>hypothetical protein
MELTDKNYYDKASSLEYLSVSLFKEFMKCEAAAYAELKGEWLPERDETPLLVGNYIHS
YFESPEAHEAFLDRKDAKGVSNRDKMLNKRIKKEKVLKKDFKVADTLIERMAADNNFM
QLYGAGDKEVIVTGEINGIKWKGKIDSLRLDLGKFFDIKTNKDLHGKNWIFKDGRNLPTT
FVEAYGYHLQMALYRELIFQQFGVSCEPVIFGVSKQDPPELMTIHFETEEMQDLLYDGL
ATIQEYQEHIKAVIDGKEEPRGCGMCDYCRSKSSFANNIYGALDIPLR
>Phage replication initiation protein
MSNLLFDERPLVIQPTLARMLGSLDEAVILQQIHYWLVRSTNVNDGYKWVYNSMTEWH
KQFSWLAIATLKRKFKSLEDKGLLITGNYNKAKFDKTKWYRINYDALDEMKQRLYQNDT
TRVSNCTNGEYQNDTTNTIDYQENTSDKAVAVEEQEQMDNVSNEADIFGEVQACGIQI
TPFNQQLLVDYIQKLNPELIVYAAQQTSVSAKHPNFNYFKAILERYIQLGLESVEEAKEL
EAKFEANKQQRNRSYNRKPTSGKSKWGGQLNEYDVTF
>Helicase loader DnaI
MSMMSLSSLASVKEVEETCPKHGCKLFASINPDRPAICPECRKEQVTTKDKGLVDNFR
LLIRRDIYFNKSYWGDSELALKTFNDFKGRADSPELEFRRKGYELAKDFVEPGNKVLNV
IMTGGSGRGKSLLAAAMLHFILEKSDKECLFININEYYEAVKEFKFNGGREPKISLDNIIS
ADVVVIDDLGTESVMNSDGREGGIMAQKMLYDIFNQRNHTIVTTNLTLEQLQETYNPKV
VDRIMKGINSQTSERILDFKDLKSKRAFSGLI
>hypothetical protein
MENEGKGIYVFEPQKRSKEQIEASRKWVKETLAYLRAEQARRGVKNW
>hypothetical protein

MSQYLLLVFLLAIAVALTMTATLFLTIVIREIIHVIQERF
>hypothetical protein
MYTNLDDFSLDLNMLLDVYDDKRIKFKASEELRSSLAASNVKVVDEYLTKAQTALEKAV
SVIDDLEAGLEVS
>hypothetical protein
MKTRIRQLRAQANLTQVQLGKAIGVAGKTISMWELGVSEPSNEQLIKLAMFFEVSLKYL
KCQKDRLYEIVNLDAQIVEAFVINDFEKTLEIAKSLTGPDIVNDTRVGDIVISCYENGTAD
NVWPGESHIWTQGLDDLLSIKERDNETS
>hypothetical protein
MKIDTADLSKKMISLTECLEDYCNLSLDRAETEISEGSSETVEVIESILNDDRSNLLRIVY
GKEVGNLQDKMGALRISLSEVQQIASNLSFRIYKDSEVLGR
>hypothetical protein
MTNEEKLKKQLQITTLALARILGKTFDLVYMIEEADVDKRNAFDDLVDWITKTSAKAQKA
LELVEEEYKEASNER
>hypothetical protein
MLLSVAIKSISKATGEEANKIILDIPVAYKNMFRKKAGSADDSKA
>hypothetical protein
MTARRKVTSTSHFGKKSVVDGYKFDSQKELDFYLRYIKNSGYEFEVQKNLVIVDKFTL
GSHNVRSVSYKADFVVLEGGRIKHVYDVKNGFNGYAIDDKSQLKFKLVAQRYHVPVEV
VVLRKNDFKVGVLGTTKKIEVQTRTNIDYEYSELIGN
>hypothetical protein
MKYFLVYIEASGIWTNVLVKANSLEEAEVIAALEVNNDGPK
>hypothetical protein
MDQNKFMARYDLKHRMTEKEYKSYKADLDKVIEHEALKKLMSLGEVIGGAK
>unknown
MNSLDAYVSKNCDPRLLGILSFAIREAYRYLRELLKNNPILEKSEMKKCIGHIRNGLVDV
AIKDVLLSSGIECNIQDKSVSRYPNGYTYLMIEVKGAILTTAKTASVRSVPKKALHRSRG
SILNRQYNLFDNADDLNDEYGSNTPAYILVTYGGKNYELEYINLGLPDLDVENWKAVKD
ITNVPMLISKTQKKEVKDKLNLKFTAYSQEKMGELENGEKTI
>Transcriptional regulator, XRE family
MKALAEQANISRQMISSYESGKTIPTGKNILKLAQVLDFPSSFFSREVEALTSDGMFFR
SQSAATKRARDMQKMDLTFSNSIYNLLKKYVEFPVVNLPQLYVDDYKELTSEIIKDKAT
ELREIWGIDNNSPIPNLIRIMEANGFIVTKSNMENPKIDAVSKWVLDRPFVIFTDNGESAV
RRRFNAAHEIGHIILHSEIESIYNLNATELKKIEKQANEFASYFLLPDDAFLDSLVTINLEG
FIQLKKYWLVSIGAMVMKTYNLGILNDNQYTYIQKKMSKNRWKKVEPLDNELNVEQPS
VFLEVYKMLVESGLFNNISLNTKIGLPQDYLRNMIGDCILVKPKATDDVHLRLIK
>hypothetical protein
MKAVNLNSFGYTVWKGNKIYPASFQCGYCGSYVTSNYGMALGYGEEYNFYNNKPYG
VYICTKCNKPTFIPSEGIQIPGSKFGSDVKGVSEIVNTIYNEARDCYAADAYTGTVLLSR
KLLMHIAIDLGANEGKNFKEYVDFLANEGFITAKSRGWVDVIRKISNASNHELVKNTKED
AKNLLQFCEMILKTNYEYPSSLK
>Phage terminase, small subunit
MSRIEDAEKDYLAGMKYKDIAEKYGVSINTVKSWKSRNGWQRDATKKKGAYKKQKRV
HTKKSKVAQARSPNVIDELVENDDLKDRQKAFCLYYLQRYNATWAYQKAYGADYETA
RTNGPRLLANARIKKQLTELKKQQSAELYATANDIMLSYLKQAHSDVTDVLEFKTVKRL

KWNKVPDDTGEYEDANGHYRLDPKIDPETGEQAFYYENLVLLKDSSKVDTSNIKSIRIN
DGEAVVEMYDKQKAMKELLERLPSADDTSDKPKTVVLNDIEEAKQDD
>Phage terminase, large subunit
MTNVVKLSEQINPHFYGMWNTKKPYIIAKGGRGSFKSSVISLRLVVNVLQQVQENRKA
NVICVRENATYLRDSVYNQILWAMDLLHVSDEFRAYSSPLRIVHRRTGSTFYFYGADD
PFKLKSNTVRDIVSVWFEEAANMKGPDVFDQAIPTFIRQKARYIDTVQVFFSYNPPRNP
YDWVNEWVTDKESDPDYFIDTSTYLDDELGFTTEQQLKLIEKYKENDPDYYRWLYLGE
VVGLGTNVYNFDLFKRADQIPDNDYLTDLYFSMDIGHDVSATTCGAYGLSVNGNLYVL
DTYYYSPAGKVRKKPPTELAQDVHDFVEEICDRYNMDPANMTADSADGALDNQYYSM
FGIHWHKVAKKKKVEMIDRAQDMLAQGRIFVLDIENNQVFLSEHRDYRWDEKTLNSDD
PKVIKEKDHTCDQFMYLCLDNERDFDLKW
>Phage minor capsid protein
MGVLSTVKNWFRKGGASLGMIKSLTLVTDDRRIAMDPSEYTRINVAKKYYSNDFKPIEF
INSYGDKRTRKYEAVNVTKLAARRLASIIFNEGCKVEIGDDKKANDLLESVFLDNEFYLT
FEEYLEKWIALGSGAIRPYVQDDKIKLAWITADQFYPLHVNTNDVKEAAIASKTTAVEDD
KNIYYTLLEFHEWQGDNYVITNELYRSDSADSVGVQVPLSSIDEYADTQETATLTGLVK
PLFAFFKTPGANNKMLESPLGLGLIDNARSTVDAINRTHDEFIWDVRSGKRRMVVPKS
WLKRPNANSRRRDNDTHPPMFDPDETVYQAMYGDDDAIGFHDMSVAIRVDQYSSTM
EFFLHEFENEIGLSQGTFTQSASGIQTATEVVSNNSMTYQTRSSYLTMVEKTITQLVDAI
LELAQCGELFSDGKARWTGDVQKVNINIDFNDGVFIDQDAQLKNDLQALSANALPLKQ
FWMRNYSLDEATADEWMQQLNSEKVGNDPDPSGEVGLFGGADDGTRTDVGESGQD
S
>Phage minor capsid protein
MELEQMLAKADKIADYYVKLQQKIFYLLIDNFKTTRPELINQDDPDSILEWRLRALAKIGA
LTKDTIRVVANTSGKAESYIYDLVKNDGLEVAKDINAELSDALKQNKPISPEVNSIVNSY
AAQTFRDIDNNVNQSLLSTNRARNGAVRVYQDIINQTVLEVQTGLKTPDRALKDNIYKW
RDNGIKTNLVDKAGHNWSLEGYTRTVIRTTTARTYNDLRIQSMKDFDSVLATMSSHPA
SRPACAPIQGKIVNIVPRESSRYDPEYPSIYDHGYGKPSGCFGINCGHKLYPYIKGVSH
NFQKQYDPKEAIEKQKIQQKQRYYERNIRRLKYNLDLARRQNDVSSERKFNQAIRGYQ
AKLRDLVKYNDFLTRQYDREQIISPNRAVIEQFRQDIRYNMNRKKVVDKTSVPISRPEL
NRITKSFRKSGGKILMGPEIDERLSNIGAEASIIGDDLIMFSSKAGRAAIHEELIHAKQKR
AHGEPKNNAEIYRREIEAGNILLENSDKWKLTEQEIENTRLLIKKYTKDLEDLRDGE
>hypothetical protein
MKALDVYYLDFKDTTCITVPSLEFKVGQKISDEQGNIFEIKGLSTFSGLTARRDVLNLIVQ
GKFEGNTVTLL
>hypothetical protein
MKKNVVEVCDFCGAEITDELLGVSPENDNVLICEECVEKAQKNLQDAKGNKVGVSVFG
IKYNGDTKASTVDVTLDLKLWCKPTFDNMKLVEKLGRNINDMKL
>Phage capsid and scaffold
MQREFLQNLGLSDDQVQAVLIQHGKSTNEIKGKLARAEEQVADLQNQIGDRDKQLKKL
EKTVGDNQELAQEIDKLRKENEQTAKDYQSKITKQAKDFAITNALKDAGAKNTKAVLAL
LDLDKVSVDEDGQLFGIAEQLEELQKTDAYLFTPKQIELEKKGPVNLFAGGNPSSNVAK
DPKNMSLDEQTDLYRKDPSQWQSLFGNNNK
>Phage capsid and scaffold
MTTHLSDMIVPEVFGNYVLNTALKTNRFVQSGILTPDPDLGPHLLEAGTKITVPFINDLD
GDPDNWTDTDNISVNQLTSGKQLGLKFYQSKAFGYSSLSQMISGAPIQTTIGNRFASF

WTRADEKMLLAVLDGVMGVTKVKNSKFYDATSKTPTDAAFSAKGFIAAIGLMGDLQDT
SFGAIAVNSATYSMMKLQGLIETIQPQNGAIPFEAYNGLRIVLDDDIPVDLTNKQKPTTT
SYIFAPGAVRYSSVLASTETKYDPIENGGTDAIVQKRVGTIHVAGTSVKASFAPAKGNF
PTMEEFGKSSTWEVVDGIDPRTIGVVAYKAELDPALVPGAEVATSGIGA
>hypothetical protein
MLSFSEYQELGGQVRDEEVYIKLERDATKLLNTATQMFYIRNNIDEDPDQWRVKMFKL
ALVAQIDYTNDVGASTEYDMAQKAVKSVSIDGTTVTTEGTFKDSSTKGVYNVALDYLY
QTGLLFGGVDVC
>major capsid protein gpP
MAKNTKPIAGTELTTNGAALNVVNKLYLDTTDATDLDDVTTGKWAWLALDITQITPSAN
ETSQSDADYAGNGFGSTEVTSKRYQLAVTGKRHIGDAAQDYVASKQFAIGNALHTRAL
WIDNGEAILAEVTLSNIVPTGGNANANQTFQVTIVFNGAPVAIDGKLTMSDSPDSGGTY
TATVAPQL
>hypothetical protein
MSVLNLDQKLTVDNKKTVQIGGKEYELIFNDKFAKLVTDMQLKVAEATKDFDDDAKLE
QFAKKEYAEQKEQLMAAFDKGKVVVVDALDQLLGEGEGERLYKYYNESTQALTALVA
LLNNAANDAVRENKAKNRAERRAKYKKNH
>Phage protein
MLSLTRDPLNKVIFEGKTYHLDLAFDTVIQYLQLSSDADLSDAEKTDCAIKLFLDDQDLP
SDPNFYELVFDAISKEITSEPYGNNVQGGNPFGLAPIKYFDYVQDAEAIFASFMSEYHIN
LLEQRGKMHWRTFKALFDGLSDRSYMQRIITIRQRDLSEIDDSKAQQQLNEAKNYYML
DDSSVTEEKQDQVEQTSALSAMFKAMQGQVRKGG
>Phage tail length tape-measure protein
MAADASVVIDFDVKTQQLESDKEQINKILSAIGENTGDKMDDEFKKSADNVVSEAKAVK
KDVDNELKKPTATITPKVDDSEAQKGTQKVITSLRKLPKEQKVKLDADAKKAGIDDFTS
LLKRVPKKLRTEILAQAQKDEVIDYESLLKRIPPKLLTDVQLNDNASDKLKAIRSEVESTE
TGFARLKTIVAGSFLGGAVLSGVTTVVDGLKDIAVEGANASDAMDKFRSTMQLGGFGE
EEIKKTSDEVMEYANKTVYDLDDISNTTAQLAANGIKNYMGLTEAAGNLNAQAGGNAE
TFKSVAMMLTQTAGAGKLTTENWNQLADAIPGASGVLQKAMKDAGAYTGNFRDAMA
DGQITADEFADAITKLGQTDGAKKAAESTKTFEGAIGNLQAAVVDGMKNIIDAFGKDKF
TGPINGLGNIVQDTFSKITKTIENNKKVIDALGGVFSGVFGAIGNTLSMFAMSFKANVAG
VLNNKEVTQNVESFKKSFDRVMEAIKPIEQAIGGLLGVIAGGAFSLAVDIVRSLAKGLGL
TGDEADKAKGKMDFSGVARVFSTVSQSLNIVLDFIKPVVRALGEMAGVIAKSAFATFAD
VLNAVGKALGAAFDKISPLIPALGDADKNMSAIAKHEGALKAVGVALGSIATAFTAIKAA
SAVGSVVSSAASAFMLLSGAIDIVGASLLVFGDFLAANPIVAIIGVVIALGVAVYEAYKHF
KPFRDWVNKTWDALKSFGSGVAKVGKQIYHGIVDGISDVVSWIKKNWKGLGLLLVNPI
AGAIKLLYDNNKGFKKWADSVIDTIKNAWKSATKALVKVWESIVDGIVKVWKPLSKALS
VTWDAIKKVAEITFKAIGAVILAPIVLVSATIVAIWKAVHKPLEAAWKAISKTASSVWNSIK
KTTEKVWNGIKDTVIDVSLAIYKPVKKAWDKVVQVVTDLWNGLKKTASKIWNGIKDAVV
DTTEAVYKPVHKAFTKVGNWIGDKWDNIKKFTSDKWNSIKDTIGNMTDKAHSVAKRNT
EALKSSVSEKWDAVKSVTGKAWGSVSDYVGEHSANAKNKASKNFNALKDTMSSILDSI
SQTWKRVWSGISDYFSGIWDSIKKNARNGLNAVIDFLNGGIGGINNVVHFFGGKKQTIS
PILKLAKGGRMSRSTLALVNDEESPTYREAIFRRDGSVEVPKERNVLTHLEAGDAVMP
AKQTAMLLGLPQYKGGFGDWLGKAADWIGNVSGDIGEWVADKIDQLEDALKDPLGVL
KRLFDNRKNDATAIWHTIGEGAGRYLPETAVNWFKNTLAGFKKKFDESGGNPPGAGT
ERWKPYVKKALAANGLPTSESYVQAWLRQIQSESGGNPKAVQGGYVDINTLTGDLAK

GLLQTISRTFNAYAFPGHKNIFNGYDNMLAAIHYAKARYGSRMLSVIGHGHGYANGGHI
YSKQLAWLAEDGDEFVINNRKPNADMLLASAIKQRAEINPNSFSAKIAHIIDSAKFSGSN
GYGMAPSTSQPNTVTRTAQSYEVTDYSKQIEAINHKLDIIADKQVKVDGRSFAVAYEKY
GSHERNQRNILEQRGLAVDVNF
>prophage Lp1 protein 51
MSTSNKAYGFTFNGRHSSEFGIRILDTKQVTLPQKRKSLIQLPYSSKQVDLSNVYGENV
YDERTITFPCKIAYGREDPYELYNKVTEITRWLYQPIGKSLLRDDAMPNYAFMGEVQVA
PTLEENYNFSKLTITFQCDAYRLKKRFDDVWDPFSFDLDAAQETEFDVKEFENLMLINT
GDREAELSVICENPVNLILNDQRYHLRSGVNSTVINENLILPKGESLVTIVGNGKVSFDW
TEEVI
>Phage lysin, glycosyl hydrolase, family 25
MSKGYRVTIREGWNGPEMLLNSDVYQRIRLVSATVSKSVSAYDSFTFTLDPSHKLYAII
KPYKFFVKVTRPDKNKVLFEGRVLTYSDSMDSSGVIQKTAVCEGLEGFLHDSVQPWK
EFHNTTPKDFLQALINEHNKQVEDFKQMRLGNVTVTNSTDNVYRFADDTKDTYDNIQD
KLVSRLGGELRVRNENGTLYLDYMPKIGGTATQKIKLAKNLLSSSRNIDPTEVITVLKPL
GATQERQNNQEDQTEASSPRLTIASANGGNEYLRDEELIKEFGIRARVQTWDDVTTAD
ALLQKGRAYLASQKAIKYEIQLTYVDLSYIKDNVAMFDCGDTVTVQNMLQGLEIEERIVS
MSIDLLQVQNTTMTLSNDPMAINAYRKQLLQAYNADMNAFYNLLKDQKLNTADLEKAL
KSSNSQIKSLQQALQELENKFNGQTPQPQPMHIGKIIDVSEWQGVIDWPQVIADDVSLS
IIRVQDGSTHQDLKYMENIQKCISAGGKYAVYAYFRGASTADAQQEARDFYNRTQRVV
AGKQQPVFYALDIESVEMGGAASQMRAGVEAYMNQLNTLGIPDNKIVLYIANHLYASF
NLNVARAGAIWIPSYGRNDGTVANSLRPTHPYDLWQYSSKGSINGITGNVDLSTEASD
KFKKLLQ
>Phage upper baseplate protein (TP901-1-like ORF48)
MSYRDDTPITEDDVSKLDKEISVGNVDQVAFQVAAWLREKMYGSDVREALAQWTLFN
AKITEYLINSNEAFKVDTNRVKKDLIARQTQVESRQTDLEDAFKTVIANATKDSEVILARS
STRYGDFKTVDDRIEYIEQLLGTYVPSGFTVTIKHNQNRNPNVKVRYYEYALGTEPDGI
GLGPKGSFGGTNNVDVPATVEYKDANTVLVHLPTNYRLTGAPIFEQDKWRLIDGYKAL
SFDLGTVDTTAAIKGNSGNSTSQDNNVVTAPQNLQATAINDTTEKLIWE
>Phage major tail protein
MKFYIYRGVGSSGELTKIAEVTDAKTYTDTGLQPKTIYRYAVSAYNGLRESAKSNIITVT
TAEIPIATITLAIDKTSLEVGGTAKITVTVTPPNQTSGTPTLASTNTKVATVDNSGNVRAV
GVGSTTITASLAGKTSNMLTLTVYEALVNVSNLSSSNVTASGVTLSWT
>hypothetical protein
MVVAKSLTVNSTVNLMYQEYSLGLVPIGTEPKGMFGGGNREALTAKVIANTNGQSTIEL
QSSFNKLAENTKLVLQNGVYCAFVGYKAIYLR
>hypothetical protein
MTVNITSIPRGDENGLDKINANFSEIRNKFIMNSWTDAGITPLNGVKLGEGNDADSKPL
YGYVLYSIAGQKFATICLRIGNLNNPNGPGKDVVQLPFRVHGYNISTGYDNKNVIVYNN
KVYTDAKNDTYVYGTFMVIG
>hypothetical protein
MKVVYVFDASTKAFVETKIVDEEYRLQSNETFVAPISEDGTGLYDPVWNGTTWSSLTK
EEWLAKQPANDEPDIKSGPTEAEQIQAQQVVMIANLTKQVQDLQGAVKTLVLQNAVAN
KEEK
>hypothetical protein
MYTYSFVNSFYKMGLFTKDDVKLLLEVKQFSQDDYNKMFPEDSSLTA

>hypothetical protein
MAVISVLFSSVYWLIRHGAKALNNAITIGTYPLQQQFKELTSTIKQLNNNFEEQRKSLKK
LEDEVEEHDKTILLHDEKIKRLEERK
>prophage Lp2 protein 58
MKKVFLDKDGKLNRKTVTSLVLLLIVLVQQLCAIFNLKFTGDVGQIMDLVNTLLTIGGILG
LVDGVTVDTETVKSIDKTANEALKLASQHQGKGGKHD
>Phage lysin, 1,4-beta-N-acetylmuramidase (EC 3.2.1.17) or lysozyme
MTKLNRKWLLSLTILVGLFLMPVTVNAARYYGVDGSRYQGNTLKKVTPEDSFAISQIGG
YYNGTFIPQTTYQSQVASGIAMGLRMHTYIYMETGSNQAQTKQMLDYYLSKVQTPKQ
SIVALDYESGASADREANTDNVLYGLRRVKEAGYTPVLYSYKPYILSHLNRQRITAEFP
NCLWVAAYRDYSVMTRPDYNYFPSMDGINMWQFTSTAIAGGYDYNVDLLGITLNGYK
NGNVEHPRSETKAIEQGQKADNTSKSAIQVGNTVRVKFGVKHWANGVGMPSWVQSN
TYKVQEVSGSKLLLGGIMSWINASDVEIISVANSNPSVTGTIYYVVKSGDTLGGIASRYG
TTWQRLQALNGLSNPNRIYPGQRLKVTGNVYAQRTYTVRYGDTLSSIASRYGVNMYTL
AHKNRISNINLIYPGQKLNI
>hypothetical protein
MVSEAQKKANEKWKAANKEKQKIYGYRSQAKKFISEFADKEDLKTLEELIRLRYEKIRG
EGNV
>prophage ps3 protein 01
MYDVFKIVNWLRVKNNADMKRNENVEELTQMKAMKLLYYIQAASLVVTGKRMFNENL
VAWKYGPVVERVHEKYRGQRAIVGEITDADLDDYSELEQDTNTADILNSIYDIYGYSSA
YDLMRQTHKEKPWQETKQSEIISDEAIKNYYSGVFEVEA
>Cholinephosphate cytidylyltransferase (EC 2.7.7.15)
MISRTQFELLTSLAKTGPIKATKVQTKSVAELLAKGWLLKRDDLLEISPSGLKELEPYQV
KQAIIMAAGFGSRMLPATKDTPKPLVEVNGRRIIETLLDALVAAEIKDITVIVGYQRQKFK
SLLANYPF
>Cholinephosphate cytidylyltransferase (EC 2.7.7.15)
MKVLDRLDRGTYICEADLYVTNPAIISKYQYDSNILGYYCQETTDWCFKLQAGYVTDYR
QGNTDCYNEYGISYWTEEDCAQLRRDLPAVYAQPEGFQQFWESVPLKWCHENYQVR
LRECQKEDIVEIDSYQELLQIRQK
>Ornithine cyclodeaminase (EC 4.3.1.12)
MNGGILLKIIDFEEVKKLAKKTKATDWYNWVEEVLKNKDIFVMPPKMSIPQDSGNYYHI
MPAMYEDENIVTVKMIGRHFLKENEKRPTMMADMMLYEADTGILKALMDAEYITTLKT
GVVAAHSAIKFVKKDFDTLGLIGLGNIMTVFFITFINKLRDGNDNRKLVVKLYKHHQQEV
RFAERFSNLENITFIYCDTYEEVITDTDLVVSAVTKADKNFVADECFKQEVTVIPICTLGF
QNCDLFFDKVFTDEIEQIRGFKYFKKFKAVANVSDVLNLKKKGRENNEERIIVYNYGLGI
HDLYFAKKFYELADVNDIDYSYPKSKYFI
>IgA1 protease (EC 3.4.24.13)
MFMPKKLMGEQKRRYKLYKTRRGWQLSSLLVLTSGLLLGLQTAKADDLSSSNQAVTT
ATATIEVNNVQTQEQNSTNSTATTSQASSVAATASSEESEPTADVVVSQPELQSSQAS
NLTSQAVSPSATSEVTSEVSTATPSATTNSESEVEGNFEDEVPYDPYIGTVFQEPSEE
EKQGWNYYREDFSLPAGMRVVQQVGNDERPTIIRVGLDDKIDVISEDKSVVEPVIYIKD
PDLEYGRVVQVRAPLYPGIVDVTSYSLSWGGTVYQGKTHRLRVNTIHPGIFKIGVKPKV
GYDVTDYQTIYRANPNLEAGQIRIVQKGIDTIKQVSETLTGVKTDLGEHYGLDETVKNVL
GLIEPGPTSVSENVIEDSNFDLPTDFSSINLFTSFFTTDGSVDLSPIFLDVNDIIGVGTEIL
GYLSTSTILQTINGGQDQIIEVGTKPTVIQNIIAYNTIYEATDILPAGEQAEKQAGVAGIKT

ETTTYQLNEATGELMASKEVNTKPAQNRIVLVGTKAKVVQTEIPYQVIYEVADTLPAGH
QAEKQTGVAGIRTETTTYQLNEVTGELTASKEVNTKSAQNRIILVGTKAKVVQTKIPYQV
IYEATDTLPAGQQLEKQAGVTGIKTETTTYQLNEATGELTASKEVSTKPAQNRIVLVGT
KVKVVQTKIPYQVIYEATDTLPAGEQAEKQAGVAGIKTETTTYQLNEATGELTASKEVS
TKPAQNRIVLVGTKEKVVQTEIPYQVIYEAADTLPAGHQAEKQAGVAGIRTETTTYQLN
EVTGELTASKEVSTKASQTRIVRVGVKPKRQIQLVDYQTKYQADDTLPYGYQQVVQNG
YKGLVTTVTTYQVDKTGQVLPTVTVTKTNAQAQVIRVGTKSVQTVTEQPYLTLYLADDS
LAAGLRQEKQAGKKQVTLTTITYNLTTDGQVISHRQIQTLPGQAQIIKVGVKTKVLVVPS
AFKVSYQADNTLAHGQRVITQFGEDSLTKYITTYSLNTDGSLNASETSQIILGQAQVIRV
GVKPTRQQTIIPYQVIYQNDGQLEAGKQVLLQAGKEGLLLRTIAYKLNENGTLTPYVTTI
TLPGLAQIIKVGTKVASPLQPATTKKELEVSQPVKTQAQPIPVPLVKTTAEDKLPATSPL
PTKAVTLALARVDGEIAQNLSKAELGLGNSDRDYQITGLVFKEVDQKLDQSVHQLGTIK
AHHPRMLVLDSAVFESKWNWLLIGVLTIIGIICFKEDKKPEQ
>FIG00751876: hypothetical protein
MLAMTLMYSSLAIIIFICRALFWENQKQLLVSLVLLLLEMLVVFTLVYFLLPSRNLLSLLLG
NAVWGGIFLMLTSISGKVTADQQVKNWVATTLPASLVALSLLIAAGGELHSILSVKPTYN
SISVKQVSSKQAPTFKRGETPIALAPKTVLNRVRKSVSDLPNSQYFKIAGTVQAQYLHG
KAVYIVPVEYQGFFAMLKAKTIPGYFMIDATSQNATPKFIHKPYKYTTSAYFGRDTERKL
YRNNPQWLKLGDGGAQLEIDNDGNPYWVETVYKSAFLSHRINYQKLRVIVMNAVTGT
TKTYKLSNLPKFVDEGITSDVAAELNNNYGSYQHGFWNQFLGKTDMEEPTNNGPEDG
VTSIFNANGTISYFTDFTNPNTKSDSALGYSMINARTGQLTYYKANGIMDSSGAKSNAN
QNYKAQQWTANMPILYNIDGRPTWIMTILDKTHAIRGYYYLDAEDQSIYGTGTSPISALD
DFRQALVNSGTKAANTPDSKLKQLTGTIDRVAIVSNKNKVMFTLQNSPVVYTIDTNDFA
KANLLRSGDHVSFKANLVNGQSIGNVSKFTNHDLK
>Glycosyltransferase
MANLLALCPYVAYRNEELLKDCVLVPYAFMTELGYDVTVMTAQKEEFTYLDQLPGLKL
DIVPTCPLEQWADYLVTYLKDYSKDIDVLFLFGMYGTYYPVVKYFKERYRGKIILKADAN
SQWVNDIINPNLSIFDYLIKNCDIISCEGSAMQNYLAKKWRRPVELIRNGSLKRDFNLEL
SPKIKDKVIINVGSQEAGRKNTPLLIQAFMNVCFDFPEWKLELIGPVDSTITELVASIYKT
HPALRKRLLLLGPIKDKAQLERHYQKARIYLSTSLQEGSPNSIAEALRNGCYVLTSKIDI
WQDTFAPKAGASFEHGNLAELVSKLRYALSQPQLLEDAFELNRAYSLKELIYEDEVRR
YQLLLDYQNTSTRQD
>hypothetical protein
MTKTNLDYYRFEYYSQSYSPLGRIKQDELYQQVLEYLRQQLKAPTEFSLAKRLMTQYF
SYLRSSPSPQIIVANSHKFDALLSQIAHAFNNQNLFFHLAINQNQTIAGANQQTYNLLPQ
QWDTLNQLPYISSLIIGPDFSFDAASLDALKAKLNRDGHLFELTATELHLERLGQTKGIL
GTESQAQELKTDFNRRLELLTTLTAKDKLKAILALEADYYRYYDYLDLASVELIEQINQLK
NETLSALSKEES
>Glycosyltransferase
MKNLVALCPYVAYRNEQLLKDCVLIPYTFMTHLGYDVTVMTAQKEEFTYLNQLPGLRL
DIVPTCPLEQWADYLVDYLKNYPHKIDVLFLFGMYATYYPAVKYFKEHYQGKIILKADAN
SYWVDNLINLDNPQYDYLIHHCDIITCEGVAMQNYLAKKWRRPIDLFRNGSRDVDLNH
HLTFNSKEELIVNVGCQEAGRKNTPALIKTFAQLAPSFPTWHLALIGTCEPAIQVLVERY
YQLNPQLKKQLLLTGPIADKKILEDYYAKAKIFVTASYSEGGSPNVAAEALRNGCYLITS
KIDAWQDTFAPEAGASFDHDDLVALTAELRRAMSDSKLCQQAFTLNRAYVQAELSYE
QQVSRLETLLKVRGTLNDKN

>glycosyl transferase, group 2 family protein
MLVNTNPKLTALKLLADQYQCSYYTYEWDDDFSAARNFAASQAQYDWIMTLDSDEYL
LTGSPLPFFMENNLSKTVGTVARTSFTSSEGHTGTSHEIEARLYSRKHFKYRGLVHEQ
LVARSAAPTEIDYVPSKIGLGHTGYQNPTVLTAKSYRNEQLLKKELAANPNDPYIYYQL
GKGYLVRAEYELALSAFKKSFANQKHATTYSLDCLKGLLLTYQTLKLNNNLVAAAKAYR
PLYEGDSDSFYFIASLFSLAGLYNQAASYYETCLSLNNSQFVEHRADGYQIWLKLGEAY
RKAKAIEQAEYCYQRASDKMDVDTLMLTGNSYFKQANYEQASMFYLAALKHDPRGLT
LLELVCQLCLTYLKVKNYDQALKAALIFKDELSGQADYLFTLGNIYLAQGKFEAALENYR
LTILRGNSRDEGINSWLSYYNQGVIYEYQNKLGLAQKAYQKCRNYEPARKRLQNLYPT
KKG
>Glycosyl transferase, group 2 family protein
MDLSVCIITKNENEKLKRCLKSLAFLEADIVVVDTDFNDETKKLVTELNGRYFTFEWCD
DFSAARNYSLEQARYDDIMVIDSDEWLDEKEANFNKAALESFMRRNLKTLACCYQVSI
RPNIDYGKNYAIVPRIFSKKYFRYEGTAHEQPVSIIENVDTRTIMINLKMYHDGYNGAEV
LKGKAKRNLDLLLKDLETKPDRVYLLFQIGRSYLGLGEYKLALEYLERTLNHRDFHYNY
PYSYSACDSYLQTLLALGKTELVLEKMAYFAKKFPNLADFWYDGGEIYLKLNKPKEAIA
SFKRATKCTRTEVSDENSDLAYYKLGQIYAYLGDIKQAKNYYEKCQPSFELAQKELLVL
RQAELKQAITIYILDHFNDAEILKLTLASLKGLPFEKVVVSQRKYSENDLAGDEVVQLTP
AVFAEYLKVATNYGLLIHSGERLVDFDLNQLKELLTSKKALSGYSLISSPNQIKALGYRS
YYHELRLVGGVSHDFYAGLLHPNAEAEEELIPLTIESPLEVRASLGFETTASVYANLLFM
QGNYQAALDHYQAYFMACDYGLEESLHSAANAILTSVFLGQLDWLATYLPQWSHYYQ
DFAIYQFSLGYYFKALGEEEQAEQAFTEAKRLANSLNPLINAENLGALYNAR
>Glycosyl transferase, group 2 family protein
MQGNSYQMGLVALKDGKVQLAKELFTVSNDSRAKDQLARLEKRKSEQMISACIITKNN
QATIEECLASLYKYNVEIVIADTGSTDKTVELAQQFTDKVYHFDWIDDFAAARNFAAGK
ASFDYILSLDSDEYFIEGSREDLEAFFRKEDTSCIVGSYELINVLGKEKRRHDRSSDLVR
VYNRRYFKYYYQIHEQLDKIANSKTKYTKTCAINIKGVHYGYQDKKVLIKKSKRNLRLLL
KQLDITPNHPYFIYQAAKSCTILENYSQAFKFYQHFFEVPQDHNLEWVRDALVTYLKLC
MDTKNYIRALKSVNVFQNQYLDSFGMLCLFAEVYMANKQDQEAEQALKLALNLKPSKY
GEFGCVKADAYEALARLYERQGKWVLVQEMLAKAEDETLQLKLNRHLTMIVLADEEAD
KLTACLQSLAVLKANYLIIVKGTNPKTIELTRQYSEDIYHFSGRAADLGNYCLTKLKTDY
GLIVHANEKLLGADLDVLKQITSKKAGIVGRVRIMVPVKDDELGRQTHELRLLVKAYSKF
NGPKLSDITNKKPKQLLRQVELPLILAKNLAYEEYYQSELTKQTQLEASLKEKASPIGWY
ELAESYYYLGNYVAASQSYFEFLSGPLDLSKEETLTAVTNLISALIEEKDYQRALEVVNT
FKPYYQELSDFQFVAGRAYMNLARFSEALLAFESAQRLANSDSIGINTYLSYYQQGVIY
EVLGKKQRAIGAYRLAGRYKLAQAGIKRLLAEK
>Flagellin protein FlaA
MRINTNTAAMNTYTKYTAAANAKASALEKLSSGSRINSAADDAAGLAISNKMTAQISGM
KQARANAQDGISLIQTAEGALEETSKMLNRMRDLTLQAGNGTLQDDDRGEIVLEMKTL
VSNINAINDNTKFNGMTLFGDSTTATKTFTFQVGANADETATAKIEEMDASGLGMDDLV
KRLNKQGGKDGTKDFTADYARDQGELLKAGSDTPSVVPSEFEKGALSILDAAIKKVSD
QRAKLGAAQNGLSHTINSLEATTTNLAAANSRIKDVDMAQEMTSYTSSNILLQAATSML
SQANSMPQSVMTLLQNIG
>Flagellin protein FlaA
MRINTNTAAMNTYTKYTAAANAKASALEKLSSGSRINTAADDAAGLAISNKMTAQISGM
KQARSNAQDGISLIQTAEGALEETNKMLNRMRDLTLQAANGTLQGEDRKEIALEMSTLI

SNIKAISDNTKFNGKNLFTGDAITFQVGANAGETASVSISKMDPEGLKFDKTLKKDIDAL
ATANDNTTAGTVDSAGKAAFDTNFKTNLDALDTAIKMVSDQRAQLGAAQNGLAHTVN
SLEATTTNLAAANSRIKDVDMAQEMTSYTSSNILLQAATSMLSQANSMPQSVMTLLQNI
G
>Glycosyl transferase, group 2 family protein
METDLLAVSACIITRNEAEKLRRCLACLSKLVADIVVIDTGSQDETAAVVAGVEVARYFY
YEWPDDFSLARNYALKQAKYDQILVIDSDEWFASAKLEQVRTQIADFFAKADLRTAVG
SFVRLNQTESGWQASQYLVTRLFSKKYFRYQGKLHERPMPLAGEKKRKEVKLELELL
HDGYLKPEVFEQKAKRNLLLLLKEQGLPAEAAKNPYVLYQVAQTYRQTGRKQMALVY
FRWWLLAETNVGQSYFQAGLATYLELLLECEQVEFASQLLTAFGDLQLDADNSYLAGV
IYSQVHDFDKAVAHYRQALALGNAKQADRSGQLGQLMLANALELSGKPNEAVVHYFA
ARLNGLSDLKRLKERQLVADYYLALDLGQSHLKLGLFDEEGNLVASRIEAVQLKNYQQ
VEATVMTIIANYQFLISGVGIALFGRVNTQKGTIGLVTPLGYEQPLVERLEAKFNLPVVVI
NDAKAALLAEQTMGSLKAVSDGALVTLGTSIGGGVILANRIHEGRHQLAGEFSYLVQNK
RFLSESASAVKLMTALRDLLGAAKEMTDHQVFLAALENPLAKELIQTYSYNLATVLVNL
QVSFDLERIAIGGGLGQEEKLITLLNQSFLEIMGSLPNNPLYRPELVAAKLGNQAGLYGA
YQLLKERMKR
>Glycosyl transferase, group 2 family protein
MRVSVCVITKDESKKLARCLESLKSLQGLVAEVLVLDTGSSDDSVQVAESFGVRVYQQ
EWREDFAQARNYLAQQAKCDLVLFLDTDEWVAGVDVPALEALLAQVDLQHTVGRLVR
LNLNQTSQTQSLEPRLYLKSYFNFKGKIHEQVNRQDGGPVYYQDIPLVIKHDGYASQE
LLLAKAQRNLDLLLAELAKQPDDPYLLFQVGKSYYSQKNLELAQKYFEASLKLKPAYSL
GYVFNCLECLVLILLATEQVSAAQRWLATYPEYEHDGDYIYLQALSFFVEKNYDFAKKC
FNLALKQPIVKVAGRNNFLAHYYLGLVYKEEGALGLAKQHLSLAADKFSPAKEELLQLL
Q
>Mobile element protein
MAHKGYRYTLDEKLRYIKLVKKGATTIAVEKEYGVDHHNLNRWLERYDEHGVDGLLPR
PVKRYSTEFKMLVVKKHLEGASLGQLRQDYDISNEGVIQRWVNLYSDGKRLKIKGNKQ
AKGKQSRKTTLAERLEIVNWTIDNHYNYNLAVEHFNVSYTQVYSWVRKFSDGGKTALI
DRRGKRKNG
>Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-) / Penicillin-binding protein 1A/1B (PBP1)
MTKEHTPTSLLKKVGYSLAVICLLLLICGISLCGYYILKAPKLTSAYVTSTASSTIYDNQGK
KLLSLGVDDREVIKPNQLPSQTSAAITSIEDRRFYSHHGLDVIRILGAALANFKASDNSL
QGGSTLDQQLIKLSYFSTKKSDQTLERKIQEAWLAVELDRIYSKKEILALYANKVFMGN
GIYGMQTAAKYYYGKPLKDLSLAQTALLAGIPNAPSSYNPYANQTLATQRRNQVLEAM
LQTNKISNSAYQKAVATPITDGLLAYDSQRNQTSHKHALWADAYLKEVVADLRAKGYD
PYEKSLKVYTNLDTAAQHRLYQLANSDQLNYPNDQMQVAASVINPKSGAVVAMIGGR
KQGELTFGLNRAVQATRSNASTAKPLLDYAPAIEYLGWSSQHELDDSKFVYPDTNKVL
HDFDGTYLGKMTMQEALVQSRNVPAIRALQAVGMKRAQSFIEQLGFTYKHKLELQNGI
GLPSSSLQNAAAYACFANGGYYYQPSYITSIELPNGQTKSFKTKPKRVMKASTAYIVTK
MLEQVISSPSGTGHDAAVYGLYQAGKTGTNAYPSDLKPASFPDNALMDSWFNGYSRN
YSISVWTGYDKPYESGHYLSEADSKIAQKFYQDIMTYLAQGQVNKQWTKPASVKTIKA
NNQTYYAPK
>Lipoprotein signal peptidase (EC 3.4.23.36)

MPIYICLSLAVVLLDQLIKYFVATTLPLNTSQPVISGLVSIAHIRNYGAAWSMLIGQRWFFI
IISLVAVAVIGYYLKKKWTNWGYGLGLALLLGGTLGNFIDRLRLGYVVDMIQLDFINFPIF
NVADCALTLGVIVLILTMIKDEQLND
>Ribosomal large subunit pseudouridine synthase D (EC 4.2.1.70)
MIEAVVENESGRLDKVASQLFGNFSRSQLKELLASGNLLVNGQQKKPKYQIKPGDKLT
LVEPEVKQLDLVAEKMDLDIVYEDDDVLVVNKPQGMVVHPSPGHESHTLVNALLYHCP
LSTINGTFRPGIVHRIDKDTSGLLMVAKNDLAHQSLAAQLKNKTNTREYLALVHGNIKEA
EGTIDAPLGRSPKDRKKQAVVADGRPAITHFKVLDRFGDYTLVACRLETGRTHQIRVH
MKYIGHPLAGDPLYGPKKTLPGAGQYLHAAKLGFKHPVTGEELLFEAPLPPYFAEMLK
KLDKRRKIN
>Uracil phosphoribosyltransferase (EC 2.4.2.9) / Pyrimidine operon regulatory protein PyrR
MAKEIYDSMAMKRALTRMTYEIIEKNKGIDNLVLVGIKTRGIYLAKRIAARLKQLEGADVA
VGELDISLYRDDRHVATTDQPVVKDAKLAVDITDKHVILVDDVLFTGRTIRAALDALMDK
GRPKKINLAVLVDRGHRELPIRADFVGKNIPTSLNEQVAVYVEEVDGKDGIDLKDLTE
>Uracil permease
MTKNNTEFRNDEAILDVQDKPSFQHWLVLSIQHLFTMFGSTVLVPLLIGLNPSIALFSSG
CGTLVYILCTKAKIPAYLGSSFAFIATMQLLMKDYGYPAVGQGAIAAGLVYVIVALVISKF
GSNWVDKILPPVVVGPIVIVIGLSLATTAANDAMLHNSNYDLKYFAVALFTLAVTIIFNMF
LGGFASMIPVLLGIICGYILAVLVGIVDFSSVAQAKWFSLPAFDVLGVNYHFKWYPAAIIT
MAPIAFVTMTEHMGHVMVLNKLTKRNFFKKPGLHRTLLGDGLSSVLAGFVGGPPTTSY
GENIGVLAMTKVHSVWVLGGAALFAIIFSFVGKISALIQSIPTPVIGGISFLLFGMIASNGM
RILVDNKVNFELKRNLVITSVILVVGIGGTYLKLGDFQLTSIALSTLFGIILNLVLPQQAVSE
RKN
>Mobile element protein
MLCVQKVRLYPNEAMKQVLDDLCDYRRYCWNQGLALWNDMYDASLILEDKKLRPSA
RKVRDELVANKEDWQYQLSARCLQLAISDLGKAWSNFFKKSLPDWGKPKFKTKKAAR
QGFKTDRAQIVNGKLRLDKPHGVKTWADISFKGAKSLGGDLKVVSVYRENGKYWASL
PFEVKKAQKAKTGQQTAVDVNVGHFDYPEGQVKTLPNNLKALYKRIKHYQRLLARKRV
TNGKGATQANNYVKTRAKLQRDYRKAANIQHDLVQKFTTMLVNDYDRIAIEDLVVKQM
QMSHVASKGLHRSMFGYFKQVLKYKCAWYGKELILADRYYPSTQRCSACGHVKTGE
DKVGLDGNKKHKTKHNEYICYDCGAVMDRDENAVRNLLALL
>Aspartate carbamoyltransferase (EC 2.1.3.2)
MVSLKHFVSVENLGETEVMALLKRAQEFKAGKQLELSQPVYAVNMFFENSTRTHTSFE
MAESKLGLKELAFNPAQSSVNKGESLYDTLLTMDALGVDLAVIRHSKDNYYQELIDLKA
DQHLNLGIINAGDGSGQHPSQCLLDMLTIYEEFGHFAGLKVAIIGDITNSRVAKSNMQLL
TRLGAEVYFSGPKYWYSHEFDSYGQYLAVDELLPKMDVVMLLRVQHERHAGDQNEA
DFSAAAYHEKYGLNQARYAKLKPQAIIMHPGPINRDVEWASELVEADKSRFYQQMQN
GVFIRMAMIEAVMRGRKLGGLN
>Dihydroorotase (EC 3.5.2.3)
MKTLLQNGLVYHDGQLTNQDVLIEGDKIVALGQNLGQLAPDAKVLDVSGKLVSPGLVD
VHVHYREPGFTHKETIKTGSAAAAHGGYTTVCAMPNLNPVPDQPELLSQLVEKNQAD
GKVHINQYSAITKGLKSEELVDFKAMKEAGAFAFSNDGCGVQTAGVMYQAMKQAAAL
DMAIVAHVEDNSLLFGGVMHAGKRAEELGLPGILGISESSQIARDLLLAKETGVHYHVC
HVSTKESVELVRLAKAHGINVTCEVAPHHLLLADEDIPANDGYYKMNPPLRAKADQAAL
VAGLLDGTIDMIATDHAPHSIDEKTGDMKQAAFGITGSETAFALLYTKFVKPGIFKLEQL

LDWMSLKPAQAFKLAAGSLLPGSLADIAVFELDQASEIKESDYLSMGVNTPFTGQTVY
ANTYLTLVAGQVAYQKEAR
>Carbamoyl-phosphate synthase small chain (EC 6.3.5.5)
MERYLILDDGTVFKGEGFGAKNTTTGEVVFTTGMTGYQEAVTDQSYANQILVFTNPLIG
NYGVNLDDYESLEPNCKGVICREVARVSTNWRKQDSFPHFLEQMNIPGISGIDTRALT
KKLREHGTMKGSIIDATDNLEHAFDQLHATVLSDRLVAQVSTTKPYPNPGNKHNIVLVD
YGSKHSILRELAARDCNTIVLPYTATAEDILRLKPDGVLLSNGPGDPEVMKDAVEMIKEV
EKHLPVFGICLGHQLLALANGAKTFKMKFGHRGFNHPVREIATGRITFTSQNHGYAVD
PASIDNTQLIVTHKEINDGTVEGIRHKYYPAFSVQFHPDAAPGPSDAVDIFDDFMHMIDI
KKGEKTNA
>Carbamoyl-phosphate synthase large chain (EC 6.3.5.5)
MPKRTDIHKIMVIGSGPIIIGQAAEFDYSGTQACMALREEGYEVVLVNSNPATIMTDTEIA
DRVYIEPLTVESVSRIMRQEFPDALLPTLGGQIGLNLAIALAKTGILDELGIQLLGTKLAAI
DQAEDREKFKQLMQDLNEPVPPSKTVGTVAEALEFAKECGYPVIVRPAFTMGGTGGG
LCDNEAQLKEIVANGLDLSPATQCLIEKSIAGFKEIEYEVMRDSANNAMVVCSMENFDP
VGIHTGDSIVLAPTQTLSDREYQMLRDCSLKLIRALKIEGGCNVQLALDPNSFNYYVIEV
NPRVSRSSALASKATGYPIAKMAAKIAVGLTLDEIINPVTGTTYAEFEPALDYVVAKIPR
WPFDKFAKANRRLGTQMKATGEVMAIGRNAEEAIQKAVRSLEIDEKDLLADYVRAASD
DTLEEKIVHAQDDRIFYLTEAFRRGYSLEEIHELTKINLYFLDILKHITELEVEICKHPHDA
RTLKQAKRYGFSDYTIAKLWNVSPAEVRAFRKAKQILPVYKMVDTCAAEFESHTPYFY
SSYDQENESQVSDRKSVIVIGSGPIRIGQGVEFDYATVHSVKALQKLGYEAIVINSNPET
VSTDFSVSDKLYFEPLTLEDVLNVIDLEQPEGVIVQFGGQTAINLAEGLEQNGVKVLGT
TVEDLDRAEDRELFDQVINELGLKQPIGKTATTHEDVLKSAQEIGYPVLVRPSYVLGGR
AMEIVYNETELEGYLAQNAPAEIEHPILVDAYLEGLEGEVDAICDGKDVLLPGIMEHIEH
AGVHSGDSMAVYPPQSFSQAVKDQIVDATKKLAVALKCVGIMNIQFIIHKGEAYVLEVN
PRASRTVPFLSKITGVEMAQVATRVILGQSLQEQGFESGLYPESDTIHVKAPVFSFSKL
ADVDSFLGPEMKSTGEVMGSDHTFEKALYKAFSGAKMTLPDNGTILLTIEDQDKDEVL
PLAKRFSKIGYRILATSGTASFLREKGLHVLTVSKIDQADADHEGILEVLMAGKVDLVINT
MGHDQEVASDGFVIRQTAIEHNVPLVTSLNTADALLKALENRSFATDAL
>Dihydroorotate dehydrogenase, catalytic subunit (EC 1.3.3.1)
MRLAVNLPGLNLKNPIMPASGTFGFGDVPSAKKFDLNQLGALVIKTATVEPRVGNPDP
KIAVIDNGVLNAVGLQNPGIKAVVNEKLPALKAKYPNLPIIGSIGGSTVADYVQVAKELSD
SQMVSALELNISCPNVHEGGMAFGTSPEVARELTEKVKAVATVPVYVKLSPNVTDIVAI
AKAVEAGGADGLSMINTVLGMHIDVKTRKPVLGNVMGGFSGHSIKPIAIRMIYQVAQAV
DIPIIGMGGVRSCDDIIEMYLAGASAVEIGTAHFYDPLICPHLIEKLPAKLDELGISSLEEL
RAQVRKELKNEN
>Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23)
MKIKRPLIALDFPDGQAALTFLKQFPKDEQLFVKVGMELFYSAGSDLVLQLKDLGHDVF
LDLKCHDIPHTVESAMRVLGKLHVDLTDVHAAGGLQMMQAAKEGLLDGAAGYKVPKL
VAITQLTSSDEALVKNEQLSAVSLNESVLNYARLTKKAGLDGVVCSAHEAAAIKAATSP
DFLRVTPGIRLADNSKDDQKRVMTPDQAANNDASAIVVGRAITKAADPVAAYDLVKKL
WEGEQ
>Orotate phosphoribosyltransferase (EC 2.4.2.10)
METIAKDLLDIKAVKLSPQAPFTWASGIKSPIYCDNRLTISYPAIRKAIAKGIAALIKEKWP
QAEVIAGTATAGIPHAAWIAAELDLPMVYVRSKPKDHGRGKQIEGVLPAGAKTILIDDLI

STGGSVLKAVAAARNEGADVIGVAGIFSYQLAAATENFAKEKLEFVTLTNYGELLDAAIK
ADYVSADQRSVLEQWRQDPAHWTGA
>ATP/GTP-binding protein, SA1392 homolog
MTQKEVNLYIISDSLGETAFNLAQAALVQFPDAKFQVEKFPLIRTSSILQGILNKAKRNQ
AFIVHTFVEPELSALANDFCKENNLACIDAISGLVNTMANLTGLEPLHKAGINHNTDTEY
FNRIEAIEFAVNYDDGKDPSGFTKADIVLLGISRTSKTPLSLYLANRGYKVANLPLVPKT
KIPDEIYQVDKNRIFGLTNDVEVLNDIRRARMISYGLNPDTAYSSVDNIKSELDFATDLY
RKLGCLQINVAHKSIEETATLIIESLAQNKIDAIGDNQQV
>hypothetical protein
MLTENDGILVYIPQKNDNPWTSGEYDYYNFKKDLEEVENSGISYSVGKERLV
>hypothetical protein
MFPNQGDLIMIGTEPHKGREYGGHNKQLGNIRQIQNTIS
>hypothetical protein
MGMLITSTNRQDPNFYRDFFDFASKISGKIILWQIPTYDFSARHGEIIGSVPASLLDDLLK
RAHYILSR
>Transamidase GatB domain protein
MSLLESLQKDMVAAMKAKDKAKLSTVRMLKAAVTNEQINVGHDLTSDEEVSVLSRELK
QRKDSLEEFKAAGRDQAVADLEAEIAVVESYLPEQLSVDEVKAVVEETIKAVGATSKAD
FGKVMGVLMPKLKGKADGKLVNATVKELLQ
>Phosphate starvation-inducible protein PhoH, predicted ATPase
MTVEKLTEKEFQLANPNLQLGLFGINDSNLRLLEEGNDVLIKPFGDRIKIEGKADDVATV
IEIMRQLEALLMQGIHLSAADIVSAQKMAERGTLEYFQDLYTDTLLKDNRGRAIRVKNY
GQRQYIQMIKNNDLTFGIGPAGTGKTYLAVVMAVAALKRHEVDRLILTRPAVEAGESLG
FLPGDLKEKVDPYLRPIYDALNSILGAEHTERLIERGVIEIAPLAYMRGRTLDKAFVILDE
AQNTTRAQMKMFLTRLGFGSKMVVNGDKTQIDLPKGHSQSGLVDAEKTLQGINHIGFV
HFDAADVVRHPVVSAIINAYEQQKQVDEKKG
>Metal-dependent hydrolase YbeY, involved in rRNA and/or ribosome maturation and assembly
MDLEVYDETQKVSPEMIQLIKDVLDFAGNYISLPANTEMSVTLMDNERIHEINKEYRGV
DKPTDVISFAIEEEDPDELPIILPEDSDFEIPKNIGDIMVSMDKVYEQADYLGHSPERELG
FLVVHGFLHLNGYDHMKAEDEKVMFGLQREILDAYGLKR
>Diacylglycerol kinase (EC 2.7.1.107)
MASNGSPNNWKNRHFYQALKFALVGLVTAYQEERNVRFHTLAVILVGMLGWYFKVSG
SEWLWLVLAVFAVLSSEIWNSALENVVDLTTGGKRDPLAKKAKDMAAGAVLLAAIFALI
VAAVIFLPKVMMLFY
>GTP-binding protein Era
MTDTFHSGFVAILGRPNVGKSTFLNQVIGQKIAIMSDKAQTTRNKIQGIYTDDQAQIVFID
TPGIHKPHSRLGDFMVESALSTLNEVDAILFMVNATQKRGRGDDFIIERLKNVKKPIYLVI
NKIDQIHPDKLLEIMDDYRQTLDYAEVFPISALQGNNCPELVASLIETLPVGPQFYPADM
VTDHPERFIAGELIREKVLELTREEVPHSVAVVVDRISRSDEEKVQVQATIIVERNSQKG
IIIGKGGKMLKNIGVKARKDIELMLGDKVYLELWVKVMPNWKDRQIDLRSLGYKQDDY
>DNA recombination and repair protein RecO
MALHQGEEFRGIVVSRKNYRERDMLVKILTDRFGFKTFFVRGVAKRGFSMGAAILPYA
YGTYLGSISDEGLSFITNPKEVEQFKQLSQDIELNAYAAYILGLAEQAFMGQYSPSWYL
NVYQALSLIDQGFDPAVITNIMEVQFLACFGVMPNLRTCVICQQATGAFDYSESYGGLL

CEKHWYLDPHRLHLDQRTIYFLRLFSAVDLSKINSIELEAKTKVNLRKTLDEIYRQQVGV
YVKAKRFLDQMAKWHL
>ClpB protein
MNNEQLTQAVQDALAQAQGIAQVRKHQEIDIAHLFKFLIQPGEMARDIYHEAGVDVAKL
EAEIDRQLDQIAVVEGNVSYGQNLSSNLYALLQVAEKLRQEFGDDFIAVDTLLVALMTL
KYQPLSQYLQKRGLDLAKIKALVLKLRGGEKVTSKNQEEGYKALEKYGVDLVKQVRSG
KQDPIIGRDSEIRDVIRILSRKTKNNPVLIGEPGVGKTAIVEGLAQRIVRKDVPENLRDKTI
FSLDMGSLIAGAKYRGEFEERLKAVLKEVSKSQGQIILFIDEIHNIVGAGKTEGSMDAGN
LLKPMLARGELHLIGATTLDEYRQYVEKDKALERRFQKVIVQEPSVEDTISILRGLKERY
EIHHGVKIHDNALVAAATLANRYITDRFLPDKAIDLVDEACAEIRVEMNSVPNELDQANR
HLLRLQVEESALKKEDDPASKKRLEEMQAELAQTKEEVQKLKLRWETEKKDIQAISDK
KEELDKAHHDLEAATNEADYEEVARLQNAVIPALQKELQTLETANQAEGKLVEESVTQ
NEIASVISRLTGIPVAKLVESERSKLLNLPASLHQRVIGQDEAVDAVANAVLRARAGLQD
PSKPLGSFMFLGPTGVGKTELAKALAENLFDSEKHMVRIDMSEYMEKYSVSRLVGAAP
GYVGYEEGGQLTEAVRRNPYTIVLLDEVEKAHPDVFNILLQVLDDGRLTDSQGRTVDF
KNTILIMTSNLGSDILLDRANEEAKVDDQTKQAVMKLARQHFKPEFLNRIDDIIMFTPLGL
NAIEEIVEKFITQLSKRLANQDISLEISDPAKLWIAKTGYDPAYGARPLERFITNHVETPLA
KEIIAGKVMPHSKVTITLGQDNNLIFATETTK
>hypothetical protein
MKGNQHIIVNCAIASIGMVALHKYNADAISPQLSIIAPLAGIMLGSILPDIDIENSKASQIFG
TSMAVSLRKRAFENYYGKKYEDDKYKNIETYIEKDIQQATILAHEDPNTSTKVIDVDFTNI
AHRTYTHGIYGILTIGIISLIIAGLTPFAGMGPSVIMTRLAVGLFMGYILHLVEDYWTSNSI
DWLYPFRNKDSFVLKAVYNEQNEPVIIHMFRPKRVPTKNPNYWYAHPSGKYYNADSD
NTYLFVFICFIIAAVINFVYFKQFYF
>hypothetical protein
MGEQRYKFSILTNPNLLSYFRLQTPLPEELEERFIAKGKVLGLVGSDSNIAYTDNEDIHR
TIKVGGNIVKV
>hypothetical protein
MQTAPDNEVLVLGASSANQAVADLQDTVVSAWSRQDANEVLPSGTIISGANTVSLGSA
SQQFADLSAANTKVSSAVATYGKTSSQATQARTMLKQAQQVEVSKGIAPYVVRSAEK
VNAVNDALVKSEFVLRHGKWSLAEQESISKLDLDSFDMKLIEGMSLERDKKTTIMVFFL
FLP
>Mobile element protein
MESFWSHFKDEYYYDKTFATYEDLVAGINDYVTYYNTKRYQWKLNSLTSVEYRNQVA
>transcriptional modulator of MazE/toxin, MazF
MIEVHQGDIIVVDAESYAGREMSGHNVVEGNIRRHFLVVSRYEYNRRSNLVCGLAITHK
HVDSPFRFPIVDFESGTNGDALLLQLLTYDFVARNGKVIGHIHDEGQLNRIIKQVSNIFS
KEIG
>Transcriptional regulator of rhamnose utilization, AraC family
MINNKLLNQLLALTPLEEQQRRNQTFIEDLPQAAFDSKEKSSRTLTNYFFKNQDIYLSKH
NRFAAYPNHKHTFFEMNYVLRGQAKEVIAGKTVTLNQGDLILLDIGTSHSIQALGENDLL
INILFQNNNLSVDLLKELKGQHNVLVNFLLAKASSKPDYIVFSHQANPKISQLLDEIITEYY
EQAEFSELVIKSYLNILISQLIRSLSIKPSPKESAQKKLALQLLADIASDYQTISLTQLAKRY
SYNRNYLSNLFKAEIGQTFSAVLTEQRLMQAHKLLTKTNLPVNTIIEQVGISNRSFFYKK
YLAYYHVSPKQSRKQASAIFD
>Rhamnulokinase (EC 2.7.1.5)

MNKHYIAVDIGASSGRLMLANLEAGKMTLTEIHRFKNGFKQVAGHDRWDIDHLGAEIIA
GLAKVKQAGIDEVTLGIDTWAVDYVLVDEAGYKLADPISYRDCRTKDAISELT
>hypothetical protein
MVTKKQDETVTDTQSVQVDDQAQLQANFLRAVAKLPTVVVDTEEKARAFFEDDGGY
>hypothetical protein
MVSNAYKYKKLWVIDDYKMVLKDGLREFKFKNSRLKPAGHTERTKGSVALFRTKADLA
FGRGFIATSLEAIGDNADHAPSHWTANPFRYLTYKEGTQSVISHTEDNLRQINQIFIDIDT
KENDINVLVLDFNELGIIPTYIIETPRGYQVFWVLETPGYVNSTHKVVSVLKNVRANLVK
FAQEQGYQVDSGANCFGVSRFPTEDNVIFNIRTEYKIADLIAVSKKLSQKFGQNTMALH
SMKVKASQNASNKRQIDSEWYQLLFTQGASILGRKGCYGRDNAIFTTALANYSSGSPY
EVALGKMLEWNCRLDNPLTDQVVEQKVKSAYSGNYQAAAEEYAQNLIETYISNEAGSL
FSRTPAYWHKFKKARAERKYSHLKEWRDDLIAYIKQKTTAASPYIKLSTSELVDALGIPE
RSLARLRSALVEDTTLMIIARPGVATTFALINYVAVYLLTEMRAKKLALKQLLTESIKRNPI
SSTVSAFLKKHN
>hypothetical protein
MISNLCQLPLHISMLMVQEFQNRLKVNSVCVSETYIVVKGIYISGKGHGIKNHLFLASKQ
RASEIADGTYCLTFAKTGAKVSSYPAIIQVTDLVQDGTSVKRIITQYGFSFLPPIVEVLEE
DKLFLKQLSYLQSRWESYEHLVAETYLDNVVEKKCYLQEAVVVMSILDFSYREIAGALG
ISKSTVQRYLKQESEYRAQFGVLAPYLKLADHKLISNQKMTIKELYMILKANGYEGSYP
RLAVHVRHLKEVLAESGLIGIK
>hypothetical protein
MNFFEWITDMFSIRMVRIFAWIVLFSVVAVPLVRRHSQYQGQALSEVEYTLHKPQDRN
FNLVFYKRGCPYCQAAEKAVISEAKRNKNKTYFINLDKARGLAVSEQYHVHKAYTIIKIR
GKKVTRYYTAKVVDGKIVANQAEISQAFK
>hypothetical protein
MATTQTYGNGTGKIDLTKQVYDAHSQTMLLQFVTTDQTNSSGMGINAKNLHWKLFVK
HKSPQIKMTVLPVVDNKITVLINNLASNYEALALDIKNTTPNVDDANVELSQSSSSMSSA
KTKKKTSTLPGDAVMFVVANDSKHLSHAKIANTKREQLAKQVIRQEIAQQNKEKQRLQ
KAIQILKKGIKANKESLKELDSQSQYLTGQDQADNYDSVRELTNDITDKQSQINNTKQNI
KVVNQRLRLLNKKLEAVANNQFDYGSKTKTYKLD
>hypothetical protein
MENTKQLETKIEAAERWLQKLKDKLDKQYAELGRLYLKHQEMTTGRSLELTEVIADLKD
KLKQGEDDKTQAAFFEGEENKAAEHSEV
>hypothetical protein
MHVEIEVGDDTLVLQNNLEYIVDCKNDKELYEEIKGMFIDTADWFKKNNRLDDLVNANE
VFAKREEVFWGE
>hypothetical protein
MQMKCLLNGKKFFGVSKAMSAQAKQNQGSTNNPANKGINKNVDAIRPNKK
>DNA repair protein RadC
MKDHTALKDYELIYELLASLQTRAQAESNLKVMEANNSYNEINLEPTVADLIYKPWLLEK
LDLTSRQARALNLIRECVDRISKSKLTVGLQITSSSELGEDLRNRLKNKPQEELWGLYL
NNKNVVIREECISRGTLNTSEAHPRDIYRRAIECNAARIILAHNHPSGDVSPSQQDFAMT
ARVKQVGELVGIELIDHLIVGGDCYYSFAEKSQLN
>hypothetical protein
MNSKLKAKAKKELRFFDGLIIGVGLGYLSSVACFLGLGIINGLITFGAEKILSFFFLGWAV
QYPLKFLVAFLVMWCLAEALYFIGIVTNRHRKFLEELVKSNQ

>hypothetical protein
MLSESEKERRWQAAGLENDTIFGLVMGRKDICKLFLESAIPDVDFADLEITTQKEIVNHY
QSKTIRLDVLAVDEQGNHYDIEIQVKNEHNVPQRSKFYHSMMINRMLEQGENYQKVKA
TYVIFVCMFNVNIEEALSHFEMINRNNLREKLNDGTHTILLNVNADSSKLNENLQGLVQV
LRRESVNQSQLGQELSKALKEVKGDMLVKNQLGYSSDDMVLYGRQENIQMMIADKEE
LGLSEKTIVRFIMKTLYLTEKEAEELYDEMLEDLEKQKQN
>hypothetical protein
MALQSEFLLRGISKRQKELIKKSTKKLGYTSQSEFILATLEKEFMGKFAEENGGRWGEK
LSELVKNSELLLSNSEMILIQEDKILEKQRNLKLLYENWLNELLDAEEGEEND
>hypothetical protein
MTKRIDVKGLTEEEVQSLKYLAKKLHAGSMNQLILSQLRTAIRIGGLDISDTKREKNFQII
ARSQLKILRYFEEKRKLNSEILEELKQTQEVLENWIGLVAEQAVKNGEEK
>hypothetical protein
MRDYFTKIDFKTQLGLKNDAESSRSNSQLLEIDLKGKKQIEVAKHSNFTQTEFEKTKHS
WIKFSGVTVTGKLKEVNLVVTPKSLNNDKNSEKELKDIRPLGWHDGQKELVKIPIVSKS
LINFTLDKNRVFTGTKETAETLNQVMSKVYESVLTSGQP
>hypothetical protein
MVRINGSTDRRITLNPKSLQNLEYLAKYVESQNEDSLHKMILEGNFEWGVSTSSLIRFAI
NFTHEMLVNAPEQDYQKRLEELRTKGADVGNNQKTATSLEYKDLKPLFNNNELQTYLL
LSIYQMIMDLNWVIGELTSYHGEQEDNTQRELIGKISQLIQMDTARGQTIKSTRQNKKK
HRNNQ
>hypothetical protein
MRYQGLELMKKLQSKANRACLLIDVNRDLGTFDEETWSFIEPGEQLVNPITHRAYTKK
GQALIEELRAEYLKFLDTMRTEFKTSKQADDLEIVRYLSGKGLKPNLVPQV
>hypothetical protein
MQIPPTAETLSKQIDDLVSAVIDNPRYRSQTIEESSRTTEKKANRHEHKQPDRFKLQKE
LTSYLVKRHFGLVNEVEATKAADWAGRLNKLSDTKLTTTLSDSRKATNNMIAVITNRTK
NVEVTLDYVQNYSFGRSM
>DNA topoisomerase III (EC 5.99.1.2)
MYTLVLAEKPDQAREYAKALSKDIEKQKSGTLVIRNSQYYPGEIHIVSARGHIMDYDLKP
DKRWSFAKLPIVNPDFTFRVADSTNAKYRFKQIKQEANRANLIIIGTDADREGERIAYSIL
SKLPMAIRKKQRNAYGLIQ
>DNA topoisomerase III (EC 5.99.1.2)
MQTPAVKLVYDNDLAIKNFVSRPFWKITGLDEANQVTFKSKALGKIMDKAKAEEILAMIS
PQAQVVEVEHEDKQVSSDKLFNLTKLQAYGGKHWKKSGKQILDLCQSLYQKKYLTYP
RTDCQYITNLEFAYLRENLATYQRIVNLSFNAANLQARSKYVNDKKVQEHFAIIPTDVVP
NLTELTADERLVYETVVKRTCLMFAADYKYKLTRVIIKAGGVLLGAEGKQVVDLGFKQY
LSESNSEPNQLLPNYQVEQQLQIKPQLTEDKTRVPSRITEAKLIGTLFPKYSLGTPATRA
GIIERILEKKYVTKNNKGELYPTDKAEVLMNYLKANSVADPNTTKIWEEGLKLIGQGKLD
GKLFVDKIQEEIVTQVNEIKAKYGK
>hypothetical protein
MREDNRKKLTHLIQLAKQYEITTYDSFLKKSQTTPKLKPELNQLRENVVVGFSLYLYFE
KLLEKGETR
>hypothetical protein

MKLSGIVKEVKLVKLGEIDAFLVGLEIEDKVIHFPIATKYLANVKLEEENEFEFNCYPGGY
EPTLNAKPAENQVEAKAEPKLIELTDDEIKAMYGIDSETHEPKVLVPKIPINLVQLTDQDL
VKQAKNQSNTKKRANSKKHQPILI
>hypothetical protein
MVQLENKLANEGSEAVDTKQNQTESTTNTDKEGENENFEFPKHKNNQEPVTTAEDVE
EADEDPDEDPDPMDYNSVF
>ATP-dependent Clp protease, ATP-binding subunit ClpC
MLAEFIPAEDLLLDEQTELTTEESTTPLLDKYTTNLSAKVAANPASYEAFGRDSEIEAVTI
ALLRKNKNSPLLVGEAGTGKTAIVEGLALRIFQGQVPPKLQALTVRTLELSTLMNENED
GIFITKFRGLIDELVATKGENLLFIDEIHTLVGTGSADGNALDAGNVLKPALARGEIQMIG
ATTIDEYHESIETDKALERRFQQIEVKEPTANQAQVILRGVAPNYAHYHQVKISEQALNE
AVNLSIRYLPDRYLPDKALDLLDEACTIVSAQNKAVVTEESIAEVLKRRTGIPVTTILKDE
ASRLTNIESKLSRRVKGQAPAIQAVANAVTISKAGLQYEDKPLATFLFLGTIGVGKTELA
KALAEVVFDSEHAMIRLDMSEYSQPGSSLKLIGTRKEKGVLTEAVKHNPYSIVLFDELE
KGDREVHDLLLQILDDGRLTDGTGRLINFKNTIIIMTTNIGAEKIRMNAELKGEVDQLTER
EYQQFIEAMEIELQSEFRPEFINRIEYKIVFKMLSEPVIREIAIKDLEILNNRMKKQGAYLV
YDEDLLDYLTFNGTDVNNGARPLMRLINRKVLAPIATELLKLPKLNSNQQLYRVKVVVD
GKKPNGVDELLDHRKISIQIQSVPNNSVPNK
>hypothetical protein
MLNQLRLAYFTGLAAADPFTAAKGLFDGWAGKFRLVATGVFVLVLVISGIVYSAGGREL
KGAAKKKMGDTIIGVIVVSGGAVFVALLLAFVQQNGLN
>similar to Golgin subfamily A member 2 (Cis-Golgi matrix protein GM130)
MAWIKRKHRASMALEDETFLLLEWGKDFNDKSEEVLKVIKKLDQAQEENKRQTLIEKTL
PYVTFGNAEALFSNLKADLIDNGDVEQIKLARLQVCYREQGEKEVTTKNEAFIKPFIIDH
NYNGILKALIEEVLNSKRFNDIEYEDKRDYFDKIYFAYSSSTGISSELLPIFPSEAEVKNA
EGKITVEFPEKVLAETPKAPVEDHQFQAEIPVVEPTEVKPDAKTQSTASENNELSSAIEA
ASEVDIPSLDLDEVEQQIAARKEEIANTDNVVAERESNRSRTTSRQGTTTVTLPPLRAN
ETEVNLDFTPPKVKSDFIAQKPTIKFPRFEEARWSKAFQAHEAEYVQWKLDQKRSEFN
RELAQKEEQNQRLANQAADQQLQSFEQVELEALNHKYEKLDHRADLKQSIVVQVRKQ
QDAKLEKALTELNQAEKQQLTALKNQYEHEVAQTKSNYDKKRYETATSHDQATLKLAQ
DEYAKAYYDASAKLQADLNEDLTNLAQNKLAKQSAINQQLEAVGQTVGQTYYKQYKA
ELSKFEAQFQAEHRLALKRWLSNTQTREQLAQSSTDAKKMNAQIVQLQQERDNLNQV
NSQLRNAVLDEKKKAVDEKINLLSNLSSAPTQTLTEQKTISPNSEATQTTESPKKSNSS
LKKVLLFTASTLVFSGAGVAGYAYHQQTQAKLANLNRQVVKVQQINDNLSKNQTQLEK
EKSKAQSKVKKLTKEVTKLQSSEKQAQASASSATSSSAINE
>FIG01115343: hypothetical protein
MATIWAILIMLCKLIFFIPLKIVELLLFSKWTCLSIIAYLVARLIYFCYQTKPSLKTLFNLAKV
EILKLGRLLKAVWKHILEIGWQGNLSVGIALLSLTVGNLFFQLIYKLPFVNLARKRFQKES
QPFLRFRTWRSWLVMGLVGSLGALMLTSWLVCLLRGVIPLIYNLGKASFDIRSLLLSNL
FKFGMFNQAPILAWPIFLTLEVVVWKSAWINWEQYRDYNHNESGDDRFAKLSELKRQ
YKRIPDKAKTYPGHSGVPVIHVNKNNLPGWALNSQMKFRARWFSKWFLWSEYILGLN
KVPAGYYYLEDDTINALITGMTRSGKGEGLVNPAIDIASRAEEKSSLILGDAKGELYQAS
YKILRKRGYNVQVLSYMDMDWSMSYNPLALAIDAAKHGYYEKTQAYVNSVAEAIYRKS
LGKGAGNEKYWEDTSIALFNAITMALIDRANETSQKATNPEIDAWDTITIRNVAKFLNDL
GSDSVIDLEKSTRNRIVMKSKLTDYFDELRQVNVIRFSKFREMADINFRASDFASEETK
GNVYSSMMSGLNLFLQDNIARLTSKNSLDLRSIGNPRRLSIRFKSSSINNQANRYAHKP

ALISFYSEANGPKQGKGRLGGGRLRSARKKYFVKESKVLLDGRGYLNFNLQPQLPKRF
YLEIELDGQTYKWVGNKIYQGTDEYTGEAILTGINLREITKPTSKPRALVDPAEIEFVYSE
KPIALFMVTPPDKPEYYDLVSLMIDQIFNANYELALASDSRKTTVRVQFILDEFANIPKIP
NMATKLSIGLGQNICFMMFVQNLEQIEDKYGKENTASIIGNCSLNILIKSTSAKTAEAYSK
ALGVKTITKREKGTNILNEANPHINTRNPEQRLLTANQLSKLESGEAVIIRGVKAQDKAG
RKVTPDPIFVHGKTELPYRYMFLADEFDQSTTVGDIPIESKHRDLDLNDVAVKADNAWE
KMRDWHNRLEKANDCELAGRHDKVA
>FIG01119425: hypothetical protein
MEFKSFTELAKKMAGDKSLLQNFDFNGLFNGDITGKPAELLQQFYHQWGNYLDPVSA
VMSVVYIVLGWIAKALFVIANSLENVFNSLFKLFGLFGYLGNDQTLIGKLYLGMQVVGVS
LFVLLLTVRVIFGLLGKNIKYKDTINNLLLVTVVVSVLPLLINVTGKQAQEYSLTKMVQSE
AKTVRGMEGGAGKDTVQTLAIQPFINNVVDLKYLALNNFDTSKFKLDNAGYLKSDSGK
ASYLSTDPASINYPTRLKFGSTLGVTNPKIAENWEKKNPDLKGLKGLFEHKINDDDSGI
DSMSEHRFFKGMNALEAVYLRYKVNWWALYAQLVTVIILLGLMSIKLVKSLFELVITGIIA
PIQGYSSVDSSKKFKELLTTMIGTLAGMFFEVVILRVTLEIMRDAPAMIAGNLTGFEQTIG
SIIIYLGTFFGAMQGVTIIERWLGVWCKHTATYGGYDGGQYFRWCR
>hypothetical protein
MAANTLGGAAKSAGQTAWGATKFAGAVATGLPSMARNVGSNVGQGLANNIGQQGVA
NTAKGAIFNATDALGNKLDKATNKVKGGLNEAYQQGKETAETTTKNPYPTAEDQLAGK
VISDEKQPIHQPDEELRAHNEQAITDNFQQEQAEVEREQQLAEAPQEQELGLSESPSQ
AASGLEDTETVGGLDTTENNAEAGGLAEAATGLKTAENNAETDQQSDLADEQDGMGL
TTSEKTMPYKQRTGPNTARYGTKQGF
>hypothetical protein
MNSEANSVTAITSEMAGPGVSGETMQGQAPIQAQAPIQAQAPTQGVRQQTITGGRVV
SESTVGVNYHPTEAYSPEVESYTPTPGLQTTATTAEVTAPEARTVVSTPQQQFSSEPS
SGYQSPYRQVIDPQASSEIAQTVPEKENGISPYRKVITPDTNLNNNHNHKS
>restriction enzyme BcgI alpha chain-like protein( EC:2.1.1.72 )
MSFLKPTNFNSFIEKVKTLKLTADEIEKLKNQREREIDASLTKLNNDIYQNEKGLGENDR
VYLVAASIIATLGVPGKVKPLTKEDLTSSLEENERDGDILIRKVNSFLSKKEIPKTKKDLIV
RTLKNTLTTENINNPVNGESQLKRVFSKIVDDLGIYYKIGLKTDFTGKLFNEMYSWLGFS
QDKLNDVVLTPSYVAKLLVKLARVNKDSYVWDFATGSAGLLVASMNEMLDDARNKITS
PDELCKKEAEIKAHQLLGLEILPSIYMLAVLNMIMMGDGSSNIINEDSLTQFDGKYGYGK
KDEQFPATAFVLNPPYSAEGNGMIFVKRALSMMTKGYASVIIQGSAGNGKAVEINKEIL
KSNRLLASIKMPSDLFVGKSSVQTYIYVFQVGEPHENDYMVQFIDFTNDGYSRSNRRK
SSNNLRDTDHAREKYQEVVDLVKYGKSKLHYLSESEYYEGHIDITKGNDWNQTAPIEL
KPTLNDFKQTIGNYLAWEITDLLTNQSSRDDKNQINDLLTQKLEEIEWGDYRLGDLFDIT
KVPSFNKNQLTEGNDFDYVTRTSSNQGILQSTGYINKENLNEENVWSLGLLQMDFFYR
KRKWYAGQFVRKIIPKFTVTENNAKFFSVILNQIKPILLPVLVRNVDETFKNQIIKLPTKNN
QIDFDFMDLYIAELEEERIALLYNYLK
>restriction enzyme BcgI alpha chain-like protein( EC:2.1.1.72 )
MVKKIQSVEPMVADKVNGWLRSYGLDYKLEQESLNSEIDHALLEYASKSGGKGGNRP
DSKILLQDKNTDYWPILIEYKGFKDKLEKLNKNGQIDNRNSKNEPKYSNIKSFAVNGAIH
YANAILHYTNYTNIISIGVTGYKDELGNLQTQLVFTLYRKIILV
>hypothetical protein
MTIKYDPDFISPYEHLDLPHAPVIFPGIKPGEFNFAAEWVVVGYNLPEQKVIIRGTEENW
QLQFTDWNTPEELEQNMRDFPLDLVEYNAEVVTKLEAEEITPQQATKLYAEPFFITPEF

FIEDEWLESHPSPAEEEIGWITDMKNFLKNVSAEDEARLDEIREQANADSDATNDDKQ
QKH
>hypothetical protein
MATTPNYDKYYKKYKNYGLRLNTERAVDQQMIDFIEKGKAEMGLNNFFRKIISEYMRV
HDGKVDAFVSLENRLRNLEVVTLKAKGESENDH
>hypothetical protein
MEKKRIIVAGAVASAMLVGVNTVKADNVSATPTTDTNNVGTVTQDTTKADAQKAVDQA
QAAKDSAQAQATSAKDSLDSAKAQEEAAQTTVDTAQDKVADAEKVKAEATPGVIDQT
KKDITKAESAVTDAQANQTTQEQTVADKQADVTKAQDAVNQAQADVNQTQAGVNQA
QDAVNQAQANLDGTNVEQVTNDAKQAQATVDQDKQDLAQANNEVTDATNKVADANN
EVSQAQTNVNTAQKAVNDAQAAVNTATNKLNDAKTQAANQAQVVKQAQNAYDSAKQ
AVDNFNTIVLPDGYSITYSNDEAYEVTKNNYQGMLINKYKHNDADKLVKVNPAHLTDEQ
QKELTLWVAGLLNPIRKGFTIKYSDGTTKEIESFETVTDKSLEYARNVAKYYDANAVNG
EINWDHDENALDKAEKELNNGSMHSESIGMIDSTITDMDSLKEQIYNNILSMLFNDAGS
NEGHAAQLLGYYPETEINGVTLSSAAKDYFGMSISSSPIKGWEWGNYIMLHFENYALD
KGFEDRDTDKTTVYTIPTVDALKATADKKQQELLAAQKTQTQLNSVVDANQADLNTKN
ATLTNAKNDLTAAQNNLNEKQASLATAQTNLQAAKQKVSDLETKLANDTKANEAAQQA
LSELNASVADKQAKLASAKQDLSAKQAVLTAKNEVLAQKQADLAKAQQAYDLAVKNLT
DAKNEVSQAQAKVKELQAKLAAYENADTILLSARTELKKAQEELASAQAKVASAQAEY
DVASDNLSKANDELRLAKDKLQVILDREELQKQLDELAKQKQDKQSQGSTSKTDNGV
VTDTKGNIVDNAGKVVATKVVKSNVVAPKSVMLGNARQLPQTGESQKANVTLAGAILT
AIGSFFGLVALGRRKGNDGSK
>hypothetical protein
MKKKILTVVKCCLIIFSFIATAFNTVTSIADSIRYHSLVINWLSLMSVAPLTFALYQEIDFIY
VWCNKTMSFFKNKTVNFKVSFYLQNNSDNFVNVINNNLNKVTKDCHLSFDNGSPQHF
SNDMYTKFYFVTNNNLKFSVAISNDNLPSSEIETNNFSITMDFQISYRDTSACWELCKK
FREKFTSRLVNENERIDLNINLSNSNQNPFYRLTVKYVSANKIDDFRLAFSEKESLKVTL
NKHSIYATSVDIQDLDRVVKEYIPLSSVN
>hypothetical protein
MIDDGFLLPGWLDPKKNESTSNTVRLMKWLYRRIKNMSKDVN
>hypothetical protein
MVNMTIIFFVTAMILVAILFYLVIYAFVMDLNLAKEVIVNAIKLAMSRRSRNEFIKEYHRLL
NTFLKNTATEIVQTKDKVLAGDLGIADIDEVKVFKKLLIKAITDKNYEFMSYEDVKDYLNE
ERVFSFMSQIYLNTRLYSNSQTMGKVNDKSLSALANLVFAEDIEGNRIDNFQKMYNAIA
>hypothetical protein
MLLLIYFLSILGAVLVLLTSVVKALNSFLASITVTLTELVNVQQAFTRYKNEQRKRKNLKK
KNRR
>hypothetical protein
MFKDIIKVIISIFQFLQALKELILAITHMLEELADLRHAVEKLKEAWNKWKRRK
>hypothetical protein
MFLKKWHRKRTKLLMLAGAGFILGSGLACFNQHVSADDVPQTINIGPTTYAVSYGLGY
REVNRGPHNVDVLREAGSNEYIFCIQWAKNSPSNMQLQRKWDADPRVQWLITNFFSG
NRYRSLSMGDEADYWLYQVAVHIIASPNDR
>hypothetical protein
MPRFRADIRQKAQDLVDEANRHNSQNDSQVVLNSASTSFNPGEISVGRDDIRDGQYK
KNFSFNTQNVSNAHVWLEGAPGGVSLSGKDGAEVNFNDVWNGTALQVNIPLSARAT

DNEYSFKVKANGTWNKQVRVAVIYVNSNNSVQNVAKNQIKAIDVADSAYSEMNVKVY
PAKARLDFTKKGTGFKLIGLDNGFTATSKADNNGRVNFENIPLGSRYHIEETEQPNKYY
VPNYVTDINDFTGFSPHWWVSTP
>hypothetical protein
MVEKATRGTLTHVVHWEAKLGWQHIDVEEAEKIYDGKTVRSSEQVLMCSLCHDYVTL
TRSSAKRPAYFKHKRMEQKSVKDCPDRSGSKEGISHSKVSLKDAKLPLRLEEHGSKY
SLKLGILPEYSNISLSDERLIITTDKSSIKQKISLESLDYDKVTYFELRDVAAKDYHLKFEN
EDINDYGLPHSALGVIDGAFFDEQTKRRIPFNGQVQLGKEYYFVISKSSSKYQDWKRL
RGAKHTTTKELGNLNRNILVFSIKANKLNFEVASFFFILGVNLVNKLDKVYPIWPETIRSQ
DRYFPNSNELILYNPKKYGIALQASRSANADIDENVFVPEPLISNDRRRALITSKNSYITD
YVFLWQDDTKLTKRKSVKNEVTVYDSDRNSDKYQFEIYTYNELPKKKLVIESKKQEILT
VILYEASIPVDIREGSKVRLRDDEINPGQEIRIYDGLVLLKRLKFADNDSDLDNKLYKQLL
SVQNSEKIKFKSNMAIWGIKLADYPKCRRWLQECKIKREIPRVAVKILQKEFRR
>hypothetical protein
MMILGKIVHQDGDESRKKIEFLAKVSEGRLKPTSGYRNLKMGFKSEVGEIRFVRLKGDV
YSAKDIQIPTIIIDTRNDIKRVLADLRYENTFSLKNVRWLIGRYQNSIFSGVLLTNQDIKKT
QAGFALNEGVVSLESYSITHDETIRILVDEDMMTFVSRLELPPSEGVYMPFSWEETFTK
ALAKIFSRKELMKIGAKTKESRRVIQTFINNLDRTLVVDYLANKLEVSTSEIMQKFEEFLR
NKQDELDYKDADSQLIIEIAKNSDLLRKRINSEIEKEWEQENKEKIAKNSEELEALRKKR
DDIESITKIKQEQLAEIRKEVAIKTKLSESLEQRLQTAKEDISDFRAQYLLLNGDSNNQPT
NSEEKSLVGYSAGEEISLNKKSVKVANGTKLIEKLADEFVANGIADDDNALILAGSLYAS
LQSKKPLLLIGPMGIEAINILSACLYKQVADELDCSVDYNSTLLESIKKSKNKMLVVKGFF
GSDWTSKIGEILTTDKAAFFVHPYKEELIMEARSIYYYLTPLLMDLVVENIVSVNYAQLTN
LKRSSTYKDIKVASSFIEDNLELSLHPIMKKNLQLLTEIEFSENEELLQKLELSLAYLTENN
EFLDAVDGMDLLKGIWYE
>hypothetical protein
MNEVVIRRKIRSDISNRLKIEELEDESFEQYEFRLVYSFLGKLLLANLWNQSEEDSEEGI
SKASLTSILSDTMVGYRALFPTMSSEGFLLDENDLIDKMLNDYLETGFIKKKSGKFSPVP
FEQATSEEVTFVRGASVKEKVNFSGLGTYYDANENDSSIFPKSAEQLFMLPEYTLKQL
YDYFNKQNFSESIPKEMLLSNSEFLVTWPTKANKWWDHNFLGKDKKLNLMRVGSAKG
KQLYYLFRGTNYAKGFQLSSKLNLTNEKGYYMIRLALLNERGMVPTIEYVDKDNYVEIK
SIFELPKRETAFLRVYSWPILNSESLLMDKGVFNACRVILERIGYIMKEVFQ
>DEAD/DEAH box helicase-like
MKYGASNIQQEMRNDLKGYIRTQYLGKTPILAENLEKELSQEGILYKEPYIESSPAYKTV
AGGFNHAKLLDWEKDFLSELADRNLGVFKQPFEHQVEALEAFEEGKDVFVSTGTGSG
KTECFLWPLLVKLAKEAKDGKSWSKRGIRTIVMYPMNALVSDQITRLRNIIGKDEFAEIF
SKLNSDTRRPQFGMYTGRTPYPGAKPKKQDDIDLAEKLKELITGNEQYTETLRKEGRIP
AKANLEEFINNLEKGKHITAADDAELITRFEMQETIPDILITNYSMLEYMLFRPREAKFWK
ETMAWLNEEKENKLLFVIDEAHMYKGSAGGEVAYLIRRFLHKLQIDRDKVQFILTTASM
PQHHEEEILDFACNLTAANNRESFSMIRGVKEELPKDKLNDLPINKLLEINLEAFNQEDD
VDILKQLNSLWGVDFENVEAAQNWLYENLYKYRPFNQLFTLCRGKAVSLGDIAKEIFPD
ENEDREITATNVLLEIATLAKNNKGDVLFPARMHMLFRGIKGVYACCNPHCTRATKGG
GITLGRLFVSQEIKRCPDCNSVVYELYSDRHCGALFYHGYIAKDVQSELDKEYLWHNP
GKITKGMTGTIKEVNLFIPYDGFNSRRIKGVAPCYLNIENGYVYFNEDYPGEKSSYLKLY
YDTKASNNVNDDVITFRKCPQCRKNLASGINNFDTRGNQAFYELVKSQFNAEPLVKKN
AQLPNQGRKVLLFSDSRQKAAKLALDMSETSQRAATRQLLALAIKYLNTESVSINLDNIY

GYYLKAVYEKGTRNILVDVDQKEYSKVVKKIDKGKTLDKKFKLDKASDFVKESILRLYCS
PYNNYYDTAISWLEPTEDSLEDFIDNLEDNDIEISSNYKNYLKEVLTIWMMLICTNYMAL
GSSISDTLRLKISEYTDKGLKESELENLPEIIRGALKKVGVREDRIEDFIDTLKDDYLSKS
HTGEWYLKLDAIKPCLDLEHTWYFCKKCLQVVPFNIDKKCPYCELGELHELSDKLLENL
DYWRKPIIAVLTSQNEKVSLIDTEEHTAQVSNKDVRKKYWSTSENYELRFQDILKDGED
SVDVLSSTTTMEVGIDIGSLVAIGLRNVPPTRENYQQRAGRAGRRGASLSTIVTFCENG
PHDALYFNDPVPMFEGDPRVPWIDVQSDKIIQRHLNLVILSEYLLAKGRSIDALKINDFID
DYYSDFLHYLDLWKVEESSILLNEKFDEVSFKDNLKKDLRTIKQKLQNHPEIYNIDKHNS
GRDVKKILDVVYEEGTIPSYSFPKNVVSAYLGDNEVDRSLDFAISEYAPGRTVVVDKKV
YQIGGLRSSANTDRSSTRKYLADSTHVKSVYSCDNCSWFGLEGDRVKAGDKCYFCGK
GTISKSGKPMIRPWGFAPKNGTQLSSSNVKEEVTFAQEPIYSTLPNNGKDIKVVNDYQ
HLKIANRKNQRIIMMNTGGNGRGFMLCKACGASMPGNDIKVLQNIQAPDKKSCSHPE
MINVDLGYDFITDMLVLEFEFDSKKINTTYNNIWYHRAGVSLAEAIRLAAGKILDIDFSEL
VTGCRYRNDKQGCFLDVYLYDDLLSGAGYSNSLSKIIGQVLAEAEKILKQCNCDDACY
NCLKHYRNQYVHGDLDRKAALELLAWGRDGEIATSIPQKQQLELLEPLHDILLERGLKL
DNLKINNKEVVIYPSMRLPEKDGKIYLSEYLMKYALPLAVQEINSNF
>hypothetical protein
MDSVKYQPSRHIMDFHLAGFAYCDGLDVIEELKLGQTVQLVREEGNPHDPDAVAILYK
DTKIGYVPAQYNGLFSTLLYYGHGDILEARIQMVNLESHPERQFRVVVKVTDARK
>hypothetical protein
MKKITGLYHEYLPNISGMAGWYYDYDTNLLDSKSKAAPFAKKLLLFHAETQEIFTVYEA
SEQQVISNFAVPEYYQGNLYFLVLEKATEQILIMSYEIVSRQVTEVARIAATGINFARLAF
YIAPVLLAVQDDLNHRVEIYYPQRLSLPLAEDESFECQEGEKFYFSKWLADESLARRNA
YLVKDAQGKTIAEGIGRITRFENGEFMMI
>FIG01114684: hypothetical protein
MDNLAKVLEDDEKFMALLKIIQSFELKDCWLCAGTIRNYIWNVLSGKEGFSDAHFSDVD
VIFFDKKLSYEQTLALEARVKRLYPGYRWEIKNQYYMHIHSPNTERYTSSRDAVSKFPE
KCTAIAVRLNSNHNLELYAPYGVGDLVNFKVTPTPYYASDAERRLVYNNRVKKKAWRK
FWPQLVVELI
>Transcriptional regulator, ArsR family
MDKDKQLAALKQGFDELADLFVALGDETRQHLIMTLFDLPCKQQGGARVGEITAKTHL
SRPAVSHHLKILRLAGVVGMRKEGTKNYYYLETEGTRWQQLASLSTDLVELLKYLNEG
KVK
>ThiJ/PfpI family protein
MNKVLVILTNVAHYGTSSEQTGLWLAEATDFVAKVQAAGFQVDYASAKGGAVPLDPR
SLKKIYRSQEVDAIYQSNDFQTRALRASLPIARVNPKDYRGIYFTGGHGVLWDFPNNAD
FEAVTRQIYQQGGYVMSV
>Bifunctional protein: zinc-containing alcohol dehydrogenase; quinone oxidoreductase (
NADPH:quinone reductase) (EC 1.1.1.-); Similar to arginate lyase
MQSFEVVKPGGKVVSVAGLPDAKFAKAYGLNFIWQGLFKLASHKITRASHKAHAEYEF
LFMRLAGLQLQRLAELVVTGRLQIRVAKEYPFEEIQAACNYVATGHADGKVIIKLTD
>hypothetical protein
MAKEKIYTKTALKREFGLTDKCLAYLPPADKIWYGRYKSQRYPAWSESLIQDFLAQSEV
KLLLVMATKRRAKRQQAAQKAVATKTNALLAELADIKVQVERLPKKKLKRLTQASLDD
WYMYREFNYHSYDEFAYHEATDWDIVNYIRHNLTNYDDELTRLANRVGRQEGFKLISR
KIYAEILKVYPEYRTTIEKQLQEHLSR

>Phage integrase
MVYNSQPLRSEDQINDVLFYLRRTRYPERDTFLFLLGINTGLRCSDLLNLKVGTLRHSK
RPEIVEIKTGKKKTLFLDAIYPKVEQYIKDKADDDWVFPSSQNHGAKPESVNNVYKLFK
RIDANLGNIGFTTHSMRRTFGLWYYRKTHDILYLMRLFNHSSEEITKRYIGITDEELGDS
LREFKIGF
>Serine hydroxymethyltransferase (EC 2.1.2.1)
MDFKDLDVELFDAIKLEEERQARNIELIASENIVSKNVLAAQGSVLTNKYAEGYPGRRY
YGGCQYIDIVENLAINRAKKLFGAKYVNVQPHSGSQANAAAYLSLVEAGDTIMGMDLA
AGGHLTHGSPVNFSGKTYNFVSYGVDPKTEMLDYDAILTLAEKHQPKLIVAGASAYSR
TIDFAKFRAIADQVGAKLMVDMAHIAGLVAVGLHPSPIPYADITTTTTHKTLRGPRGGMI
LTNDEQLAKKINSNVFPGTQGGPLEHVIAGKAAAFGEALAADFKTYGQNIIANTQAMAQ
VFAASQKARLVSGGSDNHLLLVDVRGFGLTGKDAEKLLDEVNITVNKNSIPFETLSPFK
TSGIRLGTPAITSRGFDQADCAKVAELIIQVLENPTDDQVLASVKNQVQALTDTHPIYAA
DSLR
>FIG00750646: hypothetical protein
MKIVVSSDNHLDVNRISPAEALAFQAQYLNQIKADYYLFAGDMFNDFLKTKSYLEKLND
LITGQVFFIAGNHDMLKNVTFSQLESPLSPCYLHNQAYDFPNSAWRLIANNGWYDYSF
SPQLDVEEVARWKKAYWVDAGVNQPLSDLERINLVITQVQEQLTAAKTADKRVIFLTHF
VPHHDLLWARPAHFSKPRYERVYEMVNAFLGSQRLADLLETYPNVYYTFYGHVHGRH
PALTHGQLTYFNQAVGVRRRHEWQAADFENQWLASLQEIKIN
>Peptide deformylase (EC 3.5.1.88)
MIREIIKDEAVLQQKSVPAGPGDLQIVQDLKDTLLAHKEHCVGMAANMIGFHKNIIIISLG
VFNLVMINPKINKKFKPYETEEGCLSLVGTRKTKRYHSIEVTYLDEQFKQQRQTFSDFP
AQIIQHELDHCQGIII
>Signal peptidase I (EC 3.4.21.89)
MSKFKEIMGWLWPIILGLLIALVIKTYFFSLVRVDGTSMYPNLQNNELVALFKQGKIKRG
TVIVFDAYGVDANAQPKDKYVKRVIGLPGDKIEYKNDGKLYVNGVRQSQSYITKKQQV
AGTLTLQANEAKGVTLGSGQIFTVPKGKYFVLGDNRAASNDSRYYGFVPTNKILGTVK
TGFWNDKAQVINHFVPAN
>Xylose isomerase domain protein TIM barrel
MQIPLGLKASADLRQVKDRLQYKPEVFEFFTAESDFTNDGLKRLQVAIEEVKAAGIKQIV
LHHPMRYRDTFTELIAPQEQCRELYYFIDKSTNDLLQLAFDYNLQVLVHGSYSRQTQTF
ISMYPNLKAAQAAAFKRLDYFKQLGQAHIMFENSISPIFYYGEQDSDEEILAKNYRLAFD
TSHCFIKVKANNEKLLASLTRLKDQIVHYHLVDSYGQTHDSLTLGQGLIDWRRVLPRLN
PAATSIYEINLKNQLDASEQVKSHEYLLNLL
>DegV family protein
MKIAVVTDSTAYLPQETVEKYNIKVVPIPFIIDGQVYNEGEDISTAEFYEKLRTSETFPST
SQPAVGELIELYQSLGKEGYDTVISIHLASTISGLHQTLVNLADELTDIKLVPYDSQITVM
LMGNLVIEAARLAQAGASLDEILGRLDALRATTNEYFIVDDLQNLVRGGRLSNASAFIGN
VLKIKPILTFDNDSHKIVAFEKVRSTKRALKRVEDLFAEEMAHINYPVRIIVIHANDLPAAE
KWQAKLKAQYPDCQIDISYFGPVIGTHLGEKALALAWMRDIDKA
>Fibronectin-binding protein
MTMSYDGLFTHAMVQELRAKLVEGRVMKISQPYPNEVILTIRSQRKNYPLLLSANPSYA
RAQITAIPYTNPPVPTNFTMMLRKYLEGARLVAIEQLECDRVINFAFSNRNELGDKLPLL
LSIEIMGRHSNVILINQAENKIIDTIKHVGLDQNRYRTLLPGSTYRLPPKQAKQNPLTASP
ADFEQLVATYPNREVLAKELQANFQGLALDSALYLADYFHQAGVSVASLAALWAKLAQ

PTPSLLTEPKLNFSAFDYPSTTASTRFASLSELLDEFYGTKATRDRVKQQGARLILVVK
NNYKKNKKKLGKLKRELAATENADDYRLKGELLTTYLYQVEPKQKTVTLPNYYANEQP
VTISLKPQLSPAQNAQSYFKRYTKLKNAVGYINEQISLTEQELAYLDSIQAQIELAEPKDL
PDIALELESGGYLKKPKQKAKRKPKLSQPAKFKASDGSEILVGKNNLQNDRLSLKTAAK
NYWWLHVKDIPGSHVIIANPHPSEETLLEAANLAAYYSKAQGSANVAVDYVQVKHLRK
PNGAKPGFVIYEGQQTLFVTPDKELVTRLKAE
>Topoisomerase IV subunit A (EC 5.99.1.-)
MTKGSDFLAHNNIQELSLEAVMGERFGRYSKYIIQERALPDIRDGLKPVQRRILFAMNE
DGNTFDKPFRKSAKAVGNVMGNFHPHGDSSIYEAMVRLSQNWKLRDPLIEMHGNNG
SMDGDPPAAMRYTEARLAEIAGEMLKDIEKDTVEMVLNFDDTEYEPTVLPARFPNLLV
NGATGISAGYATEIPTHNLGETVAATIHLLEHPQASLDDLMQFVKGPDFPTGGILQGLD
GIKQAYQTGRGRSILRAKTEIEAIRGGKEQIVITEIPYEVNKSLLVKRIDELRLLKKVEGLA
EVRDESDREGLRIVVELKKNANAQGILNYLFKNTDLQISYNFNMVAINNKRPEHVGLKTI
LEAYIKHQREVTSKRTQYDLAKAKKRQHIVAGLIKALSILDQVIKTIRASKNKGNAKANLI
ANFAFSEAQAEAIVSLQLYRLTNTDVTALEKEAADLAQKIARYEKILASQKELDKVIKKEL
TALAKKYASPRLTEIQAEIETIQVDTEVTVAQEEVMVLVSHDGYVKRSSLRSFNASGDG
ENGLKDEDYPVLQQTANTLDHLLIFTNKGHLIYRPVHEISEERWKDTGEHLSQTVGLGP
DEIVLKAYIFNTLEAEGLFVFATKNGFIKQTPLSDLKPGRTYKSRAASFMKLKQADDELIL
IAYGSQAELANKEVFCVTYGAYGLRYDLSEVPVNGAKAAGVKCMDLRDDYLTSVMLT
NEKQAVALVTNRGAFKKMKVSEVAKTTRARRGVQIMRELKRNPHRIVLARIITEDKVQLI
TDHQEEITISYAAHSFGDRYSNGAYILDIDSQGLPLTWQPLSEAKEE
>Topoisomerase IV subunit B (EC 5.99.1.-)
MAKKPNYDDDSIQVLKGLEAVRKRPGMYIGSTDARGLHHLVYEIVDNAVDEALSGYGK
QIDVTIHKDNSITVVDYGRGMPVGMHSLGIPTVEVIFTVLHAGGKFGQGDYKTSGGLH
GVGASVVNALSSKLTVNTVRDGIEYEEDFVDGGKPVGTLRKLGKTKRQSGTTVTFTPD
ATIFSTTKFNYETLAERLREAAFLLKGVKITLTDERTDTSDTFLFAEGIKEFVSYLNEDKD
TLGAVMYFEGSKDGIEVEVAGQYNDGYSETIMSFVNNVRTKDGGTHEAGMKSAWTKA
FNEYARKVNLLKDKDKNLEGSDVREGLAAVVSVRVPEELLQFEGQTKGKLGSPEARTI
VDSIVSEQLTYYLMENGEFAQELVRKALKAREAREAARKARDESRSGKKKRKVERNLS
GKLTPAQSKNAKKNELFLVEGDSAGGSAKQGRDRKFQAILPLRGKVLNTQKAKLQDIIK
NEEINTMIYTIGAGVGPEFKLEDANYDKVIIMTDADTDGAHIQTLLLTFFYKYMRPMINAG
RIYIALPPLYKLQKQVKKKAQIRYAWTDSELELAQKELGKGATLQRFKGLGEMNADQL
WETTMNPETRTLVRVKIDDEALAEKRVTTLMGDKVEPRRKWIEKNVQFTLEEDGSLLD
TAAAEGAHQENISVSSADLKQASLFDD
>Acyl-phosphate:glycerol-3-phosphate O-acyltransferase PlsY
MDSKIILMLVIGYLLGSIPSGVWIGKIFYGKDVRQFGSGNMGTTNTFRVLGKKAGIIVLLM
DMFKGTLAACLPALFHSNANVLLIGLAAIIGHVFPIFAGFKGGKAVATSVGVLLIYNPLFF
GISWAIFLITLYLTSMVSVASMLGFTLTAIISLFFGDWILIIVAFALTIFVFIRHWGNIERIKN
GTESMVPFGLGYSRRQKQVK
>LacX protein, plasmid
MAVFLENDHLVATIKEHGAELASLRSKETGLEYIWAGDAKYWGRQAPVLFPIVGRLKD
DQYQVAGKTYHLGQHGFARDRDFAIREQTATKVVLELKADEASKELYPFDFCLQLSYE
LTASGLVVGYRVDNPSQEPIYFGIGGHPAFKTPLTPKESFSDYELTFSADYRLPMIPLTN
GLTDIDSARLSAPTATVSRELFKDDALIYDLEQKPAKVSLTSKTSGHGVSLSVTDAKFM
GIWSTYPKEGQFVCLEPWWGLADTTDSNGDFTQKYAINRLDKAASFNHQYEITVF
>ATP-dependent hsl protease ATP-binding subunit HslU

MKATDKTPKQIVAELDQYVVGQAKAKKAIAVALRNRYRRMQLDPAMQEEVTPKNVLMI
GPTGVGKTEIARRLAKIVGAPFIKVEATKFTEVGYVGRDVESMVRDLAEVAYKQEEDRL
YKEVRSQAAQRADKRLIKVLVPAKKKEQKEQQPNLMQMMRDLQSGKLPDYLTQAPEE
QEEVTDAIRNERLSVSEQLKRGLLENTEVTIQVEEANQTFSPQNDMLGSMGIDLNETL
GALRPKKMVKRTMTVKEAREVFIKEESAKLVNSADLADGAIKRAENTGIIFIDEIDKITSK
SKQASGEVSREGVQRDILPIVEGSQVKTKYGLLKTDHILFIASGAFHESKPSDLIAELQG
RFPIRVELEELKAEDFYRILTEPTNALIKQYMALIGTDNIQVTFTKEAIQKIAEIAYKVNHET
ENIGARRLHTILEKLLEDLLYEGPDMQMGEITITEAYVTDKIGSIAGDKDLSRYIL
>ATP-dependent protease HslV (EC 3.4.25.-)
MTVSFHATTICAVRHNGHCAMAGDGQVTMGEKVIMKNTAHKVRRIYNDQVVVGFAGS
VADAFNLEERFEKKLNEYSGNLQRAAVELAQEWRSDQALQKLEALLIVMDKDTILLVSG
SGEVIEPDDNVLAIGSGGNYALAAAKALTKHASDMSAIEIAKAAINIAGDIDIFTNHNVIAE
EL
>Site-specific tyrosine recombinase
MEKEWLAEFKKYLLVERQYSAKTALAYEKDLQNFVTFLEESGGAASFATISNLDVHAY
MSYLYEQNYKTSSISRRISSLRTFYRFLMKNNYASNNPFAYVQLKRRPRSLPHFFYEQ
EMNALFAATQGDEPLTVRDNALLETLYATGMRVSECTGLTMEAIDFSMEAMLLHGKGS
KDRYVPFGRYCKQALERYFELARTPIMAKYHKDHPYVFVNHYGDPLTPAGITYILNSIIK
KSSLETTIHPHELRHTFATHLMSNGADLRAVQELLGHSSLSTTQIYTHVTPEHLQRDYR
KFFPRA
>Ribosome protection-type tetracycline resistance related proteins, group 2
MKEIVTGIIANVDAGKTTLSEALLYETGALRNLGRVDKGDAHLDTASLEKQRGITIFSHQ
TKLTTGNTSLNLLDTPGHLDFIEETEAVLAVLDYAIMVIAAPDKLSQQTKTLWKLLAHYQ
VPTFIFVNKIDASTESQASILARLKAELAEGCLAFSQPDLEEIALQDEVTLNQFLETGTVS
ASQVLDLIAKRQVFPVYFGAALKLRGIKDFLAGLDQWTKVNQAQSDLSLRVFKISHDTQ
GQRLTWVRMYGGSLQPKDVLFGEKLNQIRVYDGAKFELCPKLTTGQVAALTGLTKTY
PGQALRIENTQPVMKPVLSYQVDLKGHDIHEVLPALKLLEDEHPQIQVSWSSQLQRLA
VQVMGEIQLEVLQALLKERFNLEVGFDQGSILYQETITAPVEGVGHFEPLRHYAEAHIYL
EPNPNAGLSFASDCSLELLASNYQSQIMTALKQKEHKGVLIGAALTDVKLTLVGGKAHL
KHTEGGDFREATSRAVRQGLMELRQRGNCQLLEPWYDFVIEVNQDLIGKVMTDISRLA
GKLDSTKSKGNILVGKAPVATMRSYPTSLRSLSHGEGLIELSFAGYYPCHNQAEVLAAS
DYQPESDTDNTPQSVFCSHGAGHTVNWQAVPQTMHCPYYRWQVD
>tRNA:m(5)U-54 MTase gid
MVATVNVIGAGLAGSEAAWQIANRGVKVRLYEMRPVKQTPAHHTENFAELVCTNSLR
ANQLTNGAGLLKEEMRQLNSVVMQAADKHNVPAGGALAVDRDSFSKAITAAVKNHPN
VEVITEEVTSIPSGLTVVATGPLTSDLLAKEIVKFTGDDGLYFYDAAAPIVAKDSLDMDK
VYLKSRYDKGEAAYLNCPMTEEEFTAFHKELVNAEMAELHDFEDEKFFEGCMPIEEMA
SRGAKTMLFGPLKPVGLEDPKTGKEPFAVVQLRQDNAVGDLYNIVGFQTHLKWGEQK
RVFSMIPGLENARFVRYGVMHRNTYLRSPEMMTATYQTKARSDLFFAGQMTGVEGYV
ESAASGLYAGINAARLALGQEPVVFPTETMMGAMAHYITHASKKNFQPINANFGIVPRL
KQKIRDKRERNLQLSQRALTILDAFKAEKTL
>DNA topoisomerase I (EC 5.99.1.2)
MPTQSATKATKTKKKTSKRKVKKNLVIVESPSKAKTIGKYLGSRYKVMASVGHIRDLPK
STMGIDIENNYEPHYISIRGKGPVIKELRKEAKKAAHVYLASDPDREGEAIAWHLSHLLD
LDPQDKNRVAFNEITKETVKNAFKTPRAIDMDKVDAQQARRVLDRLVGYSISPLLWSK
VKKGLSAGRVQSIALNLIIKRQEEIKNFKPEEYWTIDAEFSKDKKQFAASFYGLAGKKKE

LGNNDDVKEILAPLDKNQPFTVTAVKKSERLRKPNQPFTTSTMQQTANNVLKFKTSKT
MMLAQQLYEGISLGKGGAVGLITYMRTDSTRISDQAKHQAASFIHNTYGEEYAALKPQ
AVKLSEGAQDAHEAIRPTSVLRTPAELEPYLTKDQLKLYSLIWSRFVASLMTPAVIDTVR
VDIEQNGRTYRANGSQMKFPGFTKVYQESKRKDNLLPALAEGDQVTLVANQPNQHFT
LPPAQYTEASLIRTLEENGVGRPSTYAPTLSVIQKRYYVKLVARKFEPTELGEIVNHIME
ECFPDILNVGFTAELENDLDSVEEGKRQWVKVIDEFYQPFAKEVSAADEKLAKVKIKDE
PAGIDCEICGAPMVIKLGRFGKFYACSRFPDCRNTKAIIKEIGVTCPQCKEGQVVERKS
KKNRTFYGCSRYPDCDFVSWDKPVGRDCPKCQHYLISKKTKQGMQVKCSNCDYEEAI
QK
>Rossmann fold nucleotide-binding protein Smf possibly involved in DNA uptake
MDLRIFLMLLHLCEGIGNANELRVAAFLRIYQRLPTKNELTQLLKLAPEKSHKLIAQLECA
ELVEQLRANLQVSQVVTILDQDYPLKLRESFRPPTVLFYQGKLALLKKPCLAIVGARQA
DAYALAALTKLLKPVLAKGIVTISGLAKGVDALCHQLTLAANQSTIGVIGTGLDSAYPRQ
NRNLQAKVAKAGLLLSEYPLGAQPLRHHFPERNRIIAGLCDSILVVEARHHSGSLITANL
ALQNNRNVLAVPGEITSDISCGCNELIAAGAKLVSQPDDILAEFANQLI
>Ribonuclease HII (EC 3.1.26.4)
MTEKLTIKAIKDRLASADLTAAELASFRQDERKGVISAVNSYDKKQARLAQKQLEFEKR
LEIERSFWRQGITYIAGIDEVGRGPLAGPVVSAAVILPPDFALSEVNDSKQLSPKKREEL
YGKILEQALAVGIGVASNKVIDEVNIYQASRLAMKEAVANLQIQPQQLIIDAMTIDSTIPQL
KLIKADAKSASVAAASIVAKVYRDHLMQFYERVYPGYGFVQNDGYGTKQHLAGLAKLG
PSPIHRHSFEPVKKALQK
>50S ribosomal subunit maturation GTPase RbgA (B. subtilis YlqF)
MTTIQWFPGHMAKALRQVKENLKLVDIVLELVDARLPESSRNPQLKELLANKPSILVLTK
MDLADPKQTTGWLKYYEEMGQPAVAVNSTQGSLKVIEKKIKEVLADKLAAKEAKGIMN
QRIRVMCIGIPNVGKSTLLNHLVKKNVAQVGNKPGVTKSQQWLKAGSSLQLLDTPGIL
WPKFEDPLVGKKLALTGAIKDTLYNKDDVALYALEHFNDTNPQALMERYKLTQANLDD
TTVELLLELTRKMGFKDDYDRASERLIFEIRRGKLGRYTLDFVPVSAAEVKND
>Carboxyl-terminal protease (EC 3.4.21.102)
MAKESKPKKQAKKVSLWTLILTGVISLLIGAGGTYAFTYKTVRDAQTVNVSMAKIQAVY
QTLYYNYYRKVDATKLEKGAIDGMLKALNDPFSEYMTKSEANSLNNTISGSFVGIGIEV
QKDGDYIKVIAPISKTPAAKAGLKAKDVITAIDGKSIKGLALERVIAKIRGKKGTSVTLTVK
RGQTSFTKTVKRASIPVATVYGNLDAKDKTVGYLRVTTFSTNTAKEFKQEVKRLRQKG
AKSFIVDVRDNPGGLMTTALQMASMFVKNGKNIMQVQGRDGQVEVYKSSKDYSKGF
KITEPTVVLVNGGSASASEIFASALNESANIKLVGTTTYGKGTVQTTQEFKDGTELKLTI
DKWLTANGKWINKTGIKPTIKADYPAVASLAAINTKKTYQEGQVSNQIKSLQKMLAYLG
YFKGQANGYYGADTKEAVTNFQSANGLATSGQADKQTVVKLETQVATKVTEHDNAYE
AGLKALGE
>FIG00743134: hypothetical protein
MFVRQSFYRYLMTLRNPNSQDELAQFANNAFYDQAFPKQASDYDEVSQYLELNGSYL
PSMTIFDQAWEAYLEKMA
>YfaA
MAEKRAQAKQINYWKWAFAILLALVIATVLFIGHKLTTPSVDQTRISQQTKQVDKNVDL
DVSMNKQQLSAMINYYLRKSQKKSKIKYRFILDKSAILMGTTKVLGSNVSFTLYAKPSLD
AKGNIVLDAKSVSVGSLNVPANFVLRYVKNNYDLGKFATIDYKKSRIILNLAQVSAKQGI
KVNGQKFDLKQNEFLFNVALPLK
>FIG006988: Lipase/Acylhydrolase with GDSL-like motif

MSKVKEGLIFLGVVLVLVGAFFSIMRLTSSKTHQEVSPTVSKKQAQVKAKKVPSISLYAL
GDSLTHGVGDTSNRGGYVYLVKQKLEHQNQVKVTSHNYGKTGDRSDQIQARLEKSK
QMQAELKQADVITLTVGGNDMMKVLQDNFLLLANNKLAGAMPKAKEDYRLKLTSLFKT
IRHYNEDAPIFLISIYNPFYVYFPTLTDLQKYTNQWNVLAKQVAKRDGNTYFVDVNQQL
SEGQYLGQDKKQLIKESKLNLNSVSSTKLEDILSDQKEKNNYLSDADHFHPNDKGYRY
MTKELYQVMMAHKDKWLKEDK
>Dihydrofolate reductase (EC 1.5.1.3)
MLAYIWAEDRAGHIGNEGKLPWNLPADLAYFKKQTMGHPMLMGRKTFDSFPGLLPHR
LHLVLTRSVALKKQAATTENLAVFNSVEEMLTWLTKHQAEEVFVIGGASLFELLKDQVD
VLYQTKIDAVFAGDVTFPEIDYQKFDLISVTPGQVDERNKYPYEFRVYKRKEN
>Thymidylate synthase (EC 2.1.1.45)
MTVVEQPYLQLIRDILEKGHEKTDRTGTGTKSLFGYQMRFDLSQGFPLLTTKRVPFGLI
KSELLWFLRGDTNIQYLLKHKNHIWDEWAFKNWVQSSDYQGPDMTDFGLRSQADPA
FNEVYQAQKQAFCDRIVNDDEFAAKFGNLGDVYGAQWRHWQTRQGETIDQIQNVIEAI
QKTPDSRRLIVSAWNPEDIPTSALPPCHTLFQFYVNDNKLSCQLYQRSGDVFLGVPFNI
ASYALLTSLIACQTGLEVGEFIHTLGDAHIYKNHFNQVEEQLNRTPFDAPSLWLNPAKK
DIMDFEMSDIKVLDYQSGPAIKAPVAV
>COG0488: ATPase components of ABC transporters with duplicated ATPase domains
MQTLRVEGLTKTYGEKTLFEDLNFIINEHDRIGLIGTNGSGKTTLLNALTEVDPADKGALI
KPNDYRIGYLMQQPKLDDEKTIMEAIYDGAGPVFKTIRHYEEVLTAYSANPTDEKLQAR
FDEAERLMTKEDAWNVESEVSTILTQLHIKDMDQKVATLSGGQRKRVGLAQVLIQAPD
LLILDEPTNHLDFDSIVWLEEYLASYKGALLVVTHDRYFLDQVATKIMELSFGKLYEYTG
NYEAFVSQKAERVEREITAEHKQQQLYKKELAWMRTGAKARTTKQQARINRFNELKD
NLNTLQVDADVDINLGQQRLGKKVLEVKDGSLTLGDHKIIQDFNLLIQAGQRIGITGVNG
AGKSSFLNALAGDLPLDSGQVIQGETVKMAYYRQQTEDIPDDKRVINYLNEVGQSVVN
KDGERISTTSLLEQFLFPRFMHGTLIRKLSGGEKRRLYLLKLLMEQPNVLLLDEPTNDLD
IGTLTVLEDYLDKFNGTVITVSHDRYFLDKVADRLLIFEGAAKIDHYRGRFTDYLAQTNE
KQSQAQKKKAKTAPTKASSANQSKVKTKLTYAEQMEYDQLEEQIDKLESKKAQLEAE
MGQVPGTDYGRLGELQGQIDQLDEEILSKMDRWDELSQYVD
>tRNA nucleotidyltransferase (EC 2.7.7.21) (EC 2.7.7.25)
MRVTNLPQAFKRALPILEKIQAAGFEAYFVGGSVRDTILNLPIHDVDIATSAYPAEIKELF
AKTVDTGIEHGTVMVIDHGIGYEITTFRTESTYQDFRRPDKVEFVRSLKEDLKRRDLTIN
ALALRPDGEVIDLFNGLADLESKTIRAVGQARERFHEDALRMMRTVRFASQLDFEIEAQ
TLAAIKENAALLAKIALERINVEWVKLMLGKNPKAGLADFMATGLFAYCPQFKEHQAQL
NSLLAIENLHLDNEAEVWTLLAACFDYHPEQGGQLVHAWKGSNELKAQVTKALRALEA
MKAGDLSPKLQYLTGSEMLAVACRVARLLGLADYSSMAVLLAYEGLPIHSRAELAVNG
KDLITLLNLKPGPILGQLLDQIEEAVLNRQVKNDKLALLEYGRSQLEE
>4-hydroxy-tetrahydrodipicolinate reductase (EC 1.17.1.8)
MVKVLVAGFKGRMGNTTVQMVLEHKGFDLVGVFDPMASEKNLNENPAYSQTDVPVF
TNLADIKTDAEVWIDFTRPTSVFENTKFALEHGISPVIGTTGMTDDQVKELQALAKEKQ
VGGLIAPNFGISAVLLMQFAKQAAKYLPDVEIIEMHHDNKLDAPSGTALNTAKMIAEVRK
EHYQGNPDEKETLPGARGANYEGMRIHAVRLPGYVAHEQVLFGGKGEALTIRQDSFD
RISFMSGVAVATEKIRNYSELMVGLENIL
>contains DUF161 family domain
MVASDEKEKISFLDLLMITIGCGLYALGFVKVNMANSLAEGGVTGITLIVRYWLHIDPAY
TTLALNIPLILIGWRYLGKRALIYTIYGTGILSAWIWVWQRLDFQLNIHHDMFIAGILAGLIG

GLGSGIVYRFNGTTGGSDVVARIFEKERGIAMGKSLLAVDTLVLVASLSYIDTRKMMYT
LLASYVFSRVVNFTLEGAYAARGMIIISQFNEAIAQDIMKEMDRGITYLDIEGGYSREAG
RAIYCVVSPGEIITVKRIIARYDKRAFISIMDVAEVTGEGFTYRVTKKKLFR
>hypothetical protein
MNIYLPNEETFAAFKETAAMIANKNLTSYESLPTFFSRMKNEVAREEA
>6-phospho-beta-glucosidase (EC 3.2.1.86)
MANFRKDFLWGGATAANQVEGGYDEGGKGLSVTDIQTAGSLTSPRYLTYTLNGKKDK
IAGAPGEGLPKGAKGAIFDDEYYPNHVAIDFYHRYKEDIKLFAEMGFKTYRLSIAWTRIF
PNGEGSEPNQAGLDFYRRVFEELRKYNIEPLVTISHYEDPLALSEKYGDWSDRGMIDA
YVNYAKVLFTEYKDLVKYWLTFNEINSALLLMEMFNQGRGISDEAYQHAYQKLHYQFV
ASAKAVKLGHEINPDFMIGNMICGIANYPGSPDPKDILANRYAWEQNIHYCGDVQCKG
EYPSFAKRLWNEHNVSLDITDQDLADLKAGKVDMYTFSYYMSNMVTTHEVEDKASGN
FSAGAKNPYLKYSDWGWATDPTGLQYYLELIYDRYHLPMMVVENGLGAVDKLEDGKV
HDDYRIDYLRQHIQAMDKAVANGVDLRAYTTWGCIDCVSAGTGQMSKRYGFIYVDRD
DEGHGSLERLPKDSFYWYKKVIASNGADLD
>UPF0246 protein YaaA
MQIIISPAKQMKACDDFEVLGQPRFLNESKTLMTYLQALSLEELQALWRCSDKLALVNY
QRIRDFSFERQLSPAVFTYVGLQFQALGADVFTQAQLNYVQKHLYILSGLYGLLRPFDG
INLYRLEMQAKLKGFSSPTLYQFWADKLYRELFAPGELVLNLASKEYAKAVSPYLQAG
DRFVSCRFAQRQNGKLVVKATAAKQARGDMVAYLAKVKAQGVEAVKSYQGYGYHYS
EADSTASELVFC
>Site-specific recombinase XerD
MDFPYADRFSHYLKHDRLLADNTINDINHDVADLFNYLRHFNVLYQVNPDLDQLEETDI
RAYFNMLQVERNIKNTTYNKVLTHLTTYFNFLFQNKLTTSLPTIGLKGLKRQTDDSLPLN
WVDQLPLLLSDDSLSYYGRLVLLLTSHFYTVKEMLQPGFYQVLDRESWLDFEQAFLAE
FKNHQADLQARQNSLELFLKQRLNLANPTLTLPALHKYLKQDQAQLSFSFKPSQLYQN
AVCYFLVSHQNLSDQQLATRLRLDLTSLNYYRQLNYRLNQN
>probable membrane protein b2001
MIRSWAKREYDYITPYDFSDRISKLTNPNEEAIPYNRVRHFSAESGILTEDEGVFAISYD
AHPHDDEVAFIEYGILVTTKGILIKNRVKDEASKGKQYKVEEIAIPYQDVFFVKLNRDNVI
MILRKGKLLPTIIKGESEARAKLIVNILKEAINSGWSLNCSRELGKLKEAQDNAIVEELAD
LDRKEQVAQAEATAKINEQAAAFDDASTKAVTLGMQADVAKGFKLNQINDRFGGGQG
HGHAAEQAGQVKNFWLGKNPVALGGTHEKHGADRVIGIINKKYIQTKFYASPSKTIGAV
FSSETGQPLYMNKGKMMIIEVPKDQHSKCVKLMAQKISEGKVPGETNPANAANYVKK
GALTYRQSQIATKSIFERNSTIQLNGQTVEVGFGKKLLYSAGLDFATGAALALPGSLVSF
VWVYANSVWQGTDSDEALNRATKAMVKPILTTAVMYTVASQFSKSLKAAAGNTNNAT
ALTVATGNMIKKVAGTSKIDTEIIASKTLGTLAVVATVGPDVTRFLVGRISPQQLTKNLAS
TGAGMVVGGWLGSVVPVVGTAVGAAVGGVVAKKIMDKITPDDNIIMSRILREEFIEIVLA
TPLSQEEFEDIIDHTFRDEKNYPKLLVNMFSAGGRSKEGLKKSREYIDEVLTNMVTAKF
AERDLPESKIIEAVKQEKDNVFVMA
>transposase
MVLTTGGIALTLGSLEPTFYPFASLCCVFTEQKCYNISKESEVNLMEIQITIKVKLNLANA
EIASSFTNTMEQYRLACNYVSEYIFNHDFDMKQSRLNKVLYTNLRSLFMLKSQMAQSVI
RTVIARYKTVKTQMKRNPYKYQDINTSEWYREIRDLTWLQKPITFNRPQVDLQRNRDW
SYLSNGQLSINTLNGRVKVNPVCHGFDQYFDGTWKFGLAKLLKTGGKWYLHISATKTV
ADFDNQTVKHVVGIDRGLRFLATTYDEQGQTAFFDGKAVMRKRAKYQKLRAKLQAKG

TKSAKRRLKKLSGRENRWMTDVNHRLSKTLVQKFGSNTLFVLEDLTGVSFERTDLPKV
LRNQNKSWAFYQLAQFLTYKAYLNNSEVIEVSAQYTSQRCPKCGTIKKSNRNHNLHEY
HCSNCGYRSNDDRLGAINIQLLGTQYISGQEQPKFELTTNA
>hypothetical protein
MYKYNKVYQELADILNDKDVEKIYKNFRGMQVNFPMRLYSRESVKKEIAKQKGEVDIK
DLAVQTGYSAYTIRRIISQLPLTKVRGL
>hypothetical protein
MSKQNLQDKFEIDSRTAQRDFANVRNFIGEYLNDYDLKYDAKDHSYYLSSGGLSLSKA
QCLVLIKVLLASRSLARKELKPLIKSLLALLTEQDRRDIKPIIQSELANYVELKNLKDKRLL
EAIYTFSKLILQRETITIDYTTQRHKYVTRTIAPVALTFSDYYFYIVSYNPQFKKNIIYRLDR
IVSWKYADKKIPLEQRTHFNLGDEKKKFQLMYPGTPITVQFKYTGIVEAALDKFPISKTIK
KIDEQTFLLEVTTYDTGALMWFLSQKDMVTIVSPPEFRDEMKATIQRMLANYED
>Substrate-specific component RibU of riboflavin ECF transporter
MNHFSTKKYVLVACLAAISYLLMFFSFAVIPMVPWMKIDFADVPILVGFFTLGPAGGSLV
AVVRSLLYLLVEGVSVASLIGVSTNLLASFAFCLPLYFTLRKQETKFVAYVRAISLATVSL
TFFLSLANWLVITPLYIAVLGLKLNVPLATMVLAGVVPFNLIKGALVGTIFALVYAKLHAW
LGQKASQLSH
>Ribosomal large subunit pseudouridine synthase B (EC 4.2.1.70)
MAERLQKVMANAGVASRRTSEKLILAGKVMVNGQVVTELGTKVTGADVILVNGQAIDQ
EKKAYYLFYKPRGVISAVSDDKGRKVVTDYFEEIPERLYPVGRLDYDTSGLLLMTNDG
QLANALMHPRFKVDKVYVAKVKGLVQTEDLKRLRQGLVVDGKKTAKAKAKLLKQDKA
KGISLVQLTIHEGRYHQVKKMFEALGYRVNKLKRECYAFLDLSGLTSGDYRKLTAAEVK
GLKQLVQG
>Segregation and condensation protein B
MLSNQAKIESLLFVSGTEGISLSELSQLTGLLKPAIVEQLGKLQAKYQADKTCSFELLKA
DERYRLATKKGFASLVKRYFESPTMTELSGAALETLAIIAYKQPVTRLDIDEIRGVQSSG
MLQKLLLLGLIEEGGRLELPGRPIVYKTTEAFLDYFGLESLAALPALPKVEDEETRTDGT
DFLELFQQTLNKDEGEQ
>Segregation and condensation protein A
MANESEQTKLTFKLASFEGPLDLLLHLIKQNEMDIYDIPMVTITSQYLAYLHQMQTLELDI
AGEYLVMAATLLNIKSKMLLPNQPVEVEADYEDPREDLVQQLVLHQTFQLAAEKLKDF
ATERQKSVGREQAQVPPDAKLGQLKAGATTVANLQQAFARLLAKKQLVKPVKRQVYQ
ERYSLKDTLVALTQKIRQAKAPIYFDSLFADQTDLEQVVTTFLALLELVKQGTVVAKQEV
TLGPILMISGEKDAF
>riboflavin biosynthesis
MLYKYKDSYKKIAMGLLSFIPDLKDISHLQTELDWYQKDASRFLYLWKNENGDFSGVV
GFELVNKQVIVRHIALSPAERNEGVTYKILDELATLYPEAKMMGSLEIAAIITKWEQRKD
GE
>Tyrosine recombinase XerD
MHYLVVERGLAANSVASYQNDLKKFSAFLAEQKVELTAVDKKVVLDFLEAQTKAKLAN
ASILRLVTSLRKFFQYLVEEGKIEVDPMLLIDTPKKKASLPEVLSPKEVEALLKTPKTGEA
LGLRDRAVLEVMYATGLRVSELVNLKLADLHLELGIIQTIGKGQKERLVPIGDVASKWVK
LYLKQSRPKLVKQKRNPYLFVNFHGNQLTRQGIWKNLKALVKQAGIEKNVTPHTLRHS
FATHILENGADLRIVQELLGHADISTTQLYTHLSKKRLSAVYKQNHPRA
>S1 RNA binding domain

MLADLIGKQISAQVIDENEHYVFAQTEGLTFRLAKAELKKLPKMGATISGFAYENEEHEL
QLTKTLPKSGIGRYAWGKVVKVNKSLGVFVDVGLPNKDVVLSLDELPTLTQLWPKVGD
ELMVALKVDVKGRLWATLATEEMFKQISLRATKKLQNRDVSARVYRLKLVGTYVITTDY
QLGFIHPSERVDEPRLGQVLSARVIGVRDDGSLNLSLRPRAYEAISDDAQMILAALSHS
LDHSLPFTDKSSPAEIKAYFGISKGQFKRAVGHLMKAGLVKQSEGKLLLNQTEGDIY
>putative membrane protein
MQTWLFLALIMLIAYLGKNMSLMIATGFVMLVKLLPFADKLFDFIHAKGINWGVIIITIAVL
VPIATGEIGFKDLLAAFRSPAGWIAVLCGVLVAVLSRNGVNLLASSPQVTVALVLGTIMG
VVFLKGVAAGPVIAAGITYCIISAFNLNF
>hypothetical protein
MQNAAYLIRSQRAYERVLNELAAQGYRYANGQPLETKPVFTNRIVLVTENGLVTKRSIE
KYNGDAMYRLTNHHLYVVD
>Formate--tetrahydrofolate ligase (EC 6.3.4.3)
MKDIEIAQQSEMEPIVKVAKRAGLSEDEIEQYGNYKAKIKLPLAARPQTKGKLILVTAINP
TPAGEGKSTVTVGLGDALHLLDKKVMIALREPSLGPVMGMKGGATGGGYSQVVPME
DINLHFTGDMHALTAANNTLAALIDNHIYQGNKLNLDPRRVIWKRVLDINDRALRHVVIG
LGGLTQGIPREDGFDITVASEMMAVLCLAKDIADLKKRLGQIVVGYTYDKEPVTVNDLG
VTGALALLLKDAIKPNLVQTLAHTPAIVHGGPFANIAHGCNSVLATQTALKLADYTITEAG
FGADLGAEKFLDIKTPVLGKTPDTIVIVATARALKMNGGVLKTDLDQENVTAVAEGYKN
LQKHIKNMQSYNVPVVVAINRFTADTDAELAKIQELCQADGVKAIIANVWAEGGKGAVE
LATEVLARVEEEHKFSPLYASDDELTIKIDKIVKNIYGGAKVEYGPKAKRQLKEFAKYGW
DKLPVCMAKTQYSLSDNAKLLGAPQGFTIHVREFVPKLGSQFVVALTGNVMTMPGLPK
VPAADGMDLADDGTISGLY
>pore-forming protein
MKFFYQPEWTTSITSWSYTLIILLSAIITWLEFTYFQIWSALLFGLFFLTAFLQVFRRQVFL
KEDGLVLRAVVPFNNKKLSYQEIVGVRKFKNGLEIKTKFRTYQVLMPLARRELCYQKLN
SLL
>Ribonuclease HI, Bacillus nonfunctional homolog
MIKLYTDAASKNNLSSAGILLVHDKKQIQLSQPLSASDNHQAEFQAAIAGFKQALSQFA
NEETVFFYSDSQLVIDSLNKNYSKHYQELLAELNSLIDSFKLVVWQWVPEKDNKGAHQ
LALQGLKKLTK
>Tripeptide aminopeptidase (EC 3.4.11.4)
MSNYPGLTDRFLKYVKVETRSDENSTTIPSTESQVEFAKQLAAELKELGLSNVRIHPTN
GYVLATLPSNLEAGKSAKKVGFISHMDTADFNAKNVNPKIVDNYDGQSVIKLDEAGEYV
LDPAEFPNLKNYQGHTLITTDGTTLLGADDKAGVAEIITAFAYLLKHPEIKHGTLQLGIGP
DEEIGTGADHFDVADFDADLAYTVDGGPLGELEYETFNAAQAEISFTGKNVHPGTAKG
VMVNASLLAMDFQSRLPQDQVPELTAGREGFFLLLSVEGSVEEAKLTYIIRDHDRQKF
EAKKALLNEIVAEMNQELGQKRVKAEVKDQYYNMREVIEKDMTVVDLAKEAMENLDIK
PLIFPVRGGTDGSKISFMGLPTPNLFAGGENMHGRYEYVSVQVMEKAVDLIIEIAKLAA
K
>FIG146262: hypothetical protein
MVLVKELVTEFERFAPKELAMPKDPVGLQLGFLEQEVTKMMVTLDVRPETVAEAIANG
CDFIFAHHPAVFRPIQPFDLAQPQNQMYAEILKHNITVYAAHTNLDCAAGGMNDWLAM
ALDLQAVKPLREDGLGRIGSLAAEMTVADFAAYCKQVFKLQGLRVISADLNAKVKRVA
VLGGAGGSFYEAALAQGADAYVTGDISYHTGHDVLAAGLNVFDVGHHVESICKPHLKS
LFEQIIRQKGWNLPVIESKLNTDPFIFC

>Putative tRNA-m1A22 methylase
MDARKLSNRLLTAAKYVKAGAVLADIGSDHAYLPAYLALNKQIKFGIAGEVVKGPFENA
QHEIKKEGLVGKVEARLADGLAAIELTDQVDTITICGMGGPLIASILEKGKAKLVNHPRLI
LQPNVGEKEVRKWLMQAGYAIMAEEILAEDGHTYELIVADYAPAEAHLNEQALMFGPL
LMQAKNPVFISKWRQEILKQEEILTHLEQAKVVPLAKKAQILATIKEIEGVLNGIS
>RNA polymerase sigma factor RpoD
MAEEKKNNNFDEVAYNKRLREVIRAYKPKKEISYTELTNQLVKPFGLDADQMDKLIGRI
EDAGISVVDENGNPSEASLMAADKEAKEAKEDAEKEEEETASTVKINDPVRMYLKEIG
RVNLLTADEEVQLALRIEEGDEEAKQRLAEANLRLVVSIAKRYVGRGMSFLDLIQEGNM
GLMKAVEKFDYRKGFKFSTYATWWIRQAITRAIADQARTIRIPVHMVETINKLIRIQRQM
LQDLGREPTPEEIGAEMDMPTEKVREILKIAQEPVSLETPIGEEDDSHLGDFIEDQDATS
PAEHAAYELLKEQLESVLDTLTDREENVLRLRFGLDDGGRTRTLEEVGKVFGVTRERI
RQIEAKALRKLRHPSRSKQLKDFLE
>DNA primase (EC 2.7.7.-)
MSLIPAEVIDRVRTSVNILDVVGQSVQMHKSGKNWFGLCPFHPEKTPSFSVNEEKQIFT
CFSCHRGGNVFKFLMELNGVTFPEAVKQVAQIAGIEIATDVGQTAAEKNSPRQKLYQL
YNRAADLYHHVLINTKLGQEAFSYLQDRGLSLELIKEFRIGFAPEQPLLEELFKKEEASD
YQLVRKSGLFLEWQDGSLHERFVDRIMFPINNQAGQVVAFSGRLLHANPQAPKYLNS
PETAIFNKSKVLFNFDKAKGEIRLKQRVILFEGFMDVLAAYRSGVKNGVASMGTSLTDE
QVYQLKRLTNQVCICYDGDDPGQLATKRALELFESQGSFNLSVINLPENLDPDEYVKK
YGEEKFAQAVAANQVGALDFYLQYFERGKNLATENDQLLYLRQVLEVLAPLKDKIQQD
LYLNRLSKRFKLEKLNLEAQLRDLQTKLKVQTKPLTDLKTSAVVPPVEKQAKVSQVITA
QRLLLYRLLHEQGVFLRLMADAQFHFFTEIYQRSYLVAQAYFAQHEDYNSAEFLDFIKE
EPIRQLVVSLEMTEFPPQALPEEIDDCIIIITKLAPLLAKLDQNDQEFERAMSLGNTEEAQ
RLLVEKIKLKQQIAQVKQALH
>Glycyl-tRNA synthetase beta chain (EC 6.1.1.14)
MAHTFLLEIGLEEIPAHVVTPSVEQLVTKVKNFLKEERLEFEAVKAYSTPRRLAVKVLGL
ADKQADIKEEAKGPAKKIALDADGNWTKAAQGFVRGQGMTTDDIFFKEIKGVEYVHVE
KFIAGKQAAEILPKLKEVAMSLKFPTMMRWGSNDFEYVRPIKWLVALLDQEVVPFAILD
VKTDRLSQGHRFLGKPVSFASADDYPEALVKEMVIADAAERKNMIRNQIATLAKENDW
VVDVDEDLLEEVNNLVEYPTVFAGKFADKYLEVPDVVLITSMKDNQRYFYVRNKAGKL
LPFFIAVRNGNKAYLENVVAGNEKVLTARLEDAKFFYQEDQKQTIAQYVERLKTVMFH
DKIGTIYQKMQRVQLLAANLGQKLGFTEEQLADLKRAAMIYKFDLVTGMVGEFSELQGI
MGEIYARMQGEKETVAKAMSEEYMPTSAEGKLPASPVGALLSIADKLDSIYAFFAAGMI
PSGSNDPYALRRQALGIVRIALDQNWNLAVTDLSQAIKAAYKADASLFEKIVPTDSQAE
MESFILDRLHQVLAADKYSHDTLDTVVAVKANSFVTSQKAAAVLTKHRADTDFKETIEA
LTRVVRLAKKAPNFAADAKVDASLFENESEVALQAAVAQVKEQFSADVAADYELLRKL
RQPISTYFDQTMVMAKDEVIKNNRLLQLSQLADLTTIFGQLDKLNVK
>Glycyl-tRNA synthetase alpha chain (EC 6.1.1.14)
MTKKLAMQDIILTLQQFWAKQGCMLMQAYDTEKGAGTMSPYTFLRAVGPEPWNVAYV
EPSRRPADGRYGDNPNRLFQHHQFQVIMKPSPSNIQELYLDSLRALGIVPEEHDIRFVE
DNWENPSMGCAGVGWEIWLDGMEITQFTYFQQVGGLEVKPVAAEVTYGLERLSSYIQ
DVNSVFDLEWADGVKYGDIFKEPEYEHSKYAFEESNQEMLFKAFDVYETEAKKQIANG
LVHPAYDYILKCSHTFNLLDARGAVSVTERAGYLARIRNMAKAVAKVFVAERKKLGYPL
IKDEKLRQELLKEDK
>Pyruvate kinase (EC 2.7.1.40)

MKKTKIVSTLGPASSDLDTIVKLIEAGANVFRFNFSHGDHEEHLGRMNLVRQAEKVTGK
MVGIMLDTKGAEIRTTVQKEGKLHFNIGDEVRISMDASIEGTKEKIAVTYPGLYDDVHEG
GHVLFDDGLIDMQIEKKDEATKELVCKVLNEGTLGSRKGVNAPGVSINLPGITEKDSDDI
RFGLDNEINFIAASFVRKPQDVLDIRELLEEKHMEHVQIFPKIESQEGIDNFDEIIKVSDG
LMIARGDMGVEIPAENVPLVQKSLIKKCNALGKPVITATQMLDSMQENPRPTRAEASDV
ANAVFDGTDATMLSGESANGDYPVEAVATMARIDVKAENALRQHREFSLNRFDKTDV
TEAIGRAVAEAANNLNIKVIVAATKSGHTARMISKYRPDADILAVTFDDRTRRGLAINWG
VYPVMAEAPASTDEMFALATEEAKKAGFAKEGDLILITAGVPVGEKGTTNVMKIQLIGS
KLVEGQGVGDETVIGKAVVATSAKEAVEKAEEGGVLVVKTTDKDYLPAIEKSAALVVEN
GGLTSHAAVVGISMGIPVVVGAANATELVADGEVITVDSRRGIVYHGASNAL
>6-phosphofructokinase (EC 2.7.1.11)
MKRIGILTSGGDSQGMNAAVRAIARSAMHAGLEAYGVNYGYKGLVEGDVFKMDSLND
LDGIINRGGTILRSARFPEFAEEETQLKAIEQLKKLGIDALVVIGGDGSYHGAMKLTKHG
FNSIGVPGTIDNDIPGTDFTIGFDTAVGVATEALDRIIDTARSHQRIFVVEVMGRGAGDIA
LWSGISAGADAIVIPERDFDAKEIAAKLTDNKKRGKDYGIVVVAEGVMSAQELKDQIDQ
YGDFDSRAVTLAHVQRGGMPTAKDRVLASRFGDYAVQLLLEGKGGLAVGIRDNQLVA
KDIIDTLENHKHKTDVSLADLNDRIRF
>DNA polymerase III alpha subunit (EC 2.7.7.7)
MFSPLQVISSYSLLQSGLRIEDYVQAAKKLGYQALALTDYNVMYGALTFYQACLRANIK
PLVGITLQLTSSQSDLITSDIVLLAKNAAGYQNLMKLSSLKLAKPEGVSLTLEQVDANLS
DLIVILPPNLTYLIASNPAANLTWLTALKARVASNLYFGIEANFDEFSKQTLIKLAEQLDCK
LVASLRVEYLEPKDILQQEVLAALKSGRTLKEADLQGKQRGKHYLKSPTQVEAEYEQA
GLTEAYANVAGLCQQVDLQLEFTKTKLPHFKTPSGLDSESYLRKLCVEGLDVKLRKAA
VSATKPYFDRLESELAVINKMGFDDYFLIVADITNFARSQKIMLDPGRGSAAGSLVAYSL
GITDVDPLKYQLLFERFLNEKRAQMPDIDIDLPDNRRDEVISHLKKQYGQNHLAQIITFD
TFGARQALRDVGRVLELNSFELERWSQAIPRVFKISLKEAYEQSRSLRNLVNDSPKNK
LLFRLALSLEGLPRHYSKHAAGIILSDTDLVETVPLHMGNDGIWLSQYAKKQVEAVGLL
KIDLLGLRNLSLLEQALSLVKRGYQTELDVRQIDLNDELTLQLFARGDTAGVFQFESSGI
RNVLRKLRPTSFEDIAAVNALYRPGPLTNIDEFIARKHGLKQVTYPNDNLKSILGLTYGII
VYQEQVMQVASTMAGFTLGEADILRRAISSKNDHLILELKDQFIKGAKKLGYEVSVAKQ
VYDYIERFANYGFNRSHAVAYSKFAFQLAYIKAHYPAAFYTALFNSLLASPKIREYVIEA
KKYQVNVKGPNINLSQKGFTLRQGQIYFGLASIKKIRRDFCDAILLERRQAGNYKGFQD
FLERIPEKYRQAALIEPLIYVGAFDEFEPNRKQLLTNLESQLKTIELTDSPVLQSLLETVD
RPVEPFSLEEQLDKESEYLGLVISGHLLQKYSWIKTFYPVTAVGDLQVNMKTRLLVYIR
KIKVIRTKNNEEMAFLTGSDESGEVDLTLFPTEFRRYGLRLKKGQSLLVSGEVQEHNGL
SFIVKQIKFADLMKKTCWYLRLTPENNALAKLKLTSTLRANHGETPVIIYDESAGTKRLL
HERYWLSTDAGLEAELINLLGQENVVLKK
>hypothetical protein
MMGKQLTILFWGFIYGEVLGYIVSALSGYTFDAKLSGLIPMIGGFIVINCLSYFVKDSQEK
>ADP-ribose pyrophosphatase
MEQVTGLFCNDSGYQTPKLDTRAAIFKDDKILLVHFKLPTFLRFYLT
>ADP-ribose pyrophosphatase
MLKISKKGGMAADQLLDWIIELQSIAQAGLTYGKDIYDRERYQRPAEMMAQKVVFLWN
K
>contains bacterial regulatory deoR domain

MKLAHHFGIAFDTARRDVLKLTSTGQAVRIHGGLMEVKLNNVPGFVTRIHIQSPIKAKM
AKLASQEVHPGGLYFIGASTTLMQMCELLGRVNTTVVTNSIDNAIRLTLHELPAVELLG
GKIDKGNRYTYSLETLAKLDNYTFDAAFLGTSKIQADGITIVDAYDATIIKKVIARSKKVVL
VVQKYKFEAQNSSPYKSADLKDIDVLITDEPLSAQYRAYFKPDLKVKYVKD
>drug resistance transporter, EmrB/QacA family
MNETVSRKTNISVASVGILTFTGILIETSLNVTFPTLMSTFNEPLAVVQWLTTGYLLMVTL
MMGATAFLLKRFKARRLFLVASSLAIIGGILCLFAPEFWLLMLGRLLQAAATGIATPLMF
SVVFNKVPAKKWGTYTGICAMLISLAPALGPTYGGIMNHYLTWRYCFVFVVILLGLALV
AGLLNFEDNETYPQTKFNLLSFIYLISFLTSLEFGLQQLSSRLGLAVGLIILSLVCLALFIFR
SLTKANPLIDLRILKDSYVSWRLVDYFGLQFTNISLSFIIPIFAENYLHVNSLVAGLMLLPG
SLAGAIIAPLSGRWYDKSGANKPIPLGNFLMLLGAVLFAGLTKQLTVALIVIFYIFTRAGF
NTAFGNVMTDSGKAVPLAKKADQNSLFSMSQQYAGSLGTAIISGILTFAQKGSNTSLQ
AIKVGSQLSFGFLAIISFIGLLSVIILHKRHVPAKKQFDK
>FIG00754979: hypothetical protein
MTRKIGIFSDVHGNYTALKAVYEDALALGVDDFWFLGDLFSPGPGAQDLWELFTQMN
PSYCLRGNWDDLLVNSIAGRLPTDKESKIYFAKLAQTLAKRLAPEVVMEIERWPLAKQV
ELAGLKVGLSHNLPTFNHGQALYPLADQANFDQLFESKPDLDLAVYAHVHHPLMRYSS
QEQLIFNPGSVGQPFSQHPIFQADLRANYLVLTLAEKKLTEVDFRKVDYDHQAELKRA
QAVELPYLDLYEEQLESGRIHSHDYELLAKINTERGYRQDVLAYNQAKKRFKN
>membrane protein
MREYSFLLILIVLIREISLPYFMGFFSANSKIQLLSELGSNKFPFHQAFDRWEFIDGILFML
GAYYIYLWAQRQAASRYAKPLALLVALYGLGDCLLTSMFVYSGSTFANWHSFLHSIAS
VLGYLGLYLANLLIAKIKHDNRFLVAVIGVSQLLSLGLNLLMESPLVTKASTVGLLQCVAL
DLMYLPLLILACLELHHKLALIRVRN
>Substrate-specific component PdxU2 of predicted pyridoxin-related ECF transporter
MTSLTKKIALLGIITAANVVIARTFLIPIPFTHGNINLCDAGIVLAALLFGSRAGGVVGALSG
LLLDLISGYPQYMFFSLLAHGTEGYIIGQLGQKKQWLGLLAGTLVMVTVYFGADSFLYT
VPTGLVGIPLNLLQGVIGVTVGYLLVVGLNKRLAF
>Pyridoxal kinase (EC 2.7.1.35)
MVEINKTFLLAADLAGIGSMSLNSAVTIMQIFGVNLALSPTVLLSTQSEGFGKTAKLNTA
VFLQHSFSHWHQIGQEFSGALVAYYEDEQVGALLKQWLVTTKPPLVIYDPACADQGAL
YPNISVELLSQKKELLTLADVICPNLTEASLLTGIKSAESKPELYMKLLQALATSNPRADI
VITGISLGEKRGCVWLDQDKLQYALTPTLASHFYGAGDVFSALLSGFLLRGLSFSKAIR
VTTNLLYSTLVANLNSDRADSYGLDLSSLLAEFILRKDDFNDIIN
>Carbonic anhydrase (EC 4.2.1.1)
MSKLDYQLQDQWPRLSNLESPINIATPTCQFKSVADQPLTVNLTSGPVKKKNQVNGPQ
FLLTGTIQLGKRTYFLQKMHFHDGSEHYLNQRPAKCELHFVCQDAKKHTLVLALLANIS
GEAPKRNWDSLYKGQATTSAFCNFFEKQLAYYQYQGSLTTPPLLPDVTWVVLAQPATI
TPSDYQVIHQDYPDNHRQLQDLKQRTITYYAY
>hypothetical protein
MAIFLQVMVPVLLIYLSGFLLQKALKLDIKPISTLAMYLLLPFLVFKTFYERKLDSNFGYIL
VITLVIMLALLGLGLIIANILHYNQKKLNAFLLTTIFSNSGNFGVPLALFAFGNSAMKYAMP
IMIIHTIYMGVVGIYIAASGSGKQQTIGYALGSVFKQPMNYVIVPGILLNRWGITLPENLM
KWIDMISKTAIPIIMLALGMQLANVVASKIDWTSVGFASFVKLIVAPLLAYLICLAFPINPLL
RNIIVIMTAMPSAANTTIYAIQYNAEPDYVSACTFVSTILSFISLTFLLNLVV
>Glycerol dehydrogenase (EC 1.1.1.6)

MTTNCMSNYSYGPNCFDELPAILAGYGFTKVVFIGGKRALASSQTEVKQVLSQANVEV
TASLVYGVESTQANIDRLVATEEVAAADVIFGFGGGKALDTAKMVAMELAKPIFTFPTIC
SNCSAGTAVAVIYHDDGSFDRYGFPPVPLHVFINTQVIANAPAKYFWAGIGDGISKAPE
VAHCVAHAKETGMELPHTATLGAAIAASSKGAFYQYGPQGIADVKNNLASRAVEEVAL
DIVVSTGYASNLVNQKEFYFNSCHAHAFYNATTAIKREGEYLHGQIVSFGVLVLHAYYQ
EDEELQRVLAFNHRLGLPVTLAQLGLSQADIPKLVEVAMTTNEYKHLLFEPDQFAAAIK
QADELGQALS
>hypothetical protein
MLEKIRLRKPLDQKVAVGDMITYQTTSFIIINILDVGLASWGKESIFAVYTCLVQQLNSPN
LSENYTTTQTELAYELNEQRNISRVGDIIYDENTGIWVQVNAILAIRYEDEKIYVKYEFDPI
PEWSPKEISKLQDKRRLQLMKLVKHEQKR
>1-deoxy-D-xylulose 5-phosphate synthase (EC 2.2.1.7)
MSELKQLANEMRQLVLEKDAAISGHVGPNLGAMEAMIAFHYVFDSHKDKIIWDISHLAY
GHKMLTGRKEALAVPPNTRIIGPRYRRKFIKELILKLFDFN
>5'-nucleotidase YjjG (EC 3.1.3.5)
MRFKTLLFDIDDTLLDFDANENQALARLFASLGIELTPQVKAEYKAFNQSLWKKLELGQI
SRDELMAHRFATFFKEHFNLNVGNELNKRYLGYLAEGHQEISGAKQLLQNLQLAGHEL
YVVTNGVKFVQEKRIADAKLRPYFKNVYISEEIGYQKPARQFFKHVFADIGKIDLDKTAII
GDSLSSDIKGAVNSGISSIWFNPKQVKATGITPNYQVTSLTEIEKIV
>Histone acetyltransferase HPA2 and related acetyltransferases
MIYLRPATMEELPLIMQIFTEAKALLKADGSSQWQSGSPNEETFKADIAANEGYVLIVD
GQVGGYLALKAGIDPNYLVIREGSWHLQEDYLAIHRVMISSQFRGKKLASYLFSNLLSL
AYAQNKFYLRIDTHPKNNRMQHLIKKAGFDYCGLIEISGDPDPNRYAYDLKLGGR
>Glutamate racemase (EC 5.1.1.3)
MDLQTVKRPIGVFDSGVGGISVLKEAVKLLPNEDFIFFGDSKNAPYGVKSVEQVYQLSK
KIVTDFIAQDAKAIVIACNTATSAAAKRLRSEHPEIPIIGLEPALKPAVLAKENSHVLVMAT
PLTLKEQKFADLMKRFESQAKISKLPAPDLVEYVEKGEVASPKLKKYLRSILAPYIGQID
SLVLGCTHFPFAAPAIAEIVGPEVTLYDGGEGAARELRHLLVKNQLRRTDKRQGKVTFK
NSINTEAELNLCRTLFSL
>Aspartate aminotransferase (EC 2.6.1.1)
MPIKDKFNHKIEKIAVSDIRQFDTEVSSIPGIIKLTLGEPDFNTPEHVKQAGIKAIEDNQSH
YTPNPGIMPLREATAAYFNEKYHLNYKPTQVITTIGATEAISVALQTILNPDDVVIMPTPV
FPIYKPITEINQGNCIMVDTSADGFILTADKLRQVLAENQAKTVKALVLVYPSNPTGVTY
SRKDLEELAEVVKEAGIWVLCDEVYAELTYNGQHTSLAEIIPDQVLLVSGLSKSHAMTG
WRIGYLFGPEDFVEQAVKAHQYMVTAPTDNVQFAALEAMTNGKEDSQVMKAEYLKR
REFLQAELAEAGFEVASPDGAFYLFAKIPAQFEPDSWAFVRRLAKEAKVALIPGISFGP
GGEGYVRLSYAASMENLTEAAKRIKAFVAAN
>lactose transport regulator
MSQEKRIELIKQLLAERQELSTKEIMEEFGISQDTARRDIVLLTKRGEVKRTHGGILPLDF
GRSVPNFQSRLSRFTKEKTQIALEATNYFRPHHVYFVDSSTILLKTCQNVNMPLTIVTH
SLDNCMALAENNKVTVKVLSGTLNHENRFFYSNRAMGELQNVAFDTAFVAASGIDEN
GIYLLDQGDAEIVGLAVERARKVILVAEHQKFVNKSYYRICSLDKISLFITDEPLTKAQRA
MFPETTEIKVATRSDR
>Biosynthetic Aromatic amino acid aminotransferase beta (EC 2.6.1.57) @ Histidinol-phosphate aminotransferase (EC 2.6.1.9)

MKEEIKKLKSYVPEEPVVVVKERYGLDRLVRLSANENPFGTSPKVKEAVVNWTFSESN
RYPDGYATKLRQAVSERLGLPEEQLVFGVGLDEILELLSRTFLEKGDEILVADPTFSEY
ALHAQIEGARVTKVPVKPTDGHYDFERALTEISSKTKLIWLCNPNNPTGVYERSSQIAA
FIKQVPATCLVLIDEAYIEYVTDEKPATSLDLLNDYDNVAVLRTFSKVYGLANYRVGYIA
MAPKLAEYMQTVRLPYNLNSVAQVAAEAALADQTFVADSVAKTVEGREYLEEFLTNKQ
IKHFHSQANFVYLKYPAALELAQKLLENGYQVRTGLQLDWLRITIGELADVKAICALIDK
DLN
>Phosphoribosyl-ATP pyrophosphatase (EC 3.6.1.31)
MQDLEALYQLVSERKQNPKEGSYTDYLFTKGLDKILKKVGEETTEVVVAAKNHDNDEL
RYEVADLLYHLMVLLVEQGLSYEQVKVELASREGLMSQFKDRPEIKNL
>Phosphoribosyl-AMP cyclohydrolase (EC 3.5.4.19)
MREPDFSKGLLTAVVVDECSKDVLMVAWMNEESYRQTLATKETWFWSRSRQELWHK
GATSGNTQRVKKIMLDCDQDTILLYVNPAGPACHTGHTTCFFEQIL
>Imidazole glycerol phosphate synthase cyclase subunit (EC 4.1.3.-)
MLAKRIIPCLDVADGKVKKGVNFVNLIDVGDPVEIAAAYEKQGADELVFLDISATTEQRK
TMVDVVSQVSKEVFMPLTVGGGINSVADMQALLKAGADKVSLNSAAVANPDLISQGAE
LFGSQCIVVAIDIKTDPVTLEKYVYTHGGKKRTDLKALEWAQKVVALGAGELLVTSMDH
DGTQAGFDVDFYQTLTNLVTVPVIASGGAGSLADFSEVFLKTKVTGALAASVFHFGKLS
ISQVKAELKTKGVTVR
>Phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase (EC 5.3.1.16)
MILPAIDLQNGQSVRLYQGDFTKETIISPAPIKQAKEINAAGVYALHLVDLDGAKAGSPQ
NLAVIKKIRAEFTGLIEVGGGIRSLDTIETYLAMGIDRLILGSVALKDPAFTKQVLAKFSPE
RIVIGVDGTDGKVAVDGWLKQSEVTMATLINEMQQAGAKHFIVTDVVKDGTMTGANM
DLLTSLQSQFPQANIIASGGIRNLQDILALQAAGIADAVAGKALYEGSLTLKEIAEVNGC
>Imidazole glycerol phosphate synthase amidotransferase subunit (EC 2.4.2.-)
MLAIIDYGAGNTYNVAKAFAYLGYESQLTSDPAQIAASSALILPGVGAFNAAYRKLHAD
GLVTAIKAEVAKGKPLLGICLGMQLLFDSSEEYGNTPGLGLIPGRVIPIPANNLLVPQVG
WNQNKLLQEQGLFKRVKDEYTYFVHSYYAKCEQQYVVSYVDYGVKVPAIVNYQNVYG
MQFHPEKSGQVGLGLLKDFMEGNGLR
>Imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19)
MRTATLTRQTAETKITVSLNLDQESGIDIQTGVGFFDHMLTLLAKHGRFGLKVMAEGDL
EVDAHHTVEDTGIVLGECLKQALGQKEQIERYGQATIPMDEVLAQVVIDLSGRSYLVFD
ADLTNPKLGNFETEITEDFFQALAFSAQMNLHARIIYGRNTHHKIEGLFKATGRALRQAV
TINPAIKGVNSTKGVI
>Histidinol dehydrogenase (EC 1.1.1.23)
MIKIYQEDLETMKRIVADYTEQVVDFEIEAKVQAIITQVRKRGDQALIDYARQFDGVSLT
DLKVAEEQLEQAWLQLDDSLKTSLKLAAQNITSFHELEKEKGFVDASQAGIVRGQKVT
PLKRVGLYVPGGTAAYPSTILMNAIPAKLAGVEELVVVTPPQKEGINQAVLAAAKLAGV
DAIYQVGGAQAIAALAYGTETIKPVDKIVGPGNIFVATAKKQVFGTVAIDMVAGPSEIGIL
ADDSAQASEIAADLLSQAEHDKRARAMLITNSLSLAQAVAKEVERQLVNLPRKEIAQAAI
MDRSFIAVMPQIAEMFDLMNTIAPEHLEIQLTEPTAYLGLVKNAGSVFLGRYASEPLGD
YLAGPNHVLPTGGTARFSSPLGVYDFVKRIQFVQYSKEALAEVEKAVTTLARTEGLEA
HARALESRFTNYEGE
>ATP phosphoribosyltransferase (EC 2.4.2.17)

MTLKIALTKGRVEKQVVDLLKKLNIDVSPLENKQRKLIIPLADGQYEFILAKGPDVVTYLS
NGIVDLGIVGSDILMEQNVSGYQLLDLNVGKCQFILASTTDFNPTSAGRKVIGTKYPRVA
QEYFARLGQDVEIIKIEGSVELAPLVGLADAIIDITETGTTLRENNLQIYDWLAKISTRLVV
NPLALKQKKNEIFQLVDQLAQVI
>ATP phosphoribosyltransferase regulatory subunit (EC 2.4.2.17)
MTKYKLPLGFRDYFGSVAQSKVQVENYLTALFTALGYSQIETPLLEDQTVFEQYSLTNE
DVYRVLGPNGEVLVLRPDLTLPIARFLATTNIELPQKFAYIGSIFKRKRELMVFSNQEAQ
AGVELVGYASLKAELECIYLVNKLNTELLENKLYLELGHANFAQEVLAKLAIPLQEQEAIK
QALFTKDLPKYEQLIQAYSKQELFSFLKLWPRLFGSLTEVAADLAKVKLPADLQAKCQTI
LNLGKLVAKLPNQEVRVDLTSSAPQDYYTGVTFKAYLPTESAYLVSGGRYDKLLNTFQ
AQQVPAVGMGIDLDLLAQLVIEKDQPGRQVKKQLYCETKQLKQALAIVEQTPDLTLSLA
TSLQAARLLAKKKGEKLQILTDKGELVDDIENSVD
>Histidinol-phosphatase (EC 3.1.3.15)
MKIDGHTHTELCPHGNHDRVEKMILKAIKLGFTDYCLTEHAPLPLNFKEQFAGNLSGLK
TAALTFNQVDEYLKLAEDMKAKFAGKINISIGFEVDYLQGFEEEIRNFLDEYGPYTDQNI
LSVHYMPGLKQQFYGIDYDPIEFATGFGPWLEMPDALYGKYFTQIKNSLLADLGQYSPI
RIGHMSLIKKYQDYFGLDSNWDEENLKQIKEILDLIKEQGRQLDFNTAGLRKPYCNDFY
PGKTIINLALEREIPLVFGSDAHSLEEVGAAWHFYDDYLN
>Manganese-dependent inorganic pyrophosphatase (EC 3.6.1.1)
MKELVFGHQNPDTDAIVAAKAYSYLQNKLGKDTEAVALGEPNEETKFVLAHFDEPAPR
VIESAKAEVEAVMLVDHNEAQQSVADIKDVTVTNVIDHHRIANFETAGPLYYHAEPVGC
SSTIVYKEFASNGVEIPAKLAGLMLSAIISDTLLFKSPTTTADDIEIGKKLAAIAGVDYETY
GLEMLKAGTNLAAKTEAELIEADAKSFEMGGLTVRIDQVNTVDLEDVFSREAAFRAAIE
AENKAQDYDLFLLMATNILDSNTRLLVVGEGQDKVEAAFGVKLVDGKADLPGVVSRKK
QVVPPLQAAFD
>Transcriptional regulators, LysR family
MKTKQENIFSSKTLSYFLQLAETMSYTQAAQILGITQPALTQQIKKLERTVGAPLFYSIGK
KLHLSDAGYTMLDATHQIYETLNRATDEIQQSTSATQGEINIGLLASIEVSVFTDFAIKYF
EENNGIKLTIHSLTRKEIWERLENNTIDLAVMYLPDESIKNWKPYESKRILTEDLLFLHAE
DKLAKRRRVKMKETLNNKWVTYPPQYYLNDLIRESYKNALVDKPISVAHFTKPAQIFNF
AKETNVYTALPSSFVTAHNSAMNGLKTALFEPEVKFDMSFVFRKGKDQIPRIESFLYAF
ERYLSQKDYVSRLTEISEKY
>TPR-repeat-containing protein, putative component of Menaquinone-cytochrome C reductase
MMTYSEETLELLASGQYEEAKKKFAWALRKDDDDMLYSLAEELYSLGFSTMSKRAYQ
KLLERYPNEDGLKTALADIAIGEGNDDEALTYLADIAPESEAYVEALLVSADLYQTQGMF
EVSEQKLLEAYRLVPTEEVIWFALAEFYYNIKKYEQATKFYLELIKAGTFELAKVNLVQR
LGVSYASDGHFEQALGYLEQIPKEKLDPDTSFQLGFTYQQLGKNEEAKKQYSALRETN
PDYATLYPALAEVLATEGNYAQALITVQEGLGVDEYNPNLYQQAADYAQKLGQAAKAS
DYLKEALALEPDNMTLVLKLSNLYLLLGRDKDNIAFLKSYLASNESDGQLYWNLGKAYA
NLDDYQAALEYYDAAYQELDKNSSFLRDASFFYRNAGQAKQALTAVKSYLETNPSDG
EMLELESELENW
>DNA-binding protein HBsu
MANKAQLIENVAEKTGLTKKDATVAVDAVFGSIQDFLKDGEKVQLIGFGNFEVRERAAR
KGRNPQTGAEIEIPASKVPAFKPGKALKDAVK
>GTP-binding protein EngA

MANPVVAIVGRPNVGKSTIFNRIAGQRISIVEDIPGVTRDRIYAKSEWMGQKFNMIDTG
GIDIGDEPFVTQITEQAQIAIEEADVIVFLVSAKEGVTDADEKVARILYQTDKPVILAVNKV
DNPELRAEIYDFYSLGFGDPVPVSGTHGIGTGDLLDLICKAFPKREAVTEDQTVKFSFIG
RPNVGKSSLVNAILGENRVIVSNIEGTTRDAIDTKFVTEDDTEFTMIDTAGIRKKGKVYE
STEKYSVLRALQAIDRSDVVCVVLNAEEGIREQDKHVAGYAHEAGRGIIIVVNKWDTLK
KDNHTMKDFENLIRQEFQYLSYAPIIFVSAKTKQRLANLPALIKRVNTNHERRIASAVLN
DVVMDAIAHNPTPTDNGKRLRVYYATQVAVKPPTFVIFVNDPELMHFSYERFLENQIRA
AFDFEGTPIHLIERRRK
>SSU ribosomal protein S1p
MSDTEAKDLLAALDSVQEVKIGDVVKGEVLAFDDSKQAIVGIENAGVEGVVPARELSVK
EDELHVGDTLELVVISKIGSDKEGGSYLLSQRRLEARKVWDDIEKKFEAGQTITAPVTQ
VVKGGLVVDAGVRGFVPASMVSDHFVEDLNQFKGQELEFKIVEIEPSENRLILSRRELV
KAERQAKKEAVMAKLVVGDVVEGKVARLTNFGAFIDLGGVDGLVHVSEIAFERVDKPA
DVLKVGQDVKVKVLAVDADKDRISLSIKQTLPQPWDKVAEEVAEGDVLDGTVKRLTSF
GAFVEVFPGVEGLVHISQISHKHIATPNEVLESGQEVKVKVLSVNASDHRLALSIKALQE
KPAVEKKEAAPKAEEKVTLPEEETGFTLGDLVGDELKNND
>Cytidylate kinase (EC 2.7.4.25)
MKKKLQIAIDGPASAGKSTVAKLVATKLGYIYCDTGAMYRAVTYAAIQAKVALDDDQKL
AELLTDLTIEFIPGQPEQRVIVNGQDATKQIRMPEVTNNVSVVAAQRSVRTDLTKRQQE
IAAAGGIVMDGRDIGTTVLPTAEVKIFLIASVEQRALRRYKENIAKGINTPLAQLEEEIAQR
DYKDSHREISPLTKASDAVELDTTPLSIDEVVAAIMAIIEKKQAN
>LysM domain
MTSHNKKPWENTFEDEKDSKGNYSRTVNRRNTKRNTVLTTSLLVIFIVLILGTIGLYFVS
QRTASEEQASSSVTIASTTKSISSKKTSHKEKKTTTKASSVKETASSSEASSQATSQAE
ESSQAQSTESAQTSSQQEQSQAEASSQEQSSQEAASYATVGAGQGIYRVAANNGLSI
SELLQLNPGLSANSTLTPGQTVRVK
>ATP-dependent DNA helicase RecQ
MANLEQFLKQTYHFSSFRPGQKEVINALLAKQSMLAILPTGNGKSLCYQFYGRYTNKR
VLIVSPLLSLMQDQVSQMKYQGYQQVIALTSEQDYLTRQKVLANLGAYQYIFLAPEILQ
NQLVKQALQTLKIGLLVVDEAHCISTWGPDFRPDYLQLGKLQRLLGQPLTLALTATATP
KVKEDICQQLGLSKVYQRSVNRENIYLEVLKVASQSQKDQKLLELVNYLQGPGLIYFSS
KKLADRLANWLGQRSQLRVAAYHAGLSHLDRYSIQQQFLQGQLDLVCATSAFGMGVN
KEDIRYVIHYHLPSDLNSYVQEIGRAGRDGKQSCAILLYQAGDEFIQRQLLENTLPAPLE
ISQYFMSKAQGQSKISSEKQRLLASFNDQVAGEQRALTFFAARKKERLTALKQMATYA
KTSGCLRQYILTYFAEKPGQHNDYNCCHNADYWQTFAQLGLKRKSRDERSNAWAGE
FEYQTILARLFTNYNKSEE
>hypothetical protein
MNYYWLLFFDANRPRRPAVLRQLLANRQKGSALYFGMLANWLQYIGLFSGFDEAKFD
REIQQLVTAAYLEAGLAGLTLSTKGVEFINENDLKLELAQPKLFRIFPMELVANLILLAVQ
VASEASYQNKQYYVVTDNVYSQYLFKRWLTQSNYGLTGLQEELVPSLATFLANEPPQ
KAVIFAQKLRGHNFPGQTNEQLATIHGKFPFEIEQIWLDLLSRFAYYLYQGDGKLAELM
ALTQPNFEPVRASRFKTINAFMAGNTIKEIASHLHIKENTVLDHLYEAYIWHGKPDLLKL
VSAKEQELMTKLFVETKSQEEWSYQDLVAVQPEIAFYKFRIFEIARGRKRWQT
>Manganese transport protein MntH
MSNNETKHDKEKFIAYANGPSLEEINGTVKVPKGKGFWRTLLAYSGPGALVAVGYMD
PGNWSTSITGGQNFQYLLMSIILMSSLVAMLLQYMAAKLGIVTQLDLAQAIRARTGKTL

GIILWIMTELAIMATDIAEVIGAAIALYLLFGIPMFLDVFITVLDVLVLLLLTKVGFRKIETIVV
ALILVILFVFTYQVALSNPDWGAVFSGFIPSAKTFSTTPTVNGMTPLMGTLGIIGATVMP
HNLYLHSAVSQTRQIDHDDEEAVAQAVKFSAWDSNIQLTAAFVVNCLLLIMGVAVFKTG
AVQDASFFGLYDALSDTSTLSNGVLISVAKSGILSVLFAVALLASGQNSTITGTLTGQVI
MEGFVNLKMPLWARRLVTRLIAVVPVLLCVAFTSGQSTIQEHQSLNTLMNNSQVFLAF
ALPFSMIPLIMMTDSEATMGRFKNSLIIKVLGWFSVISLTVLNLKGLPDSITDFFGEHPTA
TQVSVAHGIAYFLIVAILLLLAWTIVEIKKGDKHLAKMNSNHKA
>Universal stress protein family
MSDRINFEITPRSFTKILVAVDEDDSNSSVRAFNYAVTTAKRLGAPLGIVSILEVQDFNIF
DTLSPDIMKKRRNQVSDDLARYVNKARSFGVKEVFAFTGEGKPSTVIIEDILPEFQPDLL
VCGSKTKPVTGRKQIFIGSQASAMAQNAPCSVIVVR
>Mobile element protein
MQKKVDKFYDFLDTILNPSGKLKKAIGYAKKFHTRLEKIYEIGELPLANNPVEQAIRPATL
VRKNSLFAD
>Mobile element protein
MLVNYHTPRHVYIVCGNTDLRKGIDTLAILIADNFGLDLYDDSLFLFCGRRNDRFKAFY
WDGEGFILLYKRFDNGRLTWPRNSEEVKALDSQQLNLLLKGLNPLPQSKVHRAKKG
>hypothetical protein
MEKQEIVKFKLTLKRKPKAKNRTNKLRLATEIKLDKVKISLYHGADVELMNDIWELVKKH
AG
>Sakacin A production response regulator
MSFDKLKEYLPYRGEKYIKPEKAGNYRQYMMDLRDLARQARDEFSLISKSFEDRVKPF
KAERVSQWMNQSQLCRPHFWCYFRLPSDGLDDPTLAIRLYGESDNFGISVEVSFVER
RRSENSLEKQNKVLNLLPFGAMYYFVQKNGISFKMDATEENRKNLLKQVKSGEVRKVL
VKQDIPIETDHSLERLIDDLLKSFDELLPFYKETKK
>transcriptional regulator, Crp/Fnr family
MVQHICVTLVPLFHHLSEDEQIKINRLAQHKQFKKNEIVFRPGDENLDIVALGSMKVFQL
STNGDEQLLRVVEPGGYEGENQLFGIKNESLFGKTLETTEICRLSKQAFDQVMLRNPQI
ALKLFELSAQKMIQVEKQAHILSMERVEERLANYLLDLSKAANDYRFSLPMKMIDIARYI
GTTPETLSRKFRFLEDNRLIKRSGRRIIILNRDGLTDL
>hypothetical protein
MLKQFFEANDEMLDLYTNAITKAHGKNHPEVFDVRKVYETIQHKIRDNNLDLSNEFSEL
RALTKDYTIPDDACETLTKTYQLLKQFDELSQE
>Copper chaperone
MQKIEGAIKSVTGVDKESVKVLFNASKAKVNFDASMTNINEIVKAVQAIGYDVLKASEK
>Probable cadmium-transporting ATPase (EC 3.6.3.3)
MQKWLLRYRNWLTAITGVLIILAFSSKWIFSSEQGSAYLLFVASLVGGLPIFVQAYQALR
IKVISIDLLVTLAILGAFVIKEFEESAIVAFLFLFGAYLEQRTLAKTRSAIKNLVEMVPETTF
RKLHNGDFEEVSVEEVNEEDILLVKTGGKIPVDGEVIFGSGTANEASITGESIPVSKGLR
DKVFAGTILENGTIQIRTEKIGEDTTFGKIIELVEEAQDSKSKAERFIDQFSKYYTPSVLVL
AIIVWLISRNIELAVTILVLGCPGALVIGVPVSNVSGIGNGARHGILFKGSDVITNFSKVDT
MLFDKTGTLTYGNPQVAQTIYYSSDRSLAESLLASVERESDHPLAKAIVNYYSGSVVEV
IDSTEVVQGGGVVARIQNHQVLVGNSYLMNQYNIPFTNQMKSDIAGMEAEGNSIVLTAI
DGQIALVVGIRDQIRQGVKEDLETLKNMGVKNLILLSGDNQGTVDLVTNQLGLTEAYGH
LLPEDKAEFVRKRQASGEIVAFVGDGINDSPSLALADIGIAMGNGTDVAIETSNVVLMN

SDFHRIPHAIGLAKATRRNMIENIFIALLVVVVLLISVLSSSWMNMAIGMFVHEGSILVVIL
NAMRLLKYKRK
>amino acid transporter
MNELGQKQHYISWPVIAMVGFVTIIGFDDILYPLQNQGLSVVFSWIFMVFFFVIPYEMTV
AHLGSTFNGHQEGGLASWVRRSTHSDALGYWTAWMYWAACLPYIVDVANSTIVSYS
WLILGNNHLDQYMSSFWFGVWTFVIILAFILLENFYKKSLELMATIGGVSMFLMTVLFVV
MTAWTLLTGGKIATHPFNLEAFIPKFSTSYFATTGLLMFAVCGAELVAPYLTKMRNPNR
DFRKAMWLIALMTAFLTVFGTFSLAIFFNANDLPHDLKMNGSYYAFQLLGEQMGLGKG
LMYLFALVQAIYMMAQLAIFLDSTSWVFAADTAAKFMPKWMRKRNRNGRPIHSYVLTT
GLSLFLLLMSGTLPDINAVFNWLLNLNGIVFPFKNCWLFLAFIAVRYQQDRFTSEFVYIK
NRMGALAMGSWCFLFSLVCAVMGFLPQDVTFGTKAWDHQLLMNFLSVAVLFGAGMI
MPLLARLEKRKASVQE
>FIG001721: Predicted N6-adenine-specific DNA methylase
MKKYNLIATAAAGIEAIVGNELRHLGYEVQVENGRVRFQGTIEDVLETNLWLRTADRVK
IVVGEFYARSFEELFEQTKALPWDEFLPLDANFPVAGKSQKSTLHNVPSVQAIVKKAIV
TKLSQVYHRTTRLPETGAVYPLEVAINKDRVLLTLDTSGSSLFKRGYRVAKGGAPLKEN
MAAALVLLAHWYPEKPFVDPVCGSGTLPIEAALIGHNIAPGFNRDFACENFAWVDPKL
AQMVRDKADSQANYDIELDIWGYDIDGKMIEIAKQNALEAGLAASVNFKQMALKDFKTT
KEYGVIVANPPYGERLSDQASVRQLYKQMGQVYQPLTTWSKYILTSDLEFETYYGQKA
SKRRKLYNGSLRTDLFQYWGKRKPRN
>Cell division protein GpsB, coordinates the switch between cylindrical and septal cell wall synthesis by re-localization of PBP1
MENVNLTPKDIVNKNFKPKMRGYDPTEVDEFLDMVIQDYENFNKETQRLQVENDRLVA
KVDELTKQLAVGNSGKTAQPTSSTTNMDILKRLSNLERHVFGAQLNGDNNQTNRF
>FIG005686: hypothetical protein
MRLWLSGYRSYELGIFNEQDLKVKVIKAALKERLVAKLDQGLEWVITGPNLGVEQWGL
EVVLELKKAYPDLKLALLEPYADFSSRWQEAKQVKLAQLRSEVDFFGLVSKKSYQGIR
QLQNYQNFCLDHSDQALFVYDELYPGKLKFAYQASKKKANYPCELVTMDELQTFAAD
YQENYGSW
>RecU Holliday junction resolvase
MVMHYPNGQPYRTKQLSLKKSPYPTATLYGNRGMTLEEEINESNQYYLKQGIAVIHKK
PIPIQIVNVDYPKRSAAVIKEAYFKQPSTTDYNGVYKGHYLDFEAKETKNLTAFPLANFH
AHQIEHMKACQKQGGICFTILRFTSLKEVFLMPASLLFQYWDNQQANNGRKSIPYEIIKK
QCAHITYQLQPKIPYLTYVDKLLEQIL
>Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-) / Penicillin-binding protein 1A/1B (PBP1)
MTKQDQQYSRVARRRQAAHKQKHSQQSKKKWSFKKTAVTVALSLVALIFVGFIGGSLL
FFAYAKSAPALTTSKLTSDSSTTIYDAKGHKILTLGTEKRQAIKEQKVPQQLKDAIVSIED
KRFYKHRGVDPYRIMSAAVNDVTGSSSLQGGSTLDQQLIKLAYFSTKKSDQTLKRKAQ
EAWLAIKMDKQYSKDQILTFYINKVFMGNGIYGMQTAAEYYYGKHLSELSLAQTALIAGI
PNAPSSYNPYSNVQLATQRRNEVLNAMVANKKISAAKAAQAKAVSVTDGLVSQDAHT
NSSQTKQAKWSDAYLKQVIKDLYDKGYDPYASSLKVYTNLNLDIQKRLYDMANTNEYV
YYPDNKVQIAATIINPNNGKVLAMLGGRKNGDVTYGLNRAVQTSRTNGSTAKPLMDFG
PAIEYLSWATYHQLEDTEYNYPGTSIALHDFDKAYKGTMTMQDALMQSRNIPAIRALAT
VGLTRARNFVNKLGFNYKSVEYSNGIGMPSSTLQNAAAYAAFDNGGTYYEPSYINKVTI
ADGESKTYESSGKQVMQASTAYMITYMLQQVITNPSGTGYRANIPGFYQAGKTGTNA

YPDDIKTNFPDNATMDSWFNAYTKHYSVSIWMGYDHPYRDANYLDTATARIPIAFYKSL
IANISQGKSNTNWKRPSNVYTKTLDGQKYLYLAGTSDPYTSQKKHSNSSTHSSNSVAN
SQDTSSISSLTTTSSVLSSSSSSSQSTAESSSSSTASANTNSSQSQTPSQPSNPTSSA
ATTTTTTNPNQ
>Endonuclease III (EC 4.2.99.18)
MLDAKETLQAIQIMGKTYPTTGSSLVADTPFHFLMAVILSAQTTDKAVNQISPALFARFK
QPEDLANAKPSEVEPYIQTIGLYHNKAKYLVACAQGLVTQFDGQVPKTHKELMSLAGV
GRKTADVVLAECFGIPALAVDTHVGRIAKRLAMVKEDATVNQIEQTLMKKLPKEMWITA
HHRMIIWGRNQCLARAPKCQTCPLLFMCQEGQKRTQS
>Chromosome replication initiation protein DnaD
MENDWEFFVDAKQVTISSLILANYQKLGLTEEELVLFLQIDSQLQAGKNIPDIEKISQAM
GKASTDIYSLMHKLVEKKVLAINTVSDENGRETNQYSFKGLYQKLVALKRQEEMTTSQ
ATLSVSRAEIYKSIEVEFGRALTPMELQTIDMWFSQDHYQPELVKLALREAVLNQVYNL
RYIDRILISWEKQNIKTAADAQKLLVKQREVKAAKWSKTTSQSKAKDKPEIPVYKWSNR
TEGEH
>Asparaginyl-tRNA synthetase (EC 6.1.1.22)
MQTISIIDAKNHVDEKVKIGVWLTNKRSSGKIAFLQLRDGSAFFQGVVVKNNVSEEVFK
RAKEVKQETSMYVTGTIHEDARSHFGYEIEIEDIEIVGESENYPITPKEHGTDFLMDHRH
LWLRSKRQFAIMQIRNEMIRATYEFFNQEGFIKLDSPILTGSAPEGTTELFHTEYFDRDA
YLSQSGQLYAEAGAMAYGKVFTFGPTFRAEKSKTRRHLIEFWMIEPEMAFMHQEESL
EIQEKYVAFLVQAVLDNCDYALTILGRDKEVLKEYTKLPYPRISYDEAIKLLQESGEFPD
VKWGEDFGSPEETYLANHFSKPVFVMNYPKAIKPFYMKPHPTRDDVVVCADLLAPEG
YGEIIGGSERAVDPAYLEAQIKQAGLDVADYEWYLDLRRFGSVPHSGFGLGLERAVTW
ITGESHIREAIPFPRLLNRIYP
>Aspartate aminotransferase (EC 2.6.1.1)
MNFSQRAANIKPSATLALANKAKEMKAQGIDVINLSVGEPDFKTPAHIKQAAIEAIQAGK
ADAYTAATGILPLREAIATYTNNDYQTNFTAKNVAVTVGGKLSLYALAQVLFEARDEVLI
PLPYWVSYGEQVKLAEAKPVFVAPALGNDKVSVVELEAKRTDKTKACIINSPQNPAGLI
YSREELEAIGNWAVEHEIILIADDMYSKLVYNGNQFVSLLNLSDAIRKQTILVNGLSKAYA
MTGWRVGYTVASEEVIQKLAAFLGHATSNLAAVSQYAALAALTGDQACVEEMRQAYE
ERLNTIYPQLVALPGFKFTKPQGAFYMFPNIRETLELTGFKSADDFAAALLEQAHVAVV
VGSAFGMPDCIRISYATDLASLQEAVKRIKKFIEEYRH
>hyothetical protein
MKRSRQSRFGWRKLLISLGVILALYGSWYYYHQTQAPLTVAKDEAYALADKYGHLKTH
TAFYRYNRQETYFAVAGENQRGQKVYVLIAKRGGKINIYSQDKGLTAQEVQSFVKARY
RPKKILNLALGLYKKQPVWEVSYLNSKNQLCYDLIAFKNGRIVKSIQNL
>DinG family ATP-dependent helicase YoaA
MKSLKASNTYAVVDLETTGTNLDGSQRIIQFACVLVEDGQIVNQFSTLINPLRPLSPEVE
KLTGLKQKDLKKAPSFDEVAASIYALLQGTIFVAHNIQFDYRFLNLELERVGYPALELKGI
DTVQLAQILLPTQLSYRLGDLGKALKIEHDHPHHADSDAYVTAKLLLLLLAKLKELPRDL
LVTLNKLAGELVYDTGVCFKEALAQKDETEILDADLLQVAGIILKRPHFESGAKLTMAYP
LTEQEKRQQFTEIIDYRSEQVRLMDDIEEFSHSKARYFLAQAPTGLGKTLAYLYPAAYL
ALGGQKVVLAVPTNTLQEQVLTTSIPIIEQLVGKLRVVTLKGSQNYLNLEKFWQSLRQA
QGKYTRIVQMKILVWLSQTETGDLTELHLLKTKDSFFEQISHQGIEGLDKTSPFYQVDF
VRRQEVATKNADIVITNQSYLLNHASDFASLGALLILDEAQHLIDLAASYNRQVLDLDAIK
ILADSLLVKMESQVSLSFEQLVATGFMTRYSYQQIKHYTQVIDHSVMDVRSRLIQYFYR

KEAGNEIYLSNNKLRGFVKAHLGEIKQIGFAFEKLLNLNSKLEGEFIQASRKQALTQEEE
GLLLDYFTLVTSFNKELSQWSQLDLELIETKQNKGITWLSLNRSQANAHLRLNFGLLEG
KDYLATNIYQKFAKVILLGASVLTPKTKSYVLNQLDLPADLKIHTYKSQFNYRKQAQILVA
RDAPTVDQLEYSQYLVQTINQIVTESPRQTMVLFNSLEMVAKVYQGLVELGLTAKRDV
LAQGINGGVNKLKKRFALSQTMDSILLATGSFFEGIDLPEDKLELVIVTRLPFTPPDSLY
NQVRYQRLKAAGKNPFSVLALPEAIIKLKQAFGRLIRTENDRGVFVLLDSRFVTKRYGS
EFNQAFPKECSLQAVLTKQVVKEVKDFFES
>Mevalonate kinase (EC 2.7.1.36)
MEEKNTVLASGKSHAKIILIGEHSVVYGQPAIALPLSAVTVTATLKPLPANNQLIKSSYYD
GPLTTAPQSMLGITNLIKFMLAKANQPQLGFCLEINSQLPAERGMGSSAAVTIAIIRTFAA
YFAWNLSRPQLLKLAQLAEKETHKNPSGLDAATSASLKPIWLIRNQEMYPLKIALDAYLL
IADSGIKGQTKQAVSLVKEKLVNDPSSTQGQIQALGSLAHLARKQLAHNDVLGLGQTF
NQAQSHLQALGVSSKELDNFISLSLNNGALGAKLTGGGRGGCFICLVKDQATAQNLAM
VLKEHGVTQTWLQPLTNLEEFSNEENGPRPH
>Diphosphomevalonate decarboxylase (EC 4.1.1.33)
MKKTARAHTNIALVKYWGKKDQELIIPMNSSLSLTLDHFYTDTSVEFKEELTADSFSLN
GVSKDAQQITNFLNLVRKLAGITTYAQVDSTNHVPTKAGLASSASAYAALALAASSAAG
LTLSKKDLSRLARRGSGSACRSIYGGFVEWKMGTGDLDSYAVPFQEEVDWDIRMLAIV
LNDAEKKISSRRGMQTVVATSPYYPAWVKTCHQDIGELKAAIEAKDFERLGTIAEENAM
KMHALNLSASPHFNYFQPETLVAMQLVEELRAAGIPCYYTMDAGPNVKVLCLEKDVAQ
IKAKLAEAFNQEQLLLAAPGPGAHLLA
>Phosphomevalonate kinase (EC 2.7.4.2)
MIKVKAPGKLYIAGEYAVVENGYPAVLVALDQFVYVSIEKSASLGSIVSKQYEENSIYWR
REGQELVFDNRDNPFHYILAAIKLTEEYARSLGKVLDCYHLKIDSQLDSPDGKKYGLGS
SAAVTVATVKALCAYYDLKPTKEILFKLAAIAHFEVQGNGSLGDVAASVYGGWLAYHSL
NRDWLILARQKYSLKQLLAMDWPDLSIEELTPPKDLTLLIGWTGSPASTSHLVDAVGLS
KAEKQAKYQEFLKASKSCLTNMIAGFKTNSPAEILTQVRVNRDLLVKLASFSGLNIEPPL
LKTLIETATDLGGAAKTSGAGGGDCGIALLTDTNQVSTLINRWQAAGIEPLAFNVYKE
>Isopentenyl-diphosphate delta-isomerase, FMN-dependent (EC 5.3.3.2)
MDKQAHRKDEHVFIAEKFYRPQAQTKLDQVKLLATNLPELALADVDLTATFLGQKIAAP
FFINAMTGGSEQTGKINARLAKVAAKTQIPMAVGSQSIALKLPSLSSTFEVVRQNNPAG
FILANLGAHHNLINAKRAVKMINANALELHVNLAQELVMPEGDQSFYWLANLKEINSHL
DVPLLVKEVGAGMTPRSLVLLKEAGIKYVDLGGAGGTDFVAIENERRENKELGYFNEL
SLTTAESLLAAQTFQADFSFTATGGIRTALDIIKCLVLGADNVGLAGHFLHVLLRQGDDA
LVTEITNLKDQLKRLMVMLGCRTISDLKTVPYLLSPELINFKEQVTSLKF
>tRNA and rRNA cytosine-C5-methylases
MELPKDFISKYQRLLATEAEEFLAAFDGPVEKGFRLNPLKENYQAVAYSLAKKAPHVKT
GYYGEVSGRSLDHQAGYVYSQDLSAMYVAEVADAKPGEWILDLCAAPGGKSTQLVEA
MANQGLLVANEINKKRAAILAENLERSGASNYIVLNESPQNLAKNWGPIFDKIVVDAPC
SGEGMFRKDHQATSYWDKDYPAECAKRQKEILKSALELLKPDGKLIYSTCTFAPEEDE
QIVAWLLAEYPDLRVLPIKQYPGMDQGRPDWADGNPDLEKTIRLFPHHFKGEGHFVAL
LEKVGQAVPKKVSKKKGKAKQAKGNLSLEQLELWQAFSQELVGTDLFASADLKLVGD
YLYYSPKTWPRLANLRYLRPGLLLGVFKKKRFEPSYTLALALKPSQVKRVLAVGETEW
ENYVKGNVLTTTSSQQNGWYLLTCQQKGFAFGKKVNQQIKNFFPKALRFY
>hypothetical protein

MVKKNLLVVVYVAFLAFIAVATFYKIQERDLVIQLNNHNLSAAAYKVKVKSSQSLAQLNE
AVSKSKLTNLQLHYQAKDNSQVTYFYGKGKFATPPMITGNFFSEDDFKSDVAVAVVGK
DWQKKLYKPQGQAYIKLNGTYIPVIGIMGDSYKSDLDKQIFIAQSQQQTPMMKAADYKL
VVDAKQALTKKDLRTVFKQAKITPLVTNVFIVPQSTWVAEHLRELAVLALVVIGLALELG
LWVASFRRKYQAAKFAKLSLKKYAFEEWKFYTLFTGLGLAFGTVLGILRYSFTSVTAPIL
YALGIYFVASLGCYILLKKKMSEFERNEKR
>probably aromatic ring hydroxylating enzyme, evidenced by COGnitor; PaaD-like protein (DUF59) involved in Fe-S cluster assembly
MTEKRSNQAIKDDILAGLEEVVDPELGIDIVNLGLVYEVNLDEDGHCEIQMTLTTMGCPL
TDLLADMVERQLVKIPEITEVEVKFVWEPAWTMDKMSRYAKMALGIH
>Putative hydrolase of the alpha/beta superfamily
MITVTEKTVATLPVLELIEADLKTSKVPLVFFYHGWTGCKEKVLTQGYELAKQGFRVVL
PDAYYHGSRASEPANKHQLEFWHIIQHSVTEFPKLVAAYNQVGYDKLGVSGLSMGGIT
TCALLATYPEIDAAVCLEGCPHPLAFAKQLVAAIPGANQLPAELLAQEYQQIASYDLALQ
PEKLAGRPLHFWHGTADEIVPYQLTADFHQAVSKKATEKVTMTTSQGVGHKVSYDTTL
EMAAKFAQYFKK
>FIG002540: Haloacid dehalogenase-like hydrolase
MKRKLIAIDLDGTTLNSQSQISSKTKAVLNKAVKAGHVVSIVTGRPNRISEHFYDELGLK
SPMINFNGAVGHLPHRAWDKEYQLTFNKDIAFELLAAKKKLGIKLLAAEGKNLSLADQP
NDIIPDFFPAKLKPEEFLNKSNLTTDPSAMTMLVDRQKLPNVVAELERSFGDKINVGVW
GGPTPILELAPKGVNKAYGLKFLSQTLAIDQKDIIAFGDEHNDAEMIDYAGWGVVMQNG
TDQLKKLANDITPLDNNHDGLASYLEDYLDLAE
>UTP--glucose-1-phosphate uridylyltransferase (EC 2.7.7.9)
MKVRKAVIPAAGLGTRFLPATKALAKEMFPIIDKPTIQYIVEEAKKSGIEDILVVTGKGKR
PIEDHFDSVPELEQNLEAKGKDELLRLVEETTDINLYFIRQSHPRGLGDAVLMAKDFVG
DEPFVVMLGDDIMDDKVPLTKQLINRYEQTHASTLAVMKVPHDEVNKYGVIDPAAQTH
PGLYDVKQFVEKPDVDKAPSDLAIIGRYLLTPEIFDMLETQKPGAGNEIQLTDAIDRLNKI
QRVFAHEFKGQRYDVGYKFGYLKTSIEFGLAHPEVKADLTAYIKELAAKLK
>tRNA-Trp-CCA

>tRNA-Phe-GAA

>tRNA-Asp-GTC

>tRNA-Pseudo-TGA

>tRNA-Leu-TAG

>Adenosylhomocysteinase (EC 3.3.1.1)
MELLTAIKNYQPFNEQEALDKELILQQLTSNDQVFLRQSKLAHITVSAWIISPDYQKVLM
AYHNIYQSWAWLGGHADGETDLKHVILKEIAEESGLTAVEFVTPSLFSLEVLTVDGHQK
HGHYVSSHLHLNVTYLLQASDDLPLRIKADENSQVAWLKLSELKQKVSEPWFYDHIYS
KLMTKVKTLC
>hypothetical protein

MVGEIFMSRFKKKYIAVRVSYLNGKQVELQLPKDLQKPMWHYIHEHPHDWQQLLLGAL
INTPVGKYRNRKVPLMKVGKICAVFIKNKALPNHSRGQFITADKWQSPLINPWQAAFKQ
NVRFLQHDYPPLHKYLIAKDCLLWWLKTKWRP
>Stage V sporulation protein K
MIYRLYAEESRLYLQETSIETVADLEQLTMAYADGIYFKAHTENVINLYLRAKAMIYLKQA
FFEEVISPNVRLEDGALATVDQLELADGDASLLELKADATSHIIFAHQATGLKQAEAETR
FQTEDKLVGLSEVKASLATILQKARFYQLRRQRGLNAEGVNLNALFMGPKSSGKATSA
RYFVQQLAQAKLVDEQRFIKVQAKELTDERGCERLGEYLEDTRGGVLLIEGAQQLLGQ
TELLDLLLQGQTPVIILTGPKKELTKLNSVPGFTKHFPNQLNFTGYSQNELVEIGKKLLA
DWQYRLADEDYYTQTLAHLCQQANFNQASAVADFNHRLREQMSQRVMATADTDLETI
TNEDIKAVADQLGTNHATKDAYAKLNSLIGIAEVKRAVKNFIDLNVINQQRKQAGLAVAN
TSMHAMFLGNPGTGKTTVARILGELLYQKGLIRKPTLIEVSRSDLVAGYVGQTAIKTRKK
LEEALGGVLFIDEAYTLSTGGSNDFGLEAINEILKFMEDHRDDIVIIFAGYTKEMQAFLAM
NSGLESRIPNTFTFADYTIAELVQIGLLDLKKRNYQVEQAAYAYVVSESYQKSNDHSNG
RWVRNLNERLVRLQASRLAQSDGQGDLELITRGDLEKLLATN
>hypothetical protein
MARFIVGPNDWNFRDAYQQASDGDMLELQDNARVDLGSSVFQINKSLEIVGQMTPTK
DLTCYIDGAIAITNQAQVTLRRIIFRAEIDRVMLSVDNASLKLSQVIIYNGYQDAMTKVSI
WANDADVAATASIFKAISSDTGSTLKLSHSHLDLTNTSLDKVRLFADTDSNLKLAQIDSF
GLVNQNALNLENSHAKLKNCSFKIDDDVTNQDKAATVYVKNTSL
>UDP-glucose dehydrogenase (EC 1.1.1.22)
MKRIKAKGAKVIICEPTLADQATFFNSPLVNNLASFKAQSDVIIANRYSLELADVKQKVY
TRDLFGKD
>CDP-glycerol: N-acetyl-beta-D-mannosaminyl-1,4-N-acetyl-D-glucosaminyldiphosphoundecaprenyl glycerophosphotransferase
MKKIIYLWLIKLVSLFYVNKPKSGTVYLMSFGNNLAFIKALAKKDELLVCYQAQCQAAAK
QLEASGIKTMLIQDGFGFAFKKLGLILTARKLFCDNYYAFLGGCLLDHHQVKVIQVWHA
NGAIKTFGLEEPRTATRSASDQRRFRKVYDQFDDYVVASPKMGQVFANSYQVSPTKM
RVLGYPRSDRLFKQGWQKETKARLLAAYPNLRTKEVILYAPTYREVRDNSGQVALNLP
PDFDEVVANLAPKQVLVIKLHPHLKASAVKLKQRYSHNPKVLWLDEFSTEEVLVITARLI
TDYSSVIFDYSLLPNAGQLIFYCYDLEAYAAYEGLQSDFVSWLPGPLVKSAAELSRILAQ
PCQLQDLVEFNRLWNTQNDGQATERVLAYYYPR
>Integral membrane protein
MDELQKLWKRHVYVMRASTAFIYGILVSIAMNFFWTPGHIYSSGITGLAQLITTLTKHTP
FPLSTALLLFVLNMPLFVLAWFQIGHRFTIFTVIAVFASSFMIKMLTPLTLTTDPIICAIFGA
AINGFGTGLALKNGISTGGLDILGLTIRKRTGKSIGTINIIFNFFIVMAAGIVYGWPYAFYS
ALGLLVNAKVMDMVYTRQQKMQVMIVTTKPKSVIDCVQNHINHGITIIHGAEGAYMHVE
KTVLFTVISRYELPELRQAMRESDPKAFVSIADNVHIMGHFEEPSV
>Peptide methionine sulfoxide reductase MsrA (EC 1.8.4.11)
MSQQVETAIFAGGCFWCMVQPFDEQPGIISVRSGYTGGHVPNPTYEQVCSHTTGHTE
AVKITFDPTIISYADLVEIYWHQTDPTDAMGQFQDRGDNYRPVIFVNSDEQRKIATASR
QRLAESGQFSDPIVTQIEDAKPFYEAEEYHQDFYKKNPARFEAEEGLRNQFKKEHWEK
>Aspartyl-tRNA synthetase (EC 6.1.1.12)
MKRTTYCGLVDKKYLDQEVTLKGWVQKRRDLGNLIFIDLRDREGIVQLVFSQEFDQAS
LEVAETLRTEYVLEVQGFVRARKDGAVNPNMKTGEIEVEVHQVKVLNQAKTTPFYIQD
DINVSDELRLKYRYLDLRRPEMQKALMIRSKITQAVHRYLDENHFVDIETPYLTRSTPEG

ARDYLVPSRVYPGHFYALPQSPQLFKQLLMGAGFDRYYQIARCFRDEDLRGDRQPEF
TQIDLEMSFMTAEEIQELTEGLLKAVMKAALGLEIKTPIKRISWNEAMARYGSDKPDLRF
EMELKDLGQAVAGAGFKVFDQTLANGGQVKAIAVKGGADKYSRKQIEAKQDYIKRFGA
KGLAWVKVTEDGFSGPIAKFFDEEHVAKIKAATDASAGDLLLFVADKAKVVADALGYLR
VEIAKEQDLIDPNEFAFVWVVDWPLFEYSEEFERYIAAHHPFTMPNEEDVALLETDPHK
AHAQSYDIVLNGYELGGGSIRIHQRSVQEKMFKALGFSEESAKEQFGWFMDALDYGF
PPHGGLALGLDRFAMLLAKKDNIREVIAFPKNSKASEPLTNAPSTVAAKQLEELSLEVE
TNEE
>Histidyl-tRNA synthetase (EC 6.1.1.21)
MRYQRPKGTADIMPDQSAKWQYVEQTARALFRKYRYQEVRTPIFESFEVFSRTSGET
SDIVTKEMYDFYDKGDRHITLRPEGTAGVVRAFVENKLYGPEVIKPVKTYYMGPMFRY
ERPQSGRLREFHQIGVEAFGVDNPTLDVEVMAMAVELLESFGLTSLKLALNTLGDADS
RAAYRQALIDYLTPFKDQLSDDSKERLTKNPLRILDSKDEGDQAIVANAPKILDYLTAEA
KEHFQVVRDLLDQLGINYEIDTNMVRGLDYYNHTIFEIMSDAKAFGGKWTTVCAGGRY
NGLVEQLGGPETPGIGFGLGVERLLLILADQDLDLEDNLDVYVVGIGQATNAESLKLVQ
SLRKQGFSAERDYLNRKPKGQFKTASKLNARFTITIGESELESGTANLKDMQAQSESK
VSLAEIYADFAKLAQAKVSEE
>Xanthine phosphoribosyltransferase (EC 2.4.2.22)
MKLLEDRIKEDGIVLEGNVLKVNSFLNHQIDPDLMAAIGQEFAKLFKDQGITRILTVESS
GIAPAIMVGYALHVPVVFARKKKSSTLNEAVYSAEVFSYTKNVLNTISVDQKFLTASDKV
LIIDDFLANGQAVLGLIEICRQAKADIVGAGIVVEKSFQKGADLIKKAGVHLESLARIEAFV
DGQVKFVEE
>Xanthine permease
MMTTKTTNPNQAKAAILGLQHLLAMYSGAVAVPLLIGSALNFNAQEMTYLVSIDIFMCG
LATLFQLLRNRYFGIGLPVILGCAIQAVEPLKMIGKQFSIGTMYGAIIVAGIFVFLIGGQFS
KIKRLFPPVVTGTLITVIGLTLIPISLQNIGGGDVTAKNFGDLQNIGVGFFTVVVILILSVWG
KGFIKSIAVLIGLILGSLLAACLGMVSLAPVSQAAWFHVPQPFYFGMPQFEWSSSLTMIII
ALVSMVESTGVFFALGSVLKKDVTATDLKLGYRAEGLAVILGGIFNTFPYTTFSQNVGLL
ELSGIKTKRPIYWSAGLLMLLGLLPKIGALATIIPTPVLGGAMLVMFSMIGIQGIRMLAQID
FSDERNILIVATALGLGLGVSSYPTIFQFLPKTVQLFLGNGIVVASISSVALNLIFKGKDGL
DQA
>N-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28)
MAESRRARLKKQAGFEHHQLVLLAIILILGLSLVYKTISYFKQVVVETNTVALTKGPGLEY
QKLASLTKGSRLKVLDHKYHWYQVKTSTGQIGWVPDWVLANKYRLKPQSLSEATIVLD
AGHGGSDSGALSNSDKKEKDYTLKYIKQLASKLKEQGAKVYFTRDSDKFVSLKARPEL
AEKLHADAFISIHFDSSPNPNEASGITSYYYHKTSSKPLAYYISSHLDNLGLDNRGTEFG
DFLVIRDNTIPAVLLELGYINTSQDFALITTANYQDSATDGIVTGLKAYFNAKANNN
>hypothetical protein
MTNKRLLSWALSLILLVAIGGLAFSLYHKSRPQTTYLTPPKISLAQIKQQANLVVEAKILK
QKSDQLKLKGVTYTAYQVKVTKVSQGSLSQTQLTIYKAGFRDPKSGNYVLYRGDSLPL
VNHRYRFYLTKVKGHYLANSPKAMTNLKEN
>hydrolase, haloacid dehalogenase-like family
MYQTYFWDFDGTLFDSYAVMAPALVYGFKRQQVEVSKDLAYQLMRQKSRGEAFKSL
GKQWPQVDLKKAYADYKSYERAHLQSLQPFNGAKEVLVAIVAAGGQNFLLTHRNDEA
WQLLAKANLKQYFTGGVSGADNFARKPNPASLKYLCEKYQVNPARAAMVGDRTLDIE
AGHNAGMAGILFDPDDLIEDCEPDLRVTNLSAILV

>Esterase/lipase
MALLSLEKAARTLCEKTGQFVPTTIGLHVIKERLAFEQAQPIYLHPVKKMTETIWVSDEV
GQVVVRIFLPVDLQEAEQVPPLLYLHGGQFITGDAKTYDKLCRELTARARVCLLFVEYS
NLFEAKAPTQLLQAQAVFDSLASLALKYPLDLKRVLVAGDDVGGTLALNVALRAKASNL
PIYKLLLFYPVTNTNFDTSSYVSFAGGYYLTREQMKWSLESYRGDLDLKDLLLAPLQAS
FEQLAKLPETLIMTAEADVTRDEAEALARKMRDAGVDVSQIRFQGIIHDFVSLHILDKTN
ACRLAMNIATDWLAYRR
>D-tyrosyl-tRNA(Tyr) deacylase (EC 3.6.1.n1)
MRVVLQRVKEAKVEIAGQVNGQIKQGLLLLVGFTDGDGQAEIDYLARKIVNARIFSDEA
DKLNLSLLDIGGEILSVSQFTLYAQVKKGNRPSFTEAQDPKLAQLNYEKFNQALASYGP
KVATGVFGADMQVSLLNDGPVTIIYEA
>GTP pyrophosphokinase (EC 2.7.6.5), (p)ppGpp synthetase I
MEKEVIWTAEDVINICRQYMNEKHVAIVQKACDFATYVHKEQRRKSGEPYIIHPIQVAGI
LAELKMDPETVSAGFLHDVVEDTKVTLGDLGELFGHDVEVIVDGVTKLSKIQYKSHQE
QLAENHRKLLLAMSKDLRVIIVKLADRLHNMRTLSHLRPDKQRRIANETLEIYAPLAERL
GISTIKWELEDTSLRYLNPQQYYRIAHLMNSRREERVRYIDEVIVEIKDAIKDLKIDCDIYG
RPKHIYSIYHKMKDLHKQFSEIYDLLAVRVITKSIKDCYAVLGAIHTRWTPMPGRFKDYIA
MPKANMYQSLHTTVIGPNGRPFEVQIRTEEMHRVAEYGIAAHWAYKEGKTQGVNQTT
TGDKLNLFKEIIEFQKEATDASDFMESVKGDLFGDRVYVFTPKGDVFELPKGAGPLDM
AYTIHTEIGNHTSGAKVNGKIVPLDYQVKNGDIVEIMTSQNSAGPSRDWLDLVHTNKAR
NKIKRFFKLRDREENIVQGRDILHSALEEAGFKPSEVLTNDNLKRAYERRHLTTADDLY
AAIGFGEIQPTGVVNILTEEIREKAEAARKKEAERAILEDHQELEQKDKQTKQTEKKKAE
EGVLIEGIDNMLVRLSHCCSPIPGDEIVGYITKGRGVSVHRVDCPNIKSDDEANRLISVT
WNVIPDNRKFYDADLEVAGYNRSGLLNDIIMTISNNTKNMSSINGRVDHNKMAIITATVG
IKDLEHLERVMDAIKRIPDVYTVKRTIH
>hypothetical protein
MFGFDAIFAVVVGLIIGKLNHSIFFTLFFPLMFLLTYGLFFDSYVRYFGYIYLVLTILAFGIT
KK
>Ribosomal RNA small subunit methyltransferase E (EC 2.1.1.-)
MQRYFTNQEIKTEYKLDATTYHHAVVVLRMKVGSQFELVDAKQTVYLMELSELANKEA
KAKVISKLDSQVELPLAISIACGLSKGDKAEWIVQKGTELGASEFIFFKGDYSVAKWDQ
KKAKKKVDRLQKIAQGAAEQSHRTKIPTVRYVESLKELTTFEGAKFLAYEEAAKQGETK
ALVKLTQELRTKAQSLLAVFGPEGGISPAEVTLAKQAGFVLGGLGPRILRAETAPLYLLA
SLSYALELE
>Ribosomal protein L11 methyltransferase (EC 2.1.1.-)
MEWTQLLVKTSMEAKEAVANLLMENGAGGLQEIETDEQAVELVTYFPAEQNLQELALE
LESKIKALTTFGLNPGQATVSLAQVDDASWQTAWEKYYHPTRLSRYLTVVPAWEDYQ
AGDDLEKLVYLNPGKAFGTGTHPTTRLTITAMETVMRGGETVIDVGTGSGVLAIVAKLL
GAKSVLATDIDEEAIASAKTNLELNPQALPIELKVNNLLDGIEQQVDLILANILPEFLLPLM
PQAYANLNEGGKVILSGIISEQKAKVVQALTTAGFELEASFKLGDWYGLVAQKPRKGE
>hypothetical protein
MAKENLLTYLQAGLAKLDFDGDLDLSWQKAERTFTLELTFYAQNLAGQEIIDASGVDSD
QEIISFVDAILLYDEQKPAEFDTADYLACLPFAGKAGWNKQQAAGFLAYLQEVVDNGQ
ADLLDFLNEESETIFELKFDEEQLAKWVAKQNLSGNLAYPKF
>hypothetical protein

MENLSKRLVDKCVEAFIMGLEIYNKPTIRYRVEGFSFFICNAWELMLKAELLNRGISIYFK
DKPGRTISLEKATQLIYPDKNTRIRLNLEKIIELRNISTHFITEDYEFKYAPLFQACVINFTN
ELSKFHGRDITKYIPQNFLTIHANYEPLTNEEIKLKYPAEIAQKFIRQANEIDVLNTEYNSD
KFAINIAQKLYITKNRDKADFTVKVDKNSNNNVAIIKDLKDPSNTHKYSYSNVITAVNERL
KKKNIALNYEAGFNQFVLSLFIDFYDVKNNSKFAYKHVVGNQSSYTYSQQLIETIVNEIK
KDPQHFVESLKRAKK
>Inosine-uridine preferring nucleoside hydrolase (EC 3.2.2.1)
MAEYKMILDLDTGIDDALAIAYALATPESDLIGIVSSYGNVLVERACQNSLDLLELLGHPE
VPVFKGLSHSSTTDSFSVMPISAQIHGKNGVGEVEVPKAKRQVETQSGVDFFIEAAHK
YQDKLIIVPTGPLTNLAAAIQKDPTIVDLIGNVTLMGGALTVPGNVTPVTEANINQDPEAA
DIVFRSNLNLTMVGLDVTTRTLLTNKETASWRALGTLAGQKYADIVDYYIDAYKVTSPH
LGGCALHDPLAVAVAVDPSLVDLLKLNMKVDTKAPFAGRTIGDETRINDPALHTKAAVN
VDKERFLDLFMTKLTALFKAN
>hypothetical protein
MAYLGIILFCFWLLIISHKLTAGPKNRSFSYARAFLGLRLWYQNPRILLLLIALACLIFFSPL
KLVYLVFALAAYLTAFLCGRNFWNRIGPAWPGLILTLSSFTLAVITTVIYFHM
>O-succinylbenzoate synthase (EC 4.2.1.113)
MKITNYKVYEIDLPLRTPFTTSFATIKYKKTLILELVTSEGVKGYSECSAFDIPFYSEEFRA
GALALLTEQLLPQLLKVEIKHPDEIYSLFAYVRRNNMAKAAINCALWDLYAKEQNLPLFK
ALGGVKNKVEVGVSIGVQDSPEALVEVVRGYIEAGYRRIKMKIKPGSDYDYLKAVRTEF
PQAMLMADANSAYRLKDAELLKRLDDLDLIMIEQPLEPGDIVDHARLQAQLKTPICLDE
SIVSLDDARKMVELGAGKIINIKVARVGGFGAAKQIQAYAGQHGIQCWCGGMLDSGIAR
AGNLAIATLANYTLPNDIAACERYYAEDIISPLIKLNGTYVNVPTEPGLGYEVDYVNLKKF
TKAVYEG
>hypothetical protein
MEQTSKKQKVTSHAYTTLALIFIAYIICFIDRSAMNIALAYVGKDFGLSPAALGGVASAFF
FSYSLMQVPGGWLVDRFGSKKMIVFSLAMWSLFTVLTGFAWSVASLLVIRVLFGIGEG
GYPTASLKQIAEEFPTDAKSQATTVAISSNYVGSALAPVIIAPIIATLGWRNAFHIMGLIGL
VYVLIYWLKERPLTTNVTKSASRKFNWKLVDKKIWSFFFIIFGLNTITKGLDTWMPTYLL
TARHINLTGIAWLVPLPSIAAGLGAASAGFIMVRFFKGKERWLISATAALTMLFMYGMY
SSRSLAAIIIFEVLAYLVKSISFATTFGYVAQIIGAKTYGSSVGVVNFGGQVAGFVAPLIIG
AIVSATASYSMAFLFLVAAAGLALVSSLTVRK
>N-acyl-L-amino acid amidohydrolase (EC 3.5.1.14)
MFKVNGSQIEEFKIYLAKFIALKGISENGEGVVASLKFLTTFIKEQLNGEVELLPTSTNPAI
YTELRGKNSETVLLYGHYDVMEADDLGWQTPPFVLTEIAGRYYGRGCGDNKGQLLAQ
LLGIATYLREYGELPINLKILIEGQEEQGSRDLNKLVQKYQTSLLADIDYVLVIDGSKNVN
GKNVIRLGNRGLLAFELGVKVATRDLHSGNFGNVSENAALKLIGYLNKCYDFQTKRVKI
PEFYADIQPATTQELEWIKKLPLPQGQTGSNVAYYHKLMFEPTFNINGMQSGYTGAKL
KTIIPHKAFIRCDCRLVWGQSPELIRSQLEQLYHEELARGELEIKFLAEIAPSKTDSSSPY
LMELIKTIREVTGDCLIEPIMPGTVPNYVWTDTLKVPTFTVSFANFDQNNHSVNENISRQ
AFWNGIEIIYHFLERKKRYYGTNI
>transcriptional regulator, LysR family
MRLTQLEYFKDVAQTLNISQTAAELHVSQPAISRSIHELESELNLTLFIRKGKKLSLSQE
GLVFKQTVDNIFNQLTTGLQQLEQLKQHAQSRIIFTIANSTPLMVEIIEQLHHSLPETTILV
RVNSHSLPASNPEFIDYYYQLTDHPLAGYQSVPLLTERIGILVSNSSSLAKLSALSLADL

DTGLIKLSTSPFQALVDRYLTNHQVALPTVITTGDRKIICDMVSHNLGIAFIPEFSWHAYL
NSQQVCFVPLKDEQLQRTIYLNYIQPKESEFHHKVVTSIQKYFTKYQTRKEHFQ
>FIG00742360: hypothetical protein
MNRELFIKALLHSALTATPLKVPGNSSNFSLSNTWSIPNTSLVGKSCYFLGSSITYGAAA
NGVSFPDYLAKQYGLTIIKEAVSATTLADKYIEISWPYHNNSEKLAIANFFTGNQAQKAG
YSYLARFKKYLLKLPAPDLFVCQLSTNDSRNGVPIGKITPNSQLTNFDTTTTLGAIETLC
ATIKQNWNCPLTFYTCPRKDESYHQLIRYLQKLQSKWHFNLIDLANNVPFLAKLYSNPI
YMADDAHPTMAGYKEVWTPFFYHELAKIIGA
>Mobile element protein
MDQIRADQAELPKAQRYKIGDLLSRINLAKATYHDERKRIANPDDKYCEVKKVIRRICEQ
GKVRGRLTYGYRRVQARLDQLGIHLAGATIRHLMSELGVQVEIYNHHRNGKYSSYHG
RVGSVADNVLQQQFNAKDPYQVIHTDVTQVRLANQQWAYISAMTDEASKEVLAFQMS
SHPNQELIMNTLNELLDHLPDDAQPIIHSDQGWHYQLKYYTKKLKDHHFIQSMSRKGN
CHDNAPIESFFHLYKTELLQGLPPCNNITELRKLSLTYIDQFNNERISLKTKGMTPVGYR
NHALAV
>hypothetical protein
MTKYSPEFKAQVVVEYMTDGLSSYTLEKKYKVSARQIRKWIQQYQLQGIKSFKQHHRK
RIFSAKFKWHVINYYQTHAESFAQVAAKFDLLPSQISIWRQAFIKYGYDALKPHRKGRS
QKVKKQTKKQKQKLVEKKEIDQLREELARTKQELYNTKMDRDILKKSLALFGPSKPGK
KRK
>Mobile element protein
MRFDKTQVQNYDDLCGHLTIVMMTYDLLAWQELQEKDERTIGDLFYIMGETMPDLKLT
DVLVWFIKLLNRLVESEPMAPIKQLNETVDKFISSLPQSIASQLQKA
>hypothetical protein
MNKKPYSDRRWYDADFETVKNLKPLKHSEEYKKNAKAFFSGLYDREFKEEDLPIELQT
SKKEDTNC
>hypothetical protein
MDVLQTKLSQTIMAGITTGQSYDEMDRHLKTQVVKDVPIISVHPNCRCAVGAYWVD
>Integrase
MKNIEKNQLGKYRYYFKYKGKNYKGSWCNSLALAREQLFEVQAKVRKDGFIDNKQITF
RQVHDWFMEDHKLRVKESTLNKEIGDCKNHVLPKFGDYPIASITVDYCQSVYDEWAIE
LGRHDIIKMYASRVFKEAIRRDIISKNPFDYVKKPKEHKKPKEREFLERDEFITFMQAAQ
KFDIEKYLMVRLLAYTGMRRGELFGLTWADIDFNSHTVDINKGFTLGLDNKKKIDTPKT
DSSYRVLKVDTETIELLAKWKQQQALRIRTNNLRVKRDDKQFIFASQKKNEFTHPNKVY
KVINDLIKVTGIKKHITPHGLRHTWTTLFVEGSGEENLMIAQKQLGHRSVRTTQDIYTHL
TEDKKSKAIDIFLEGFK
>hypothetical protein
MGNRIKDLRLISGITQKELALKIGITQPALSNYESGRTPNSEITGKLADFFGVSVAYLLGL
SDSPKDPFFLDPVAIALDDRKLFEQMVDTTSQWDGSDYSELVPSAVKEKIGKNLHWFS
HSYEALVKHYDTPDAKIETEIEPGIERLLDLELDIIEKLSYFHYKMNNELRSGAKLDSSLE
YVLEFIAGFLEREYRWK
>hypothetical protein
MELQVITPEEIQVIIADELEKAVKDINTKGTTKQWMNKTELANYLGIGRPGLNKLIEKGM
PKVMFEKTYKFYVPDVNKWLTNRK
>DNA primase/helicase, phage-associated

MKFKELTKEKQWGVYRLKPRDNEKLGKIPVDPKTGIEGNQSKPETWLTYYEAKTYIEH
YNGLSVYMGNGLWLLDIDNIRADIDDYQSGNPKADNKVVAARKLTKGSYAEVSVSGN
GIHILFIYNGKEPRRRERLNNVELYSNGRHCAMTFNSLDPITDAEIKELTADELLKLIEFY
GLNEPKARTPTKKPNSVILTNCQIFRDEQGHYCRRLNADEVIEEMRRYNGPAWALFSR
GVETYENNDNSSADLALCNNLAYLSACDPYVMDEVFRQSALYRPKWDRLATDRATNK
TYGEVTIQKAINGNEKCYYKWVKPDFGYKFGFN
>DNA primase/helicase, phage-associated
MNNKALEIIKKGITDISSLEAVSALINENVAIPKIKPFNFAFIFSQLQKADNFRKPQLANAL
DTITPTYIRVYCVSKKEKESVIYTACWELLPKPALWLFDEEHHIKTYSKLEHGMEYMPD
NGAWRPLMKNEKNNLLSSFLRENLDRWKLPISLRASKQELGFQNDTMSFFKNGFTFE
REVDLPKEMAKRAHLVAFKNGTYNFKTGKMQEHSPANMIINAHDYNIEQGNTPETDYL
LTEMIGEKNILLLKQYIGYMFYRENPLQKVLFFYGPGGEGKSTLLNYIEENIIGRDNVSHI
TPYDLDGNDKFAAYSMRYKELNTFAELPQDNLNSRTIASIKSLTGNDEMSIQNKGEARI
NERLYAKHIWSANRLPGISTKEINNALLDRFTVIKFINGNTRPGGENASFWDRHNMKKV
KAERNAFVYSCLEAFRKMYFVDHTGFSISDEVKKYTSEWLNDNNTVKNFIDECCNQET
LGNKKAVAAYKELYQSYRKYCIDTLGSQPLNLKNFISELNRLGDSYEYRGKCKGWYYS
ENAKDPTTGKRKRALIGIELQTPPNLGFLKEL
>hypothetical protein
MYRSMILDFTNRLVFALPTKDNYKDMYDQAFSFARDHDGIVIICLDSHDLFTRRNHLIDH
FDFDTKAITQFEVEAY
>hypothetical protein
MQDIKITRLNKRAKLGGYETGVDENYSPIETFVEHATVWCAELAMTYKEVITNAGGVRK
GYKTIAIRHFPNLDDEWVVILDDKTYTIETISQESGIHGFDILDLKLKKD
>Phage terminase, small subunit
MAKNVMAITKGHETNEQKAIRQQALNEVKNAPMVAANPPTFLDEVAKAEYKRLIALLAD
SVSALDENILASYCATYSAIVACQKAINADGVVMTENKINPAVKQMDALTKTLRSLANDL
GLTPQARLKMAWNNAKQEDKSDPFKALVKS
>Phage terminase, large subunit
MNRALKYAYDAINGKKIVNKQIKQACQRFIDDLNRVDFPYYYDEAYADKVINFIEMLPTD
NREQLKLIDFQVFLISQLFGWRTTNGGMRFKTFLYSCARKSGKSFIIASLALVYLFLEPG
YSKQVLFTAGSLQQAHLAFDKAKNQLQNVSRVSSYMRRRFKVTSEKIFDTETDSFAVP
LATNGTNNLDGYGADLAVMDECANISNKVFEDTKKVLESGMIQNENGLLCMISTAGFN
MTSSFKAQCDYACDILSGKVTDDRYLALIYSLDDPEEIYDPETWVKANPILANKKIAETM
QAKIQADLDTATKQGDISNVLVKNFNMWTPGRVDAYIDPKDWNKNLITPPNIYGKDAYI
GIDLSTRNDLTSISWLIPLDDDDLTYYADSFSFVGTANGIERKERTDEFNYRLAAKRGE
CSITELESGIIDVDAVFSWLMNFITENNLNVLGIGYDPWNSNNLIARLEKQLALPLVGVR
QGGLTLNVPTRQFKDLLLDGRLKHRDNKLLEYAFNNAVVKLVNNNWILNKVDRMSGK
HIDPAAALINAYVLGMNYFDEQLKNKERNEYFESDEFSF
>hypothetical protein
MKAMNFLFNQGTLLSLGLIMITIATYGLYGFFKALGVAGLFLVVIAVILNSERK
>Phage portal protein
MSFFINNQATTGSADPVEQVLIGFNDNANFNISNNVLSSSSLFTVIQIIANDVASCKLNC
DTPNIGNNFSKRRLLFPVVSNMLLYGNSFVENLGSGNFRLVPNDHIEIEYDDVTNQVRY
KYNNHEQKARYIPNEKLLHFFEFTKNGVFGLSRVQALKGTLKAEKELNNLLGGYLTSGI
HGTTVIKSVNSDLSREAREKVRKAFDEATTGDKALNSIVVGPDLDVSSLELNTDVLNVI

DKVNEVVIRRIASTFGLPIEKLGLENVHSNALQSDKTSSSYLRGTLQHIFDCITSELTEKL
GNKNFSFDTSALTSVDPKEMLEQAANAYNNGLITLDEARQMINLDPVGNTYFKE
>Phage head maturation protease
MLEKRFLINAELRANNPQKGAEDSPDAKETAEKTVQGYAVLYEQESELLTGKDGLKFT
EVIHKGALDNVDFSNLVMIYNHDYKDVLASVKAGTLTVELDDKGLKFKATLPGTSVSDD
VYKNIQAGNLDSMSFGFTVSDDGSAWTQTDDGNYKREVNSIANLYELSVVTLPAYPDT
NVNVDTRSLITNTKENKDMFKSVNTNDEVTPVQAMEAYIRAAGNLDAETRSALTTESG
KVVMPKEVITPAYELKDDGLNLAEFATVENVSMGSGSLPIVPADDTTTLVTKTEAEELS
DTDLALKEVEYKVETRAGRIVISDEAKDDSAVDIIALCKKQLKRLVRNTDNQNILKLLATL
PGKTAANLDDLKTIFNVNLNSNLKKTSMWIVNATAFNTLDQLKDNEGRYLLQPDVTAPS
GFSLLGQPVFMVSDRAWKAAITGKANMILADMSQAVAVFRRNQVSYNWEKFDSYAQ
GLAIATRDDYKLIDTEAGFAITFGAAA
>prophage Lp3 protein 19, head-to-tail joining
MVTVEDLKNSLRIDNTEDDKLIQTFINSAEGYIKSAVGKADGLESDDRYTVAVIFLAGMY
YENRAETEVKIPYQIRSIIGQLQAEYQ
>hypothetical protein
MDAKKTAKNVRDYFNKQYDKDCMLAGTTLKAVRLNGMPATRNYENHNEVAFINAAEA
EQRLEWLRQALDLIEDNFKNDDTDYVKLIELAYVQKLRQVDLAQSWGVTRRKLYDLLL
SACIILAGAMKQVSGNRINWIA
>tRNA-Ser-CGA

>Glycosyltransferase (EC 2.4.1.-)
MKILSVIIPCYNAQDYMEHAIETLLTGGERLEILLVDDGSKDNTAQIADNYEQRYPEIVRA
LHQENGGHGKAVTTGLHAATGFFVKVVDSDDWVDENALREILDFLEKESLKKDAVDM
LISNFIFDKQGVKHKKAMEYSSFLPTDRVFSWDEVHFPLGKYLLMHSVIYRRDVVVGQ
AQLELPHHTFYVDNLYVFEPLRYVKNIYYLDVNFYHYFIGRDDQSVNEQVMIGRIDQQL
KVNRLMINFYAKNEAQLDPALANYMFRYLEIITTVSSILLLKDGSKESLRKKTALWNYIKR
KDAHLYRELRWGLFGTGVNLPWKLGRKTALGVYKLARKMYGFN
>hypothetical protein
MSLYQLYAGCFFFVEEVDKTSFQYLLYLRLPYYLADRVGLDPADENLLPALNYSFDYQL
NSTHDRFTQLLAHIHLNHHNFTLDLLKMPELCFITKHRCKEVKVTDEDSEYFGQVMQE
ALPDPEAVYLDPNEIAYKLTPEDMPLVEWTLNFEPLDTYEELINYYHNDN
>hypothetical protein
MAIDDTYCMSPEEERKILDRINKKIESNFEFKKREEATDKYLKDHNII
>hypothetical protein
MDNKNQDVEVEKINEEVENEQDNAGKIEEDKAAKTVEKLQKRLEKLTGDKHDLEEELV
NTKSCLKHISRVKRP
>hypothetical protein
MKKLSEEDKTKQAQDTKDKELASLHAELVRTKALSEASEVLKEQGLNVTADILNNLHSS
>hypothetical protein
MNAANPKFAKYSNGPIEGINSKIKLLRRLSFGMPNFTNMKKRIILWLNFKPKKKNSKKIA
LIS
>hypothetical protein
MDMAMEVSFDYKGINYFIEPDAKSDKWMIFCSLKPDVPLFMTMNEVLDMKIDDMPLKE
VLPLVTNAIY
>hypothetical protein

MNVTYRDYNDELELRKIEKEFVLEIADSLTPDNSDLNGDGLKLEEIWDYA
>hypothetical protein
MKLYKWKKGDGKQLLGCLVAVLVGGIIALPLLILGVTSIDDIIIPVFFIVGFITFIGGIFKWIF
RL
>Mobile element protein
MLVNYHVPKHVYIVCGNTDLRKGIDTLAILIADNFGLDLYDDSLFLFCGRRNDRFKALY
WDGEEFILLYKRFDNGRLTWLLT
>hypothetical protein
MGKKFKGFLNKEYVGSIEEVNKHSDQFYKDFSLYIYDKVQVRKEYTKLSRFIDLLKISDV
YITDMGDILNLKYDSLLDWILILNVDEEINNKINRMRNELKLLQNIKSEYLYDSSSKLHDFI
QKCNYLFNNEYLKQNISNIENIISITSYYYKQIDDVEINLLNKQNENPNLKFYDFYFFEIVQ
MWKFSDAKISKNNDGKIYVNNINDISNAFVVTRNLNKDNRNLARNLSLSKKMGSHRIDD
YLIGLGEELFLGLDESDKKVYLRLDLLKNTMLYDITILRWVECLRSLRKSVIYKYEKYRKC
FFNINELNIEDVETIQIAKILSFHNDFLDTPFLVLDKEIFVYAPALIIGDIVHYLNIAIKYSRDK
SIQNMRGNNFENTIGRLLDYHGNDFCKAEEDVYGNTIKLIYTENSRQHEFDLIASDDDE
RVVQIECKTFMDPFSYRDYRIGIDKMFSGKKNNGYLETDFMHFESLKNNGHNVIINNIK
NHSDISKNKGVSTYFNKPRNWDDMYMMFISNYIFPSRMVKNWSNKYGINFVHWFEFN
RLIKNIPLDKDWFIVNGVRVLNSEIETNITDKFNRNNIKEISDGQYISELVNHRSQKCKLK
FNEKEWIPGVFVKCYE
>Mobile element protein
MDQIRADQAELPKAQRYKIGDLLSRINLAKATYHDERKRIANPDDKYCEVKKVIRRICEQ
GKVRGRLTYGYRRVQARLDQLGIHLAGATIRHLMSELGVQVEIYNHHRNGKYSSYHG
RVGSVADNVLQQQFNAKDPYQVIHTDVTQVRLANQQWAYISAMTDEASKEVLAFQMS
SHPNQELIMNTLNELLDHLPDDAQPIIHSDQGWHYQLKYYTKKLKDHHFIQSMSRKGN
CHDNAPIESFFHLYKTELLQGLPPCNNITELRKLSLTYIDQFNNERISLKTKGMTPVGYR
NHALAV
>hypothetical protein
MTKYSPEFKAQVVVEYMTDGLSSYTLEKKYKVSARQIRKWIQQYQLQGIKSFKQHHRK
RIFSAKFKWHVINYYQTHAESFAQVAAKFDLLPSQISIWRQAFIKYGYDALKPHRKGRS
QKVKKQTKKQKQKLVEKKEIDQLREELARTKQELYNTKMDRDILKKSLALFGPSKPGK
KRK
>hypothetical protein
MLKQVLVSPKEYGNDKAVAASEYNAVITLMSGKQIRVSGNAYTYKDALVNLYENNQST
FEEASVKSLLLELS
>Mobile element protein
MGTTILSFSDRIVIETLRHEKRSLRYIADYLGFSKTTIFNEIHRLKGEYHATSAQADHETK
LSYRGRKCSLTANLKRLIEDKIKIQKWSIEQVAHVVRIAYKTIYNWVDQGLLDIRMTDLP
DHGIRRKRAKETRGTFSHGRSIEDRPAEASDRNIAGHFEADTVLSGKRKGQAVATFVE
RKSRLTIVKRLGGRDSTSMTKAILELANQLGTNLKTLTVDHGKEFANYQSIEKNTGVPL
YFAHAYSPHERGSNENRNRVLRRFIPKGQPIEEISDDELIQINWYLNSRPLKCLNWRTPI
EIFLLNLRH
>Membrane protein involved in the export of O-antigen, teichoic acid lipoteichoic acids
MKLIRNYLYNAGYQMLALIIPFITAPYISRTLKPHGVGIYADTNAWIQWFVLLASVGIALY
GNREIAYVRDDKEKLSRTFWEIQVVKMIMTVISYIVLCFFLNVYTQYKNYMWIQSLNILA
VAFDISWFYMGIEDFKRTVIRNTLVKLMSLALILLFVKSPNDIGVYIFLTTLSVTLGNFTLW
PRLRILLVRVGLNTLRPLKHLKTSLALFVPQIATQVYLILGRLQ

>UDP-galactopyranose mutase (EC 5.4.99.9)
MTKYLVVGAGLFGATFAYEAAKRGHEVKVIEKRDHIAGNIYTKEVNGIQVHQYGAHIFH
TSNKKIWDYVNQFAEFNRYTNSPVANYNGEIYNLPFNMNTFNKLWGVATPAEAQAKIE
EQRSILGDKRPENLEEQAISLIGTDIYEKLIKGYTEKQWGHKAIDLPAFIIRRLPVRFTYD
NNYFNDDFQGIPMGGYTQIVEKMLAHENISVELNTDFFANKEAYLKDYDKVVFTGMID
QFFDYKLGELEYRSLRFETEVLDEDNYQGNAVVNYTDAETPFTRIIEHKHFEFGKGEK
GKTVITREYPANWQRGDEPYYPVNNKENNTLYKDYKELADREAKNVIFGGRLGQYRY
YNMDQVIMAALVTVEKEFEG
>Phosphoglucosamine mutase (EC 5.4.2.10)
MKYFGTDGVRGIANESLSPELAFKLGRCGGYILTKHANGHTPRVLLARDTRISGQMLA
QALTAGLLSVGIEVYDLGVMTTPGVAYLVRTEGMDAGAMISASHNPAQDNGIKFFGND
GYKLSDEKELEIEELLDAPEDTLPRPSSEGLGTVSTYSEAAMKYLRFLQNTIPDDLAGM
KVAIDGANGATSALVLQLYANVGAEFVTMATQPDGININDGVGSTHPEALAKFVVDQE
AQVGVAYDGDGDRFIAVDELGNIIDGDKIMYICGKYLASQGKLNQNTIVTTVMSNIGLYK
AMAAHGLNSVKTKVGDRYVVEEMRKSSYNLGGEQSGHIVFLDHNTTGDGMLTSLQLL
NVMKQTGKKLSELAAEVTTYPQELVNVKVTDKQEALSNAKIKAKIAEVEETMAGEGRV
LVRPSGTEDLLRVMAEAPTKKLVHQYVMEIVNVVEAEMGI
>Glycine betaine transporter OpuD
MKNKLVKQIDWLQTLLPLISIALLVIVFLLYPQKSTQVLHAIRSELGNKFGWLYILLGISMF
TTTLVVAFSKYGHIVLGNKQAPIYGNVKWGMMIFTSTMSADIIFYALTEWTMYAREQFI
QEKVGGVGLWTLTYSLFHWGPIAWSFYIILAVAFAFMIHVRKKDKQKFSEAVRPLLGNK
VDGVWGKLINLIAIFALIAGTATTFSVSMPLLSAALSKVFGFHDSSFLAVIILLIIATIYTVNV
LLGMDAISRLASYCVVLFLLLLGFVFLCGGKELFILNNGIEAIGNLVQNFIGMATWTDSLK
TSHFVENWTIYYWAYWLVWCTATPFFIATISKGRTIKNMILGAYSWGLSGTALAFIVLSN
YGLGKYLLDGLAVEQLINKGNSYIEIAIKVLDTLPAHNIVLMLLILTMMGMYATVFESITMV
VSFYSYKELSEGQLPDRRIRAFWAITFIILPLALIFMKQSIYSLQSVAIIFAFPISIIVLLVVVS
FFKDAKKYLNNN
>Choline permease LicB
MNKVKKAIIMAAGFGSRMQPLTFKTPKPLIKVNGKRMIETVIEGLHENRITEIYVVVGYLE
EQFKYLEEKYSQLKLITNPYYDTANNISSLYVASNVLDDATIILDGDQIIKNSQILKAEFDY
SGYSCSKVNEETDEWVLQLDSHDVITSCNRNGGKYGWRLYSISRWTKDDAISLKKQIQ
YEFDKKNTRDVYWDDIPMFYYFENYKLKGYRISATDIEEVDSLTELIQIDSSYQEVIEDE
K
>Cholinephosphate cytidylyltransferase (EC 2.7.7.15) / Choline kinase (EC 2.7.1.32)
MNKKEAAIIRYLKNNSSKATQRKIAEGVGLSLGSVNNTVKKLVANNILDEELNILDDSMF
EKPKRAIILAAGYGMRMVPINTEYPKGLLEVKGQPLIERLIKQLHMVGVHDIKVVVGFMK
EQYEYLIDDFGVKLIVNQHYSDRKNLYSLNLACQQQSLEKTYIIPCDIYAKENPFKKEEIS
SWYMVTNQLTKNSDVKVSRNFDLKTIRTAEIGNKMIGIAYVSKTDGKYYTNRLQDFVES
SEHDKEYWETILNHNGTYQVLANIMSVDKVYEINTYEELRNLDSNSNHLNNDTISLLSEI
LNVDNNSIKDITVLKKGMTNRSFLFSCEDKKYIMRIPGEGSNELIDRQKEAKIYELLGKE
KITENVLYINPDNGYKLTEFVAGARNCNIDDNSDLEKCMETLRKFHSLKLKADFEFNLY
KQIDFYEKLRDANSNYRDYDFVKERVLGLKDYIDTLTKDYILCHIDANPDNFLIDKAGNV
MLIDWEYAAMQDPDVDIAMFAIYAMYDRKQTDNLINIYYQNQCDDETRTKIYAYMAICG
LLWSNWCEYKNSLGVDFGEYSLKQYRYAKEYSKLVYEILEGKENE
>acyltransferase

MNGSKHNFDRQSRRIDSLTGARFLAIMIIVFSHFELLKDYGSFGRTYWNWWHNATMG
VDFFFMMSGFGMMLSSIRRDPSGSKPVGGLHGSLKFGIHHVKKIYPIYFTMLLIGIPYAV
LTAVFEYGGNLIKQIIKNVIFFFGDLTLLQTATGWEAFSHSLNGVCWFLSSLFCIYLVSPF
IMQWLKRHVKTVRTAIVGVVICIICSYLLAILFTNIDDKTWFFDDLCYGSPYRRVFYVIPG
MLLAQVYTFHQQDESYCMPVLFSSGVFEYLSIGFSMVWFFWWHQFSETYWVYIVNMI
VVACDLYAIAIRRGKISHLFETKPMIYLGNISMYIFITHYLIRMYVDFIVRMTGIGSLPVAIC
EVVIILLLTFVVSAWLNKKNRKLAYFDRIVDSIRPKKKTQIE
>ISSth1, transposase (orf1), IS3 family
MNRLELLYGYQKTESGLKVYADMHHVNQNTMTRWIRQFLLHGLAGVRHRSFNHRYSL
NTKIEAVKEYQAGFPGEKIIEKYDIRSLSQLKQWSIQYNNDELKTKSRKRVRKMGRKVN
FEEKCKIVQWVLDHDNDYQGATIQFNVSYQRVYSWVRKYHEATDSWEALKDNRGHR
KPQKALEEMTEKERLEAEIKRLKKINREMELEIAFAKKLVEIRNRGVKRPDDIKRFKN
>Mobile element protein
MEAIYQIEKRHQNCAGHKKVTAILNNENGKNQEIPEYTFSFDFRINIKRVRRIMRKHGIK
ADVRQKKRSHKKEQQQYQEDNLLERKFIQIAPNLVWVSDTTELTYGRNNKVRLHVVLD
LYGRYALSYHISPTETAEAAIEAFKRAAKQAGTFAPMIHTDRGATYTSNDFNNYLTSNN
SKHSLSAPGTPADNAVIEHYWGDFKYIWMAHHSHPQTLAELKDLVEQGVEYFNTVEIS
SKRNNPTAGDFRNEAV
>Mobile element protein
MVQLDETPFQVLDLDLDQSHGYFWSACSTEEFSPHQVSYFHYAPSRSGKVITEILGDD
FSGGIMCDGFSGYSDNRLPEAYWGTCLVHINRQFKRLLDPKITRFQGQSIASDAVRILA
KVFHIEKRLKYSSASEKAAQRRQHLKAMIDDFYQLIEEALTKSPLKPLRNAIKNALKLKK
RVYQMFKHGELPLHNNHNEQLIRPTTLVRKNSLFAKSTAGAKANAIWHSIVKTAKLNHL
DVFKYLETLLSAFTKRETPEIEAYLPWAREIQESCKA
>transposase
MAEKITLEEALRENKELKVKIAELTEMVNYLQKQLFGKKTEKLSAGQLDLFDDKEQETA
PTNDSTVEKLTTTVTSHQRKCKSKESRKQWLDCLEQVEELHQLSKQELTCDQCGQM
MTVIGKHEQYREVKLVPAHLCCKVVCTETARCEHCQDPETGNDVLVQSKTPQPLFPH
SYLSSSVIAEVLFEKFGLAVPFTRQTQYWERLGLPITSKHMARGLIEAGERFMNVCVKK
SR
>Mobile element protein
MAMVIAENYGLELDNDSLFLFCGNRNDRFKGLYWDGEGFILLYKRFENGGLRWPRHR
EDAVALTKSQIKVLLEGISPLPQKRIRPAIKGPGY
>hypothetical protein
MKQDFVEVDIPRTKKPAAKRSRLAATVSLHQVKINVYQEADSALVSEIIKAVAAYAG
>Lipopolysaccharide cholinephosphotransferase LicD1 (EC 2.7.8.-)
MSEEITLKESQQLVFEILKFYKQICDENNLTYYIAYGTLIGAVRHHGFIPWDDDIDVLMPR
KDYNKLVEIMSNIDGRYKLVSYENNNKFTAPLPKIIDTKTELVQKYDFYERVPLGVYIDIFI
LDNAGNDFQEASSFYRNVYSIYRKWLKSDLKLFLNNKSKFRSILGFVKRLPYKIHGISYY
LRCYTEYSSSYIDVDTKYVAATNAGTPEAERNVFLASDFGEGIDILFEDVYFRAPNNYQ
KLLEVEYGNYMKLPPKEKQVSHHMYELHWKN
>Glycosyl transferase, group 2 family protein
MKCSCVVLNYNDSDTVIEFIQRIRDYKNIENIVIVDNMSTDGSFERLSNLYSEESKINILQ
TDENKGYGYGNNYGVSYVYKTFEDKYTLISNPDVIFSEEMLNDLLDTIDNNNVSVVSGV
QFTKEDVEIADKAWKVPNAFEYIFEGSKLAKLFKVSSLYKGEHFKSTLSRVDCVPGAML

LVNTKDFLDIGGYDDRMFLYCEETTLGYKLMKAGLSTLLVNNADYIHIGSTSTGKELATK
KAQKRLIYDSRLFFLKNYLRVNRLTYSIAKWNYNRIVQKMK
>UDP-N-acetylglucosamine 2-epimerase (EC 5.1.3.14)
MYKVMKKIKVMTIFGTRPEAIKMAPIVKKLESDNRFEAITTVTAQHREMLDQVLDIFKINP
DYDLNVMHNRQTLTDITTNVLIKLDEILADAQPDIVLVHGDTTTTFAASISAFYHQIPIGHV
EAGLRTWEKYSPYPEEMNRQMTDVLSDIYFAPTSQSKENLVMEHHHNNVFVTGNTAI
DALELTVDDKYNHEILAKINPSNRVILVTMHRRENQGEPMRRVFEVLKNKVKKHQDIEII
YPVHLSPAVKEVAFDVLGNVERVHLIDPLDVVDFHNIAAKSYLIMTDSGGVQEEAPSLG
KPVLVLRDTTERPEGVQAGTLKLVGTDTEKVDSALEKLLNDNEEYERMAKAKNPYGD
GKASDRILQAISYYFNQSDRKPEEFEI
>FIG01117555: hypothetical protein
MKKKVLFLMNVDWNWIKQRPHFIAEGLSDSYDVTVVYRYRYRRQGYQKRREKNFSG
NVRKLYLIPKLDKFKFLKPINDFIRKYTIRKLLKDEGFECIYVNSPKQEMLIPKWYTGKVIY
DCMDDHIALAENKKDRENCYVLEKKLLDKCDELVVTSNTLKKVLMERYPVMNINNITLIR
NGYSGEIFDVKKMNTKYKDDKFIFAYFGTISEWFNFDVIIESLKKFKNIEYWIIGPILGKVQ
FPKDSRIKYLGTVEHDHLYEATKKANCYIMPFKLNDIVVSVDPVKLYEYINFNKNIMTVE
YDEIERFRPFVNFYKDTPSYLDAIEKLLTDGKIKYSNEDRIEFLRNNSWQKRVEQVVSVI
EN
>Membrane protein involved in the export of O-antigen, teichoic acid lipoteichoic acids
MSRGKIKIHSVKYNFFMNIILKTSTFIFPLITFPYVSRILGPGPNGDISFATSVISYFSLLAG
MGIPSYGVRKCAENRDDINKLSKIIQELLIINIIFTVISYIILICMVIFIPKFSDNASLLYIMSITII
FSTFGVDWVYQAIEQYDYITVRNISFKIVAVILMFVFVHKPSDYVIYAGITVFGNVGSNIL
NLLKLPHFVTFKRYEVYNLKQHMNPIVMLFLYNATTTIFTNLDQVMLGFMSNSKEVGYY
AATVKIKNILASLITSLGAVMLPRMSYYLGKGDKKKFSSLVRKSFNFIFISAIPVTMYFFIE
AKPIIGFLAGPKYYNSVSILRNVLPSIFFIGLSSVTAWQLLIPLKLEKYTVVGSIFGAIINVV
FNMLLIPIVGGNGAAFATTVAEFCVLLTHLIVLRKRIVESKLIDSNEFIRAVVAAVVATLML
ILFNSKVIFSNLLVDCVVTGFIYFVIYIVMLLILKENILIGEMKKFLSNNNIGG
>hypothetical protein
MVPIRRELHILNLLFVILVMAEYKFFMLIPSTSVIYKYNVGETQLLKDILCILILFVGLNYIRN
FLNSLIYVKWLFCLGIVYALSMTMQGMLYGNYSAALRMLSNLLYMPLAYVVCILILKIND
SASFFLKTLIFSNVIFSTLVLIQSVIYNKSSKLFLTIFWVDNVPYMRNGRIRITEAAAIILLMF
IISYCMLFISRELWKINFANCILSLITLLYSSQTRQLTLIAYVVMILVIIYKLMQYAQRSITGFI
FFIFFIIVAIACFLIFILFNGFNSVDIFSLVGISKTEKSYVARYGAVHFFIHQLHWLYGTIVEF
HHGQGIAGLYYTADIGVLGNMVDYGILSGIVYVLMIVDGLIRSLFIDKRGSILLVSMIIMLM
MQMGSISVNYNTYMPLAIVFSLIDFYIYKKNQVANQSVFDDIFW
>Alpha-L-Rha alpha-1,3-L-rhamnosyltransferase (EC 2.4.1.-)
MNNHKRTLILMAAYNGEKYIKQQIDSIINQTDKDWDLWIRDDGSNDNTVNIIKNIQKYDN
RIRLFQNTTDVHGAYHNFYALIREARNSNYEYYAFSDQDDLWKEDKLKQEIDIIRNSDE
PALVYADMEVIDENNHVIMDSIDKVMGISNFGKENVFFAHAYIWGGTTVFNRKLMSMLE
KVVGSRLPNESVMAHDSLLMKLALSTGKVHFCNTPTIRYRRHSANVTAGDRINMNLKKI
MKRVLDKNNSVPAVHALVYGETLTFIYQMKINNIDSKINLDKMQEAILNGGIKGVGYLLKI
NFKRKQVARTIGTYLIMLTGKYKKYLFEYKKELKF
>dTDP-4-dehydrorhamnose reductase (EC 1.1.1.133)
MKYFVTGVAGQLGHDVMNELNKRGHVGVGTDLAPEYAGIQDGSYVTTAEYVSLDITN
NDEVQRVIEEVNPDVIVHCAAWTAVDAAEDEDKQAKVRAINVNGTQNIANAAKKIDAK
MVYLSTDYVFDGQGTQPWEPDCKDYKPLNVYGQTKLDGELAVANTLDKYFIVRIAWV

FGVNGANFIKTMLKLSENHDELTVVSDQIGTPTYTYDLARLLVDMTETDKYGYYHATNE
GGYISWADFAKEIFKQAGKDVKVTPVTTAEYGVSKAARPFNSRLDKSKLVEKGFEPLP
TWQDALKRYLEKITI
>dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)
MFKNIIVTGGAGFIGSNFVHYVVNNHPEVHVTVLDKLTYAGNKENLAGLPADRVELVVG
DICDKELVDKLVQKADAVVHYAAESHNDNSLQNPEPFIQTNIVGTSVLISACTKYGVRF
HHVSTDEVYGDLPLREDLPGHGEGPGEKFTPETPYKPSSPYSSSKASSDLLVRAWVR
SFGLKATISNCSNNYGPYQHIEKFIPRQVTNILSGIRPKLYGTGKNVRDWIHTNDHSSAV
WKILTEGKLGETYLIGADGEKNNKEVLELILELMGQPADAYDQVKDRPGHDLRYAIDAS
KLRNELGWQPQYTDFKEGLKATIDWYTNHRDWWEAQKAAVEAKYAKNGQ
>dTDP-4-dehydrorhamnose 3,5-epimerase (EC 5.1.3.13)
MGKLVVKETRLQDVKLITPAVFGDHRGFFTESYSDRDFKEAGIDFDFIQDNHSLSTQAG
VLRGLHFQRGKAAQTKLIRVATGAVLDVIVDLRKGSPTYKQWEGYILSESNHRQLLVPK
GFAHGFVTLTDNVNFLYKCDNYYDAAADGGISFKTPELNIDWPIDLDKAITSEKDAKQP
TLTEFEKDNPFVYGEI
>Glucose-1-phosphate thymidylyltransferase (EC 2.7.7.24)
MKGIILAGGSGTRLYPITRGISKQLIPVYDKPMVYYPLSTLMLAGIRDILVISTPEYTPLFE
ELLGDGSEFGINLSYKVQEQPNGLAEAFVLGKEFIGTDSVCLILGDNIYYGSGLSKLLQE
AAQKESGATVFGYHVNDPERFGVVEFDEEMHALSIEEKPAQPKSNYAVTGLYFYDND
VVDIAANLKPSPRGELEITDVNKEYLRCGKLDVKLMGRGYAWLDTGTHDSMLEAASFI
ATIQKRQNLKVACLEEIAYRMGWISKDKLVELAQPMKKNDYGQYLLRLAQED
>hypothetical protein
MISFEELHKKTLPEKKKKFVKVDFVSYYLWRPLCDFISILLMPTRVTPTMVTIFSFYTSLL
SLIVFILIPNKLGALLGYLLIWIWNISDGIDGNIARYKEQFSKSGDLWDATAGYMALVSFY
FGAGIIASNELSTFNVPYINDKYYVIMGAIAAICTIFPRLVVQKKNVVYGDEAVKNLKDRS
YFSLTRIAMVNINSINGFAGLMLLVAIIFNITNIFTICYFIIQIVFFVGAMTVTLRSV
>dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)
MSVLDKDFQTLFDSTLEYEKFKDANFLITGATGLVGSLLIKSLLTLDDKFNLNLKIFAVVR
NLQKAENIFGELCRTKSLIFIKADLGSDNVVVKDDIDYIIHAAAITTSKILVEKPVEAMDIAI
NGTKKVLELAKKKQVKKMVYVSSMEIYGQINSVGKTSEPELGYVDLSKVRSGYPESKR
MCELMCNAYASEYNVNVVSARLAQTFGAGILPGENRVFAQFARSAMNNNDIVLHTEG
KSEGNYVYTIDAVLALFVLLLRGEKGEAYNVSNPENHVTIREMAEVVATNFSENSKVVI
DVPSDSKKYGYAPDTHLWLDNSKLVKLGWTPQNNLVESYKKMIVWMKEENI
>2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (EC 2.7.7.60)
MNIAAIFAGGVGTRMHTKELPKQFLKIHDKPIIIHTLELFEENPEIDMIVVACVENWIGYLN
ELISKYNLRKVQKIVKGGKSGQESIYNVLKAAEELGNRKEDIVLIHDGVRPLINQKTISDN
IASVKKNGSAITSVKVKETILIVDEDESITEVPDRAASRLARAPQSFYLDDILTAHERAIAE
NKFDFIDSCSMLQYYGKKLYLVDGPQENIKITTPDDFYTMRALLDAKENAQIYGIEE
>Poly(glycerol-phosphate) alpha-glucosyltransferase (EC 2.4.1.52)
MKKILMVCEAFGGGVFTYVSQLCNDMVDNFDVYLAYAVRPQTPKNYRDFLDKRIHLIE
VDGFNAGNLLNIANDMKVIKALRKIEKEVQPDIIHLHSSVAGGIGRIAFKGKNNTVLYTPH
GYAHILMEPSRKTKVYGMMEKFLGKYSSATILTCCESEDDVTSKLTKNHTYIETGVNLK
DLSESLDGIRPVENEKFTVFTLGRVTGQKQPKVFNRIAELVPEARFVWIGGGEQENLLT
APNIEITGWKDRKEALAMAKGADAYVLCSRGEAIAMSLIENMYIKKLVLVSNVMGNRSV
INDGVNGYVCNTPEEYADRIRSAMQNFPTRLPENAYKDVLDIYNTDVMKEKFIRYYNSL
>Undecaprenyl-phosphate galactosephosphotransferase (EC 2.7.8.6)

METKQITLDTTKIDQRYVYRTLKRIFDFCASLLALIPLSIIFLMIALLIKLDDGGPVFFSQTR
VGRHGKLFKIYKFRSMRVDAEDLLEKLMEKNQVDGPMFKMKDDPRISRVGKFLRKYSL
DELPQLINVLLGDMSLVGPRPPLPREVAEYTDYDKQRLYVTPGCTGLWQATERNNVG
FAGMVKLDLEYIQNASVWLDLKIILMTVKILFVPNGAY
>Manganese-dependent protein-tyrosine phosphatase (EC 3.1.3.48)
METLVDLHCHILPGIDDGSRDVETSLHLARQAVLDGVTHILATPHHLDNDYVNHARDVE
QLASDFQAKLDQEGIKLKVFPGQEVHINGDLLDHYDDLLGIDADKHYMLLELPHNDVPR
YTKQLIFNLKKKGTTPVIVHPERNKVIQDHLGMLYEFIEAGALAQLTSTSYLGGFGSHVE
EISKQLVEHNLVQILASDAHALKGRGFVLSEAFDKLAKEFGQDKAQSFETNAERLLNGQ
RVIARDFSKVEKKKRFFFF
>Tyrosine-protein kinase EpsD (EC 2.7.10.2)
MGLFRKDKNQDANTMETGVKMITVTDPTSVDTEQFNTIRTNIKFSNVDKDYKTLMVTS
SNMSEGKSTVSANIATAFAKQGLHTLLVDSDLRRPTINATFGIDSPQGLSNYLSERNFDI
NSIIYKTSVKNLYVMPSGPIPPNPSELIGSKRMAELIKKLSEQLDLVIFDAPPVLSVTDAQI
VSTNVDGTILVVRANKTEKNAVKEAVRLIKQVGGHIIGTILNDVEVKGGGYYGYYGYSK
DSKKKK
>Tyrosine-protein kinase transmembrane modulator EpsC
MVKENEETVIDLGRLLLTLKKNLFSIIAWAVVGLIISLVVLFIFIEPKYSATTDILVNQKNDN
VQTQYAAQQADLGAVTTYEDILKRSVILSPVLKEVRARDNYKGSLGDLQKSVSVSNET
NSKIISVTVTDKDAYTAADIANTVAKTFKEKIVKMMKIDNVTIVSTAKPNATPVFPKKTLG
ALVGIVLGALIGIAIAIVRELTDKTVKDMDFLTDEVGLTSLGTVFHIEDDDAFAAVEVLNE
NREATRTKRV
>Diaminopimelate decarboxylase (EC 4.1.1.20)
MTEALEVNQAGHLTIGGVDALKLAEEYGTPVIAYDVSKIKQQIADFKKAFIEAGVDYKIS
YASKAFSALAMYRLVADAGLGCDVVSGGELYAALKGGMPPADIEFHGNNKLPAELEM
AVDSQIGCIVIDNFQEIKFLDKLLRRKHTTANVVLRVTPGIEAETHKYISTGQSNSKFGFD
VQSGQAKEALASLIKNPRMNVEGIHCHIGSQIFQARGFVMAAQKMVDLLKEWRDELGF
EAQVLNLGGGFGIRYTKDDQMIATEDFAKDIIAAVKERVAAHGLKLPAIWIEPGRALAGE
AATTLYTVGSRKDVDGVCHFLAVDGGMGDNIRPALYQAKYEAFLAEKPHAINEQIVTVV
GKYCESGDVLVRDCPLPNTKPGDILAMPATGAYGYTMASNYNRNPRPAVVFCEDGHS
KLVVRRETFEDMLACEVTE
>4-oxalocrotonate tautomerase (EC 5.3.2.-)
MPLVHIELIEGRSQEQLKNLVADVTKAVVANTGAPAEHVHVVLSEMQKNRYAVGGTLK
SDEQ
>Dipeptidase
MTVTKQSLSACTSILVGKKASLDGSVYIGRNEDAKSAWPKHMCVHPRQEFTSEQVFTS
KDNGFSMPLPKVAYKYTATPEWTDEFGLFEEDGINEFGVAMSATESAYANSEVLSCD
PLVANGIGEEAMVTVVLPYIKTAREGVLRLGQIVSEYGASEANGILFADDQEAWYLEIG
SGHHWVAQRIPDDAYAVVANQLAIEEIDFADTKNFLTSAGIQAFVEANQLNPSLDGTFN
FRKIFGTQTLSDQIYSTPRVWYGQKYLTPSVEQDPMSLELPFIRKPDRKLGLDDLSYVL
GSHYQNTPYDPLNSKDPLLKKKFRPISLAKTQESHILQLRSDFSQDIKGIHWLALGVTA
QSVFVPFFAGANDVHPAYKVGKATYDPTSAYWVYKQAGILVDAHYQAFAKELSDLQT
ELMVSFKTAVRKVVAAVKAGEFDGDLADFLTTESIKLQDLGLKKYQALSASLITQATDM
SPLNFKTDANL
>hypothetical protein

MEILSQEHQLTLKSEQLSLSLDPKRPCPGPILISHLYDELFDLLAKLEDQTVYLHLDLLLY
LQKYYRFKGWRLPDLNFVALPLGYPFQLADLTIEAFNNDDNAFGSLVCIISDGQSKLGY
APKFVPHGQHKKRIKLWKKAFSQADLDYFCLSQDLVAATNDFRALTEVGRQKSLTKYL
QHLEADQSGQILLSYEKPERILTMQKTALAAGFNLKLAPAGQAFLQALFPFEEPVSQTE
ATRDLIVVSTPSATTFDARASLAIQPVMAPKEAKEFLNVIKAKEVIYL
>ISSth1, transposase (orf1), IS3 family
MSRKSKYSVEQKLNILNEAVHSSFKRVAKKYGVNESTILTWYRLYKYQGIDGLRSIRSN
RSYSKEFKLFLVEQYRNSDDSLELFAIKHGLKSKSQLLNWIILYNESKLKAYTPRKRDSI
MPGRKTTFEERLTIIEDLIKHDVNYNWAVNKYHVSYQQVYGWYQKYRKSGNDPQSLR
DRRGKAKPKAEWTELDKLRAENRLLKAEMKRNEMEVAFAKKLIEIRNREAKKSSNTKS
>Mobile element protein
MLEAILKLEEKHKWTLGYLGMTTQLAFENKLSFKAGLKRVTNCMRNHGIKANIRKKKHN
RIKRHKEYINDNLLNGQFDRQNKNEVWVTDTTEVLYGSEQVRKVRVHVVMDLYGRYIL
SYNISATETATSAIEAFKRAFKVEPGVQPMIHTDRGAAYCSRAFNDYLAEQNCIHSMSH
PGHPWENSPIERWWNDFKLVWLAKRARSKSLSELEQSVQQAIKYFNTERAYASKNGL
TAEKFRAQAA
>Translation elongation factor LepA
MNYMDFEQMKERQKHIRNFSIVAHIDHGKSTLADRILEMTDTVSKREMQDQLLDSMDL
ERERGITIKLNAVELHYHAKDGETYIFHLIDTPGHVDFSYEVSRSLAACEGAVLVVDAAQ
GVEAQTLANVYLAIDDDLEIVPVINKIDLPSAQPEVVRQEIEDMIGLDASEAVLASAKAGI
GIEEMLEKIVHDIPAPAGDLQAPLKALIFDSVYDDYRGVVLSIRVVEGTVKPGDKIRLMN
SGAEYEVVEVGVNSPKPLRRDYLMAGDVGYLTASIKDIQDTRVGDTVTSATNPTAKPL
AGYREMHPMVYSGLYPTDNAKYNDLREALEKLQLNDAALEFEPETSQALGFGFRCGF
LGMLHMDVVQERLEREFNMDLITTAPSVTYQVIQTDGSLKSVSNPSEMPEASSIKEIKE
PYVKATIMVPNDYVGPVMELCQRKRGTFVTMEYLDDYRVNVIYDIPLSEIIFDFFDKLKS
STKGYASLDYEIEDDRPSSLVKIDILLNGNKVDALSFIAHKDFAQERARVIVARLKEVIPR
QNFEIPIQAAIGAKIIARTNIKAYRKDVTARIHTGDPDRRAKLLDKQKRGKKRMKAVGTV
DVPQDAFMAVLKTDEDYNKK
>hypothetical protein
MKGDASKQAALNQEMLTYFNDGTVTGKFSAALDRAVREVKNDVKKGENYMTIEEYAA
RQSAYAREEERAEKEAQTIKGLVTLNLDKDQIIKFLIDNFNLDKQEALAAYERVMATA
>hypothetical protein
MQNKNHHDIEKRMAIYQGNLEKQALYAGQSFSDRRKTVVLFLCDHDVYNLNRVHYQLI
PQLVGYPEILIDNGETNVIVNLKGDASKQAALNQEMLTYFNTGTVTGKFSAALDQAVRE
VKNDVEKGENYMTIEEYAARQSAYARKEGREEGREEGRTEERAEKEAQTIKGLVTLNL
DKDQIVKFLIDNFNLDKQEALAAYERVMATA
>hypothetical protein
MNLRKEIALKWQQATIKDHLVFSWVFNDHPELLEQLLKMWLPNFKVRHFNYHQA
>hypothetical protein
MKYEVPSIAINKHLKQDILKIETQKVLIKSESFKEKVVDNIFKDKEVRTALKMLSEV
>Chaperone protein DnaJ
MAEQRDPYEVLGVSKDASADEIKRAYRKLSKKYHPDLNKEPGAEAKFKEINEAYETLS
DPQKKAQYDQYGFAGNQQGFGGGGFSGFGNGGFDFGGSGFEDIFSSFFGGGGQSQ
SRTAPRQGRDLQYEMNLTFEEAIFGKKTNIEYSREALCKTCHGSGAKAGTSPVTCHKC
GGRGFIQVQRQTPLGRMMTQQECDVCHGTGKEIKEKCTTCYGSGHVREKHEVQVTIP
AGVEDGNQMRLQGQGEAGTNGGPYGDLFIVFHVSPSDKFERDGAEIYYTLPISFVQAA

LGDSVKVPTVHGDVEMKIPAGTQTGTTFRLREKGAPRLRGNGNGDQKVTVTIVTPKNL
TKAQKEALKAFADVSGEDPAGDDSFFEKWRKKQRKKR
>Chaperone protein DnaK
MAKIIGIDLGTTNSAVAVLEGKEAKIITNPEGNRTTPSVVSFKNGETQVGEVAKRQAITN
PDTISSIKRHMGEAGYKVDVNGKSYTPQEISAMILQYIKDYAESYLGEKVTQAVITVPAY
FNDSQRQATKDAGKIAGLTVERIVNEPTAAALAYGLDKTDKDEKILVFDLGGGTFDVSIL
ELGDGVFEVLSTNGDTHLGGDDFDNAIIDWLVANFKSENGIDLSSDKMAMQRLKDAAE
KAKKDLSSASQAQISLPFITAGANGPLHLETTLTRAQFNQLTEELVERTKIPVRNALKDA
GLTNADIDEVILVGGSTRIPAVQEAVEAETGHKPNQSVNPDEAVALGAAIQGGVITGDV
KDVVLLDVTPLSLGIETMGGVFTKLIDRNTTIPTSKSQVFSTAADNQPAVDIHVLQGERP
MAADNKTLGRFQLTDIPAAPRGVPQIEVKFDIDKNGIVNVSAKDLGTNKEQKITIKSSSG
LSDEEIDRMVKEAKENEEADKKRKEEVDLRNEVDQLIFQTDKTLKELDGKVSDDEIKQV
KDAKEALKKAQEANDLEDMKAKKDELTKIIQDVSVKLYQQAQAASQAAGADGQADGS
AKADDNTVDGDFKEVDPEDK
>Heat shock protein GrpE
MAEKNEELKQTEEAKQVEATEATEKPTQAEVKEDPVAEAEAKIKDLQAKLDEMEDKYL
RSQAEMANMTNRFKKEQASLLKYAGQSMATAILPAMDNLSRALEIEVDDEASKQLKKG
VEMVAQNIEAALKENEVTKIEALGKPFDPTLHQAVQTVPVEEGQTPETVVQVFQEGYM
LKDRVLRPAMVVVAQ
>Heat-inducible transcription repressor HrcA
MIVVLTDRQLLILEAIIRDYTEIGQPIGSKALQEQLPIHVSSATIRNEMAVLEHQGYITKEH
SSSGRVPSLKGYRYYVDNIVKPAKLDKQSLKSIRNSFGNEFLKVDEIVATSAKILSDLTS
YTAIMLKPDSTDLRLEGFRMVPIGHQQVMVILVASDGTVESQIFNIPPEISGDELEAVMR
LINDQVVGLPLNQVIAKLHATLPLLTRYLQQPAGFLDVFGNVLDKAVKEQVYVSGRMNL
LNFAHSNNVEQLKSLYSLVNQVPNISNLVDNDKGISVRFGDELTNKLLLDYSLISATYDD
GHNGQGLIAILGPTNMPYAKMIGLLGAFRKELTARLIDYYQNFDNFN
>Phosphoglycerate mutase family
MAITVYMVRHGQTMLNRYNRLQGWCDSPLTEKGIEDARLAGRHLADIKFDAAYHTNT
MRTIRTSGFILKANHHTPNLKAAPLTYFKEQGFGYFEAADASQVWLIAGANHGTRTYN
DILRDYSFAQTRDFLHQADPFGDAESDAAYWQRLNAGIDYLRQNHTDGQKLLLVSHSL
TTRSLVDHFAPELGAYNLGPKNGAVTKLEITATDVKVSYYNHYLDGEEY
>hypothetical protein
MTKYSPEFKAQVVVEYMTDGLSSYTLEKKYKVSARQIRKWIQQYQLQGIKSFKQHHRK
RIFSAKFKWHVINYYQTHAESFAQVAAKFDLLPSQISIWRQAFIKYGYDALKPHRKGRS
QKVKKQTKKQKQKLVEKKEIDQLREELARTKQELYNTKMDRDILKKSLALFGPSKPGK
KRK
>Mobile element protein
MDQIRADQAELPKAQRYKIGDLLSRINLAKATYHDERKRIANPDDKYCEVKKVIRRICEQ
GKVRGRLTYGYRRVQARLDQLGIHLAGATIRHLMSELGVQVEIYNHHRNGKYSSYHG
RVGSVADNVLQQQFNAKDPYQVIHTDVTQVRLANQQWAYISAMTDEASKEVLAFQMS
SHPNQELIMNTLNELLDHLPDDAQPIIHSDQGWHYQLKYYTKKLKDHHFIQSMSRKGN
CHDNAPIESFFHLYKTELLQGLPPCNNITELRKLSLTYIDQFNNERISLKTKGMTPVGYR
NHALAV
>hypothetical protein
MMLTIFKSYGIKPPLDEFFEAEVKDYITDSYLCYLQLLPLLLLGTMALAGLKLNWLAGLS
LAAVGINLGLILALRSVVSWRYRRLQAGRPVWQFTRTFRNKTAKKMFNLPSQLDEFEA

NYLTDLMVKASVKTELGNLVLTGLTIPCLVIWPQVAWGGYLILVLLNCIVFQTGIQLKAG
DAGLLNVKPTSKYNPPAKKPNKLRLLGNLVVMGLLYWAVVDGKFTLVGLLGTGCYLVL
MGAFWYGLIKADRNLQTYVDKHQPK
>Transcriptional regulator, PBSX family
MNSRVKDYRLKRGYSQASLAYQVGIARQTLSMIENGRYNPSLDLCIRLAEALETDLNTL
FWKVGKDDANNF
>Ferrous iron transport protein A
MQTLAMAKHAGKFIITKIQGEKALSRRLSEIGMTRGKIITVLSGNQADSGLVIIFQGQKLA
LNASMAACIYVASVSDPAQATLTSLAKVKLHTPCLVAKITGDKAVRQRLMDMGLTRNTL
VTIDKVAPLGDPLELNLRGYKLSIRNFEAQSILVQEVDVS
>Ferrous iron transport protein B
MTTIALAGNPNCGKTTLFNHLTGSSQYVGNWPGVTVEKKEGVLKKEPDIMLVDLPGIY
SLSPYTLEEVVARKYLLDQRPDKVIDIVDSTNIERNLYLSLQLLETGLPTIICLNMLDLLTK
SGKKINLQKLSYLLQVPVVAISALKKDHLPDLLKEIKATPKAYSYPRYDERLESALSMIED
LLTDELAANERRFYAIKLFEQDEVIQAQLALSQAKQAELAQIIATAEKLFAEPSDSIMVNA
RYDLIAKIVSMCVINPDDLKLSLSDKIDLVVTNRFLALPIFALVMWAVYYLSIQTVGTIGSD
WINDQLFGNYLPHLAASLMTALHVAPFMKSLVLDGIIAGVGSVLGFVPQIMMLFFCLGL
LEDCGYMARIAFVMDRIFRRFNLSGKSFIPMLISTGCGVPGIMATRTIEAEADRKMTIML
TTFMPCSAKLTVIALIAGTFFTGQSWVAPSAYFLGMLAVVCSGIFLKKTKLFAGDPAPFI
MELPAYHLPLLRNTLRYVYSRSVAFIHKAGTLIFMSCVVIWFLSSFNWKLQLVTNQEDS
LLHYLGLALAPFFAPLGFGDWRASVAVIAGLVAKENCIDTLRIAFGSATNQAFTKDLLAS
YSSLAAYCFLAFNLLCAPCFAAIGTMYKEFGDVKWTLRAIGYQTGFAYLVALLINQWGL
VASGHYSWFNLLLASLATGLAIWALVFKKGDLKNQHELSQGGI
>hypothetical protein
MLATIILAILIFGYTGYVIYKRFIKKGASSCQDCQEVGCPLVDPAKLKRNLK
>Substrate-specific component PdxU of predicted pyridoxine ECF transporter
MQNDSLHKLTLAALFSALDYILALFQFRIPSPVGHPFVDLGFNFVALGVLFLGYKYGLLS
GAVGLLVFDLLNGYANHAYLTVMEVILLTTGTYLAWILLKKNVSLPAIITLGVVSGLLKMG
TGYLRYLIEALVDAGLPLPKAAVTALAAFPADFATGVSMFIVVPLFFLALNKVSHQLHWN
WYQQYQ
>hypothetical protein
MYESKTVLEQEFTSLADLSAFLDKEEVARTDVINIQKLKASYLLIYWG
>FIG00749852: hypothetical protein
MDANLDRKTALKDVLPTAFGYVGIGLAFGIVGKASGLSVLELGLMSLIAYGGSAQFIMV
SMMLAHSPVLSIVLAAFLVNARMILIIKRAQTSDFNQGSSPLNYLIYFYHNSL
>Mobile element protein
MIIHEEGYVYNFNFHLVFVTKYRRKIFDTQEKVDAMKAILQRQATISEVTVSQLEVMPDH
VHMLISFKPKYAPSNIVKNLKGASARIWFKQYPETKKLLWGGHLWSPSYYMGTLGDMS
KEVVEKYIESQYTEAMRRQLKGYYGKNR
>Mobile element protein
MVKKKGNKFGYTRQYHISLSYQQIDLYEKAIKTQNELYQFALKYLYKTYGRKHIGRPLP
FGQGINYLNNKIKAMFIKEKYDLKRWNVKKLGLSSHAADEFLKTIFTNFSQYRKRLEQT
GQMSAVEKYNLRMNITKDKDGKHKNPKHRSWYRKGSMNFLRNNNSFRTITSQKLPN
GNTKLVSPHHLNVADFGEIMVFENLKHLNFKEIALTKIKRLPDNTFRLQITFTREKKRSS
QNKVVGADWNMFNNEVFRTSENKTIAIPKDIVKKANELAAAKDKFKSLRDREYNKRGK
TTLWQKYQRKQAKLSAKRANILTETYRKLVHELVDEFDTIIIEKIDAFEMRKRSWSLNKA

QNTGKNKRLALIKPYELSKIVESLVNKQNKTLIKVDPYKTSQVEYGTEYEEKHELRATNK
DGKRIYVSAYTGKEVDRDYNAALNIKEWGLHPEKHVKLRDYPKLKAGNLVEII
>FIG00753879: hypothetical protein
MSISSYVLLTILLAGFSMWLAKVLPFLLLKKFTLPKWTLEYLSFVPLTIMSALWFSQLFEQ
HLGHLPILNWENLLASLPTLLAAIFSKDLLVTVIVGIISLACLRFGFF
>Asparagine synthetase [glutamine-hydrolyzing] (EC 6.3.5.4)
MCGFVAVDKLDYDVKEFEDILQTDCHRGPDMTWVSDFGTTKFGFNRLSIMDLSADGM
QPFKALDCMLVCNGEIYNYQELEKLTPDFTYQSTSDCEVLIPLYRKYGLDVMCKLLDAE
FALVLHDQISGKVLAARDPIGIRPMFYGYVKGSGEIAFASTAKSLFKFCDQVMPFPPGH
YYDGHKFVCYHDASFSPKMISPSLDEATKEIHDRLVAGVKKRLHADAPVGYLLSGGLD
SSLVCSIATKLLPDKKLPTFAIGMDRNPIDLKYAREVADYLGTDHTEFIMTREDVLGALR
EVIYALETWDITTIRASIGMYLLCKKIHETTDLKVLLTGECSDELFGYKYTDYAPSPAEFQ
AESQKRIKELYMYDVLRADRCISSNSLEGRVPFADLDFVNYVMSLDPQMKMNIYGKGK
YLLRKAFADQDYLPDDILMREKAAFSDAVGHSLVDDLKEYAEEKYSDQDLKDAQVKYD
YHAPFTKESLLYRDIFEEFFPGQAHWIKDYWMPNKTWEALKDVDDPSARVLKNYGDS
GK
>Sortase A, LPXTG specific
MKWRQKASLILGICLLVVGLVLIFNRPIQNLMINHATKTELKAKITKPKKTKETSFDFKQV
KSIDSAKTLKNTLNKSQGMIGKIAIPAVALKLPIYYGISDASLYRGAGTMKPDQQMGQG
NYALAGHHMRDPQLLFSPIVRSKPGEIIYLTDGKKLYTYKITDKQIISKYAVEWIDDVANQ
KLITLITCDDQGVDRYAIRGSLVKVQKLTAGLSKKYF
>GTP-binding protein Obg
MFVDQVKIKVKAGKGGDGMVAFRREKYVPNGGPAGGDGGKGGSVILKVDQGLRTLM
DFRYHRIFKAKSGQNGMIKGMYGRGAEDTYIAVPEGTTVSDADTGEILGDLVEAGQEL
VVAKGGRGGRGNIHFASAKNPAPEIAENGEPGIERNLKLELKVLADVGLVGFPSVGKS
TLLSVVTSAKPKIAEYHFTTLVPNLGMVRLDDGRDFVMADLPGLIEGASAGVGLGIQFL
RHVERTRVILHLVDMSGMEGRDPYEDYLKINAELKSYDPSILERPQIVVASKMDMPDAS
ENLARFKEQLANDHTLAKQPEVMEISSLTHQGLEALMLRTADLLAQTPKFETVAEKELK
QAGYYKFEADETEAPFTISRDADATWILGGAKLEKLFKMTNLEHDESMMRFARQLRG
MGVDDALRERGAKNGDLVRIGDFTFEFVE
>Excinuclease ABC subunit C
MATQHIENKLKLLPDLPGCYLMKDLNSRIIYVGKAKNLKNRVRSYFKSSHEGKTARLVS
EIRDFETIVTSTDKEAFLLEITLIQKHKPYYNIKLKQGTGYPYIKITNEKDPQLKIVSQVKK
DGAYYFGPYPNVYAATETMQLLQKVYPLRRCNGYQKRPCLYYHMGQCLGACFKEVP
ASEYQAQIKKIKSFLNGNVLGIKRDLTKKMEQAATNLEFERAAELRDQIHYIEVTVEKQKI
ISNDNTPRDLFAYYMDRGWLSIQVFSIRQARLMKREKRLFPCIDTPQAEMASFILQFYN
QKNRLLPKEILVASDLDKDVLREILQVPIRTPKRGAKRELLELAEKNARLVLEEKFRLLEL
DERKTTGAMDQITSALGLPKGHKIEAFDHSHIQGTDIVSAMVCFIDGRPSKNDYRKYKL
KTVSHADEAASTREVIFRRYSRLLKEKQPLPDLILMDGAEIQIEAAQDVLQNQLGLAIPV
AGMVKNDKHKTADLMTMTYQKIGLDPKSEGFYLLQRIQDEVHRFAISFHRQVRSKNSL
ASRLERIAGVGPKTRIKLLRKYSSLKKIADAPLEELTALGISATVAQTIKLSLKGQD
>Translation elongation factor Ts
MAKITAAQVKELRDKTGVGMMDAKKALVATEGDMEKAIDFLREKGMAKAAKKSDRIAA
EGLADVEVKSNVAAIVEVNAETDFVAQNDQFKALVKRLATLVAENKPADVEAALKISTD
KGTVNDEIIEATQVIGEKITLRRFELVEKADNENFGAYLHMGGKIASLVVVEGADEATAK

DVAMHVAAINPKYVNRDQVPAEELAHEREVLTEEAKNEGKPEKIIEKMVEGRLNKFLAE
IALDDQEFVKDPDQTVAKFVASKGGKVKSFVRYEVGEGIEKKEDNFVEEVMSQVKG
>SSU ribosomal protein S2p (SAe)
MAVITMKQLLEAGVHFGHQTRRWNPKMKKYIFTERNGIYIIDLQKTVKLIDDAYDFVKNA
AADDGVVLFVGTKKQAQDAIAEEAKRAGQYYVNHRWLGGTLTNWNTIQKRIKRLKEIK
AMQEDGTFDRLPKKEVALLMKQRDRLEKFLGGIEDMPRIPDVLFIVDPRKERIAVKEAQ
KLNIPIVAMVDTNADPDEIDVKIPSNDDAIRAVRLITAKMADAVIEGNQGEDEAPVNEET
FADSADEKVESIEDIVEAVEGDNDSKDAE
>D-lactate dehydrogenase (EC 1.1.1.28)
MNFIAFSVRDDERPYFEKWAQVNNVKVDLHSEEISEDNLDTIKGYDAVIGLQTGTYPLG
LFAKMQEYGVKDLSIRNVGVDNIDLAAASQKGVTVTNVPAYSPNAIAEFAVTQLMQML
RQTRTYRRKFAEQDYRWAPYISKELRSMTVGVVGTGRIGRAFIQIVQGFGAKVVAFDP
YHNPELEAAGLYVDSLEELFAKADAISLHMPATDQDKHIIGQEALAQMKDGVYLVNTAR
GSLIDTKALIAALKSGKVAGAALDTYEAETPIFNHDLRNQALTDETFKELISLDNALVTPH
IAFYTETAVQNMVEIALDSAKSVLETGQASTIVDA
>COG2827: putative endonuclease containing a URI domain
MLLCGDKSFYGGFTTDVWARLKKHQQGKGAKYTKSHLPVKLIHYEEFETKRAALQAE
YAFKHQSRRQKEAYLKAHGLKDFT
>tRNA (adenine37-N(6))-methyltransferase TrmN6 (EC 2.1.1.223)
MDKNLVQANERVDELYANDIKIIQSPEVFSFSLDAVLLAHFAKPYTKANGRIVDLCAGN
GAVGLFLTAKTKAQINLVEIQPRLADMAQRSVKLNHLDEQVTVYNIDLKQTNDYLAKDS
VDTVTCNPPYFADLPTSKKNPNQHLMLARHEITTTLKEVVAMTSGLLKMGGRFYLVHR
PDRFLEIMDVLRAERLAPKRVQFVYPKAGREANMVLIEAIKDGKDGGLRFCEPLTVYD
QEGNYQPLIKEILYGKG
>1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)
MFYSFIRALVRVIVFIINGNTHYLNKDKLPKGNYILVGPHRTWYDPIFFALAAAPKKFSFM
AKEELFSNPFLRYILVHANAFPVNRANPGPSVIKTPVKWLRKRDLSLIMFPSGTRHSQE
LKGGAVLIAKMAGVPIVPAVYQGPLTFKGLLARKKVTIAFGDPITIERKQKLTDEYQEQVI
AQMNQAFNNLDQAINPDFVYIDQTTKKN
>FIG00742188: hypothetical protein
MSTGVWVLIVIIALVGGFVAGFFAARKYMEDYLKKNPPINEQMIKSMMLQMGQKPSEK
KLHQMMTAMKSNYGKK
>FIG00742757: hypothetical protein
MENKGIPQELIDRINFLAKKKKTTGLSEDEKIEQQSLRETYLAMFRENFRSHIEMLQVYD
KDGKEVTPEKVKEIQRKKGLRDD
>SOS-response repressor and protease LexA (EC 3.4.21.88)
MSRISEKRQRAILQYIYEFINEKAYPPTVREICKAVELSSTSTVHGYLDTLTKNGYIARDL
TKPRAMEITPSGLEFLGISPHEGFIPMVGTVTAGSPILAEQETSDYFPLPPSFETSADLF
MLTIRGESMINAGILNGDQVIVRQQASADNGDIVIAMTEENEATCKRFFKEDGHFRLQP
ENDTMAPIILDNVTILGKVVGLYRNFI
>DNA repair protein RecN
MLQELTIRDFAIIDKLDIDFQPQMTVLTGETGAGKSIIIDALGLLAGGRASVEFVRLGANK
AMIQGQFVLPKDAKTFKILAELGIEPDEVLILERDLYQSGRNTCRINGHLVNLSSLRRVG
ETLIDIHGQNEHQELMHPETHIKLLDEFDPKLQALHKDYLAKYQDYRQLKAALAKREAD
EKAWAQRLDMLEFQVQEIEAAALRANEEEQLLEEKNKLDNFQTINTALETGYQILTGEN
LDVVGNLGNAMSELDKIANLSPELEQISEKVADAFYSLQDVSRDISNQLDQMEWDEDR

LNQIEQRLELIHQLKRKYGDSVAQILEYYEKIRAELSEMQATEIDASKQEAAVEAAYQAA
LSLAQKLSAKRKQSAQVLEGAVKQQLSALYMDKAVFKVKFFPVDELTENGTERVEFYI
QTNLGEAMGPLAKIASGGELSRIMLALKTIFAQKQGVTSIIFDEVDTGVSGRVAQAIAEKI
SQIARNSQVLCISHLPQVAAISDHHYFIAKHIKGGRTETQLTMLEPQQRVTELARMLSG
SEITPLTLEHAQELLALATKIKQDQQAH
>Arginine pathway regulatory protein ArgR, repressor of arg regulon
MRKQERQKLIRQILQDNNVYRQEDLVKLLEQQGVWVTQATISRDVKDMQLIKLPTLNG
EYVYSVPTEKKMNTKKKLAKLISDTYVFSNCHRDLCLIKVLPGNGPILASLIEQMKYEQV
FATMGDDNTVMIFAYSAKAAEEMHEIILDLAKKGN
>RNA binding methyltransferase FtsJ like
MKKERVDVLLVQQGLFATREQAKRAVMAGEILGENEERLDKPGMKIDVETKLHFKGTK
MPYVSRGGLKLEKALKVFKAPIRNKTVLDIGSSTGGFTDVALQKGARLSYALDVGTNQL
AWKLRQDERVVVMENTNFRYSQLADFTSGQPNFATIDVSFISLHLILPNLHTILKTGGD
VIALIKPQFEAGKEKVGKHGIVRDERVHLEVIEDIIRLAVSECYDVVGLDFSPITGGQGNI
EFLIHLVKSEKPGQILLTETPDKVVKQAHQTLKDKH
>Octaprenyl diphosphate synthase (EC 2.5.1.90) / Dimethylallyltransferase (EC 2.5.1.1) / (2E,6E)-farnesyl diphosphate synthase (EC 2.5.1.10) / Geranylgeranyl pyrophosphate synthetase (EC 2.5.1.29)
MSKESWNEFAAYYLPKVEACLQDYLSKLAAYPRLKEAMMYSVNAGGKRIRPLLVLAVI
KTLNGQLNQASLQVAASLELIHTYSLIHDDLPEMDNDDLRRGKPTNHKVFGQALAVLA
GDGLLTTAFELLGQVKLPAETVVQLLNYLAYAAGPQGMVNGQVGDIEGEKTALSLAEL
EQVHAGKTGALLVYACQAGSVICQAKQPIQMALENYGRNFGLAFQIYDDILDRVSTEE
QLGKAVNKDQAEHKNTYPGILGLDGAYTKLAQTIAAANQALVPLKEAGYDVSLLAELLE
YFKVKD
>Exodeoxyribonuclease VII small subunit (EC 3.1.11.6)
MTEKQATFEENLARLDQIVNQLEQGNVPLEDALEQFKEGMKLSKTLEASLTKAEQTLT
KVINDEGEEVNFERDSDE
>Exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
MATSSQYLTVSALTKYLHAKFTRDPYLQKVYLTGEISNFRKRPRHQYFSLKDDGAKIDV
VMFEYSFNRLKFQPENGMKVLVTGRVDLYEANGNYQIIIDEMQPDGIGALYQAYEQLK
AKLAAEGLFSRPKRPLVKFPRRIAVITSESGAVIKDIITTTRRRYPLVQLVLFPAVVQGDA
AADSLIGRLKEVNARGDFDTIIIGRGGGSIEDLWPFNEEKVARAIVASQIPVISSVGHETD
TTIADLVADQRAATPTAAAELATPVLQEELVKLKTQANRLIQAFSQLIRYRQQELAKLMN
SYIFKQPQRLYENYAQKLDLLESRQTQALAVIIQQNKQLTANLTQRLLQQTPAQRLALE
KQALARQKERLIQAVTHYLHDYETHLLRSIQSLEMLSPLKVMGRGFSYVTVADQVIKQS
KDLHVNDQVTLHLVDGEATAKILSVTKDEK
>Methylenetetrahydrofolate dehydrogenase (NADP+) (EC 1.5.1.5) / Methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9)
MTVIMDGKKLAQEVKAELKQRVEELWTSKQIVPRLAVILVGEDPASQVYVRNKRRAAE
KIGIKTVDNLLPATVSEADLLDLIYEYNEDPNIHGILVQLPLPAHIDEAKVTHAILPSKDVD
GFHPVNMGHLFMNMPHSLPCTPRGIMELLAAYGISPAGKEVVIVGRSNIVGRPMVALM
LNADATVTVAHSKTKDLRAVTKRADILVVATGRPEMITSADIKPGAVVIDVGVNRLADG
HLVGDVQRQTVEGVASYLTPVPGGVGPMTIAMLMKQTVEFAEKSEQDGY
>Transcription termination protein NusB

MRVLAFQALFALSSNPEADIRAVCDELVAQLTSADEVVVPPYLYELVEGVITNQAQLDA
EIEKYLNAKWSIARLNKTDLLILRLATYEMLFVEAVPAKVALNEALQLAKEFSDDKSRRFI
NGVLSNLLK
>Alkaline shock protein
MAEDNKIVLASNEPSLGEVRIAPEVIEVIIGIAANKVEGVYSMRGSLANSFSELFGRQSR
GKGVKLDQTNGELKVDVYAFLNYGVSVPKVALAIQEKVRQQVLFMTGLELTQVDVHIQ
GVVPEKQEPTVDPNNLFGEEDGEKK
>Translation elongation factor P
MISVNDFKNGLTIQVDGELWRVVEFQHVKPGKGSAFVRSKLKNLRNGGVQEKTFRAG
EKVEQAQIDRKKMQYLYADGDNYVFMDTETYEQLELPGARIQDELKYLKENMVVSIIMF
GSETLGVELPNTVDLEVKETEPGIRGDTSSGGSKPATMETGLVVQVPFFVNEGDILTIN
TADGTYISRA
>Aminopeptidase YpdF (MP-, MA-, MS-, AP-, NP- specific)
MTDRVEKLRLEMARTSVDCFLVTEPANIFYFSGFTGDDGLLLITEQNKYVITDARFELQL
KKQAETWQHVITRNYLQAACELVAKEKLVALGFEAQLPYWQYDYLDENLTSDLVALPN
LLEKIRSVKDGDEIAKIKAACKLAGQGYDYLLTKFQQSTQWTERQLATELDYFMKSNGA
SDKSFETIVASGLRTAWPHGTASDKLVKKGDLITLDYGYYLNGYTSDVTRTFAYGSQP
AEVRTIYQVVDEARALTIQAIKAGVSGKELDRIGRDFISQAGYGEFFNHGMGHGIGLSIH
ELPNVGARYDDVLQAGQVITIEPGIYLPNIGGVRIEDDILVTEAGYEILTNFGRNYQEL
>2-dehydropantoate 2-reductase (EC 1.1.1.169)
MKFTVVGAGAMGLRYGVLLQEAGNEVEFVDAWEPHVNKMKENGGVYVTRDRQNRH
LVPVKVYTPEEYTNTDADVLIFFLKQMQLADMLERCKHFFTPDQYVLTGMNGMGHVE
KLLNYFDGSKLIAGTALIATILKAPGEVDFVGKVGAGSMNLANYTEEPDEMTHKVVAEL
AKANINPTLTTNFRGTLMAKVVFNSVVNTLCTLFEITMGEYAAFEKADELSLALINEAYD
VCERAGVKMLNSRQEELDTVNYVSKVANPLHYPSMYQDMSKNRPTEVDYINGYLVDL
GRKYNYEAHTHAFLTNLVHLAEHTRQQRQN
>LSU ribosomal protein L27p
MLLNLQYFAHKKGGGSTANGRDSRSKRLGAKRADGQTVTGGSILYRQRGTRIYPGVN
VGMGGDNTLFAKVDGVVKFERKGRDKKQVSVYPVAE
>FIG139598: Potential ribosomal protein
MIHAKFTKKGDLITAFEMYGHAGYADSGFDIVCAAVSVLAINTVNSLEQLVKIEPLVVED
PTEGGFLRLEVGTKDLDKPGVQILLAALELGISDLVKQAEYQKYISLD
>LSU ribosomal protein L21p
MYAIIKTGGKQLKVEAGQAIYVEKLAANAGETVTFDQVVFVGGEKTVIGTPLVEGATVT
GTVEKQGRAKKVVTFKYKPKKHSHTKKGHRQPYTKVVIDAINA
>Aspartate-semialdehyde dehydrogenase (EC 1.2.1.11)
MNEKLRVGVLGATGMVGQRFVAMLENHPWFEVTTLAASPRSAGKTYAQAVDGRWK
METPIPEAVKDLKILDVSEVEKVAAQVDFVFSAVSMSKDEIKAIEEAYAKTETPVVSNNS
AHRWTPDVPMVVPEINPEHFKVIDYQRKRLGTKRGFIAVKPNCSIQSYAPALSAWLKFE
PYEVIASTYQAISGAGKNFDDWPEMKGNIIPFISGEEEKSEKEPLKIWGQLDEAKGEIVP
ATSPVITSQCIRVPILYGHTATVFVKFKQNPTKEELVAALESYQGLPQSLNLPSAPKQFI
QYLSEDDRPQVAKDVNFENGMGISIGRLRKDSVYDWKFVGLSHNTARGAAGGGVLSA
ELLTAQGYITKK
>tRNA-Arg-TCT

>Dimeric dUTPase (EC 3.6.1.23)

MDITKMLTFARNYPLRPLVPSKTFSEKEQLAGQYVLVIYYLGQLMGSIQANQAALQAYC
NALQAFLQIANERKWTYKMLVDDELLAQIEQKWASQSLNTVYLILEQQLNKSYFENNQ
AAFVHAWHIFLKFGLVDLGFTAGEIEAQYLQA
>hypothetical protein
MTETKNETTVFYPDGHHETKGHAASLTAKGRLIEYNQQLKANLAAGIRFTEILDILSETK
DTDELLEAALALSTYQLDSAYLVYPQQYSKSDFYLIFLQRLLQLHQTNTMILESSERKKE
LHHSFEGISQAGYFLFATSETDPAGAYYYEKETEQRLFYINFKRHLLSFNSQALTNLLVV
NYGKRYQFKAIKNLANLLIAVGKCFKEDYGYEVDFSYLDAANQALYPLTSSKLPQQALD
RLFIKASQAGYMLTTGIDEEAILDLGPTIRLSLFAHGENTWVMQVQDQANCLSWLDLLF
KYEFLKDWYLENLDTIEIRTDSRYFKG
>FIG00749923: hypothetical protein
MANTNQFSLSDPKQTQATLAGALANKIKYLQTLEEAIANKDDRLVYQLIDGGRYAKEIT
RARHVGGDENEQLIYDVHPEISDFLSHKLIAYLREIYPFFYFEETAVGQFQIYFGNWWG
RRVFGKLDVLNVKFIFDEGEYGKLARAFALEAENKRLNSDKIKQISEESDRLQALIDDQD
KRDAKKEELREEMKRTNQERSGFWKVAEQKEAKQRLVEELTELAELDEKANRAYKQI
RENEEKVLALSKEDTLIGYEKQSIVAKFGSFEMFEDKVASLYRDYIADLIATKGVRADD
>Hydroxymethylglutaryl-CoA synthase (EC 2.3.3.10)
MKIGIDKIGFATSNLYVDMTELAEARNVDPNKYLIGIGQSKMAVIPPSQDVVTLAASAAS
KILTADDKDQVELVLFATESGVDNSKSTAAYLSELLGLKANIRAIELKQACYGATAGLQL
AKGHIALNPQSKVLVIGADIARYGLNTPGEPTQGGGAVALLVTANPRILALEEDASYLTK
DIMDFWRPLGHIEALVDGKYSSNVYLDFFQDVFSAYQSKTGLQLQDFAALLFHLPYTK
MGLKALRLAIEGADDQTVTRLSHEFEAARSYNRQVGNLYTGSLYLSLLSLLENSDSLRA
GQRLGLFSYGSGAEGEFYSGILQPGFERMLCGHAAALKARTKVSVAEYEQLYKAGLTL
AADIDFDWQNDASAYVLTGIKANQRQYLAK
>Predicted nucleoside phosphatase
MQAAKLTVFGIVQGVGFRYMTKQVADELGVVGFVKNRVDGSVYIEAQAEPLTLYKFIS
QIKAGPSPAGHVDHVSVTEMKPKDYQDFRVEFY
>Inner membrane protein translocase component YidC, OxaA protein
MKRKKYLTLGSLFVMVSLFLSGCVQRTASGKPYGFVYEKLAVPTQHIMDALSNWLGG
SGNSYGWAIILITFIVRLVLMPVMLSQLKKSTIQQEKLGSIRPQLNDIQKRQKEAKTPEER
AQISQEMMALYKENNISMTGGIGCLPLLIQMPIFAALYAAIQYSELLSKSTFYGINLGQRS
FTFVILTFLIYALQGWLSLQGVSEEQKKQMSSMMIVSPVMLAFMTYISPAGLGLYFFVG
GIFACIQTLLINIMRPRIRANIEAELKAKPAKTVIAKESQPTVVKQPAKASAPKTHNRKRN
AGKQRHKHK
>Two-component system histidine kinase
MEVRKTPREDKHFFSLKLKWSVATGLGVLLIFLVFSVLLYQSFSSLLLRQEQQYASEAL
STAINKLAVNDNELTADKVRTDLSSALAEVRSTARPSVLYSDSVFVTLSRKNIGLSVYNL
SGDEIFASRKVPVKFHKSGSNVDHLTKIDGVSVFVTSQVVKSKATHQVIGYVQVTNRLI
DYDNTKRKLMLIFLVFGLTAALASGMLAYGLSSWLLRPVDLLNETIGKINDDEEGDALA
NVRVPVLNQHDELSELSLLFNDMLDRMQRYIEQQQQFVEDVSHELRTPVAIIQGHLDLL
NRWGKDDPEILSESIAASLQEITRMKSLVQEMLDLSRAEQVEVQFAQEETDAREVGLQ
VFNNFQMIHPEFTFVLDNDLKKPTITKIYRNHLEQVLIILMDNAVKYSRERKEVHITMAKD
QRHVQIVVQDFGEGISEENLAKIFDRFYRVDKARSRDKGGNGLGLAIARRLIEGYHGKL
AVESVVNQGSVFRITLPLVTEPKETENKA
>DNA-binding heavy metal response regulator

MSKILIIEDEKNLARFVELELQHEGYDTFACKNGREGLEVALEKDWDAILLDLMLPELNG
LEVCRRVRQVKNTPIIMMTARDSVIDRVSGLDHGADDYIVKPFAIEELLARLRAVLRRVD
LESEQNASKQTTVTYRDLTIEKENRVVRRGDEIIELTKREYELLLALMENVNVVLARDVL
LKKVWGYETQIETNVVDVYIRYLRNKIDRPGEDSYIQTVRGTGYVMRS
>6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)
MDKPQIGVIGMAVMGKNLALNIESRGYTVAVYNRTGSKTEKVVADHPDKNLVPSYTIE
DFVKSLEKPRRIIMMVKAGKATDATIKQLMPYLEKGDVLIDGGNTFFEDTIRRNQELDN
SDINFIGMGVSGGEKGALEGPSLMPGGQKEAYDLVAPILEQIAAKAEDGKPCVTYIGPN
GAGHYVKMVHNGIEYGDMELIAESYNLLRNVLGLSNEEIAKIFEDWRSSELNSYLIDITA
DILTRKDDLGSDKPIVDMILDRAGNKGTGKWSSQSALELGVPQSLITESVYARYISTFKD
ERVAASEVLPAPKATIGELDKQAYIEKIRKALYFSKIMSYAQGFEQLRLASEHYDWDLKF
GQLAMIWREGCIIRAQFLQKITDAFERNADLKNLMLDDYFKDITAKYQDAVRDVAALAIK
AGVACPGFTAAITYYDQYRSAVLPANIIQAQRDYFGAHTYERTDREGSFHYEWYHEE
>ATP-dependent Clp protease ATP-binding subunit ClpX
MYENTDNNGPVTCSFCGKSETQVSSMISGPGVYICNECVELAQSIIESESKAQVQEEL
TEVPTPKEIVDILNQYVIGQEEAKRTLAVAVYNHYKRINSNIEADETELQKSNICLVGPTG
SGKTFLAQSLAKILNVPFAIADATTLTEAGYVGEDVENILLKLLQNADYDVERAERGIIYI
DEIDKIAKKSENVSITRDVSGEGVQQALLKILEGTIANVPPQGGRKHPQQEFIQIDTTNIL
FIVGGAFDGIETMVKNRLGEKTIGFGTDSKQQFDESKSVMQQIIPEDLLKFGLIPEFIGRL
PILTALEKLTEDDLVRILTEPKNALVKQYTKLVELDGVKLKFEAPALKEVARLAIERNTGA
RGLRSIIENTMRDIMFDIPGREDIEEVIITKATVDGTRKPKVVKKKVS
>Mobile element protein
MDHSYSNTKPHQKGKHLSKDDRITIQVMHSIGCSNRAIARELNCSPSTIGYELKRGTVS
LYTGNVKRYKAAKGQMTYERHRRECGRKSLFLHRSRFINYVSHCFHKRGWSLDACV
GYALAEGIFPKDQVVSTKTLYNYVNLGLMDIKNIDLPEKVKRNTKARRARANKRILGRSI
DDRSPRIDTRKDFGHWECDLVLGHKTKDDDVLLTLCERKTRCFFMIKIEDKTSASVMK
AFDKLREFYGSKWNRIFKSITTDNGSEFADLSNLEQVSKTLVYYAHPYTSCDKGSVER
HNGLIRRYIPKGDRIDKYSVEDIAKIEVWCNSLPRKILNYKTPEECFDIELDHIYRHR
>Cell division trigger factor (EC 5.2.1.8)
MSAKWEKQEGTNEGKLTFEIAPEQVKEGLDIAFNRVKKSLNVPGFRKGKVPRQIFNKM
YGEEALYQDALNAVLPAAYEAAIKETDIKPVAQPEIDVESMEKESAWVLTAKVTVQPEV
ELGQYKDLEVTKQNTEVTDEDVDAELKRRQQQQAELVLKEDAPAENGDTVVIDFEGK
VDGEAFDGGSAQNHSLELGSNTFIPGFEDQLVGHKAGETVEVKVTFPEDYQAEDLQG
KDAVFTTTIHEVKTKELPALDDEFAKDVDEEVETLVELKAKIKEELAKNKAEFAKNQVQE
EAIAAAVKNATIGEIPAAMIEDDVQRQLDQYLSGMQQQGISPEMYYQLTGTKEEDLRK
QFEDGAEERVKTNLVLEAVVAAEGLKASEEEINNEIKNLAAQYNMDEKAVRSALSDDM
LDHDISVRKAVELIADSAKQTLEPKEEAND
>Translation elongation factor Tu
MAKEHYERTKPHVNIGTIGHVDHGKTTLTAAITKVLAEKGLAEASDYASIDAAPEERER
GITINTAHVEYETETRHYAHIDAPGHADYVKNMITGAAQMDGAILVVAATDGPMPQTRE
HILLARQVGVEYIVVFLNKCDLVDDEELLDLVEMEVRDLLSEYDFPGDDIPVIRGSALKA
LEGDPEAEKQIEELMNVVDEYIPTPERPTDKPFLMPVEDVFTITGRGTVASGRIDRGTV
KVGDEVEIVGLKDDIVKTTVTGVEMFRKTLDEGEAGDNIGALLRGVDRTQVERGQVLA
KPGSIQTHKKFKGEVYVLTKEEGGRHTPFFSNYRPQFYFHTTDVTGVIELPEGVEMVM
PGDNVTFTVELIAPVAIEKGLKFTVREGGRTVGAGVVSEIDD
>hypothetical protein

MSAISISEFEQAQTCYEKKDYLKARQILSKLYLQKQTLRTNYMLFQTLVATADYSAAYQ
LASDYLNDYLARNGWFKQYLQVGVKAGQNIKLWQLVSQISPYLNEAEQTLVVKTLLET
GEDTQLSKSFSHLGAFELKQQRRIYQDAYSLAKEVWLQGVIPILVDQDVHPLIRNTALS
DLQKLAYSKRVKIRTFFEEKLELVPSQLVAFEDDPVVQAQEQLFTKKVNEGKLDALWQ
QAELKLVLMLLYPDFTKAKNLLGDYQQWYYLLVDENSKATLEKQVMILRKKVEKSLAT
WEKAWQ
>Ribonuclease J2 (endoribonuclease in RNA processing)
MANIKIIPLGGVRENAKNMYAVEVADSIFILDCGLKYPENELLGIDIVIPDFTYLRENANKI
VGVFLTHGHADAIGALPYFVSEFNVPVFGSELTIALAKIACENEEQARKFKDFHIINAETE
IDFTSATVSFFKTTHSIPESLGIVLKTTEGNIVYTGDFKFDQTATRGYQTDLGRLASLGQ
EGVLALLSDSANAENPKEIVNEHEIYDFISETFEYRRGRIIVACVASNILRVQQIFDAAAA
NERKVALTGHDVEKVVRTAMKRGKLTLPTDDILVSLDEIKKLPDEQVVILETGRSGEPL
KSLQKMALGRHRHINIHEGDLVFITTTPSHAVETMVAKTRDMIYRAGGDVKAISDEFNS
SGHASRNDLQLLLNLLHPQYLIPVQGEYRLLAANAQWARELGMAAENIFILAKGDVAEY
SNGKMLVVGSVEAGNTMIDGLGVGDIGNIVLRDRKILSEDGIFIAVVTIDRRKKKIVDSPK
ITSRGFVYVKTSRDLMAESAEIVTKVVQSNLDNKEFDWSHLKQDVREQLNHFLYEQTK
RHPVILPVIMEINQYHGHKKAKNKKSKN
>4-hydroxy-tetrahydrodipicolinate synthase (EC 4.3.3.7)
MKFDNTEILTAMVTPFDEKGELSLTRLEKLINHLLATGTEGLLVNGTTGEGPTLEHDEKL
ALIKAAAKIIAGRVPLMVGTGSNNTKQTIAFTKEVAAIEGVDAALVVVPYYNKPNQKGMI
AHFEAVADASDLPIFIYNIPGRTGVTMEVATVAKLAQHPNIIGIKDCTGNVGLAKLVEQT
PADFLVYSGEDADALAAKAVGAQGIISVASHIYGLEMQAMYQALAKGNLAEAASLMRA
LTPKMETLFSYPSPTPVKAALNHLGIEVGATRLPLLALNDVETKEVLAKLEL
>SSU ribosomal protein S15p (S13e)
MTKSEIMAKYARHEGDTGSAEVQIALLTADINKLNEHIRQHKKDFASQRGLMKKIGHRR
NLLAYLRKTDVQRYRELIKSLGLRR
>SSU ribosomal protein S20p
MPIIKSAIERVRTNEKANAKNSAQLSQMRTAIKKFEKAKTAGAENTEELYRAAVSAVDK
ALSKGLIKANKAARDKSRMAARLAK
>DNA polymerase III delta subunit (EC 2.7.7.7)
MKVSELEKKLAQGECAPVYLVLGAEEYSINQAKEAFLKLVPKEEREFNVGAYDMETTP
LVEALNDALEIPFFGDKRLVFVKRPLFLTAESKKSKVEHDIDGLMNYLDNPEPTTVLVFL
APYVKLDERKKVTKKLKKAAVLVDAAPASEGEIRQRLKEELQAKGYTITAEALEELLVRT
DANLTRIMAELPKLTLLAMETREITVAQVERMVTRSLEQNVFKLVDLVLAKKVQAALGL
YHDLLEQKEEPLKLNAILLGQFRLLIQIQVLAAKGYAQGNIAQVLKAHPYRVKIALRNIRQ
YRFALKDLASAYLGLVAVEKKLKSTQQEAELLFQLFMLEYVK
>Late competence protein ComEC, DNA transport
MYADQTKVNGNLVTLEAYCPQIKQKVQAFYQLKNQDQQQWWLENQANLQVDFTGKI
NGINVPTNENEFDYRRFCETKKIANSIYIQAASPQLEKKPGLVGKLHEVRRRLMLYFAR
LPQPLGNYAAVLLIGYQPQVFADTLDQIRNLGLLYLFCLSGMHVFYLLLVIKRICRLVKLS
QEATNLLTLCIIPVYALLGGASLSLLRASGMVFIANLMKLLGVRTWNSLNSWSLVVLVNL
LVNPAVVFSLGAQLSYLLTLMILIAPKLSDIKLGVYLNLLSLPLILWQTYKWNLLTMALTILI
IPIFEWLILPSVVLGIMCYWLRPLCNQVLTACSLGFTLLKRLPMEVVFGKPPLIIVCFWLG
LTLLLLVKPKKRLVICLLLTYGLTFMMIHLPLTSEVVYFDVGQGDATLIRDKFNRTVTLIDT
GGKVAFTKEKWQVRQSKTAGQTVIANYLLSKGVGRIDQLVLTHQDSDHVGNFPSLSQ
VITIKQIFIPAGMEKLPSFQGRVRQARLGQTKIWPVTVGNSATLRQVKLLHPFTEGKGT

NECSLVLWYNLKRTSYLFMGDLDRANELVLLAKYPNLTATVLKVGHHGSKTSSDPRFI
GKIKPQLVIISAGRHNRYGHPHPETLATLAAANIPYRLTAQRGMIKIETDFFGTKLVDFA
QQTGTSSFLD
>dCMP deaminase (EC 3.5.4.12); Late competence protein ComEB
MTEKRIPWDQYFMLQAVLLSLRSTCNRLAVGAILVRDKRIIAGGYNGSVSGDDHCLDA
GCYIVDGHCLRTIHAEMNAVLQCAKFGVATADAEVYVTDFPCLQCTKMLLQAGIKKIYY
LRNYHNDDYAVRLLKLKNVAVSQVKIDKEYLHMVAASQLKISLDD
>Late competence protein ComEA, DNA receptor
MLTELMEFWEENKFKLCLLGIVLLLGGGYFAWNRVTKPGEDSLMAVTSSSGVSSQLSA
TVSQFSQNKTATRIYVDVKGAVKKPGVYQVPANVRVNYVLAKAGGLTADADLKQVNLA
SQLVDQMIVYIPQQGETPEINSPILIKNQAASAETETTSEVAGTEGQGETDASKINLNTA
TKEQLMELTGIGDKKAEQIIAYRQEHGKFNSIDDLKNVSGIGDKTFASISSRLTV
>Lon-like protease with PDZ domain
MTKLLTKKVRYLILLLVGLLLILFLPLPYYIELPGKAVPVSNYLTVTRPSPKKKAGELRLVY
VSLYRASPATYLWSYTNRFASRIPAKEITGENNNSDYNKIQAYYMQDALNEAEYVALKA
AHYQVKRKYQGIYVMSLLKNSTFKGKLQVGDVITAINGKTYQSSGEFVKAVQGLKYGS
RVQVSYLRDQTPKVSWGRIVKLAGSNKTGLGISLADKSRIQTPVSFKTNVDDIGGPSA
GLMFSLELYNQLTAKQLLRGRNVAGTGTISVDGKVGPIGGVDKKVIASAKAKADIFFVP
VQSEPGAISNYQLAKKTAKQIKTKMKIVPVASFSEAVTYLEEN
>Phosphopantetheine adenylyltransferase (EC 2.7.7.3)
MKAVFPGSFDPLTNGHLDLIKRASGLFEEVIVAIATNTAKQPLFSPQEKMTLVNSAIKGL
PNVSAVAIPSDLTVNVARKLGAKAIVRGVRNVQDFEYEQGIAAMNKQLAPDLETILLFA
RPEYTLLSSSLIKEVARFHGDLAKFVPAEVGLALKKKMESDD
>Ribosomal RNA small subunit methyltransferase D (EC 2.1.1.-)
MRIIAGEFGGRRLKAVPGMKTRPTTDKIKESMFNIIGPYFDGGRALDLFAGSGGLSIEAV
SRGVDQAVLIDRQYQAIKTIKENIAVTKNEAAFTVIKGDAKKVIYQLAQKKEQFNYVFLD
PPYKQQEILKMIAELVQLELLAPNALIVCETDQNANLPTSLAGFSCYKQVDYGITELTFY
VYQGEV
>hypothetical Cytosolic Protein
MAFEVRKRQGVFVYLYHLKQSKQLRKYGSIQYVSRKMKYVLLYLDRSQVEETVAKLKQ
LRYVKEVLVSKRPELPTNFDNSLGKFKLTEEDQEKFRDKE
>Pyruvate carboxylase (EC 6.4.1.1)
MKKILVANRGEIAVRIIRACHELNLKTVAVYAKEDEYGIHRFKADEAYLIGQGKKPIEAYL
DMDDIIRVAKMTGADAIHPGYGFLSENEEFAEKCAAAGITFIGPSVRQLAIFGDKIEAKE
VAHKAGLKTIPGSNAPVKSIVEVREFANKYGYPIMVKAALGGGGRGMRIVRDDEELEE
AYNRARSEAMQSFGDDELYVEKYLLNPKHIEVQILGDKHGNVLHLFERDCSVQRRNQ
KVIEFAPSISLSEERRQEICQAAVRLAKSVNYQNAGTVEFLVTDDDFYFIEVNPRIQVEH
TTTEMITEIDIVQAQIQIADGKDLFKDLHLPKQEDMSYRGVAIQCRITTEDPENNFMPDT
GKIETYRSPGGFGVRLDGGNAYTGAVITPYFDSLLVKACVQARNFPEAVAKMDRVLTE
FQIRGVKTNIEFMQKVLNHPTFQKGQAHTTFVDHTPELFHFERKPSTTNHLLKYISDVT
VNGFPGTTHHKKVFVPDIKLETDFEETPAKENAKSVFDRNGVEAAMNWVKQQEKVLIT
DTTMRDAHQSLFATRMRTKDMLPVIENYEKAFPNVFSAEVWGGATFDVAYRFLNEDP
WERLKIMRKQMPHTLLQMLLRGSNAVGYKNYPDNVLKEFISRSAQNGVDVFRVFDSL
NWVVQMEKPIEFVRDAGKIAEGTMCYTGDILNPAERKYSLDYYRSLAHDLVGAGSQIIG
IKDMAGLLKPKAAYELVSALKADLDVPIHLHTHDTTGNGVATLVEATRAGVDVVDVAAS
SLSGTTSQPSMSSFYYALEGNKRQPDMNMTNVEKVNRYWGGIKPFYQDFMNGTTSP

QTDIYQTEMPGGQYSNLQQQAKALGIEDFELVKKTYREVNELLGDIIKVTPSSKVVGDF
AIFMIQNNLNSENIFKLGKTLDFPESVVDFFAGDIGQPVGGFPEKLQKLVLKDRKPITVR
PGSLAAPVDFDRVKSELAEKIKHDPTPEEVLSYILYPEVFLGYEANVERFGSMYALDTT
TFYQGMRPGETVHIDFAPGRSIIVRLDSVSTADDAGNRSIFYSLNGQTIQIMIHDKSKEA
KVKSIPKAEPTNPNHVGATLSGSVLEMLVEKGQVVKKGTPLVVTEAMKMETTIKAPFD
GKVTHIYVKDGDVLESQDLLLELTPVKSK
>Cell division protein FtsW
MEGGIIRLIFWISRLVIVSMNNLKNKLKYFDLYLFIPYIVLSIIGIIMVYSASSINQAYSGLKS
NYYLVRQGAFLGVGLILFFFASHMNQRIWIKKISITLFYILVLLGLIVAVFMNEAINGAKG
WIVLGSFSIQPAEVCKLLLAIWLARYFAKQEQGKLRGMHYHPIKSNGWLIAGIFPLLGLIL
ITPDLGSLVINSAIIFVIILAANDWSSSQKIKALVGSLVLFYGIINVLRFWNPFSGSRLAYVY
QRFVAYFDPFKVATSSGKQLVNSYYALSNGGLLGRGLGNSIQKRGYLPEPYTDFILSIIA
EELGFIGVLVVLGLLLWLTLRIFLVGIRSTNTYDSLFCYGIGTFMFVETSLNVGAVCGLLP
ITGVTLPFVSYGGSSMLVLSVALGMVVNISVSQRRNMEFDPLLKPRRGKRK
>GTP-binding protein TypA/BipA
MKTRDDIRNVAIIAHVDHGKTTLVNELLKQSDTLDEHTEIGDRALDSNDIEKERGITILSK
NTAVRYKDKQINILDTPGHADFGGEVERIMKMVDGVLLIVDAFEGTMPQTRFVLKKALD
QHLTPIVIINKVDRPGARPEEVVDEVLDLFIELGADEEQLEFPVVYASAMNGTSSFDSDP
AKQEHTMDPIFNTILDTIPAPVDNSDEPLQFQVAMLDYNEFVGRVGIGRVFRGKIKVGD
QVTVMKLDGSQQNFRVTKLFGFFGLKRLEIQEAKAGDLIAVSGMDDIFVGETVCPVDH
PEALPVLRIDAPTLQMTFMTNDSPFVGREGDHVTARKLEDRLKVQLHTDVSLRVEDTD
SPDAWIVSGRGELHLSILIETLRREGFELAVSRPKVIYREIDGTLCEPFESVIIDTPDEYS
GSVIDSMAQRKGEMLNMETGNNGQTRLTFSTPTRGLIGYDSQFLSMTRGYGILNHTFA
EYKPVIRNWEPGRRNGALVSINQGKATTYAIMGVEGRGTIFVDPGMEVYEGMIVGQSS
RERDISVNITKGKNMTNVRSSNKDQTATIKKPTHLTLEESLEFLNEDELLEITPESIRLRK
RILETNAREKAAKAAKKATK
>Inositol-1-monophosphatase (EC 3.1.3.25)
MSWQAIDNEVKSWLKQARELILTKMDSDLQVSTKSGPNDLVTNLDKQVEQFYISQIKQ
KFKHSRILGEESCKEEGLLSFDGLFWVIDPIDGTMNFVKEKENFASMIAIYKQGKPYMA
YIYDVMKDRLVWGGPALGAVYCNQTQLPAPKQHALKDGLVAISCPMVLTNYRNVVEVI
KASSGCRMVGSAGIEFINLILGKHCAYLSYLRPWDFLPGKILGETLGLSVKTIDDEKVDV
VSSSVVLVATKSAHTEIIELVNR
>FIG006438: hypothetical protein
MRNNVEYPLLAEWTTTEIIKVSEFYLQIEQANTVGVERVKLLKTYQDYLKIVPAKMVQK
QIDKQFEAASGYSIYRTIQAAKQGTKKMVKIN
>Peptide deformylase (EC 3.5.1.88)
MITMDDIIREGNPTLRARAQKITFPLTAEIKTLAQEMMEFLENSQNPELAEKYKLRAGVG
LAAPQVDKSLRMTAVLVPGEEGQPAIFKHVLINPTILSESVQMIALSEGEGCLSVDREV
PGYVPRHDKIKLRWYDLDGQEHIERLKDYPAIVVQHEIDHLNGILFYDHINKDNPFATTD
DMLLLE
>hypothetical protein
METLSQEQTDKIIRLVLIKEGLITEDQEVSSTVLSDIWGQGVLVFSYELVVQTNDGDLSA
TRRQFVKDLQTVCSAQKLQGLPGYPPLMVTDFWVDERQSLHIDVANIANKATAQYVH
DINKVE
>Protein of unknown function DUF1447

MIFKVYYQPSKKVNPRRENTEVLYLDTDSEVNARLLVESNTEYNVEFIEGLDEATLAYE
KESPNFKLTEF
>Ribonuclease J1 (endonuclease and 5' exonuclease)
MKKLNVKNNETAVFAIGGLGEIGKNTYGIQFQDEIIVVDAGIKFPEDDLLGIDYVIPDYQY
LVANQQKIKALVITHGHEDHIGGIPYLLKKVNIPIYAGPLALALIKSKLEEHGLLRTTELHEI
NEDTVLKFRKTSISFFRTTHSIPDTLGVAVKTPSGTIVCTGDYKFDLTPVTNQPPNLQKM
AKLGSEGVLCLLSDSTNAEVPNFTKSERWVGRSIRRIFEKVEGRIIFATFASNISRIQQA
ASAAVNKGRKIVVFGRSMEAAIVNGRELGYLDIPDDALVDAREINSLPANKVMILCTGS
QGEPMAALSRIANGTHRQVAIQPGDTVVFSSSPIPGNTLSVNKVINELEEAGATVIHGK
VNNIHASGHGGQEEQKLMLRLMQPKFFMPVHGEYRMLKIHTELAEKCGVPTENSFILE
NGDVLALTKDSARIAGHFNAADVYVDGSGVGDIGNVVLRDRRILSEEGLVVVVATIDMK
KREILAGPDLLSRGFIYMRESGELINQGRRRLFRTIRRTLKNDKVTEAMLRDAIIGDLQN
FLFEQTERHPMILPMLITL
>Transcription regulator [contains diacylglycerol kinase catalytic domain]
MKNFFVIVNEDANDGQGKKNWQLVKQKLTERQVSYTYQLTNNNAKDLAYSFVKKTTT
SELTKTVILVVGGDGTLNETLNGIKAAKLDYDVPLAFIPSGNNNGFAQGIGISFDPLTALE
QVLTTDKASYFDIGFFEETTHNQTGYFLNDFGIGIDAYIVSLNAHAPHHKYLQKFHLGIF
SYIWNVFEAYVNQEAFNVTVRIGSKYEFYKHAFLVNVANHPYFKKNIALSPVANAGDHK
LDLIIADNMNFFKFLLLSILIYFRKQLKLPSVHHYKEKDIHLIINSLEFGQIDGQEQGNKYY
DIFFKASKYPFWFTALTPTDSQEKSA
>transposase
MNNSIKTILGIKDPYLKLDEKNFDNPIEDQPNQIIVHLIQTYPMHCPKCGHLMCKNGYKT
VNCLGPELHFKPTIWSIKKQKYICKASSSCPEVVTKLAAVEDIHYRNHISLAIKQRAMML
LTKNESQSDLAKELNVSDWTIRRVITNLDQFFKPNYHWLPRHIAFDDFKSGRFAPSGM
SMILMNIENKRTLDIILSRKNSYLRKYFLRYDRSARLAVQTVTVDLYTPYRHLIHELFPHA
LIIADHFHIVAQAYRAFNKIRIQVMNRAGAGTHKWRALKHFWKLLLTPANELKYDNYWS
RRNFSYAQLTDVEVIHRLLSFDNELKRAYEYYQNLILVIAHRSKKELKNLLAIKWTQLPQ
ALQKVQRTLRRHKQEIYNSFKYDTYTNGPVEGTNNKIKVIKRTAYGFRNFFNFRIRILLA
LPNTYIAINWRNKQTAHAKVQAQAA
>RecD-like DNA helicase YrrC
MVKALEHDFEWAKSQITVVGNFSDIKEGGRYRFFGRLVAHPTYGEQFRVDNYQLEVA
TTAEGLIAYFASENFPGIGKKTAAKIVAKLGTDAISKILASQSVLADLALSAKQRQTIVEVL
QANNGMEQIIVGLNSFGFSGNLANKVYNAYHEQALDIIHENPYRLARDISGISFSRADQI
AEELGFVSNDPNRIEAAIFQVLYEANDQEGHTYLTLRQVVSEAQAYLEAGRNEEIDPNE
IADMMIKLAEQRKIVPEGKLVYLKRNYYAELKIAENCYRLVKDKDQPEFSATKLAHALRI
VEKKLGIDYDHDQLEAIKAAIVAPVFLLTGGPGTGKTTIIKGIIEAYCLLYDLNPQAIKLAA
PTGRAAKRMSEVTGLTASTIHRLLGLNGNEEDEFEPKEVTGNLLIVDESSMVDTHLME
TLLEAIPSGMQVILVGDRNQLPSVGAGQVFSDLLDSDVLPKKELSKIYRQGEGSTIVTLA
HKVQEGKLPADLLVKQADRSFINCQPNQVPQVIEQVVKIALAKGIKQDEIQLLAPMYKGI
AGVDALNELIQGIFNPRKANKKEVTYLNQQFRIGDRVLHLVNNPEENIFNGDIGKIVGIE
DSKKEQQLIVAFDSNEVTFDTKNWNNLRLAYCLSIHKSQGSEFPVVILPLVRQFSRMLT
RNLLYTAITRAKQKLIMVGDPSAFKLSVETLSLNRQTSLKQRLTEVFETDQSRKEDADT
VQVAAEQITEGQVTQERILTMALIEGNQIEPLIGLDGLSPYDFMEKVN
>hypothetical protein
MTKREEKEKQAQVALKQLIAQVDKDPYNPEAYYQLGMALTEMKSYPQAEELFKKALN
VFEGQADKQALLHYGLGNVFYASEFYPEALKEFQVVKDSKLANEALLMLAQTNYAQK

QYQSCVAFALTVFEQDATNTTALKLVADSLMALGSFAEAKQYYQKFLELSPANVEVLF
SLGLVKLVLGEDAKGLFTKVKQSDAKYYAKMQTRLQEVEAFIKKTQPKDKAGD
>Phosphoglycerate mutase family 5
MTKLYFVRHGKTEWNLEGRYQGAKGDSPLLGQSYAEIEQLATYLKQFKFAKIYTSPIKR
ARVTAQKLAASLDQNVKLEINDAFKEFNLGQMEGMYFTDCERLFPEEVDAFRNHPDK
YDATKIGGESFEEVFSRMTPKVKEICRRYPNENVIIVSHGAALCAEIRHLLGVPLEQLRA
KGGLNNTSTTILQTNDGEYFKCLAWNKHDYLKRKLDKTDMV
>tRNA-specific 2-thiouridylase MnmA
MKDKSQIRVVVGMSGGVDSSVCAYLLKQEGYDVVGVFMKNWDDQNDSGVCTVTED
YQDVAKVANQIGIPYYSVNFEKEYWDRVFTYFLDEYKNGRTPNPDVMCNKEVKFKAFL
DYAMELDADYIAMGHYAQLKRDEDGTVHLMRGADDNKDQTYFLSQLSQEQLQKAMF
PIGGMNKQEVREIAKKAGLATAEKKDSTGICFIGERNFSKFLSEFLPAQPGKMMTVDGE
YKGEHYGLMNYTIGQRRGLGIGGDSKSNEPWFVVGKDLKTNTLLVGQGYHNEHLYAT
SLDASKLSFVAPIEERGNEFHCTAKFRYRQTDSGVTVRFNEDKTKVDVIFDEPVRAITP
GQEVVFYDGAECLGSGTIDHAYQDEKLLQYV
>FIG00743560: hypothetical protein
MAYTEVAKVLGAQDSYRLSPEIKRYSLRDVGFVETKGGKFQLERSLDPNTPFGQAFKL
KVVVDASLKKFKLATVTANGLKEVNVFTGKNAEVNVEQLNFILADLVVHNILVKL
>Cysteine desulfurase (EC 2.8.1.7)
MNYIYLDNAGTTPMAPEVIAEMTTKMETTFGNASAVNYYGRQARFILDTSRHLIAQSIN
AQNDNEIVFTSGGTESDNTAIIQTALARKSEGRHIITSAVEHEAVLKPMAYLEEIGFEVTY
LPVDENGLISIVDLKKALRPDTILVSIMMGNNEVGARMPIHEIGQVVAASNAWFHTDCV
QTYGLLEIDVQADKIDLLSTSAHKLNGPKFMGFLYRRDGLNFKPLLKGGDQELKRRAG
TENIPAIAGFAKAVELANQNKAKHRQDYLAFKQQFIDELTALGIEFEVNGRHNEDELPH
VLSIWFKGCPADALLTNLDLAGVSAAAGSACTAGSLEPSHVLTAMYGKDSPRVSQSLR
FSFGITNTKDEISQTVAALAKIIGRLKRK
>5'-methylthioadenosine nucleosidase (EC 3.2.2.16) @ S-adenosylhomocysteine nucleosidase (EC 3.2.2.9)
MKYGIICAMEEEIKELRAQLMNANEENIANMIFYSGQINDHEVVLVRAGIGKVQAGVTTA
FLIENFKVDAVINSGSAGGIGTGLAVGDLVLSTGAAYHDASATVFGYKPGQLPQQPQIF
EADQELLQAVSEAASQAGLQVKPGLIVTGDQFVSSQAQIAAIKEIYPQALCCEMEGAAV
AQVAYQFDKPFLIVRAMSDVGDEDAGQSFDEFIIDAGKKSAQMILNLLA
>hypothetical protein
MTEEQVETKAENLSRTEMRKQQKELLEEAKITKLKRRLNYVIAFLVTVLILILLFMRFINF
>ADP-ribose pyrophosphatase (EC 3.6.1.13)
MDFEEKVISNQNLYDGAILKLDLETVSLPDGRQAKREIVRHQGAVGIIAITPANKIVLIRQ
WRAPLGQVTLEIPAGKIEPGETPSQTAVRELNEETRFQADKLELLTEFYTSPGFADEKM
YLYHAVNLQAVADKLAQDDDEFLQLEELTLAEVQAQIAAGTICDAKTLMAIMIWQAMQ
>Diaminopimelate epimerase (EC 5.1.1.7)
MKMLKVHGSNNDFFILPAANLTKPLSEAELKLLAQKLCDRKSGLHGGADGVLYVTSGH
DASVAGKMRVINADGSEASMCGNGLRTVARYLSTQTGLTDFKVETMHADLEVKAAPA
FAKGVPAFDVEISPVSFKAEDLLMHIGKEQLINETVPQLGSNLKFTAVAVPNPHLISFVD
HETLMGEELGRIATYLNDGQNPLFPDGVNVSFVEVLATNKIFVRTYERGVGYTNACGT
AMSASSLVYSLLYPETSVEDLITVINPGGMVKTKVHQRQDGSYWMSLIGNASFVAEVE
LEQADAMQGDFTKATWNETGEQAAYEEFVANLKK
>Isoleucyl-tRNA synthetase (EC 6.1.1.5)

MRVKDTLNLGKTKFPMRGNLPVRELERQKAWEENKVYQQRQKLNEGKPSFVLHDGP
PYANGNIHMGHALNKISKDIIVRSKSMMGFRAPYVPGWDTHGLPIEQQLKKAGVDRKA
MSVAEFREKCREYALEQVDKQRTDFKRLGVAGEWDNPYVTLDPKFEAAQIRVFGKMA
EKGLIYKGKKPVFWSWSSESALAEAEVEYHDVTSPSAFYGEQVVDGKGVLDTDTYMV
VWTTTPWTIPGSEGITIDPNFNYSVVKPAGSDKKYLLATELVKQNAELFGWESYEEVKV
VKGKELDKVLAQHPFYPDRKLVTMLGDFVTLDAGTGLVHTAPGFGEDDFNVGMKYGL
DIFVPVDDKGYLTEEAGPDFAGVFYEDANEISLAKLKEAGLLLKAMDYEHSYPFDWRT
KKPIIFRATPQWFASVDKIRDEILAAIDEVKFFPDWGQKRLHNMIRDRGDWVISRQRVW
GVPLPIFYAEDGEAIMTKETIEHVAKLVEEFGSNVWFEREAKDLLPEGFTSEHSPNGVF
VKETDIMDVWFDSGSSHQGVLALRDYLKYPSDMYLEGSDQYRGWFNSSLITSVAVSG
KAPYKQVLSQGFTLDGKGRKMSKSLGNTIVPSEVIDKMGAEIIRLWVLSVDSSADVRVS
MGSFQQISESYRKFRNTMRFLLANTSDFDPKQDTVAYKDLESLDKYMLVLANKFTKEV
LDAYANYDFMTIYKKLINFLTTELSAFYLDIAKDVVYIEAAASQKRRSMQTVLYEVAVRL
AKLMTPVMPHTTEEVWEYLKEPEEYAQLSEMPAVEYYDGQDQLVENWTRFMKVRSD
VLKALEEARNAKLIGKSFEAHVNLYVSEGVKAELDSLNTNLRQVLIVSALDVLPLEEAPA
DALKLDAVTVKVEHASGEVCPRCRMIKEDVGADQDLPKLCAHCAQIVRANYPEAVSAG
LED
>Cell division initiation protein DivIVA
MALTPLDIHNKEFHVKLRGYDQDEVNDFLDQVIKDYEQALKENERLTDSLKQNQEKLK
YFNDLKDSLNQSIIVAQEAADKVKANSQREAEIINREAQKQGQDIIDQANEKARHILDEA
SKKAKKLAVETDDLRKQARIFRQKLQVMMESQLEVVKGSEWDEILKQGDTSSFEEIQE
AIRKDESLDKDGSSEVQSVADDISFSNGDSQTDTENDNETVVVFPDASDDTDDNEFAD
NVYINN
>FIG001583: hypothetical protein, contains S4-like RNA binding domain
MVNPTIYQHFRVEERPLVDEFMALIAQAQTEYRPILTHFLNPRERYIIQTLLGNSDELKL
EAHGGYLAAERKRMLIFPSYYEPQASDFEVTPLEISYPQKFAELSHGQILGTLVNSGLN
RNVIGDIITDGERWQVFTQTNMKDYFQSQIERIAKIKVKLVELDPAQVIEPLDEWQQASI
TVTSLRLDAIIASCYHISRQRAKELIKSRKVQLNWMLIEKADYTLDSYDIISVRGHGRIRL
DEVLGTSKKGKLRLGVSILGK
>hypothetical protein
MASICEPYLEIFRFIPPLGGIDFSPLVAFFALGIVEKGLLFFLSLIL
>FIG001960: FtsZ-interacting protein related to cell division
MNNMLKRFFGDDYDEEDYYEEDETPVTSNSDTSRDTKVMSLVGKNKGQTVEKKIMIF
EPRVFSDVKQIARRLLAGEATLVNFQRMDDDSAHRVVDFLSGVVFAVDGEIRRVGETI
FLCTSSSFTIEGDLSGFGQSKNSFN
>Cell division protein FtsZ (EC 3.4.24.-)
MSFDITLDSNTEESQAVIKVIGVGGAGGNAVNRMISDNVQGVEFIVANTDVQALKHSNA
EVKIQLGPKLTRGLGAGANPDKGQKAAEESEEAIAQALEGADMIFVTAGMGGGTGTG
AAPIVAKIAKEQGALTVGVVTRPFTFEGQRRAKYAAEGVAQLKENVDTLVIISNDRLLEI
VDKNTPMLEAFHEADNVLRQGVQGISDLITSPGYVNLDFADVKTVMENQGAALMGIGT
GEGDNRIEEATKRAISSSLLEVSIDGAEHVLLNITGGSDLSLFEAQAASEIVTQAASSDV
NIIFGTSINEELENSVVVTVIATGIDKKNESKKPKRILNERPRASKQETDGDTGSKDKFN
DWNISRELDNGPKVTPKTRTNRFENVEKTQFDVFGSGSAQSVDSDDDNLDTPPFLRH
RH
>Cell division protein FtsA

MDNSGLYVGLDIGTTSIKVIIAEYVKGQLNIIGSGNTLSAGLNRGVIVDIDQVVDAIKDAV
KQAEDKAGVDITDVIVGLPANMLQIEKCQGMISIGDDANRAKEITDEDVQRVTEAALVQ
NLPPERDVIDIAPDEFVVDDFEGIRDPRGMVGVRLEMKGVMYTGPKTILYNTKKCVER
AGLRVQEIVVAPIATAQLALDEGEKDFGAIVIDLGGGQTTAAVVHDHKLKYTYVDQEGG
DFVTRDISVVLNTSVENAESLKRNYGLALASGASSEETFAVQVVGKQEPRMVTEEFLA
QIIQARLEQIFGKIAQALGKVGALDLPGGIILTGGMADLPGIAELAEEIFGVNVRIFRSNE
MGLRIPSYTQVLGLVTYVALQNEVALVVKHTIDKKLGTSNVNAYKNEQVKDRQAEVQS
EELDDYEEAADYPNEEKKGGFFSKIKNMMAGLFDED
>Cell division protein FtsQ
MKKKVNKPVTSEAELTPWEKAQAKRAKQQTPKFAKQKQKKRQKRKLPLLAEVKRRRV
EKTGGSLLVIFATIAIVSLYFILPISRVAQVKVEHLATKDAQVVVKASKIKQRQTLLSVLLN
KNKAEKTIVTNVPFVNRAKVTMAGSQIVIKVNQFRRVAFLEKDKHYYSLRQNGSTNLSE
TTNISGNVPVIVYKANTNPTLLAQQLNLLPNRLLGAISEIHFTPTKSDPTKLLLYMNDGN
QVIANAETLAKKLAYYPSIVAKMNYQGIIDLEVGAYSYPKNSNE
>UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase (EC 2.4.1.227)
MRLLVSGGGTGGHIYPALALIEALKKKDPTAEVLYVGTTRGLETRIVPAAGIPLKTLKIQG
FKRSLSLENLKTVYLFLKSVHESKRIIKEFKPDVVIGTGGYVCGAVVYAASKLKIPTMIHE
QNSVAGVTNKFLSHFVDKIGICFEDAASQFPEKKVILTGNPRAQQVALMKKTGRLYSEY
GLEKNKKTVLIFGGSQGARRINEASTAAISAWKQADYQVLFVSGQGHYEKIMAKLADK
KLPENVHIVPYIADMPAILPEVDLIVGRAGATSLAEITALGLPAILIPSPYVTNDHQTKNAQ
SLVTRQAAVMIKESELDETSLVAAINKLMTDDKERMQMAANCKQMGMPNAADKLLAVL
ESLIK
>UDP-N-acetylmuramoylalanine--D-glutamate ligase (EC 6.3.2.9)
MKKITTYQGKKVLVLGLGKSGINAAELLAKLGAKVTVNDKNEPSDLNQVAALEAQNIKV
VTGSHPLALLTDIDLMVKNPGIPYTNVLVAEAQKRGLKIITEPELAYEILEGQMIAITGTN
GKTTTTTMITLMLNQASAKQHAYKAGNIGISASKQAQTLGQDDILVTELSSFQLLGISAL
HPHIAVLTNIYEAHTDYHGSRENYVKAKMRITMNQTAADYFVVNWDNPEWRELSKQS
QAQVVPFSRNNLSQKGAYLKDGILYFRDEAIMAAADIKVPGTHNVENALAALAVAKLM
GVPTAAIVEVLKTFSGVKHRTQYVRTINDRKFYNDSKATNMESTEKALAGFKQPVVLLA
GGLDRGFTFEKLVPFFKDLRGLVVFGQTAELIAQAGKAAGVKHIIRTENAVTAVPLAYQ
LSQPGDIVLLSPACASWDQWKTFEERGDKFIEAVTKLAQEVEEK
>Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13)
MDIIAMLIPLVVSFILTTLFTPLFINYAHKKKQGQMIREDGPTWHAKKSGTPTMGGFVFN
FVIVITALLTAWLQDNLNSAVIILAFILLLYGFLGFFDDSIKLWKKQNEGLKPWQKLLGQIV
GALIFMGVYLHEGLPLSLALGGYELNIGFIYVLFVIFWLVGFSNAVNLTDGIDGLVAGQA
TIAFGAYAVIAYVQAQYAVLTFCLAVMGALLAFLGFNHKPAKIFMGDMGSLALGGALAA
VSILLHHELSLLLIGIIFVCETASVMLQVAYFRKTGKRLFKMSPIHHHFELSGWSEWRIDL
TFWGVGLLAAVIVDLIVIL
>UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase (EC 6.3.2.13)
MLASELISKLKFKEVENFTDFEVALVTQDTREVVADSIFIAIAGHTVDGHAFVDQALQSG
AKLIIAQKPIKVGVPVVYVENTYKAMAILSSAFYNYPSQKLQVIGVTGTNGKTTVTHMID
AIFRDQGQKTGLIGTMYRRIGDEVLKTNNTTPDIITLQRTLAKMATVNVSEVAMEVSSVA
LIQGRVWGVDFDVAVFTNFSQDHLDYHKTMANYAHAKSLLFAQLGNSYGGKEKVAVF
NLDDPMSKKFMQDTSAQVLTYGVDSPAMIRAENIDIHSQGTDFDLIVFKQRYHLHVKTV
GLFNVYNLLAAFGAAYVSGIATEDIISALTKFPGVKGRLQLLSHKQVSAIVDYAHTPDGL

LNALETINDFAKGKVYCVVGCGGDRDHDKRPKMAQIAVTNSDLAVFTSDNPRTEDPQ
AIIDDMLKGVEAGTAKVIVDRSQAIRWALEQAQPGDVVLIAGKGHEDYQIIGKQKYHFD
DVEEVTNYFKQA
>Cell division protein FtsI [Peptidoglycan synthetase] (EC 2.4.1.129)
MNMKNKRTLNKQKMQLKRKRLGIGLFFTVFLIFVIFIGRFFYVGVFHKVAGIDLNKKVAQ
LYANKTQIKAKRGSIYDTNGNTIAEDTTTYSVYVVLSKSAVYYGKRAYLADKDKEKAAR
VLSKNLPISYEKVYRILNPDNPDLYQVELGTAGKGISLETKKAIEAANITGINFTPTVARLY
PNGNFASYFIGLTATMNDEISGTMGIEKVYNDFLKETNGTRSYMKDNSGTEIPGTKSKY
KKAKNGADVYTTLDSQLQVYMENLLTAAQKKYQPQSMTAVLMNAKTGAILAASQRPTF
NAQTKSNIPVWRNELLAETYEPGSTMKIFTTAAAINSGHYNGNAYYNSGALTIDGSKIY
DWNRAGWGSITYDQAFIRSSNVGMALLEQQMGKKLWLNYIQKFGFLKETNAGLGPES
KGSISYNYPIEQANTAYGQGIDVTVFQMMQGFSAIANNGKELRPYMIKKIVSADGNKVL
KQEGKKVIGQPIKATTAKQVRNLMTQVVTDEHGTGTAFKSDQYEVGVKTGTAQIASPS
GGGYLTGDTNSIFSVVGMAPIDNPKYILYITVKQPKTFAGKTSGEMLAEVFNPLMERSM
KETAASTSKKAVTISNYKSQSSSEVAEKLQAAGLVVTVLGNKKQIQAQSTPANQELFVG
SRIILKTGGTMTMPDIKGWSRSDVVKLATILGLKLKIVGAGYVSEQSVVANSNLSGVKEL
TVKLS
>Cell division protein FtsL
MAQPALKEWTETTNPSPVHKINKQKRLAKKVTLVGKDFIYLAAVVIIAFAVMTSVVKSTV
ALTNAQHELQTIRKQQTKLQVQNTNAKQEISELSSRSRLSKIAASAGLKLNESSTRNVT
K
>rRNA small subunit methyltransferase H
MTEFEHVTVLLNEAVDNLNIKTDGIYVDCTLGGGGHSRKILEQLGPKGHLYAFDQDQT
ALDYNQENLRDWLVAKKLTLIKSNFQQLKQKLAEHGVSHVDGVLYDLGVSSPQFDVAE
RGFSYRQDAILDMRMDQSQALTARKIVNEWSYGELVRIFTRYGEEKFAKSIARQIERKR
QEEEIQTTGQLVELIKAGIPAKARRKGGHPAKKVFQALRIAVNNELGVLEDSLGQALSLL
NPHGRISVITFQSLEDRLVKTMFKEKTSLPELPRGLPVIPSELELEYRLVNRKPITPSQA
ELEANHRAHSAKLRVIEKV
>Cell division protein MraZ
MGEYRHSLDTKGRLIIPVKFREKLTPTFVLTRGMDGCLFGYPQSEWQKIATKLQSLPLT
KKDARAFVRFFYAAATECEFDKQGRINIPSMLRQHAQLTKKCVIVGVSNRFEIWAEEN
WLDFCEQTATDFDEIAENMIDFEL
>Aspartokinase (EC 2.7.2.4)
MKIIVDKFGGTSVQDEEARKKAMAHVAHAVSEGYKVAVVVSAIGRNGAPYATDSLLNL
VGGEKTALNGRELDMLVSVGEVISTAVFTQMLKQNGYKATALTGPDAGFRTNNDYQN
AKVIDVQTDKIMAAFEEVDIIVVAGFQGLSKEGHVTTLGRGGSDTSAALLGAALGAERC
DIFTDVSGMMTGDPRLVKKARFLKEVSYEEVANMAHEGAKVIHPRAVEVAQQAHLPLR
IRSTWDGVNELGTLISNRNLSISHYRTVTGVAHQVGLTQFSIATSDVTADKVFNLLAANK
LSVDFINISPDRIVFNLNREQGEAAQALFTKEKIKAKVVADCAKVSVIGAGMLNTPGVTA
KMVAALSKQGIQILQTTDSDTTIWVLVHGKDLPKAVNALHDTFLLND
>hypothetical protein
MWPKKLLVMDIISPLYLVPLAFFSQEIWGLSLSLYLIFSLALLGLLMTLKQILVKGDIILSRF
LRRYWLLVDLVLSLSLLVVLVSRIIMFFN
>Aspartokinase (EC 2.7.2.4)
MKIIVDKFGGTSVQDEEARKKAMAHVAHAVSEGYKVAVVVSAIGRNGAPYATDSLLNL
VGGEKTALNGRELDMLVSVGEVISTAVFTQMLKQNGYKATALTGPDAGFRTNNDYQN

AKVIDVQTDKIMAAFEEVDIVVVAGFQGLSKEGHVTTLGRGGSDTSAALLGAALGAER
CDIFTDVSGMMTGDPRLVKKARFLKEVSYEEVANMAHEGAKVIHPRAVEVAQQAHLPL
RIRSTWDGVNELGTLISNRNLSISHYRTVTGVAHQVGLTQFSIATSDVTADKVFNLLAA
NKLSVDFININPDKIVFNLNREESRLAQKLLEAQKIEAQVVPDCAKVSIVGAGIMGTPGV
TAKIVAALSKQKIQILQTTDSYTTIWVLVHGKDLVAAVNALHDTFLVKEA
>hypothetical protein
MKEKKKKTTLQKITMVVVWIMLIFTLLSVILAAVGGVMGY
>Hydrolases of the alpha/beta superfamily
MIKKRWLKYGLVGLMILLVGGILVGSNYLVSYAFTRGVPGPSNHKKAAVLKADQVWLK
NVKQEQWQQRAANSSLRLVATYVPAAKATNKTIIVAHGYHENHTSMASYIRLFHNLGY
NVLAPDNRGSGKSQGKYLTFGWLDRLDYVNWIKKVIKYQGQNSQIGLFGVSMGGATV
MMVSGQKLPSQVKAIVEDCGYSSVEAELTNQLKAQFNLPKEPFIMLASMIAKIRVGFDF
RQAKATAALRKNTLPLYLIHGDSDHFVATKMVKQNYAAQAGTKKLWISKKTGHAMTYY
NYPHEYQQRVGAFFAKYLQ
>Cardiolipin synthetase (EC 2.7.8.-)
MAIINLVVLGVFIVNDIIAIITVFREKRDIAATWAWLLVLTLIPVVGFIFYLFAGKKISQEKIFD
LKTQERTGLAQLVSLQKEQWQEQELMPKEILTVESRQVTKLFLETDEAVLTKHNQVTL
YTDGQEKFAALLADILAAKKYVHVEYYAIFSDEIGTKLKQALITVAKRGVKVKVIYDSLGS
RGQRHGFFKDLEKVGGQAEPFFGHRQSPIHSPRFNYRAHRKLVIIDGEIGYIGGFNVG
DQYLGKSKYFGNWRDTHLRIVGNAVIALQSRFFMDWNATIKGDKKRTKLTYADSYFPL
SPVRGTTSIQIVSSGPDNENEAIKLGYLKLISSANKTIDIQTPYLIPDDSIYETLSVAALSG
VKVRIMIPSKPDHPFVYRATQYFAKALLEKGVEIYQYEDGFLHAKTFIVDDELASVGSAN
VDFRSFKLNFETNAFCYDRALVQKLQASFEQDLAKCSRLDEAYFRQQSVWLRFKQQF
SRLLAPIL
>amino acid ABC transporter, amino acid-binding protein
MKKMKYVWKILALLIILTSSLAGCENVSKRANQTDTWNKIEKRGTVVVGLDDSFVPMGF
RQKNGQLAGYDIDLAREVFKRLGLKVSFQTIDWSMKETELRNGTIDLIWNGYTKTKAR
EKTVAFSHTYLKNEQILVSKKKNQILRPADMKNKLIGLQSGSSGYTAYNEYPSMLKQYP
KQAIQYDTFTNAFLDLNANRIQGLVIDSVYANYYIAHLKDSASYRTLKVGFPNEDFAIGM
RKGDKTLRHKINLTLAQMAQDGSLNKLNQKWFGENNQISFKDVKINKK
>amino acid ABC transporter, ATP-binding protein
MLELQNLTKSFGKRKIIDGLNLQVADGEILCIVGESGGGKTTLLRCISGLETIDSGQLLID
KQPFDPSSSTSEAVIGVVFQEYNLFPHLTVLENITLAPKMVLKKDEASAQAEARQLLAK
LSLSDKAELYPYQLSGGQKQRVAIARALAMQPKILCYDEPTSALDPNLSQAVKQIILDLK
KSGMTQIIVTHDWDFADEIADQILKVKPLKF
>ABC transporter membrane-spanning permease - glutamine transport
MDYIVEILPSLMAGAKLTLEVFCWTLVGSVPLGLVVALGLMSRFKVVKGILRFYVWLMR
GTPLLLQLIFVFYGLPILASWLILPRYEAALVAFILNYTAYFAEIFRGGLQSIDKGQYESAR
VLRLTYWQTIRKIIAPQVVKIVLPSIGNEIINLVKDSSLVYVIGLGDLLRAGNIATARDVTLV
PLIMVGIIYLMMTAILTFILGKLENHYSYWK
>Septum site-determining protein MinD
MGKAIVITSGKGGVGKTTTTANLGTALALQGKKVCLVDLDIGLRNLDVVLGLDNRIIYNIV
DVVEGRAKLHQALIKDKRFDNTLFLLAAAQNTDKDALEPEEVKQLVNELKPEFDYVLLD
CPAGIEQGFMNAVVGADEAIVVSTPEIPAVRDADRVVGLLEKYDLEQPPMLVISRIRKH
MVNNGQMMDVDEITHHLGIKLLGIVFDDDAVISSTNKGEPIVLADNNIVSQCFKNIARRL
EGETVPLLTMEEAQQEGFFSKIAHWFKRA

>Septum site-determining protein MinC
MQTVILKGHKDGYEITLKDDADFTLITSELRQLLEGLKQDEGSKQTTITFKVNTGARLLN
TWQKKELEKVFGDYPYFAIHKITASVIDKQEAWDFMENHNIHINAATVRNGQVLELTGD
VLFVGAVHQGGVLQTSGSIYSLGKIEGIVHAGYDNNSRAIIAGEISKAQQVRIGDLVDIVE
EKTIPTSRCLVYVNDLHTLTYADISELKALRPKLFVKIGGF
>Rod shape-determining protein MreD
MRYLFPIGLFLALFLDGALSGHWANYLFTQTTFMESRLVLLWLVMAVCYSDARHIFVW
ATVAGLLFDSYYTGILGLFTVMLPLIVYLTREIMKYFTPSFIVVLLIYLIDITVLTILFYWANV
LIGFTSVSIADFIGQNLGPTLIYNLACYVILFMPLHWLFTKFA
>Rod shape-determining protein MreC
MQKAFFNKKLVIILITLIISFLLIAFSISIRNNRNTPSFVQKIGNESASLVNRVVDLPLTAISN
LETAISDLHNTYEENELLKKKIDSLAATKIENQTLKSENKQLKQQLHLNKSLTDYTKVTA
YVLARTPSAWQNQIIISKGTNAGIVKNSAVVSDQGLVGRVTEVNATNSKVELVTTKNDA
ANRFAVQLTNDKGETVNGLITGTSGKFLVMGQITNNAKVKKNTKVVTSGMGGTTPQGI
YVGKVVKVSTSGDSSNTKIYIQPAADMNKLNIVTVAKKTD
>Rod shape-determining protein MreB
MFGIGTKNIGIDLGTANTIVYAEGKGIVLREPSVVAKNTKTGEIVSVGSEARDMIGRTPE
SIVAIRPMKDGVIADYDTTVAMMKYYIEKACGRRAGKPYVMVCVPSGVTEVEKRAVIDA
TRVAGARDAYVIEEPFAAAIGAGLPVMDPTGSMVVDIGGGTTDVATISLGGIVSSRSIR
MAGDKIDESIIYYVRKNYNLLIGERTAEQLKMDVASASVEFAKELDSQTIRGRDLVTGLP
KTIEIKAEDVATAIQEVVEEIITAIKETLEETSPEIAADVIDHGIVLTGGGAQLRHLPEVIAE
ATKVPVFVASEPLDCVAIGTGESLKSIDVMKKK
>Cold shock protein CspG
MLTGVIDTFETEKGYGFIKPDDSRDKIFMYYTALKQGDFKTLSPGQKVKFLIMEGKKGP
QAVMVELID
>DNA repair protein RadC
MSLINVYPKNYLQAKVAKFSSAALTDRELLTFLFEQHYSQNKSKQLVDKFFKEYESLGY
LEYFTLTDWQKFLADADKAAVFASLCEFARRYRSQRHLNLGQIYSSKEVGEFLIERLGA
NKQECLVLLVLDNKNQLLDMKTIFKGTLNSSLVHPREIFNFALRYSAARFMVAHNHPSG
NLQPSQNDLAMTKRLQQAGEIMGIELLDHIIVGNDNYLSMREEGIIN
>hydrolase, haloacid dehalogenase-like family
MDLDLVLFDMDGLVFDSETVYFEANQIAAKELGLANYTFDYYQQYIGAGDDAMVAGMI
RDYGDEKIVAQFMERSKALVHPLVKQGELKIKPGFHELVSYLKAKQTKIYLASSNFISEI
NFFLENTGLSGTFDGIISADDVAKAKPAPDIFLAAWEKAGCPPKEKCLVLEDSLNGIKAA
NNAQLPVAMVPDCIAPNDYARQNTVAIIADLAKIKNL
>Dihydrofolate synthase (EC 6.3.2.12) @ Folylpolyglutamate synthase (EC 6.3.2.17)
MAITTYEEALAFIHGRTQFKKSPTLDRMRYLVEKLGNPQLQLKMIHVTGTNGKGSTTAY
LRDLLMSQGFKVGTYTSPFITRFNERISVNGQMIADNDLVALVNEVYPLVKELDAEGMG
PTEFEIITAIMFTYFSRVKVDYAVIEVGLGGLNDSTNIITPLVSVITTVALDHARWLGDTVT
KIAFHKAGIIKKQTPVVVGRLPKEALDVVKQRAQKLQAPLYEPQSNYQVTNDNYHGWG
EAFTYRFEDKTYRLQLQMLGSYQIDNAACALTAFLILAHKEGFVPTGRELKAALGRTTW
PGRFEKLTSEPLIVIDGAHNEAASLELKKLLQTRFADMQVKILMAVLADKQADKMIANLA
SLPNVELTLTTFAGPREVTSLKHYQTQFPDLRTEENWQAAFLKIIKELNGDDMLLITGSL
YFISDVRNYFIEEG
>Phage terminase

MQLLTLKIEGSNDMSNHKQSVIAPYQKTLSDEAVKFYERTGLSCYEWQKNLLDPIMAV
DEDGLWVHQKFGYAIPRRNGKTEVIYIKGIWALEHGLNGTISVSLIKKVKCK
>hypothetical protein
MEEINKTKKYRIESVYYEFSVLKIVDEYTHEQYEKIAALNSKWSDYDFDKTDGYIYFDDL
EKELVPPELTPADRKRFIEYLEKEIEVVNK
>tRNA-Leu-CAA

>tRNA-Cys-GCA

>tRNA-Gln-TTG

>tRNA-His-GTG

>tRNA-Trp-CCA

>tRNA-Tyr-GTA

>tRNA-Phe-GAA

>tRNA-Asp-GTC

>tRNA-Pseudo-TGA

>tRNA-Met-CAT

>tRNA-Pro-TGG

>tRNA-Arg-ACG

>tRNA-Leu-TAA

>tRNA-Gly-GCC

>tRNA-Thr-TGT

>tRNA-Leu-TAG

>tRNA-Lys-TTT

>tRNA-Val-TAC

>5S RNA

>Large Subunit Ribosomal RNA; lsuRNA; LSU rRNA

>Small Subunit Ribosomal RNA; ssuRNA; SSU rRNA

>Cell division transporter, ATP-binding protein FtsE (TC 3.A.5.1.1)
MPKTIVEVKNLKKVYGKNEVLKDISEKVEQGQVICVIGPSGSGKSTFLRCLNVLEKPTA
GKVIFEGSELTNISDKELNVLREKMGMVFQSFNLFPNMTVLENIKLAPMKVKGVSETEA
ETTAKELLAKVGLADKASQFPQSLSGGQQQRVAIARALAMNPEVMLFDEPTSALDPEM
VGEVLKVMRDLADSGMTMIVVTHEMGFAKEVADQVWFMADGYIQEVASPDEFFSNPK
SERAKDFLNKIL
>Amino acid ABC transporter binding protein and permease protein
MKKLKQLLILFTLLVTLFIPVSQIKAADNEGAATNSSQTTDDSLQKVKDKGTLVVGLSAD
YPPYEFTAKVDGKTKYVGIDIDLAKQFAKDMGVKLVIKNMNFDSLLVALETHKVDAVISG
MNPTTERQKSVDFSNIYYTGGKYIVINKKDKTKYTTKDSFSGQTVGAQNGTLEYSLIQK
QMDGAKVKGLAKLNNLILALQSGKVAGVLMEEATAKAYVQNNSNLYAFNSKVKVNTNE
TGTAIAFPKGSASLVAAANKTIADVKKADLIDQQYVPNAAQYMQTASKQNTMLNYWTY
FAKGLEYTLLITVVAVFFGFILGVLLALMRLSNGKILHSLAVAYIEFIRGTPLMVQVMFVY
FGVGAIIQSLPALVAGIIAVSLNSAAYVAEVIRSGIESIAIGQTEASRSLGLSQKQTYQYVV
IPQAIKNIWPALGNEFITLIKESSIVSVIGVGDLMYQTQLVQSATYKGVMPLFITMLIYFVV
TFTLSKVLNYFEGKMKHA
>hypothetical protein
MSLKVNQQVDLETLFNKFALDPIEPPKKKADSKQAEAADKSEKKPRSKN
>FIG00743467: hypothetical protein
MEFNEYQEKANRTLYGNEQVLTNCALGLAGEAGQLIYLIKDYTFKGKELDRDKMIHEM
GDVLWYLSQIAQWADIPFDEVAKNNIETLNKRYPDGYSPAK
>GTP-binding protein EngB
MQVHNVELTISAVKPEQYPDKGYPEIALAGRSNVGKSSLINSLINRKNYARTSSQPGKT
QTLNFYNIEDELYFVDVPGYGYAKVSQKEREKWGRMLETYFEKRQELKGVISLVDARH
EPTKLDQQMVEFLHYYNLPLLVVGTKLDKIPKSKWNKAESTIKKTLKLTGKDRLVLFSA
VTKQGKEDIWAWIEQQAQLESGAN
>COG1399 protein in cluster with ribosomal protein L32p, Firmicutes subfamily
MKWSLNELRQAGSEPIMVNEELDLTSALKKRRPEIISASKVAVQGLFSVDKLGVLGYFK
VQVELTLPSTRSLEPTVLPLSFEVTEHYVSYHEHDLSGFEDTDVVIVLDKDILDFDRVIE
DNILLQIPMQVLTKAEQQAQSKQMPSGQDWRVVSESQLEQLRKDEKTLDPRLAKLAD
FFGENQED
>FIG007079: UPF0348 protein family
MSKLEAVGIITEYNPFHNGHLHQLNVIKEKAPTACIVVVMSGNFLQRGEPACLDKWTRA
KEALVNGVDLIIELPPQACIQPADRFAFNGVNGLASLGVDYLVFGAEHAEYDFMNYASQ
VQGLKGEFKKYNQSYAASIQAAITQALGHELAEPNDLLALAYAKANLKLGSPLKLQPIQ
RIQAGYHELALKAGQQIASASAIRSNYQQGRLAELRTYVPHETYADLSANSLVFWDDF
WPFLRYKLLTTTPKELANIYGMAEGIEYRLVQKANELAYGSSFDDWLHAVKTKRFTYTR
LARLAVAMLLNLQTEDVTSYNEAPYFRLLGFTKRGQEYLNAKKKQFTYPLVTKVTQKD
AKGPLAVDYRVGKLYQSLTGKEQDLKKAPLRIF
>FIG145533: Methyltransferase (EC 2.1.1.-)
MIYTTFAQLYDELMDPDMYANWLDLVKTQLPANANLLDLACGSGRLAVSLAQNGYQV
AGLDLSEEMLALADYHARQAGVELELYQGDMTDLSQLENYEYVTCFADSLCYLPTQA
DLQACFQAVYQHLEPGGKFLFDVITPYKTDVVYPDYMYNYSDEEQAFLWNSYSGQDE
HSVIHDLTFFIYDDKKAAYERLSELHYERTYPLVTYKTMLAQAGFKQVEVSADFARKEA
AKTDERWFFICQK

>Ribosomal silencing factor RsfA (former Iojap)
MDSKELLETVVKAADSKRAEEIVALDVANVSLLADYFVIMQANSERQVKAIADEIEEKVA
AAGVQVRDIEGKNAANWVLLDFGDVVVHVFRTETRQFYNLEKLWAEAPLVDVSDWVK
D
>Hydrolase (HAD superfamily), YqeK
MNELIYKEGIFPSNKREELLTRVAASVSEYRYKHILGVEQAARKLALRFNEDEELASIAA
LVHDYAKERSDADFKQVIMANGLDQDLLNWNNFIWHGVVGAEIIKAELNINDERVLNAV
RKHTVGAAVMTTLDKIIYVADFIEAGRDFPGVELARKLAEKDLDLAVAYETKHTLEYLLA
QNKTIYPAAILTYNAYVAGKNMEDFNG
>Nicotinate-nucleotide adenylyltransferase (EC 2.7.7.18)
MTVCTVKTPQVAVLSQVDYQGPKKKVGLLGGTFNPPHLGHLVMADQVLDQLDLDEVL
LMPDNLPPHVDTKETIAAKQRLEMVKLSVEGNPKLGIEDIELKRGGVSYSYDTVKYLKQ
LHPENDYYFIIGGDMVEYLPKWYRIDELVKMIQFVAVSREGFKRTSKYPLLWVDVPTIG
VSSSLIRQKVAQGCSIKYLVTDGVAQYIDKEGLYRE
>FIG004454: RNA binding protein
MAELRGKQKRYLRSQAHEMRPLFQVGKDGLSANWLKQIEDAIEKRELLKINLLQNTMV
EVAEVKEFIEAHSQIQVVQTIGHVLVLYCPAKKKEHRSYSLEVAKL
>GTP-binding protein YqeH, required for biogenesis of 30S ribosome subunit
MTELNEALYCIGCGAKIQTVDKNQPGFTPKSALEKALDTNEVYCQRCFRLRHYNEVAD
VSLTDDDFLALLNRIGEADALIVNVIDIFDFNGSLIPGLHRFVGNNPVLMVGNKEDLLPKS
LKRSKLQDWVRQEANKQGLRPLEVVLTSAKKNHSIDRLLALIDKYRQGRDVYVVGVTN
VGKSTLINQIIKQKTGISDLITTSRFPGTTLDRIEIPFGDGHDLIDTPGIIHRHQMAHYLSAK
TLKLASPQKEIKPKVYQLNEGQTLFLGALARFDYLAGPKQGVVAYLDNNLMIHRTKTRN
AASFYDKHAGELLAPPSKDELAEFPKLVRFEFKVDIKSDLVFAGLGWITVKAGTVVAGW
APEGVAVVLRRAMIYGGK
>FIG001553: Hydrolase, HAD subfamily IIIA
MFNKFKPTWMIESIYDLQPAELKARGIKVVLTDLDNTLIAWNNPLGTPQLREWLEVMKK
AELPVVVVSNNNHQRVKKAVAPFGLPFISRALKPLGRGIKLAKKQFDVQDSELVLVGD
QLLTDIAAANHTNIRSILVKPIVQTDAWNTRINRGIEKIVKKQLLKQGDLKESWGHTLND
>LSU ribosomal protein L20p
MPRVKGGTVTRKRRKRVLKLAKGYRGSKHIQFKVAKQQVMKSYQYAFRDRRKTKSNF
RKLWIARINAAARMNDISYSKLMHGLKLAEIDVNRKMLADLAVNDAAAFTAIVEEAKKAL
AK
>LSU ribosomal protein L35p
MPKQKTHRASAKRFKRTGNGGLKRSNAYTSHRFHGKTKKQRRQLRKASMVSASDMK
RIKQMLSQMK
>Translation initiation factor 3
MVNDGIRARELRLIAVDGSQLGVKTKQEALQLADQANLDLVLVAPKAKPPVAKIMDYG
KYRFELQKKQREARKKQKVVNVKEVRLSPTIDTNDFNTKMKNAQKFLAKGDKVKVSIR
FKGRAITHKEIGRDVLNRFAEETKEVATVESKPKMDGRSMFLMLAPKAEK
>Threonyl-tRNA synthetase (EC 6.1.1.3)
MSQIKVTFPDGAVKEFAAGVTTKEIAESISISLAKKAVAGKYNGKLVAYQAGLTEDGALE
IITKDSDEGLVVLWNTASKVLAAALKQLYPEMHFGQGQATEHGFYFDTDNQAGQVAET
DFEKIEAKMKELVKANLEIETVSYSKEEALKLVAGDPYQTELVEEQAQEGQVIGKKLGN
FVDFSNVVTLEKTGAVKVFKILSVAGAYWKGASSNPMLQRVYGTAFLKQKDLDAELVR
QQEARERDHRVIGNKLDLFFVDPKVGAGLPYWMPNGATIRRAIERYIIDKEVAWGFQH

VYTPVLANLDLYKQSGHWDHYREDMFPPMDMGDGEMLELRPMNCPSHIQVFNHHKR
SYRELPLRIAELGMMHRYEKSGALTGLSRVREMTLNDGHDFIEPDHIEDEIKTLIQLMTE
VYHDFDITDYRFRLSYRDPENTEKYFDDDEMWEKSQSKLKAAMDDMGLEYFEAEGEA
AFYGPKIDVQTRTALGGEETLSTIQLDFLLPERFDLSYIGADGQEHRPVMVHRGIVSTM
ERFTAYLTEIYKGAFPTWLAPKQVEIIPVKNELHFDYAKQLQSKLAAHHIRVSIDDRNEK
MGYKIRQAQVNKVPYTLVVGDNEKAEGTVTVRKYGEEQTQAMSQADFMAMLSRDIES
YSRKEN
>Helicase loader DnaI
MENVGKVLQEKMRTTDMRQRYQKLVTKALADPDVKTFLQAHRAELTQEDIERSASKL
YEYVKIKENIAQGKEIPMPGYSPSLIISQHRIEVNYVPTAELLAQREQLARQTRVRTINMP
KSIRRARIEDYDQTKRQDVLMAALAFIDAYQEQPQEFHKGMYLQGKFGVGKTFLLAAIA
NELAANGFKSTLVHFPSFAVEIKSAIGENNTKQLIDEIKKAPVLMLDDIGAEQLSSWLRD
DVLGVILQYRMQEELPTFFSSNLAMAQLESEYLTVNNRGEAEPVKARRIMERIRFLASE
YLVEGPNRRQK
>Helicase loader DnaB
MLRSGREIVAGWKDLDVQAGFIITAKTWLNLEMVSQALPLYQPLMGVSAYGVYAFLSA
NVDYKPLLSKRKAHKELLVNLGVDLPTFYEARIKLEALGLLKTYMQEDSLGKLYIYEVQ
QPLMGQSFFTDDLMAALLLETVGQRQFEQLERKYNPKQLKHPQAVEITKNLLEVYKLT
QTDNLYTPKVIGGQKVATSKVVANLEDDLDKQFLQQLLTKSFINAQVLVEHYDALKTVH
LLYGLDELQIVQLLEQAANVQTGQVDLLKFKQLANQQFESRLNLKASIQTEVAPPAKVD
NSGMTTEDLQLIAACRAYLPLEFLEKLKTDMGTFATKNERYLVQNLVERQTLPNAVINIL
IHYLLSDQNKPYFGSNFESIAADWAKKQISNPEDAIRQVRSFYQRKDKQAKQTKFQPV
KRRVIQKESLPAWAKPDYQEKDTPDDPAKSKQIEQLMAKINAKKKSES
>Ribonucleotide reductase transcriptional regulator NrdR
MKCPQCHNNSSRVVDSRPADDGHAIRRRRECEQCGYRFTTFERIEVTPLLVIKRNGTR
EEFSRDKILRGIMRAAEKRPVGMDEITKIVDQVEAKVRAVGGNEISSQAIGQYIMDILVD
VDEITYIRFASVYRQFKDMRAFADELNELMDREHH
>Dephospho-CoA kinase (EC 2.7.1.24)
MTYILGLTGGIASGKSTVSAYLAQNGALIIDADLIARQVVAKKSSGLKQIVAKFGGEILTA
SGELDRKKLGKLVFSNKELLKALTDITGPLIRAEILREIEAAKKAQVKLVVLDIPLLFETGY
QTLCDKVMVVTIPSKLQLERVMKRDNLSAAEARKRIANQLPASKRNELADVLIDNSKSV
AETYQQVLKWLKIEALC
>Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)
MPELPEVETVRRGLMSLVAGSKVSQVEVLYPKMIIGASDEFKQRLTGQTLERIDRRGK
FLLLRFSNEWTVISHLRMEGKYFVRQADAPVEKHTHVIFTLADGRQLRYNDVRKFGRM
QLVQTGSEQTVKSLAKLGPEPTAAEFSFATFYEQLQKKKKAIKPALLDQNLVVGIGNIYA
DEILWLSKVHPATPCNQLTSKQAQAIHDNTIEELSKAVEAGGTTIRSYTDAFEQTGSFQ
FALHAYGQTGKPCPRCQTPIEKIVLAQRGTHFCPHCQQVK
>DNA polymerase I (EC 2.7.7.7)
MAEKKLLLIDGNSVAFRAFYALYNSVDRFTNHDGLHTNALYGFNTMLDKLLKTEEPSNI
LVAFDAGKKTFRTEKFAAYKGNRNKTPKELTEQLPYFPDLLAGYGIKSYQLVNFEADDII
GTMAKAAENQGYQVTIVTGDRDLTQLTTDKVTVAVTKKGVSEVEYYTPEYIAEKFAGL
NPRQIIDMKALVGDSSDNYPGVTKVGEKTAIKLLTQYDSVEGIYQNIEQMKQSKLKENL
LNDREVAFLCKDLAEIRQDAPVEISLADTKWAGVDDEKLRAFYQEMDFKSFLSKMQVS
EEKTAASLEYTVLTAANLDLISERQPKEVAFYLEMDGPNYHKAELTGFSLKLDDSYFVS
RNCQLLKQTALKELLAKQAVDVFDGKRTYVGLKRLGLDLARVDFDLLLVSYLLDNSEN

SNDLGILAQAHGYQEIASDEVVYGKGAKFKLPSDDAVYFKHLVHKVKAIAALKADLYQQ
LKANQQEQLYFEMERPLSLVLAEMELAGIKVDTDCLESMQSKFKERLAEIEQVIYQEAG
KEFNINSPKQLGTVLFEDLGLPYAKKTKTGYSTAAPILEKLQGQAPIVDNILQYRQLAKL
QSTYVEGLLKVADEADHKVHTRYTQTLTATGRLSSVDPNLQNIPVRLEEGRKIRQAFVP
SKPGWEIFSSDYSQIELRVLAHISGDENMQAAFKSGLDIHANTAMNIFNLDSPDEVTPN
LRRQAKAVNFGIVYGISDFGLARNIGITNKEAKTFIDNYFAAFPGVHQYMQKSVALAKK
QGYVETLFKRRRYLPDIHAKNFNLRSFAERTAMNTPIQGSAADIIKVAMIRMNEMLAKE
NLAAKMLLQVHDELIFEAPKEEIPILERLVPEVMDSAVKLTVPLKVESKHGRSWFDAK
>Integral membrane protein, interacts with FtsH
MNNQERMIVSNEQGLNRFLSKMYAFMATAVGVSALVSYLVMGVFYSQVVTLFATNKL
AIWILFILQFALVIGISFKADRNPVASATMLYLYAGFEGLIFGIIIPMYTAQSVTLAFASAAT
VFVVLALMGTNTKKSLATIGRHALAMLIALIVVSIVNLFLQSSVIQFAFSIIGVLIFSALTAY
DAQRFKLMYLQYGHQVNTTNLAIMGALQLYLDFINLFIQLLNIFGIGSNKD
>MaoC like domain protein
MTEDVHRQGKKIDELKEGDSLTVTETISEREILLYLGLTNDTNPLYMQPDYAAKTEYGQ
TIVPPILLTGIITSSISKLLPGPGSEIVNMDVNLISPVYHNQIITFDFEVIKIDTMKEVVTISVE
GVDEDDERVLDAVIMVRPPKLEQLALDETSSLEGVDINE
>Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA (TC 3.A.1.5.1)
MQKKKIVGLVMALGLAGLFLTACGKQQAASKQQELRLSASAPLDTIDISKATGYGQTG
NVFESFYRLGKNGQPTAGLAKKATVSADGLTWTFTLRDAKWSNGDQITAQDFVYSWR
RSLTPKTASPYAYLFSGVKNADAISKGKLPVEQLGIKALDNKTVEISLEKPIAYFKVLMAY
PLFGPQNQKVVEKYGKKYATNSKYMVYSGPFVIKDWTGTSDKWSFQKNPNYWDKKQ
VKLDKISYKVVSNPNTGYELYQQDKLDMTPLSSQQVKNYKNSSEFKQYPYSYVAFLAY
NFNATNSANKKALNNQNIRRALSLALDRKVLTKKVFGDGSQVPTGFVANDLAKNPQTG
VDFSKEQAVTKTVTYDANLAKKYWRKGLAETGLKELTFNILASSDTPTNDALTQYLQS
QYTKVLPGLKVKVQNIPGKSALERARKGEFDIYVSGWGGDFNDPITFLQIPLTGTSYNY
GNYSNSEYDRLIAKATNEDANDKLARWNDLVAAAKLFNSQQGVSPLYQQTTAYLQKK
RVKGIIHNTAGTQWNYKYAYLK
>UDP-N-acetylmuramate--alanine ligase (EC 6.3.2.8)
MNMDTTYFFVGIKGTGMSSLALILKDKGCKVAGSDIDKYTFTQRGLEQAGIEVLPFSAQ
NIKEGMTIVKGNAFGEDHEEIKRAKELGLKVQTYPETVEMLVTETTSIGVAGAHGKTST
SGLLAHVLSGVAPTSFLVGDGTGKGISASRFFVFEADEYRRHFVAYHPDYTIMTNIDFD
HPDYFTGIDDVCDAFETLARNTKKGIFAYGEDKHLRKLKYDVPVHFYGTEADDEFRAE
NIKRTTKGSTFDVYHFDEFVGSYQIPMFGEHNVLNSLAVIAVAYFEKVDQDEIKRELLTF
KGVKRRFSEKRVADMVIIDDYAHHPAEIKATLDAARQQYPDKKIIAVFQPHTFSRTIALM
DEFAASLNLADKVFLTEIFSSIREKDGKVSAKDLGAKITKGGAILSMDNMSPLLDFHDDV
VVFMGAGDITKYEHVYEELLSNLSLKNN
>Cell division protein FtsK
MEKYTGPAFYQTKTKVAGENNRFKAKHETQGDKRVATNLTRKTAYNNSRGFHPRTFP
PFRQQGPVVDLKAKYQKLSQQLKKAKTDFILFQDESIASQQELELFDLRQPVVDFETPS
KVTTPVKQAEMEAAVVEASLAKAVEDSSEAVVVFEPHELSEQPLTQTETETVADSQPE
AAQVAAVIEEEAAKKEEPVVEFEKVQEKQPVKTGDDKQPLATTKPVVQTPAQPKQASII
SDVSGEAKQAAGGYERPALTLLPAPISENTEQVTTWVSEQATKLDETLKSFRVDAQVV
NWTIGPAVVQFEVKLGRGVKVSKITNLTDDLQLALAVRDVRIEAPIPGKSSVGIEIPNRY
ARPVNLAEVLGSQAFLEAKSPLTVALGVDLFGNPQVTDLQKMPHGLIAGATGSGKSVFI

NSMLVSLLYKASPAEVKLILIDPKAVEMAPYQGIPHLLAPVISDAQAAAAALKWVVAEME
RRYELLVAAGVRNIEGFNTRAEESGQYGLKLPYIVVVIDELADLMLVASSEVQDYIARIT
AKARAAGIHLIVATQRPSVDVVTGTIKNNIPTRIAFMVSSQVDSRTIIDTAGAERLLGRGD
MLYLGNGSSQPVRLQGTYIGDEIDQVTDFVRSQGQPKYAFDPENLKKKLTEDEKQDKL
FPRILDYIVNEEAISTSKLQREFSIGFNRAAGIIDQLEEKQYISRSNGSKPRKVFLTQEGL
NKLRAQG
>Phenylalanyl-tRNA synthetase domain protein (Bsu YtpR)
MLISSYNPSQMGDVLVTIINPDKSTQASEQKQDVTRIYEPKTDLTLGYNFFKLGEYLPDL
KGQGQVFLTTAQVATLNDHLEAVGFKAELEADLSPKFVVGKVLEMTEHPDSDHLHITK
VKVDNEQVLQIVCGAPNVDVNQHVVVAKVGAMMPSGALIWPGKLRGVKSDGMLCAA
RELALPNAPQKRGILVLDADEFPVGQAFDFEKGRQLFIN
>FIG009688: Thioredoxin
MEMLAKMNKAEMEAKLANGAYVLFFTADWCPDCRFIKPAMPEIEAEFKEFTFIQVDRD
ENLDLCQELMVMGIPSFIVYRDGKEVARFVNKDRKTKAEVTAFLSAAK
>tRNA (guanine46-N7-)-methyltransferase (EC 2.1.1.33)
MRLRNKPWAKPLIEANPQWVVTNPSEYRGKWQERFEKAQPLYIEVGMGKGRFIVEMA
KKYPHYNFIGLEMQTVATGIALKKQLVEQLPNLQLVCANGAGLTEFFATDEVAGIYLNF
SDPWPKNRQEKRRLTYKSFLAAYKEIMVADGQLEFKTDNRGLFEYSLASLNNFGMTFE
KVWLDLHADEEAVVDNVMTEYEEKFSSKGQPIYKLAAHF
>aminoglycoside phosphotransferase family protein
MDFDLENGWRILPIDGDTGTAYMGIRQQEKLFLKRNTSPFLAALSIEGITPRLVWTKRM
TTGDTLTAQEWLNGRCLHKAEMDSEAVSALLYRLHHSALLKRMLKQVGGQVVTPEEL
LQRYFRGLPQDLRHHPLLKTVANRLRAKQPQVKRSNYEVCHGDLNHKNWLLSDQGN
LFLVDWESAVIADPAYDLSMLMCQYVPRNNWQQWLQTYQANGQEQPSDDLMLRINW
YSFCYLLLQIKAKHQRGHFHEMNQDIIKLDEVVKKQTDI
>ABC transporter, permease protein EscB
MDELWQKRRQAYHKRLLKYLRYVLNDHFVLALLFICGAVAYNYNNFLKTYALSQGLGW
TRIIIALILGLALSFGRLATFLEPADLVFLAPKESDLKHYLKKSFVYTSLVCLCEQAVIYLLL
CPYLKIILNFSLGQLICLGLSQLLLKELGLYWDYLRGFAPLKQVGFMKLASLILVVLASYF
PLGGLILSGGFWLGCLYGLKVNQQVVAWKFLIEKEAARMQGLYRFFNLFTDVNLVTTS
AKRRKYLDWLLPTRPAYTFLYLRGFMRTSEYLGIYLRLIVLGSLLLIFIKNYYLACGVAGL
FLYLIGLQLFSLSDRYQEIVFVYLYPLSDKQRLREFSGILLRLLVPVCIIFSTLVLVTGQNV
YLFSCLLLELVLESYLITKPYFVWRQGRVRKN
>ABC transporter, ATP-binding protein EcsA
MTLEVTDLVGGYSQIPVLKQVSFTVAKGQLVGLIGLNGAGKSTTINHIIGLLRPQKGSIKI
NGVSLSEDSKRYKQSIAYVPESPILYEELTLKEHIDLTIMAYNLDEKQAWQRAQKLLTTF
RLDKKLDWFPANFSKGMKQKVMIVCAFLTDADLFIIDEPFLGLDPLAIRDLLALIAQRKA
QGAAILMSTHVLDTAQNYCDKFVLLNEGQVTASGSLAELQAKYGNSQATLADLYLQMI
ERN
>Histidine triad (HIT) nucleotide-binding protein, similarity with At5g48545 and yeast YDL125C (HNT1)
MEDCIFCKIIAGEIPSTTVYEDEVVKAFLDITQVTPGHTLVIPKKHVPDLFSYDEELASQV
FARIPKIARAIKAANPEVVGLNLVNNNGQAAYQSVFHSHFHLIPRFGAEDGFAMKFTDN
SANYTPTDLALIAKSIETKVEK
>hypothetical protein

MTKPITKILCISALAAGAYLKRDQILTTGRQLKQTVDENIAAVDDWLAALQKVNASLTHL
AQEVDAALPVISELGEAVAKFSEESSSHIEALTDFDLKK
>Foldase protein PrsA precursor (EC 5.2.1.8) @ Foldase clustered with pyrimidine conversion
MKKWLVACAGVLMTFSLAACSKTVATTSGGKITESEYYSSMKKTSSGKAVLEQMILDK
VLEKDYGSKVTSKKVDKQFNEQKEKYGSTFSTVLSQYGYTNSSFKKAIRSNLLLEEAV
KDHTKITNKMLKKQWAKYQPEITVAHILVSKKSEAEDIINQLKADPTYANFKKLAKKNST
DSSTASDGGKLPSFDNTDTSLASSFKKAAFALKQGEYTATPVKTSYGYHIIWSVKNPG
KGKMSDHISDLKKQIIANKESDSEYMQKVVAKVLKGGKVSIKDKDLKDVLSNYLSTSSS
PSSK
>3'->5' exoribonuclease Bsu YhaM
MTKKIFDYELNENLDLFVLIKAAQARLAKNGKQFIAFEFEDKSGEISAKYWDASDEEVK
RYQSGQVVHLSGKRENYQGNPQIKIYKLRLATEQEPNDPSLYVQGAPQSVTEMQAEL
EKFVFKIHNSTWNRIVRVLLQKYQAQFFTYPAAKKNHHAFQGGLAYHTLTILKLAQGVS
QLYPQVDESLLYAGAILHDLGKVIELSGAMSTQYTVAGNLLGHIVLVDEEIIKACQQLGID
TAKEDAILLRHMILAHHGLLEYGSPVTPHILEAEILHHLDDLDASIQMVETSLSHTSPGEF
SERVFGLEGRSFYQRKQGKLE
>Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)
MEAKTLKDKILTGLQKVASSLGSFIKRKWYRYQLTRWLIVIFLSLAVVSSAYLTFIAKTTD
VKGLKATLKQATQIYDVDGDKAGYLYAQKGTWVSLKQISPNVTNAVLSTEDRNFYKEY
GFSLTGIGRASLLYVKNKLTGQSGISGGGSTITQQLVKNAFLSQEQTFSRKAKELFLAV
QVENDYSKKEILTMYLNNAYFGNGVWGVQDAAKRYFNVNASELSVPQAATLAGMLKS
PSAYNPADHPQAARQRRNVVLGLMAANGKLSQSQLSQYQASAMVVSNGYDYVSGYK
YPYYFDAVIAEAIDDYGLSETEIMNKGYKIYTSLDQNYQTQMQTAFANSSLFPYNASDG
TKAQGASIAINPQTGGVAALVGGRAGSHVFRGYNRATQLKRSPGSAIKPLAVYTPALQ
NGYYYDSQVKDTLKSYGTNKYTPKNWNGVYSGELPLYEALALSKNTSAVWLLNKIGIQ
KGYASVKRFGISLSKSDNNLSLALGGMKEGVSPYEMAAAYTAFANQGVKYQPHLIRKI
VDSSGKVIVDNSKVDSKRIISKKIANEMTSMMIDVYKMGTGTSAKPYGFTIAGKTGSTE
SAQRIDGEDNDHWYIGYTPDVVVATWVGFDSSKYSLQNQGLRAGAGLFKTEMEGILP
YTPQTKFDVSAASTMVNSKTASSSSSEQVWSKLEDAGKDIGAKAASIKDKAGSYLKGL
FGY
>Arginine pathway regulatory protein ArgR, repressor of arg regulon
MKKVERQRKIEQLIKQNVIATQDELMRQLEANGIKATQATISRDIRELQIIKEQDSNGVV
RYVIFKGNNQNEEEKLIKAIEETVVDIRQVEFMNIIHTLPRNANVLAAILDGLSLPAVAGTL
AGYDTVVIISGSKDEASQINSLFTKHVAEQ
>oxidoreductase, YhhX family
MLKVAYIGNGKSTNRYHLPFALKSGRVEVKTIFARHNNSTWERLANVNYTNNLDDIYTD
NELDLVVVATPSAFHYQYAKECLEHGKNVLVEKPFAQTVAQAKELFELTQAKGLFIQAY
QNRRFDSDFLTVQAVIESGKLGQLLEVENNYDYFRPEVPQSQSELTWENSFLYTHAC
HTVDQMLAYFGKPNQVNYDVRQLLGRGRMNDYFDLDLNYDNLKVSVRSSYFRLKSR
PSFVLTGTKGCFVKVSEDRQEQDLKHFYLPTKAHPDFGLDRPEDYGTLTYVDAAGNY
HEEKVVSKVGDYSRLYEGVYQSLVNGAEKLVKDEETLAQIEILERGFKLLR
>Choloylglycine hydrolase (EC 3.5.1.24)
MCTAITFTGTSNYFGRNLDLNYSYGEQVVITPRNYEFNWRQLPSAKEHLALIGVGIVVA
NYPLYFDAINEAGLGMAGLNFPGNAYYGKVTVGKSNVSPFEFIPWLLGQAHNVGEAR
KLLADLNLVKINFSEQLPLADLHWLIADKDESIVVEATKTGLHIYDNPVGVLTNNPAFNY

QLENLKNYRQLSARTTPNTFASGLDLPVDATGFGSIGLPGDLSPKGRFVRASFAKLNA
LKGTDPLSDVNQFFQILATVKQVKGLNWADEHSCEYTVYSDCYDLQAGVLYYLTYESP
QLHAAKLQGQQLATSQLITYPLKNQVVVDWQN
>CAAX amino terminal protease family
MESLDLIKQKPPKSWWTPILVIVLTFVLAGGLGTGLLMWGFSLISGQAAFDEKTASGAL
AALFGFLFPLLGISLYNRLVNKRSLRALGFAREKWLLTYVKGFALGTVCLLAIVLVATLF
GGFKLSVTPQPDWLLLLGLLVGFMVQGLTEEVLCRAYLQNSMAAYKGQKWGIIVSALL
FTLLHASNNGLAFLPILNLFLFGLVFSLLYALTDNILLVGAAHSAWNFMQGPVMGVKVS
GNQLATSVLNAEATNRLNFVSGGAFGIEGSILTTIVGLILITWFFKTLERK
>Predicted tyrosine transporter, NhaC family
MKKFSLKESVVMLIILLVILAGGVQQGLSPETPVLTVIALLILVAKIHGATWEKIHNGIKDGI
STALIPIFIFILIGILIAVWIKAGIIPALMVVGFKLISIKFFVPSVFLVCALVGVSIGSAFTTISTI
GIALFGMGITMNINPALVAGAILSGAIFGDKTSPLSDSTNLASAIAGADLFAHIKNLMWTT
IPAFLVSLLVFILIGRGKTAPDFTKIETTVRVLESHFSISWLAALPVVLMFVCAWFKIPAIP
TLFINIFVSTGIIFLHQPHVALTELATLMEQGFVSQTHNKVVDALLTRGGIANMMGTVSLI
VTTLALGGLLMELGIIQTAMEPLIKRLKKPGNLVLATILSGIGVNLFVGEQYLSVILPGRAF
KDAFGRVKLAPLALGRVLEDGGSVINYLIPWGVAGSFAASTLGVPVLEFLPFVIFAWMS
PILSVLSGFTGIGLKMQK
>Cytoplasmic copper homeostasis protein cutC
MIKEFCAENYLDIPTAIKNGANRIELCDNLAVGGTTPSKGIIQASVAYCKARQIPVMTMIR
PRKGNFYYTNTELAIMKTDLLLAKELKTSGFVFGCLDANNCLDKQALLELINLADGLEIT
FHMAFDSIPQDKQFAALDWLADHHVTRILTHGGDAKQPLSQTLPHLKELIAYAKGRITIL
PGGQITYANATKIATSLGVNEVHGTKIIPL
>FIG00633510: hypothetical protein
MPHKANLPTSDRELLFLSWPIFVELLMGVIIGNINVWMISHFNELAVASISATNQLIGLSV
NIYGFITIGAQITIAQFLGANRNKELPTIINTALFGGLGIGLALSLLFFSFPKQLLAFMNLPP
NIISLGTDYARIYGGGLFIAALNAVMIATLRTHGYTRQALIIPMAASLFTVCGNFVALFSPF
GLPNLGVRGLACSALLGNLIAFGLASQLLKKYIHFDLLSARPRLISAKMLKQILKLGLPSS
GESVSYQGAQVVVTMIVASLGANVLITKSYVASITQFVFLTAAALSQGNQIIIGRNVGAK
NFTQAYQRGLRSTWQNVLISTSICLATFCFIEPIMHIFTHNPAIISLARWVFLVEILLEAAR
AVNMTLVGSLNASGDVKFPLICSLTVLWLISLPFSYLLAIGLKLGLVGVWLAYTIDESLRA
CLMIHRWRKGQWRQKVIID
>unknown
MLRKTKNFLKAHGVAYEKEHVNPLMVPEKNYVLKFEKNDKGDYQNRFIVEHTYTWTG
RMKINKITLRLHGQVHPREFKTESDLLRYLKKHAHRYAADVKKQHHKKKH
>tRNA-Glu-CTC

>FIG009886: phosphoesterase
MRYLVVSDNHGDREILVKLLDKYHANVDYFFHCGDSELRADDELFDVYQVVGGNCDY
NPAFSEKQLFEMGASGDKVLLIHGHTKGVSYGMTDLQLAAKEEGANLVFFGHTHQLG
CVYHDQTLFLNPGSISFPRGEFASFDGTYAVVESDEKQLIVQYYQRDFEPIPNLRFVYS
R
>Nucleoside 5-triphosphatase RdgB (dHAPTP, dITP, XTP-specific) (EC 3.6.1.15)
MNENTILIATNNPGKAREYAELFGAYGLEVKTLKDFTEIPAIVEDGKTFTENARKKAQTL
SDALGVAALADDSGLVVDALNGEPGIYSARYAGDHDDAANNRKLLANLKGLPAEKRTA

HFHCSIVVTRPGKEELVASGDANGLILEEAKGAGGFGYDPLFYYPPLKKTFAELSQSEK
NEVSHRGIATRALMKKFANWWEE
>hypothetical protein
MTNNFYKTSMAANQTNSTLAVTILRDALIPELTGEATGILYWVGKRLARQFALAKDEDL
PLFFEQTNWGTLKRIKGKANQQVFTLTGEMVATRLKLNDQADFKLECGFLAETVQNQS
GFIAEALVDSVNSKKGIVTIIVQVDTKDPLDFELYEQPETLELVTPKKEATE
>Amino acid permease
MLTQAKEAKKIAWWTLALMAFSTVWGFGNVVNGYIYFGGTKVIFSWVLMFLLYFVPYA
LMVGELGAVFKESEGGVSSWVDATMGPKIAYYAGWTYWACHVAYISSKGTGGLRALA
WAVFQSVDKFDAMNTMLVQGVTFILFVFFCWLASRGLSVLKLLATIAGSTMFIMSILYIL
LMLAAPVLNPHANLLSVDLSFKHLLPSFNVNYFTSLSILVFAVGGCEKISPYVNKMKHPS
KDFPKGMIALAIMVMVCAILGTIAMGMMFNPKDIDANFDSYVSNGAYWAFQKLGNYYH
LGNSLMVIYAWTNVIGQFSTLVLSIDAPLRMLLGSKEAKEFIPTKLLKQNKYGAYTNGIW
MVGILSGAIILVQSVVPGATKVLTQLVKLNATTMPLRYLWVFAAYIALRMHKKHFQGSY
QFVKNDKLAIFFGAWCFLVTALCCLLGIYSPDPFTLALNILTPVILTALGIILPIIKQHEKNS
>Thioredoxin
MVKALTDATFEEATANGVTLTDFWATWCGPCRMQSPVIEQLDQEMGDKVAFNKVDV
DENQATAAKFGIMSIPTLLIKKDGQVVDTLVGYHPKEAIEAKLNQYL
>Recombination inhibitory protein MutS2
MNAKGIQTLEYDKIKHRLATYVQTEMGQKLVAQLQPSSNIDEVNHWLAQTSDGADVFR
LKGGIPLATLVNIKPYLKRLEIEASLNAKELAAVGKVLRATSEVKSFLRQLASEELKLEQL
YRLGDQLETLPAVSKELLVAIEGDGHVTDEASPLLKSLRQQISVTESTIREKLADYTRGK
AAKYLSNAVVTIRNDRYVIPVKAEYRGQFGGVVHDQSSSGATLFVEPQVIVTLNNRLKQ
QQAAEKEEIRRILRELSQKLAPYTKELAQNAQLLGQLDFINAKARYAHDLKATLPLVSSA
NDVYLRQVWHPLLDQKTAVRNDIAIGKDYQAIVITGPNTGGKTITLKTLGLVQLMGQSG
LFIPAFEESRIGVFSNIFADIGDEQSIEQSLSTFSSHMTNIVEILANLDERSLVLFDELGAG
TDPQEGAALAISILDAVGAKGAYVVATTHYPELKAYGYERPETINASMEFDEKTLKPTY
RLLIGIPGRSNALDISLRLGLDPSIVTAAKQLTSQDSQDLNQMIADLVSKRHDAEENAIEF
KQKLRAAKKLHQELSSNFEAYLKQKEHLLETAKKEANEIVAQAQVKADEIISDLRQLRLA
GGQNIKENELMDAKGQLNALEQPLHLQKNKVLRRAKAKKDFHPNDDVLVKSYGQRGV
LLAKVGNHAWEVQLGILKMKIDEADLEKIPAAQAEPKQKNVVVRSSKTSHVSPNLDLR
GKRYEEAMTEVDRYIDAALLAGYGSVTIVHGKGTGALRQGITSYLQANSAVKRFEFAP
SNAGGNGATIVYFK
>Possible colicin V production protein
MILSLIIVILLILAFRYGLRRGLVLAMLNFVAYAMVLALSLLLAGGLGDFLATLLPSLNSES
TSFLVWSNVNLSQVFYRILAFWIIAISLGLVFRFGTGLIAGVFKLPVLSQLNSLLGGGVAC
LIMYGLIFFGLVLLATWPNQDINALVSNSTLAEFILEQTPLFSQELLSFLVR
>conserved domain protein
MSEQVRRFKVEIDGQTYTVIGKHSPEHMKAVVDTVEEQLNQIKEMMPSLSKEKAAILIA
LNAVSDQLEKQAQLNQLNDKLPK
>Ortholog of S. aureus MRSA252 (BX571856) SAR1694
MAADANNQEELITLVDDQGNETLYQVLFTFESEDYGKSYILLVPAGSEPEEQVDVLAFA
FNPDENGDAKDAELFDIESDEEWDMVEGVLDTFLNDENMR
>Putative Holliday junction resolvase YqgF

MGLDVGSKTVGVAVSDLLGWTAQGVEIIPIDEEAGVFGLDRVQELVAEHEVVGFVLGL
PKNMNNTLGPRAEAAQAYGQMLATKFNLPIDFQDERLTTVEAERMLVEQADASRKKR
KQVIDKLAASLILQNYLDAKGKLLSEI
>FIG01055109: hypothetical protein
MSSLDKTMYFDFGNSRQKDVKETLTTVYNALEEKGYNPINQIVGYLLSGDPAYIPRLND
ARNLIRKYERDEIIEELVRSYLDKGEK
>Alanyl-tRNA synthetase (EC 6.1.1.7)
MKQLTSAQIRQMYLDFFKSKGHDVMKSAPLVPKDDPTLLWINSGVATMKKYFDGSVV
PKNHRITSSQKSIRTNDIENVGHTARHHTLFEMLGNFSIGDYFKEEAITWAWELLTSPD
WFAIEPEKLYITVYPKDTEAKRVWHEVVGVASDHIYEAEDNFWDIGEGPSGPDSEIFYD
RGQEFSDLAEDDPENYPGGENERYLEIWNIVFSEFNHKPDGTYEPLPHKNIDTGMGLE
RVVSVFQHARTNFETDLFLPIIEKTAELSAGKKYGQDKEADVSFKVIADHARAVTFAIGD
GALPSNEGRGYVIRRLIRRAVMHGQKLGIKGSFLYQLVPVVGKIMEAAYPEVLQEADYI
EKVIKLEEERFQVTLKDGMNLLDDLIAKVKAEQGTVLPGKPAFKLYDTYGFPLELTEEY
AADAGLTVAKAEFDQEMAEQKERARNARSNAKSMGVQNALLTDIKTESKYVGYEQLT
ANGVLKDIIVADQLVDQVTTGKAQVIFDQTPFYAEMGGQVADQGVIEDKAGQVVAKVV
DVQHAPNGQNMHTLEVLQPMQVETSYRLVVDQTFHDKVKKNHTATHLLDQALRDVLG
EHTHQAGSLVEPTYLRFDFTNLGAVTAEDLAKVEAIVNDKIWANLKVETIVTDLESAKK
MGAIALFDDKYGEKVRVVKAGDYSMEFCGGNHVANTNEIGLFKIVSESGVGAGVRRIE
AVTSREAFEFLANRDRLLQETAAELKIVQVKDVPAKVKQVQANLKELEQQKQALEAKF
ASQQAQSVFEKVQTINGKTLIAGQVQVAGMNQLRQLADQWKAKQLSDVLVLATATAE
GKVNLITAVCPDQVKAGLKAGDLIKAIAPLVGGGGGGRPDMAQAGGKKPEGIQAALSQ
AAKWLETK
>DEAD-box ATP-dependent RNA helicase CshB (EC 3.6.4.13)
MKFTEFNFKPFINQALAEINFLEPTPVQAKLIPLIRKNKSVVGQSQTGSGKTHTFLLPIFN
AIDPDNKSVQAVITSPSRELAYQTYEAAKQIAKFTDGEILVHNYVGGTDKKRQVEKLQSI
QPQIVIGTPGRILDLVKSNALDIHNAHYLVVDEADMTLDLGFLKETDAIASSLPKDLQML
VFSATIPEKLKPFLNKYMTNPVIEVIKNQTVISPTIKNWLLSTKGRDKNQLIYQLLTLGEP
YLALVFTNTKSRADELTDYLRAQGLKVAKLHGGIQPRERKRVMREIQNLDYQYVVATD
LAARGIDIEGVSHVINDDLPEDLEFFVHRVGRTGRNGMEGIAITLYAPNEEEQIEELEEL
GITFSPKALKNGELVDSYDRNRREKRRTRKAMVAPELKGMISKAKKNRKPGYKNKIKR
AIKKNEQQRRKIERRQEMRALRRANKQKNSF
>FIG146085: 3'-to-5' oligoribonuclease A, Bacillus type
MMAIETDILAQIKKYDTIIIHRHQRPDPDAYGSQVGLAEIIRASFPKKKVYQVGEMVSRF
DWLVSQDEIEDSVYNDALVIVTDTANRPRVDDERYTKGKYLIKIDHHPNDDVYGDICWV
NTQASSCSEIIYDFYAACQAELTLPASAARVLYAGIVGDTGRFMYDATTPHTLRVAAAL
VETGIDAAQLNRKLDSISEPLARLSAYVLENMVVTKHKAAYVILTREIMAKYNLSDGGTA
PVVPLLGKLEEVVCWTVLVEQENHSYRLRIRSKGPVINELAKEYHGGGHPLASGAMLK
DASAIETYVEKLDQIAASAAERE
>DNA polymerase IV (EC 2.7.7.7)
MAGRKIIHLDMDSFYASVEMRANPDLRTKPLVIGRDPRHSGGKGVVATANYVARAYGI
HSAMPTAKALELCPQAVFKKPDFDKYRQVSSQIHQLFHSFTDKIEPVALDEAYLDVTEN
KRKLYDEVQLAHLLQQEIYEQIHLTSSVGISYNKFLAKLASDYNKPVGLTVIKPADVESF
LFPLPIEKFRGVGQKTVLKMHELGIYTGADLFAYSDLELVAHFGKIGHFFYQCVRGIDD
RPVEWQRKRKSVGRERTFDSPLTTTEEVATFLKQAANQLVVSLEAKQLRGKTLVLKVR

NSDYETITKRVSQAEFFKNDASEFAYQGQKLFEQIQQDEIDVRLLGLTMTNLAKSEETV
QLSLWD
>Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)
MKKNEKALVIIFGGTGDLAQRKLYPSIFQLYKKGYLKENFAVIGTARRPWTDEYYREVV
KDSIDGMYDEQTVATEFASHFYYQSHNVTDTNHYKTLQALAQSLDEKYQIGGNRLFYL
AMSPRFFGTICDHLKSQQIMTETGYNRVIVEKPFGRDLASATELNDEISKSFPEEDVFRI
DHYLGKEMVQNIAALRFGNSIFNSMWNNRYISNIQVTLSEALGVEERAGYYETAGALR
DMIQNHVLQIVALLTMSAPVSFTDTDIRNAKINALKCLKIYNPEEVKENFVRGQYGPGN
GLIGYREEAQTAEDSMTETFVAGKLMVNTMDMAGVPIYVRTGKRMKTKATRIDVVFKQ
VPNNVFEDSNLRHDVLTINVDPEPSLSLQLNSKTIGQGFDFDTHHLNLDYKLSPEQRAA
VPEAYEKLILDAMQGRASNFAHWKELQCSWKFVDAIRQAWNQNVDQVAFPNYDCAS
MGPKASDDLLAKDGNHWEFNC
>Mobile element protein
MIKLPKTTEYLRVRRYRLVATNDLVAKFERNIAVENKVYNYVLKYLEKTYGVKHLKRPY
PTNKKAKLFLAKDVLIPKILKDLYGLSKWNGKKVRIHSQALRDEYLVSILTNFGEYRKNLI
SASKMSKQNKKDYQNNLHGNNPKHKAWYRKGSLNYLRPNQSKKCVSLPSNGQAKIIS
AHFIKIQDYGTIQVVENINNMKDTKIVTTKIKRKANGQFELQIVFKTINERKEIVNMIGADW
NMKNNKIWHTSDNDELYLSKNVSDKADKLETEINKLKSQRDLITWLSRTSKRIVKLKEKI
RYLSSKRSHILTAEYNEMAKKLLFKNDLVAIENLDAKEMRIRNQSSLNQNKAKNRKLAKI
KPYEMGRLLIQMANRLGKTVILVDSYKTSQVEYGTEYQEKHGVDDREWVSKYTGAVIK
RDLNASKNILAWALNPKGHIKYKESLKLYKEGKIKTAIKPLDLITVN
>transposase-like protein A
MLFMHHPEIKEKLWGGHLWQPSYFVCTVSERSEDMIREYIRKQKKSKGKI
>Multi antimicrobial extrusion protein (Na(+)/drug antiporter), MATE family of MDR efflux pumps
MEKLFEKEAVSKVYFKLALPVVLSMIVSMVYSLADTYFVAKTGNANLVAGVTIGAPLFS
FMLAVGDIFGLGGSAIVSQLFGQKNYTLSKRLASASFYATIVISVGLTLILLALEKPVLLLL
GAKPATYQYAADFYRVLVAGSTFITVSLVPGNLIRTEGLAVQSMVATMAGTILAIILDPLF
LFVFNWGAAGVALANVLGYALNTSLLIYCVKTKSRYLSFNLKLAKLSWADWKRVIEIGIP
ASITNFMQTFGMALLNNYLAPFGAASVAAMGITQKIYGIVILVLVGFAFGAQPLVGYCYG
AKNVQRFKEVLRFDIMVEVVYALVFSVILMILAPSLVSCFMNQAEIIKEGTYMLRACLAT
TPFIGLIMVLTTVFQSASKAMSALLMSVSRQGVFYLLVMLAFSSLFGYHGIIWAQAVTDI
LTAIMGIYLYRRSFKDFN
>4-hydroxybenzoyl-CoA thioesterase family active site
MTPYQHTVQYYETDQMQISHHSNYIRWMEEARVAFLACIGFPYEKIEASGLISPVVSVK
CDYKAPTTFADVISITVTVKQFSGVRLVLGYRMTLPDQTVVCLAESSHCFTTKAGNIVN
LKKTNPDFYQALVANLKAD
>Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9)
MGDLLTGKKILVMGAANKRSIAWGCAQMMMEQGAQLIFTYQNERLKKSLHKLVADEN
ALIECDVASDESLQTAFATIKDRFGKVDGIVHAIAFAPKDALGGDIMNASRQDFAIAQDIS
AYSLLAVAKYGQEILNNPASIVTLTYFGSERAVPNYNVMGVAKASLEAAVRYLARDMAK
YGVRVNAISAGAVKTLAVTGIKGHSELLEMSEERTVDGQSVTTREIGGTCAFLMSDLAT
GVVGDTIYVDKGVHLV
>Acetyl-coenzyme A carboxyl transferase alpha chain (EC 6.4.1.2)
MVSLFERKQKVAKSAAEVVKAARSEDKITAPELIKGIFDDFFELHGDRTGADDPAIIGGI
GYLGNQAVTVIAIDKGKTPEERIAKHFGGPTPNGYRKALRLMKQAEKFKRPVITLINTAG

AYPGRTAEEGGQGQAIAQNLLAMSDLKVPIIAVIFGEGGSGGALALALADKVYMLENS
MYTVLSPEGFASILWKDSRRADEAAEVMQVTPKSLLKQGVIEGIIEEPSSHKKVIKNVA
RVLTTQLEPLLALTPRELIARRQARFRKF
>Acetyl-coenzyme A carboxyl transferase beta chain (EC 6.4.1.2)
MQLFDELKTLGHKHVKANKAAQQKVPTGLWISCPKCHQSFYHKDLGRYQACPSCQY
GFRITARERLEWLVDDFTELDDQLQTKDPLDFPGYQAKLAKAQAKTKLNDSVLTGLAKI
GQTSFALGIMDPQFIMGSLGTITGEKITRLFEEATRKNLAVVLFTASGGARMQEGIFSLM
QMAKVSQAVKNHANAGLFYLTVLTDPTTGGVTASFAMQGDIILAEPHALVGFAGKRVIE
QTIHQRVPEDLQAAETVLKEGFVDAIVPRAKQKETIAFLLATNQKGGLTNG
>Biotin carboxylase of acetyl-CoA carboxylase (EC 6.3.4.14)
MFSKVLVANRGEIAVRIIRSLKELGIKSVAIYSTADRESLHVQLADEAVCVGTARPQDSY
LNMQNILAAAVGTGAEAIHPGFGFLSENSRFVEMCEACGLTFIGPKSQTIELMGNKANA
RMQMQASNVPVIPGSDGFISDVTEAKAVAKKVGYPVLLKAAAGGGGKGIRRVNDEDE
LTTAFEEARREAQNGFGDNRMYLEKIMENVKHIEVQIFRDQAGHTVYFPERDCSVQRS
KQKMIEESPCSVITPAQRKELGQIAVRAANAIDYLNTGTIEFLMDSEHHFYFMEMNTRIQ
VEHTVTEMVTGIDLVKAQVKVASGEDLPFKQADIKANGYAIECRINAEDPAKNFMPSVG
KVKYLYFPVGNLGMRIDSDLYSGLVIPPYYDSMIAKVIALGHDRHEAIEKVKRLLNEMVIL
GITTNQAIHLDILQAPTFLKDQVTTDYLEREFLPKWKEGLDSATV
>3-hydroxyacyl-[acyl-carrier-protein] dehydratase, FabZ form (EC 4.2.1.59)
MVLDVTEIQKIIPHRYPMLLIDRVEELEAGKRAVAIRNVTAHEQVFNGHFPGNPVLPGVL
IVEAMAQTGAVALLSMDEFKGKTAYFGGIKEAKFRKVVRPGDTLRIEVTLEKIRNNVGL
GKAIATVAGKKACTAELTFMIG
>Biotin carboxyl carrier protein of acetyl-CoA carboxylase
MDFKEIKELMAEFNESPMRELEIETDGFHIHLSKNENTFKPQGVLPQAPQAETPSPKVE
EASEQAEAASYIKAPMVGSIYLQPAPDKEAYVSVGTKVHKGDVVCIIEAMKMMTEIKSE
VDGIITKVLVTNEELVEFDQPLFEVQEG
>3-oxoacyl-[acyl-carrier-protein] synthase, KASII (EC 2.3.1.179)
MTRVVITGMGTVSPIGNDTASFVENLMAGKLGITEITKFDASATGISVAGEVKNFDPKLR
LGKKAAKRMDLYSQYAVHTALEALEQAGINEENTDPYDMGVIIGSGIGGLTTIQEQVIK
MHEKGPKRVSPMFVPTAIANMAAGNVAIFANAKNICTAVVTACASATNAIGEAYRQIKE
GRAQVMLTGGAEASVNEIGIAGFAALTALSTQTDPKLASLPFDKDRSGFVLGEGAATLV
LESEEHAKKRGAKILGEVVGYGATCDAYHITSPDPSGEGAARAMQAAIAEAGIKPDQV
GYINAHGTATHANDEGEAKAINLVFGKDSKVLVSSTKSMTGHLLGAAGAIEAVATVAAL
QMGKLPANIGLANQDEKCQVNLVTQNAQAAEIEYAISNSLGFGGHNAVVAFKKWSE
>3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)
MDLKEKVVFISGSSRGIGQAIAQAFAKAGSNVILNARTAVKPELVAQIESYGVKAAVVLG
DISQASDVKRMVKEAYAAFGQIDILVNNAGITNDKLLIGMKEADFASVIDVNLKGTFLMS
QAFLKKMYKQKAGVIINLASVIGLHGNVGQANYAASKAGVIGLTKSLAKEGALRNIRAN
AIAPGMIVSDMTAVLSDKVKADIMTEIPLKRFGQADEVAQTAVFLAQNDYITGQVITVDG
GMTI
>Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)
MEKLVLLFSGQGAQYSEMGLDLYQTDSLFKEKIDQASQASGLDLVEILANKAGQLEKT
AYVQPALVAVSLGIYAMLARDFDLDIRGMVGLSLGEYGALMASGALDFEAGMALLKDR
GAYMQADADANESMLAAVLGPDVAQVEAVCQAISTAANPVGIANYNSPKQVVIGGSP
QAVLKASEELTTAGAAKKVVPLKVSGAFHTPLFKNSSHKLAARLEAVKFKEPAVTVVSN

TLGQPFTKANLAETLSKQVCLPTHFADCVKYLLATKEVSGVLEIGPGETLSKFVKQIDR
KLARYHIEDLATYQAFGEKLRGN
>Acyl carrier protein
MTEAEIFDKVKEIVVDQLDVDEADVTLAADIKEDLDADSLDVFEIMNELEDDLDIKLEAD
ENVKTIQDVVNYVKKEVDAK
>3-oxoacyl-[acyl-carrier-protein] synthase, KASIII (EC 2.3.1.180)
MEKVKIIQTAHYVPDLVVSNDDLSQIMTTSDEWISKRTGIKRRHVSLGENTSDLCTKVA
SDLLAQSGWQASQLDLIIVATMSPDAYTPATAAIVQGNIEATNAFAFDLAAACSGFAYA
FEVASQFITTGRAKKVMVLGGEVLSKVVDWQDRTTAVLFGDGAAGVLLTESDDQAAIL
ASDLQTFGSEAAKLSAGLNQPLTSFGDFKPKQISAFAMDGHAVYNFATREVPASLKRA
CHQAGLELAEVDYFILHQANERIIKAIAKRLKLPLSKFAVNISEYGNTSAASEPILLDELVK
QNKVKRGQIIALSGFGGGLTVGTQIIRY
>Transcriptional regulator of fatty acid biosynthesis FabT
MNERDKAINDSLVRIYSGIMWIEENELRKSKFNDLTMREMHAIHAISMYDHKSASEVAK
ELHLTPGTLTAMVDRLVRKGYVKRLRSEDDRRVIRLGLTKKGRVIYRAHEAFHNMMVR
SFLKDLDATEIKTIEKALWNLEQFLREHS
>3-hydroxyacyl-[acyl-carrier-protein] dehydratase, FabZ form (EC 4.2.1.59)
MRVMDAQEIMDLIPNRYPICYIDYVDEIKPGESIVATKNVTINESYFEGHFPGNPVMPGV
LIIETLAQAASILILKSPEFLGKTAYLGSVKNAKFRQMVRPGDVLKLEIKMVKKRANMGIV
ETAALVDGKKACTAELMFIVEERTEKI
>cell surface hydrolase (putative)
MTERWRKFLVMGASLFILLAGGFLVVRQGIWHANAKKYRQSQVPTLYLHGYGSSYKA
EESLVKAALNSGASSGVIRANVSKSGKVSLKGKFTANVNNPLVEVNFADNKNGDYHQ
DGAYLKAVVEKLQARYQIKQFNVVAHSMGNMAVAYYLLDNGQNKKLPRLNKQVSLAG
HYNGILNYDAPKDASLGKDGAPKEQTNSYQALLPLKKRYPKQVAVLNIYGDLNDGSHS
DGRVANASSKSLAYLVADRAASYQELVIKGKGGQHSRLHENQRVAQQIVNFLFLKS
>Galactoside O-acetyltransferase (EC 2.3.1.18)
MEQDLIKRVQAITNGGELYYDYNEEIDLARDYSFKICRQYNDLVQHQNEYHYELLAGLF
AHLGKNVRIEANFLTEFGFNLSVGDNVTILQDVSLIDCTFVTIGNNVVIGPHCGIYTANHA
EDPKLRANHYCYERPIVIEDDVVIGGGSVIAPGVKIGTGSIVQAGSVVVKDLPANSFCLG
DPCQAIRKLGVD
>Preprotein translocase subunit YajC (TC 3.A.5.1.1)
MVLGAGMAGSYSTILMFVVFIALMYFMMIRPQKKQQEKRQEMMNGLKKGDEVVTIGG
LHGVIDEVNQADKTVTLDCDGVYLVFNLAAVGRVVPAKSTAAKEETVVATDAEKTDEP
KEDK
>tRNA-guanine transglycosylase (EC 2.4.2.29)
MTTPAVTYNLIKKDKHTGARLGQITTPHGTFDTPMFMPVGTQASVKAMAPEELKEMGA
GVILANTYHLWLRPGEDLVEEAGGLHQFMNWDQGILTDSGGFQVFSLAELRNITEEGV
HFKNHLNGQKLFLSPEKAIAIENALGADIMMSFDECPPFFESYDYIKKSIERTSRWAER
GLRAHKNPQTQALFGIVQGGGHKALREQSARDLVSLDFPGYSIGGLSVGESKAEMNH
VLDFTTPLLPENKPRYLMGVGSPDALIDGVIRGVDMFDCVLPTRIARNGTCMTSKGRLV
VKNAKYARDFRPLDEKCNCYACRNYTRAYIRHLLKTDEIFGLRLTSYHNLYFLLDLMRKI
RAAIKEDRLLDFRAEFFEEYGLNKENRKNF
>S-adenosylmethionine:tRNA ribosyltransferase-isomerase (EC 5.-.-.-)
MTKQAHLTTEDFDYDLPQELIAQTPLEKRDHSRMLVLDSKTGEMKDDVFYNVIDQLNP
GDALVMNDSRVMPARIYGVKPDTGGHVEVLLLNNTKGDEWETLVKPAKRAKVGTKLS

FGDGKLMATVTKELEHGGRLIEFEYDGIFLEILDQLGEMPLPPYIKEKLDDPEMYQTVY
SREVGSAAAPTAGLHFTKELLAKIEAKGVKLVYLTLHVGLGTFRPVSVSNVEDHKMHS
EFYRLTEEAAQTLNEVRANGGRIVAVGTTSIRTLETIGSKFNGEIKADSGWTDIFIWPGY
EWKVVQAFITNFHLPKSTLVMLVAAFTGRETILNAYKHAVAEKYRFFSFGDAMFIK
>Holliday junction DNA helicase RuvB
MEHDERLISSDELAADFTDDLTLRPKFLDQYIGQSRLKQEIAVYIQAAKSRQESLDHVLL
YGPPGLGKTTLATIIANEMGVNIKTTTGPAIERPGDLVALLNDLDVGDVLFIDEIHRLPKN
VEEILYSAMEDYFIDIIVGQDATAHPVHFPLPPFTLIGATTKAGMLSAPLRARFGIVGHM
NYYSEAELSQIVLRSAEVFTTNINEDGAHEIARRSRGTPRIANRLLKRVRDFAQVAAKA
EIDKEIVDYALNLLRVDSKGLDETDRKLLTTMIKLYHGGPVGLNTIAANIGEETDTIADMV
EPYLLQIGFLKRTPRGRVVTKAAYDHLGLEMTND
>Holliday junction DNA helicase RuvA
MYEYINGKITDVSPYHIVVENNGLGYLVYVANPFSYQVESQQKVYLYQAVRDNDISLYG
FKNKVEKQLFLKLLNVSGIGPKSALAILASDDLSGLVTAINNDDDGYLVKFPGIGKKTAK
QIILDLKGKLADLGNLDQLQGQQGLALDLAAENQYVTESLEALLALGYTKTELKKLKPKL
LDYQASSTDEYLRYGLKLLMRK
>DNA mismatch repair protein MutL
MGRIHELSDILADQIAAGEVVERPASVVKELVENALDAHSSQIDIYVEEAGLKSIQVIDNG
EGISQADVLIAFRRHATSKITDRADLFKVKTLGFRGEALPSIASVSHVSLETSTQAENAG
SLVEIKGGELIKQQPSSARTGTKITVTDLFFNTPARLKYLSSRATELANISDIVNRLALSH
TDVAFSLVSDGKTIFKTAGNGNLKQAVSAIYGLTSARQMIEFSGADADFAVSGLISLPKL
TRAARSYISLLINGRYIKNFRISKAVVAGYGSKLMVGRYPLAVVNIKLDPLLVDVNVHPT
KQEVRISKEEQLEFLLTKTVSQALSQENLIPNGLENLVGKKAKQQLDFKQLEIGLNETK
GAYQFKNQPPALPKKPALNKEVTQALLGQGQEKSSPKTVVKPAALIITDKAQLKQAAVK
NWDERYQQPTTPPKDLQLEEAEVAPTEAVAAVDEEASPAETVRFPNLVYIGQFHGTYL
LAQSDDGLYIVDQHAAQERRNYEYYRQAIVKVGQAQQELLVPIVLTYPTSESLKLNQHL
DDLQALGVKLEDFGQNTYIIRSHPAWFKAGQEEATIREMIDYLLTDQKLTLAQFREKTAI
MMSCKRAIKANQHLDDRQAKELLAKLATCENPFNCPHGRPTVVKLSPKDLEHMFKRIQ
DSHQSKRKEW
>DNA mismatch repair protein MutS
MAKKKVTPMMEQYLAVKKQYPDAFLFYRLGDFYEMFYDDAIKGSQLLELTLTQRNKNS
LEPIPMCGVPHHAAKNYIDILVDKGYKVAICEQMEDPKQAKGMVKREVVQLVTPGTIMD
EGASQAKSNNYLTALHEQAGQLGFAYVDLSTGELKSTVVNNLASLQNEILSLHTKEIVV
DQSLSAQTKELLAKLKILSSHQDDYQAQAEFSFISQDLHEASEIEAVMHLLMYLKVTQK
RALAHLQKAVHYEPANYLKLDHYAKRNLELLVNSRTNKKSGTLLWLLDATKTAMGGRL
LKQWLERPLLNQAEIEARLELVAILKDNFFERSSLQESLSKVYDLERLAGRVAFGNVNG
RDLLQLKTSLMEVPQIRYVLEQLAGKQFADLLAKLDPLEEVVDLIDRAIVDEPPISVTDG
NLIKAGYDPKLDEYQDAMKNGKLWLAQLEASERELTGIHNLKIGFNRVFGYYIEVTKAN
LAKLPKDRYQRKQTLTNAERFITPELKEKEDLILGAESKAYELEYQLFVAVRDQIKDQIK
RLQSLAQALAKLDVLQSFAVVSEDAHLVRPTFTSDKSLVIKDGFHPVVEKVVGRQSYV
PNDIIMSPTKQILLVTGPNMSGKSTYMRQLALTVIMAQIGCFVPASQAALPIFDQIFTRIG
AADDLISGESTFMVEMKEANEALQNASENSLLLFDEIGRGTATYDGMALAQAIIEYVHD
KVKAKTVFSTHYHELTDLEQTLAHLENIHVGAVEKDGQLVFLHKVEAGPADKSYGIHVA
KLAGMPSELLTRADKILTKLEAHKQDQVSPVAKSEELIKEEGQLSLFAEETPVEPKAKP
VSKQADKVLKELGQLNLMALTPMEVMNAVYKWQQELE
>FIG006542: Phosphoesterase

MRILFVGDVMGKPGQEMLTSYLPLLKRKYQPQVTIVNGENATRGRGINEKVYKEILKAG
ADVLTMGNHTWDNPELQTFIGRAPKLVRPANFSANKVPGQGYTIIKVNQLKLGVVNLV
GRVFMNPSDDPFAIGEALIAKLAKQVDAIFVDFHAETTSEKEAFAWYFAGKVSAVVGTH
THVQTNDARILPGGTAYLTDVGFTGPYDGILGMKRANVIERFLTQMPTRFEVDTENQT
TLGACLIEVDDASGKAKSIQTIQINPDHPWLD
>6-phospho-beta-glucosidase (EC 3.2.1.86)
MRNKHQFLYGVSSAANQIEGAWQEAGKGVSVTDALATIAGHRVETEDYLNEDLFFGS
HQAVDFYHHYKEDIELLAKMGINSYRMSIAWTRIYPTGLEETPNEAGLAFYDRVFDELL
AHGIEPIVTLSHYESPLALAKEFGGWSNRQMIDLFVKYAETVLRRYRNKVSRWITFNEI
NCLQVPFGIMTAGGIYMGIKDEANTERLRYQALHHQLVASAKVVELAHKVNPANQVGC
MIAAMFNYPLTPKPEDVRANQKTMQVQNYFASDVMIRGAYPGYMKRYFAEKEIKLEVA
PEDVATLQAGTVDFYACSYYMTNCVAKGDAGQKTAANLVAGLVNPYLEASEYGWQID
ALGLRTYLNEVYDRYQLPIMIVENGLGAKDELVAGKVEDAYRIDYLEKHLAALEEARKD
GVDLIGYCAWSSTDLIALSTGNIEKRYGFIYVDVNNLGQGSYKRIPKQSYYWYRNYIKQ
HLNED
>PTS system, beta-glucoside-specific IIB component (EC 2.7.1.69) / PTS system, beta-glucoside-specific IIC component / PTS system, beta-glucoside-specific IIA component
MDYEQLAKEVIKGVGGSKNVSHLTHCYTRLRFKLVNVDKADKAKIEQLDGVMSVVYSG
GEYQVVIGPEVGNVYQQIVANKLVANVETDGASEADSAPVKKNFKYWFDLCLDVLVSC
FTPIIPAIAGSGMIKVLTAILVSFHWLSDKSQTYVVMNAIGDGIFYFLPFFVAITAAKRLKA
DQFISMTIAAIIMYPKLEALVAKGTIHFIGLPMQLVDYSSQALPLIFAVILVKYAGKLAEKIS
PTVLKVWLVPMITLLISAPLTLFIFGPFATYLGSLFAIFTNIMNHWGWIAVGLNAALFPILV
MTGTHNALIPLMIQMFATQGFDAVLVPSGLAANIAESGAAAAVATKTKDMKMKSTASG
AAISALFGITEPALYGVNLKLKRPFIGMILGSLIGGGIAGLIHLTAYAFVSPSLLSLPIFIGK
NSNLVTAIIAAVVTFVVTFAITWILGFKDVSDNQKAKTEGNTTATVNHKAVAVNMTVPVS
GQIISLSELGDQVFAKGLLGKGFAVKPSGDTIVAPVSGTVTTIFPTKHAIGLKTDDGLEV
MVHMGIDTVELKGKYFEDLVKAGDQVTSGQALAKMDIDQIAKAGYQTSVIIVVTNTAQF
ANVTVTPSADKREVGFTIEK
>Transcriptional regulator, AraC family
MELEDGRIGRIELITSDKHLYLDRGLTIFYVVSGQVTFSYLGQAKSVKKGELFVQQSALE
VEIMLAKAQLVMISLQADFTDIFNLNKLIISCASQDSKRVEYQHLVNLVQALIDYVSVDSA
NYQANIAYGKVIEIIQYLSANFSKQLLSKQLQIRDVIEYLRLNYNQALDLVEVAKHFGVSP
QYLSKFFKAKHGNNFNQELTDIRLYNAARQLVNEDKGIIDVAYDSGFNSIATFNRAFKK
HYQMTPANYRLSHQNKVSVVPVEQTKHSLALLAAANKREVLTVNANDKQEPLVLPRQ
VYLDSTPTKQTLTLLAKLGIERVVVTIPPITNLDAFLYSFNQQIDPLLRAGYALVLAFATPS
LSYLQEILRYFINLIGIENMAKWQVQIDLQAYQEIPKLNKILGEFKLDQGLVIAGRYRQLK
ALSAKLVPLKAMIKLYSVSGTKADTLAKVKSLKALGYQKLGVVLHGFFEANLKVLNDTE
YLGVDLVAQVLALQGQVDFVETPKLYDDLQPKLVAGQMGLATSNGLLKAAYYALRLLN
KVDPNLVYADKKTVVTKDSVGNYNLTLLNQQPILNQQVLIDYQNYPEILATQTRKVTISL
KGIANGRYRLKKRIVSRHTGNLIKVWGQLGFIQQHLTVSEGQYLAHKTLPDIFINELVVE
DNLAVFELELAANELVSLHLIKLY
>hypothetical protein
MDIENMIMMERAYARQEGWDEGRVEGRTEGRDEAREEAAKALIEDARQNHKNLADLL
ATLKVIFKDQAYIDKMLAKYYY
>hypothetical protein

MLARSVLAKRWQQATIKDNIIFNWIFGENPTLFKGLVERLFETRVGNFEALSSERSFKD
HRIFYRPRFDTYGEDDLDNIINVELQNENHHDLEKRIAIYQASLTQRALLSGQSFNKRKQ
TLIAFLCDFDLYGRNRPVYKLKTTVEGQEDILVDRGETNVIFNLKATDLAGLPADYVALV
TYFNTGRVSDEFTAAVDQEFRRIKGNDTRKEYYMSIEAILTMEKAYARQDGWDEGHVE
GRAEGRTEGRDEAREETAKALIEDARQNHKDKTDLLATLKVIFKDQAYIDKMLAKYYYE
>hypothetical protein
MQTVQAAFRQADTEELIDAYLAEYPLRMDDFADEVTIGEAKGFTRIKLRDYLERLRTLP
VKAATEGEHLFYAYHRPKEGIAEPTFALVLLTELRQKGIRAADYAFEFTDQSEIMCYYIA
DNRLTKHYLTELLVYIMYEASFFGFEQEDLQENLVDLKIRVQEAKDWATAKQAETKHDL
DWLEKLVKEAGIIGFSPEEKAYLDKVINYEGKIINHSQQTEVGQILNKLG
>hypothetical protein
MDLAEWILMLRNEARREGLEQARKERREGKVVTVDEEEKQKLFEELMQSD
>hypothetical protein
MEYTEIIEMINQFRDARGWKKYHNLKDLAISVSLEASEVLEIFQWKDSKEELSEDEKAKL
KEEIADTLIYLLYMVDKLKMNPYEEIQKKVEINKNRHWKEDS
>FIG00748724: hypothetical protein
MKKLSKPIIYKVNYPREDAQNLDDKIRKNGNIKGEEQKYLIDYPTVYVIASEKSKRYTVY
VGETNNIKRRTLEHIDVDSEVREDWQKLKEFQDAEMYIVGHEHFNKSLTLDIENRLMQ
YMSSVESVEKINNRRENSQNKYYTSDEMKKIFDKIWKQLHRENPRLFPAKKLIEDSALF
KASPFNKLTKNQLYAKDEILSSIDKALSKEKSNQLIKVIGNAGAGKTVLMSNIFYELAKKS
NIKTVLMVNHLQQEKIYRQIAKKLDLTDNAKVLKVTTFLNQTDVDFPVDVAFVDEAHLLL
TQRSQAYYGHGTNELIDIIKRAKVTLAVYDEKQILQTTQIIEDDDRLELEKFDTEYIFLND
QMRIKASKETVEWIRKFIDDGEINPVPNDESYDLRLFDTPQEMQEAIFEKNTKDRENGL
SRMVATFDWDYSSTKKERGYWEVSEGEWHMPWNLQLKPSKSQKNASSISYGELSW
PEQPHTIDEIGSHYTVQGMDLNYVGVEIGPSVKYRDGRVVFDPTASSNKKAINNRTMH
SGEKKKFGKKLLRNELNVLLTRGVHGLYIHAVDDELQQALKRVFKKGE
>Riboflavin kinase (EC 2.7.1.26) / FMN adenylyltransferase (EC 2.7.7.2)
MEIIKLSEPYQSSQLPKEPSVLALGFFDGVHLGHQKVIKTAKKIARERGLKLAVMTFDR
QPKLMYQPHPKPALYLTLLARKLELFEQLGADIAYVVTFDEHLASMGPQEFVDKYMVG
LSAQVVVAGQDYTYGKKELANMETLPSFAKGRFDIVAVDHLVLAASTSKVGSTAIRAYL
DQGQVDQANELLGYPFQTKGLVVHGFARGRTIGFPTINLEIDPSQHLPAPGVYATQVE
LLGQTYLGMASVGYNETFGDNFDLTVEIYLLDFAKDVYGQEATVAWYHWLRPMVKFA
SVAELIDQLKADEAAVRQYFAK
>tRNA pseudouridine synthase B (EC 4.2.1.70)
MDGILALYKERGMTSNDAVFKCRKLFKTRRVGHSGTLDPNVDGVLPICVGKATKVVNY
LMDSGKVYMGQIILGFATETEDLDGAVVEQQRLVRPFSDQEIDQAMQALCGDLIQIPP
MYSAIKVNGRRLYDYARKGEEVVRPKRQIHVDYFKQTKPSSFDENRGQQTIYFEVGC
GKGTYVRTLAVDLGKKLGVPAVMSDLTRLKSGGFMIGQAVSLAQLESLSADQLAQALA
PIDHALGAYAHHELSEKQWAFVKNGGFLAPSYVNGGQVVPKLVLMYQGKARAVYKYN
EPKARYQPEQMIDLS
>Ribosome-binding factor A
MAQNFRVGRLAQEIQREVNDILKKRVRDPRVSGVTITGVDVTGDLQHATIYYSILSDLA
SDAKKTQQGLDKASGLIRGELGKRLTIYVTPELKFEQDKSVRYGDHIDELLNKLRQEDR
F
>Translation initiation factor 2

MRKKHIYEIAKELNVASKELVKLAQEKGFPTKSHMSTVGENEERQLRAAFSKQAKPTT
KTQSQPEVKQAHTKKPHSQNVDQAKPAAKTNERHSGKNMDKKKQTNKGGNASNNN
KKQRNNNNNNRFNGKKNNRNNKKNNRNNRNKNTKPAVPARKNKPLPEVLVYTVGMN
VADIAKKIHREPAEIIKKLFMMGVMVNQNQSLDKDTIELLAADYGIEAEEKVEVDVADIDK
FFEEEAQSTDNLQPRPPVVTIMGHVDHGKTTLLDKLRHSHITAGEAGGITQHIGAYQVR
HNDKIITFLDTPGHAAFTDMRARGANITDITVLVVAADDGVMPQTVEAINHAKAAGVPIIV
AVNKIDKPGANPNHVMEQLSEYELIPEAWGGDTIFVEISAKFNKNLDELLDMILLEAEILE
LKANPDQRAAGSVIEARLDKGKGSVATLLVQHGTLHVGDPIVVGNTFGRVRTMLDARG
HDIKKATPATPVEITGLNDVPVSGDRFLVFEDEKTARAAGEERAKRALMKERSQKNHV
TLDNLFDTLKQGDMKEVAVIIKADVQGSVEAIAQSFNKIEVEGVRVNIIHQAVGAINESD
VTLAEASNAIIVGFNVRPTPQAKVQAEADKVDIRLHNVIYKAIDEIETAMKGMLEPVYEE
KVTGQVEIRETYKVSKLGTIGGGYVTDGYIQRDSGVRLIRDGIVIYEGKLASLKRFKDDV
KQVKQGFECGLMIEKYNDIKVGDQIEAYIMEEVPVD
>ribosomal protein L7Ae family protein
MENRQKVLNLLGLAQRANKLVTGEQLVLKQVRAKKAYLVFIASDGGQSTHKKISDKCY
SYGVALSSDFSQLELSIAIGQKRSLIAVTDLGFSKKMRQLLA
>COG2740: Predicted nucleic-acid-binding protein implicated in transcription termination
MAQRKVPMRKDIVTGEMKPKKELVRIVKNKENEVSIDPTGKKAGRGAYIFLDVAIAKKA
QAERTFDKAFGIKIKPEFYDELVAYVDHQVARQELFGNGK
>Transcription termination protein NusA
MSKELLNALDALEVEKGVKKEVVIEALEQALVSAYKRNYGQAHNVEVEFDRTRGNIKV
YAVKEVVNEVFDSRLEVSLKDALDINKAYEIGDTIRFEVTPKDFGRIAAQTAKQVVMQR
VREEERTNIYNQYIEYEKEIVTGEVERHDSRYIYVNLGKVEAVLSHQDQIPGEKYSPHD
KIKVYIYKVENSAKGPQVYVSRSHPDLLRRLFEQEIPEIFDGTVEIVSIAREAGDRAKVAV
RSHSDNVDPVGTCVGPKGERVQAIVNELKGENMDIVEWREDNAEFIANALNPAEVVA
VHFDPENDRACTVVVPDNQLSLAIGKRGQNARLAARLTGFKIDIKSESDYEELQAQLAQ
APAKEASDSEEPVDAVDPELAAPASDLDEE
>FIG000325: clustered with transcription termination protein NusA
MSSVVETVTDLVEPILASHHFELVDVEFVKEGKSWYLRVFIDKPGGITIDECVLISDELG
EQLDNCDPDPIPQAYYLEVSSPGAERPLKKEADYMRAVDKYVNVSLYQQLNGHKVYE
GYLRKVTADELTLEYMDKTRQKEVIIPRKQVAKARLAIKF
>DNA polymerase III alpha subunit (EC 2.7.7.7)
MALTKAELFEKLLEQIKWQPNESEQSYFEGAVIDKVVVHTKSRRWDFYIGLEAVLPYAT
FIDLKNHIELAFKEIASVSLAFRTTSEQLEAGKLGDYWQWVAQNSGLNSPLQVSLAKNK
PPYVDGKRVVLLAENELLQNFLKNEALGKIEATYQNLGFPKFNLNVLLDETKSQAKIEEF
KQKQALSDAELTKRAVEAIQKQNAKKKQTKQEPAPVTGDGPLLMGKAISSDQPIRPMV
EITEEERSVVIEGYVFDKEVREFKSERQLLTLKVTDYTSSFIVKKFSRTDEDKAFFAAFK
TGMWVRVRGSVQEDSFIHDLTVNAYDLQEVKHAKRVDSAPADEKRAELHLHSNMSM
MDATNSISDYVAQAAKWGQKAIAITDHGTLQAFPEAHSAGQKHGVKILYGVEANLVED
GTPIAYNEAKVDLKSATYVIFDTETTGLSANYDKVIELAAVKMKDGEVIDRFEEFIDPGH
PLSKTTIELTSITDEMVKGSKTEEEVFKLFQTFCEGCVIVGHNATFDVDFMNTGYSRHH
LPEISQPWIDTLPLGRFLYPEMKRFTLDTLARKLNVQLEHHHRAIYDAEATGYIYYAMLK
DAEANYDIKFHDQLNDHMGEHDAYKNEHPFHATIIAKTQAGLKNLFKLASEGMTQYFY
RVPRIPRSVLAKYREGLIVGSACADGEVFTAMMQKGYAIARQKAQAYDYLEIQPKALYE
PLLEKELVKDNAQLEEIIQNMVKLGHELNIPVVASGDVHFLNPEDAIYRKILIHSQGGAN
RLNRLKKLPDAHFRTTDEMLADFAFLGEELAKEVVVTAPNEVVSWCEEISPVKDKLYTP

NMPGANEQIRDLTINKAKELYGDPLPDLVQARVDKELKSIIGNGFAVIYLISQKLVYKSN
KDGYLVGSRGSVGSSLVATMTGITEVNPLPPHYRCPNCQYSEFYTKGEYGSGYDLPD
KNCPNCQTLLVKDGQDIPFETFLGFKGNKVPDIDLNFSGDYQPIAHNYTKVLFGENNVY
RAGTIGTVADKTAYGYVKAYERDTEQVFNTAEVDRLAKGATGVKRTTGQHPAGILVVP
DTMEIYDFTPIQYPADDLTAAWRTTHFDFHSIHDNILKLDILGHDDPTMIRMLQDLSGID
PKTIPTDDPGVMALFSGTESLGVTPEQIQSKIGTLGVPEFGTRFVRGMLEETHPATFAE
LLKISGLSHGTDVWLGNAEELIKSGTITMAKVIGCRDDIMMDLIHYGVESDVAFKIMESV
RKGKGIPDDWQAQMREANVEEWYIGACLKIKYMFPKAHAAAYVLMALRIAYFKVYFPL
VYYSAYFSVRADDFDLVAMAHGKETVKARMKEINDKGNEASAKEKNLLTVLELANEML
ERGHKFKMVDLKRSAASDWLIDGDDLLAPFTAIPGLGLNVAKQIVAAREEGEFLSKEDL
SKRGKVSSSLIEFMTVNNVLDDLPDSNQLDLFENLF
>Prolyl-tRNA synthetase (EC 6.1.1.15), bacterial type
MKQSKMLIPTLKEVPSDAEAISHQLMLRAGYIKQISAGMYSYLPLAHKVLLKLEQIIREE
MDKIDAVEMLTPAVLPAELWQESGRYESYGPNLFKLKDRHERDFILGPTHEETMTTIVK
DAIKSYKKLPLTLYQIQMKYRDENRPRFGLLRGREFLMKDAYSFHADMDSLDKVYRDM
DVAYQNIFDRVGLNYRSIIGDAGAMGGNDSKEFMAIAPIGEDTVVYSDGSDYAANLEM
AQNKRTLQPSHEIPKELTKVETPGCKTIDEVAAFLKTKPENTIKTLLYVADDKLVVVLLR
GNDQLNEVKLTNFLGADFLRPATNDEAKQYLGADFGSLGPVGISEDLQVLADLDVKAM
ANASVGANEDGYHYINVNLNRDFKVDEFVDLRTVNEGELSPDGQGTLKFTRGIEIGHIF
KLGTRYTEAMDAQILDENGRQKPIIMGCYGIGVSRLLSAIVEQHCDEHGMIWPKEIAPF
DIHVVPINMKKAEQAKLAQELTATLETAGYSVLVDDRNERPGVKFADSDLIGLPVRITVG
KKASEGIVEIKLRQSGETVEIKQDELLNTLNILLANNAD
>Membrane-associated zinc metalloprotease
MIVTILAFVLVFGTIVFVHEFGHYTLAKRSGILVREFAIGMGPKLFSYRKSGTTYTLRILPL
GGYVRMAGLEEDSEELKKGMVVSLILGENGKVSKINTSQKVSLMGGIGLELADWDLVD
ELYISGYENGNEDELKRYEVDHDALLVEHDGTEVQIAPRDVQFQSAPLVSRMLTNFAG
PFNNFILAIVAFALVAVLQGGVKTMTNTIASVQPNSVAHQAGIKAGDQLVEVAGVKTND
WTKLAQTISERPGKKTTVVVKRQSTRLKFELTPKATQVNGQKVGLIGITSKLTVNRSPL
AIIGYGFKESWQVTIQVLAALKNMVTKGFNLDQLGGPVAMYSFTSQAAHYGLASVIMLL
GLLSINLGIFNLLPIPALDGGKILLNIIEAIRGKPLDPNKEMIVTVAGFAFLLILMVLVTWNDI
ARYFLH
>Phosphatidate cytidylyltransferase (EC 2.7.7.41)
MKQRVITAIVALIIFIPLLILGGTPFNLLVVAMGLIGMSEFLRMKKKLLVSPEAIIGFLTMAS
LLIPKTWLDFIPEQVTNLGMFYFFSLCLLLYTVFSKNRFSFDDAGVLILGSVYAGIGFGY
MMHARAVSLWMLLYALLIVWLTDSGAYMIGRKLGKHKLAPHISPNKTWEGSIGGSLVA
VVVVSCYLYFFPELHAYDSVTMIGLTLLYSVGGQLGDLVESAFKRYYGVKDSGNILPGH
GGILDRFDSLLFVLPLMHFTGLI
>Undecaprenyl diphosphate synthase (EC 2.5.1.31)
MFTKLFKNNNDEAEQIELDESKIPNHIAIIMDGNGRWAKRRHLPRIAGHRQGMDVVREI
TRAASDLGVKVLTLYAFSTENWKRPDDEVSFLMDLPVKFFNKFVPELIEQNVRVNVMG
YVEQLPAKTKKAVLDAIEQTKDNTGMVLNFALNYGSRAEIVTAVKELSQQVKSGSLEVE
AIDEQTISRALMTGKLAPYCDPELIIRTSGEERISNFLLWQLAYSELVFTETLWPDYQPA
NLYQDIKTYQSRDRRFGGVTEEEK
>Ribosome recycling factor
MDKILEAHQAKMTKAEEVLQRELGNIRAGRANASLLNRINVEYYGVPTPLNQMAQISVP
EARVLLVTPYDKSSLKDVEHALLASDLGLTPANDGTAIRLVIPQLTEERRKELAKEVKAT

SEKAKVAVRNVRRDMMDALKKAQKNGDLTEDDLRDLENQAQKLTDASIKNIEAITADK
EKEVLQG
>Uridine monophosphate kinase (EC 2.7.4.22)
MSEIKYKRVVMKLSGEALAGAQGHGINPPEIRKVAEEIKEVHDLGVEIAIVVGGGNMWR
GEAGAQMGMERAQADYIGMLGTVMNALALQDNLETIGVPTRVQTSIEMRQIAEPYIRR
KAVRHLEKGRIVIFAAGTGSPYFSTDTTAALRASEINADVILMAKNGVDGVYSADPKTN
QEAVKFERLTHMEVINKGLRVMDTTASSLSMDNDIPLVVFNMNEHGNIKRVVEGENIG
TTVEGK
>Integral membrane protein
MFKNKILQSFLALLFGMCLFYGGKIQADYAITNYDVNVNLQQDGSADITQQITYNFKGT
SRGVYYDQPLEKGVVATQPSIVVQENGQSIQLTPNQDGSNNSFQVIETDNALKLKLFH
NTADTTTTYYYSYHVTNFVKNWQDTAELNWTIIGNKWNVPLNNVRLKFQLPSTNVADL
KAWTRSESKATTSVSKEKGQVVVNLKQNPSASQVKTRLLFPTSVVSTNGLVVNQKHK
QAAIKSEQAWLKDNTQTKKVSNRNNLIFKAGVFGLALVIILVWALYLLSHPVKKQKWPD
VNPPHSYEIPPMSAPVAMSIYTDNVPDNNALGAYLLELAVKQKIAIEEFNTKRKDYQLTL
LDASLANSEPMIDFLFNTVGDGKKVTLHELRKYGKTQRESKQMSHIYDKWKNQVEQE
ALSYDYLSKKSAGNFYLTVFLCGISCLLLLWSLFLWKTPWMTGLSLVLILGQVGFNVYH
FRKHSPYTETGVELITQLRAFRHTLKDLNKINLKSKGDLNYWSQVLPYAIAFGYNKRVIN
ALQANFTVKELEEGLGFYYNLLFTAPIDFGDSFSTSVTRALGNYHPNNN
>Hydrolase (HAD superfamily) in cluster with DUF1447
MTVKLVAIDVDLSQDVMAIGDGGNDLPMIKAAKIGVAMANALPEVKAASDFVTADNNH
DGVAVAIEKFILA
>hypothetical protein
MKIKVRNIGNNVGIILPKELGLVSGDIIQAEKKGNLFILDTSEIAREHDRKLVEDSFADFEK
ELIVSESKMKAIFEKYGWK
>LSU ribosomal protein L19p
MSVNPLIAKLTEAQLRNDIPEFRAGDTVRVHARIVEGSRERIQIFEGVVIKRRGEGISET
YTVRKISNGIGVERTFPLHTPRVDKIEVVRHGKVRRAKLYYLRALHGKAARIPERRR
>tRNA (Guanine37-N1) -methyltransferase (EC 2.1.1.31)
MRIDILSLFPNMFQPLKESILGKALEKELLEIEITDFRDYTTQKQRHVDDTPYGGGAGML
LQAQPIFDALAATQAKAKSKGRVILLDPAGRKFDQKVAEELAKEDHLTFICGHYEGYDE
RIRTLVTDELSLGDYVLTGGELGAMVMIDATVRLLPEVLGNSESAQTDSFSTGLLEYPQ
YTRPADFRGMKVPEVLTSGNHQKIAEWRHKEALRRTYERRPDLLEDYPLTDQDRKWL
AEFAAEKTDLAEK
>16S rRNA processing protein RimM
MQYYQVGKIVNTHGIKGEVKVLASTDFKESRFKPGAKLYLFQKQTQTPLEVTVKTHRK
HKQFDMLTFVGLEDINLVEKYKQADLKITADQQEELAEGEYYYHQIIGLTAITTEGRELG
KVKEIMTPGANDVWVIARPKQADLLLPAIADVIKEIRLADGQVLVEMMDGLE
>KH domain RNA binding protein YlqC
MTEADVKQLIITIVAPLVADPEAIRIETKRDTRFLNFNLTVAKDDVGRVIGKHGRIAQALR
TLVYSVRLDDKSLRARLNIVDD
>SSU ribosomal protein S16p
MSVKIRLKRMGSKKSPFYRIVVADSRSPRDGRFIETVGTYNPLTEPETVTLKEEEIMNW
LNNGAQPSDTVRNILSKNGVMKKFHEAKFSKK
>Signal recognition particle, subunit Ffh SRP54 (TC 3.A.5.1.1)

MAFEGLTERLQNAISKLRRKGKVSEADVKEVMREIRLALLEADVNFKVVKDFVKTVRTR
AVGAEVLESLTPAQQIVKIVNEELVKVMGQEASPLNKAPKIPTIIMMVGLQGAGKTTTAG
KLALKLKEEQHARPLLIAADVYRPAAVDQLKILGQNIDVPVFEMGTDVNPVEIVRQGLAL
GKEKKNDYIIIDTAGRLQIDERLMNELAQIKELSQPNEILLTVDAMTGQNAVEVADGFNS
QLDVTGVVLTKLDGDTRGGAALSIRAVTGKPIKFIGQGEKMTDLDVFYPDRMASRILGM
GDMLTLIEKAQKDFDEKKAAEMAAKMRENSFDFNDFLDQMEQVQNMGPLEDIIKMIPG
MANNPALKNMNVDPKDIERTKAIVYSMTPAERENPDLLNPSRRRRIAAGCGRPIQEVN
RMIKQFKQMREMMKKMSNGNFAGMEGLLGNGMTGKLAKMSMNRMMRKNKKRKLQ
RLRKRKK
>Signal recognition particle associated protein
MAIEKTNRINALFEFYEPLLTEKQMTYIELYYRDDYSLGEIAENYQVSRQAVYDNIKRTE
KILEEYEDKLHLYAHFQKEVAQTDRLLAYVSENYPQDKQLMALVRQLEKLEE
>Signal recognition particle receptor protein FtsY (=alpha subunit) (TC 3.A.5.1.1)
MGLFDKLKRAFGGGQAEAKPESEEKVATENPAPQAEVEPTPAAVEPAPEEKTKAGQA
EVNQESVETESEPEVEVAAPIPVQPAEVEEKVAQKPAQEVAANEETKRYERGLAKSRK
GFGAKINALLANFRSVDEDFFEELEETLIEADVGFETAMKISDQLRDEVKFQNAKKKAD
VSDVIVEKLVDLYEAEGENEDNQLHFATQGPTVFLFVGVNGVGKTTTIGKLAHRYQAA
GKKVLMAAADTFRAGAIEQLVEWGKRVDVEVVRKPEKSDPAAVVFEAVKKAKEEDYD
LLLVDTAGRLQNKVNLMNELEKIQRVITRELPDAPHEVLLALDATTGQNALTQAKQFGE
VTKVSGIVLTKLDGTARGGIVLAIRNELHLPVKLVGLGEQMDDLRDFNPEKFVYGLFKD
LIAQR
>Chromosome partition protein smc
MKLKALTLNGFKSFADKTKIEFQPGMTGIVGPNGSGKSNIIEALRWVLGEQSVKSLRGG
KMPDVIFAGSQTRAPLNRAEVVIELDNSDGFLKQQPPELVITRRIYRDGESEFLINQKKV
RLREIVDLFMDTGLGRESFSIISQGRVEAIFNSKPLERRVLIEEVAGVLKYKKEKHRAQK
ELEETTGHLNRVADIITELYQQKQPLEEQASIARDYLEQKKQYDYFDLNNLVLEIKAKSQ
KRAEIQAELAKIKDLVAKHSQSASKQEQLTANLHQQQQQLEAKLDAAQERLVELTKQK
GYLSGEEKVSKQTREYQELQRKDTSDRLEVDQKNLRELKEQEAKLEAQLAQVQAQKA
DYQAKIDRLTADANIDEATLNEQVEKLRQELFEKSQDQSALQSQHAYLKRESEHSQSL
TASALAKHKELEAAYQAAKEKKHQAKADLLAAKETVAKLDQRQVELTGEVSKYQLRYD
NQKERWYEASDILKQAQTRYETLKNVATSYDGYYQGVKQVLQAKGKLAGVIGAVAEIIT
VPQEFSQALEAVLAGQLQNVVVTTDEAAKQAINYLRANKLGRATFLPRNTVKRRALTT
KQRQILASQQGIVGVASELVTVNAADQPILDYLLGTTVVVKSLAEATVVARALNHSLRIV
TLAGDIVNAGGSMSGGANQQRKIGLLEQKQQLEKLAANIDLMKSKLATVELEGNEAKQ
QLTASKQAVQQLAPKLQAAQTKLQQCQQQFDISQVELSHQEQQLRAEAASQEEQTKA
SMQAQTELKEVVAKLDQLEQAIAQLKEQLAYKQEFLNDTRTSQAKNQQKLTTLKQELA
VVEERLQSVKLQANEVANQIKRLSENIAKAKERLAILKQEQELRHGQVEDLADKQAKVK
QEHAQLEQTVAKMIEQRKTLHAQVAQAEEQLKRASELKAAAYDEQSEKARQDSKLEA
QLDGLLTDLADRHELSYEAASKLENLETDIDYVKRRLKLLQLGLDELGDVNLGAISEFD
RVNERYEFLAKQQDDLLAAKAQLEKSMQEMDGEVKRRFKQTFEEVARAFEEIFPAIFG
GGRAKLSLTDSDLLTTGVEIMAQPPGKKFQTLTLLSGGEKALTAIALLFAILKVRPVPFV
VLDEAEAALDDANVVRYSQYLQNFEGKTQFIVITHRKGTMMNADILYGVTMQESGVSK
MVSVSLAQLEQ
>Ribonuclease III (EC 3.1.26.3)
MIIGLAKMLHDNFQIDIKNYALFDEAFTHASYVNEHPKENLKYYERIEFLGDAVMQLCVS
EYLFKRYPNLPEGKLSRLRAAMVCEDSFSKFAKECHFDEYIRLGKGEEKANARNRPSL

LCDIFEAFIGALYLDQGKETVVSFISKVIFPKLDMGWFDHFLDHKTELQELLQQDGDCEI
EYQVLAAKGPDNAKQYYVAVLANQKELGRGTGHTKKAAEQAAANQALKELRG
>Acyl carrier protein
MTEEEIFAKIAKIIAERFELPEDKVTNDLNFQNDLNADSIDIVEFVLELEDAFGAEISDEDA
EQIATVGDAVKFIAARQ
>Phosphate:acyl-ACP acyltransferase PlsX
MKIAVDAMGGDNAPKTVVEGVLRARDEFADLEFCLYGDEQQVKPLIDNMERIEIIHTSE
VIDMNDEPVKAIRRKKDSSMVQAAQAVKENKADALFSCGNTGALLTAGLLIVGRVKGV
SRPGLLSTLPVLNATSGDGRFNLLDSGANADGKPEQLYQYALLGKYYAQNVRGLANP
RIGLLNNGTEPHKGSKLTLEAHQLIAADDSINFVGNVEARDLLTGVCDVVVADGFTGNA
VLKAIEGTASATLHLLKDAIMAAGVSGKLGALMLKPVFKELKGKLDQSQYGGAVLLGAK
APVVKAHGASDAKTVYYTLKQIREIVAGQVVEKFSAYLEQQRSEQTENN
>ATP-dependent DNA helicase RecG (EC 3.6.1.-)
MHGQLSDSVGTLRGVGPKKVQSLNKLGINTIADLLGYYPFRYDDIAVKSLAEVADGQK
VALKGIVASEPVLARFGRAKSRINFRFLVEHDVILVTFFNQPYLKQQLVSGKEIVIYGKY
NAKRQALSAMKLLANQGGNDFTGVYRASKEIKATTIKSLVKDAYELYHESLQDLVPVKL
RAKYRLESHQQIVHDMHFPATAKAANLARRSAIFEEFFLFQAGIQFMKKLARKDEGIAL
KYNNDELRAFIKTLPFELTAAQKRVVNEICADMRRPAHMNRLLQGDVGSGKTIIAAIAM
YAAVTAGYQATLMAPTEILAQQHADKLADLFEPLGVNVALLTSTTSNKAKLRQELLAHL
AKGEIDILIGTHALIQADVEFAKLGLVVTDEQHRFGVNQRKNLRQKGAKPDVLAMTATPI
PRTLALTAYGEMDVSIIDELPAGRKPIKTSWVRKKQLKELWGFVAQQLELGSQAYVISP
LIEESEAIDLQNAKEVYANLAQLFKDRFRVGILHGQMKAAEKEAIMAAFKKHELDLLVST
TVIEVGVDVPNATVMVILDADRFGLAQLHQLRGRVGRGSKASYCLLVADPKTDYGKAR
METMVQTNDGFIIAQKDLELRGPGDVLGSKQAGIPDFRVGDPIADLKILQIAQMEAQQV
LAAPDFLTNPENAGLIQYLHQKVGQGQKFD
>Dihydroxyacetone kinase family protein
MKVTKITETEFRKMVLVSSKRLNSNAEFINSLNVFPVPDGDTGTNMNLSLASGEKYVSE
STTSNVGELATALAKGLLMGARGNSGVILSQIFRGFSKSTAKKEILTAQDLADALSAGV
EMAYKAVMKPTEGTILTVARKAAEAARKLVKTSDDCLAVMQAAYEAADAALQTTPDLL
PVLKEVGVVDSGGQGLTFVYQGFADALSGRERQESHQPSVVEMDEMVNAEHHKSVQ
SQLHTEDIKYGYCTEIMVRLGAGRSVEHEFNYDDFYAALAAIGDSLLVIADDEIVKVHVH
TEQPGKVLTLGQSYGSLIKVKVDNMRLQHETILEKDEEQAADEVEVNKELTGDYGVIAI
ASGDGVAKLFKSLGVTYILSGGQTMNPSTQDIVDAIKETKAKKVLILPNNKNIFMAAEQA
TKVAEVAAEVVPTKTISQGLSAMLGFDKDASLEENKAAMLAELDGVKSGQVTIAVRDTT
IEGQAIKKDDYMGIVDGQIVVTNPDRQAAATQMISKMLDEDSEIVTIIYGQDGSETEAQA
IKAAALELDEDIEVEIHAGDQPVYPYLISVE
>FIG001802: Putative alkaline-shock protein
MAVKIQNQFGKIDISNDVIATVVGGAATEIFGIVGMASKNQIRDNLNEILKRDNYAKGVV
VRQEDDGVAIDVYIIAGYGTKISEVCHNVQEKVKYNLASMLGVTAKSVNVIVQGVRVIE
D
>LSU ribosomal protein L28p
MAKDFVTGRKTTFGKKRSHALNQTNRSWKPNLQKVRILVDGKPKRVWVSTRALKSGK
VTRV
>Thiamin pyrophosphokinase (EC 2.7.6.2)
MEANILVGGPKQFWPANFEQLATKHAQDFIWVGADQGAVALCELGISPNLAIGDFDSA
SSRNQFLVQVRAQKIIAARSEKDYTDTQLAVLMTEKYFPRVDKINIYGATGGRLDHLLA

NIFLVLGPEFEHLATKVRLFDCQNSVSFYLPGEYQLLKEAQMTYLAFIPLTAVKDLSLFN
TKYRLAKAKFNRPISLASNEFMEEKAHFAFSEGIVCVIQSRD
>Ribulose-phosphate 3-epimerase (EC 5.1.3.1)
MKIAPSILSADFAELKTEVKRVANAPYLHIDIMDGQFVDNLTFGPNVVKALRPHSNQVF
DCHLMVINPEKYIKAFAQAGADIIGVHIEATAHIHRVLSQIKQAGVKAEVVINPGTPVAAIK
EVLPLVDQVLVMTVDPGFGGQAFLSETVSKVKELAQLRLQAGLQFEIEVDGGVNADTI
KTVTAAGADVAVAGSFVFESDDPAKQVALLKELAQG
>Ribosome small subunit-stimulated GTPase EngC
MIRGQIRQSLSGYYDVMGEDGVLYRTRARGNFRKKGQSPLVGDWVELKAEKPDEGYI
LKIEARTNQLVRPPVANVDSAIVVTACCQPTFTSNLLDRQLLMLESNGIKPLLYFSKYDL
LPAAKQAEMQEVMAYYNQLYPVYVQTKAADEQVFEDLLADLVGQVVVVMGQTGAGK
STLLNKLKPDLALETGQVSKALSRGKHTTRRVTLLDIKGSLIADTPGFSSFDILNISKEEL
PLLYPEFRQRQGACKFRACLHVNEPQCAVKEAVAKKEIRQSRYDNYLQLHEMILNQKP
KYR
>Serine/threonine protein kinase PrkC, regulator of stationary phase
MQAGYLLNERYKIIRTLGEGGMATVYLAEDIFLNREVAIKILRLDLRDDSAALRRFHREA
RALTELSNQHIVNIYDISEKNGLQYLVMEYVPGTNLKEYIESHHPLTYPRVVEIMLQILSA
VKEAHRHGIIHRDLKPQNILIDTEGNIKITDFGIAVAVAEDTMTRTNTLMGSVHYISPEQA
RGSIITKQSDIYSLGIILFEMLTGHVPYEGETAVAIALKHYQSEMPSVKDYDQQIPQSLEN
VVLHATAKDLASRYQNVDEMELDLSNALDQSRLNEPKWHGRQQVEAETMVLPNIKDE
ASQSPKKATKKRRQVLLGGLLVVSLMVIAGLVYLLMSPKEVTVPDVRGMTIAEAQESLN
GEKLALGKIRRHYSSKYYSGQIITTNPEPDTTVKQKSKIDVVLSKGQRKEPFGDYRGQS
YQTVAKKLTKKGVSVTKTGQYSNKVAKGRIISQSISAHKKVSFNETNVIFTVSAGAKSE
RLRDLSNYTEKSAQDYANDLGLKLVVNEADSSSLPGMVIAQQPQAGTKVKAGDTLTVT
IAKNDGSSYTSSNAASSSTDTTTGATTFSFDVEVPYAATNAGSNKVQIYIADSDHSYSD
VYQTATITKDTNFSLTFTLKPEETGKFKVVRDGQVIAENNNVDH
>Protein serine/threonine phosphatase PrpC, regulation of stationary phase
MKVSYRSDVGKARKNNEDSVGYFKNKAGIVLGIVADGLGGHQGGEVASEMAVSHLGY
QFEQTNFYEAQAGLKWLQQVVSAENEVILEKAAQYHDLHGMGTTLVCVLVFADEFVLA
NIGDSRGYLLRDGQLLQLSEDHSLVNELVKQGKISPAEAQSHPQKNIITRTLGISTSAKL
EANIYPSKLNDYLLLCSDGLTNMVSEAQLQQVLEEAGDLASKCQQLISLANQAGGLDNI
TLLLVHLDSEVSS
>Ribosomal RNA small subunit methyltransferase B (EC 2.1.1.-)
MPNRKLKASPRYLAVEILTRVFKQGSYSNLALNQTITKYQLDKRDAHLLTTLVYGVIQR
QLTLRYWLKPFIKKPAKLEAWVEVLFLVAIYQLEYLDKIPKRAVFNESIEIAKVRGHEGV
RRFVTGVLHEIERQGLRSTALIADETERLSVEASCPLWLVEKLIGEVGLVKARSILASIN
QAPEQTVRVNTKLVSPTQAREKLQAEGFEVATSPVTPLALRVSGGFVPASHTFKAGEV
IVQDESAMLAVESMAIAGPEQVLDACAAPGGKTTQIAMGLTTGTVVALDIHKHKVKLIAE
NAKRAQVSNLVTPQQLDARMVATEFKAGQFDKVLVDAPCSGLGLLRRKPEIKYSKSEE
DSQNLSKIQLAILDAVAAGVKEGGYLTYSTCTILSQENQRVVTAFLAKHPEFEQVKTKT
AYNLKADRQELGLTIYPDDYLSDGFFIANLQKKKLGDD
>Methionyl-tRNA formyltransferase (EC 2.1.2.9)
MTSVVFMGTPAFAAPILEGLIQTNDYEVLAVVTQPDRPVGRKRILTPSPVKEVALKHGIE
VLQPQKLSGSEEMQRVISLAPDLIVTAAFGQFLPTKLIAAAKVGAVNVHGSLLPKYRGG
APVQYAIMNGDQETGVTIIYMVKKMDAGDMLAQAKLPIGPHDDTGSIFAKMSILGRDTL
LATLPKLVAGKLVAQKQDESQVVFSPTIKPEEEKLSLTLNAKQIDWKVRALRPAPGAYF

ENFNGKRTKLWDVTPLAETTSLPAGAVVAVTKHELKLAAADGSVYQVNELQPAGKPR
MQVTAYLNGVGQGLREGQVIIGAE
>Helicase PriA essential for oriC/DnaA-independent DNA replication
MKTQYAQVIVDVATMQTNQPYTYVIPPELAGQVEAGMRVVVGFGRGKRRVQGFVVAV
SAENPLTSSAKLKPILEVQDLRPVLNQEALQMADWLARYTYAFKITCLQAMLPAAMKAA
YTKTVSLTTSEVTPAVREMFAQKSELALAKLSPEQKKLLVPLSKQGKIEINYHVKDRAK
SKLVLAIKRSLSETELQAAKAKLRKGAKKQAQLLDYLQTLPLDKEVFQANAQAKTGLGA
AIFKAGAEKGWLKQSKVESYRNPYKQEVVKTEAKSLLPQQAQVVQAINRASDEGKAQ
TFLLQGITGSGKTEVYLHAISHVLAQGKQALLLVPEISLTPQMVSQVKARFGKEVAMLH
SGLSDGEKYDEWRRIERGEAQVVVGARSAIFAPLPKLGLIIMDEEHEASYKQDDMPRY
HARDVAIWRSEYHACPLILGSATPSLETRARAQKGVYGWLKLTHRANGKALPAVSLID
MKKEIMTAPVQDISQVLLAAIKQKLAQKEQVVLMLNRRGYSSFAMCRECGFVLKCPNC
DISLTMHLDSHSMRCHYCGHEQGIPSLCPSCHSKRIGYYGTGTEKVETELKELLPTARI
LRMDIDTTRKKGAHEKILRQFGARKADILLGTQMIAKGLDFPNVTLVGVLNADTALELPD
FRASERTFQLLTQVSGRAGRAQKQGEVLIQTFNPENYAIKLAQKQDYERYFRMEMAV
RHRGGYPPYFYTIKITTSYRDEATAAKKIFQIAGQLKQVLSPTAQILGPSPRSVMRVNNR
YYYQLVIKYKRESLLEDYLQHLLLISQKEERQGLRIVIDREPINFI
>Phosphopantothenoylcysteine decarboxylase (EC 4.1.1.36) / Phosphopantothenoylcysteine synthetase (EC 6.3.2.5)
MEGKRIAVYLCGGIASYKVVSLVRMLIKNKAQVRVAMTKEAQKFIQPLTLATLTKHQVY
TDLFAENNQDFVPHIDLADWSQAAIVAPATANIIAKMATGLADDFVSTALLASTCPKYVV
PAMNEHMLLNQATQRNLARLQADGIKVLETKRGFLAEGYEGQGRMPEVEEIYAWLVQ
QLEVKQDLKGKRVVVTAGGTHEEIDPVRYIANRSSGKMGYALASAAAQRGAEVTLISG
PSALPQPMGVKFIGIETAAELETAVKRAYPTTDIVIMAAAVADYRIKKVASQKIKKENAS
WHLELVKNEDILFGLGQTKRNQLLVGFAAETENLLANAQAKLTKKNLDLLVANDVARK
DIGFASDSNEVTLLTKNQEPKHLKRASKQVIAGAILDEIVRLM
>DNA-directed RNA polymerase omega subunit (EC 2.7.7.6)
MILYPSVDDLLEEVDSRYSLIMLASKRAHELDAGELPMLSEYESVKSVGKALEEVVAGD
LKIKQD
>Guanylate kinase (EC 2.7.4.8)
MLSGPSGVGKGTVRKAIFAQDNNKFDYSVSMTTRQMRPGEVNGKDYYFVTKEEFEKE
IAEGGMLEYAQYVDNYYGTPLKYVNEMLDEGKDVFLEIEVKGAMQVREKVPDGLFIFL
TPPDLMELRQRLVNRGTDDLDVIDMRMKKAVDEIEMMQNYDYAVVNDEVEAAAEKIKT
IIRAERWRVKRFLPDYKKQLGDVLK
>hypothetical protein
MSLLQHKLLQAPFIKHLKTTLSQQVENFFAPGWQETATESEITTPVQIMPLFQVELFINK
AIIKRQPVQLALKTPTGIEQLTGILHYQSVKGQQLLILQQANVMRPLKTMDVLYIQLSPTT
NQLTNYRELA
>Glutamine synthetase type I (EC 6.3.1.2)
MAKATYTKDDIRRIAKEEDVKFLRLMFTDLYGTIKNVEVPISQLDKLLDNKLMFDGSSID
GFVRIEESDMWLYPDLSTWMIFPWGNEHGKVARIICEVYNADRTPFYGDPRNNLIRILD
EMKELGFTDFNIGPEPEFFLFKLDPETGKPTTHLNDHGSYFDLAPMDLGENCRRDIVLE
LEKLGFDVEASHHEVAPGQHEVDFKYADALTACDNIQTFKLVVKTVARKHGLHATFMP
KPLERINGSGMHLNMSLFTKDGNAFFDENGQEQLSKEAYYFLGGILKHARNFTALTNP
TVNSYKRLVPGYEAPVYVAWSGHNRSPMVRVPVARGASTRLELRSVDPSANPYMAV

AAVLAAGLDGLKNKIEPAQPVDRNIYSMTEDERKAAGIVDLPSTLHNALKAMREDEVIK
DALGNHLYSNFMEGKRLEWDAYRLEVSQWERDQYLEMY
>Transcriptional regulator, repressor of the glutamine synthetase, MerR family
MKEKELRRSQAVLPIGTVMKLTNLSARQIRYYEEQGLVIPERNEGNRRMYSLNDIDKLL
EIKDFLADGINMAGIKRIYEEQA
>Aluminum resistance protein
MMDWKAKFPQELQAKIRKVEAKIAPKLAEMDEQALYNQQRVLELFRKHKVSEEDLVSS
TGYGYDDIGREKLEKIFADYFKAEDAYVRPQLVSGTHAIATALFSVLRPNDTLYYLTGM
PYDTIQEVIGVAGDNPGTMKEYGINFAYTDLLANGKVDYQAAKEKLLTDKSIKVVAIQRS
RGYATRDSFVVAEIKEMIEFVKQVAPQAVIFIDNCYGEFSETTEPTFYGADLMAGSLFK
NAGAGIVKGGAFLVGRHDLIEGAAARLVAPGVGKGEGPTWGYMRDFYQGFFMAAHT
TVQAIKGSIYTAALLAEMGLEVSPAWDAPRTDLVQTVIFHQPEPMIKFCACIQHYSALNS
YVDPVPSYMDGYEDQVIMASGSFTEGSTVELSSDGPLREPYALYIQGGLSYEHVKIAIS
NAVNETFYQG
>tRNA dimethylallyltransferase (EC 2.5.1.75)
MQVYKGLDIGTAKIKPIEMQGIKHHLIDCREVNERWSVAEFISEAKQIMADCWQRGKLP
LIVGGTGFYVQALLDGYQLGQDNYEETEIRTELEEYVKAHGTQALWQQLAAVDAKAAA
KIPVNNVRRVIRALEVYRKTGKLFSSQEDEAQNSFEPLVLGLTTERTVLYERINHRVEL
MLADGLEAEARSLYEHGNLEMPAALGIGYKEFYPYFQGEYSYERAVELIKRNSRRYAK
RQLTWFRNKMDVKWFDLIEHPEQEAAIMKLIEGWTRK
>FIG00743062: hypothetical protein
MKLFGKGFLVGVATTLGTLAGLAFAFKKTVVEPIEEREAELDEHRRRALRKSRSAHQG
>Rhodanese-like domain protein
MVGAINVLYVLDIILIIILAWILGVQLYTWLLGKKVAKIVEKDEFKAGMKKAQLIDLREPDS
FKASHILGARNMPFSQFKLYQGSLRKDMPVYLYDQGKSISIRAAAKLHKAGFKEIVILNG
GFAKWDGKIKKAKL
>Glucokinase (EC 2.7.1.2)
MDKKLIGVDLGGTTIKFAILTAAGEIQQKWSVETNILDDGAHIVPDIIDSINHHLALYDMKP
EQFIGIGMGTPGTVDIKAGTVIGAYNLNWKTKQDVKSAIEAGTGIKFAIDNDANVAALGE
RWKGAGNNADDVVFMTLGTGVGGGIIANGQMIHGMGAGGEIGHINVEPNGYLCTCGN
KGCLETYASATGVVRVARDMAEEFAGKSELKRMLDDGEEISSKIIFDLAKEKDTLAAMV
VDKVCYYLGLAAANLGAITNPAYIVIGGGVSAAGKFLLNQVEDYFQKFAFPTVKTSTKLK
LAELGNEAGVIGAASLALAFK
>FIG00743953: hypothetical protein
MKTLYDVQQLLKKFGIYVYVGKRMWDIELMALELDHLRQAQVIDEPTFVQAKLVLTREH
RIEEKREQEKLK
>GlpG protein (membrane protein of glp regulon)
MQIKRPLVTISLIVANIVVYILMTVAGGSTNPLVLIKFGAKVNEYILMGQWWRLITPVFIHI
GLTHIVVNMVTLYFIGMQIEYLFGASRYLLIYLVSGLCGNIASFVFNPDSVSAGASTALF
GLFGAYLMLGEAFRENVYIRLMSRQFLALVILNIAFGFFGNIDLAGHLGGLFGGFLIAYVL
GVPKVGSVSKLKRITAAGALIVIFGLMLYSGFNLI
>5-formyltetrahydrofolate cyclo-ligase (EC 6.3.3.2)
MQTKVEFRKEQRQRLIDQADTWQKEKEELALYLKLFATKKWQEAKVVALTLSQPEELD
TKPMIMAAFQAGKRVCVPKTLPKRQMVFVELTPEVQITRSKFGVLEPTANYHQLAKDQI
DLIVVPGLAYTPAGDRLSFGGGFFDRYLADYPGKTIALANKSRYFTQVQWEVEPTDIQV
QEVINLEENKRAN

>D-lactate dehydrogenase (EC 1.1.1.28)
MTKILMLSVRLDEQAAIKAWQDRHPGVEVTTAAWELRASTVDQTAGFDGLVIQQRYNI
EAEVYPKLKALGFKQISARTAGFDVLDLNLAKQNNLVVTNVPAYSPRSVAELALAHTLR
LTRNLELFDDRASRQDFRWAGLQAKEIHSLTIGIIGAGRIGGTAARLFHALGAKVIANDL
VENPELKDVLTYLPFEEVLKQADVVSLHVDLNETTHNLIDEHALSLMKPTAFLVNECRG
PVVDTDALINALENKQLAGAALDTLTGEENFFNFNLEGQALPSQQLVKLRAMPNVVITP
HVGFYTNIAVQNMVDISLDSALAVLEGQKPTTQVN
>Cell division protein FtsI [Peptidoglycan synthetase] (EC 2.4.1.129)
MEFFRKVRKTNNKRKQKKSVIPFRLNFLFLIVFLLFAALIGRLAYLQILHGTMFESEVSSS
ETETETQNVQRGQIYDSTGKVLVSNKSARAITYTKPLNVTNKEMYAVAKDLAAFISVPT
KNLTATNKADFYLAQHEKSVNKHIPGAAKMTPSALYKAQLDYVKKNKLYGSLSPNQKEI
ALLYSKMSGAYALSTVYLKTEGVTDKEMAKVGEHLSEMAGVKIGTSWRRDYPNGTSV
KSIIGTVTNEKTGLPDDSINQLLAKGYSRNDSVGSSYLESRYEDILKGTRKTIDVETQNG
KILKETTKYGGKKGDNVVLTINEKFQNEVQKFLEDGFNGIKGSNPYAPGAYVVVINPNT
GGIYALAGVSRNLSTGKVTSNALGTINQSFVMGSVVKGAQVMGGLTSGVITPTNNSLY
DEPIKLKGTAPKTSWFNKTGGSGMSLTATQALEVSSNSYMMKLTMMEAGYKYVPGGA
LSMPSSIFDKLRQNFNQFGLGVKTGIDIPGEGLGYAGASNQANIGKALDLSFGNYDSYT
TIQLAQYVSMIANGGYRLQPHVLEAIRESNADGSLGRTVYQAQPNVLNVYSASGAEWN
VVQNGFYQVVHGSMSLRTGSAYASLKPEVSAKSGTAQTFHGTAETITNSLIMYGPSKD
AKVAMAVVFPGMGTSSSSSVVASVAKQTYEAFWKDVITTEGFTDSSNEGN
>transmembrane protein Tmp5
MQKVRSYFKEHYPLLLSFWGPLLLMLGYFIFRKMYPFGTSSILTVDMGQQYVDFFQFF
RRTLLTDPSNFFYSFSKAIGGEMLGEWAYYLCSPFNLLILLFPGKSIMAGIIVITLCKYAC
AGYCFGCLLLKRKIQTGILVPTFSTAYAFMGWAVANQLNLLWLDVLVFLPLIILGLLDLM
DTNRGYAYTFWLAALLFINYYMAYMVCLFLILFFAWYALTSFTDVKKLGKQVGLFVHRS
LLAAGLAAFLLIPTFISLKASKGQYTTQTIHFKWEYFPPKMLTKLVLGAFDFKQMPSGYP
NIFVGSLVLFAFLLFFFNRRFKLSSRLAALALTGFLLVSMCFEPLDLLWHGFQFPVWYP
YRFSFLLSFWMVFLAAQAFNSPTASPKWWQIWGCFFIEAGILGYSYTVMKKVSYLNRT
VLIFSALFAVATLLLLSIKPTKTKNLRNLCLYLLVLGELVANVCFSLNNISYLTAKEYSNPS
QALTQDAQSVHKLDNGFYRTGQTYSRTRNDGLALNFNSGSYFSSALEKAIPDFYGQIG
NPDGDNYVSYSNGTLISDSLLDMKYFFNQKDSSELEHKSADLNLAALTEKQDLNLYRL
VKDNVYSRIYQNPYASGIGYLAGPSLKSLPSLYDDPVTYQTNWLNAVTNTWPTTKYFT
AKNFNEVVFQNTSQTINLTGATLKRKKANADAQVIFKFVPQTNNSYYLTLGTGLNGDNV
SLFIGNRPLSYYSTFRHTTIVNIANHAKGQEIVITARFKKQTLSLNNFVLYELNNPLVISKL
KQVQAHSWQLKQFNNRKLSGTITVPNKNQLFTTTIPYSQGWQAKVDGKAVKTFKVQGI
FLALNLPKGKHTVSLSYWPPYFNLGLALSILSLILIFIPTLLKRFLKQN
>Adenine phosphoribosyltransferase (EC 2.4.2.7)
MAVDLNNYVASIPDYPEKGIIFRDLSPLMADGEAYRQATDRIVQFAKDKNVDMIVGPEA
RGFIVGCPVAYELGVGFAPARKKGKLPRETVKAEYTLEYGTSALYLHKDAVKPGQRVLI
TDDLLATGGTIGATIDLVEQLGGTVVGCAFLIELKQLNGRDRLKGYDVLTLMEY
>Single-stranded-DNA-specific exonuclease RecJ (EC 3.1.-.-)
MLTSKYEWQIEAELDKNVIAKLAAETKLDATLVQILYRRGYQTATAIENFIHPSDSELND
AFLLYDIDKAVERIQTAAFAGEKIVVYGDYDVDGITATAVMYSTLEQLGADVSYYVPDRF
KDGYGPNKAAYQRLIEAGAQLIVTVDNGVAGHEAIDYANSQGVDVVVTDHHELPPTLP
AAYALVHPRHPKGNYPFGGLAGVGVAFKVATALLDEVPSDMLDLVALGTVADLVPLVK
ENRLLVRYGLEVLRTNPRLGLKSLYELAQLNPEELTEQTIAFQLAPRLNSLGRIAQAKLG

VELLTTEELTAANQLAASCQAYNQQRQELVEQITAQALAKVAANDTNAINLVYERGWH
EGVLGIVASRLVEETGKPSLVLSVNEAGLAKGSGRSVAGFDLFASLDHQRDLLDKFGG
HEMACGLTIAEASLPKLQVQLNAEAKTQGLTAQTKQPLHIDAELKLEQITPTFYQQLQS
LAPFGTENEEPLFLITGYDSLAAKAIGKENNHLRLELSGSGAGPSISGVAFSLGNQAVTT
IVAHSEALRFVGTLSENKWRNQVSLQIMIKDLELDQDKSQGVIIDQRTNKLRKSLFTKVA
SYIFFDAKLKEKVTPLLPSGANALLAAEIDDAYQAQELYLVDCPQSLTQLEALLPRLEFA
KLYLLFYQQDNAYLLGLPTREQFARVYRFSASHKNVAIYQRLDELANYLKIRRELLIFILQ
VFFELGFVKIENGLMNYVPNVKAQALENAECYQARKLRMQAESTLLFSNSTSLTQYLT
KFLKH
>Short-chain dehydrogenase/reductase SDR
MAGYKELKNLRNRVVVITGASGGLGEQLAYQAASKGAIVVACARRLDKLQEVVANCR
QISGRLAFAYTLDVSDPNQIDRVLAEVETNVGPIDVLINNAGFGLMENLLQIDEQTIESM
FRVNVLGLIYLSKYTALRMAKRHRGAIINIASMAGKMATPKSTIYSATKFAVLGFSNALR
LELKPLGIMVTTVNPGPIATAFFEKADKTGNYLKSIDKIVLQPEMVAKKVIRTIGTNKREV
NLPYSMEIASHLYNAFPTLGDYLAGGIFNKK
>Ribonuclease Z (EC 3.1.26.11)
MELEFLGTGAGSPAKFRNVTGIALKLLEERNSIWLFDCGEGTQHQILRTNIRPRKIDKIFI
THLHGDHVFGLPGFLSSRSFQGGDQMEPLTIYGPRGIKEFVNCALKISQSKLSYKINFV
ELDKEGVLYEDDRFVVSFKYLEHRITSVGYRIVEKDYPGELLVAKLQAANIPSGPLYGQ
LKAGQVVTLDDGRIVDGKDYLGPDKKGRIVTILGDTRQTASISELAQDADVLVHESTFG
KGEGKLARSYYHSTSLQAAKVAKNAGVKQLLLTHISARYVGPLVKQLEQEAKSVFANS
RVVKDFDVIEVPFPKS
>hypothetical protein
MQTTTTNDDSENVAILQKRLQAKVKELTNKTEALKAMTAKATNEVDAKSKADAKSKLA
ELQAEVTDLNQEVTSYQDQISQLNKQEQLSLKATSEQLGLVSQKAAVALDQSKVAVN
>Transcription regulator [contains diacylglycerol kinase catalytic domain]
MSPHYHFIINPVAGSGHAKIVWPKLRQILQTRNISYSFDESRYPGHTEKLCRQYALNHP
QAVIVLLGGDGTLHEAVNGLHKETKVSATLAYIPCGSGNDFARGLSIPADPVKALDKILA
AKAPRLVDIGHYTEKFRNESAFFTNNIGIGFDAQIVYLTNHSPLKRWLNKYHLGFLSYAA
LFFKALKKQPTFKLEVTNPKGHQVFEQAFLCTTTNHPYFGGGVNLMPTAKINDGKLDL
VVIEKKRTWPFLRLFLLMILPGHFHIFSRRVHHLQAPKLKLAVSSAEHGQADGEEMGTE
PFEISFSINQQAFLY
>Acyltransferase
MNRRLKHSRYITGFDGIRTLAVIGVICYHILPATMPGGFLGVPIFFVVSGYLITDLLLQEW
DEYGKIDVKAFYLRRIRRLYPALVTMVLTTAAYITVFQRNLLTNLKGIVISNFLYVYNWFQ
IAHHESYFDKFGGESPFTHLWSLSIEGQFYFLWPLLIILALTFIKKRVKIFDILFVLAGVSA
LAMAFIYHDGIDPSRLYYGTDTRTFSILLGASLAFIWPSTALKRQLDRRLSLMVDVVGLI
SLAIILTCFFKMDSQSATIYRGGMFLFSLVAMIFIATVAHPGADMNRLFTNKVFTWLGKR
SYGIYLYQYPVMIFYESKIAVGNHPLIHALIEILLIVVISALSYRYIERPLQRYNYRRLGVDL
KGLLRKQSQEKWLLVVPVIVIVTAVVGMFLPAKDPGSEQSKQLQETIAKNTKQAAQKN
KQIGTTKKATKDSKNTNEAATSQSPKDFQLTTGLTKQQLAKAQTLQITAVGDSVLADTS
AKLTEIFPQMYVDGKVGRQVYDTIPVLKSLAANGKLANVVLLVEGTNGSFNDNQLAEIM
AILGDRQVYWVNVYVPTKSWQNTVNNALTKATKKYHNLTIIDWYNYSRNHSEWFYND
QVHPNVKGQPYYANYVARKILK
>hypothetical protein

MLVKDFLTATKGIDYSRELFWQAPTKQLIPLAEIIVNSDSVLLIADATLQLPALTLLSVHN
DLVQISQSLPILLRTATKPKLHFYGYHLQNNQLIIR
>cell surface hydrolase (putative)
MRRRKNIASLLFSLLLILALAWPAYFWTKDNVANLTKWHNSKLSPIIMIPGSSATENRFD
SLVSLINKQSFKKHSLLKVKVWNSGKITYQGKINPKDNEPFIVVGFENNHDGYDNILKQ
AKMFNKVFTALTNRYDFNNFKGIGHSNGGLVYTAFLEKYFNEDDAQIKKLMTIGSPFNF
AEKSTANKTQMLASFIKNKVNLPSNLIVYSVAGTQNYESDGLVPLDSVEAGKYVYQNQ
VKSFTEITVTGVDAQHSSLPQNKQIVSLITEYILDKHTLDGKGKLNPSPNQSDNRN
>PTS system, cellobiose-specific IIC component
MVTKLVAWMAQKKEQPTIKAVRRTLVMIFPFVLLGSFMQILAQSVFSQTGFLNSIFQLS
DKVPAFEKIYYFFTNLSFFTIGISALLAAYQVANYQAKYLHKDGAMAGLTGLVCFLILVIR
PLDGTGRLNFNWALLGTPGLLLGLLVGCLVGNLFNLLGKKPARQVERTDDVLARAFAA
LPVITLALLLCLLFNLALAFFNGQDAVTNFFAGMEAFADNQDRVGTVVGLGIISNLLNWL
GLSGPYNTIGTVYGDQKYLDNLNYALSHGSAVNVPYKYTVTTLYQSFGMFGGVGATLA
LLIAISLVSNSHRALRVSRWSFLPVFFNVNSPLMIGLPLLFNPLYLLPFVLAPCFNMLIAS
LAIKWQLITVLTYPVPVGTPGVLVAYIGTNASISSLVLGIVLLGIDVLLYIPFVKVADAVSC
KLTGGI
>Poly(glycerophosphate chain) D-alanine transfer protein DltD
MSLKKRLWLILGPAVIAIVALLLFLVTPYKLKVNQEVLPGAAISQSANIFKGNAIKQAALSE
NYVAFMGSSELSRMDPFHPAVLAAKYKRPYRPFLLGAAGSQSLSHFFAMQGINSEISH
KKVVFVISPQWFTKQGIDKNAFSYYYSNLQAVTWLKTAKPTVMNRYAAKRLLMMPSVH
SDERIAHAVDTIAQGQALSKGQLTYVKLKYNELSREDELFSSLHMNLRTQKLAKAAKQL
PAEYSVEHLDKLAMEIGEKKTGNNPFRISDRFWNKRLRDNYRQLKGRQAKFNYLASPE
YSDFELVLNQFALDKTDVLFVIPPVNRQWSQYTGLSLKMLAQFDNKIKYQLQSQGFNNI
VDLSKDGGEKYFMQDTIHLGWRGWLKLDQAVKPFLTKKQPAPSYRINDYFYSQKWQN
LQGNY
>D-alanine--poly(phosphoribitol) ligase subunit 2 (EC 6.1.1.13)
MKDTVLDILVDLTGSEEVKQDLDINLFDNGLLDSMGTVQMLLELDNQCGIQVPVSEFER
SEWDTPNKIIAKVEAMK
>D-alanyl transfer protein DltB
MPNLQPYANPVYFVYLLLGLAPIVVGLLLGKRWRLYETAISLVFLFLIFDADKWQQGVAL
IIYVLFEVGLIFAYLKYRRKHNQTWVFVLTVCLAILPLVIVKVTPAIMAGRNSLLGFLGISY
LTFKVVGIIMETRDGIIKELKLLEVLHFILFFPTISSGPIDRYRRFIKDYRQIPSQSKYLQLL
EKAVWYLFLGFLYKFLLAYYFGSVLEPRVAHLALIHGGFSKYLVAYTYVYSMDLFFDFA
GYSFFAVATSYLMGIETPMNFNRPFISKNIKEFWNRWHMTLSFWFRDYIYMRLVFFLM
KKKLVKSKIALANLGYLTLFLIMGFWHGLTWYYIVYGLFHASAIIICDMWLRFKKKHQAHI
PSNWLTKYIGIFVTFNVVCFSFLIFSGFLDKLFFNN
>D-alanine--poly(phosphoribitol) ligase subunit 1 (EC 6.1.1.13)
MNLIEQIDLWAQNEPTRLAYDYLGQEYTYQQLKQASDAVAHKLLTMNLSAGPIMVYGG
QEFATLATFLGCVKAGHAYIPVDSHSPKERLEIINQIAQPVACVKAEEVELTVEIPVIEPT
ELEVILKQMVAPIAKEHYVSGEENFYIIFTSGTTGLPKGVQISHDNLMSFVKWMQLDFG
LESGATFLAQAPYSFDLSVMSIYPSLSMGGKLAVLPRQITDNFKELFAALPKLSLDAWV
STPSFMDICLMQPTFDQEHLPALKTFLFCGEELTHKTAAMLKQRFPAAKIYNTYGPTEA
TVAISAVEISREILDNYGRLPIGYVKRESEIVLVDNDLQPVEVGQPGEILLLGPGISKGYL
NNPEKTAHAFVEYQGRAAYRTGDLGQFDDQGLLFYKGRLDFQVKMHGFRIELEEVDH

HLDKVSLINQATTVPKYDKDHKVNQLIAYVVAEKNDYPTELALTQAIKKELAETTLNYMI
PQRFKYVDSLPLTMNGKIDRKSLMQEVNS
>hypothetical protein
MRELKKLWHNSVFNFIMRTIFYFVILLGLVYLYSYSGINQPHFIYNEF
>Na+-H+ antiporter
MDFSLVLVTLAAFATPMILSRFKIGILPTSVAEIIVGIILGKSALDIVHTGPLLSYASTFGVLL
LMFLSGMEIDFSLFKKSAGPQTPLAKKKANQEPTVSPLKAAVGAYILATVTAVILALVFH
VTGLFSDIGLATILFMTVSLGIVISLLKENELLSRPFGQAILLFAVLGEVVPMLLLTLYSSL
VAGDGASIWLISILFIVAAFLFRKFRSFFTFFDTINKSTTQLDIRLAIFVIIMLVWMAEKVGA
ENILGAFVAGIVFKLLQPAEDTEQRLDSIGYGFLIPFFFIMTGVKLDIPALLTSPTTLVLIPL
FFLAYVLAKLPAYLALRKYFSSKNALAGSFLASTTITLVLATLTVAEELKVVTAQQSGAFL
LAGLLTCIISPLLFNKLYRPEPTDKKKTKVHFIGANLVTISAAQQLTQSWYEVDLYTDSQ
KNYATYDSEANVHLLTRMDKATAIASGAFDTDILVLGHVNFDINYQLALAAKEYGVERVI
ARLDNPDPVINSQQEAECEKQGIEYYSSFDTRVGMMRSMIESPSIMQILTSGDARLYEL
TVNNAKFDGLELYKLPFIEEITISRIFRNHKSIAPHGSTRIQLGDHIVFSGSRELVSQVRA
TLEKLNE
>FIG002344: Hydrolase (HAD superfamily)
MSSIIVIPAVAIITLLLGFAGGIWYRKRHYEEKLDQATQTAAGILAQARKEADAEKKEVILE
AKEESHQYRSEVESELKERRSEIQRQENRLLQREELLDRKDNGLEKRENVISKREKQL
DVEQQKIVAKRNKVDLLVEQQQTKLEEVAELTQDQAKELIIKDTREKLSHDLAVMVKDS
EAEAQAKADRIAKDLISQAIQRSAADMVSETTVSVVNLPNEEMKGRIIGREGRNIRTLET
LTGIDLIIDDTPEAVVLSGFDPVRREIAKMTLEKLIQDGRIHPARIEEMVEKSTKEMDAKI
REKGETAIFDLGLHSMHPDLIKTVGRLNYRTSYGQNVLDHSVEVAKIAGVLAAELGEDV
TMAKRAGLLHDIGKAVDHEVDGSHVEIGVELTKKYQESKTVINTIASHHGDVEPESVIAV
LVAAADAISAARPGARSESLENYIHRLEKLENISNEFNGVQKSYAIQAGREVRVIVKPNE
ISDLQATLLAREIKEKIENELEYPGHIKVTVIRETRAIEYAK
>RecA protein
MVADERKAALDVALKKIEKNFGKGSIMRMGEKADTQISTVSSGSLALDEALGVGGYPR
GRIVEIYGPESSGKTTVALHAVAEVQKNGGTAAYIDAENALDPAYATALGVDIDQLLLS
QPDTGEQGLEIADALVSSGAIDIVVVDSVAALVPRAEIEGEMGDSHVGLQARLMSQAL
RKLSGTINKTKTIALFINQIREKVGIMFGNPETTPGGRALKFYSTIRLEVRRAEQIKDGSEI
IGNRVKIKVVKNKVAPPFRIAEVDIMYGEGISKTGELIDMAVEKDIINKSGSWFSYGEERI
GQGRENAKKYLLEHAEMMNEIYLKVRKAYGIADADEEAQLELKLGDKQDKPDAKEKTA
KDKK
>Competence/damage-inducible protein CinA
MHAEIIAVGTEILLGQIVDTNTRLVGQVLADLGIDVYYQTVVGDNEVRMRAAIDLAAKRS
DLVILTGGLGPTKDDLTKQVVAAYLGKQLVEDEAAMLKIKRHFEISQRKMTANNRLQAL
YIEGSKPLANETGLAVGDFYQAEQGPDFMLLPGPPSELRPMLFKVALPLLKQAYRQDQ
ILSSRVLRFFGIGESQLASQLDDLIENQTNPTIAPYAKDNEVTLRLTASAKNEQEAQALL
DGLESKIAKRCGQYLYGYGDDNSLAQVVMTKLREKHLTITAAESLTGGQLQAALTSIPG
ASQAFMGGFVTYANYAKEKLLAIPAEVIDKHGSVSEQTAILMAEQAKQKLGADVGVSLT
GVAGPDSLEGQPVGTVWIGIAYRNKAGYAQKFHFPRQRKYVQARAVLTALDLVRKELL
SEN
>CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase (EC 2.7.8.5)
MNLPNKLTIMRIILIPIFMLVLLLPLNWGQVNALGTPIPVTQLLGAIIFAVASITDFADGQIA
RRHHLVSNFGKFADPLADKMLVMTAFIILVEMGKVPSWVAAIIVCRELAVTGLRLIVVEN

DGQVMAAAMPGKIKTFTQMFAIIFLFLNNVLFANIGFPIGQVLLYICLFFTVYSGCDYFYN
ARFVFADSFKSGK
>Transcriptional regulator in cluster with unspecified monosaccharide ABC transport system
MSIGQTLKDARSDKGLTLDDLQQNTKIQKRYLIAIEEDNFAALPGEFYVRAFVKQYAEA
VDLDSKELLLQLAEQTGKDKEATPEEVVDSRTKKRKLTESNKAPSKFETLYTYLPTIIIVV
VVVAIVGSIYLVTLNNKKADAQIDSSSSKVSVSSDVSSTTKKASSTKKASSEKSVSSEKK
TEVATKKAKSTKKSSKKTSSKQKITSTQVDASTFTYELTGSKAKTTIKLGADTTAWMAV
SANGSQQWQGTLSSGASHTVTLPTGTESITFNLGNSKATTITINGQKFNFLKEDATTTV
RTIVIDIKN
>3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)
MKWALILGASGDIGQNIAYDLAKKGWSLYLHYYQNETKLMNLQAKLQADYPKQDFLSI
HYDLCDSHHLEQITTNIFSLDAVIFTQGTTYYGLFHDLKAEQLATLWQMQVQTPLLLLQ
KLEEKLARSGQGRIVFIGSVYGGAGSAMEVAYSTVKGALSAFALAYSKEVASLGITVNV
VAPGAVDTQMNQVFSKADKDSVSAEIPTGRFAAPAEISYWVSALLAQKAGYLTGQTIY
VSGGWLK
>FIG009210: peptidase, M16 family
MEKIEYQQFGETLYRVTLQNGLQVNLLPKNDFHKTYGVMTANYGSLDNTFVPAKGTEF
VKVPDGIAHFLEHKLFEKEDYDAFELFGKYGAEANAFTSFTKTSYLFSATNNVKKCVEIL
LDFVQQPYFSEQTVAKEKGIIAEEIKMYDDDPSWRLYFGMIGNLYPNTALSIDIAGSVES
IGQITTEDLYTCYQTFYQPQNMSLFLVGNFDIYEILATIEANQAQKQFPAPQEIKRPKLEV
DSSGSDIIPYRMLELNVNRPKTMVGIKGLDEVSTTKAGLKYKLAVELLLYLLYSESAPEY
LKLYQEGILDDSFGYEFTFERGFHFAVISGDSDQPQALSNAIIEIAENAVGSLGNRQAEF
ELAQREFIGRYIQAMNSLEAIANRYSADLFEGATVFELVPVIEELTLEDIIAAAKSFIKPEAI
SVFQILPKG
>FIG001621: Zinc protease
MQVQIVPGVTLDFIASKQFKTSRINLTFVTAAVSKKTVALRTLIANMLEVSSQKYPNQKA
ISDQLAFLYGATFGTSVNRRGNLHLVNFEMRVVNDHYLKEKQQLLTEAINFLQELIFNPL
VTNQAFDQAMFTLQQKNLIAYLESIKDNKQAYALQKLQQAYFEDPVHQIPPYGDKENL
AALTAAECYAYYQEMLAHDEVIITLSGDFASDEVLAAIAQLKFTPRTVGKYQLTYKQTAR
NKLVAYEEQQDLNQSKLDLAYHFPVEYRGKYHYAALVFNALFGGSALSKLFTNVREKA
SLAYYANSSFDSLRQVLFVQTGIQADKKQQVLDLIEQQLAALVAGDVEAQLLANIKQELI
TDYEIRQDSQATALIQATMDQLSQSKVTASEWKAAIMAVSVADVQAVAALAKLEVSYFL
KGQIN
>Cell division protein FtsK
MAKARRKPHPRRKKNKKVKFKIELSLQVSGLCFIFLGIIGLFNLGYVGTLMGNFIRIFVG
DSYQLGLGLLIILGAYLVIFNKQPKLKYNLWLGASLVYGAALLALTAYFFGKLDIHAHFIS
VTWQYLQADIIDASLGSSVGGGILGAILYSLTYFLIAQLGTYILSGLVGVLGLCLLFKVSFT
QVVATVGKLGYLILQGLGIVGNFGKELVLKAKPSLEKATLPTLKAKQPDQEQLKESKIVA
AELEKPVATKPKPVVADEVEEVSVWDEAPTPSVKEESLSEFTVENRVNPNYKLPDYSL
LSEVGATDQSSERKAMQKNKKILEETLASFGVDAKVVNAILGPAVTKYELHPAIGVKVS
KIVNLSDDLALALAAKDLRIEAPIPGKSLVGIEVPNQNVATVSFKSIISAQKRNPKKPLEV
PLGRDISGNLVTADLAKMPHLLIAGSTGSGKSVAINGIITSLLMNCTPDQVRLVLVDPKK
VELGVYNDIAHLLIPVVTNPKKAAGALQKLVSEMERRYECFADTGQRNITGYNEMIQRQ
KADGVESDSQPMPYIVIIVDELSDLMMTASNEVEDAIVRLAQMARAAGIHMILATQRPS
VDVITGLIKANVPSRMAFAVSSGTDSRTIIDQYGAEKLLGRGDMLFAPMGSNKPTRVQ

GAYISDQDVERVVEFVKQQQEVSYDTSLEVSDDELAEGVNGKDEQDELFQQVLEFISQ
EQKCSISLLQRRFKIGYNRAARIVDQLAAAGMIGPQEGSKPRKVFLPKQEDN
>tRNA (cytidine(34)-2'-O)-methyltransferase (EC 2.1.1.207)
MYFINVAVRRAHTEWLRFFYAFRKELGLMTNHIALFEPLMPANTGNIARTCAGTNTHLH
LIEPLGFSTDDKHLKRAGLDYWDKVEVTYHKNLPAFLETIPDLSKLYIVSKFANKDYTDV
DYQDTEADHYFLFGKETTGLPEPFMRQNPEKCIRIPQNDANIRALNLSNTAAIVIYEVLR
QQGFPKLERTHTYENDKLK
>Transcriptional regulator
MKKMAKKAKLKKTNDERILDQHHIEYVETHFDWINKGKDALAEAATMGVDPKSILKTIVL
QANKDPKDYLVVCLPLEYELDLKLVAQQLNKKQVHLADNKNLINITGYVHGENTPIGIKI
RKGFPIYFDERIKVFEKIAVSAGKIGRSVLLKQTDLVKLVAGEYIRVAE
>GMP reductase (EC 1.7.1.7)
MPPVFDYEDIQLIPNKCIVKSRSEVDTRVKFGPMTFNIPVVPANMQTIIDEKLAIWLAQN
GYFYIMHRFDEDDRLDFVKRMHSLGLFASISVGVKPKEYQLIDDLAAQNLVPEYITIDIA
HGHAQTVIDMIKYIKQKLSGVFLIAGNVGTPEGVRELENAGADATKVGIGPGKACITKIK
TGFGTGGWQLAAVRLCAKAASKPIIADGGIRTNGDIAKSVRFGASMCMIGSLFAGHDE
TPGEIIEEDGKKYKVYFGSASQYQKGEYKNVEGKKLLLPYKGHIADTLREMHEDLQSSI
SYAGGKTLAALRKVDYVIVKNSIFNGDRY
>Proton/glutamate symport protein @ Sodium/glutamate symport protein
MKRIKSWRLSLGWQILLGLAIGIILGACFYENIKFITFATNLGDAFINLISMVVLPIVMSSLI
VGIANMGDLRKLGRIGVKTLIYFEILSTIAFIIGMVISNLAHLGSMVDLNQLTKTDISTYLKT
AKTTHANLGSVLMSIIPTNIFASLGNNDMLPVVFFSSIFGLGLASIGEKGQVVIDFLNAVS
ATMFKITGWVMHVAPFGVAGLIGATVAKLGLASLKPLSLFILMAYATMIFFILVILGITSRL
FGFKITEQLLVVKDELILAFTTASSEVTLPRLMDKSVKLGVDQSIGAFVIPTGYTFNLDGS
AIYQGLAVIFVAQAYHINLSLSQQLTLLFVLMLTSKGMAGTPGASFVVLLATVGTVGIPT
SGIALIAGIDRLVDMGRTVVNVIGNMTATLVIGKSENEFNQETHDRYVASFKK
>PTS system, IIA component
MIKSKITYIGPKAVDEKDQMLILFDASAGKELRAVSVIQEFEGPKDYDLAVGKQVVIGSK
AYKITYLGDMVASNLLEIGHTVLQFKAVPDSPQHNAVYLEPQELPDLELGMEITFG
>6-phosphogluconolactonase (EC 3.1.1.31)
MQEPVLFGTYTKKSSHGIYRAFLDTETKTLSKPELFLEIQNPTYLTLSPQKTLLTISSQET
LGGISSYNDTNQALYNEVLFDGPNPCYVSLDEDRQLVYSANYHKGEILSYQLAADQTL
TLADRVVKTGSGPLPEQTAAHMHYSDLTPDKRLVAIDLGADTVYTFDVAPTGKLTEVA
ALKVPAGFGPRHLVFAPNKQYAYLVGELSSQVATLRYHDDGHFDYLGALKTIPADWTE
HNGAAAIRCSQDGKFVYVSNRGHNSLAVFTVQADHSLTLSQLISTAGEFPRDFALDPS
EHFVVVANQNTDNASLYERDTTTGQLTLLQADIALPEGVCVCFE
>Mobile element protein
MDQIRADQAELPKAQRYKIGDLLSRINLAKATYHDERKRIANPDDKYCEVKKVIRRICEQ
GKVRGRLTYGYRRVQARLDQLGIHLAGATIRHLMSELGVQVEIYNHHRNGKYSSYHG
RVGSVADNVLQQQFNAKDPYQVIHTDVTQVRLANQQWAYISAMTDEASKEVLAFQMS
SHPNQELIMNTLNELLDHLPDDVQPIIHSDQGWHYQLKYYTKKLKDHHFIQSMSRKGN
CHDNAPIESFFHLYKTELLQGLPPCNNITELRKLSLTYIDQFNNERISLKTKGMTPVGYR
NHALAV
>hypothetical protein
MTKYSPEFKAQVVVEYMTDGLSSYTLEKKYKVSARQIRKWIQQYQLQGIKSFKQHHRK
RIFSAKFKWHVINYYQTHAESFAQVAAKFDLLPSQISIWRQAFIKYGYDALKPHRKGRS

QKVKKQTKKQKQKLVEKKEIDQLREELARTKQELYNTKMDRDILKKSLALFGPSKPGK
KRK
>Substrate-specific component QueT (COG4708) of predicted queuosine-regulated ECF transporter
MKKLTLSDLVKAGVVAAIYVALTLVNPLSYGAVQFRISEFLNNLAPFNKRYVWALTIGCV
IANLNSPLGVYDVIFGSLGTLVMTGLSYFLAKKVKSVPAKLAITVIICTLMSWSVALELYL
LTDAPFWFTYLTVALGEFGACLLGAVIVWILNTRIDLSK
>hypothetical protein
MTFEEVLPQLKAGKKAVRTKGWSGFEEYIFVVSADTYQGQAVNPYLMIKTKEEPSLSQ
YMPTSCDVLADDWELVND
>3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)
MINSYPDLTKKVVIITGVASGIGYAQAKAFLTNGCQVFGLDKQTTAKITELTATFGSNFS
FSQVDVKEEPQLEKFVKLVQAKCSRIDILLNTAGILDGYRPSLETSLAFFDEVIATNLRSV
FILTNLVLPQMLSQKHGVIINMASIAGLVAGGGGAAYTTVKHGIIGYTKQLTYDYASAGIR
ANCIAPGAINTPMNAADFAGDGAMAKKVAALTPAKRWATPEEVANLSLFLASQSADYI
HGSVLPIDGGWLVK
>Magnesium transporter
MVETNMTTHDREIEIKALLTKQKATAFREAYLDLHVYEQAQIFVDLDKTQRARLYRYLT
PEEVGDMFNAIEEEPQDVVGYFKEMPLKYAGEVIYEMYTDNAVDILAYADPQDLNEYL
HHVPQEAANEIREMLHYEDKTAGAIMSTEYVAITANQTVRSALQVVKKEALDAETIYYIY
VIEDENKLQGVLTLRDLLTNDDDVLISDLMNTPVMSVRVDEDQAEVAQTIRDYNFLALP
VTDFDDTLIGIINVDDIIDVIDEESAEDYSGFAGVDVENTPANPFRAAWSRLPWSLGLLIL
GLATAVVIEHYQALIQDDSTLIIFIAMIMGAAGNAGTQSLAVAARRLTNKEDESESFGKLI
LAEVWAGLLVGLASGLVFALIAGWWQHNYLLGLVLGLALMLSIATANLLGCLVPWGLD
RIGIDSSVVPGALIATVSDVSSVILYFGLVQVFLKWMV
>Similar to ribosomal large subunit pseudouridine synthase D, Bacillus subtilis YjbO type
MTSMEIKWTYQEQAPIKLKTFLKQKGLSRRLLATIRNDNGQVEVNGKSGRIVDKLYPND
QVWLRLPVEKNRKQSLAKSYVPLKIIYEDRDFLVIDKPARVATIPSRLHPVDSLVNRVAG
YYDLRGYQGIIPHVATRLDRDTSGLVLFAKHRLAHALLDQQLRQHEIDKRYWALLTGPV
TWTKLEVNQPIARLATSTMKRGVAADGKAALTDFEVIHRLGKATLCQVKLHTGRTHQIR
VHSQYLGHPLVGDELYGGNIELPLQRQALHCQQLSFYHPFLEKQLRFVSELPADMQG
YIEKNR
>NAD kinase (EC 2.7.1.23)
MRVAIFANMGRKTQAVKQKLTRKFQERHFIIDEVTPEVVVTIGGDGTLLSAFYHYRQQL
ETIRFVGVHTGHLGFYTDWRDDEIDDLVISLQSDNRQSVSYPLLDVLVHYNNQDRPQR
FLALNESTLKRASSTMVTDVYIGGELFERFRGDGLCISTPTGSTGYNKSVGGAVVHPD
MEVMQIAEISSINNKVFRTLSSPMIIASNDWITLKPMASRENKQDLLLTVDSNSYHDHNI
RQIRYKIAKERIKFAKYRHTHFWHRVQEAFIGDEYGD
>GTP pyrophosphokinase (EC 2.7.6.5)
MIEDWQQFLEPYEQTVAELKLKLRQLRNQYRQRGQHSPIEFVTGRVKPIASIVEKMQR
RNIRLDRLEEDMEDIAGVRVMCQFVEDIYLVVELLKQRYDMQVLEERDYIANQKASGY
RSYHIVISYPIQRLDGQRVVKAEIQIRTLAMNFWATIEHSLNYKYQGAFPPELAERLRRS
AEAAFMLDKEMSDIREEIKEAQVLFQTKQKPKVEDDK
>Putative uncharacterized protein
MLEMYLFINPLGAVCYQAEQDILTLVARSDKQINFRFIPLLTLKTVTSVMKTRHIPLGDLE
QRNKLAKELYQAALDCKAALFQGKKKGRAFLLYIQQAFLNSPTSYNEALALSAAKACKL

DVDMFLEDRQSEFTEHQFQQDQQLAAKMNVTDHPTLVLYNLEGFDCGVAVTGCDSF
DVLESIIHGNIPRELILEHQTRQRLHPLKNQKKSSLPPC
>Oligoendopeptidase F (EC 3.4.24.-)
MTEVKKLPQRKEVPVELTWDLELIFANSGEFDTAVTELEQQATTFLTYQGQVGQSAQT
LLTALEAWLDLDRALEKIYVYASMKNDQDTKNAEYQALFTRVQKLVAEVSAKIAWFEPE
LLALSEETVATYLQVEPKLQQYQHLLEVLLAKKAHILSDKEEALLAGASDILGAPENIFGV
LNNADFKFAQVSDEAGNKVQLSQGLYGKLLESVNREVRKEAFEKLYLVYHQFRHTLAT
TLATHVRMHNFEAKLRGYNSAQERALASNSIPVKVYTTLLEQVEQHLPLLHRYVELRKK
VLGLTELHMYDLYTPLTGKANLAYTYEEAKQESLKALAVLGDDYLEHVKEAYASRWIDV
VENEGKRSGAYSSGVYDTAPYILLNWQDNLDNLFTLVHEMGHSMHTYLTTHNQPYQY
GDYPIFLAEIASTTNENLLTQYLLETQKDPQVQAYVLNHYLDGFKGTVFRQAQFAQFEL
WLHEQEASDQALTAEKLADYYGKLNQHYYGSAVVNDQQIAYEWARIPHFYYNFYVYQ
YATGFAAATTLSQNILSKDPAKLTAYLDYLQAGSSNYPLEVMKQAGVNMEESAYLQEA
FKVFEKRLNQLEGLLLK
>Competence protein CoiA
MPIKLKRTIAIIVPGCHEKLILKRGLKKVAHFAHPRHGCQLFSEGESDEHLEGKKFIYHEL
KQAGWQVQYEAYQSELHQRPDILGQKAGCGLVAFEYQCAPISFQKLVYRSRGYKQG
RLAYIWILGKRYFLKGRLTQQNAMFMHWAPNLGFYLLFLDIARKRFEIDYAIQQADFLPL
KYLRHYAYSFQELVAFLHARHEICYYPLTAKQIMTQEYKFQQQIYYGSKSLASLRAYCY
TKGLTLADLPKQIFIDSYHPPLYRKPLLVLRSFQLLDYQDTQAVLKDSLYLLPFVRIPPHL
ASGGLQLT
>Negative regulator of genetic competence MecA
MEMEKINENTIKVSLENDDLTERGITVLDLLGNQKEIESFFYSILDEVDTEHEFRETDAVT
FQLMPNLNGLELFITKMDPDSDEEPFSPLKELEKKGKRVQQVDINDLGDADDITDFIRK
QLGAALEKSESKMLTANDSETKRLQAKSDYLDDQSNMQKRYLVSFANFEDFIQLVKEL
DVDGVASNLYKYDDKYYLELIFFVDQVNENEIKDILAIAYEYGEVVSFSAAVIQERARLIM
KRNALEIARGYFG
>Regulatory protein spx
MVTLYTSPSCTSCRKARAWLKEHDIPFTERNIFSEPLTVNEIKQILQMTEDGTEEIISKRS
KIYQDLHVEIDELPLKDLFELIQNNPGILKRPIIMDDKRLQVGYNEDEIRRFLPREVRALE
LRKAQELAGFY
>D-alanyl-D-alanine dipeptidase( EC:3.4.13.- )
MQTQLSDFVNLQTLTEDLIIELAYATPNNFTGQVVYDFTTAIARRGSAQKLAEASKILKK
QGYRIKVWDAYRPVSAQLKLFEVYPDPTFVAKPNPNFSHQKGVTFDVTLTDLAGKELV
MPTAFDDFSEKAKRSSKHLWSPEALANYNLLNDAMTQAGFIGYPNEWWDFRDSQMD
EYGPLSANPNDY
>Cardiolipin synthetase (EC 2.7.8.-)
MKEHEIQLLTKQKNGLGRILFSRLGLVVLLLLLQLVWIFSLFKWFSQEFPHYTVAVVVFD
FIMMLYLVNSKIDPSAKTTWLLLIAVAPLFGSVLLLYTQSDIGHRALRKRIRQLVAESKHS
LTEDNLALKKLKLEDADTYGLAKYLQRTNDNFPVYQDTKVTYFPIGEAKFKEMLHQLRQ
AKRFIFLEYFIIAEGKMWGEILAILAQKVKEGVEVRVMYDGMCEFTTLSVDYPQRLAKLG
IKAKMFSPIHPFVSTYYNYRDHRKILVIDGQVAFTGGVNLADEYINELERFGHWKDTAIM
LEGKAVKTFTQLFLQMWNITEKEVDFSKYLEVESKVPAKTSGYVIPYADFPLEEEKVAE
NVYMDILNRAHNYVHIMTPYLILDGKMERALTFAAERGVDVELILPGIPDKPIPYALAKTH
FKQLLQAGVKIYLYTPGFVHAKVFVADDKKAVVGTINLDYRSLYHHFECAAFMYKTDCI
PAIAADFEETKAKSTPVTLASLQQEPVTTKVVGALTKTIAPLL

>Ribonucleotide reductase of class III (anaerobic), activating protein (EC 1.97.1.4)
MIEVKMKCKDLQARAKAPRLPRNPEPKEWLASNLSQQKIADYKPFNFVDGEGVRCSL
YVSGCKFSCPGCYNKVAQNFNYGYDYTKALEDQIIADLGQSYVQGLTLLGGEPFLNTQ
VCLQLVKRVRAEYGHQKDIWSWSGYTWEELQKESADKLELLSYLDVLVDGRFLEAQK
DLTLQFRGSANQRIIAVPASLKARQPILWQAEF
>Ribonucleotide reductase of class III (anaerobic), large subunit (EC 1.17.4.2)
MEVQTKKVVTGLILPTKFIKCDGEEMPFMLYKLEYLLDVLGLKADQNFLLSQMFQALGI
EEVVTAKQMRAIFRHVLLENGFDDEVAKYDQVYQERQVKWQEATDPKVRLEKLQAND
PRLVHENANKDSRVFNTQRDLTAGTVGKTLGLKMMPEHVAKAHLRGDIHYHDLDYTP
LSPMTNCCLIDFKEMLTNGFKIGNAEVESPHSIQTATAQMAQIIANVASSQYGGCSADR
VDEVLEPFAKKNYQKHLVEAKEYISTEAKQIEFARKRTKKDIYDAMQGLEYEINTLFSSQ
GQTPFTTLGFGLGTSWIAREIQKSILKIRIAGLGREKRTAIFPKLVFTIKAGLNLKPGEPN
YDIKELAVKCATKRMYPDVLMYDKIKEITGSFKAPMGCRSFLQGWRDETGKEVNSGR
MNLGVVTINLPRIALFAKGDQELFWQIFAEKVALAKEALAFKVKRTLDAKPENAPILYQY
GAFGKRLKPNQSVAELFKNERATVSLGYIGLYEVATTFYGPNWEHNQAAHDFTVEIVR
RLAQACKQWSQESGYHYSVYSTPSESLTDRFCRLDLQKFGKVKDVTDKEYYTNSFHY
DVRKHPSPFEKLTFEMDYPRYAAGGFIHYCEYPNLKQNPKALEAVWDFAYDKVGYLG
TNTPIDRCYKCGFEGEFKATARGFECPTCGNHDPKTCDCVKRTCGYLGNPLQRPMVH
GRHVEIASRAKNLSPEMREE
>Hypothetical protein ywlG
MTLDLEKIKTQVQTGLKELLAQANLEAGDIFVLGCSTSEIQGEHIGKDSNIEVGRTVIAAL
LEILRPLQINLAVQGCEHLNRALVVEKAVAKENQLEIVTVYPSLHAGGAGQIAAFEQFTD
PVEVEHITAKAGMDIGDTSIGMHVKFVQIPVRTSVKEVGAAHTTYLRSRPKLIGGKRAK
YEWNPFN
>Ammonium transporter
MLNEANTTFIIISTILVFFMTPGLAFFYGGLVSKKNVVNTMLSVFIICGLAVIIYIAFGYDLS
FTGNLFGIVGKLQHPFLSGLDLSKVTLPKQGLDTGTYLSFQMMFALITPALFVGAIVGR
MRFNYLLAFIVGWSILIYYPLVHMVWSPKGILANLGVLDFAGGTVVHINAGITALLLSLFL
GKRLKKETSHYNLPWVLCGTAILWIGWYGFNAGSALGINGLAVNAFLVTTVAAAFAML
TWMGLDIYLSGKPSLVGVCTGALCGLVGITPAAGYVTSVGAVLIGFLCSLVSFVFVTYV
KPKLSFDDPLDAFGCHGVSGIIGSVLTGAFATKTVNPTISENGLCYGGGWHLLGVQLL
GTLVTIGFVSLMCLLLIPALRLVIKMRVSQEEELIGLDRSEHGERADYLVDEQHTIERYP
HEFRGQLERLTNKYN
>Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49)
MSTQGCFAETEIRKANPIFSQVRTTVETAFYANNMVAVNDVTTAYKLAAANPTTIVTDL
PIIHTTELGLPADAKMLVDNHGAVVGRTAAARHLIGDPKENSEELIKVLRDALYEGSKRK
FYKTQVVVGLDKDFMLKAHLAVPVGFEMNLLSYLLNFQAINEEYQELYANSKQYPEGDI
YLYADPEWQDDKYPNGLVIIDAKHNVAAVLGLRYFGELKKGTLTLAWATAHRNGYTAC
HGGAKAFHFKDKPDKVFAFYGLSGSGKSTLTHAKHGGKYDITVLHDDAFVISRQDGSS
VALEPAYFDKTNDYLPGTRETTYFTTLMNVGVTLNKEGKKVLVTEDLRNGNGRTIKSR
YSSLNRVDKEASPMDAIFWIMKDDSLPPIVRLENPTAAATFGVTLATKRSNAENVVGAA
DKNALVIEPFADPFRAYSLSEDYHDFKELFAKRNVACYIVNTAAFNGIDINKEVTLGVLE
AVVNEQEGWQQFGKLEDLSYMPIEGYPVDFTDQDYLLKLKTRLEIRRDWVKDYNEAH
PTDTLPAEILTELERLLQALA
>Nitroreductase family protein

MSNSVINDLKKRRSIYALGENVKQSPAELATVIKDAIKHSPTAFNSQTVRAVILFGDKSN
KVWDIVEKRLETEVPNEEAFKKTQEKIATFRAGFGTILYFTDTDVVKNLEENFPLYAANF
QDWSEQAIGGAQQAVWVALANEGVGASLQHYNPLIDEMIAKEFNIPANWKLRAQMPF
GSIEAPAGEKEFMADEDRFRVEE
>ISSth1, transposase (orf1), IS3 family
MSRKSKYSVEQKLNILNEAVHSSFKRVAKKYGVNESTILTWYRLYKYQGIDGLRSIRSN
RSYSKEFKLFLVEQYRNSDDSLELFAIKHGLKSKSQLLNWIILYNESKLKAYTPRKRDSI
MPGRKTTFEERLTIIEDLIKHDVNYNWAVNKYHVSYQQVYGWYQKYRKSGNDPQSLR
DRRGKAKPKAEWTELDKLRAENRLLKAEMKRNEMEVAFAKKLIEIRNREAKKSSNTKS
>Mobile element protein
MLEAILKLEEKHKWTLGYLGMTTQLAFENKLSFKAGLKRVTNCMRNHGIKANIRKKKHN
RIKRHKEYINDNLLNGQFDRQNKNEVWVTDTTEVLYGSEQVRKVRVHVVMDLYGRYIL
SYNISATETATSAIEAFKRAFKVEPGVQPMIHTDRGAAYCSRAFNDYLAEQNCIHSMSH
PGHPWENSPIERWWNDFKLVWLAKRARSKSLSELEQSVQQAIKYFNTERAYASKNGL
TAEKFRAQAA
>Inner membrane protein
MAQLVKTKSFWLAILTTLICSLGGLFLSKLPYLSLIGALVISLLLGMFLQVSPKVVAETKG
GIGFISNKFLRLGIILLGFKLNLMTLINAGVKTITLAIVAVTFTICLTYYLCHKFGAEKELALL
TACGCGICGAAAVMGVSPQLEDSDLSAERKRENEVLAVAVVCVMGTFFTLVEIFLKPLL
GLDATQFGTVAGGSLHEIAHAVATGGAGGSVSLQAALLMKLSRVLLLAPVALIVGIWYQ
KTSKVTNQSQSKKKLPIPWFMGGFILASVAGTFLPFNQAMLDFLVQAAYLFLGMAMAA
LGMSVNFRVIAQRGRSVFAAAFISSTALLVLVTIVSKLFF
>Transcriptional regulator, LysR family
MLEQKYLTFIKVAQLGSYTKAAKQLFISQPAVSQQLTSLENELGVKLFTYHNRQLKLTC
EGEKLLTYVQSLKVQTDKFVTNLHHPSSLQTSLTIATTMSLANEITPNIIAKIQANNTFTK
LNCKIANTKLCLELVQAGKVDFALIEGNFPKEKFSSHIITEEKFIPVCSPLLGLNAEQTYR
FSDLLNYPVIYREEGSGSYNIFKTILASRNIAPSDFKFNYQVGSPTAIKALLLKGLGISFIY
QTVAKQELATKQLIPLKIENLNITHPIYLVYSKNSYFAKDYLNILGL
>tRNA-Leu-CAA

>tRNA-Cys-GCA

>tRNA-Gln-TTG

>tRNA-Gln-TTG

>tRNA-His-GTG

>tRNA-Tyr-GTA

>tRNA-Phe-GAA

>tRNA-Asp-GTC

>tRNA-Val-TAC

>tRNA-Glu-TTC

>tRNA-Pseudo-GGA

>tRNA-Asn-GTT

>5S RNA

>Large Subunit Ribosomal RNA; lsuRNA; LSU rRNA

>Small Subunit Ribosomal RNA; ssuRNA; SSU rRNA

>Lipoteichoic acid synthase LtaS Type IIa
MRSALKKVRSGLNTRTGFFCLVLVLFWLKTYLVYQTEFTLGISNGMQKFLLALNPLPFG
LLLFGIALYFRGKKSYWIMMIVDLIESLWLFANILYYREFSDFISMAIIKSSGSVSNNLSLSI
SQIIHTSDFLVFVDVIILTLLLAFKVIKVDYSPAKRKGALLITVVSILLFAVNLSLAEKDRSQ
LLTRTFDNNYIVKYLGLNFYAGYSGYQAYHESVTKSQASSKDMDKILTYLKKHRADAN
MAYYGKAKGKNVIIFHLESFQQFLINYKVNDEEVTPNINKFYADSNTLAFDNFFNQVGQ
GKTSDAETMLENSLFGLPSGSAMTTYGTTNTFEALPAILNQQGYATAAFHGDVGSFW
NRDNTYKSWGYQYFFSENYYQVKSGYKIGYGLKDKIFLKQSTKYLEQLPQPFYAKLITV
TNHYPYEIDKKNQAIEKTTTGDNTVDGYVQTARYLDQAFGEFISYLKKTGLYDKTMIVV
YGDHYGISNNHKKAIAQLLGKKSVTSYDLAMFQKVPFMIHMPGLQGGVNHTYGGEIDV
MPTLMDLLGLKNTDTVQLGNDLLATKRNQTVVFRNGDYVSPTYSKIGSTVYNNKGQVV
TKKLTKAQSKSVAQKEAYADKLLSMSDQIITGDLLRFYQPKWFTKVNKSDYNYKKAEAL
KQLDQAQVKEQTSLFFNNNKKSTESLFVTDAPELK
>hypothetical protein
MENSQLVAIISRLDAMMKAAGDEVQSRRFEKDGQERGFVTYDPASGTFELEELASHQ
KFEFDNIDLTAMEIYDLLDY
>Integral membrane protein LafC, accessory function in glycolipid and LTA synthesis
MTKKNKIVLVIMLMVGLVIFYFSVKQISYHSLMDAIAHLHWSWLVVAFLLMASSLLLEAR
VVQVLIRQQLPTYSFWAALRVPLVEQLFNGITPFATGGQPAQLVALLQSGVDAGRATS
SLLMKFIVYQAMIVVNFIICLALGFQYVADRIHGMIWLLVFGFIIHFAVIVSLLMIMFWYRL
TKKLLKLMMLVLKKVMTNSNRYKQLEQTLNEKVDSFYQESLKLKANKQVLVRISCLTLF
QLLIYYVIPYFILLALGVTNANILLVTTLHVLIVMVISLFPIPGGAGGAEYSFSVVFSTFVQQ
GNRLVVAMFLWRIITYYLGMFSGIIALAFEPKKTSK
>Glycosyltransferase LafB, responsible for the formation of Gal-Glc-DAG
MFSSATKIKGQGVASAYLELIALLKKYCATEIEVSINNLKRADISHYHTVDFRFFCSTFLP
FRGKKIGYVHFLPETLEGSLKLPRLAKVIFYRYLITFYKRMDALVVVNPSFIPKLEAYGIP
AAKITYIPNFVSQTEFYPQAATKTELRQKLGLKDRFTVLGVGQVQARKGVFDFIELAKN
NPQLQFVWVGGFSFGKITDGYAELKKVVANPPSNLIFPGIVERQQMNYYYNAADLFLL
PSYNELFPMSILEAFNCRTPVLLRDLELYHAIIAGYYEPAKDRAEMDIKIKQLAANPDWL
TTLKDKANRAATYYSEKRLAQIWKEYYLAQAKRD
>Glycosyltransferase LafA, responsible for the formation of Glc-DAG
MNIGIFTDTYYPQVSGVATSIKVLREQLERQGHNVYLFTTTDPKVKKEVYERNIFRFTSI
PFISFTDRRVAIRGLLQAYRISSKLQLDIIHTQTEFSIGLMGKLVAKMLKIPCVHTYHTMY
EDYLHYVAKGKVLRPKHVRLMTRSFCQNLSGVIAPSDRVLAKLQEYQVTTPVEVIPTGI

DVSKFQQPEYLNYRAKYGLTATTPVILTLSRVAYEKDIDRLITAMVEVKKRVTGAKLMIC
GDGPARADLEAQAKLLGLKQDVIFTGEIDNDSVAHYYHLANVFVSASVSESQGLTYLE
ALAAGIPVVAANNPYTNGLLNHPSLGATFETPAQLVSEIVTYLNNPPRIADQAARSEKLA
EISAQTFGQKVSEFYKESQAYYLNERQMAKRSSEIG
>Phosphoenolpyruvate-protein phosphotransferase of PTS system (EC 2.7.3.9)
MTKVLKGIAASDGIAAAKAYMLVQPDLSFTETSIDDPEAEVKRLDDALAASKKELELIKA
KATENLGPEEAEVFEAHLTILADPELIGQIKDKINSDKVNAEAALKSVTDMFISMFESMT
DNAYMQERAGDIRDVTKRVMSHLLGVTLPNPALINEEVVIVAHDLTPSDTAQLDRNFVK
GFITDIGGRTSHSAIMSRTLEIAAVVGSGNATSEVKDGDMVIIDGIEGDAIVNPSEDQLA
EYKQKAADFAAQKAEWKLLKDAKSVTADGKEVTLAANIGTPDDVVGANDNGAEAVGL
FRSEFLYMNSAELPTEDEQFEAYKKAIEGMNGKQVVVRTMDIGGDKELPYLPLPKEEN
PFLGYRAIRISLDRQDIFRTQLRALLRASAFGRLAIMFPMIATVNEFKQAKAIFEEEKANL
VAAGVKVSDDIEVGMMMEIPAAAMIADKLAKYADFFSIGTNDLIGYTMAADRGNERVSY
LYQPYNPSILRLVKRIIDASHAEGKWTGMCGEMAGDQVAVPLLAGLGLDEFSMSATSIL
KTRSLMKTLDTTKMQELADKALDCETAEEVTALVEEYTK
>Phosphocarrier protein of PTS system
MAKEQQFHVTAETGIHARPATLLVQAASKFNSDITLTYNDKSVNLKSIMGVMSLGVGQ
NADVTIKAEGADEEEAIAAIAETMSKEGLAE
>hypothetical protein
MDKDYQKLLEQLVSGELAELDVEQKEFQAFQPVLMEFHQRNSIVGRAKRGGGVSYHY
SRDGI
>ATP-dependent Clp protease ATP-binding subunit ClpA
MNLLMLCQICHQNPATIHLTTNVNGQQTTLDLCQKDYQLLKERSSQNGGLSALMQDPF
GFGNLDDIMRSLSGQANGANPSMPGGPRPNPQAGRNGNGGNSILDQFGLDLTNEAK
NGHIDPVIGRDKEIARVIEILNRRTKNNPVLIGEAGVGKTAVVEGLAQKIVDGDVPQKLL
NKKVIRLDVVSLVQGTGIRGQFEQRMQQLMTELKQQQDTILFIDEIHEIVGAGNAEGGM
DAGNVLKPALARGELQLVGATTLNEFRSIEKDAALARRLQPVQVDEPSIPETIKILEGIQ
SKYEAYHHVKYSEDALKAAVTLSARYIQDRYLPDKAIDLLDEAGSKKNLTIQAVDPNLLS
EKIKNAESQKQAALKEEDYEKAAYYRDQVNRFEKMKENSVSDEDAPLVTDQDMEKIVE
EKTKIPVGELKAKEQSQLKNLAASLESHVIGQNQAVDKVARAIRRNRVGFNKSGRPIGS
FLFVGPTGVGKTETAKQLAKELFGTEDSMIRFDMSEYMEKFSVSKLIGSPPGYIGYEEA
GQLTEQVRRHPYSLILLDEVEKAHPDVMHMFLQILDDGRLTDSQGRTVSFKDTIIIMTSN
AGSGDSVVNVGFGAEASGKSNSVIDKLTKYFKPEFLNRFDGIIEFNALSKEDLTKIVRLM
LADVNDMLAPQGLSLHVTAPVEAKLVDLGYDPKMGARPLRRVIQEQLEDRIADYFLDH
PTEKQLVAKVENGEIKITNQTTETSN
>hypothetical protein
MIMTDVTNSYHRFINEELRSSNAHFIKVYSLGNSKVVYKKKFGHAEVVISNKIRPVTQKE
IDFVLKELASSLEVTPTIDNANHNLVQITWDLAS
>hypothetical protein
MLWSRIFSNLTTNFNIRRRQRGRHFATKKHYIPTPKAAPSPKTVTPSTSDATLLIIYKDE
AHQALADPELVVGQLGAKFNLRFKNFKHYRLVKIQGFTSTFVDSYGIITLTYQRITAANI
WVFCIDLDSRRFLTKPHFVNGLLGQSYALASPTIKNYRLLLAKGPIRGTFSDKQAVVTY
YYRNNNIQDITYQPYFLEMLDFITCYDNPLGHLTSVTLAKGTVWKTFQTILLTDGQIWHC
MGGSLWIKETATAFKKITGPTKKSALNYQAAIKLPYSQPINQLATINFVPHKQVTSYERP
FGRPSQQFTHNQQVLITATLQSYNITWYQLNHTSWITSQYLKISAKKGN
>Neopullulanase (EC 3.2.1.135)

MPNIHYNSWIAKYKQPFGAIAKDQAVTFGITVASQTNPQVELLLRKDGEMSATSYQMT
YKEADFYQVTLSLEQAGLYYYYFKIELDENGQRACYFYGADYDGYGGPGRVSAGDNIT
AYQLTCYQTEVTAPKWYQGGVCYQIFPDRFYNGNPNQEITGKKKNSFIYATLEDTPYYI
KDAKGEVVRWDFFGGNLKGIKAKIPYLKELGIRGLYLNPIFLATTNHRYDTNDFMEIDP
MLGTSADFKDLLDELHRNGMYLILDGVFNHVGKDSKYFNARKLYYGADQGAAQDENS
PYYSWFKFTHYPDEYASWWGVKDLPEVDKANPDYQAYIYGAKDSVISKWTNLGVDG
WRLDVADELPEYFIEGIRKNLDGFANEKLLLGEVWEDASNKIAHEARRHYVEGSELDG
VMNYPLRKCILDLLINSHQRPIWSVARELMSLRENYPRTFFFNALNNLGTHDTERILTAL
GNDERKLIAAYGLLFMFPGVPCIYYGDEAGLTGGKDPLNRAFFPWGRENLTIQKLVKK
WVHRHNDSEALQKGELIIGGNDNCFIILRYTETKLAFYVLNLSQQTIDLGQESFAFDHDN
HNLGETVVKQVATTQLQPFEGYFWQGNRS
>Glycogen phosphorylase (EC 2.4.1.1)
MELTKENFVEKFEERLETKYALDVKDATAYELYDTLCSLVKSAYTRKWRQTWKEYVNK
DEKQIYYFSIEFLPGKMLKSNLLNMGWLETVQAGLAELGIDLEDLVKVEPDMALGNGGL
GRLAAAFMDSLASDGFPGNGNGIRYKYGLFKQKFVNGYQVELPSVWLKSGVGNAWE
VRRDSKAVMVRMGGHVHLEQKGDYLEPVYEDGMIIRAVPYDTGMVGYHNQHVNTLR
LWEAEIPESEETKYPTIEARRRVEDLTAVLYPDDSTEDGRRLRLQQEYFFVSAGVQSIV
RHYEKYGKPFTELGKYIGIHINDTHPAMCVAELMRILLDEKHMSWEDAWQQTVQVMSY
TNHTIMAEALEKWSIQMMNEIQPRLLQIIAEIDNRYVKEWAGKVDGSIIERTRIISGGQVH
MAHLAIIGSHSTNGVAKLHTELLKEDVLHDFYTLYPERFNNKTNGITERRWLQLANPSL
ARVLDETLGTNWRSDIKQVNKLLDYADDQAVLDKIGQAKRENKVRLAAVIKEKTGIEVS
PDAIFDVQVKRLHAYKRQLLNLLRIVKVYLDLKDNPNLDMVPRVFIFGAKAAPSYRYAK
SIIKCINEVANLVNNDPSIQGKLKVVFLEDYGVTLAEKIIPAADVSEQISTASKEASGTSN
MKFMLNGALTVATLDGANIEIKDKVGDDNIFIFGLTKDQVYEYYANHSYHGSDYYYNDP
VIHRVLDAFVDGTIPNISNEGHEIFDSLTYLNDDYFVLGDFEDYLRAQAEVDEAYRDRN
RWNKMSLINTANGCFFSSDDTIKKYAAEIWNIKSVTEGKEVGTNAKYPL
>Glycogen synthase, ADP-glucose transglucosylase (EC 2.4.1.21)
MKILFAAAECAPFFKTGGLGDVVGALPKYLAKKGHEVKVVLPYFAGMKEEYKQKLAYL
FNFRVNVGWRSQYCGLMHLEQDGVEYFFVDNEYYFKRPGLYGYYDDGERFAFFQQA
VIEMMEKVNFIPDVLHVNDYHTAFIPFLLREKYGWINAYQQIATVLTIHNLEFQGQYGRD
VLPELFGMDTGRYDDGTVRFGDAVNFMKAGILYADKINTVSPSYAGEIQTPAFGCGLD
GVLRMVNWKLSGILNGIDYETNNPKTDPNIKHTFTKNNLKGKAVDKAELQQKFGLPVR
KDVPLIGIVSRLTYQKGFQLVVSEMENLLQFDVQVVVLGTGYPDLEHDFWYFGDKYPD
KCGVKIDFDVNLAQEIYAGCDMFLMPSAFEPCGLSQMISMRYGTLPIVHQIGGLKDSVE
AYNPIEHTGTGFGFEEYNSFYMMETIKQALGVYTNEPKVWKKLVTNAMKRDSSWQTA
SQSYEALYEEIIRN
>Glycogen biosynthesis protein GlgD, glucose-1-phosphate adenylyltransferase family
MCAIIGNEHEYTSLLPLTKHRPLSTLYFDCKYRIMDFALSSVVNANIRNVFMIMNEGRIK
SVFDHLGGGREWGLDSIGSYQYISFYQDILRRKAEGKTYFDDVIDYLQKSKASYTVFIS
NKILANVDLRAVLKIHQAQNNKITAVFKRVARNNIASDDQILTLGENGTVLGTKKFNDLS
EKDDTYNLAMGIFIANTDWLIKALDAGQNAGAPTSIEEFLVSELTKVKSSAYEYTGYMS
NIYSVKSYYQANMDMLNPEKFNSLLFSSQKVITRTKNEVATYFTEDSEVRRSQFATGC
VIKGKVEDSLISRRTVVENGAEVIGSVLMGSANVGENTSVKYAILDKNVTIDANLKIEGT
PEKPVVIEKDTHVTADVIGG
>Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27)

MKNEMLAMILAGGQGTRLKKLTKNIAKPAVPFGGRYRIIDFTLSNCVNSGVHNIGVITQY
QPLVLNNHIGNGASWGLDRLDSGVTILQPYSSNEGSKWFEGTAHAIYQNIDYIDSMDP
EYLLVLSGDHIYKMDYEDMLQQHKEKNAALTVAVIDVPWDEASRFGIMNTDENNRIIEF
EEKPAEPKSNHASMGIYIFSWPRLRQVLVNGFTKGVPMVDFGKNVIPAYIKSGDNVFA
YNFSGYWKDVGTIDSLWEANMEFIDDKGGLDIRDHSWRIYSKNPIAPPQFITDTAKVNN
SMIVDGCFVAGDIEHSILAANVQVKEGSVVKDSIVMPGASIGKNVKINRAIIGENAVIGDN
AEIDGSDEIAVVGNSEVIGVKVNED
>1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type (EC 2.4.1.18)
MGKFNMNDLKNGLATFTTGENFHAYNTLGAHKATVDGKDGFIFSVWAPNAQNVSLVG
EFNKWQAEGIPLTKDQAGVWSVFTTIPKEGQMYKYNVTQANGRQIFKIDPFAQRFEKR
PGDAAILYDLKDFHWTDGLWRGRQKRSNYFERPLNIYEVHASSWKHHADGSPYTFKD
LTDELVPYLKKMGYSHVEFMPLMEHPLGASWGYQLIGYFALSSYYGTPEDFQYFVNA
CHNENIGVLVDWVPGHFCRNDDTLSYYDGTPTYEYYDPDRANNIGWGALNFDLGKPQ
VQSFLISSALYWIETFHLDGIRVDAVSNMIYLDYDVGPWKPNKDGGNRNLEGWHFLRK
FNSVVKLKHPKTIIMAEESSAGTKVTGMLEEDSLGFDYKWNMGWMNDVLKFFEMDPV
YRKYHLNLLTFSWMYRMSENFILPLSHDEVVHGKKSLMHKMWGDRYRQFAQLRTLYL
YQATHPGKQLLFMGSEWGQFLEWKYDDQLEWVDLQDNLNASMQHYTASLNKFYHD
NRALWETDHTENTLEVIDADNTDETVLTFIRHGKRKKDFLIIAMNLTPVERRNFKIGVPY
PGKYQEVFNTEMKEFGGTWTKHNVDSLSQDESFKQFEQVITTTVPALGAIILKPVDINL
RKKKKRPTAKKEVKKKK
>hypothetical protein
MKSNETLTLIEEITRNDGSRYFEIANMVQNGRAELAAVRGFIKQVRILQLNIPRSTNVIKY
EQYVNENFEMPTEQMDHFEEWTKTPEAQKLVDAILRENHIA
>hypothetical protein
MINTKQFQLSLFLGLVITLIHLLLEFTQLNSGTNLTFDTLLLFFHTKFFLLPILVWFLEYLVN
TSICYLFFRLLTHLIKLFQ
>Phosphoglucosamine mutase (EC 5.4.2.10)
MKYFGTDGVRGIANQSLSPELAFKLGRCGGYVLTQHASGKTPRVLVARDTRISGQML
ESALIAGLLSVGIEVFTLGVMTTPGVAYLVRLQDADAGIMISASHNPASDNGIKFFGSDG
YKLSDDQEEEIEVLLDQASDNLPRPAAEGLGRINDYKEGALKYTQFLEQTIPDDLTGLHI
AVDGANGSTSALVSRLFADLNADFETMATSPDGININAGVGSTHPEALAKFVLEKGAQ
VGIAFDGDGDRCIAVDENGEIVDGDKIMYICGKYLSQRGRLKKDTIVTTVMSNIGLYKA
MEANGLNSVKTKVGDRYVVEEMRKNGYNLGGEQSGHVVFLDFNTTGDGMLTALQLL
NIIKQTGKTLGQLAAEVTTYPQELVNVRVTDKKAALDNEKIKEVIVKVEDKMGDDGRVL
VRPSGTEDLLRVMAEAPTKEEVHDYVMQIVAVVEEEMGASEA
>Uncharacterized secreted protein associated with spyDAC
MKVRQAFNSKLFYRLISLFFALLLSFYVTSEKGNTTRSGNSGTFGSSLVSNKTITLETDL
KLNVNSDKYFVSGYPDVVKVKISGPSALVIAAKNTQNFEIYADLNDLKPGVHTVRLKDS
GLSSELTVKIIPEKITVQISKKKTKKTAVQVRYDSNRIASGYAAGEAYSSQAYVTVTGAA
KDINNLDKVVANLNLPRNTNETYAKNVALQALDESGKVLNVVISPANVTATVPIYTASAS
KKVPINLVARGEGVSNMKYSFSSDTKTVTVHGTKEALSKLNRLDVPVSITGVTADTVEN
VRLSPSQSGITAISPTTITVNVSVGQKETSTAQKQTQASSQAETSSASSTKETVATDDS
SSEATSSSSTSSQLSSSSSSDEVEN
>Diadenylate cyclase spyDAC; Bacterial checkpoint controller DisA with nucleotide-binding domain

MAINWGDLLTLRNFVNLIDICLVWFVIYELMMLLKGTKAVQLFKGVLVIIVIKLISWYVGL
QTLSWIMDQVINWGIIAIIIIFQPEIRRGLEHLGRGSVFTYNKTGNQEEKMVEALDQAIQY
MSKRRIGALITIQQNTGLEDYIETGISLDADVTGALLINIFIPNTPLHDGAVIIKDNKIAVAA
AYLPLSESNLIPKELGTRHRAAVGISEVTDALTIVVSEETGGVTVTKNNELIRDLTQQDY
LKLLRNELVPRQEEVKTNPVTRFIEGMIKGGGHKH
>FIG028593: membrane protein
MLYVYLTVVRHNSFYSFLLLNTFLAYIPIELSFHIKVKQTGLLFYPLLIGWLLFYPNAPYIL
TDLFHLARMTPYNPTTGLMSFNLHMWLNFTNLVSSALAAALIGNWSLQYVADVICLRF
KLHNNFGRLSIVTLLLLLSSVGIYVGRFLRLHTAYLFLNPNWVIEQLLNMWSIRMLAFVC
FMFLLQLIIWLSTYLSYLAFRNMDSKFN
>UDP-N-acetylenolpyruvoylglucosamine reductase (EC 1.1.1.158)
MTELAQEFKVKFPKIKILAAEPLAKYTNTKTGGPAECLAFPKNVAEIKSLLVFAKENQVA
VTVIGNASNLIVKEGGILGLVIILTDMNEIKVEGNFILAQAGASLIKTTQVAYQAGLTGLEF
AAGIPGSIGGAIFMNAGAYDGEIKDVVKSVEVITRDGQLKTYQNNELHFAYRHSRLQEE
DDIVLQATFSLRAGDRYAIRNRMEELNFLRASKQPLEYPSCGSVFKRPVGHFTGKLICD
AGLQGYQSGGAQVSKKHAGFIVNVDHATASDYLDVIHHIQAEIKAKFDVDLETEVRIIGR
DS
>Exonuclease
MDFVAFDFETANAKRNSACSLALSVVRNGQIVDEFYSLINPQAEFAPTNIRIHGIRPADV
KKAPTLAELWPHIQSFFQPNQLAVAHNASFDCSVLKRTLESYDITVPQFLVLDTLKTSR
KFFADFPNQKLDTICQRLNISLENHHNALADASACANILLYQANHFGSSALVDFIKTYPK
>Cell division protein FtsX
MKTLTLKRHLTESFKSLRRNGWMSVAAISAVTVTLLLVGVLLSVLLNINKVAHDVERDV
QVRVVINRGTNKEQIKQIKADLTKIKAVKKIKFSSRKQELKNVIGVYGDDFRLFQGDSNP
LYDVYIVETSSPDQTIKVAKQAKKIKNVYDARYGGVAAKKLFKAVATIQKWGLTISVLLLF
VAVFLISNTIRITILSRKDEIGIMRLVGATNSYIRWPFVLEGAWTGLLGAVIPIILVDFGYN
WAYAVLSKSFEGTAYRLLTPGQFLWEIDLLLAVIGIVIGALGSGLSMSKFLKI
>Cell division transporter, ATP-binding protein FtsE (TC 3.A.5.1.1)
MIEFKNVVKKYDNNITATNNISVKIKPGEFVYVVGPSGAGKSTFIKMMYREEAPTSGSII
VNEYALESIKSREIPFLRRELGIVFQDFKLLPRLTVFENIAYALQVIEKTPEEVEQRVSRV
LDLVGLSAKRDNFPAELSGGEQQRVAIARAIANEPSILIADEPTGNLDPETAEGIMEILER
VNKQGTTVVMATHNSFIVNEYKHRVLEINGGKIIRDEEEGVYSNEDANA
>Peptide chain release factor 2
MAQPGFWDDQESAQKLINQNNELKAKYDNFKELVKLVEDLEVLLEMVAEEPTDQELLA
ELEQTYAAAANKLQAYQLSILLNQPYDKDNAILEIHPGAGGTESQDWGSMLLRMYTRW
ADQHGFKVETVDYQAGDEAGIKSVTLLIEGTNAYGYLRSEKGVHRLVRISPFDAAGRR
HTSFASVNVMPQLDDSVEVELRPEDVKMEVFRASGAGGQHINKTSSAVRLIHIPTGIVV
SSQAQRSQFQNRATAESMLKSKLYQLELEAKEKEKAKLQGEQLDISWGSQIRSYVFHP
YSMIKDHRTGYETGNTGVVMDGDLDPFINAYLQWNIHKASGDNDK
>Protein export cytoplasm protein SecA ATPase RNA helicase (TC 3.A.5.1.1)
MANFLRKWVESDKRELKRLGKIADKVEAYASTMEALSDEELKAKTPELKQRYQNGSSL
DELLPEAFAVAREGARRVLGLYPYRVQIIGGIVLHEGNIAEMKTGEGKTLTATMPVYLN
ALSGEGVHVVTVNEYLSSRDASEMGELYNWLGLKVGLNIASKTPAEKREAYAADITYS
TNSELGFDYLRDNMVVYKEDMVQRPLNFVIVDEVDSILIDEARTPLIISGQATGSTMLYT
RTDRFAKTLTEDEDFKIDLESKTVSLTEEGIKKGEKYFGLTNLYDPDNTALNHHLDNAL
RANFIMLRDKDYVVQDGEVLIVDQFTGRIMDGRRYSDGLHQAIEAKEHVEIEEETKTMA

NITYQNFFRMYKKLSGMTGTAKTEQEEFREIYNMEVVTIPTNRPIAREDRPDLLYPTLK
SKFKAVVEDIKTRHEAGQPILVGTVAVETSELLSHMLDQEGIPHSVLNAKNHAREAEIV
MNAGQRGAVTIATNMAGRGTDIKLGPGVRELGGLAVIGTERHESRRIDNQLRGRSGR
QGDPGVSQFYLSLEDDLMLRFGSERIKAFLERMKVADDDAVIQSRMITKQVESAQKRV
EGNNYDTRKQVLQYDDVMRAQREIIYSQRQQVIMEDKSLKGVIMPMIKRTVEHTVQVH
TQGEKNEWNLQAILDFAVSAMVNEDSIELADLVGKTPDEIVAYLMDRAEANYKLKEKQL
YDASQMLEFEKVVILRVVDSRWTDHIDEMDQLRQSIGLRGYGQLNPLVEYQSDGFRM
FEQMIGDIEYDVTRLFMKAEIRQNIKR
>Ribosomal subunit interface protein
MLKINVRGENIEVTGAIREYVEKKLTKLEKYFENTSNSTAHVNLKVYTDKTYKVEVTIPL
PYLVLRAEETSPDMYASIDLVTDKLERQIRKYKTKVNRKSREKGFASFEFVPQAEEKEE
EDKKFEIVRTKRVSLKPMDSEEAVLQMDMLGHNFFIYEDAETQGINIVYRRKDGRYGLI
EAGDED
>Competence protein F homolog, phosphoribosyltransferase domain; protein YhgH required for utilization of DNA as sole source of carbon and energy
MNKQCLLCQQNLTYHFDLKWFLTLQPLVDKYLCERCATNFIRLDNRLTCPGCGRLSET
GKLCSDCLRWHQINPELLANKALYSYQTEAMKDYLERYKFLGDYRLRYVFQRQWCDF
IVKNYPPKQGWLYVPIPVDQESKSQRGFNQVIGLAELLPLTEVLEMKQTHGRVKQSKK
NRRQRMQTKQPFNYSGSEGEILRKKVVLLDDVYTTGRTLYHARVVLKTHGAALVRSVT
LSR
>ComF operon protein A, DNA transporter ATPase
MDASILYGRQIVLGEKDLEGLNNLQVKPALVVGKRGVICQRCGQTTSLKLARLPTGKYY
CPNCLLLGRLTSADRLYNLPEPHLFEHQKNSLTWQGQLSKYQEKCSAQLQTVVKHGG
KHLLWAVTGAGKTEMLFETINWALQTGKRVGLASPRVDVCNELFPRLQAAFAKTPMIL
LHGHAKESYRYTQLVVATTHQLLRFYQAFDLLIIDEVDAFPFVDDQGLHYAANNALIPG
GTLVYLTATPSRELLAAVKKKELSVSYLPLRFHGYLLPTINLVAVAKLRSQLDEGILNSKI
KKILLRWQEEGYQFLVFVPRIALLKPVYQAIANLLKPELKGQTVYANDGQRIEKVAKMR
AQEYRYLVTTTILERGVTFPGINIMVLAADDSNFSTAALVQIAGRAGRNKARPIGDVYFF
CQDYTRTVKAACKQIKFLNRKGKKLKKDE
>FIG000605: protein co-occurring with transport systems (COG1739)
MENETYLTIKSPGSYKQTIKKSDFICSLARTETEAEAKAFIAQIQAQNKKANHNCFAYTL
GLNDEIQRESDNGEPSGTAGVPILEVLKNLSLHNVTAVVTRYFGGIKLGASGLIRAYSST
TSQAINEIGIVKRVMQRELALQIAYNQQGRLENYLKQNNYHILNTDYGVDVTIKIAVDLE
NVEPTEANLIDLLNGQIQITPGPNFYNEIPYQPHH
>Undecaprenyl-phosphate N-acetylglucosaminyl 1-phosphate transferase (EC 2.7.8.-)
MYKILASLFATMLLAALLTPFIRQLAFKVGAVDNPNARRVNKIPMPTMGGLAIFLSFNLA
NLFLLRNQYPSSQLYSLILAQLIIIATGVIDDIYELKPRQKMFGIFLAALVVFFVAKVRMTT
LTLPFLGVIPLGWLSFPITIIWILAITNAVNLIDGLDGLATGVSIIALTTSAVTGYFFLNVTNT
FVSIMMFTLVAALIGFLPYNFHPARIYLGDTGSLFIGFMMSVFSLYGLKNATFITIIIPVIIM
GVPITDTVYAILRRWLNNKPISQADKHHLHHRLMQMGLSHRQTVLVIYGIALIFSFISLLY
PLSNLWGSVLLTVATLLGLELFVELIGLVGDGRQPLLDGIKKLVLMTSARERRDRDKED
RHK
>transport protein
MWRYLKRLLIGQPLKTAEEGGQSLTKFKALALLSSDALSSVAYGTEQITTVLFALSAAAT
WYAMPAAGIVLILLFAITMSYRQIIHAYPSGGGAYVVATKNWGTGAGLVAGGSLLVDYM
LTVAVSVTAGTEAITSAVPALRDHSVAISVIIVLFIMTLNLRGMRESASFLTVPVYFFVIMI

VGMVVWGLINIATGHLAYHAPAAIGASVPGMTVILFLKAFSAGSSSLTGVEAISNAVPNF
KEPKRKNAANTLAIMSLILATFFAGITFLSYYLGVMPNPEQTVLSQIGALVFGKGIFYYLL
QLATAMILAVAANTGFSAFPMLAYNMAKDKFMPHLFMDKGDRLGYSNGIIFLAVGAIALI
FIFHGQTNLLIPLYAVGVFVPFTLSQSGMIIHWFREKGKNWVLNACFNLLGALISFALVIF
LFWLHFDNVWPYLVIMPFLLFMFYKVHDHYVKIAAQLRLTMDEEAQFHKYDGSTVIVLV
SSVTHVTTGAINYARSIGDYVIAMHVSFDTNPKKEHETAKEFKQAFPDVRFVDLHSSYR
SVTKPVLRFTDVIAKNAANRNYSVTVLVPQFVPKKPWQNILHNQTALQLRRALNLHENV
IVSTYSYHLKK
>Heat shock protein 60 family chaperone GroEL
MAKEIKFSEDARSKMLAGVDKLANTVKSTIGPKGRNVVLEQSYGSPTITNDGVTIAKGIE
LEDHFENMGAKLVAEVASKTNDIAGDGTTTATVLTQAIVNEGMKNVTAGANPVGIRTGI
EMATKAAVEGLHKMSHTVASKSDIAQVASISSASEEVGNLIADAMEKVGNDGVITIEES
KGIDTSLDVVEGMQFDRGYLSQYMVTDTDKMEADLDNPYILLTDKKISNIQEILPLLQEV
VQQGRALLIVADDVDGEALPTLVLNKIRGSFNVVAVKAPGFGDRRKAMLEDIAILTGATL
ITDDLGLQLKDTTLDQLGTAHKVIVTKENTTITEGAGDSAQIKERVEQIKKQIAETTSDFD
REKLQERLAKLAGGVAVIKVGAATETELKERKYRIEDALNATRAAVEEGFVPGGGTALV
NVIPAVASLKAEGDVLTGINIVKRALEEPVRQIAENAGLEGSVIVEKLKEQKAGIGYNAA
SDEWVDMVEAGIVDPTKVTRSALQNAASVSALLLTTEAVVADKPEENPAPQAPAAGM
GGMM
>Heat shock protein 60 family co-chaperone GroES
MLKPLGDRVILEVQKEEEQKVGGIILAANAKEKPQTAKVIAVGEGRVLDNGQKVAPSVK
KGDLVLFDKYAGTEVKYEDATYLVVRENDLVAIVD
>CAAX amino terminal protease family protein
MANSKNALALVLTYFLTLNFANLASYWWAKNPYIYLFLTIINLLGAGLMLGLNHRYPYHN
QFERRPAPWSVSLLYLGLGLLGLLISQKLALGFESYVLHLPSASANTQSLLTITKSYPYY
ALTLVLAAPLMEELVFRKVLFGTVSDFTNKLYPLLGSTLLFVLAHNDGHFLVYAFMGLIL
SGVYYKTGRISISVGCHLGMNLLILLLSQTI
>Redox-sensitive transcriptional regulator (AT-rich DNA-binding protein)
MATENIPNATAKRLPIYYRYLNILADSGKKRVSSTELAEAVKVDSATIRRDFSYFGALGK
RGYGYEVEALIKFFKKILNQDKLTNVALIGVGNLGHALLNFNFHKSNNVRISAAFDVNDE
ITGTIQSGVPIYPMSDMKEQLTVQQIEIVILTVPASVAQDVTDEIMETGVKGILNFTPLRIS
VPDTVRVQNVDLTNELQTLVYFLDHFGTKDED
>ABC transporter ATP-binding protein uup
MILLQAQAVARHFGADVLFSNVNLDIQDNSRIGLVGKNGAGKSTLISMLLGDSEPDEGQ
IVRKKGLTIGYLAQTTGLDSDRTVFEEMKSVFAYLQEMEVKLHQLEEQIATSTDHQSAA
YEQLTKQYDQLLHDFNAQNGYGYENEIKAVLHGFHFDEDMYDHAITDLSGGQKTRLAL
AKLLLEKRDLLILDEPTNHLDIETLTWLENYIQGYQGALLIVSHDRYFLDKIVTEIYEITHG
HSNYYKGNYSYYIKEKAARLAQEWKSYEKQQAEINKLEDFVNRNLVRASTTKRAQSRR
KQLEKMDRLERPEGYEKGPHFSFAGEKESGNVVLTVSDAAIGYNQTTIAAPINIDLRKH
KVMAIVGPNGIGKSTLLKSILGKIPFIKGSAELGTNVTVGYYDQEQQNLDFKKTVLNEIW
DEHPTMPERDVRSILGSFLFVGDDVKKIVGSLSGGEKARLLLTKLALEHDNFLILDEPTN
HLDIDSKEVLEQAILDFNGTILFVSHDRYFINKVATSVLEIEPTGSTLYLGDYDYYVDKKK
EQAELAAAKASEQSDQTAQATSSVSDSKLSYQQSKERQKQERKLQRAQADAESLVE
QLDAQKSEIEVAMTAPENLNDLGKLQALQKQLDQVTSDLTAAETAWEEASLALEEFLE
QSTESPTKVGGL
>HNH endonuclease

MSEIQKVNYIYVVDSNGTPLMPTSRLGMVRRWLKTGQARWFGNSRKIIQFVRPVTTNT
QELTLGVDAGFHLGLAVVGNKREYYASESLRKSEKDRITSRRELRRTRRNRLRYRQAR
FDNRRRKASWLAPSVQHRLDFTIKEIKRLYTFLPIAKLVVEVAPFDNQKLLNPNIKPWEY
THGKTHGFKTIKDYLLARDNYRDALDGKQYPASQLRVHHLVQRKDGGTDQPDNLVLL
SDINHTQANHNNGILAKLRENRQKTLDYRGAYFMNILATRLSDYFDNYTTTQGYLTANL
RQKYQLEKSHLNDAFIIAGGTDTTLRTNNVYSRQKLRNNNRVLQKFYDAKYIDSRDNKT
KPGKELSSGRTRRSRELNYDNLKQFRKTKVKKGRVSIRRNHYQLRPHDVVLNTKTNKI
ETITGVQNSGTVIKFQTGKTCSIKNVVSLYHVNGMLEKKIENI
>DNA-entry nuclease (Competence-specific nuclease) (EC 3.1.30.-)
MGIGIIVLLILVLLLNLKQLKKIKPVVLMALIAVGFIAGIALSQNNSPVEALNQTKTQLISVLR
SLTEKNAKQERKNESLKAKLSSQAQKLGITLEDTDSKQTASANEDLAQLSYQGQQVIT
VNDNKPGFSAADLSTAKGPWQHYGDLDNLNRATQADALLNQSLMPKTKRERLTVNPT
GWHNKRISSGWLYNRSHLIGYQLTGQNNNWKNLITATRSLNDPGMVKYEDQVADYLK
ASPKHYVRYSVRPVFYQNELLARGVQMRAQSVGDNSVSYNIYIFNIQPGVTLNYNNGT
SRLN
>hypothetical protein
MLILTLGFVGTVKHSPATIQAKTNLTSLSKKQLRARISSAKKKSLSLKADASSLRVLSTSL
ASLEASANSQATSQVQASSEVSQAQTSSVTTGQTGQIVGNRNTHVYHVPGQARYSM
KAANAVYFNTEAEAQAAGYRKAQR
>LSU ribosomal protein L7/L12 (P1/P2)
MALDTEKIIADLKEASILELNDLVKAIEEEFGVSAAAPVAVAGAAGGDAAAEKSEFDVEL
TDAGSAKVKVIKVVKDITGLGLKDAKGLVDGAPSVIKEGVAKEEAEEIQAKLEEVGAKVT
LK
>LSU ribosomal protein L10p (P0)
MSEKIIAKKAAIVDEVAEKFQSAVSVVVVDYRGLTVEQVTNLRKDLREAGVEMRVIKNT
YLKRAADKVGYEGLDETFTGPTAVAFSSEDVVAPARIMAKYAEDIEALEIKGGMIEGKV
ASLEEINALAKLPNREGMLSMLLSVLQAPVRNFAYAVKAVADSKDEDAA
>ABC transporter ATP-binding protein
MTRTLIEIRDLTKSYGKQIVFKDLNLAIQKHSLTTIYGPSGIGKTTLLKIMGGLATYQAGN
VLFNGQAVPAKGARATRYRREHLAYVFQDYGLLEEQTIAANLKIGLAYQKLTKREKNLK
MNEALKQVGLSYRLKTKVKLLSGGEKQRVALARALLAKGELVLADEPTGSLDLNLRNQ
VAELFIKLRSLGKTVVIVSHDPYFKEISDQVIELK
>hypothetical protein
MSAKKWLLGLIVLMSMLTGLTSFYLNQYQQQVQLATQQQDLSKHFTVLRLPADLLINS
ATKQRLIFSALNKSASKYHLNYLEKCADVGYSYDRLGLKPTSNYHNHYTFYVHHLATS
MITGNFGDLGTQKDYYRYTYPYLKGYQVTLKPLNEAKPLRAQYEGIYFVETRNQARYD
AFLADFNQQINHKFKKHYASGDYQLRDEEQLLIPHFYFSDTLASINDYLLLFSLLALVIYY
FLEQASIKLYKLNGYSLPLTVITVMKVPGLLLTVTTGLSLTVLLVAGGVAESFIVLNILFW
QLLVSLGTLTVLYIVGLDELKEKLLEIGSFLGLFSLKFLVVALSFVSLFNIAEATISGFKSV
QVNQVKKDYLCFYPLQVGHNPVGDVDNYLRGGKRLYRLAAKEKILYIDSGAASYKQPA
DVARAKTSVTMNNYYLDAYPLKDIHGHKITVKPSESRLVILFPKDRRNQERAALTYIKKY
VQESNKPLKVIYTDPKYNQASYDALNQRTMADEIVYVVTSESLTTYLNILNGSPRDGLM
IPLVGSEKGLRQKWLKRMSQVNLEDNFPQLIRYSQADRAEFRQELTDLGEHVVENLFY
LLTLILLSGYSLGSYLKLFAREIRLKRSYGYSCFKAYQSYWLLLLGQYLCVGTYGMVRVI
QGADGLNLLVVLAIISLIELAVTGLALRKVEAEND
>hypothetical protein

MEIPNDLPSGSIELDDNDRLVNVSSQRLAEIIHGSEEVSQNSDKRFGTMLTAVNFADGR
WVYYSERTWNLYKVGHSNYLNRNRSHASYAQVGSAKSGWVWRRAGVWSFATAKGK
GTFVAKYN
>LSU ribosomal protein L1p (L10Ae)
MAKKKSKKYLEAAKQVEAGKNYTANEALALVKKIDFAKFDATVEVAFNLNVDTKQADQ
QLRGAVVLPNGTGKDQKVVVFAKGPKAKEAQEAGADVVGDDDLVAQIQDGWLDFDV
AIATPDMMAQVGRLGRVLGPKGLMPNPKTGTVTMDVAKAVSESKAGKVTYRTDRDG
NVHVPLGKVSFSEEALEGNFKTIYDVIAKARPAAVKGQYIKHVSFASTFGPSVTVDLASL
>LSU ribosomal protein L11p (L12e)
MAKNVVNVVKLQIPAGKATPAPPVGPALGQAGINIMGFTKEFNARTADQAGMIIPVVIS
VYEDRSFDFITKTPPAAVLLKKAAGVEKGSGEPNMKKVATVTKAQVKEIAETKMQDLN
AADVEAAMRMIEGTARSMGFVVED
>Glyoxalase family protein
MYNQETVITSYLLNVTNLATMTDYYHHHLGLNILEKVSKQVILGLAGRPLLILKEQVGQT
RQQVAGLYHTAFLVPDQASLGGILRHLLVTNTPLLGGADHGYSQALYLQDPEDNGIEIY
WDKPLSDWDVRANGQIIGVTEELDGTRLLELADNRTGLPSNTKLGHIHLQVTDLKANR
TFYQYNLGFELKSSMKQADFLAYGLYHHHLAANTWTRAKLTIRSRNQLGLGTYALTVP
DLDLVREQLDINNYPYRNLATKLALTDPNGIELIIS
>Phosphoglycerate mutase (EC 5.4.2.1)
MVKLVLIRHGQSIANKENVYTGWNDAPLTELGVKQAILAGQALAQNELEFTDIHTSLLG
RAIKTANLVAEACHQLYLPLHKTWRLNERHYGALRGLNKDDTRKIYGKHQVALWRRSF
YAVPPALAKPDLERRYANLDSRSIPTAESLAQALDRIMPYWQDEIAPKLLADKNQLVVA
HGSTLRALIKYLESISDEGIDGLEVENCQPIIYELDGQLKIVKKEII
>transposase
MNNSIKTILGIKDPYLKLDEKNFDNPIEDQPNQIIVHLIQTYPMHCPKCGHLMCKNGYKT
VNCLGPELHFKPTIWSIKKQKYICKASSSCPEVVTKLAAVEDIHYRNHISLAIKQRAMML
LTKNESQSDLAKELNVSDWTIRRVITNLDQFFKPNYHWLPRHIAFDDFKSGRFAPSGM
SMILMNIENKRTLDIILSRKNSYLRKYFLRYDRSARLAVQTVTVDLYTPYRHLIHELFPHA
LIIADHFHIVAQAYRAFNKIRIQVMNRAGAGTHKWRALKHFWKLLLTPANELKYDNYWS
RRNFSYAQLTDVEVIHRLLSFDNELKRAYEYYQNLILVIAHRSKKELKNLLAIKWTQLPQ
ALQKVQRTLRRHKQEIYNSFKYDTYTNGPVEGTNNKIKVIKRTAYGFRNFFNFRIRILLA
LPNTYIAINWRNKQTAHAKVQAQAA
>hypothetical protein
MKKLGLKKVAVMAAAAFTVLGVATSASADTQSDLDAVKAKIAETEKKVSEGQVALSKL
ESQEAKTSQEIATLKKNIAARSEQLADQARAAQLNDAGSMLEFVANSDSVTEAVSRTV
TVATVVRANNQTLADQKADKEKLAQKQTDLQAAAEKQKKQNKELAQARADLAVQKTD
LEAKKASEDEAKAKAEAAKKAAEEAAKQVAAAKSTDDAEKAANQAEAVAASVSSSSN
NGNNNSNNQAASQTTSQASSQASSQASSTSAAPAASSVNSSSVLAMADSLTKMHIPY
VWGGTSLSGFDCSGLTQYVYGHFGKSLARTAAAQAAQTTRIPVSQAQPGDLLFWGE
GGGVYHVAIYVGGGSYIHAPTEGQFVKYGSISGWAPSFAGRL
>Permease of the drug/metabolite transporter (DMT) superfamily
MKKTIIYICLSTFLFSSMEIALKVAGATFNPIQLNLIRFFIGSLVLLPIAQASLSKLNRKLTR
ADWKIFWLTGFCCVIVSMTLFQLAVAATKASTVAVLFSCNPVFALLFAFLLLRERMSRA
SFISVVVSLIGLVVIIDPAHLSQPVGIVLGLLSAVTFGFYSIVSRYGSNKLGLNGVVMTAY
TFVAGTIELLVLSLVTHLPFVARSLAAIPALKTFANLPILANITIGSLPLLLYLGICVTGGGF

AFYFMAMENSDVSTASLVFFIKPGLAPILAAILIHEKILPTTIIGIVIILIGSVITFVGNRTVED
TSVVNKLPEEKEVVAEEDKLGHYTTHLKH
>Malate permease
MQVFYHSISGVLVILVMIIVGYILTERKWFDGRSSGLIARLVTQVALPCYMLDTVTSKFTA
HELLKMLPELKYPALSMTILMALAMAIARIFRVDKSRRGVFTSMFFNSNTVFIGLPINEAL
FGNQSIPYVLVYYMCNTTFFWTIGTYFIQRDGRGQKARLDLRQVIRKVFSPPLLGYILG
VILVFLKVKIPTFIQSDLHYIGGLTIPLSMIFIGIAVSNAGLGRLTLHKDNLLILGGRFLVAPA
LMALLVMQAPMPLLMKQVFILQAAMPVMTNAPVVAKLYGADAEYAAVMVTETTLMTM
VVVPLLMVITQNLT
>Mobile element protein
MDQIRADQAELPKAQRYKIGDLLSRINLAKATYHDERKRIANPNDKYCEVKKVIRRICEQ
GKVRGRLTYGYRRVQARLDQLGIHLAGATIRHLMSELGVQVEIYNHHRNGKYSSYHG
RVGSVADNVLQQRFNAKDPYQVIHTDVTQVRLANQQWAYISAMTDEASKEVLAFQMS
SHPNQELIMNTLNELLDHLPDDVQPIIHSDQGWHYQLKYYTKKLKDHHFIQSMSRKGN
CHDNAPIESFFHLYKTELLQGLPPCNNITELRKLSLTYIDQFNNERISLKTKGMTPVGYR
NHALAV
>hypothetical protein
MTKYSPEFKAQVVVEYMTDGLSSYTLEKKYKVSARQIRKWIQQYQLQGIKSFKQHHRK
RIFSAKFKWHVINYYQTHAESFAQVAAKFDLLPSQISIWRQAFIKYGYDALKPHRKGRS
QKVKKQTKKQKQKLVEKKEIDQLREELARTKQELYNTKMDRDILKKSLALFGPSKPGK
KRK
>Exodeoxyribonuclease III (EC 3.1.11.2)
MKVMTLNTHSYIESQPLEKLAQIATYIKEENLDIVALQEINQLITSKEIEGDDYFCPLPNQ
VPLHEDNFAYLLVSRLRELGLNYYWSYSYTHIGYDTYQEGLALLSKEKLISTEQVTSNP
SYVGKLTRKILVGQTSQATFVVGHYSWAGKNGFDYEWAKTKAALKGQKQPLVIMGDF
NNPPESSAHQQIREANLGLVDSFQVAAKRYGEHTVVKAIDGWQDNEGQLRIDYVYLS
DTFKVSKYQVIFDGQTSPQVSDHFGILVEIE
>PTS system, maltose and glucose-specific IIC component (EC 2.7.1.69) / PTS system, maltose and glucose-specific IIB component (EC 2.7.1.69) / PTS system, maltose and glucose-specific IIA component (EC 2.7.1.69)
MKKIFSFEFWQKFGKALMVVVAVMPAAGLMISIGKTIPLINPHVTALVTTGGVVENIGWA
IIGNLHILFALAIGGSWAKEKAGGAFAAGISFILINRITGSIFGVTSEMLNDPKAYTHTLFG
TKIMVKGFFTSVLEAPALNMGVFVGIIAGFIGAMAYNKYYNYRKLPDALSFFNGKRFVP
FVVILWSLICSLILAVVWPTIQGGINNFGLWIAESQDNAPILAPFLYGTLERLLLPFGLHH
MLTIPINYTALGGTYEILSGPQAGTHVFGQDPLWLAWATDLVNLKNAGDMSQYHYVLT
HWTPARFKVGQMIGATGILMGMTLAMYKNVDPDKKHLYKSMFFSAALAVFLTGVTEPL
EFMFMFAAVPLYVVYAFIQGAAFATADLVSLRVHSFGNIELLTRTPLAIKAGLGGDLLNF
VWVSILFGVFTYFLANFLIKKFNFATPGRNGNYDLDTGDNASGTGAPVEGNDQAFAIIN
LLGGKANIEDVDACMTRLRVTVKDKAKVGSEDQWKAVGAMGLITKDNGVQAVYGPKA
DVLKSDIQDILDSGMEIPNVVVDAAAAKKETVAKGAQAEVVAVANGEVMPITEVNDPVF
SGKMMGDGYAVKPADGKIVSPVAGQVKSLFPTKHAIGLETAAGMEVLVHMGIDTVSLE
GKPFTLMVKEGDQVEAGQELAQVDLAALAAAGKEDTMIVAFTNVDQVTFTLDKTGKQT
AGTVIGQAKG
>Cell envelope-associated transcriptional attenuator LytR-CpsA-Psr, subfamily F2 (as in PMID19099556)

MNNKNQRDYRPNDIPPRREHRGHSSQHVLISPHRKRKRLFTILLVLALVIFYGGGAYLY
HMLANAKSAVNETYTATNIKKSRNVSKVLKAKKPISILLLGTDTGDLGRTDKGRTDTLIV
ATINPSTKRVTLTSIPRDLQVKVPGSIDSYDKINAAYTIGGVSTAIKTVQNTFHIPIDFYVL
VNMGGLTKIVDAIDGVTVTPPLTFSYGDAHVTKGKKVTLNGKEALSYSRMRYDDPLGD
YGRQKRQRQIITAIIKKSLSVSTLSSYEELLNTLEENMKTDLTYDDIMTIVSEYKAAGQSI
KSYTVQGEEAMIDGASYQVASAAVKQKNSDRIRNELGLEPSDAEFTGRVDYDDSYDS
TSTTTSTYDNTYSDSYGDSYGTSTYGY
>hypothetical protein
MESFDGPDSYYKSLLSVTNTLALNCMQANLSLTGPQTEQLKVVLDLLLEYGWEPEYYV
SDVLDLPYEQLWY
>hypothetical protein
MKKAFTIATLRKKISSSQPKEIENLLVELYRTNTTASRQIDVFFNSNSLEDEY
>FIG00750050: hypothetical protein
MKYDKFNQQYGLMFGKSKLVFIKTGAAGSIYGHKNKYLELASKIQNERGYAVVVSANP
VGSPLNLQEELEKVSTYLTDIKEIILIGTSRGGLLVLQQGYLNTKVSRILAINPPLAINWHK
TKKGLINFSGAKVQVVFGQYDPSVDYSELIERLEGLETDCSSQIISKADHNFKGKLDIFQ
QLVMQFVLKE
>hypothetical protein
MEFIQYNQAGKIAKRCICELNQLKCPPLGQKVRATLNDEKQYIGFWDTFFGFDTKIVEIS
KYDLDENTGELRSDNSLVTFIPTDKIFRLEAILYSNPRWGVAPTNKFTFSKPVRRDFKS
DFFKNWSVK
>hypothetical protein
MQKIKLMFEFMHGPIWPSDPFTGQPMTDIKLVDNDPDLKSWNTRCSELWDECYEFNS
HEQACYFNKVTLANNKQELEDLLANIKKRLVELNQEEFILEDQVTDCLNGGS
>hypothetical protein
MKIKTSEWQWTRKPKAYTITDDKIEITTNPHTDLWQRTYYHFRNDNAPVLQVKTTDKYF
SFVVKTEFDSKVRFDQCGVVMYLNSENWLKSSIEYENEEFQHLGSVVTNNGYSDWAT
TEIDANIKQMWYRLSRREDDFCLECSEDGVKYKQMRICHMAKANDEIQFGIYACSPEE
SSFKATFTNFEVTDCKWLTHDGQAPDEN
>hypothetical protein
MKERAEVNSDDLRTEREKQYRRIIKRHQAAWLVLFIAAIAPTDYSLAHQLIWLHYLFVILF
VATTISMGYWHYRLKNFLQENGK
>Site-specific recombinase
MRTIFYRLSNQAIDCLQYIRELKELHVAVFFEKENINTMDTKGEVLITIMASLAQQESESI
SRNIKIGLQYRYQRGQVIVNTARFLGYDKDDDGNLVINPEQAKVVKRIFYECLTGKSAIE
IARELTKEGLKNGIGRTKWHSSGIIKILRNEKYIGDALLQSKHPIARYIKIPS
>Mobile element protein
MDDLTCPDCQHELKEIGSFCARQELLYIPAQVKRIDHIQHSYKCQHCSDKAPADKIIKAP
VPKAPLTNSPGSASLIADTIYQKYVLKVPAYRQEKDLRRMGLPLDHKTVSNWHIKVCEY
YLSSLYELLRKELLKLDVIHADETPYRVLDSERAKDYIWTFLSGKHAEKPIALYHHGSRK
GAEAWDFLTGFSGYLHCDQYPGYLRLSQQDVTLVGCMAHARRKFRDSLPKDKARES
DATSVAKQGIHYCDQMFSLEHSWKDLSAEERYEKRQSELRPLLEKFSDWCSKKSISVL
PSGKLGTDFQYCIKHMDKFMNILKDGRLELSNNRAERAVKEIVMGRKNWLFSQSSTGA
KSMAIIMSILETAKQNGLDQFKYINYLLDKLPNELSLLDTQRLEAYLPWAENVQLHCK
>Mobile element protein

MTEKETLALIETLKANNEELLKQVALLTEQVAHLTNKLYGKSSEKRAVNPDQLSLFEED
NASQDDEDLPSHYRNYYL
>Mobile element protein
MINLAELGRVFIVYGKTDMRRGIDSLAYIVKKSFNLDPFSGCVFLFCGSRRDRFKALYW
DGQGFWLLYKRFENGKLAWPNNEDEVRELSDEQVMRLMQGFTIDPKINTARARELY
>L-asparaginase (EC 3.5.1.1)
MVSKNNYQLYQQKHVDGKQRWGIRKLSVGVASVLLGTTFMLYGNHAVLADTVTSPSD
NVTRSTTTQGGNKDKVTEGTTSTPQTSGDSTDKGQGKTLTETPTGTDKQANGQKVN
QQVPTTDTEEKVVATNHQDTPQGQNPTQGQDTPQNTTNVDKKDTEVTPANDATTPT
PQNTANVGKKDTEVTPANDATTPTPQNTTNVDKKDTEVTPANGATTRKTNAKFKVLAA
RPVMKVAGTASLPISNQDIKLDPQPMLTEIINKPTDNWVYNNLKWYQDTSTAKVKEILQ
NHTANDESGRYYFAGVANYNEHYHAIYLLARSNNLNDNSLYVTILHTGLGKNIQEEVVA
PGESKKVEYSGTTHTPIFTNYDGTSASIDLDGIEKGDNIYGMVVGFAYGHNTGIKGDPA
SMGNGFVMTPIPTKMTTTIHYIDQATGDEIAVPKSFEGVAYQKYTITGEAPTIDGYTLKK
SPETTGYISPYKVGESYDFRLDKHVVIKQTVIDSQGLIRITAYYDGEAINNTTRYLGTTSG
VNDLFSFISHGKSYTYINQITSTNDGIVYYYAKDGSEDKSEVRVHYIDVTGNKSSLFVPG
DGEEVATDKISGKLGENYNYNVNLPTDYNLATNQANTVNGTYTIDHHDEYVYVVKKTS
AEKLDPTVKDKTKVDDPTKLTDGEKKEVEDNIRANNPGLPEGTKIKVGDNGDTTITYPD
TSVDTIPGKKLVEEKNSAEKLDPTVKDKTKVDDPTKLTDGEKKEVEDNIRANNPGLPEG
TKIKVGDNGDTTITYPDTSVDTIPGKKLVEEKNSADKLDPTVKDKTKVDDPTKLTDGEK
KEVEDNIRANNPGLPEGTKIKVGDNGDTTITYPDTSVDTIPGKKLVEEKNSAEKLDPTVK
DKTKVDDPTKLTDGEKKEVEDNIRANNPGLPEGTKIKVGDNGDTTITYPDTSVDTIPGK
KLVEEKNSAEKLDPTVKDKTKVDDPTKLTDGEKKEVEDNIRANNPGLPEGTKIKVGDN
GDTTITYPDTSVDTIPGNKLVEEKNSADKPDPTVKDKTKVDDPTKLTNDEKKEVEDNIR
ANNPGLPEGTKIEVGDNGDTTITYPDKSVVTILGNKLVEEKNSAEKLDPTVKDKTKVDD
PTKLTNDEKKEVEDKVTEANKDKFPEGTEVTVDDDGTVTINYPDGSQDTIPGDQVVEA
KTDADKTNPTVLGTKIGVKDPNNLTDEEKQQVKESIENANKYQFPEGTQVTIGNDGTA
TITYPDGSQDVIPGSQLVTKQGSDVANTGDVKNNGQGTTASIQGVTADTTAKKLPQTG
EQDTSATAATGLGLIILSLLGLFGLGSKKRKED
>Mobile element protein
MTLSEGMSREFLIMRIPDKTSASVMQAFCQLKAQYSEHWNDIFKTITTDKGSEFADLSK
LEDISKTLVYYAHPYTSCDKGTVERHNRLIRRFIPKGETITDYSLQEIMDIETWCNSLPR
KILAYHTPDEIFEKEIDRIYQAA
>Site-specific recombinase
MDAKGEVLITIMASLAQQESESISRNIKVGLQYRYQRGQVIVNTARFLGYDKDYDGNLV
INPEQAKVVKRIFYEWLTGKSAIEIARELTKEGLKNGIGRIKWHSSGIIKILRNEKYIGDAL
LQKTYTVDFLTKKRVKNDGQVPQYYVENNHSAIISRATFYQVQKILDMRREGFTTEGG
HHHGYTNAYCFSSIVFCSRCKDIYTRCV
>Site-specific recombinase
MGRTITEAALEEASLKAMNELIQQHQLADKQIAANILKVTKGTTGPSLDDLDQQLEDQQ
LLLLSMSTHNKNVEQLTEQVQALRKQREQLIQQEIDHDIKRSNLKNIQSFFQTYQGGLT
KFDEKLVRLLIEKITVFKSKIEFTFKDGEVITIKA
>RNA methyltransferase, TrmA family
MAKLKLPVYKKEELTVRIEDLTYQGLGVAKVEGYPLFIKEALPGEEVVVKVEKLGKHFG
YAKALKWLSTSADRVKTKAKYLQTGIAPLMHLSYPAQLAFKQKLIKDLLAKAHLDQLEV
AQTLGMAQPYAYRNKAQVPVRQVRGQLEIGFFKQGSHDFVSLEDFYIQDKRIDQVLVT

VRDILRKYQLSAYDEKSHTGVVRHLMVRRGYYSKQMMVVIVTRTKKLPHASEIARDILA
ACPDVVSLLHNVNEAKTNVILGSKTELLAGQAEIEDQLDGLTFKISAQSFYQINPTQTER
LYQTAIKQAGLTGKEKVIDAYCGIGTISLSMARHANKVYGVEIVPQAIADAKRNAEVNKL
TNLTFEVGEAESWMATWQAQGIKPDVIMVDPPRKGLTESLITSATAMAPKKIVYVSCN
PATLVRDLQSFIAKGYQVTAPIQPVDQFPQTTHVESVTVLERKK
>Transcription regulator [contains diacylglycerol kinase catalytic domain]
MEKRCRIIYNPTSGREAMKNDLVEILNVLENAGYETSAYATTPEPKSAQNEAKRAALAG
FDLVIAAGGDGTINEVVNGIAPLEKRPKLGIIPAGTTNDYARALKIPRESPLEAAKVIAKG
QTVKMDIGQAGDQYFVNIAAGGLLSELTYSVPSQLKSLFGYLAYLVKGAELLPRIRPVN
MRIEYDDGVFDGQASMFFLALTNSIGGFEQIVPDASLDDGKFTLIVVKTSNLLEILQLITM
VLNGGKHVNDPRVLYVKTKKVKAEAPDSKMMINIDGEYGGDAPMTFRNLKQHLEIFAN
TDKIPNDAITNREEENEKDARELFVKGVEDYSKDKNE
>Aspartyl-tRNA(Asn) amidotransferase subunit B (EC 6.3.5.6) @ Glutamyl-tRNA(Gln) amidotransferase subunit B (EC 6.3.5.7)
MNFETTIGLEVHVEMKTKSKAFSPAPVQYGAEQNTNTNVIDWGYPGVLPEVNKGALEF
GMRAALALNCEITKDIHFDRKNYFYPDNPKAYQITQAQTPIGTNGWLEIELEDGSKKKIG
IEEMHVEEDAGKNTHNPDGYSYVDLNRQGTPLIEIVSKPDIASPAEAYAYLTRLRQIIQF
TGVSDVKMEEGSMRADVNISVRPIGSDKFGTKTEMKNLNSFDHVRKGLEYEAKRQQQ
VLLAGGKVQQETRRFDENTGETILMRVKEGASDYRYFPEPDLPPIHISDDWINEVKASI
PEMPDKRRERYVNDWGLSDYDASVLTQTLEMSDFYNETVANGADPKLAANWLMVEV
NAYLNANKLELGQTALTPAHLAGMIKLIQDETISSKIAKKVFKEIIANDTDPKAFVEAKGM
VQLSDPAKLQPIIDEVLDNNEQSIEDFKNGKDRAVGFLVGQIMQKTRGKANPKVVNKLL
MASLKER
>Mobile element protein
MLRVQKVRLYPNETMKQILDDLCDYRRYCWNQGLALWNDMYDASLILENKKLRPNER
KVRDELVANKEDWQYQLSARCLQLAVSDLGKAWQNFFKKSLPDWGKPKFKSKKTAR
QGFKTDRAQIVDGKLRLDKPRGVKAWADISFKGANDLKGELKVVSIYRENGKYWASLP
FEVKVTKKTTTGQKTAVDVNVGHFDYTEGQVKTLPNNLKILYKRIKHYQRLLARKRVAN
GKKATQANNYVKTKAKLQRDYRKVASIQHDIVQKFTTMLVNDYDRIAIEDLAVKQMQM
SHVASKGLHRSMFGYFKQVLKYKCEWYGKELILADRYYPSTQRCSQCGHVKTGTDKV
GLDGNKKHGTKHNEYICYECGAVMERDENAVRNLLALM
>Aspartyl-tRNA(Asn) amidotransferase subunit A (EC 6.3.5.6) @ Glutamyl-tRNA(Gln) amidotransferase subunit A (EC 6.3.5.7)
MVDYFKTDLENLHQALVNKDLSSAELVKETFAKIETIDPKIEAFLALNKEQALAKAEELD
QKGIPADSLLAGIPVGIKDNIVTKDLTTTAASKMLANFKPIYNATVMDKLDQAGMINVGK
LNMDEFAMGGSTETSAFKLTKNAWDQTKVPGGSSGGSAAAVASGQILASLGSDTGG
SIRQPAAYNGIVGMKPTYGRVSRYGLIAFASSLDQIGPLTRTVKDNALILNAIAGHDKRD
FTSSAKEVPNFAEGLDGNIKGLKIALPKEYLGQGVAFEVKEAILKAAKTFESLGATVEEV
SLPYSKYGVPVYYIIASSEASSNLQRFDGIRYGFRASDVKNLEDVYVRSRSEGFGDEVK
RRIMLGTFSLSAGTYDAFFKKAAQVRTLICQDFAKVFSQYDLIMGPTAPTPAFGLGEDI
DDPVTMYMNDILTIPVNLAGLPGMSLPAGFANGLPVGLQLIGPQFAENTIYKAGYAFEQ
ATDFHKAVPTIG
>Aspartyl-tRNA(Asn) amidotransferase subunit C (EC 6.3.5.6) @ Glutamyl-tRNA(Gln) amidotransferase subunit C (EC 6.3.5.7)
MMAITASEVAHVAGLAKLSFNETELEKFTGQMDEIIKMVQELSEVDTEGVPVTTHVTDA
FNVMREDVAVKGLDRDLLMKNVPESQDGYIKVPAIIDESEEG

>Putative pheromone cAM373 precursor lipoprotein CamS
MNKKLNLMVTLGLSTLLLTACGNLGSSSGTTTNTNEESASSGYQTTGQTNDSEYTGVII
NGRYRTNKSRGVGVSQNSENLLNLKSFESGLTDISKEYFSTKDYVFQEGQTLSKSTVE
DWLARKSSKNPDGLNPEDNGKNGENERNPMYVQQIEEQDYMKEDSKGNLSLAGITV
GIGMNRKDYYQKEEYGATYVTTISKEEMEKQGKIAAAKVVARLRAKVGNDVPIVVAMY
QQAEMDSLVGGSFYAYTEVKAGSSEITDWTDTDISSYVFPATSTKSVPNDNDEAAFES
FKNKIQNFFPNLAGVTAQAKYKGKNLQGLKVTITTQFYSETEITSFTQYVAKAAKTYLPS
GIPVDIQVKGSDGTVQSFLAKTAKDSSYYTHVFTSY
>DNA ligase (EC 6.5.1.2)
MEEKQAKSRLAALRQQLMTWSQEYYVNDNPSVEDAVYDQTYRELLTLEQQFPNLVTA
DSPSQRVGGQVLAGFTKVEHEIPMLSLGDVFSKEELAAFIDRLNEATKAPFAYNCELKI
DGLAISLTYENGVFVKGSTRGNGIIGEDITANLKTIKAIPLTLKEPVSIEVRGECYMPKDS
FLALNKEREAQGQAVFANPRNAAAGSLRQLDTKITAKRNLSVFIYYLMEPEKLGVTSQS
QALAKLAELGFKVNPDFRLATDMETINNYIDEFHAKRGELAYEIDGIVIKADDFKVHELV
GNTVKVPRWAIAYKFPPDEKETVIHQIEWTVGRTGVVTPTAVMDPVEIAGTMVARASL
HNPDYLREKDVRLGDTVKLHKAGDIIPEISQVVVAKRPEDSEPYEIPTTCPECGAKLVHL
DDEVALRCINPKCPALIKESLAHFASRDAMNIDGLGPKIIERLYEAHLISDVADLYKLKLE
DLLALDKFKEKSATKLLTAIENSRHNSSERLIFGLGIRHVGAKVARLLAERFKSVTALSQ
APSEEIAQIDSLGQVIADSVVTYFERPEVKQLLAELEEVAVNFDFLGQSASEIAQANEDS
FFNEKKVVLTGTLDQLKRSQAKQWLEAHGAKVTGSVSKKTDLLIAGHDAGSKLTKAQE
LGIEIMDEQQFIAQMEAIDGE
>ATP-dependent DNA helicase UvrD/PcrA
MGARDLLASMNKEQAQAVQTTEGPLLIMAGAGSGKTRVLTHRVAYLIENLGVMPWNV
LAITFTNKAAREMRERINTLLGDSSGQEVWVFTFHALCVRILRRYADRIGYQRAFTITGS
SEQRTLIKHILKDLNIDPKKYDGRAILATISNEKNKMVTPTDYQAKATSPYEQIVAKVYAE
YQRRLRANQAMDFDDLIMETIRLFKEDAEVLEYYQRKFKYIHVDEYQDTNEAQYQLVS
LLAQGYHNLCVVGDADQSIYGWRGADMSNILDFEKDYPDATVIKLEQNYRSTKNILAAA
NAVIKNNDNRRAKTLWTENPTGAKINYYRADSAASEAFFIIKTIKEQLAKGRAYGDFAVL
YRTNAQSRSLEEVFVKANIPYKLIGAHRFYDRKEIMDILAYLRLVTNPNDSLSFERIVNE
PKRGIGATSLAKLQDFATMHDLSLFEACQSIELVGIRGKASKALSEFADLIEKLHTYQET
ATITELTEEILKRTGYQQMLAENQSVENESRLENLKEFLTVTKQFDDNWQPSEEADDP
FVEFLGDLALVSDQDGVEDEPNEVTMMTLHAAKGLEFPVVFLVGMEEGIFPLARANED
ENELEEERRLAYVGITRAQEELYLTNASSRMLYGRFQSNPVSRFIEEIDGDLINYVGLAT
KQGSPKPLSFAERMNHQSRGGLRTVKNTNPFRQARPASPQFSQTKSEANVATSWKV
GAKVNHKKWGLGTIVKVSGSGDETELDIAFKEQGIKRLLAAYAPISKA
>Phosphoribosylaminoimidazole carboxylase ATPase subunit (EC 4.1.1.21)
MEEEILFPGATIGIIGESPNGIMLEQAAHKLGFDVIAYGPNEEAPTLRGADVKVVGTYTD
QAKLQDFAQRCSLVTYESETIPAQTIAYLSRFTKLPQGSETLEIVQDRLLERTFFEQSNL
NIAPYATIVNLDDVYQAISSIGYPAILKPIQKGFSKKIVIKKQTDIAKCADIIDQGTFILESMI
PYQKELAVTMSKDKNGDIKFYPLVEAEYRQGKLHQVLAPAQVDSDVANEIRRLSELVV
KQVKLVGIMTLSFFLTETGALYVKRLVTGVNSLGYVFSRAANVDIFEQHLRVLANMPLA
QPELIQATGMVMIEQDKREALRTQWLLKTNWHYQFYRYPKSMTTLNWGHVLVTGESS
QAIKEQVAATGIWDKLEE
>N-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28)
MVKRRSRRQGNLVNYLLLLILLGLLFCLVFPFFNKQEEKEPQTEVRKEFIARIAPLAQAE
QRKYRIKASITIAQAALESDWGQSQLASKYNNLFGVKGNEQDGTLLATKEYVNGQWIT

VEDYFVVYPSWKAAVEAHSKLFVDGTGWDHNHYQAVLEASDYKQAAQALQERGYAT
DPNYAAKLIELIEQYQLDSYD
>Substrate-specific component PanT of predicted pantothenate ECF transporter
MNTKQKLVKRVIIALFAAIIIIQNFIPFLGYIPLGPLNLTIIHITVIVIALSMGPLYGGIVGGIWG
VITWIRAFVWPTSPLATIVFVNPLIAILPRILIGIVAGGSFSWLKKRLAKQSSAMVLAAILG
SLTNTLLVLGQIYLFYRGKSQAMYALNVNELLPYLLGVIATNGLPEAILAGLVSPLIAGPL
RKRLKNN
>transposase
MVLTTGGIALTLDSLESTFYPFASLCCVFTEQKCYNVSKESEVNLMEIQITIKVKLNLANA
EIASSFTNTMEQYRLACNYVSEYIFNHDFDMKQSRLNKVLYTNLRSLFMLKSQMAQSVI
RTVIARYKTVKTQMKRNPYKYQDINTSEWYRETRDLTWLQKPIIFNRPQVDLQRNRDW
SYLSSGQLSINTLNGRVKVNPVCHGFDQYFDGTWKFGLAKLLKTGGKWYLHISATKAV
ADFDNQIVKHVVGIDRGLRFLATTYDEQGQTAFFDGKTVMRKRAKYQKLRAKLQAKGT
KSAKRRLKKLSGRENRWMTDVNQRLSKTLVQKFGSNTLFVLEDLTGVSFERTDLPKVL
RNQNKSWAFYQLAQFLTYKAHLNNSEVIEVSAQYTSQRCPKCGTIKKSNRNHDLHEYY
CTNCDYRSNDDRLGAMNIQLLGTQYISGQEQPKFELTTNA
>Malolactic regulator
MNIKSLEYYHKLVQEKNFSKVAAYFDVSQPTITLAIKRLEAEYQTTFFIRDRSHKELVVT
DMGYQFDQHIISILNELEIAAKEIKRGSEQKILFGLPPIIGNYYFPKFTPALLKAGVMEQLE
TFEGGSTDMLSHLLRGNIDLALLGSLEPIKQPKLKVVQLASNPFKIIVSKQHPLANKGAV
AFSDLKKEKFIDFSQGFVHTTALKQLAKAANIRTKIIYQTNDLHIIKALVAQNIGIGFLTDLA
ILPSDELVALDILDEDQPTFITSLASRTNHILTEQQAAVFKVFQENQ
>NADP-specific glutamate dehydrogenase (EC 1.4.1.4)
MSYVDEVYNRVVAQNPSQPEFHQAVKEVLESLRPVIEADEAKYRKEALLERLTTPDRQ
LLFRVSWVDDQGQVQVNNGYRVQFNNAIGPYKGGLRLHPSVYLGIIKFLGFEQVFKNA
LTSLPIGGAKGGSDFDPKGKSDREIMAFCQSFMTELYKYIGADIDVPAGDIGTGGREIG
YLFGQYKRLKSTYEGVLTGKGLTFGGSLARTEATGYGLLYLVDALLKDHGHDLQGKTV
TVSGAGNVAIYAIEKAQQLGAKVVTCSDSTGWVYDPEGVDVALLKEVKEVKRARLTEY
AANRPSAVYHEGRGVWTVKTDVALPCATQNELNLDDAKQLVANGVVAVCEGANMPT
TLEATKYLQENKVLFVPGKASNAGGVAVSALEMSQNSERLAWSFEEVDGKLKDIMVNI
YHNIAAAANKYGLDGDYVAGANIAGFLKVAEAMEAQGVV
>Fumarate reductase flavoprotein subunit (EC 1.3.99.1)
MQTKIQTANIDSLEKEYDLVIVGSGATGLVSALQAKRLGLKVAVIEKMASLGGNTMRAS
SGMNAAETEVQLANGVIDDFTAFYQEILQGGGKLNDKELLEYFVTHAPLAIEWLAKQGI
VLSDLTISGGMSQKRMHRPQSMAPIGAYLIKGLLARCVAEAIPLFTGTKLVAINKAAGKI
TGITVAGTDGKEVEVGVKALLLATGGFGANANLIAQYQPQLANYKTTNHAGASGDGLL
LAQKLGAQLVDMNFIQVHPTVQQDFDHTYLIGEAVRGEGAILVNKNGERFVNELATRK
RVTAAITALNENGAYLIFDQTVRQRVKAIEFYDKVGLVVTGASLAELATKLALPKENLSQ
TLSNWNQAVATRTDCQFGRKTGLVTLNQAPYYAIHIAPGIHYTMGGLHINAKTQVLDED
GQVISGLYAAGEVAGGLHGNNRIGGNSIAETVVFGLQAGKQAAKYILGK
>hypothetical protein
MVSSLQKSMINLAQVLRLREHQKSFKELEKLGAIDVYYLEAIYQLKESTIGTISKLLGQST
PNTNYHIKKLINLGLIEKRIDDVDRRVTHLLVTDKYSELINSDEDFWARMQDRLESEVAP
TDLAIFQRVLRQTIEIIGEENRQDSVIG
>Transcriptional regulator, TetR family

MSQLQEKRKNSLALSFQELLRSEPLEKITVEQICRKADVHRSTFYRYFQDKYDLLRYAF
EHFLIARIDEEDLIGSTISMISKEKQLFRNISVNNDSSFLLWQMLEMLSERLLVSAKNGKL
ESTPWLARELNSSQYPQLAADMVAGAFLTLLYKWINSNYQMEPKQLANFLGQLGKIE
>hypothetical protein
MLTEAANFQCDLTEVDRINAKLRPFLTEISASENRIL
>Negative transcriptional regulator-copper transport operon
MNERVEISNSEWEVMRILWTLQAATSRQIIDLLAEKLGWKEATTKTYISRLVTKKYVKTE
KRGRAYFYLPAVEEEATITEQLLASFAKICQKHIGKTLANVVREIPLSQADIEEVSRILAA
KQKTAPLELECDCLPKGHKAKHC
>L-lactate dehydrogenase (EC 1.1.1.27)
MRKLAVIGLGHVGATVAYTLVTKGLVDELVLIDKNEDKVVAEQYDLLDMLARIDTSTKIKI
QDYASLKDADVIITTFGNIRALSDGGNRFGEFDHNVAQVKEVGQKIKESGFNGVIVNVT
NPCDVIATLLVRYTGLPHNQVVGTGTFLDTARMQRAVGQALGEDPKNVTGYVLGEHG
ESQFTAWSTVRVKDHPMAEVAAEKQLDLPDLDNQARRGGWLVFHGKKFTSYAIATCA
VKLAQAVLADAHLACPASVFVEKYGCYVGYPAVIGANGVEAVNHVPLTKEEEDKLTSS
ANFIAEKVASVK
>ISSth1, transposase (orf1), IS3 family
MSRKSKYSVEQKLNILNEAVHSSFKRVAKKYGVNESTILTWYRLYKYQGIDGLRSIHSN
RSYSKEFKLSLVEQYQNSDDSLELFAIKHGLKSKSQLLNWIILYNESKLKVYTPRKRDSI
MPGRKTTFEERLTIIEDLIKHDVNYNWAVNKYHVSYQQVYGWYQKYRKSGNDPQSLR
DRRGKAKPKAEWTELDKLRAENRLLKAEMKRNEMEVAFAKKLIEIRNREAKKSSNTKS
>Mobile element protein
MLEAILELEEKHKWTLGYLGMTTQLAFENKLSFKVGLKRVTNCMRNHGIKANIRKKKHN
RIKRHEEYINDNLLNGQFDRQNKNEVWVTDTTEVLYGSEQVRKVRVHVVMDLYGRYIL
SYNISATETATSAIEAFKRAFKVEPGAQPMIHTDRGAAYCSRAFNDYLAEQNCIHSMSH
PGHPWENSPIERWWNDFKLVWLAKRARSKSLSELEQSVQQAIKYFNTERAYASKNGL
TAEKFRAQAA
>tRNA-Thr-AGT

>UDP-N-acetylglucosamine 2-epimerase (EC 5.1.3.14)
MEKIKVMTVFGTRPEAIKMAPLVLALQADERFEAVTTVTAQHREMLDQVLEIFKIVPDY
DLNIMHKEQTLTEITANVLTKLDQVIKTVKPDIILVHGDTTTTFAASVAAFYNQVAIGHVE
AGLRTWNKYSPFPEEMNRQMTDDLADLYFAPTEESKANLLKENQAAAKIFVTGNTAID
ALKQTVSSDYQHELLAKLDPGHKLILVTMHRRENQGQPMARVFRTLLEVVEQRPEVEV
IYPVHLSPKVQKLAHEILGNHERIHLIAPLDVVDFHNLAAKSYFIMTDSGGVQEEAPSLG
KPVLVLRDTTERPEGVSAGTLKLVGTEPAKVKAAMLELLDDPNVYQQMAQAQNPYGD
GQASQRILAAIAYHFGLSEQRPKEF
>Aminopeptidase S (Leu, Val, Phe, Tyr preference) (EC 3.4.11.24)
MTMTNFEEKLAKYAQLIIEVGVNVQKGQAVVLYISVTQQKLAHLLVEKAYQAGASEVIV
KWSDTFSNRQFLKYASAERLSNAPEYIVKEAEYIVAKKAARISVVSEDPGAYAGIDQKRI
AQNQKTMGQALLSVRKATQNNDLSWTVVGAADEAWAKKIFPDLATQDAVSRLWEEIF
ETCRINEADPIAAWKAHDQQLRSKAEWLNQEQFSALHYTSPLTDFTVGLPKNHIWEAA
SSLTPDGISFMANMPTEEVFTAADCHRIDGYICSTKPLSYAGNILEDMKFTFKDGQIVEA
TAKKGQAVLDHLLELPGARSLGEVSLVPDPSPISQSGITFFNTLFDENASDHLALGAAY
PTNVQGGTSMTDEELLAHGVNVSQTHVDFMVGSSDMNIDGIKADGSLVPVFRNGDW
A

>tRNA-Arg-CCT

>hypothetical protein
MKVKTKLNVSAEFLYQRLIQTALDDLSQELGYQAAASELAGAKYQAGNEEHKAAVTVL
AAEQNQSYHYMVELAQGQLMRSYDLKKLGANQVELVYQSQLVAKNWLAKFELAVSD
FFLGWLKSYNYKKLLKQLELSYSL
>ATP-dependent RNA helicase YfmL
MNQKFEAQFKEAGFTKLAPIQEAVYPLLAAGKDVLGLAPTGSGKTLAYSWPLLESLLP
GDGVQLLILAPSQELAAQLTDVIRPWAKLLNLNVTSVIGGANVKRQVEKLKKRPEVVVG
TPGRLLSLANDKKLKLHTLTSIVIDEADELLGQEQTLTTCRELVSKAPAETQLAFFSATK
NQILDELHKWFGVEVTVVDVRQQDHTQGEVTHYLIETPTRKRVEALRKLANVDGFYGL
VFFKQAATLKDVAEKLRHHHVPLATLSSEQRQVDRQAALRALRKREIRLLLTTDLAARG
LDIKELPAVVNYDLPHSHNTYVHRVGRTGRMGAKGMVVNLGNEHDLRLFKQLIRKEAY
QFMTAYLYEGQIVTELPEKSSPKVKKPKKASKPVQAVTEKVRKHKKKRKRDQKNKGK
RRAK
>oxidoreductase, Gfo/Idh/MocA family
MLRLGIVGTNWITKRFVDAAHESGLFELRAVYSRRLEKAKMFAKDYPSVQDFQTDLEQ
FMASPSFEVVYIASPNSLHFSQAKAAIKAGKHVIVEKPICLNPKQMQELRELLAEHQGQ
FLFEAARHTHEPNYLAVKKAVAKLPVVQGATLTYMKYSSRYDAYLAGENPNIFSPKFA
GGALQDLGVYMVHCAVGWFGKPLNAQYFPTKLANGIDAKGVAILTYQDFTVTLNVGKT
ANSYLPGEIYGLKETIWMDNAADLKQVELRTPTATRTLSQTPAENELLAEAVAFGKVLA
NPDLPENKAQAQAWLEQTVTVNEVMYQLRQSSNLYFEDEN
>FIG00749603: hypothetical protein
MKRKSGIAVLGLSILAGVNKVIHTRKKAAITTTALPLIYAGTWFFTDQNSQHQHKLEVTM
DLDILLDGRKLAGTVEKIDERELLFLDNYGYHLRIDAIDAHPVSVYDEADNQIYRFTLQ
>FIG053235: Diacylglucosamine hydrolase like
MIDLHEILDRMNPNQKINYDKVMQKMTQRWAKENTRPKILLHSCCAPCSTYTLEYLTQ
FADITVYYANSNIHPRAEYQRREYVQQKFIHDFNEKTGNQVAFLAAPYKPQEYFQKVR
GLEKEKEGGARCEACFDYRLDKVAKKAVELGFDYFGSALTISPHKNSQVINKVGIEVQ
NFYDTHYLPSDFKKNGGYQRSVEMCNEYDVYRQCYCGCVFAAKDQGIDFSKIRREAL
AFMAGKDGARDFPEIRFNIQN
>hypothetical protein
MAFYDDYLAGIDNLEHRQKFAQVLKWVEAHYPNLEGRIAWKQPMFTDHGTFIIGFSVA
KAHFSFAGEAKIITVFEKEIQQAGLSYGKKLVRVGFD
>Na(+)/H(+) antiporter
MATSVFFGALVILLAVLAEKFAGKFKLPALILFMFIGMLFGSDGLFKISFSNYRLADQICS
LALVFIMFYGGFNTKWKVAKSSAPKSLLLSTLGVIITAGITTVLAVYLLHFKLLEALLIGAV
LSSTDAASVFAILHQQKVNLKYNTASLLELESGSNDPVAYLLTALIIGFITGAKQVNIPLEI
LLQVSMGLWWGFILAKLAVYLLTQTNLIAEGLETIFLMGMMLLCYALTEVTAGNVFLSIY
LLGLIIGNAQIANKQVIIPFLHGITSLAQILIFFLLGLLTFPHRMLEIIPTALIIAIIITFIARPLAV
FTLLLPFKAKINQCLLVAWAGLRGAASSVFAIFALAHGVSLNYDLFHIVFLVSLFSIALQG
SLLPWVATKLDMIDESDDIRKTFNDYQAESAITLMRLFVPKDHPWVNRPIKDVHLPSGS
LALMIKRNSKTIINKGDTIILGGDILILSVPAYETKGNETLNEINIDANHPWAGQKIATLNLP
DNMLIALIIRGEENLIPDGNTVIEPSDTVVLYH
>Flavodoxin

MILMNDGLVIYFSQTKRTEKIANLIGQAYKTSPYPIQVTVPYSSKDLDWTLPNSRANQE
MENPAARPDLKPLKLELTGKLKIFLGYPLWWSQAPRAINTFLDTYDLSGYKLYLFCTSG
SSSIDKSVRELSKTYPQLNIQAARRFANDTTLAQIETWLKSL
>Lipid A export ATP-binding/permease protein MsbA
MGPHGHQMGKKPQGVKNFWGTVKRLVNFMGPWKWSFLVVLIFAVGSVIFQIRTPKIL
GEATTTIFAGIMKGYQQMKMGQQLSAYPIDFTKIKHILLVVAVMYLISAALAFAQQVIMA
NVSQKIVYKLRKDLKEKMQRLPINYYDTHENGDIMSRAVNDMDNIAGTLQQSLAQLVT
SVVTFVGVFFMMLSISWKLTLIACITIPLGLIVVAVVAPRSQKYFAAQQKSLGLINNQVEE
NYAGHTVVKTFNKEAATIAEFEKENENYYRSSWKAQFISGFIMPLMNLVKNIGYVLVAV
VGGIQVAKGQVTLGNVQAFLQYVNQFTQPITQIANIANTIQSTVASAERIFELLDEEEMA
QVQAQADEPKFAANKVTFDHVKFGYSDDKILMTDFNLAVKPGQMVAIVGPTGAGKTTL
INLLERFYDIKGGAIRLDGVDTRKLSRAGLRKHFAMVLQDTWLFTGTIYDNIKYGREDAS
HEEIIAAAKAAHVDTFVRQLPDGYDTVLNEEASNISQGQRQLITIARAFLANPEILILDEAT
SSVDTRTEILIQKAMEKLLAGRTSFVVAHRLSTIRDADNIIVMQHGSIVETGNHSQLMAQ
DGFYADLYNSQFAH
>Lipid A export ATP-binding/permease protein MsbA
MIRLMKKYLDWWAVLGAVIFLGIQVMCDLYLPNLTSRLIDNGVAKGDTVYIWHVGGQM
LGIAFVGLIAALFNVYFAATQAQKMGMRIRSNIFKRVLSFSNKEIDDLGTSSLITRTTNDV
LQIQNVVIMMLRMMIQAPLMLLGAGFMAFTTERRLTSVFAVSIPLLLIAIALVMGSAVPLF
KKLQKQIDKINLVFREGLTGVRVIRAFNQDAFEQERFDQANKDYTLTGIKVFSIVSLMFPI
MTLILNGTNMGIVWFGGKLIGSMDMEVGKLVSFMTYAAMVLFSFMMLSMIFVFIPRAQ
AAASRINEVLQKENSIVDAPVSKQVKPQAQASLEYDQVSFRYAGAEKRVLEDLSFKMK
AGQTLAVIGGTGAGKSTLINLIPRLFDVETGAVKVNGQDIRRLSQADLHAQIAFVQQKA
VLFKGTIRSNMQFGKADASDEEIWQALEIAQAKDFVSKLDGGLDAVVEQNGDNFSGG
QKQRLAIARALIKDASIYVFDDSFSALDFKTDAKLRQALSADERISRGIVVIVAQRISTVT
QADKIIVLDEGKVAGMGTHEELKQTNPVYQEILHSQLREEEI
>Transcriptional regulator, TetR family
MPKQTFFNLPAEKRQRLLDAAQVEFRTYSFSEASISHIIERANISRGSFYQYFEDKADLY
AYYGQMMKKDYGAKLVKILQANQGDYFASFRSYINLIVDEVFQGDDALFYAKMMEARS
FEVFHKFHQKKAPRLEDERDARFIYRFVDKSKLNLSDFEDFKMLNNLVVGGIFHTIMQG
LVMKQAQGKADLERLKQNYKRMLDLLQYGAGKTEGKC
>ABC transporter, ATP-binding protein
MAVLELKNVTKTYKLAGGGSFQALKGISASFEAGELVAIVGESGSGKSTLMNTIGGLDS
DFGGQIIYEGKNLATYSKQELVDYHKKSIGFIFQNFNLIPHLSLLDNVAMAMTLSNVDQK
TREARAKELLKQVGLEAHIYKKPDAISGGQKQRVAIARALINDPDMIIADEPTGALDAET
TTTVLEMIKDIAKSGKLVLMVTHSERVASHCSRVLRIDNGQLLEDFRQEELVIEDSQNK
EQVAVKNMSWYNAFRLALLNMKAKFGRNALVAVGSSIGIMSVVLMLSLGKGVTSYVKS
TMNSYTNPRVTEVHKNGTSNQKRMATQTNGNSAAQAQANQDMLKALTGTGNDSGFS
QKEIKKLASLKDVTKAQKGFSTISLGTTTVKYKNKEASIVQLQTMSSAVTDSNIRQGKRP
GKNEILIDAGTADILGKNMIGKKVTLHLTVDSKQIEKTFKVAGTYSLSASGSSTATILARY
SDLASLYKETGTNLKPNVIYLTAKDKDATSRVKDKVKDLGYSGSMQEAMTDMFSDMLT
IITDVLTGIAAISLIVSAIMILVVLNISVVERTKEIGVLKALGARRKDIRRIFVSEAFLIGLSAG
ALGDLVTWILSLIANHFSQAMFEINVVAITPQYVISGLVISILISMLAGMMPANRASKLDP
VDSLRKE
>hypothetical protein

MTPTERRQALGHKAMNLYQVQMYQLEQSSELVTFEPDKIPTDYNYQTKIYVKNENLLY
RLHQFGSQVKVGCMNFADPLIPGGNFMLGVNAQEEAICRNSFLYPELKKYRHTYYQQ
NKLYPNDFFYSPALIYANQVKVIRDEREDRIVMGRLVDFVTIAAPNVTAMKAAKIKLNTS
RIYTDLENKILRTLRMFKNNAVTHLVLGAFGCGVFGNDPKMVAHAFKQVLARLEFQGA
FAEIYFDILENPVAEQAFRTTIL
>Bifunctional deaminase-reductase domain protein
MAAKIIISLLVTVDGFIAEENRSSNQLEEAFFIRQPAGVELRQPQDCDVILVGHNTFESY
RQELTVGSYSDLPVVVLASEPLKLTAAEQNVSCIKGEIKAVITYLERIGQVVWVMGGHE
LINELTSQNLADEIRLVTLPLEFGTGISLFNTNLPAPQVELASLKEYGDMTYSIWRKKDD
AN
>Xanthine/uracil/thiamine/ascorbate permease family protein
MNALNSFFNLDKLGTNPKREILAGFTTFISMCYILFVNPTVLSAAGMDKGAVFTATALSS
AIGCVLMGVIAKYPIATAPSLGINAFFAYSVCIGMKIPWQTALAGAFVAALIFILITVFKLRE
LIIDSIPQDLKYAISAGIGMFIAFIGLQGGGIIVKNESTLVSLGSFHGQVWVTVFGLIAIAIL
MVLKVPGAIFIGMVLSTIFGIATDLIKMPTAFVSAAPSLKPTFLVAAKHITDINSFQLVVVV
LTFLLVAFFDTAGTLVGLCTQAGFVKDNKIPRIGQALMSDSSAMLVGTLLGTSPVGAYV
ESSAGIAVGGRSGLTAVTTGILFIFGLFFSPLLAIVTTQVTAPALIIVGVLMAQNMARVDW
SKLEIAIPSFIIIATMPLTYSIADGLALGVILYPITMLAAKRGKEVPVLMYVLFFVFIGFFWVI
GRG
>DinG family ATP-dependent helicase CPE1197
MPYSVGVRELVEFILKSGDLSAGGLQSPNSALEGTKIHQAIQTTWPSETECEVDLKQTV
NLANQSVLIHGRADGLKKADNHYTEVLEIKSSAPSFADLKENTLTLYWAQAKIYAQLLM
QADPSLNKLDLTLIYVQVTTGYQWQKTQSITRKEAQTFFDQVTNEYAKWLQLKQDLAN
RQQEMLTKLKFPFPNYRQNQRELAAVVYKTILNQKKLFVQAPTGTGKTISTIFPALKAM
GQGKVNRLFYLTAKQSTRQVAQDTLNLLAQSEPLVTTITLTAKDQITFEEEKDLADEEN
PYLLGYYDRIKPALMDILRNEWLITKELVCKYARKHQVDPFEYSLDISLFCSVIICDYNYL
FDPMVYLQRFFSKPDNSNCFLIDEAHNLVARSRDMYTKEISLQTLTTLKDKLKQTSGHR
GLINRLQPIIATLTELKQTLVKHHQPYLVMEDELIALAKSCRKFCDFTSNWLREHPEEAL
AEEILTVFFDLYSYLKVGEFYDPSFRTKLSIEADDLLVKIFCLNPSKLLAKSMALGGSSIL
FSATLSPLAYYQEVLGGTDSLCYQLPSPFDQERLGLVITANIATTYRQRQASLDAVVATI
ASLSQSKSGNYLVFLPSYAYLTRSSPPLTNATPWLGPLFSSPK
>DinG family ATP-dependent helicase CPE1197
MTAKERSEFLANFKPTKESLVAFALLGGIFSEGIDLKGDCLNGVAIIGVGLPGISPENDA
LRTYFDHQEKAGFEYAYQLPGLNNVFQAAGRLIRSASDYGPVILVDSRFASSRYQRFY
PNHWSHGRIVYNTTQLQEYLASFWQKFTS
>LSU ribosomal protein L31p @ LSU ribosomal protein L31p, zinc-independent
MKQGIHPEYHEVVFMDSATGYKFVSGSTKTSNETVEFEGKTYPLIRVEISSDSHPFYTG
RQKFTQADGRVDRFNKKYGFTN
>Transcription termination factor Rho
MVEKNDLNDHGKRKPRQLKKQGTTDHYLGETQDEVALSDAERQAITEERRVAIAELQE
KPLADLLSEADRYGIDGADMSKQELVLAVAQAQADEEGRYYAEGVLEIVRPKDKDPFG
FLRLVNYGPSKEDLYVSYSQLRLGLRQGDWVAGVARPAKRGDKSGALLWVDTVNGK
RPDQVRHRPPFNSLTAIYPNEQMKLTTTPARLSTRVIDLFAPIGFGQRSLIVAPPKAGKT
TLLKEIAQGISKNYPNTKLIMLLIDERPEEVTELERTINGEVVYSTFDQRPENHVRLADFV
LERAKRLVEEKEDVVILMDSITRLARAANITERPSGRTLSGGVDPAALYRPKRFFGAAR
NVEEGGSLTIIATALIDTGSRMDDVIFEEFKGTGNQEVKLDRQLAERRIFPAIDLKASGT

RKEELLLQKQYLKAIWQLRDQFGQVADRRENSHNRTVDMAQTTRQVLTMLKAATTNR
KFCSNVLEQKAATSD
>UDP-N-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7)
MKKMIVKGGQTLAGEVTIGGAKNSTVALIPAAILADTPVKFDSVPDILDVHNLMLILKSLN
VTSKFGHGELEIDPTKIKEAALPGNAIRSLRASYYFMGALLGRFGQATVGFPGGDNIGP
RPIDQHIKGFKALGARVEEGENEVHITTGPEGLKGARIFFDLVSVGATINVLLAAVRAKG
TTIMENVAREPEIIDLATFLNNMGAHIRGAGTDVIRIQGVDRLEARNTHTIIPDRIEAGTYL
SFAAANGKGVLVKNVIPEHLDPFIAKLVEMGVNLEIDEDQIYVPGGDFLAPITVKTSPFP
GFATDLQQPITPLLMRAHGKSTIIDTLYPERIKHIPEMKRMGGKIESKDGVITVGNSPKLT
GAVVNASEIRAGVSLLGLALMAEGTTVIEHADHILRGYDRIVEKFRNLSVNLEVVTETV
>hypothetical protein
MEKTSQLSQLSKRIVEFEKKHHLNDTDFALLSRLSVEKVHALKAMYALPDEEEMQMLE
EAFASKD
>CTP synthase (EC 6.3.4.2)
MTKYIFVTGGVVSSLGKGIVAASLGRLLKNRGLKITIQKFDPYINVDPGTMSPYQHGEVF
VTDDGTETDLDLGHYERFIDINLNKYSNVTTGKIYSEVLKKERRGDYLGATVQVIPHITN
MIKEKIMRAGTVSDADVVITEIGGTVGDIESLPFLEALRQMKSEVGSDNVIYIHTTLVPYL
KAAGEMKTKPTQHSVKELRGLGIQPNILVVRTEEPISQEMRNKIANFCDVEPEAVIESM
DVDTLYQIPLNLQAQGMDEIVCQKLKLDTPQADMTSWIELRDRVQNLEGEVKIALVGKY
VNLQDAYISVNEALKHAGYHINSKVDITYINSENIDADNVAETLKGYDGIIVPGGFGDRG
LEGMIQSIRYARENNVPFLGICLGMQMACIEFARDVCGIEDANTSEVEPHGKNSVIDLM
ADQEDLENMGGTQRLGLYPCKLKPNTKTAQAYGNQDVIQQRHRHRYEFNNKYRELLA
SKGLTFAGTSPDNRLVEVVEVTDHKFFVASQYHPEFLSRPQRPEGLFQAFINAAANV
>DNA-directed RNA polymerase delta subunit (EC 2.7.7.6)
MELEKFKDANKNELSMIEVAHAILSQNGDVMAFADLTNKIQEFLGASDQEIRDRLAQFY
TDLNVDGSFISLGDNLWGLRSWYPFDSIDEALVHSDDDEDDDRPRKKRKKVNAFLAD
ASDDDVIDYNDDDPEDEDLDADYDSDDDDDDDADPELKAYSKDLSDIGDTDDDEDDD
LPDGIEGQLTELSDDDEELD
>DUF1934 domain-containing protein
MKEMCAERMIKVAELSNGIPVTIHVVTQNIQDNQVVDYDEVFEGQFFQMGASIYLRYQ
ETAEQEALVTFKITANGEVQLTRKTDEMNLRLYFVAGKQVNAKYVTPYGVVPVTSRTS
KLQAEVDDTLRGTIGQINIDYTLNSGQQLLGRYKIRLQFKG
>Deoxyguanosinetriphosphate triphosphohydrolase (EC 3.1.5.1)
MITIVCKNIKYLPKEKVLRDPIHNYIYIQHQVILDLINTREFQRLRRIKQLGTSNLTFHGAE
HSRFTHSVGVYEITRRICDIFVRNYPSHNPADGLWNDDERLVALCAALLHDIGHGAYSH
TFEHIFNTNHEEMTVKIITSEQTEVNAVLRQVSDDFPAQVASVITKTYPNKRVVQMISSQ
IDADRMDYLLRDSYYTGANYGSFDLTRILRVMRPAKEGIVFDISGMHAVEDYIVSRFQM
YQQVYFHPVSRSMEVILNHLLQRAKDLSRTHKLKSALGTYYLKPFFEGQADLEAYLRL
DDGVLNTCFEAWRDSDDEILADLASRFLDRHPFKSVKFTATTANLLDPLRQMVAQAGF
DPNYYTAENDSFDLPYDSYSSQTDKGQIKLIQKDGSLSELSAVSPLVLAITGKFLGDKR
FFFPKEMLGTDTEVDLLVNPAFRTHLKNDHLI
>HMP-PP hydrolase (pyridoxal phosphatase) Cof, detected in genetic screen for thiamin metabolic genes (PMID:15292217)
MTIKMIATDIDGTLVNDNKEISPKTLAALKAARKKGVFVVLCTGRPVSGVKPYLEQLGLT
QDENYVITYNGALATNIASGQVVVSNTLDYSDFTKLDSLSHELNVHGQAIQPSSDLYVT
TPDISYYTVLDSFYTKMPLHYRPQAQIPTEFPMAKYMWADEPAAIDQALPNLPTAIKED

YYTVRSEGWFFEFMHKDAHKGKTTLELANKLGIKPHEILTAGDQQNDLTLIQAGFGVA
MGNAIPEIKDLAQYVTADNNHDGLGLAVEKFVL
>hypothetical protein
MRKLQTKLWLKLDGAVLLLALATALCGRWLGVAVQVSNVMFLVGLLLICLMIINILLHAS
LLAGWFKKKQAGESDEEYQARKIDVKSVASKKNKPIHFQSFAANCLVIGLGLIILAIVITI
>Hypothetical membrane spanning protein
MSAITLLSLILLILMLLVGGKHGLSSFISILLNFALLFIAVILIAGGFPPLGVSLFIGICILAITIY
MGNDDEAATNTAFIATILVVGVMLVLIVPADYLAQVQGFGNEQSEEIEAFNLLIGVNFES
LLIATTVLATLGAIAEASIAVASGLEELLSQKPNLTLAQLYKSGKAIGFQIMAMTFNTLFFG
MFGGDLALFILLARLNASFGYYLNAKIFVGECLAVLFSALAVILVIWITIYLTGRKLNTQR
>Hypothetical membrane spanning protein
MIIIKKSWVIEIVVILLSGLSLYWVTSHDARLYHQDLGKIISVKTGPKQVVKDEYNNQDYQ
VKQRLKIRLLNGTKRGQEMVVTNTYSGSKAFDQRLKLGQEIFVDFHHEGGKLTASFSH
SKRDTILVMLAWLVIVLLYITMRFQGIRALASVVANFVIFLIFVKLDIALNLTYFFWLFALSA
LIFTALSLVMIMGFSKQCLVTFSSIVLGTSLGLGLGYFALWLTNNQGVHYEALDMATQS
PEQLFFAATLIGLLGTVMDAATDIVATLFELKESKPELSRRQLFLSGRQVGRAIMGPLIN
VLFLIFFAETFAMAVLYLRTGNTFAYTFEWTMSLGVVQSLISGIGISLVVPTASLLSAYVL
GEE
>Biotin--protein ligase (EC 6.3.4.9, EC 6.3.4.10, EC 6.3.4.11, EC 6.3.4.15) / Biotin operon repressor
MKEADLMSTSSQLLTILGSQSQPISGQELADQLNISRTAIWKAINNLKDAGYQITSQPRT
GYFLEDNGYLDVGLISRYLPTNFNFPLEIHQTIDSTNIRAKQLANKSDLTTPYIIIANQQTN
GYGRYGRNFTSPSQSGIYLSILLANEQTEFNPGLLTTVVALAMCRAIEKKLAASPKIKWV
NDVIVDGKKVCGILTEGISNLETQSLSNIVVGAGINYLTADFPEEISQRAGSLRNYALKAK
VSRNEFIATYLEEFFKLYQTYQTGDFMPEYRAHSNIIGKEVTITQGTKSFQGNVVTIDDD
GAIVLADGRKFTSGEVTKIRAN
>Glucosamine-6-phosphate deaminase (EC 3.5.99.6)
MRVIIVEDKVAGGKEACKVFANALANGAKVLGLATGGTPETTYQELVASKLDFSKLVSV
NLDEYVGLGADDPQSYNYYMKQHLFNQKPFAKSYLPNGLAEDEAAEIERYNQILAANP
VDLQLLGIGENGHIGFNEPGSPLDAKTRKVALTESTIKANARYFEREEDVPRYAYSMGI
GSIMQAKEILLEAFGAKKAQAVKALVEGPVDANVPATILQKHPNVTVIVDKQAAALLTKK
>ABC transporter, ATP-binding protein
MITVSNLSLNFSGRLLYDDVNLKFTPGNCYGVIGANGAGKSTFLKLLEGKLEPTSGSISI
GANERISSLSQDHFGFEEYTVLETVIQGHKKLYQIMKEKDALYAKPDFSEEDGLRAAEL
EGEFAEMDGWNAEADAATLLQSLNIPDSMHDQKMSELAERDKVKVLLAQALFGQPDIL
LLDEPTNGLDVHTISWLEDFLADYPNIVIVVSHDRHFLNQVCTMMCDVDFGKIKLFVGN
YDFWLESSQLAAKLQANANAKKEEKIKELQEFIARFSANASKSKQATSRKKQLEKITLD
DIKPSSRKYPFIKFEMERELGNDLVKVENVSKTIDGVKILDNISFILRPGEKAAIISRNDLA
TSVLMQILAGELEPDSGTVTWGQTTSRSYMSRDLNANFAQSDLSIIDWLRQFASKEEN
DNTFLRGFLGKMLFSGEDIEKQVGVLSGGEKVRCLLSKMMLQKANVLLMDDPTNHLD
LESITALNDSMVSFPGSIIFTSHDHQYIQTIANHIIEVSAKGVVERLDTTYDEFLAHETVQ
AQVAALYE
>transposase
MMEQYRLACNYVSEYIFNHDFDMKQSRLNKALYTNLRSLFMLKSQMAQSVIRTVIARY
KTVKTQMIRNPYKYQDINTGEWYRETRDLTWLQKPIAFNRPQVDLQRNRDWSYLGSG
QLSINTLDGRVKVNPICHGFDQYFDGTWKFGLAKLLKTGGKWYLHISATKTVADFDNQ

TVKHVVGIDRGLRFLATTYDEQGQTAFFDGKIVMRKRAKYQKLRAKLQAKGTKSAKRR
LKKLSGRENRWITDVNHRLSKTLVQKFGSNTLFVLEDLTGVSFERTDLPKALRNQNKS
WAFYQLAQFLTYKAHLNNSEVIKVSAQYTSQRCPKCGTIKKSNRNHDLHEYHCTNCGY
RSNDDRLGAMNIQLLGTQYISGQEQPKFELTTNT
>FIG028593: membrane protein
MEFSQHTLENQKLLGYFVLSIFLNSLGNGLTVAINLGSALWTAASVNLNHAFAVSLSLM
MFLPAFLVIIVNSLILGKLSPKRIIGNLVFLTPFSLLVGYFSAGLSQLGINHLNLLVRTGLD
CLGVVLIAVAISLYQRVNWMLHPVDDLMQIIRFKYFKGNATIAQLVVFMPPVIAIIVAVLV
SHHIYAVNVGTIFALLFQGSLVGLADKYLFPSLKHQNLDK
>Phosphinothricin N-acetyltransferase
MTPIAFRIAHSSDAAEILQIYAHYVEQTAITLDTVVPSIADFQTKITTTLERYPFIVAVQAN
KIVGYAYASAFNYRNAYDWSAELSIYLLPTMRKQGLGSRFYSLLEDIAKAQNITNLNACI
TYPRTKDDPHVTLASVNFHEKRGFNLCGHFHKCAYKFDRWYDMVWMEKQILPHSDS
PAPIKPFNQIAKQFFS
>hypothetical protein
MSENVFFNPGQAIASDFDFNKAYVAAQIYHHKAKKPVLVVQEKDGQPFVIFDEQAALD
SEKEEAKRYSLVKRVTESD
>hypothetical protein
MVQLVRITDDNLADDIYKEVTAAGVDFFSSDKQYLWAITFRGKIVGLVSLYNLSRTSADL
GYWIKKTARRQGYAKQAVAQVVFLAAHDLGVQTIHAVVRRDNLASISVIRANDFELSLK
MGAYCEYVLKLASKKAGK
>Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)
MSEQNAGSNLKIFALNSNKPLAEKIAAAVGLPLGKTSVDRFSDGEIRINIEESIRGDEIFIV
QSTSAPVNDNLMELLIMIDALRRASADKINVVIPYYGYARQDRKVRSREPITAKLVADMI
EKAGADRVIALDLHAAQIQGFFDVPVDHLMGAPLLADYFLTNGLEKDAVVVSPDHGGV
TRARKLAEFLKAPIAIIDKRRPKANVAEVMNIIGSVDGKRCILIDDMIDTAGTITLAAQALK
DAGATEVYVCATHPVFSGPAIERIEKSPIKQMVVTDSIQLPEEKQIDKLVQVSVGPLIGA
AIKRIHENKPVSPLFKNRFRGIEADQD
>N-acetylglucosamine-1-phosphate uridyltransferase (EC 2.7.7.23) / Glucosamine-1-phosphate N-acetyltransferase (EC 2.3.1.157)
MTAKYAIILAAGQGTRMKSKLYKVLHPVCGKPMVDHVLTQIEKNDIDEIVTVVGHGAQK
VEATLGTRTQYVLQAEQLGTGHAVLQAEALLGDKEGMTIIACGDTPLFTAKTFEELFEY
HKSKGAVATVLTATTDNPFGYGRIIRNEIGIVEKIVEQKDATKEEAAVTEINTGVYCFDN
QELFHALHQIKNDNAQGEYYLTDVMEIFKKEEKIVAAYRMKDFSESMGVNDRLALAKA
TAVMQKRINETHMKNGVTLIDPATTYIDVDVKIGADTVIEPGVQLKGHTVIGSDCHIGAH
SEIRDCVLKDGVTVTASFLEEAVMHTGSNIGPYSHLRPQADIGEGVHIGNFVEVKKAQI
GKNTKVGHLTYVGDATLGKDINVGCGTVFVNYDGINKHHTNVGDYSFIGSGSNIIAPVQ
IADHAYVAAGSTVTEDVDAHDMAIARSRQVNKKGYFDHYPVAQAAAEAEKAENK
>PurR: transcription regulator associated with purine metabolism
MKTRRSDRLIDMTRYLLERPHTLISLTFFADRYESAKSSISEDLSIIKRTFKYRGIGILETIP
GAAGGARYIPSISEEEARKFITDMTKELSEKDRLLPGGYVYLSDLLGRPDVLRQVGRVI
ARQYIDKQIDAVMTVATKGVPIAQSVSSYLNVPFVIVRRDSKITEGSTVSVNYVSGSSE
RIEKMELSKRSLKRGSHVLVVDDFMKGGGTVNGMKSMIEEFESKLVGITVFAEGAFEG
HRMVDEYTSLLKVDKVNTLDKAIHVSPGNYLEKVFEQTDSKAD
>membrane protein, putative

MTISKKDFTLNILNGISLGVVVTLVPAALTGQLMKALLGVFPLGAAVISMTNFTMSLLAAV
SGFCVGYLFKFNMIQTSAVAAAAMIGAGVIKQTIQGFVLSGTGDVINTALTITVAVLVAQ
FIGNKFKNYTVLLLPLLVIVIAGGIGLTTLPFIKMVTTFIGMAIKNLVTLQPILMGALMGMVF
AVLIVSPISSVGIATAVGLSGIAAGSANLGIVAGSFTLAIMGSSVNSLGTTLAHFVGTPKI
QMANMLEKPFLFIPVMINAAVLGAIGAIFNIQGTAMSAGFGFSGLIGPMAALGAMKVDG
AALVLVTILFLILPILLGFLAKYVFVKKTKRIVPKDLLIQV
>Polysaccharide deacetylase
MSKGKIILWTLVALVLVTGIGFACKKYADNQHQKQVAKITAMEKKASQKATKTNKQAKT
SSKKAKVTWVKEAKRVSFPILMYHSISESNGNTLRVPANEFDQEMKWLKDHHYYTLSD
KEAYRVLSKNEVPAKKIVWITLDDGYRDNYTNAFPALKKYHLKATINLITDQIKNGQANK
LTMDQIKEMQASGLVSFGSHTVSHSDLAALSPSAQYAELKQSRTWLNKNLHQNTIYAC
YPAGKHNADTLAATKRAGYKLATTTIPGLANNEGGLYGLSRVRVEPNLSLSGFQSLLE
NN
>4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (EC 2.7.1.148)
MEILEKAPAKLNFSLDTPFRHADGNHEWQMVMIAIDLADYVQVRTIPTSKQITVLTDSG
FLPCDRRNLAYQAALKLQQDYQVDQGVEIIIKKNIPVAAGMGGGSSDAAAVLRALNQA
WNLKLSKAKLAQIGLEIDSDVPFCVYSEPALVTGKGELINPLGDFPSLWIIVAKPKASVS
TPTILKQIDYSNLAHLDVTGLVSALKEQDYAQITAKMGNVLEPITSKKHPDILRIKTKMLK
FGADAAQMSGSGPTVFGITQKYSRAQHIYNSLRGFCDEVYLVRPLSATSKKEGS
>Methylated-DNA--protein-cysteine methyltransferase (EC 2.1.1.63)
MDYTCVTIADHSYYLAASSTGLCFVGSPDGIFQELVSFYPRATLTPTTSNPILLAAQAQL
DAYFKKATHSFNVPLDFTVGTAFQQTVWQQLTQIPYGQTLSYQELATRLAQPNASRAV
ASAVAKNPLLIIVPCHRIIRADGSLGDFRAGKAFKQQLLELEQADLNSK
>SSU rRNA (adenine(1518)-N(6)/adenine(1519)-N(6))-dimethyltransferase (EC 2.1.1.182)
MDNNLPEIASPTRTRAIMETYGLTFKKSLGQNFLTDINVLKKIVAAAEVSEADDVIEIGPG
IGALTEQLAKSAHQVLALEIDDRLLPVLKETLAPYPNVEVIHQDILKANLKELIAEHFDGQ
HRLKIVANLPYYITTPIIMHLLATELAFDQIVVMMQKEVANRLAAQPGSKDYGSLSIGVQ
YEMDCQIAFIVPKTVFVPQPKVDSAIISLTKKKAKEYQPQDERFFHQLVRGAFMHRRKS
LWNNLQGLYGKQAEVKEKLSQALAQAEIEPSVRAERLTIADFVRLADCLLATGVKVK
>Ribonuclease M5 (EC 3.1.26.8)
MTEEKMKIREVIVVEGKSDTKRIQLAVNADTLETRGSAISEETLEQIEELHEKRGVIVFTD
PDYSGEKIRKTIREAVPGVKHAFLNRRDAAPNHKGSLGVEHASLEAIRAALKNLYTEDV
AAPIQISRADLALAGLISGPEAKKRRRLLGEYLHIGYTNGKQLYKRLNLFQISPEEFAAAL
KEINEGEQNG
>Putative deoxyribonuclease YcfH
MKIFDSHTHINARFFAGDEYDAIKRAQAFGVTAMLLVAYDDYSATQMLKLCEQFPDIYG
AVGIHPDDALNYDDAKENWLETILAHPKVKVLGEIGLDYHNEVDHQVQKEVFQRQLQL
AAKHQLPVTIHNRDAIADTYPILKASELGRFGGIMHSFNGNQDWAKKFLDLGMYLSYS
GVVTFGNAKEAKQAALSTPLDRMLVETDAPYLTPLPFRGRANEPALTRYTVEFLAQEL
GLSAASLASRTYQNTERILRIND
>Methionyl-tRNA synthetase (EC 6.1.1.10)
MAEKPTFYITTPIYYPSGKLHIGNSYTTIACDVLARYKRLQDFDVFFLTGTDEHGLKIEQK
AKALEMTPQAYVDKMAADIKALWKLLDISNDKFIRTTDDYHERAVQNIFEKLLAKGDIYL
GKYVGWYSVSDEEYFTESQLAEVYRDENGKMIGGKAPSGHEVQLVEEDCYFFKMSK
YADKLVEYYNAHPDFIIPENRKTEMLNNFIKPGLEDLALTRTSFNWGVKVPSNPKHVVY

VWIDALCNYITALGYETGDDESLFKKYWPADVHMVGKEIVRFHTIYWPIILMALDLPLPK
HIIGHGWLLMKDGKMSKSKGNVVYPEMLVERYGLDALRYYLMRAVPFGNDGIFTPEDF
VGRVNYDLANDLGNLLNRTVAMINKYRAGQVPAFVSGVTAFDADLESLAAETIKAYEA
EMEAVKFSNALDALWKFVSRSNKYIDETEPWALAKDEAKAQELDSVLAHLVASLRLIAV
LLQPVMTKTPKEIFHQLGLGENLPAIKGLAYADLAVSQVIAKGQPIFPRLEVEKEVEYIKS
KMTANEKAKGRKAMAQKAAQGFDPTTTTLKLTKDAIGFDKFDKVELKVAKIKEVSKVE
GADKLLKFRLDAGDEGDRQILSGIAQWYPNFEELVGKNVIAVTNLKPRKMRGEVSQG
MLLSAEFGETVQLITVPENIPAGSLVG
>NAD-dependent protein deacetylase of SIR2 family
MTQLNDLLTTAKRIVFLTGAGVSTASGIPDYRSKNGLYANKRRPEYLLSHTCLVKEPEV
FYEYLKANLYYPKAKPNIIHEKMAAFTKAKRATIITQNIDGLHAKAGSDLERLVEFHGNLY
DLYCQKCGQTVTYKTYLENMYHQGCGGILRPDVVLYEEGIKQARLNQALTALSQADLV
VVCGTSFVVYPFAGLIQYRSPKAKVVAVNKEVISLPSDGIMIQGDAVKEFAQVDNP
>hypothetical protein
MVEVVLNEAVTDLVKQVQEIHPGEINFKYHDKKVGYLRHDQAQEYVDNGQITVEVFDL
TNPNYTIAHELLHFLLQFKQLPRISFNVTSGNQALDERIMVTATELYDSLLHFRVYEQQR
QLGLVDESVEELYFKGVLAILKPEPKDQRDSWMVLRTLTLFDCLIFFKGQEESILPKLQE
LYPESLQAAQKLYQIATAKDLESAFALRRAIVKVFAAFDEQLQAWNLMKMNLNEFVTLT
AVLSERQSRLQIAQVFEIYHSVFQDNTVFKDAYIGRYQGDQQNTFVFTVPKGQDKEVF
AKVYESKVSDYLQANGFYYLKR
>Hydroxymethylglutaryl-CoA reductase (EC 1.1.1.34)
MSGFKKYYQKPWADRLQVLQTEQLLTADQLASLKAAAVNPELGETMIENFITSYSLPE
GLALNYMIDGKDYVVPMVTEEPSVIAASSNGAKLVKQAGGFKTNLTDRLMIGQVVLEN
VVAPTELAAKLKKAEEQLLQVANAAHPSIVKRGGGARFIRVRILAADLISLDLAVDVKQA
MGANMLNTMLEAVADYVRTNYQQEILMSILSNYATSCLVTAHCQIPLALLATKQLAGKE
VGQKIAQASRLAQLDVYRATTHNKGIMNGIDAVLIASGNDWRAVEAGVHAYAARDGQY
RGLSEWQVVGDNLEGKLTLPLPVGFVGGSIGIVPLVKLNHQLLQLESAEELEKICVSVG
LAQNLAALRALVTEGIQKGHMHLQLKSLAVSVGAKQTELAPLISALESAGVRDSNSAKQ
LLAELRRKED
>amino acid ABC transporter, amino acid-binding protein
MKRFLTLLLATCCLFLGLTSLTATANSLKVGVLANNPPYTIQNHSGFDLALSQQLAAKV
KVIPQKSNSQLITALKKGRVDLIICDVSLLSKQMSHSQSFLHPANVLFTRHDSKTKSILKL
NQKKVGYLKNNPHHALLTTLATKPQGFATESALLTALNTGKIKAAILTDYQYNQLLAANP
QLVTAQDETDQKQVGQVLTKISDPKLTSQSLVVASFKKPRLQKKINHKLKQLQDNGTL
SKLSMKYFHQDWTYQ
>Exopolyphosphatase (EC 3.6.1.11)
MENVVILDLGSNSTRLSINEVTNGSFREVRRQKVMTRMAEGMGKTLGEKVLKKAAIER
TLAALTDFKKEYQDLPNLKVKGIATAAVRAASNSQAFLNQVKELTGVDVEVLSGDAEAY
YDYVGVKNSLKLNDFILCDMGGGSFELAVVKDGQAVNLASIPFGAVSLTERFNAQDLL
SAKDLFRFQGFVQGLFNEQLWLLEGRYLPLVLLGGANRTVARTEMMRRGQKNLDEFH
GVKMSARSFLNIYTDWLGLDFSQRQERLGVEADRADIILGGLTPIAQLIQQLDIPELIFSE
SGVREGILYSMVTE
>Tryptophanyl-tRNA synthetase (EC 6.1.1.2)
MTSKKVILTGDRPTGKLHLGHYIGSLRNRVEMQNSGKYESFIMIADQQALTDNARNPE
KIRNSLIQVALDYLSVGLDPAKSTIFVQSQIPALAELNLYYLNLVTVSRLERNPTVKAEIQ
QKNFERSIPAGFFTYPVSQAADITAFKAELVPVGDDQEPMLEQTREIVRTFNSIYQKQV

LVEPEGVFPPKGMGRLPGLDGNAKMSKSLGNAIYLSDDADTLRKKVMSMYTDPNHIH
VEDPGQVEGNMVFTYLDVFDDDKQKVAELKEHYQAGGLGDVKIKRYLNEVLEAKLAPI
RAKRAEYEADIAAVYQMLKEGSAKANEVANQTLAEVRAAIGVNYFD
>Cobalt-zinc-cadmium resistance protein
MRNIDRHLEKEYQLWEKIQSQELAKLTAALHHLFLNLGAYFLISIIEYVLAIISHSQTLRAD
AFNNLSGIISTILLMVGLHIARDIDDDDIAGLTPMPKISLHKRGNDQRVQFTRFRYETIFTL
VTGIIMVAISLGIILAGISNLLNPAKHLIPDRMALLGAAIASVIMLGVWYYNKQAGKKLQNA
ALIASAKDSLSDAWMSIATFVSILGALVFKITWLDGAGGVLVGLFILYSGISIFWESSLNL
ADYFDPEAEAQYRQVLATFPEIIEVDELKAHYNGNIITLDLVLVVDSQMTVLSSYQLAEKI
EQLMMHRFGIIDTDVSFIPTDKTKF
>DNA helicase
MHPEQPAEQVRVDKVEQVIAERLTTAELAHAKAKAELSSVERNYVQNARINTLEVDDR
METNAEVQQQKQLVAKNVADEEILNRQLTTLKQLKKSPYFGRIDILETGESQTETLYIGT
ASLVDDHQNFLIYDWRAPISSIYYNGTLGQVEYETPGGIQSAELKKKRQFTIVNGQITN
MFDTNETVGDEILQSVLGQQSDEYMRNIVATIQREQNDIIRDTEHDLLLVQGVAGSGKT
SAILQRIAFLLYHSRTSLAADQIILFSPNRLFSHYISEVLPSLGERNMRQLTLAEFLSQRF
QGLQVESLFERYERKESLSLGNQALQAKKEAATFMSELEKYVNALTSSKLCFTDILLDG
HVFFSKEEIYETFASLPAAMKITHRFLETKNLLIKRLKHRIIQEAQSEAVLEQVDSLSDAK
YRKYLGTKRRGAFQEVAAEVEYIAKQIVKERFAVIYDAIYNNHFIDIYEQYAAFLTTHLPE
AAATFNQNLEYHRLALEDAAPLLYLRDLITGEGQNNTIAHLFIDEVQDYSSAQLIYLKHV
FPKANLTLLGDGEQALFNPLMQAQTMLTKLATSLAAKKPRLIALNKSYRSTSEITNFMK
ALLPDGDKIIAFSRPGPKPKLIFTQNAAESLASLKKELTLRLTDSKQVALITKNLAESKYL
YEQLHREFTTHLLTDGDRALPSGLIILPIYLAKGLEFDNVISYDVSATNYPDEKSIGILYTI
CSRAMHNLTLITNGHIAPILKKVAPHLYTSQHTFKLS
>tRNA-Thr-GGT

>sigma-70 factor family protien
MENTRSEEKKSAGDIALKAYKKDPCGSNLLTLFAAYRPLIQTVINRFSFNLLDEDDLIQE
AQIVCFKAAQTYQESFNTKYGTYFIQCLVNRYNSLTRFDGADKRRNQKRDLSLEMLLE
AGEEKILGYVSTNEGSISDELNELITKQAPVLLSKFELRVFKLKILQDYSAAEIAEMLEVT
PTAVYNALGRCHRKLEGYIQGGDYYYPSSKEEAIKKPSPNSDEDNTSSS
>tRNA-Gly-GCC

>tRNA-Val-TAC

>5S RNA

>Large Subunit Ribosomal RNA; lsuRNA; LSU rRNA

>tRNA-Ala-TGC

>tRNA-Ile-GAT

>Small Subunit Ribosomal RNA; ssuRNA; SSU rRNA

>Tyrosyl-tRNA synthetase (EC 6.1.1.1)

MVNILDELEWRGAVNQVSDEEGLRKLLEDKSVGIYCGVDPTGDSLHIGHLIPFMMLKR
FQEAGHRAHIIIGGGTGSIGDPSGKKSERVLQTMEQVHHNEEALTKQMKHLFGKDNITI
VNNREWLSKIDLLGFLRDYGKLFNVNTMLNKEVVASRLEVGISFTEFTYQILQSIDFHHL
WTENDVQLQIGGADQWGNITAGIDLIHKMEGSDAKAYALTIPLMLKADGTKFGKTAGG
AVWLDPEKTSPYEFYQFWLNQDDRDVIKYLKYFTFLSKEEIAALEEKVKTQPEKREAQ
RRLAQEVVAFVHGKEAVKEAEHISAALFTGDVKDLTASEIEQGFKNMPSVEVTNAKTNL
VEWLVDVTKIEPSRRQAREDIKNGAIYINGQRVQDVDYELDPSQSFDGKFVIVRRGKK
RYFLARVK
>Peroxide stress regulator PerR, FUR family
MLRNQEMLVTAARTLKERSIKNTPQRQVILAYLMNSKAHPSIEMIHQHVQESGFSVSLA
TVYNTLELFLKNKLIIEVTPDSEGHMRYDYFERPHYHVICVNCNKIVDVYNDTFTELEKE
AAQTTGYQVYNSQYEVYGLCPDCQKKLVADS
>Membrane-bound metal-dependent hydrolase YdjM, induced during SOS response
MQYKTHVTTSLALGLPLIVSSQQLSALNVGALLVGCLLPDIDQPTSFLGKRNRLVSGVT
HKTLGHRGGTHSLLGLVIVFVMVAFISQRYFNFAGRYFPFWLCTGYLFHLLEDSFSKD
GVKWLWPLGKKGIRTGGKWLTYSTGGIGEYLLLGFTFCLLLIEIRLIWLNQLGNFLPSG
MLAHLQALTLKLQTIFSSN
>Acetate kinase (EC 2.7.2.1)
MPKILLLNAGSSSVKWKLFEIDNEAVVAKGLVERMLTPEAKWEISYGKESAQETIPNLT
YEQAAQLIITKLQELEITSLADIKAVGHRVVAGGHTFKQATLVDEEALDQIKQLSSHAPL
HNPMEAKYIELMQAVMPKTPQYAVFDSRFFTDMPEVNSIYSLPYELTEKYGLRKYGEH
GISHTYISKRAATLLGKDLANLKMITLHLGSGSSVAALKDGKAYDTSMGFTPLTGLTMG
TRAGDLDPALVPFLMAELNLDAKQVLELFNSKSGLLGLSGLSADMRDIDKAIAAGNKRA
ELALEIFINRVVKYVGSFYAELGGLDVLVFAGGIGENNRALRQKIAQKLACLGIEIDDNLN
QELAEGVISPADAKVTTLVVPTNEELAMVQEIKAVM
>Ribose operon repressor
MTTVSQILNGKGERFSQATQDKVRKIRDELGYVPDFNARNLIMGSAKTIGIIIPNITNPFF
ATFVRGVEVAAQRGDFIPLVLSGNGNPDLERQYVEQLIQRTVDGLIIASPAVNKDTIDHF
LKANGVPYLLIDQNPLDDGDRVSTDDYLGAKLAVEHLIELGHKRIAMLYCDKPSVNQAQ
RLQGYKDQLQSAGLELTDDLIVTGPLSKLGGYQASEQVLQTKASAVFAVNDEMAIGLY
RGLNERGLEVPADLSIVGYDNIDLSEYVNPPLTTIAQPVFEMGQVAFRTLLQRIEQPDA
PAQKIKLPVELVVRESSRKV
>transposase
MVLTTGGIALTLGSLEPTFYPFASLYCVFTEQKCYNISKESEVNLMEIQITIKVKLNLANA
EITSSFTNTMEQYRLACNYVSEYIFNHDFDMKQSRLNKVLYTNLRSLFMLKSQMAQSVI
RTVIARYKTVKTQMKRNPYKYQDINTSEWYREIRDLTWLQKPITFNRPQVDLQRNRDW
SYLSSGQLSINTLNGRVKVNPVYHGFDQYFDGTWKFGLAKLLKTGGKWYLHISATKTV
ADFDNQTVKHVVGIDRGLRFLATTYDEQGQTAFFDGKIVMRKRAKYQKLRAKLQAKGT
KSAKRRLKKLSGRENRWMTDVNHRLSKTLVQKFGSNTLFVLEDLTGVSFERTDLPKAL
RNQNKSWAFYQLAQFLTYKAHLNNSEVIEASAQYTSQRCPKCGMIKKSNRNHDLHEY
HCSNCGYRSNDDRLGAMNIQLLGTQYISGQEQPKFELTTNA
>hypothetical protein
MMAEKKFTKAELAQFDGQNGNKAYVAVDGVVYDVTNVAAWKDGAHHGNKAGNDVT
AGLGHSPHGERVLKKLEVVGKLVD
>TyrA protein

MAIDPKKAQDVWKDIGEHVQFTVLKLKREDQAHELEVIREFCDRYQAIIRSLKIRDSNP
QAESRLKVTFGFSNAAWDYLFPNGPKPKELETYTELKGPKYKMPASEGDLFFHIRASN
EAVVYEAQTQFMRFLRDITEVVDETKGFRYFEGRAIIGFIDGTEAPEVEKAADYALIGDE
DPTFENGSYAFAQKWLHNMDFWNKLKTEDQEKAVGREKFSDIELEDEDKFKNAHNVS
SKVEIDGEEQKIIRMNVPFSDPASGKTGTYFMGYARHWRVTKMMLQQMLDTSDFLLTF
SKILSGQLFFIPSKDTLDDITDEKI
>Substrate-specific component YkoE of thiamin-regulated ECF transporter for HydroxyMethylPyrimidine
MQTTNQKKKAWSLQDVILLAIIAIFFGIIYQLWGYTYYALAATPLKVYANDATLGVWIMA
GPLAAILLKKAGASVIGEVLAATVEMLLFSSWGAANLISGLIQGLASELGFTLTAYKNWG
ALSLTLATITSTIVTFAWDLFQSGYAAYNFSLLVSLFIIRLLSIGFFAGLLVYLIKALLNKSG
LVR
>Duplicated ATPase component YkoD of energizing module of thiamin-regulated ECF transporter for HydroxyMethylPyrimidine
MTASLITKNLSFAYQEDGPLIIDKLNFELQPASFNLLTGPSGCGKSTFFKLLAGLYPQYG
GKLLSGQVLLNNQAVAAIVPYERSRYVGMLFQEPSRQFALPTVAKQITFALENIQTPHD
QIEPKLNQVLTKLNLTALKDRPLFQLSGGEQQRLALATTLALDSQIILLDEPFANVDPLG
KQTLLADLKKLQLEAGKTIFITDHDTNGYYGLVDNLYRFSKSELTPLPLSSLTERQEQPL
TYTPLNSGQLTWQDVTVTVQNRQLFKTDSFTLPQGQLGLLSGPNGVGKSSFFKALTK
QLAFNGNIAYRKRDSRRYKAKSWAKIVALVFQDSDNQFVRLTVREELALSKKHSLAKG
YWTDAKITQALEVLDLTAYLDHSCYQLSGGQKRKLQVLSMLIMAQPVLLLDEPLASLDT
SSADKLLTLITETCQKLKLSILLISHQRSCLPGYVTYELVLNQGRLALRRDFDA
>Transmembrane component YkoC of energizing module of thiamin-regulated ECF transporter for HydroxyMethylPyrimidine
MLNPAGLTLLIIGLGLELAFIKSVSLNCLLAGCCLLYLLCYRRVKLTTIALLLLVTFPLALGS
WWSFLAFGRTNNWHNAWIYGSRIYAYLLLGASLTLTNSVEDILFNLHVHLKLPATFVYG
LLAAFNLLTSIKNQFATIRYSAQIRGRNYHLWQPELYLRIILVALKWSADLAQAMTAQGF
SEGYPRTRPSAKPFPKWQFYFIGLCLISYAWGGISGSWW
>hypothetical protein
MVYGATTMKQKDTWLTRFLILLLGVGCLLVSGYLLVRIFIV
>Aggregation promoting factor
MKLNKLLLSATAAAGFFAATTAVANADTVTVQAGDTVSEIAGKYNTTVEEIQNLNQLEN
VNLIYVGEELVVNTNGQAATTASSEAPAATATVSSQAPVQQAPAQPVAQQAPAQQQV
ASAPAQQATSQQTYTAPAATSSDEASAREFIAARESGGSYSAANGQYYGKYQLSRSY
LGGDLSAANQERVANNYVQSRYGSWTAAKNHWLQNGWY
>Large-conductance mechanosensitive channel
MLKEFKEFIARGNVLDLAVGVIVGSAFTAIVKALVDYIINPFLGLFLGSIDFSAFVIKVGSA
SFKIGSFLNAVINFLIIAFVVFLIVKAVNAAMPKKEEPAEEKVDPQVELLSEIRDLLKK
>hypothetical protein
MRSEDYLKKLTSYQIKMANLQCGLERIAQARQIIEAYQRQQLPQEVKTKLLFSEEELFN
HPELMALEDELLNLRQVVIEDFGIWHIFSQEWVADLKEYVGPGLGLELMAGNGVLRASI
SELIATDNLDWAGQDNESPQAWQRVEQLDAISALERYGLKANYIILAWAPDTSQVDWE
ILNWLRTKHWSGKFIVIGERNGATNSSKFWQEANLLKPRLLNQHHQAIDFINDEVFLVK
>hypothetical protein
MFKYIGGILLAVLILAGSAYYYHASAQAKEVKKETAFINNAQQTFTKAKTFSDKLASYQN
LKERYDYYKANGKNNQEVLSKYQQILAQDQHYFTDKLDKLLKANTFSDDQLSELSKAD

VKKAKADLKAAKKYLAKISPIYSKDELKAKQAQIDDLLNKLDQPVKKKNHKKKKKKKAQ
ASSSAIMPSKYTDEIENSSYTWRVFLCLENQNIV
>ISSth1, transposase (orf1), IS3 family
MSRKSKYSVEQKLNILNEAVHSSFKRVAKKYGVNESTILTWYRLYKYQGIDGLRSIHSN
RSYSKEFKLSLVEQYQNSDDSLELFAIKHGLKSKSQLLNWIILYNESKLKVYTPRKRDSI
MPGRKTTFEERLTIIEDLIKHDVNYNWAVNKYHVSYQQVYGWYQKYRKSGNDPQSLR
DRRGKAKPKAEWTELDKLRAENRLLKAEMKRNEMEVAFAKKLIEIRNREAKKSSNTKS
>Mobile element protein
MLEAILELEEKHKWTLGYLGMTTQLAFENKLSFKVGLKRVTNCMRNHGIKANIRKKKHN
RIKRHEEYINDNLLNGQFDRQNKNEVWVTDTTEVLYGSEQVRKVRVHVVMDLYGRYIL
SYNISATETATSAIEAFKRAFKVEPGAQPMIHTDRGAAYCSRAFNDYLAEQNCIHSMSH
PGHPWENSPIERWWNDFKLVWLAKRARSKSLSELEQSVQQAIKYFNTERAYASKNGL
TAEKFRAQAA
>hypothetical protein
MPSSSVVASSSSSEEASSSSVASTDSSASTDTTTTDSTTSQTTTYQAPSNNSYSNTTT
QSYGNYGNYGNYGGGNSNNTYPSGNTNTNGSGNTGNSAGNDTTSQDQTPSADTTQ
TSDPSSIESPDGE
>Succinate-semialdehyde dehydrogenase [NAD] (EC 1.2.1.24); Succinate-semialdehyde dehydrogenase [NAD(P)+] (EC 1.2.1.16)
MMAYTVTNPYTKEVIASYEVASQAEVDQALQRAQAYYLATKGQPVEKRAQKLLKLVDL
MEERKEQLAKIASLSMGKLIGQARSEVDLCLEILRYYAARGPQMLRAKPYLYGARQEA
FLHYEASGIVLAVEPWNFPYSQVIRVLAPNYLLGNPVILKHAGQVATCARALEELAVEA
DLGTGAFKNLFLSYQQVDNLLADCRVSGLAVTGSSQVGKMLAGKAAANLVKLSLELG
GNDAFLVLADADVKKAAHDAVKARLRNSGQVCTAPKRLILAKEIATEFCDLLKKELAQL
KLGNPLDEATTLAPLASEAALTNLTKQVTAACENGAQVLVEGGPVAEMAGYFFKPVVL
RGLTPDNPMYDQEFFGPVFQIYQAQDEAEMVTIANESQYGLAGAVYSQDKVHADELA
QKLSTGQVFINQPATSHPELPFGGVKNSGYGRELSDLGLYEFANCKIIAW
>hypothetical protein
MFESLVKKLQKIYLHSSVTTVPIPDFKETAPERRRILFTGIVQGVGFRYQSKLLADRLFL
SGWVKNLPNGQVEMVVQGAPEKIDFLISQMQHTGRIKISRIQTHLLEVDWTEAGFKILS
>3'-to-5' oligoribonuclease B, Bacillus type
MPSKTRYIKLFSHNDLDGFGAPTLFEAVQPDLFKGVVFDLTTCSAGRLDARLASFFRQ
PDIGRFTDVFIMDMTPDSDYSFQQLEQHFANHWLIFDHHESEANLRAQYASNCINPTN
PAINPSATSLVWDWLQTNKNFNRLNPARVKELGLLVELIRAYDTWDWQNDDQMPSEI
KEAADNLNQLFWFYPLSKTQEFIQAVFATGWPAYRQANALLIDTLNARRTSYIKRHLKS
VATFEQDGHSFGVVYASDYKSEIAHELLQQFKVEVALVIDDRSVSLRSNGKLDVAKFAE
DYFRGGGHEDSAGGSLNFNPVLEAEKMVVAAIKEQKQVNEQMTAAEAETDSLADQLD
PETAAKLAALFGNQTD
>NADPH-dependent FMN reductase
MKVLGILGAHKEHGVTRMMLDRVLHQVQAPNTSEIIYLEDYQIKPDIKGQPNPTIDQIIAK
LEAADVWVIAAPTYWGGLSGVMKNFFDCLRQRLVRFDRKGGTHPDKFAHKHYISLTN
CYAGKWENFLSGVTDASFRTIDKVLTAGGLIKIGEAVQTATWGLTELAPAKQKELTKLG
QKINRATRKDDQTVKRYFELFFIVAVMAFITMAIQSLCLKLLNLSLGFWSNYISFVIIFFAL
LAVTLHFFALLAVTLHFFTVVKHKRK
>Hydrolase, haloacid dehalogenase-like family

MKVIKAKAVIFDMDGVIVDSEPNIMEAKQNVLHKYGVYKPLSYHFKFMGMSFRDIWAQ
TRQELDLPASTEQLMAEYFEEVAQVRAKNGLQSIAGTVEFIKALAQNDFKLALASSSPL
ADIEEVLTTFGIKELFSAVISGAELAHPKPAPDIFLLAMEKLGVKAAESLVVEDSVNGVKA
GLAAGAQLLVFNDPRYNPNRQIEGVNLVTTMTNLKVGPEKISY
>Histone acetyltransferase HPA2
MMIRQIKPSDNPVLAQIIRASLEEAKLALPGTAYFDANLDNLSEFYLAKPRRCYFVIEEA
GQVLGGAGVAEFDSANKIAELQKLYLTTEARGQGLSKLLLGKCEQFAQAAGYHELYLE
SHHNLQVALQLYERMGYRSCQGLAGSEHSLMDRFFVKELSRK
>Phosphoglycerate mutase (EC 5.4.2.1)
MSKLVFIRHGQSEWNLKNLFTGWVDVNLSEQGEQEAKEAGRKLKEAGIEFDQAYTSV
LTRAIKTLHFALEESGQLWIPEEKSWRLNERHYGALQGKNKKAAAEKYGDEQVHIWRR
SYDVLPPLLDANDEGSAANDRRYANLDPRIVPGGENLKVTLERVIPFWEDKIAPDLLAG
KNVIVAAHGNSLRALTKYIEQISDEDIMNVEMATGQPVVYDLDDKLNIVNKEKL
>Aspartate aminotransferase (EC 2.6.1.1)
MDKETFVKEYAVERQNTDAVKWDGLEAEFGASDLLPLWVADTEFKAPKAVLEAMVKR
VEHGAFGYSMTPAAYFDAYFTWQKERYGIELHKDWLRFGTGVVQSLSTVIQFLTEEND
AVMVMQPVYYPFMRVIENNHRKLVISNLKQVDGKYVMDLEDMRQKMEANQVRALVLC
SPHNPVGRVWSEAELEAVLELCRQEQVHVISDEIHHDLIMGKRPFVSALSIKDGYYRD
NLVVLDAPSKTFNLAALQNSHVIIPNPQLRDRYDEYVAKLSSPAGSLLGKVAAQAAYSE
GADWLAGFIKVVQANFDYVKTELTQAFPAIKVSDLEGTYLMWIDLSGVIAKEEVTDLVK
KQAKLAVDFGDWFGQAGLGFIRLNLGTTPANIEKAVVALIKALKEKQA
>hypothetical protein
MENLNLWQKFITNERFRRFVVLAGLIGILYLVRSLMSIILLTFIFTFLVVRAVSWVQKYLKL
PSRLIVIVIYSLLIALIYLGITIYVPKLIVQSELMVKSVLKFYQNMPADTNKVWAYISSYINSS
EIMKQVKNGAGILLKYITSIGSMGFTFLMSLLLSFFFTVEKDEMYTFSKSFLKGPNSWLF
KDIYHFAKIFVNTFGVVMEAQFMIALVNTAITLVCLSMMKMPQIISLGLMIFVLSLVPVAG
VIISAIPLSFIAYSVGGIRDVIYVIVMLLVVHALESYILNPKLMSSKTELPIFYTFVVLVASEH
LFGVWGLIVGIPIFTFALDVLGVKPVSIKHQGYQAKVKKLGRKK
>transport protein
MTKKASLGIVFATSAAILWGFTGLLSQYLFKETDVSIGWLMGVKTLISAFIVLLAAYHTEG
NKLFEVWRHPSDYLQVAAYGLVGLAGVQLTYSLTVYDSNAAIATILQTLGVVGVIIYSAL
VFHQRPRRQEYLAVIIALIGTYLLVTRGHGLNIALSPATLRDGSLLVLCGTGLAVLPVGLL
KKYSSLTILGWGLLIGGIIFEVGHPFWVGLPTFSLINTSCIILITLIGTTLSYLLFLHSMNYIN
STVAALLDTFEPLTAAVGTMLFFNASYNWAEYLGSFLILSTVFILAVGKKN
>hypothetical protein
MKHYKNRRRQFLEEWEQSDISNDYIFSKVMAQPQLCLQLIKAALPQLKIKRVYVRSQD
GLSANVDAKDIRLDISAYGEGGEMFDIEMQTIRPKYLMQRILYYHSTMITERLYPRESYG
EIPKTYVIFICLFDWYRLGNSFYEVNLVPNGVNELKFKDGFVSLILNPYGDMTGCSEEL
QDFFDLVAGRPAGRTDFVREVKYAVKLAKRRNDWRDRYMLVKDWVQDRVDYAVEQ
NTITMQEKFAQQLIKKAMSTEQDYDSLRADLIDFFDADFADKMLAKYYQQA
>Aldose 1-epimerase (EC 5.1.3.3)
MDIYRTEVEKYKTHLVTKVTLVNDNGISISCLTMGATWQEFSVPQADGSKKNLLLSFDK
PADYYANGLCTCQSIGRVAGRIKDGRFKLDGKEIQLPQNEKGNCLHGGPHGFNHTNW
KYTTSRNDSSVSVIFQQKIRQAIDGFPGDILATVIYTLDNSNRVTIAYSALGGEVDTLFNP
TCHVYFNLSNRQDLSTHSLQINSDGYLETDDELIPTGRVKDVEASPYDFRKMKNLQVAI

AENHGFDDVFVVNGPGKATMPIAVLRDDESGDQVTIDSDRNGLVMYTMHEIEPGIYFS
RDNGKEAIGGEGIALEAQTLPDAINQANFGDIVLPRQGKKTYRIQFAYDNVLPKE
>hypothetical protein
MNETIIKKLESRMTEAKAWRLENSETGHFLDVVFSLNLEDKMRNKRNFSFNRFESEQL
NELSKLVPALENDYRLELNSTNVGLGYLPVSVDSAQSLLQEV
>UDP-glucose 4-epimerase (EC 5.1.3.2)
MSILVLGGAGYIGSHTVDRLVNTTDEEVVVVDSLVTGHRAAVNPQAKFYQGDLADKDF
MRKVFTENPEIDAVIHFAAFSLVAESMTNPLKYFDNNTAGMVKLLEVMNEFKINKIVFSS
TAATYGIPEKMPIYEDDPQNPINPYGESKLMMEKIMAWADKAYGTKFVALRYFNVAGA
KPDGSIGEDHGPETHLLPIVLQVAQGKRDKLMIFGDDYNTPDGTNVRDYVHPYDLADA
HILAVNYLRAGNESNAFNLGSSTGFSNLEILKAAREVTGKEIPAEIAGRRGGDPDTLIAA
SDKARKVLGWQPEFDDIKKIIATAWKWHSTHPNGYDDRK
>Transcriptional regulator, GntR family
MAELVYQQVIKDLKKRIFANEFLDMRLPDERSLAESYQVSRSSIKRALAKMANDGIIFKK
RGSGTFINPLYLKNESVFNYEAASNLGVTDNFQMDGKKPTTKVLSFEVTLPSKELQRD
LFLSEEDFVYKIVRLRSFGSTPFMIETGYIPIKIMPQLSRKDIENSIFKYLENQYKQSVTKS
FMSVFAEPSTENDRELLGLQVNEPTGIMEGIFFLDNGTPFEFSHMRLHYKYLKFNTFVS
VH
>Xylulose-5-phosphate phosphoketolase (EC 4.1.2.9) @ Fructose-6-phosphate phosphoketolase (EC 4.1.2.22)
MTVDYSSKEYLNKLDAYWRAANYISVGQLYLKDNPLLRRPLKASDVKVKPIGHWGTIA
GQNFIYAHLNRVINKYDLNMFYVEGPGHGGQVMVSNSYLDGSYTEIYPDITEDEAGMK
KLFKRFSFPGGVASHAAPETPGSIHEGGELGYSISHATGAILDNPDLIVAAVVGDGEAE
TGPLATSWQSNKFINPINDGAVLPILDLNGFKISNPTILSRESNENLQKYFEGMGWHPIF
VEGDDPEKLHPLMAQALDQAIEEIKAIQANARQNNDASLPVWPMIIFRAPKGWTGPKT
WDNVPIEKSFRAHQIPIPVDQNDMQHADALVAWLESYRPEELFDENGTLKAEIKAIAPK
GDKRMAANPHTNPGKLIRDLVKPDFRNYAVDTTIKGQTVAQDMTVLGTYLEDVVKENA
ANKNFRIFGPDETMSNRLAPLFKETNRQWLSPIKEPNDEFLAPAGRIIDSQLSEHQDEG
FLEGYVLTGRHGFFASYEAFLRVVDSMLTQHFKWLRKAKEQPWRTSVPSLNVVATST
VFQQDHNGYTHQDPGILGHLADKKPEFIREYLPADANSLLAVFDKVLNDRDKINLIVSS
KHPRQQFYSAAEAKELVDKGLKIIDWASTDKNAKPDVVIAAAGTEPNLEALAAISILHEK
LPDLKIRFINVVDLLKLRSPKVDPRGLSDAEFDAYFTKDVPVIFAFHGYEGLIRDIFFDRH
NHNLSIHGYRENGDITTPFDMRVLNQMDRFDLVKAAVLSLPDADKYGQIVAEMDAKVA
KHHQYIRDEGTDLPEVEAWEWKSLE
>transposase
MNNSIKTILGIKDPYLKLDEKNFDNPIEDQPNQIIVHLIQTYPMHCPKCGHLMCKNGYKT
VNCLGPELHFKPTIWSIKKQKYICKASSSCPEVVTKLAAVEDIHYRNHISLAIKQRAMML
LTKNESQSDLAKELNVSDWTIRRVITNLDQFFKPNYHWLPRHIAFDDFKSGRFAPSGM
SMILMNIENKRTLDIILSRKNSYLRKYFLRYDRSARLAVQTVTVDLYTPYRHLIHELFPHA
LIIADHFHIVAQAYRAFNKIRIQVMNRAGAGTHKWRALKHFWKLLLTPANELKYDNYWS
RRNFSYAQLTDVEVIHRLLSFDNELKRAYEYYQNLILVIAHRSKKELKNLLAIKWTQLPQ
ALQKVQRTLRRHKQEIYNSFKYDTYTNGPVEGTNNKIKVIKRTAYGFRNFFNFRIRILLA
LPNTYIAINWRNKQTAHAKVQAQAA
>hypothetical protein
MTIQNLTKSVAIYRLAYIVQLSWLLISVPFFVILLYRIFTHEYFKTLNTFSKVFLACALLIISG
VSVYLGALASSMLYFSI

>Phosphoglycerate mutase family
MTKLYIVRHGQSQANANGILQGSKVDTPLTQTGIKQAFATKEALSDVSFSKVVTSPLLR
AAQTATIIAGADRTITFDPRLVEYDYGNWDGMLEADVWAKFPQYFDAKHNLLPNSWET
SGGDTYDEVKAKIAALFGDLVAKYPDEAVLVVSHGFTIKLMLDYVLGINNLVNMNEPLN
AAVSLVELTDSSQTLVYFNR
>Permeases of the major facilitator superfamily
MAVLLLLTIYIGFIGLGIPDSLFGPAWPVIYKEFGVSLASANYVSFVISGLTVVSSLVAPSV
IARLKTAKVAALSTALTTIGLLGFAYAPNLWWLCLWGIPLGLGAGAIDTALNNYVALHYH
PSQMNFLHCFYGIGVSLSPAIMSLALQDHNNWRAGYRWAFYIQAAITLIIVLSFPLWQK
VHPEVKEEKATVHNHPGLMELLKQPLVLTTCLVFMGSVAIEITTGSWGSTYLVNAKNLS
PELAAGLITLYYVGMASGRFFGGIVANYWRCWTLIKLSVSLTIVAIGILFLPLPELVAGFA
LFLLGFGNGQLYPNLLALTPVHFGKEASQTVMGLQLAFSYGGGLLLLPILFGYIAQAFG
SGLFAAYLLVLALVTLVSMLMLLYLLKQAGIKQQ
>D-alanine aminotransferase (EC 2.6.1.21)
MQTVGYYNGTLAPLDELMIPATDRALYFGDGCYDATTFKNNVAFALEDHLDRFYNSCR
LLEIEFPLTRSELKDKLYQVIDANESDTGLLYWQTSRGSGLRNHLFPEDSQPNLLIFTAP
FGLIPFDAEFKLISREDTRFLHCNIKTLNLLPNVIASQRAKEAGCQEVVFHRGERVTECG
HSNILILKDGVLCTPPRDNLILPGITLKHLLELAKENNIPTSEAPFTMDDLRAADEVIVSSS
ACLGIRAVELDGQPVGGKDLATLQVLQQAYAKKYEADVQSR
>FIG003003: hypothetical protein
MIELGLTTWSEHQTFINENRPVTLEEYAAHLPLVEVDTFFYGLPRASTVANWLSQVPPS
FSFIPKVPKALTRHPGEQLAGLELKQAFSAFRQSVDSLVATNHLKTVLVQFPPTFTATK
ENVAYLRYFRSLLSDLPLALELRHQSWYAPGVKDSLLRFCQQYKYSLVAADEPTNTPA
SVPFVLAATNPKLAVVRLHGRNARGWANQGADWRKKRTLYSYSQQELAWLKEQVLRI
EAQVEEVVVIFNNNSAGDAAPNALAFKEMLGLEFSGLGKMPPEQLDLL
>tRNA-Asn-GTT

>hypothetical protein
MLTQPELGIRIALTTDYLSEFLPVLYKTLLILAFPPTHIDTYLDQLVSQSRTEATAKSLTAK
LDLTAATNQPEFKADLETLLAEGPEYTSDYSLSMLSSTLTISNCGRESLAQTRVLVSKTL
AKKPEKPTRAFKYITNKPNFDDYYQWEDLSFKALPSKKILNKIKRLTFN
>hypothetical protein
MNEYLIKNYLRTVISDLASPATFLYLQWNIVCDEIPDLDYALDLIHSLEQALVLFYQFLAS
CGVITQKDSNLAKFELEQALELIHSECLEEVAADQIMQAFPMDNGDDLLDLTDPDKAAQ
VLAKLYDLKDSFEQPTPPAHEKTYMKLHKPK
>GCN5-related N-acetyltransferase
MTEVKIRQIGLADVEELQAISCETYEATFGADNPRELMDEYLAQAYDLNKLKRELGTEG
TYFYFLEVDGKTAGYLKLNVGQAQSEPMGPDNLEIERIYLRTAYQGKGLAGKLLTKAEE
MAQALGKNRIWLGVWERNYRAQKFYAKYGFSRFSEHFFMMGDDAQTDYLLAKQL
>hypothetical protein
MPMHKAWAGETVSSSVTITAVVPEKYMIALDIEGQGAVEVNGVSYDKSTKIQGRGAEN
YKVTAAKGYKIASVMYGKGDKFRPVALNEKQEYKAPAVSENGNKLKVVFEKPKMIEAV
KAEVVKPHYLIMSGVFAALVGLIILL
>hypothetical protein

MKNKVITGLLGFLVVLGMATPALADTAVNTKTINEGQRNDQINSTIVSTDVNSSFTMSIP
ATATLTYGVEDSNIGNFSVYGNLKTNQSVSVTINKGDFKNVNDNSVLPYKVTDEKNKD
LTNLNYTADDIASKTVTPINVHIDRTAWQLAHAGKYQGSIQFDATLNN
>hypothetical protein
MEIKEPQKEPEESKKDDKKKLLLLLLLLLLLFGGIYIGLNWNNWFGDRPTVNIDANAQN
QKGKKQKKAGSIQIPGYASINMKANVKEQKVNLYNPKENKCYFRLSIYLPDGTKLWQS
QMIEPGKAIYNLSLNQTLKAGTYKNCTLRYECFALSKKLVPLNGSQINFTLNVVE
>Putative hemagglutinin/hemolysin-related protein
MLLVTKDVLQIKGVAERQMKKNKTITALCVSLLALSSSNSLVYAATTSASSASSQTKVA
TSSQATSSSSTSATSSQSGQAQSSQNASTSSQTVSVSSLSGNIALLNIAESKVSDNNY
KIANANDFLSWLAGINSGKLDASHNVYLSNDIDMAGKSFTGIREFKGHFYGGNHKISNL
NCSNSSSSSLIDSNSGTFENITFDGLTIKSNDHYLIRSNSGSLTNVSFKNLSLSGTRAYLI
DNNSGNINNLNLANVTLNGEAESIIHTNSANLSRVTVDTLKATGYMVDFIETNSGSITDS
QFTNMTFTGRQTTFIKVNSSNLQRVGFTGLKLDGTNSEASFIQDNSGALDTVSFDIEIK
GRQPRFIGNNNSSLTNIKFNQIDIEGYQASFIQNNGGMVSGSFVGKVEGEDACFINLNN
GRYDFDWTDKFIVKDHITNIWELNNFSTYVNNGQVKIDAVLDNNIDAVYENFTPIAYNN
NNPYQGTFDGNGKAIYNFYSNRYGASFIFNNAGTVKHLEISGQVFGEASDFVSHNGKT
GLIQDCALILDITGETYDDLFGVTHYGRTANSFENNEGTIKNIFFAGRLTNIRSENDPRN
REGTVDHFYYLDQFGVIGGKAIRVSRWRAESGEIATKLASNSTIGWMQVREYWHYDIL
SSRPGFSGDKGDQTKLMAEKPSLSINDIDRTEKSLRVTKKFDSKILDEYGGVRYELLDK
DGHVLGNGWQESPEFTNLAPETEYQVAVRFNGNGRYLPSDATIVKVKTKGIAPAVELT
KDDIKRTNTSLEVTKKFGSEYGGVHYQLKDTQGNVVKDWQKDNKFTGLTPGTEYQLV
AQYVGNDDYVASEEVIVNVKTTKEAETPELSSTDVKRGDTSLKVTKQFADTDKYGQVE
YELTDSQGNVIKEWQANSEFTGLTPDTEYTISVRYKGNDEYLESDKASVTVKTKKTAE
PQVSVDNIYGSSDSFTIKDNPDTRYGNVSYEITDNQGKPIKNGQKQVNGDFSGLDLED
GVIYQVHVGYDGNDDYASTEKVIKIMKAPTVTIASENANSLKVADLANTDKYGQAEYAR
SPSGLTANATSSTFNNLEAETTYVVYARYKGNGYYPPSAIGQTKAKTSEASYQLRIPPV
LRAGSDVGEVGVNPSTFNIGTSDSVRISIVGGVSHGKVTLTRQNDPTYQKLKTRVLLNN
VEVDNGDSDEINVIDYDKNNYQQQTSLKFTDPRDNKEDNPKTPAGFYTGAIDFKITYNK
EGNK
>Ribulosamine/erythrulosamine 3-kinase potentially involved in protein deglycation
MFSVGLLWGDYLNKSLLASLPLDGIEECLPVSGGDVNRAFKLVTKNQNYFLLVQPVCV
REFYAAEIAGLNDLKRAKIITPEVISTGKLGRGAYLLLSYLEKGEGSQADLGRTVAKMHR
TYNPSGQFGYKYDYVGTSISFSNDWRASWAELFVNNRLDILAQKLVEDTYWKIEDFST
YMQVRQLILTALSQHDSKPSLLHGDLWAGNFLFTEAGQPALIDPAPLYGDREFDLGITT
VFGGFTEEFYQAYQESYPLKAGYSKRMHFYRLYYLMVHLDKFGRSYKQAVADEMQAI
LQAG
>Fumarate reductase flavoprotein subunit (EC 1.3.99.1)
MKIIGISGTNSSDSPTEKLVNFMAQHFASQVEFEVIELKGLPMFNESNDLSNQEPLKSL
VAKIEAADGVVIATSEHNRSIPSALNSFLEWMSFTVHPFDEKPVMVVGTSVTRQGSAS
AQLHLRQVLDAPGVNALVLPGNEFLLGQAAEAFDENGAIKEANTIDFLESCFANFLRFIK
AADSLQIPDEVRFEPGDYQVKTKGHNGDLPMVVSFSENRIEDIKVDTSGETEGIADTVF
ERLPQEIIAGQTLNVDAVSGASVTSYGLIDGVAQAVKLAGVDPNILKKRPKPSKSQDLS
PLEYKTDVVVVGGGGAGLAAASRVLQAGRSTIVLEKFPALGGNTVRAGGPMNAADPD
WQKQFAALPGEASVLKEMLDYDLAKIDPEYQADFKALQGQIKDYLAGKADYLFDSILW
HRIQTYLGGKRVDLNGNEIHGDYDLVKVLTDHALESVKWLADLGVEFDESQVTMPVGA

KWRRGHKPMESQGFAFIKTLKKFIEEHEAGQILTETPVKRFLLDEQGQICGVVALNAAN
RQVIVKAKAVILASGGFGANTKMVQKYNTYWSQVDDDIATSNSPAITGDGIKLGQSVG
AALVGMGFTQMMPVSDPKTGELFSGLQVPPANYVMVNQQGKRFVDEYEGRDVLTKA
AFDNGGLFYLIADDEIKKTAYNTTQASLDAQVEAGTLFRADTLADLAKQIKVDPQTFEET
IAKYNSYVDAGVDPEFGKEVFDLKVVKPPFYATPRKPAIHHTMGGLKINTKAQVINEAG
QVIPNLYAAGEVAGGIHAGNRLGGNSLTDIFTYGRIAAKTALEKM
>Fumarate reductase, flavoprotein subunit precursor (EC 1.3.99.1)
MEKLNLVGIVGTNAEKSTNRTLLKFMQKHFATQAELTLIEVKDLPFFDKPASHEAPAAV
AAINQAVAQADGVIIATPEYDHTIPAPLTNLLEWFAYTAKPLQDKPVMIVGASYGVLGTS
RAQNHLRQILNAPELRALVLPGREFLLGHSLQAFDEEANLLVDEKVAQLESYFADFAAF
VTRVKA
>Membrane protein involved in the export of O-antigen, teichoic acid lipoteichoic acids
MKNKLLSGAFWMSFGSIFSRVLGILYIIAWLAMFHSPQQQYTAQALYNAAYTPYGLFLA
LGTSGFPTAISRKIAYYNSRKRYLDSKRLFKAGLLFMALSGLICGVVFFIFAPTIAKSSPI
MDQAAATFVLRSLVPALMILPTMSIIRGWFQGHQDMKPFGVSQLIEQFIRVVAILVFTFV
SFNVLKQGLTWAVALSTFAAFIGALGSLAYLGKYYRAQLPTYRANEAESLPASQLDIKG
LFLEIVKEAVPFVFVGAGVTINQIIDQLTFKHLMLWASNYDMNYIQNMFTFFSANPTKITT
IIISLATALSDTSLPILAGLVGNRQKIAQTLSDNFKLMITLLIPFVAILTVLAQEVNTIFFGYN
WAAGQLMAASIMLSLILSFFLDVSTLIQALGKHRLAVILLSFSLLLKLALQAPLIWLFHEY
GALLATTIAFASSLALGLWYLNRTYLTDRRANLSWLKPSLGIGIYVLGLGLLAKPLSLLSL
GHGKVAALITCVIFGLVAGSYYLVIAYLFRFKDLLKSMNR
>hypothetical protein
MSTSKFCPNCGQAVKQNDKFCQHCGAKLEQAQQGYSRTASKKTLDKKWLILGAVVLV
LLGAYLGGRSYYQPNKQLERFTTALVGKGDVKSYLVSSDQNLKITAKDAAAFKKSFAN
SQDQVATMKQHFLMGSSYAGYSWVKSGSHWLIFPAYKLQVSPAYVTLTTNKRGVKFY
QDGKVITKKQVTMANGYGVKVGPLFPGQYTFTAKGKASGKPIVVKQTKVLAAGNNDLA
LELATVNFKVEGQAGQAVYINNEKVGTLDKEGEKVIKNYPASAKMKLYLTTKAGNKEIK
SEVVDVSGYTYNYQDVNEDDLPEVAPKFKGLISKDEAEDLLESAFNEGEDGSSAASFV
GEENNDSYNELVAFFKSLHNLSSYEVEDIKSISPESTTKVKVSYAMKYTFINYDSDSER
KKIQQFTYHDAEIVYDQKDNTYKLQSVGKTGTKADWERTYSADEDD
>hypothetical protein
MKKCPNCGSEVAEEAKFCTNCGNELATGDEVKPKAVASEVQSENASEQVKESAAKQ
EAGVSLDKEKLKASSLSYWSWLVTAWKHPFKVKNETSKWSGLITLGLEGLIFGLSIMIIF
RGYYNSANSYYSQAFGGSFGDNPVGLGTLITFFLATFLSALATTGIGYGARELFTPNRT
DFLDFVNGIAHRTSPILLLNLIALLFSFTVGFTTSGVLNTMLVLCLFSIILVIWQFGILAPIYE
LENPKLDRLYILIIVTLLNGFVYGLAYMAVAAHYVSYLQQLFND
>Beta-hexosaminidase (EC 3.2.1.52)
MKKLLLTLVTLCLAWIIVPGSNDVQAKKISSKRLNQLVAKMSLEEKIAQMYLVPSSGNPA
GMAAEIKHYQLGGVVLFGADFQNQTPKQLKAKLKTWQKVSQTGLFIATDQEGGTVSR
LSANPKLTGKKYPAPQELYAKGGLAAVSKNYRQTAKTLRRLGINWNFAPVADVSNDPN
SFIYPRTLGQDYQTTATYIKKVVPAIQRQKVIASLKHFPGYGSAADTHTGFASDKRPLS
DFQQNDFLPFEAGIKAKVDTVLVAHVVITAVDDHNPASLSPAVHQLLRQDLHFKGLIVT
DDLGMGAIKDFAKEQAVNPDLLAVKAGNDILLSTGYQKGIPAIAAAVRAGEISEGQIDQT
VKRIIKLKAKLGILK
>5S RNA

>Large Subunit Ribosomal RNA; lsuRNA; LSU rRNA

>Small Subunit Ribosomal RNA; ssuRNA; SSU rRNA

>Mobile element protein
MDQIRADQAELPKAQRYKIGDLLSRINLAKATYHDERKRIANPDDKYCEVKKVIRRICEQ
GKVRGRLTYGYRRVQARLDQLGIHLAGATIRHLMSELGVQVEIYNHHRNGKYSSYHG
RVGSVADNVLQQQFNAKDPYQVIHTDVTQVRLANQQWAYISAMTDEASKEVLAFQMS
SHPNQELIMNTLNELLDHLPDDAQPIIHSDQGWHYQLKYYTKKLKDHHFIQSMSRKGN
CHDNAPIESFFHLYKTELLQGLPPCNNITELRKLSLTYIDQFNNERISLKTKGMTPVGYR
NHALAV
>hypothetical protein
MTKYSPEFKAQVVVEYMTDGLSSYTLEKKYKVSARQIRKWIQQYQLQGIKSFKQHHRK
RIFSAKFKWHVINYYQTHAESFAQVAAKFDLLPSQISIWRQAFIKYGYDALKPHRKGRS
QKVKKQTKKQKQKLVEKKEIDQLREELARTKQELYNTKMDRDILKKSLALFGPSKPGK
KRK
>hypothetical protein
MSAKKREVLSKLNETVNLDLLKKALLHESLYAGLKEHRHIAIRILSQVKSGLANDVLHLA
EVLDLVPSEVKSAVDKLVASGLLSLKEGRLELTAKGQEVFTKMAAIHDEISEQIFKDFDS
EELDQFLTYLAKIKANLAGISESELAQLKEKMRELPTHPSQFFKK
>Amino acid transporter
MQNKSTKIGLFSAVMLALGSLIGSGWLFGAGEAARVAGPAAIISWILGAVIIMVIAFNYIE
LGAMFPESGGMSRYAQYSHGPLLGFIAAWANWISLITLIPIEAVASVQYMSTWPWKWA
NWTNQFMDHGQVTGLGIVVVFFFMLFFTMLNFWSVKVLTHFTSLISIFKIGMPTLTIIMLII
SGFHTGNFGHSWQSFAPYGSRAIFEATTVSGIIFSYDAFQTVINMGGELERPQKNILRG
IVISLSITAVIYILLQVAFIGAVEPGSLAKGWAGLNFASPFADLAILLGLHWLSVLLYIDAFV
SPFGTGVAFVATAARTLASMTKSGHIPAWLGRLERRYMVPRFALLADLGLAVILVCLFR
DWSTLASVISASTLIAYLTGPVTAVSLRKLRPDFKRPVKSNWLKVLAPLSFVLTSLAIYW
AMWPTTVEVILVIALGLPIYCYYDWRYGQLSLKGQLKGAGWLLVYLVFISVMSYLGSAG
FGGQDYFKYPFDFLIIVIVSLAFYYWALASRIDGPDLVAAAKVNHQVKLED
>Transcriptional regulator, GntR family
MSGDNKKKITFPRYQQIAVALAERIVAGKYPVGSKLHARSTLASNFKVSPETARKAINM
LADLDIMEVKHGSGAYISSKEKAQEFLANYQETTSLAKMEAGISQALSKQKADLEELAQ
MVTNLVERTQLVHKKFPFSPFELVLPTTATNLERSLNELNLWHSTGATLIAIRHADELVV
SPGPYAKLAAGDTIYFVGSELSYERVCNLFNIQLNN
>Glycine betaine ABC transport system, ATP-binding protein OpuAA (EC 3.6.3.32)
MEQLEVHNLTKIFGRKRKAALNLVKQQVSKTEILKKTGATVGVYDANFKVKAGEIFVIM
GLSGSGKSTLIRMLNRLINPTSGQIYIDGKDVAKMNEEELRQVRREKINMVFQNFGLFP
HKTILENTEYGLEVRGVAKEKRRQLAEAALDSANLLAFKDQYPNQLSGGMQQRVGLA
RALANDPDILLMDEAFSALDPLIRRDMQDELLNLQEEKQRTIIFITHDLNEALRIGDRIAIM
NDGQILQIGTGEEILTNPQNDFIKEFVEDVDRSKVLTAQNIMTAPFTTNIENDGPNVALS
RMRVEEVSMLLAVDKRRQLKGSLTAEAALKARKEKLPLTEVIDKNVKTVSPDTLITDIFD
LIYDSPTPLAVVNENGRLLGVIIKGRVLEALAVNSSED
>Glycine betaine ABC transport system, permease protein OpuAB / Glycine betaine ABC transport system, glycine betaine-binding protein OpuAC

MDIFSYKLPIAHLIDQLTNWLTTNCAGFFNLITKGGSALMDFITNGLLFLPPLLLIVLLTIAA
YFGSNRHFKLPALTLLGLLFVYNQGLWTGLMNTFTLVLIASLVSVIIGIPLGIWMAKNDK
VQAVVKPILDFMQTMPAFVYLIPAVAFFGIGMVPGVFASVIFALPPTVRFTNLAIRQIPKE
LVEATDAFGSTPKQKLFSLELPMAKPTIWAGINQTIMLALSMVVTASMIGAPGLGKDVL
SALQHANIGTGFVAGLALVILAIIIDRFTQKFNQPIQAKPKLTPKQKRNHRLIVGGILAAIIL
VLGAGNFKSLTQKAKPTVNLGYVDWDSEVASTNVIAQVLRDQGFKVELTSLDNAVLW
KSVANGSVDAEVGAWLPTTHGNLYKQYHNNLVDLGPNLKGVKLGLAVPSYMNVNSIS
DLSNQANKTITGIEPGAGVTAAAKKTLKAYSNLNDWQLSTSSSGAMTVALDKAIKQHQ
EIVITGWSPHWMFAKYKLKYLADSKHTMGRSENIHTIVKKNLKTDNPKAYRILNKFTWT
KEDMESVMLDINNGMSPQKAAQKWIKAHPNQVKAWVD
>conserved hypothetical protein
MYLTVTTEAKEKLAPFADFDLLLDYDDGVGPLSKAGSCTLDGGFRIIAVKPGTYAKDYS
ETLDSDLGPIRYKDYTDMFMDNNMKLQINPVNKLLKLSGNSSGELTASVAIVDATNN
>hypothetical protein
MSKKILPLFGGIIVLGLIAGLLLNFCPGLEAWLHQITGWH
>Mobile element protein
MNHFKGKQFQQDVIIVAVGYYLRYNLSYREVQEILYDRGINVSHTTIYRWVQEYGKLLY
QIWKKKNKKSFYSWKMDETYIKIKGKWHYLYRAIDADGLTLDIWLRKKRDTQAAYAFLK
RLVKQFDEPKVVVTDKAPSITSAFKRLKEYGFYQGTEHRTIKYLNNLIEQDHRPVKRRN
KFYRSLRTASTTIKGMEAIRGLYKKTRKEGTLFGFSVCTEIKVLLGIPA
>Mobile element protein
MRFDKTQVQNYDGLCGHLAIVMMTYDLLAWQERQEKDERTIGDLFYIMGEAMPDLSL
TNVLVWFIKLLNRLVESEPVAPIKQLNETVDKFISSLPQSIASQLQKS
>Glycerol uptake facilitator protein
MNGFLGEFFGTMILIVLGVGTGAGVNLKTSYARNSSWLFITIAWGMAVTFGVYVAGSF
GSQGHLNPAVTIPFAIFGFFPWDQVLPYLLGQFLGAFVGAVLVIIQYGAQFKNAKDVSD
GNVVGIFATGPAINNSVYNFLSETIATFVFILTLLNLGDFTKGLKPLIVGLLIMVVGQALGG
TTGFALNPARDWAPRFAYSVLPVPNKGDSNWSYAWVPMFGPLTGGLIATCVQYFLM
>Aerobic glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)
MTFSIKTRKNTVDRLKNEELDLLVIGGGITGAGVALQATASGLKTGLVEMQDFAEGTSS
RSTKLVHGGIRYLKTFDVGVVADTVQERAVVQGIAPHIPRPFPMLLPIYDDPSSTFDMF
SVKIAMDLYDRLAGVTGTQYANYTISREEVLQREPNLKSDKLLGAGVYLDYVNNDARL
VIENIKEADELGGLMASHVKVVGVLHDDNGRINGVKAKDLLTDEEFEIKAKMVINTTGP
WVTKTLGLDDKNGEGEVIRPTKGVHLVVDSSRLRVPQPTYFDSGIGDGRMIFVVPREN
KTYFGTTDTDYTGDYKHPRVEQADVDYLLKAINNRYPEAEITINDIEASWAGLRPLLSG
NGSSDYNGGGSNTGKVSDESFENLIKTAVAYSKDEASRGEVEKSISSLKTASAEKTLS
PSQVSRGSSLKVGDDGLITLSGGKITDYRKMAAGAMELIRKLFKELFNKDVREVDSKKL
QVSGGHFDPTNVDETMKFYQKIAMDKGIPAADAADLANRFGSNTSRVISYADDVEAAP
GLTLADTLSLHYSVNEEMTLNLVDFLLRRTNYVLFHKDELMMKKDAFVDEMARIMEWD
EKEKVAAAKQLDETIAESSLAYLK
>Glycerol kinase (EC 2.7.1.30)
MMSDKYIMAIDEGTTSTRAIIFDKKGNKIAEGQKEFRQYFPQPGWVEHDANEIWNSVLS
TLANSFISSGIQPKQIAGIGITNQRETTIIWDKETGLPIYNAIVWQSRQTSDIAQKLIDEGY
SDTIHEKTGLIIDAYFSATKIRWILDHVEGAQERAEKGELLFGTIDTWLTWKLTGGKVHV
TDYSNASRTMLFNIHDLKWDKEILELLNIPETLLPEVKGNSEIYGQTAAYHFFGSQVPIS
GMAGDQQSALFGQLALEDGTVKNTYGTGSFIVMNTGNKPVLSKNNLLTTIGYSINGKIN

FALEGSVFVSGSALQWLRDSMEIIESAPESERAANESTNHNEVYVVPAFTGLGAPYWD
SDARGAVFGLTRGTTKNDFIKATLQSLAYQTRDVLDTMEKDTGIKIPVLRVDGGAAMN
NYLLQFQSDLLNKSVERAGNLETTALGVAFLAGLAVGFWKDTDELKELFTVGRRFEPE
MSDDAREYLYAGWKNAVEATRIFKHKINENL
>arylsulfate sulfotransferase
MYLQLKRKLFLLALGLVVVTLTASLIFVKQKASSNATSSETILPVANGQTVTTLDYSDLNI
KSVSDIYQTATQTKIAQKIKQLEASQTYSFDKMLVLANPYLTNTTGLYLHFSTTKATKISY
TVKSKGTSTFSKTLYNPNGTYAKDHTYQLIGLIAGQENTITITATDQNGQKETKNFTYTP
NKLRGSKQNQLKVTKGTSKTKLSTGLYTVIGDKSLKTRNTYLVDNDGYIRAEIPTINYNS
LCLIQTNNKLYLAVDDDQIVTLDRLGQVVQSYSLKDTNFKLHHDFAVDSQGNIIALATDT
KLKASEKRVEDQIIKIDGTSGQVSRLLDFKDLLGDLYKTATGIETLTNNKGYRDVIHANTI
QLTKDDQVIISSRETSTIMKISNLTSQPKLDYFISDSSVWQGVGTYSNLLLNKVGNFTSQ
AGQHSVTYIPTETAGVYYLEMFNNNSAIMNSRTNFSWANYPNSGGATTSNDYQSSYY
KYLVNENTGSYQLVKKISLPYSPFISSVQTKDNNVVTDSGMTATFAEYDADGKLIQSFE
TTDSTKFIYRVYKYDFNNFYFAN
>Putative stomatin/prohibitin-family membrane protease subunit YbbK
MLGIKIIRQNCQGLVETLGKYSHSVEAGLHFYIPFIQRIESVELAMHPLKLQKYSVITQDN
ADIEASVTLNYHVTDAKKYTYENTDSVESMAQLVRGHLRDIIGRMDLNAALGSTSKINA
ELAAAIGDLTNIYGINVDRVNIDELTPSVEIQKAMDKQLTADRERVAVIAKAEGEARNIKL
TTDAKNQALVETAQAEAEATRMRADAESYRIEKIRAALESVSDNYFRDQSIAAFTKLAQ
SDTNLVVMSKNEITDLGNIPVIDKMLDMKK
>hypothetical protein
MARDEEMTSFLFRLPKNLKRRLESKAQSENKSVNATLQEIVSLTLKDPPQQVEAGSFE
KRNFLGQKIAGQKIDQANGLVEVKGIYYRYLIEGNRPVDKQADYIVIEAVGNIITLRPLVG
E
>NADH dehydrogenase (EC 1.6.99.3)
MAVVNNDFRDVMLNRHSVRKFDEQVKIERSELTEMIAEAITAPSACNLQAWRFVVVDS
PAGKDKLRSFFMRFNLPQLETASAVVMIFGDTLAFESYRDLWNKAYENGQITAAKRDE
VLKTFLPLYESAPKSMLINDSMVDCSLAAMQLMLAARAHGYESNPIAGYDAKKAAPAF
DLDPERYVPVMAIAIGKPVDEGVDKVTSSRYPVKDVMDFA
>PlcB, ORFX, ORFP, ORFB, ORFA, ldh gene
MLVTTTEQIPGKNYEVLGEVFGVTTQSRNVVSDFGAGIKNIVGGEIKGYTKMLQTSREK
AVSRMIAEAKDKGADAVVMMRFDSGSIGGDMQSVVAYGTAVRFVD
>Translation elongation factor P
MVEAINLRSGMVFEKDGKLIKVLEANHHKPGKGNTVMQMKLADVRAGSIVQTTMRPS
EKVELAMVDKKQAQYLYRQDDLLCFMDLETYEQYELAASLLENELPYLIENLQVQLDF
CNGELIGIQLPGTVELTVTETEPTIKGATAASGGKPATMETGLVVTVPDFIQTGEKLVIN
TANGTYKQRA
>Exodeoxyribonuclease III (EC 3.1.11.2)
MKLISWNIDSLNAALTGQSERAEQSRAVLDTIQAQAPDIVAIQETKLPASGPNKKHVEAL
AKWFADYDFVWRSSVEPARKGYAGTLFLYRKNLTPKVTYPEIGAPSPMDQEGRIITLE
FPKFYVSQVYTPNAGDGLKRLAERQVWDQKYAAYLAELDKTKPVLASGDYNVAHEEI
DLAHPKTNHQSAGFTDEERAGFTNLLKAGFTDSFRYVHGQVEGVYSWWAQRSKTSKI
NNSGWRIDYWLTSNRIADKIVKAEMLDSGDRRDHTPIMLEIDL
>Fructokinase (EC 2.7.1.4)

MTKLYGAIEAGGTKFVCAVSDENLNILKRVKIPTTVPAETMAEVFKFFDENPVVAMGIG
SFGPIDVNRSSDTYGYITTTPKKGWANYNFLGDMKAHYDIPYVWTTDVNVAAYGELKK
GSAQGKNSCVYLTVGTGVGGGGVVNGQLLEGFGHPEMGHIIMRRHPEDTYQGNCPY
HHDCLEGLCAGPAVEGRAGGIKAFDIPKDAKDWEIEAYYLAQACVDYTVILSPEKIIFGG
GVSKQDQLFPMIRKSFAEQMNGYLKTPDLDDYIVHVALGDDAGITGCLLLAQEEAAKN
>biphenyl-2,3-diol 1,2-dioxygenase III-related protein
MKIKRIDHIVLTVTNLASASRFYHEVFDMPVLAETAESHTLRCGHQLIVLQKVDAKADLK
ASHPTSGSADLCIVAGDKMDDIVNHLKSYFVNIVAGPSQTSGSEGKMTSVFVRDYDNN
LIEVATYKG
>Aconitate hydratase (EC 4.2.1.3)
MSGKYLRRFNSNKAEYSYYDIGAYLKEQGQTVANLPYSLRVLLELTLRQSARKPELAP
FLKHFCDWDQKHDQDIAYKPARVILQDLTGVPALVDLAAIRAEVKRQKGNPSSINPDVP
VHLVVDHSVQVDMAGNDQALAYNIKKEFERNQERYHFLKWAQKLFENLTIIPPDTGIIH
QVNIEYLADVIRTDTADLAFPDTLVGTDLHTTMVNALGVLGWGVGGIEAEASMLGQES
YFPMPKVVGVKLTGKLAEFATATDLALTITKLLRTHNVVGKFVEFFGPGLASLPLANRA
TIANMAPEYGATCGFFPIDDQTLAYLKLTDRTATQVKLVEDYAKANHLFYQADVTPTYS
ELVELDLATVKPSLAGPKRPQDLVDLDQVSTNFCQYDDLADYQLKIEKNPVKLRPGAL
ALAAITSCTNTSNPEVLVAAGLLAKRACELGLKVPAYVKTSLAPGSRVVTAYLKTTGLQ
TYLDQLGFNVVGYGCTTCIGNSGKLQADLQAALEARPYPVSAVESGNRNFEGRVNPL
VKDVYLASPPLVVAYALAGTMDIDLTKEPVGTTPNGQAVYLQDLWPAASEITALIQAYV
KPALFSANYRDIFKQNEAWNELDAQASETYDFDATSTYIAQPPFVVAGTGTKELTGLR
VLAKLGDTVTTDHISPAGFIGQQTPAGQYLAELILQKNS
>Oleate hydratase (EC 4.2.1.53)
MYYSNGNYEGHARPKKPAGVEKKRAYIVGGGLAGLSAAVFLVRDAQMPGENIHILEEL
PVAGGSLDGALKPGLGYVVRGGREMEDHFECMWDMYSSIPSLEIPGASYLDEYYWLD
KDDPNSSNCRLIYNRGNQVPSDGQYQLGKLSNELVKLVLTPEDDLEGKTIEDWFSADF
FETNFWTYWATMFAFEKWHSLIEMRRYMMRFIHHIDGLPDFTALKFNKYNQYESMTK
PLLAYLKDHKVQFDYDCQVDNIVVDTAEGQKHAKELQLTKAGKSEAISLTDDDIVFVTN
GSITESSTYGSHETPAPPTHELGGSWRLWQNLAKQSPDFGNPDVFCQNLPERSWFV
SATATIRNPQIEPYIERLTKRDLHAHKVNTGGIITVTDSNWMLSWTIHRQPHFKGQKEN
EIVVWIYGLYSDTKGNYINKKITECSGKEITEEWLYHLGVPESLIASYAKEESIQTVPVYM
PFITSYFMLRKKGDRPAVVPKGSANLAFIGNFAESPSRDTVFTTEYSVRTAMEAVYSLL
NVDRGVPEVFNSIYDIRELMRAMYYMNDKQGLAEMDLALPVPKPVKKALLKKLEKTWL
GELMRQEHLL
>Transmembrane component MtsC of energizing module of methionine-regulated ECF transporter
MNEHQILGYIPKQTVIHSFNATAKLFFVLAVSFACMTTYDTRFLLVVGICSTFIFKLADLK
WNQISFVIKFVCFFSALNILAVFVFQPTYGQELYGAKQVLISAGYFTLTSQELFYLANVA
LKYFCTIPLTLGFILTTNPSAFAASLSKLKVSYKVSYAVSLALRYIPDIQADYFAISQAQQA
RGFEMSKKGKFFNRVKGALQIVLPLVFSSLERIETISTAMELRRFGAKKKRTWYVAEPF
RTRDYLLVVSGLVLFIISLILFKINGGRFYNPFK
>Duplicated ATPase component MtsB of energizing module of methionine-regulated ECF transporter
MTDAIIDFKDFSFRYESQAHPTLKKINLQIYPGEKICICGPSGSGKSTLLQCLNGLIPNQY
EGEIKGQALVAGKDITKTTMFDLSFETGTVLQDTDSQFIGLTVAEDIAFALENDGVAKSE
MQKRVAKWAKRVNLASHLNYRPQELSGGQKQRVSLAGVLIDESPLLLFDEPLANLDPA

SGQATIDLIDDLHRQEDVTVVIVEHRLEEVLTKPVDRVIVINDGQIVADLPPAELLRSDLL
MNSGIREPLYVTALKYAGVDLNEVKGIESMASLCLTDEQAQQLADWQAQQPAETTQK
QAQPLLEVADLSYQYLPQNPPALKEISAQIKRGEMLAVVGRNGAGKSTLFKALCGFITP
SGRISWRGQDLQAESIKERADKIGYVMQSPNQMISKTMIFDEVALGLRLRQVEEAEVK
RRVLAVLKVCGLYEFRNWPVAALSFGQKKRVSIASILVLEPELLILDEPTAGQDFRNYTE
MMKFITALNQAGQTIVMITHDMHLMLEYAKRALVFSEGQLLADLSPAELLTNQSLIKQA
ALKETSLYQLAKLAGLPAETDFVANFINYEERKRRDE
>Substrate-specific component MtsA of methionine-regulated ECF transporter
MKKNGFSVKTVVAIGIGAAVFVILGRFVSIPTGVANTSIETTYPFLALMASLYGPVAGFLI
GLIGHTIKDATAYGSPWWSWVICSGIIGLIFGLVGSRFDLKHGEFKLKQILSFNLVQVVA
NYLVWALVAPSLDILIYSEPANKVYLQGIVSATANAISVAILGTLLLKAYAATRTKKGSLR
KD
>hypothetical protein
MKIRVKDFLAVGEAQTFASQEFSADLKQVQADLKQALGQVVTAFLEEVADHPQNLAQV
SLTIKGDSDFKLFLETGIINLAFRDIKRVDNFFSDEEAEVPLNVYLSFEAPGLNASGYRID
EITDLASAQAKPAKFLDLVVAQFESYLAMLAASEQENTTENEN
>dipeptidase
MMMEQACTTILVGKKATIDGSTMIARNDDTFLPLTPQRFIMQPAVKNRKETWVSNKNG
FSAPLPAKGYRTTMTPNADVAKNGVFGESSINEKNVAMSATESVYGNPAALAHDPWV
ENGLAEDSIQDMVAPYIDSARDGVKYLGKLIKQYGSPEGNGVLFSDKDEVWYMEIVTG
HHWAATRIPDDSYAICANQVAQDYIDFKDSKNYLWSDGLEEFVAKYHLNPNPHGKFNF
RKIFGTDNEKDRHYNTPRVWYGQKLFNPEVSQEPESSDLPFLRKANRLLSVADLQAVL
SSHYNETPFDPLGHGSEADKRRYRPISLNRTQNSHVLQIRNDVPEKQAALMWICFGVP
TFTPYLPFYTNANDTDPSWANTPLKFDMTSAYWMYRALSMLTESHYSNFIQANRDYLT
AARQAQLAHLEATDKAAKKLEGKALTDYLTKQNYKLVADMRKRTEEQIASCITEGLTLS
KLTFNMDINL
>D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
MRKFFKLILLVLLFIGGATYYYHHQVKLNQATTLASRAQSKVSPKKTVMRTPINWRLPS
ETVAYPDLKQLNDFWIKVVISKNRVYLMDGKKAVYTMYASAGEYERQAGGKKKSTTPT
GTYYIQAERGDSFYNPNLSEGANYWTSWLNHGEYLFHTVPTDANGNYKVAEAKKLGK
KPASHGCIRLSVADAKWINQNVPRVLRSLLPIKKAKLLKVQITFQ
>cell surface protein, ErfK family
MKKGNLIQWVVAVCLVLVTAYTAVHVHHANQKNAIKMAKQAKLDKKAAKREKQLLKKI
KRHERKPLSYELSSLTMPYPNVRAKQYHNTKKHPYRLWVEVSLAKQRVYIKSGRYTLY
TMNASAAKNYDKGQKEYQKTSLGLFELDKNRGPAYFDATRGYGAKSWVSFTDGGLH
RFESVPVNKDGQVIKASAKQLGQKVTKKHNVKAYGSIRLSLADAQWFAENMPAKTPIFI
HQDPKNHFPDDYEE
>DNA-entry nuclease (Competence-specific nuclease) (EC 3.1.30.-)
MGIKDDSFNQLLEEEKKLQAQIDALQKQLIENERQKQAAIDALLRKEVLNLLGFKPTDLK
QPADVTKVTDKLSQLLATKKKVESNRRQEGQTGNYVIRENKTKNRFQQANSNQVLRL
NAPKKVENPTPNARTVTNLDEALYENIPTDDINRARNSLTEGIQKSYKLATASLEFDTER
NVFVVNDNKTRLGFTQPRTQKIPTLRVDGRGRPTKASMWIVKANLPRLSNKEAIKTDP
VGFAEGFTRGHLIARSLFDDRVFKTKNTKRTNLTPQTQWSNGVLQAEIEKEVLEAVKS
HDNVRYQVEPLYHSSGEVIPRCYHIQARSRDGKFKLNVIVPNIEPGFKINYATGTMTPA
D
>contains glutamine amidotransferase class-I domain

MRINVIQHTPNEGPGAIADWAHEHGHQLYIYHPVHFDCGLPTAAETDLLVILGGPMSVN
DKDEWILEERALIAKMMEVHKPVFGVCFGMQQIAITIGGEVSSGPKEAGWGMVEKTSN
VIPKIPVALTALHWHAEQCHFPPYIAREITTLFSSEYTENQGILVGDNVIGLQFHFETLAD
NVREIVINDKDYLRGSEFKQYPDEVMAYPVPPANKQTLFKLLDFITHK
>hypothetical protein
MSQNTHVGVANDRRPVNSKNIKKRRAQHQAVLNFLKEREEAEKEAK
>hypothetical protein
MKVNEVKKMEFLKGLVWGLLFSALAWTVLVVSVVQLYWLF
>TerC-like integral membrane protein
MSLLQNLYGPFFELENWKTVITSGQDWMIILSLVLIECLLSVDNAIVLAAQTQKLPDKVQ
QEKSLFYGLWGAYIFRFLLIGVGTYLIHFWEIKVIGAGYLMYLALSHFYHMAHPVAKQKK
KKEKKRILPLFWSVVVSIEAMDIVFSIDSILASLAISTNPVIVLIGGLIGILAMRGIAEVIMKL
MDIVPELEVTAYFLIALIAIKLFLTIPMIEIEIPAAIFGALVLISLVATVVVHQFRHQK
>Cell envelope-associated transcriptional attenuator LytR-CpsA-Psr, subfamily F2 (as in PMID19099556)
MKHDTDELRTHKYHHHHRNKKRRRLMKWVLGILAFVIIIGAAGAIKIYTDTKSAIQSTYK
QVKHKDIRKGAAADVTAGDPFSVLILGVDTGEYGRTYQGRSDTIMVAAVSKDKTSLVSI
PRDTKVAISGHGENNKINAAYAYGGVSGAINTLQSYLDVPIDHYIKLNMRGLVQLSEAV
GPITVDNDLDFTNLGYHFKKGKVTIDKNNILAYTRMRYEDPRGDYGRQLRQRIVTIALV
HKLATVGNLSKYKQILNAVSANMTTDLTFSEIKSLVTEYKGADDNIQQVQLQGTGKMID
GVSYEVVSQASLNKVQTELKDALKIK
>membrane protein
MFLATIIKHLGIIYRKELIETANKNRIFLGIVLILLAAAMVLKAFIPVSWNLADICYLVFALVV
ALYGLVNRSILVTFTGLGLVAGASNRWLDLIPLGDGSLFLVGVFLGLGCHLLFAKWLVY
EWSAKQTRYTSNQAVTDFDVIMSSSNHYLHQNDNNVNADVVMGNLNLYVDENASKPI
TLNLDVIMGNVTVYVPASWQVATQTANIMGTTKIPVEMTDYAAQPDLIIKGDAVMGSIIV
KQV
>hypothetical protein
MSGTFFSLGSGIDHLSKKTDHLLKRPVFTAKKQAVNFLPG
>transcriptional regulator
MFPQRLRALRVGRKLSQKQLAEALNQTLSEGERPNTAGQIGKWELGKTKPSYLEVKK
LAAFFQVSLDYLLAQGYESIELDDLFVSTADLRLAGHILSSEERTEVYQLIKGYLNGRHP
QKKTQVDRVDEELSLDL
>Sucrose permease, major facilitator superfamily
MEKKKSLQNPFYLESSLTFILFFFGWGIWWSFFQIWLTTKQGFSGAQVGTIYSFDSALT
LVLMFVYGSIQDKLGIKKTLLNLCVFLELLLGPFFTWVYVPMLKSNFYLGAFIGSIYLSVA
YLAVSPTFEALAERMSRRFSFEYGQSRAWGSFGYAIAALAAGFLFNINPYLIFWVGSFF
ALLLLLIMLFFKPENKQVSLDKYELKEEEAKEEEKPSFKDILGVFKLVELWKIIIFIILSWTF
YTVFDQQMFPEFFTKFFGTASSGAAAYSILNSAEVFLESIMMGLIPMMMRKTGVRKTLL
IGLAIMIIRIGGCGLVTNPLGVSVIKLLHAPETAIFILAMFRYFTLHFDTRISATVYMVGFQI
AAQVGQIIFSTPFGALHDKIGYQNTFLVVAGIVCVSAIYAFFILKKDDEDVLGQPLGD
>Sucrose-6-phosphate hydrolase (EC 3.2.1.B3)
MEEIKVTNNRFRLDYHLMAKAGWINDPNGFSYFKGYYHMFYQYYPYDSEWGPMHWG
HARSKDMIHWETLPIALFPDEKDDGCFSGSAVVYDNKLWLIYTGHHYIDEEKDLFYEDQ
NLAFSKDGIHFEKYAGNPVLKRPTDNTKHFRDPKVWQENDEFYMVLGSQGDDELGRA
LLYQSQDLKEWRRVSILAKAKNLKDEGYMWECPDFFRINGQDILLMSPQGLEARGELY

RNLNQTGYLLGQLNRKTYELKRGSFEEIDKGHDFYATQTTLTPDGRRVMVAWMNAW
DSPMYEKEDGWAGALTIPRELIVKDGKLYQQPVGEMASLRQERIVKEEASGNYQLSRA
TEFSLETMALGQLFKLSDGEHTLTITREDEKLVLERNTKDGQRSAHITDPVTKLQVFVD
ASSVEIFVNDGQVTFTERMYWKKAPTLTVQAGSQLEIYRLEENVNRY
>Sucrose operon repressor ScrR, LacI family
MEKPKRTKLEDVAKLAGVSKTTVSRVLNKRGYLSEATIAKVYAAMKELNYQPNTIARQL
FKQKTNMVGLIFPTVNNPFFGQLVARLEQILYEHGYKVLIGNSQNNPQKEEDYLNQLLN
HQVDGLIVGAHNRGLKEYQHANLPVVSIDRIMNDDIPVVSSDNYQGGKLATEALIAAGC
QTIIHTDGPENLESNTSKRQQAYEDVMRKNGRQPVTYNLDFNISTAEKIKIFRKIFEDYP
QVDGIVTSNDTDAATILSLAPLYNKRVPHDLKIIGYDGADMTRLMLPHLSTIQQPIDGLA
QVAVDLLEEKISGKKPSSPKPLQVKFIRGLTS
>lipoprotein precursor
MKKFYLLAAGLTLSLTLGACANHQSKTATSSTTSQSQHSSQSKKTNQASAPKVQTQAK
ASHTQATKPTWNETKSKQLATFMTNWGASMNQSYDAYTPDKPGDFYGASIPNATYN
GQMPLAIDSVDNQVKVAWAPNGNGPKDTYLLLACYSDTAHARYLGKHVYLFASYNNQ
PQVLVSQQNQGNERNSFVFKQTANQDLAKGFAQIFAGKNIPTKFAKSPANSSNKQGL
AVSDEEALKLFYTVQTSAGDNAEFSQRLANDADSYLFTNVSGQNVTFGMGKDVTRTF
PTNTYQVLGSPSYLGSTVFQKISSNKYLVYQVPSHFDLQFQEDASFAKQKSDQYMDH
PQEVTVGTVDEQVLAQMKAKLTH
>hypothetical protein
MTKYSPEFKAQVVVEYMTDGLSSYTLEKKYKVSARQIRKWIQQYQLQGIKSFKQHHRK
RIFSAKFKWHVINYYQTHAESFAQVAAKFDLLPSQISIWRQAFIKYGYDALKPHRKGRS
QKVKKQTKKQKQKLVEKKEIDQLREELARTKQELYNTKMDRDILKKSLALFGPSKPGK
KRK
>Mobile element protein
MDQIRADQAELPKAQRYKIGDLLSRINLAKATYHDERKRIANPDDKYCEVKKVIRRICEQ
GKVRGRLTYGYRRVQARLDQLGIHLAGATIRHLMSELGVQVEIYNHHRNGKYSSYHG
RVGSVADNVLQQQFNAKDPYQVIHTDVTQVRLANQQWAYISAMTDEASKEVLAFQMS
SHPNQELIMNTLNELLDHLPDDVQPIIHSDQGWHYQLKYYTKKLKDHHFIQSMSRKGN
CHDNAPIESFFHLYKTELLQGLPPCNNITELRKLSLTYIDQFNNERISLKTKGMTPVGYR
NHALAV
>Pseudouridylate synthases, 23S RNA-specific
MEIKNAVINVMNREINSAVFSQKELDGRDQIVQEYIEKLVKKFLNSECLTLDLAASEPAQ
YLVGQNEDFIEVAEKLTQFYFEAIKTWEEIPEGDLLFFRVEDGYGKTYTGMVKLDFSAK
YIHLIDYDDDNLVTNKISQNKTILPNPGQGVSCGIIVKDQTCFIHNKVYQGIQGKWSFLE
RILEVTEKKPKISEEVAQVKKALKTVSKKYDIEQYEAVMGAQKALYDSIDEEASLSSERV
AEQVFGDNHQAKQEFLEELDQKGIQREIALEANTPVFERKYQKQKLKLDNGIEIVVPAE
LLKNKDYIEFVTNDDGSLSVILKKIESIKNNF
>Argininosuccinate lyase (EC 4.3.2.1)
MAIKKMWGGRFEAAGAQWVDEFGASIGFDQLMAYEDLEGSLAHVKMLKHTQILSAKE
ADQIIAGLETLQQKLAAGELEFSVANEDIHMNMEVLLTEEIGPVAGKLHTARSRNDQVA
TDLHLYLKYRLETVISELKHLQEVIVAKASANVETIMPGYTHLQHAQPISYGHYLMAYYQ
MFKRDLERFEFNKLHTDISPLGAAALAGTTFPIDRHFSAKELGFSAVYENSLDAVSDRD
FVLEFLSNASILMMHLSRFCEEIISWCSYEFGYLELADQYATGSSIMPQKKNPDMAELIR
GKSGRVFGDLMGLLTTMKGLPLAYNKDMQEDKEGVFDAVKTIIPSLKVFSGMLETAKV

NKEKMLAATENDFSNATELADYLATKGIPFRQAHAIVGQLVLEGLKTNTSLQAIPLARY
QEIDPKIGADVYEDLKAYVAVKRRHSYGGTGFDQVKCQIDLAKQDLAK
>Argininosuccinate synthase (EC 6.3.4.5)
MTKEKVVLAYSGGLDTSVEIAWLKNKGYDVIACCVDVGEGKDLEAIKQKGLAVGAVESI
VIDAKEEFASEYCLTALQGHAYYENKYPVVSALSRPLIVKHLVAIAKEYGATAIAHGCTG
KGNDQVRFEVGIHALAPEMKIEDPIREVHWSREEEIDYAKKNGIPVPINKKSPYSIDENL
WGRANECGILEDPWQAAPADAYARTNALEDTPDEADVIEITFEQGVPTKLDGEELPLA
QLIMKLDKIAGKHGIGRIDHVENRLVGIKSREVYECPAATVLMKAHKDIEDITLEREMAH
FKPLIEKQIADMIYNALWFSPLMKALQAFLKESQKDVNGVVRVKLFKGNVICEGRKSPN
SLYDEKLATYTSADEFDQEAAAGFIKLYGLPQQVYAQVHQKIAGKD
>hypothetical protein
MTKYSPEFKAQVVVEYMTDGLSSYTLEKKYKVSARQIRKWIQQYQLQGIKSFKQHHRK
RIFSAKFKWHVINYYQTHAESFAQVAAKFDLLPSQISIWRQAFIKYGYDALKPHRKGRS
QKVKKQTKKQKQKLVEKKEIDQLREELARTKQELYNTKMDRDILKKSLALFGPSKPGK
KRK
>Mobile element protein
MDQIRADQAELPKAQRYKIGDLLSRINLAKATYHDERKRIANPDDKYCEVKKVIRRICEQ
GKVRGRLTYGYRRVQARLDQLGIHLAGATIRHLMSELGVQVEIYNHHRNGKYSSYHG
RVGSVADNVLQQQFNAKDPYQVIHTDVTQVRLANQQWAYISAMTDEASKEVLAFQMS
SHPNQELIMNTLNELLDHLPDDAQPIIHSDQGWHYQLKYYTKKLKDHHFIQSMSRKGN
CHDNAPIESFFHLYKTELLQGLPPCNNITELRKLSLTYIDQFNNERISLKTKGMTPVGYR
NHALAV
>Mobile element protein
MLRVQKVRLYPNETMKQVLDDLCDYRRYCWNQGLALWNDMYDVSLILGDKKLRPSA
RKVRDELVANKEDWQYQLSARCLQLTISDLGKAWQNFFKKSLPDWGKPKFKSKKTAR
QGFKTDRAQIVNGKLRLDKPPGVKAWADISFKGADDLNGDLKVVSIYRENGKYWASLP
FEVKVTKKTKTGQKTAVDVNVGHFDYPEGQVKTLPNNLKTLYKRIKHYQRLLARKRVA
NGKKATQANNYVKTRAKLQRDYRRVANIQHDIVQKFTTMLVNDYDRIAIEDLAVKQMQ
MSHVASKGLHRSMFGYFKQVLKYKCEWYGKELILADRYYPSTQRCSQCGHVKTGAD
KVGLDGNKKHGTKHNEYVCYECGAVMERDENAVRNLLALL
>Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)
MQGNAGIKLFALNSNPELAQKIAAEMGVKLCDASVKHFSDGEIQININESVRGDDVFIIQ
SISDPVNENFMELMVMMDALRRSSAGSVNAVIPYYGYARSDRKARSREPITAKLIANFI
QLAGADRVVALDLHAGQLQGFFNIPVDHLQAIFTQAQYFIDNKIDEDVVVVAPDHNGVK
MARNLAQFLKAPIAIVDKRDANRLERPEEATVIGDVKGHKCIVFDDMIDTGGKMADAAE
ALKAAGATDIYACATHPIFSGNAVEELKDSPYKEVVVTDTITIPEEKQFANLKVLSIAPVF
AKAISLIYNNQSVDSLFRRDF
>Malolactic enzyme (EC 1.-.-.-)
MTKATDILNNPFLNKGTAFTKEERQALGLTGTLPSKVQTIEEQATQAYGQYQSKPSRLE
KRIFLMNLFNTNRTLFYHLMDQHLVEFMPVVYDPVVAESIENYNELFLKPQDAAFISIDD
MDNIEASLKNAADGRDIKLLVITDAEGILGIGDWGVNGVDISIGKLMVYTAAAGINPAQV
LPVSIDAGTNNEALLNNDLYLGNKHKRVYGEKYYQLIDKFVAAATKLFPESLLHFEDFG
RSNADVILKKYQDKILTFNDDIQGTGIIALAGVLGAMNISKEKITDQTFLTFGAGTAGCGI
AKMLFDELVRQGLSPQEARQHFYLVDKQGLLFEDTEDLTDGQKQFTRSRSEFANADE
LTNLEAVVKAVHPTVMIGTSTQPGAFTESIVKEMAAHTSRPVIFALSNPTKLLEASAEDLI
KWTDGRALVATGVPSDDVVYNGVTYTIGQANNALVYPGVGFGAIAARAKIINEEMLSA

ASHALGGIVDPTKPGAAVLPPVAKLTDFSQKVAEVVAQSAIDQGLSKENITDATQAVAA
LKWEPKY
>Malate permease
MAVFLTSVSSVLVIVLIMALGYILRANGWFADSFAGNISNLITKIALPASIFTAVMKFLTRD
KLVSLSQGILFPALGVIIGYLIAWLVVKVLKIKPGRRGIFMNAIVNANTIFIGMPLNNALFG
EKAMSYFLIYYIINTVSTWAFGIFLISNDDPTVDKSKQTNKQKINWKKVLPMPLVGFIVAII
WMLLGIPVPTFLGNTLNYVGALVTPLSLIYIGIVLRDAGLNNFKFERDSILALLGRFVLSP
VILILLIKFGLHTGFDLPALMRQTLVVQSATPMLAVLPILADEAHGDVEYATNVVVMSTVL
FIIVVPVLMAILQFI
>Gamma-glutamyl phosphate reductase (EC 1.2.1.41)
MTVDLVKMGQAAKKAAYELGQLSTDEKNQALLKMAEALQDKAAEIMEANALDLDQAK
KADLKASFVDRLVLDEGRIEGMAAGLRQIAGLEDPIGQVDRAWVNYAGLEISQRRVPL
GVIGIIYEARPNVTVDATGLCFKAGNAVMLRGGKEALRTNIKLVEILRSVLTELGINPDAV
QILTETSHELANEFMALTGYLDVLIPRGSARLIKAVVANAKVPVIETGAGNCHIYVDAAA
DLKMAAAIVYNAKVQRPSVCNATEKLLVHQDVATEFLPLLAAKLAQASVELRGDEQAC
QILGSSCKAATAADWDTEYNDYILAIKVVSSLDQAIDHINAHNTKHSEAIITNDYNNSQAF
LDRIDAAAVYVNASTRFTDGGEFGFGAEIGISTQKLHARGPMGLRELTSTKYVIRGTGQ
VRK
>Glutamate 5-kinase (EC 2.7.2.11)
MANKRQFDKIKRIVVKVGTSSLTHPNGKMNLQAIDQLSYTLSGLVNEGKQVVLVSSGAI
GVGMGQLNLQERPTDIPAQQALAALGQSMLMTIYQQRFATYGQKISQVLLTHDVLDYP
QSKENVNNAINKMLAWGIIPIVNENDTVATDELDHKTKFGDNDQLSAIVATIIDADLLIML
SDIDGFFDKNPKKYADAKLIPEITEINEQTYQHAGGKGSKFGTGGMTTKLKAAQRMMK
ADKLTILANGANPSIVFRILNGDAIGTLFRKG
>diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)
MRCQKNEQPGELKTGLDPVISDEILLDNLIEMADLCIFWKNKQRRYLGCNQAFLHYFGL
TEVEEICGKADQEFDWRLGSGEEIAVENALLTDGIAVHNYKKSIYAKGQWRKIVVNEKP
IYVEDEIVGLLGYFKDITQEGREQHYLETLAYTDSLTGLYNRRAFVRTIARYEDEYERNR
IDFSLLFIDINGFKKYNDTYGHSYGDEVLRRVAACIRRVTGYTGVACRIGGDEFVIIHQD
RNQNYIKDAENKLHQAVAEIEFIQKRPCKLELAMGRAVYSEVNDREELLRFADQRMYL
NKQKRR
>diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)
MTKNNQTNELTGLLTISRFREVAHQTLENTDLRAQGVSFIYLDIENFKNYNEIMGFSAG
DEVLQFMAKTIDDEFNNRHVAYFGSDHFVILARNSEALVKTKTIVATLNAKFGQMSVNV
KAGIYTLQPDDDIDISVCCDRAKIACDSIKHKYDASYCFYTNEMGRDLWLRRFITDQFPT
ALASGHIKVNFQPIVRALTDDVCGLEALVRWNDPDYGFISPGQFVPVLEQAHLVHKLDI
FVIEEVCRAYKYSLVDSNLATVPVSVNLSRLDFSLCNIYEEVERLIKKYDVPKDMLHIEV
TETGLNEEGNFLRDGIIKFQENGYQVWMDDFGSGYSSFNVLKDYDFDVLKLDMKFLA
DFEKNENAHIIIASIVSMAKKLGIRTVTEGVETKEQWEFLKSIGCDMGQGYFFNRPAPLE
QIWESMTAKQRGNEELTNRSYYDAIDRIDLKGSNPFSYERKKQLDQGLPLAIAELRNG
SYRFLKATPGFVKLIRTLGVKSVKECEELFNDTNSNINKQINRLLENSTQTGEVSEIDFV
MNGNYANISVRYIDQDANSTAFLLEVANLSQNRNLARSKKREEAMRFIYSLFDRVELIN
LSEGTVENLYRNNMHTSSILRGGAIKEVVDNYAQENIHPDDIAKFRKFYDENTIEERIKA
EGGKKYITEYLRTKDENGHYCWKLYCVVPFPLNNNRMLLSCVFDVDAKKSLKLIGQMA

IKKDVLSDSDGKISRAILWANLLAMSQFGVYWKDTDRRFIGANQTYLSYNSLKSEQELI
GKREEELGWELNPDFRIDEDQKVLTEGEQVVNREGTMIVDGNLRNIATSKSPIYDTEG
KIIGLMGYIIDITMESRDSKGLLNLAFTDLLTGILNRRGLMDALLEYEQMYLKRKADFVIA
YVDVDNFTDFNDKNGRSYGDRLLSEVAYALRNRLGKSCVIGRIGSDEFVLLKRVKDLG
DASQLQALIRATVADIHQIGNKSCQVSASVGTALFSEVDNTQQLLRKADRKMQMDKHG
HSRLTKMFEV
>Ornithine decarboxylase (EC 4.1.1.17)
MAMEKYLEIAVSADFSQLDLSPYKRVALEAVTAPAGLAAIVVGASDLVARAQAKLLKQQ
AHLDIPVIYGDKQTPAATLAKEVLRQAQAYENEMVPAFLTDLINFADEKPTSFTTPGHH
NGQFYERHPAGMVFKRFFGDKMLTADVSDTVPELGDTLTHGGTPLTAEQKAAQVFHA
DKVYFCTNGTTSANTICASALLAEDDLVLFDRNNHKSLYNSALVMTGATPVYLRTDRN
PLGLIGPLVAADLDEAKIRAQIAKVAPKKANAPRPFRLAVLQLETYDGVFYNARWLLEKL
GSLCDYVLFDCAWGGYEQFVDLMADLSPLNYEYGPDAPGILVTQSIHKQQAGLAQTS
QILKKDHHLRGQKRYVDHKHFNNAYLKYVTTSYSYPIYASLTVNAYLADGQANKAWWE
ETLRLGIEFRKQLLKRSKLFKPLVPELVAGQKWSEIPTEVLASDRKYWQLDPKESWHG
FNKAALGEAILSPLKLTIQTPGVNLATGEYEASGIPGPVVGAYLTERRIIAGKSDLYSNLF
LLTPGESRLNMQTLLEAYLQFEKFYLEDAPLEKVLPRLVEQTNGRYQGYGLRQLCQEM
HEYYAKLKINRLQQALFLKPNLTGFVAKPQKADALFMKNEGRLFSLDNLKGKVALEGAL
PYPPGVFIVAPGERWQQLDIDYFKALVGSMEQFAGFDPEIQGVYITKQADNHLAVEGF
ALDND
>Mobile element protein
MLRVQKVRLDPNETMKQVLDDLCDYRRYCWNQGLALWNDMYDTSLVLGDKKLRPSE
RKVRNELVANKEDWQYQLSARCLQLAISDLGKAWSNFFKKSLPDWGKPKFKSKKTAR
QGFKTDRAQIVNGKLRLDKPHGVKSWADISFKGANDLKGELKVVSIYRENGKYWASLP
FEVKVTKKTKTGQKTAVDVNVGHFDYPEGQVKTLPNNLKALYKRIKHYQRLLARKRVV
NGKKATQANNYVKTRAKLQRDYRRVANIQHDIVQKFTTMLVNDYDRIAIEDLAVKQMQ
MRHVVSKGLHRSMFGYFKQVLKYKCEWYGKELILADRYYPSTQRCSQCGHVKTGAD
KVGLDGNKKHGTKHNEYVCYECGAVMERDENAVRNLLALL
>amino acid transporter
MNNLNANSLEPQRKYTPWLIIALLDFVTIISFENIFYPFQNQGLSVVVSWIFMFFAYVIPY
ELITSQLGLTFSNQEGGLASWIRRGTKSDLLGYVTSWMYWVQTIPYLVDVSNSVIVSIS
WIILGDASLGKHMSTTTFGLLTFVIILAFILLENAFRNSLEILSLIGGGAMFIMSMLFVLMT
GYALTHGAHIATQPFNLSAFKPNFSLHYFSTTGLLIFAMSGAELAAPYIIRLKNPKVEFPK
AMWALAIMTGCMTIFGTLALAVFFDANHIPHDFKMNGPYYAFKLLGQELGWGNVLMYV
FAVVQAFFMLAQLAILIDAASRVFAGDVHEKYMPSWLLKKNKHGRPIHSYTLTAGISLVL
LLLSGTLPNINSIYNWLLNLNGIVSPYKTCWVFVAFLAIRWQQNSFKSDYIFIKNRSFALA
VGFWCFLFTFVCATMGFIPQDSAVGSKQFDHELMMNFFFVFLLFGMGFIMPLLRKLED
KKKQA
>GGDEF domain protein
MGQFVIDSLLEWAEINEFLRYTIIAVIVNVMGIWLYGYFMIQTITNNYILFVTMGSLLMQFL
VTTYSSVKEKQARDFERARYNEEHDALTGLYNYEAYGKEIKRLSGSSQDYNVIMFDID
HFKQINDTYGHLEGNGVLQYFGTCLEDYTHTKVDDSALVYRFGGEEFCVIMPKMKQE
RVKKYFLNFQAELKKARYQTEDGRLVELTFSGGIASTGKSAGISNAMKKADNALYQAK
NLGRARIEVCYEN
>Integral membrane protein

MAEQTNREKNQVASQQQAQVREQEVTVRGGKFDGLTKRNEDYLFHLDKALDERGMD
YEKKNDILETMYQELKKVQKQGYTATKLYGPVQEKAKELVDGPAKKEAAPAQNQKFW
VVALDNGLIMFILFCAMYSLMGFFGGNSQTQASGGWITLLTTSVIAGIGLAYFYTAMLPE
NLNKHKHKWVHMILSTIALIAVWMVVFYLVALIPASINVVLSPIIYAILAVVGFGVRYYLKG
KLGFQRVFR
>GTP-binding and nucleic acid-binding protein YchF
MALTAGIVGLPNVGKSTLFNAITKAGAEMANYPFATIEPNVGMVEVPDERLARIDELIPA
AKIIHTTFEFTDIAGIVKGASKGEGLGNKFLENIRQVDAIVHVVRAFDDDNITHVANKVAP
LEDIDTINLELVLADLDSINKRYARVEKVARTKDKAAVAEFEVLKKIKPVLEAGKAVRSIE
FNDEEKKIVKGLFLLTSKPVLYVANISDDYIANPEESKYYQEVKQFAEAEGAQALAISAR
TEEEIAELDEDERQFFLEDLGVKESGLDRLIKASYKLLGLATFFTAGGKETRAWTFKQG
MKAPQVAGIIHSDFERGFIRAETVSFEDLDKYGSMQAVREAGRLRSEGKEYQVQDGDI
IEFRFNV
>FIG001891: protein involved in chromosome partitioning
MYNLHDIVEMKKPHPCGLNRWEIIRLGADIKIKCTGCGRLIMLPRREFEKKMKKVLETA
ENN
>Chromosome (plasmid) partitioning protein ParB
MVEQQALNSKWLAQLANGKDDQVVKLSLDDLRPNPYQPRKVFDQGALEELANSIKKT
GVFQPVIVRPSKVKGYELIAGERRFRASIMAGQKQIPAIIRTISDKEMMEIAVVENLQRE
DLNPLEEAQAYATLMEKLHLNQTQLAKRLGKSRPYIANYLRLLGLPPKVKQSLRAGELS
MAQARTLLSLNDKSKLAQLAQKTVSEGLTVRQLTELVAKLNGTNKVKPKPVKATNQYA
LYYQETAQLLAEKLGTKVEVKPQGKKAQGKIEIDYYSTEDFNRILEILQISLD
>Chromosome (plasmid) partitioning protein ParB
MAFSFFKSKKEVSAKPGVNDVVELEVQKIIPNRYQPRKLFSDESIAELATTIAEHGLLQPI
VVREYEAGKYEIIAGERRFRAVNRLKWEKIPALIKKMDDTQAASMAVIENLQREGLSAIE
EAQSYQKLMELNQLTQLQLAREIGKSQSFVANKLRLLKLNMGVQRAILNRKLTERHGR
ALVSLDDKQQVKLLKQIEEGHLTVKETEALVAKITAPKVKAQPKKKAKARTKGSSKDVKI
ALNTIRQSVKMVTDTGLKVTTSEEEAPGVHRIIIEIPLDGSDQK
>rRNA small subunit 7-methylguanosine (m7G) methyltransferase GidB
MTPEEFRQALAAQGIELTARQLEQFAIYYQFLVETNQHLNLTAITEEKEVYLKHFYDSIL
AGLAVTDLRKEELSLCDVGAGAGFPSIPLKIAFPQLKVTIVDSLNKRINFLNELAAKLDLT
GVAFFHARAEEFAGKKSPYRESFDIVTARAVARMTVLSELCLPLVKKGGRFVALKAQK
AESELAEATYAIATLGGKFSQDIELTLPKLGDERHIIVITKKKETPKKYPRKAGTPAKQPL
VAK
>Nicotinate phosphoribosyltransferase (EC 2.4.2.11)
MNPEKLALLTDLYEFTMANAFYQELPKEEGVFDIFFRNLPDNGSFVVLAGVKQAVEAV
EKFHLSASELDYLRELDLFSEEFLAYLADFQNQCEVWTIEEGTPVFSGESVVTVQGPLI
QAQLYETMLLNLINHQTLIASKARRIVASAKGRAVMEFGARRAQGPDASIYGARATIIGG
AKLTSNLLAGKKFNIPVAGTMAHSWIESFPTELAAFEAWARQNPHNCSLLVDTYDVLN
SGIPNAIKVFEQLKAKGVDANIGIRIDSGDVTELTKKARKMLDQAGFTEAKITISNALDEYI
IESILDEGAQVDNFGVGEKMITSSTSPVLSGVYKLCAVKHEDGQVEPKIKVSASKTKVT
LPGIKQVYRLYNTSGQAFADVIALRDEKLGHSLSVVKADPLATKKEHVLQDFIAKPLLRP
VLRGTKRVEELEGDVFKIQANSQAALKELPAETKRLMNPAIFPVYVTSKLNTLQKEMLA
KH
>Nudix-related transcriptional regulator NrtR

MAGVELLERPLINIVNLIWSFDRASEQVYLLLVKRASTPYKNRWALPETFMRTKESADQ
AALRLVKEKIGLELASFHTEQLATFTNPKRTKKERAISLAYMTFLPEMPALRAGYGASE
VRWFRFGYQAEKYCLVSGQKLFWASQAQTEQAYYDQPLASGQAHLAFDHEWIFQVA
ATRIKNKLDYQPNILLILGNTFTLKEARSVYAPFLMTTLKQIDNSNFKKTHGHLLKDVGIA
TAGRPGRPPRLYRLKLE
>Probable lipase/esterase
MPSFTLNEPTLLQNPRLHLSLPYAFKSERALRLHLIEPATSKGHYPLVIFSAGTNFTNDD
LSHYFYPLMQLAQQGFVIALVEVRPKDVAPFPAQTSDLLSASRYLLGHAYQYQIDPYRY
FLWGHGASASYSSLLASLTATNQAFNDEDCRMQALNYQACICFGLTSLKPTSSYFANQ
SADLFSYWGKRDLKDNLYIQVGSEADNLFLNLGLLAHPSQKESLVYYTPAKKALDDSF
FSPYLINIVLNFLKQKVKKK
>hypothetical protein
MVNLIQNYSTKKLHDREAMEKKEYVSNQFRMSDLMGSLTV
>metal dependent phosphohydrolase
MNSQAQLTTIKTFAKSKLAADKTGHNFDHIQRVVKLANRLAKAQTKPVDLTVILAAAYLH
DTIDDKLVASPKQAVNEVKELLTSLAFSSAQSDEIIHIITHMSYASNLNEKQVLSLEGQIV
QDADWLDAIGAIGITRAIYYGGAHGEKIYDPAVLPRTQLDKASYRNLADETIINHFYEKLF
KLADMLNTPQARALAASRQQVMKDFVATFKAEWDAEA
>LysR family transcriptional regulator YeiE
MRSFENDQQFIKILETIIEYGSVSKAADYLFISQPSLSKKLKQLERDLGVTLIDRKHHPLR
LTEAGEFYLTNLKRLTSQYEAVYHDLTRMASSNHGQLTLGITQSLGQKILPLILPQYHKV
YNDVNIQIVEDTSEHNEERLIHAKIDLYLGILPIFRSDVHYEQLASEPAYLLVPQSHPRYV
ADQPKQKLASVKDILPFVTDFPYISASENVGFQRLITNTLAENGITLKHFLRCSNLLTIAQ
LAVRNLGCTIVPKSVANKVKNIQPANYYELPVDDLSYKLVIAHKENEELTAQMTAFIKMA
KGLGDRLLA
>Zinc ABC transporter, periplasmic-binding protein ZnuA
MSISKIFKAFLASLTLCLILGLSACAKNQTQTKADQLKVVTTTDFYGEVAQAVLGSHGKV
TPIITSPSVDPHDYEPTTKTAKLVAKADLAVYNGLEYDSWIKKLGAKKYLTVASVTGAKI
GQNEHLWYQPQTMSKLALALAKEYGRLDPKHKQDYLANAKKYQAKLAPLTKMVAQIK
ANSRGQKVAVSEPVFDYSLEAMGYHITNKHFAMAAEEETDPSYTDIKKLQTAIKNKEIA
FFVLNTQSDSKVITNMVKLCQEYKIPVVKVTETLPAHQNYVSWMKKQYQQVLTLQK
>Zinc ABC transporter, ATP-binding protein ZnuC
MDSILTVTNLDLDIPKRRLFNHLSFELPAHKLTCITGENGVGKTTLVKHLLDDLARKNQN
HVSFQITREQVQYVPQLRNIDDEYPLSIYDFVALGFKKRILPWNSPKMKATLNQILAETN
LTKIKYQPLGRASGGEKQRAYLAQALCANPELLILDEATASLDVVSKHELLKLLTNIMKK
RGLTILFITHDPELIEQYADYELHLANQTGHLIKRGQA
>Zinc ABC transporter, inner membrane permease protein ZnuB
MFYYAFMRYAFTAGIFIALICAFMGVFVVARQTSFFTHTLSEIGFSGASFGVFMGISPLL
GMLTFTCASAILIGFSGEKLGRREASISVFSGLFLGLGILFLSLSDQQSSYATSILFGSIV
GIDAHNLHTLIGLSLVLLVTIFFMFRRLAYNSFDGQGAEYNQRYNQLVSIVFLVLLALTIS
VTAQIVGSLLIFVLLTIPASAAKYFAHSLKKMIGLTFIFALTGIWLGLYLSYLTDWPVSFFIT
AIEAVIYGGAIFKEQLATAHK
>Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA (TC 3.A.1.5.1)
MQTEMKKKLKKTSVVLMGALSLATCFPTTGVKAASKQTLRWSEPAEMQTLDPSLVQD
QVSSDVIANSNEGLYRLGKNSKVENALATSTKVSDGGLKYTFTLRKDAKWSNGDPVTA

QDFVYSWRRTINPKTASAYAYLFEGVKNASDIAANKKSYKTLGVKAVGKHKLVVTLDK
KVPYFKLLLGFTVFYPQNQHAVEKYGKKYGTASKYMVYNGPFKMTGWSGSNLSWKL
VKNPTYWDKKDVKLQEMTFQVTKSTNTSYNLYQSGKLDQTYLDATQARELKGKPDFV
SYKQARTNYLEFNQTKKVLQNTKVRQALSYAVDRKQLVNKVLGNGSLTVTGIVSKGLT
QFDGKDFSVLQDTKTGANYNLAKAKKLWQEGLKEEGVSKVSLNLLGDDDGVSKNVTE
YLQSQLEQHLKGLTVTVNNVPKKTRIAKSESGDFDLVLSGWAADYSDPVSFLDMFTTD
ASYNNGKWSNKEYDQLIADAKGKDATDKSARMNDLVKASRLLNNDQGVAPLYQLAQA
SLVKTNVKGVVYNTAGVQFNWKGAYLK
>transposase
MEIQITIKVKLNLANADIASSFTNTMEQYRLACNYVSDYIFNHDFDMKQSQLNKVLYTNL
RSLFMLKSQMAQSVIRTVIARYKTVKTQMIRNPYKYQDINTGEWYRETRDLTWLQKPIA
FNRPQVDLQRNRDWSYLSSGQLSINTLDGRVKVNPICHGFDQYFDGTWSFGLAKLLK
TGGKWYLHISATKTVANFDNQTVKHIVGIDRGLRFLATTYDEQGKTVFFDGKAVMRKR
AKYQKLRAKLQAKGTKSAKRRLKKLSGRENRWMTDVNHRLSKTLVQKFGSNTLFVLE
DLTGVSFERTDLPKALRNQNKSWAFYQLAQFLTYKAHLNNSEVIEVSAQYTSQHCPKC
GTIKKSNRNHDLHEYHCSNCGYRSNDDRLGAMNIQLLGTQYISGQEQPKFELTTNA
>transposase
MVLITGGIALTLDSLESTFYPFASLCCVFTEQKCYNY
>Substrate-specific component FolT of folate ECF transporter
MLKLTWHSPKLATKILTLTAILVAIHVVLGKISVGSDSLFKISLGFIGTALIGYFLGPWLGG
LAMVLSDLIANTILASNGNFFIGFTFSAFLSGVIAGSLLHKQDLTWYRLAFYEFCQLVITN
VFFTTLWIHILYTTPFKTLLLARLPKEIIFWPIETIVLIVLFKALSKHKLNLD
>Hypothetical protein VC0266 (sugar utilization related?)
MVENLTKIFAEIMADPQNETYTRAGIEPLYYANPQAKLLIIGQAPGKKAQDSRIVWNDPS
GDRLRSWLAVDRQTFYESGQVAVLPMDFYFPGSGKSGDLPPRKGFAQKWHPRLLAA
MPNIQLTILVGAYATKHYLNLKASTKLTTVVKNYQEYLPTYFPLVHPSPRNQIWLKKNP
WFEAEVIPALQARVRKILS
>D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95)
MMDIKTYNVIAKEGLAEFPTDLYTVNQSQKPQALLIRSQDMHKTPFGTNVLAIARAGAG
VNNIPLVQATSAGTVVFNTPGSNANAVKELVLTMLLLASRPVLDAIPWAHKLGGADVSL
QTEAGKNHFAGSELLGKSIGVIGLGAVGSRVAKAASDLGMKVIGYDPYISVEHAWQLS
NDIPRAESLEDLLAASDFVTVHVPYTEKNRHLLGKEQVEAMKQDAVLLNYSRSGIVDD
AAVLAALDAGKLRLFISDFGTPALASHKKVIVTPHLGGTTAEAEVNGAKMAARTLRKYL
ETGNIVNSVNFPTVEMAFEAPLRLTLIHQNVPNMVGRITTILANYNLNIDNMINRSREKIA
YTVIDMATVSAEQQAELTKELLHIPEVIRVRFLKNPRA
>Phosphoserine aminotransferase (EC 2.6.1.52)
MTEVYNFAAGPAVLPVSVIKQMQAELPSFKDSGMSVLEISHRSSLYEEMFHQAKSDLR
KLMKIPDTYEILFQQGGATLQFTATALNLATKQKRIGLIDSGHWSQRAGQEAQLVGVEV
DPVASGQASGYTELPYQLNLTGAYDYLHITTNNTIEGTMYRQLPATEGNLLVGDMSSNI
LGQSYDVNDFGLIYAGAQKNIGPAGLTIVIVKKDLLGQAQALPAMLDYSKQAKKDSALN
TPPVFAIYAAGLVFKWLLDLGGVATMQERNEKKAALLYDYLDNSTLFKPAVTKVGDRS
TTNVPFVTGDEALDARFIEEATAQGLLNLKGHRLVGGMRASLYNAFPYEGVEALVAFM
EKFAKTVGEK
>hypothetical protein
MNNTTKKILFGTVGAAGLFLASSASVNADTTHHVVKNDTVWDLSQKYGVSIQSIEQLNK
INTDTHLIREGQTLTIPTTGKTAAKTATKTSQVTVKAGDTLWALAQKYGTTVEKLRQLN

GLASDAYLIHVGDVIKVDGQVTTTTTTNAQATSTVSSEANQTSQAASTASSEASQTSQ
AASTASSEASQTSQAASTASSEASQTSQATSTASSEASQAASTSQAATPYVAANHVTY
TVQAGDSLYTIAQKYGVTVDSLRQANTLGATLQVGQSLTVNDPTKNPQASVAASQAE
QTPAQTTTTSQATQATQTTQAAATTNQVAQTTSQPAAQTQTSQASQTTQTSQQTQAS
QAQVATPSSYSVSALLSYAQTFTGVPYVLGGTTPAGFDCSGFTQYVFNHFGKNIGRTT
YQQQYAGTKLAVSSAQPGDLLFWGAYGSAYHVGIYLGGSSYIAAPEPGESVSVKSFTY
YQPSFAVHVN
>hypothetical protein
MAMDKKLLWLSHVNLVIMAVLTAIVGNIALAHSKMKMPAQALKTAQVAGTVTAVIATIIG
CLLVALILKGVVRLTGETMSFKDAYVTYLISRYPVIFYSLVEVGYALVKKQLISTDFSNYA
NILMVIFSQAIMYVYLDGRLSKKQNLIVNIVFAILLSAGSIYQLVK
>hypothetical protein
MNRCKRNEKKSRNEVNGKLRKAHGHGLNWQFEQQSKILSKK
>hypothetical protein
MRKVVLFWVSCLVFLGMVGSVWASNNSKVKVYNSQGKLVKVYSFKKQPRQVELVSDI
LGDAGGKAEGKDNFLFKRPRHAKLKATYVLFQNGKRSTAKLYSNREALIKVKLLKVKV
KLTKKQYRQLLDLSK
>Neopullulanase (EC 3.2.1.135)
MVAMELAAIYHRPNSEWAYMSTPDEFVINLRAKHNDLSKVEILYGDPYDKTLTEANTSK
WNYQVKELVKNLRGVSADYWQVKLSIPASRRLAYAFRIYGQESEQFLYDDRKLAPFSE
TSLHDLVAFRTPYLDPAEITGLPTWVAKTIWYQLMPDRFANGQPTNDKEGTVAWDSEE
ASATNFFGGDLAGIIDKLDYLQELGVNGLDLTPIFTAYSNHKYDTADFWSVDPAFGNKE
DLKRLVDEAHRRQMKVMLEGTFDHLSDFSLQWQDVHKLGAKSRFAKWFKVKAFPVR
YTLSADPNYAPDASYAMFANNPHLPKLNLHEKAVQDYLLEVLTYWVQQFDIDAWRIGT
ADEFPPSFSRRLVAHLNKIKPDFYLVGENNDTGINTEGRPLFNGGVNYPLVEIMKDYFL
DEKAAVGWLIEATNTQLMRHANPTLQGMLLSLDGPNSPRLLDLCHQDKQLMRALLAY
SFLQIGSPVLFYGTELGLTGASVPANRACMKWKADEQDETMLRFIKVLADFRKAYNPLI
TKGSYEWGQYSSKYDYLSFTRRLGKQRIFALFNLGYSSIKFVLPANAKLIMSQNLIETDS
RLGQYGFVIIEA
>Succinate-semialdehyde dehydrogenase [NAD] (EC 1.2.1.24); Succinate-semialdehyde dehydrogenase [NAD(P)+] (EC 1.2.1.16)
MAYKTTYPYTDEVLHEFPNATDAQVEATLAKAHALYKKWRKEGLAKRQATLHKVADLL
RADKDKYAEIMTKDMGKLFTEAQGEVELCADIADYYADKAEEFMKPTPLETQAGTAYY
LKQATGVIMAVEPWNFPFYQIMRVFAPNFMVGNPMILKHASICPASAQAFDELVKEAG
GEEGAFTNLFISYDQVSKIIADKRVSGVCLTGSERGGASVAEEAGKNLKKSSLELGGN
DAFIILDDANWDQLAKTIFFARLYNAGQVCTSSKRFIVMADKYDEFVALVKDAFAKAKW
GDPMDPATTLAPLSSKQAKEDVLAQIDLAVKNGAKVVYGNEPIDHPGYFVMPTILTDIT
KDNPIYNQEIFGPVASIYKVDSEEEAIELANDSSYGLGGTVFSSDPEHAARVAAQIETG
MSFINSGWTSLPELPFGGIKNSGYGRELSPLGFNAFVNEHLVFTPKGLEH
>Calcium-transporting ATPase (EC 3.6.3.8)
MALKYYQLDAAQVLAQLHTQKTGLSQTEAATRLQATGPNKLSDHKQTSLLTKFILQFK
DLMIIILMIAAFVALLAGEGSDAIIIFLVIILNAVFGVFQEAKAENAIAALQKMTTPKSRVLR
DQTVQVIDSEDLVPGDVVLLEAGDIIPADVYFLEANALKVEEAALTGESVPVEKQASPLP
AKDIPLAEQTNLGFMNTNVTYGTAQAVVVETGMATQVGQIAKMLTEVDQTVTPLQKNI
SHLSKVLSVLILVIAAFIFAFGTLTGHESPLDMLLTAISLAVAAIPEGLPAIVTITLALGTQK
MAKKHALIRKLPAVETLGATQIIASDKTGTLTQNKMTVEQFYTDGQLKPAAQLKFQADS

KLALAMLLANDSQQTPTGLVGDPTETALIQHFLGQTGFDWDKVCTAYPRTASLPFDSD
RKLMSVQTGELLLTKGAPDELLRRANFIEEAGQVTPLTKAKRQELAKVNKQLADQALR
VLGFAYKKLAPADQLNLEAEQDLIFVGLIGMIDPERPEVKAAVAEAKQAGIRPLMITGDH
MDTARAIAIRLGIIAPTDTKGVISGAQLDQLSDQELAKHVGEYSVYARVAPEHKVRIVKA
WQAHGKIVAMTGDGVNDAPALKTADIGVGMGITGTEVSKNAADIVLADDNFATIILAVK
EGRKVFANIQKAIQYLLSANLGEVLTLFLMTIFNWSIFAPVQILWINLVTDTFPAIALGVEP
GDKGLMDQAPRGLKSNFLSNAVGFNIIYQGLLEGLLTLGVYLYAICWPVHQDPTLIHQD
ALTMAYITLGFLQLVHAFNCKSLTASVFSKQIFANKFFNWAILTSALLLAATIFIPQFNALF
HVSPLNLSQWGVVLLAGALLVIIVELVKAYQRTKK
>hypothetical protein
MIEDAVQNQRDLDSLLTSLKAIFKDQAHIDELLAKYYYR
>3-ketoacyl-CoA thiolase (EC 2.3.1.16) @ Acetyl-CoA acetyltransferase (EC 2.3.1.9)
MSDVVIVSACRSPLGKFGGALKDVSAVELGQSVVKEALARAQIDGQLVDEVIFGNVLA
AGLGQNPARQIALKAGLPQKVAATTINMVCGSGLKAVCLGASMIKAGDAQVVVCGGS
ENMSAAPYLLPSMRFGARLNDSQVIDSLIHDGLQDAFSGEHMGLTAERVAAQFKISRQ
VQDEFALASQLKAQKAQTTGRFEAEIVPLAQLASDETIRYDTSLEKLAKLKPVFKANGT
VTAGNAAGLNDGAAALVLMAKEVAQARGLDYLAEIKAYANVGLDPQVMGLGPIEAVKK
LVAQSDLQLADFDLIEANEAFAAQALAVLDQLQLDASKVNVNGGSLALGHPLGASGARI
LVTLLYELKRRQANLGLATLCVGGGQGVALAVQAK
>ABC-type Na+ efflux pump, permease component
MSKKQIVALGVFSRRLRSASYYWMIFGPLLMCLIIGGINYFVAKSVTKSEPKIAVVANPQ
VKAALMSLPADGYRLDKKLTLTNQKQIKQALAQKKVDGVLVIKGDFKSARYLYQAGND
KNQPTQVLTAILSKLKSSLAAQKAGLSPKQWQEITAAVPLKTEVIRQSNQKSEDDKQIA
KGISFVVVVILFLVVSSYIGIVGSELGNEKSSHLIEALLASVPAKEHFSGKMLGIAYLISFQ
LVTYLLMGGLFYVGLQVAGYGKLVNDLHLEKYLTSFGPQFWLLICLIIVLSLVLYVLLAAL
LASFVSRVEDISQLSSMVGMLVMVPYIFALIAMYAPNVLAVKILSYLPFIINGPMVVRLAL
GEASLSAGWIAVGISALAAIGLYFLAVNAYQRNALTYSNKSLFVTLGQALRRK
>ABC transporter, ATP-binding protein
MLTVTNVTKKFGDLTAVCDLSFELEAGEVLGLVGQNGAGKSTTFKMILNFLEPTAGSLS
FLGHQLGPNDLNEIGYMSEERGLYMDMTIEQQVVYFARLHGLTTKEAKARLGDWMDK
LHVKGKVTSKIKELSKGNQQKVQLITTMIHEPKLLILDEPFSGLDPVNVAALIKVIKEARS
KGAAVIFSSHNMQNVTEVSDKVLMLVNGKRRLYGDLNRVRAGFERQKLYLEGDFDW
QSLTGYQAIKDDHPGKILTFESEQQARACLEQVKNSPNLEGYRLLEPSLDDIFKMVLGG
SKHE
>hypothetical protein
MIKSKKYLAAKANELVAAFAKQAPANPNLEEVIKRAQAELEARQIAPQNIVEQLVLDSYR
ICRLSHLKVNDEAFQLLKEMEALSRERSILPFFRYDPWT
>hypothetical protein
MTYVIVKDEGVGARYYRRKDNNFTMDLMLASKWDTLSRAEEILAELSKTGAQELAIEET
AKLEQQVKQEQVIKKDTKLREYAHQLLAAKDQAFYAELERVSRFLAASPYLLDEITAKIS
QEDGTVLQDLLHTIELGDYRAHEGLALLKDLRASRKKRRYYKDRQQAVNYLLAKTNAG
QVNTTSDKAQAALTKLGNLQAERNYKYRNSGLQAKYAKLLGKQPVQSRVQLFTRLRK
C
>hypothetical protein
MTKILKREANYPITIPAKVRFDLRLTANEKLLYGELYTFGGMAGIFTVSNQRLAALYHVT
VRSISAWLNKLAKCGYIELANSLDKNERIISFVPGVDPLDRDGYLYDEEVGHEEVIEAVA

ESADKYDYTVDTKVDPEVIEAEMTERFGAIPDHLTPEQKAELREFRKLAYNKLHSKNSY
DPKLRKEAQLGKALYAEEKITEFFNKLKSKETEPPLEKQIYMMKQITKQANKEYDWLHE
DEFEQETSMEENFYPLEKKSSRVEANFQQEDNFQGLENVDNTGVEENFMGKNTSRV
ENNFQSKLDNINSIVDKIIVGFADIIKNINKLSDLRTTKNLISNQSA
>DNA/RNA helicase of DEAD/DEAH box family
MKIISQKQALKDSLLNSFLDKNKFEANDEYAAKLIANTEGETMYDRIFDGVKHCQSFTF
AVAFIESGILNSLKVALKNRTVQGRILTSTYLYFNKPKMFKELLKLPNVEVRIYEQNHGK
FHAKGYLFQHEGYQSALVGSSNFTTGALQVNKEWNLLFTSYNDGQLTEDISAEIEDQW
DAATPLTAEWIEKYEKVYHLNQKFVKSASNIVPEQVDENKGNYSSIQPNKMQAEALKN
LQNLRLLRETKGLVISATGTGKTYLAALDVKNYQPKRFLFIAHREQILKKALESFHKVLG
GDKDSNYGILSGNSKVIDAKYLFATVQTLSKDENLKQFKQDEFDYILVDEAHHVGSKIY
NKVLDYFEPDFLLGMTATPERSDDKNIYELFDYNVPYEIRLQEALEEEMLVPFHYIGVS
DYTYNGEVISDDTNLKYLVSEERVKYVVKQSEYYGYDGDVLHGLVFVSRKEEAREIATL
LTKYGYPAQMLSGEDSQEEREEAIEKLRLGQLTYLVTVDIFNEGVDIPFINQVIMMRQT
ESVIVFVQQLGRGLRKAQGKEYLTVIDFIGNYKNNYLIPVALTGDQSRTRDGIREHVET
GKIEGLSTINFELVAKKRILDQLKGSKNLFTLEIKKDYQELKRRLNRIPLIADFIKNNMVAS
QAILSVLNKKGNYNDFLRKVNKDAAYSWEKGYLDTTLTFVTTELLNGKRVHELLLLQKIL
RDGKSSRAELEELVLAMGSYCNEDVLDSVERILTLDYYKPKGYDDRSEQDKTKTEFAK
LGGRSLIRIDEEKGYRFNDTVSELLDSTPGFRELFEDVVESGLLNAKEYNLRELFTLEK
RYTRSDVNRLLNWQRRVPAQSIGGYFIENGYCPIFVTYMKSEDINESIKYEDRFISNSKI
HCYTKSGRRLGKGETEKMFAGVVEGAPKLEYLLFIKRSDAEDDDEFVYLGTGKVDGN
SLKQEYRELEKKGKIVNEPIVSYDLKLDVPVPFMRYRMLTEENN
>Mutator mutT protein (7,8-dihydro-8-oxoguanine-triphosphatase) (EC 3.6.1.-)
MKNIYVVGAALIEDGKLLVTKRNSDRILGDLWEFPGGKIEQGELPQESLKRELKEEFND
EIIVGDKVTETASYEYEFGTVHLTVYYAKFLSKNFDLIAHSKVKWIDPAEALDLNWAPAD
IPAAQAISKVNLREVRFDEDH
>Lipoprotein
MKKQVIYLATTALLTLGLAACSSGKQEQAKTSSSQKTSQVKKASHKKAEQKSSSSQSA
KKQAKTKNQQASVWDQSKTQQLTTFMTNWGYTMQQSYQAYSPQNPGHFRGITIPTG
VYPNGLAPLAINEPANKQTAAWSNDGTGPKNTYNIVASYVYHNDLGAIDVYAYLFAFYN
GKPVALVTNQTEANTEGTTVFKETANPDVKAAFAQIAAGKGIPAKYASPKQKVEANTK
MTTDLALRVFWSAKKAEDANWGLDNVTFLFMRDVSNHHVYDSDNINAVFPANTYMVG
QSRASANDVAFQLIGNNKAKVYYIPGSLGMTTADVDPNDIVNNAMSHPQEVEILNVDS
ATLDALKAKLNQQ
>hypothetical protein
MTKYSPEFKAQVVVEYMTDGLSSYTLEKKYKVSARQIRKWIQQYQLQGIKSFKQHHRK
RIFSAKFKWHVINYYQTHAESFAQVAAKFDLLPSQISIWRQAFIKYGYDALKPHRKGRS
QKVKKQTKKQKQKLVEKKEIDQLREELARTKQELYNTKMDRDILKKSLALFGPSKPGK
KRK
>Mobile element protein
MDQIRADQAELPKAQRYKIGDLLSRINLAKATYHDERKRIANPDDKYCEVKKVIRRICEQ
GKVRGRLTYGYRRVQARLDQLGIHLAGATIRHLMSELGVQVEIYNHHRNGKYSSYHG
RVGSVADNVLQQQFNAKDPYQVIHTDVTQVRLANQQWAYISAMTDEASKEVLAFQMS
SHPNQELIMNTLNELLDHLPDDAQPIIHSDQGWHYQLKYYTKKLKDHHFIQSMSRKGN
CHDNAPIESFFHLYKTELLQGLPPCNNITELRKLSLTYIDQFNNERISLKTKGMTPVGYR
NHALAV

>hypothetical protein
MQNEDQHDLAERISIYQGNLSRQALYAGQPFKKRKRTVVFFICDHDPYGLGWPSYALK
VELVGCPNFSVDNKETNVILNLKTDVSQLTQAQREMVAYFNDGIVSGAFSRKLAAAVA
AVKADAIRGGIYMTIEEYAALKSAYAQESERIKTITGLAALDMSRAQIIEFLKTNMDLSEE
EAQAAYDNAMAAHA
>hypothetical protein
MDLRKKLARKWRQATIKDHLVFSWVFNDHPELLEELLKLWVPWLKVRQNCQS
>Restriction endonuclease
MVYQRRSVEKAIVLALRYLGGSASRKQVKKTIADNEIDGLDYNQVYETKEGRKGTYIPF
DLDFGFGIRNIVSVGYLEETQRGKDIVLTPLGMDDNLEGFPRLDQKEMISAYWNKKNA
ERAKRKKQQEEQEQIKSQQKNNDEDSIQNPNIDSEDTEESNWKAQLLNQLKKFTPEKF
EQFSRLLISKMGVTLDKERGIVLSGDHGIDGFGYFTSDEFRTSRVAIQCKRFTTGAVSE
PDIDRFKGVMDSFNAEYGIFVTTSYFTSQAKKKATQGSNTVTLIDGQQLADLVEKYQLH
ITPITTYTLDEYYFENGE
>Lead, cadmium, zinc and mercury transporting ATPase (EC 3.6.3.3) (EC 3.6.3.5); Copper-translocating P-type ATPase (EC 3.6.3.4)
MSYFKRFVISLVLSVPLIIQMVGMAFSWMMPAYNWIALITTTIIMALSALPFWQSAWAAF
KNHHANMDTLVAIGTAVAYFYSIYAMATGRAVYFETAALVTVFVLLGQVFEERMRNNA
SSAVEKLLQLQAKEAEVWRDGHYQLVPVAQVQVGDLVRVKPGQKVPLDGLITEGVSS
VDESMVTGESMPVTKKVGDRVVGATLNKTGTFTFRVEKVGQDTMLAQIVELVKKAQT
SHAPIQRLTDKIADIFVPIVLIIALLTYLVWYVFLGASGVDAMLYAVAVVVIACPCALGLAT
PTAIMVGTGRAAKLGVLIKDATVLEAIDGVQTVVFDKTGTITEGQPVVTDVVGQKDLLR
VAASLEAASEHPLAAAILTAAKQVDLTLSPVTDFQVVAGQGVKAKLANKEALLGNEQLT
QAYELSPQLQATLTKLKQAAKTVVIVVYDGQVLGLIGIQDRPKADAKQAMAALVKRGLK
PVMLTGDNEVVAQAIADEVGIKEVIAGVLPDQKAATIEHLQANGKVAFVGDGINDAPAL
STATVGITMGSGTDIALEAGGIVLTNNQLMGVVNAFDISQKTFNRIKLNLFWALIYNVIGI
PLAAGLFSVLGLSLSPEFAGLAMALSSISVVGSSLLLNKVKIGA
>Copper-transporting ATPase
MQILVVIVVVATIAFIAWWFFGKHDQGQSQAQVKGNSQEARITVNGGYSPETVVLKQG
VPAKLVFNRLDKSNCLKEVVFKDFGISEELPIKQDKVIEIDTSKAGEYTYACGMDMFHG
KVVIK
>hypothetical protein
MSDEQLLTFIYNLILDPSLTEAERKLLVDFKNRLKPKSDLEDELRNLADDMRHLAVKNL
GHTSFSPKMQTIYKLISQSGLLKENLARGLAAVGPML
>hypothetical protein
MTTTGHKSKNKLVQLKLPVDVKNKLDDIFAKDGTTTPQALKMIATQIANRGFSPFTTLH
YEQYSEPVSAELEAKLREDELKIMGYLPNDTDEFADEQSLNQAFEKNLK
>hypothetical protein
MSKSDYATALRNLKSPNIKTRKRALKIIKQFKLQKGQK
>SSU ribosomal protein S14p (S29e) @ SSU ribosomal protein S14p (S29e), zinc-independent
MATKAKIAKAKKREELVAKYAELRAQLKAQGDYQALAKLPRDASPVRLHNRDALDGRA
RGYLREFGLSRINFRKLALAGKIPGVKKASW
>Mobile element protein
MDQIRADQAELPKAQRYKIGDLLSRINLAKATYHDERKRIANPDDKYCEVKKVIRRICEQ
GKVRGRLTYGYRRVQARLDQLGIHLAGATIRHLMSELGVQVEIYNHHRNGKYSSYHG

RVGSVADNVLQQQFNAKDPYQVIHTDVTQVRLANQQWAYISAMTDEASKEVLAFQMS
SHPNQELIMNTLNELLDHLPDDVQPIIHSDQGWHYQLKYYTKKLKDHHFIQSMSRKGN
CHDNAPIESFFHLYKTEVKQKG
>hypothetical protein
MTKYSPEFKAQVVVEYMTDGLSSYTLEKKYKVSARQIRKWIQQYQLQGIKSFKQHHRK
RIFSAKFKWHVINYYQTHAESFAQVAAKFDLLPSQISIWRQAFIKYGYDALKPHRKGRS
QKVKKQTKKQKQKLVEKKEIEQLREELARTKQELYNTKMDRDILKKSLALFGPSKPGKK
RK
>Tetracycline resistance protein TetM
MKIINIGVLAHVDAGKTTLTESLLYNSGAITELGSVDKGTTRTDNTLLERQRGITIQTGITS
FQWENTKVNIIDTPGHMDFLAEVYRSLSVLDGAILLISAKDGVQAQTRILFHALRKMGIP
TIFFINKIDQNGIDLSTVYQDIKEKLSAEIVIKQKVELYPNMCVTNFTEPEQWDTVIEGND
DLLEKYMSGKSLEALELEQEESIRFQNCSLFPVYHGSAKSNIGIDNLIEVITNKFYSSTH
RGPSELCGNVFKIEYTKKRQRLAYIRLYSGVLHLRDSVRVSEKEKIKVTEMYTSINGELC
KIDRAYSGEIVILQNEFLKLNSVLGDPKLLPQRKKIENPHPLLQTTVEPSKPEQREMLLD
ALLEISDSDPLLRYYVDSTTHEIILSFLGKVQMEVISALLQEKYHVEIELKEPTVIYMERPL
KNAEYTIHIEVPPNPFWASIGLSVSPLPLGSGIQYESSVSLGYLNQSFQNAVMEGIRYG
CEQGLYGWNVTDCKICFKYGLYYSPVSTPADFRMLAPIVLEQAFRKAGTELLEPYLSFK
IYAPQEYLSRAYNDAPKYCANIVDTQLKNNEVILSGEIPARCIQEYRNDLTFFTNGRSVC
LTELKGYQVITGEPVCQPRRPNSRIDKVRYMFNKIT
>FIG00628965: hypothetical protein
MIIKIERKEKKEKPKKSKKPSKQRKVPMIKVGTHKKLTFVLWVLLIGSVGFGIYKNFTAID
THTVRETEIIKQQIVDTNQVESFVKSFAKDYFSWQQSQEAIDKRNEKLTHYLTEELQVL
NEEMIRKDIPTSSSVNDIQVWQVSQVNENTFEVLFSVEQVITEDKDKETISSSFHVVVHI
DESDNMVIIKNPTMSKKPQKSDYQPKQLESDHTVDTETMDEIISFLETFFQLYPTATEK
ELTYYVSNHVLPMINKEYVFEELVNPIFTRKDNQVIVNVAVKYLDQETKATQISQFELILE
KQDNWKIVK
>lipoprotein, NLP/P60 family
MKLKITVICSVVFSLFFFLLLFLVIFFADDTDSGENNKDSSIPEGGVTVSPEVLAHRPLIEK
YGKEYGIEDYVSYILAIMQVESGGTAEDVMQSSESLGLPPNSLSTEESIKQGVKYFSEL
LTSAEQQGVDIDSVIQSYNYGGGFLNYVRSHGKKYTYELAEQFSKEKSGGQKADYPN
PIAIPVNGGWRYNYGNQFYVQLVSQYLTDTSPTEFDDETVQVIMDEALKYEGFPYVFG
GASPTTSFDCSGLIQWVYDKAGISLPRVAQDQYDATQEISMEEAQAGDLIFFHSTYNA
GTYVTHVAIYLEGNRFYHAGDPIGYGDLSSRYWQDHLIGARRVIHN
>membrane protein, putative
MTQKQKKFLHYSWITLLCVGVILLLLGTLGNAVQATGLVDETIDTSNEYSKYGLNQYQL
DYYVDNSWGWLPWNWSDGIGQSVMYGLYAITNFIWTISLYLSNATGYLVKEAYALDFI
SQTSQAIGENMQTIAGITPNGFSSSGFYVGFLLLFILIVGIYVAYTGLVKRETTKVIRAVLN
FVVVFILSASFIAYAPDYIAKINDFSKDVSTASLNIGTKIVLPNTDSQGKDSVDMIRDSLFS
VQVKQPWLLLQYGSTDVEALGEERVESLLSTSPDTNNGEDRENVVIEEIEDHDNTYLT
LPKTISRLGTVFFLFVFNIGISIFVFLLTGIMIFSQVLFIIFAMFLPISFLLSMLPSFDGMGRH
AITKLFNVIMTRAGITLIITVAFSISTMLYSLSASSPFFMIMFLQIVTFAGIYFKLGDLMSLFS
LQSNDSQNMGRQVFRRPRQMMNRQTRRFQRSMSKIFRGKAGSTKQKTMTTKQNQT
NKANHTRQNERTTPKKETALSNVKKQQGGIGQRLGEKTAHVLDTKNRLVNKRKQAKQ
QIQATPTHVKYALHQGKENVTHNVTDFKQSMADTKAQRKQQRMNQQEQKRKNIAQK

RLELDKAKEKRRKENPSSPVQNNQQHVRKRQQTVTPVLSQNTKRGNHQRPLKKEQK
QKTPPKQFNQRKLNKRTDKRGQQT
>FIG00627334: hypothetical protein
MSNFSVALQLLVIDELRKEDTPISNSIADHIESFVDYDFAHLLFSDGSVTQSISLEKQLNII
QVADLVLPDSNSSFEEYTTMELLSVAMLIVISTFALDFIHSDRSIFKIVDLDEAWSFLQVA
QGKALSMRLVRAGRAMNAGVYFVTQNADDLLDEKMKNNIGLKFAFRSTDLVEIKKALE
FFGVDKEDENNQKRLRELENGQCLVSDIYGRVGVMQFHPIFEELLHAFDTRPPVRNEV
EE
>Mobile element protein
MDQIRADQAELPKAQRYKIGDLLSRINLAKATYHDERKRIANPNDKYCEVKKVIRRICEQ
GKVRGRLTYGYRRVQARLDQLGIHLAGATIRHLMSELGVQVEIYNHHRNGKYSSYHG
RVGSVADNVLQQQFNAKDPYQVIHTDVTQVRLANQQWAYISAMTDEASKEVLAFQMS
SHPNQELIMNTLNELLDHLPDDVQPIIHSDQGWHYQLKYYTKKLKDHHFIQSMSRKGN
CHDNAPIESFFHLYKTELLQGLPPCNNITELRKLSLTYIDQFNNERISLKTKGMTPVGYR
NHALAV
>hypothetical protein
MTKYSPEFKAQVVVEYMTDGLSSYTLEKKYKVSARQIRKWIQQYQLQGIKSFKQHHRK
RIFSAKFKWHVINYYQTHAESFAQVAAKFDLLPSQISIWRQAFIKYGYDALKPHRKGRS
QKVKKQTKKQKQKLVEKKEIDQLREELARTKQELYNTKMDCDILKKSLALFGPSKPGK
KRK
>Integrase
MTKIIEYTAKGKTLYKFNFYAGKDELTGKKMQVRRAGFKTKKEAQLALSRLEYKKESGE
FEPLANRNITFETLALQFMDAKKASLRKSTYLVQRNMLNRILASLGNLKMKNIRGYHCQ
NAVNRWSKKAPATVSNIKILANQVFEYAIKNGYITKNPMDAVIMPKKPAKKRENNYYTK
QELNDFLTFVKQDRPSLYPLFRLLSYSGMRIGELLALTWDDIDLEASTVSITKTLSRVND
EGRTTIEDTKTKAGRRTISLDATTINILASYKQEQSKRLFSLGILNTLPKPLIVFDRGIGSS
YSKAQGERFVRVSDIRRYLNSMYKRMPKDFKHITLHGFRHTHATLLIESGANIKAVQQR
LGHANINMTLGIYAHATKVMENEAINGLSLYMEN
>hypothetical protein
MLKFNLKRIMDDDGITIQELHERTGISRNAISLLYNNKSKGIQLDTLDKITKALGLYSFEDI
FEEYSEYDIKDLSFSIEYDKDFLNYIEDALGILRVLEGDNILCELPILYGGDDNFSIFNINR
EALAFFESNSSFDFQRAENKIDTLFKENKTDLELFLIKVFKKIVEDKGINSKWFIVNSDIG
YLDKKQYTAFSFCWDRELLFSHEENLKAYIKAKYTTFL
>hypothetical protein
MTEYLSFKDACKVLGLNSYNSLRAYIKQGLPVIVVGKSKRISKKDLDEFMAKNRKVVK
>hypothetical protein
MLEVVLSEKQDEALRNYIYSTIVNEVDRIRQDQKLNMLLYSRKNLATACGVSTSTIDKW
QKMGLKVSYVGNTPYYTADALKAFMKEREQ
>hypothetical protein
MLDRHGLLDTLLDYERYRIIDTVILFETREPSEIEEILSYLNMNRPLPLAVRARYQKRYKD
ELKDYKTADFNDLTRAFNNTDPLQLTKEIKLAWHELQANPLNLLALHNLIEAVS
>hypothetical protein
MITKTNIKSDANFIIVANYEKAKTAKQLFKDLTARGYTVYMVIDKNFKLDDGFLIDNRHN
QAKQILKALTTQQDRDVLTQCISTFTIIEHLKNK
>hypothetical protein

MLYIQNGVTAPNMTNTGRSFKDFIQTYRPRNVSINNDNKGKLKASQQLAFSGRLVDDT
HKIESVLERDLIALDLDFIHSSIKTKDDLFKLLKEPLGGLNWYLYPTISNGLGNKLKYRLII
PTPRLRPGTYQTLAKGLNYYLLDRGILSDVDPSNWKVNQLFGMPVKNQFSEEPLIKAN
TQNKDLDKETTEKLIQELTKFISKHERATKWDASPTASTKQAINKGASQGAILLNQIVTGI
PQGQRNNTMYSFMLYWLKCGMDATALHDLASNINALYFEPPLDNKELENILASALKGH
IKIIGGVYNGKK
>prophage Lp3 protein 8, helicase
MAKNKVSSKLQKEIEQHNKVVLQAKSKGKQASSIDRYDPVQVFNLPSWFKAGVFLRG
TVNKQGDRPLRGISESSTKNITLILENDSNSPKPFYDEFSETFMYLDEQGKATLWDASE
SNALKLYIEETYDYTPTPTTVENGVKAYAKTHPRNLVKERVEAVKWDSIQRVATLFIDFL
GALNEDYTKEITKRWFTGMIKRVYEPGCKFELVPILYGDQGIGKSDLAHAIFPDNFVDN
LDTMKLKDDLLKLNGKMVCELGEMLAFNASSTDQIKSFISERNDLYRAPYSSVIKDHPR
KTVFLGTTNKATFLKDETGERRFYPIRCFGEQATKNPHNIDDAYMLQVLAEAKTLYEAG
YKCYIDIKEDAGLIKVASSKQEQAKEYSMVDDEIEQYLESTVPPDYENRYFSLDLNDKR
EIYQKSKKKQPKNWVKIPFVNTKDIWAVGLGNEKEMRISGRVSDSQKINLRMATEKKW
ASTQHFKGKYIRGYLRKEP
>hypothetical protein
MVDILQTMNFIYDSTNMLDTLIARFKDADKKVKELQNELEQREETAEIYEFEGLYSLHTL
RKGLISKGVKKRQFKNFGVTTDRVRNITFVVPLNFLSTMQEYDEQVNSYNGLLRRCKR
LVKDIKENIDTFEGIAPNDLYSSGIPVKDYIVMLNTVKHVKRIYYKPL
>hypothetical protein
MPKETTDKSKKQVEYNKRWYEKNREHKRYLSKRSTARSFLRVATKEDIQALIEFANEQ
LKERF
>GMP synthase [glutamine-hydrolyzing] (EC 6.3.5.2)
MANVDMQAFDKIIVLDFGSQYNQLITRRIREMGVYSELLSHKLTAQEIKQMNPKGIIFSG
GPNSVYDEDAFKVDPAIFELGLPILGICYGMQLMAHDLGGKVESADNKEYGKANIEVVD
KDAVMFKDLPTTQTVWMSHGDLVTEVPAGFTVSATSANCPIAAMANPAKKLYGLQFH
TEVRHTEYGNDILRHFAFDVCGAQANWSMEDFIDMQVAKIRQTVGDQKVLLGLSGGV
DSSVVGVLLHKAIGDQLTSIFVDHGLLRKGEAEQVMDSLGGKFGLNIIKVDAKDRFLGK
LAGVSDPEKKRKIIGNEFIQVFDDEATKLKGIKFLAQGTLYTDVIESGTDTAQTIKSHHNV
GGLPEDMQFELIEPLNTLFKDEVRELGEKLGMPHELVWRQPFPGPGLGIRVIGEITDEK
LQIVRDSDYILREEIAKHGLDKEIWQYFTVLPGIRSVGVMGDGRTYDYTIGIRAVTSIDG
MTADFAKIDWDVLQAISVRIVNEVDHVNRVVYDITSKPPATIEWE
>Pantothenate kinase (EC 2.7.1.33)
MENQLNYYKISRAEWKEFYKEHIVPLTKGELARIMSLNDRISQADVSEIYMPLVHLIGLR
WKIQQQKLKQQAEFLGKKKPQKVPLVIGIAGSVAVGKSTTARLLQTLLKEIYPHKKVQLI
TTDGFLYSTAELKRRHILSKKGFPESYDMDRLLRFMDDVKNNLPAKAPVYSHKIYDIVS
GEYELVKDVDILIVEGINVLQLPSNQQLYVSDYFDFSIYVDAKEELIKQWYLERFDMLLD
LAAQDKSNYYYAYAIGDRQAAHAMAKDVWQKVNHPNLLEYILPTRNRADLVLTKGKNH
LINELYLKKY
>contains lipase/acylhydrolase with GDSL-like motif domain
MNRIVLLGDSITAGMTNGYPSPIFSHKLASFLPEVEFINRGVPGDITRLAKLRLQADVLE
VDPDLVVVFFGTNDVADPQTDLEDYRSNLEEICRTVGLEKCLLITPGLVGPSQEASRPT
PKMAAFAKATLELGQDLDIPTLNWQTIMEKHIPASLLQADEIHYNERAYDLLVANLVPLL
KNKLAQVKPKIIH
>DNA helicase

MASIKEQEQARMDMVVTKVKEAIAKDQANLASVEADKKAIQADFSNNVRIKTSTYSGM
METALTVRQQQQLLQERENNMQQASRRLGILEKLEKKPYFARIDFEEEGEGKEEKIYIG
LGSFTDGPDKFLIYDWRAPISSIYYDGGGLGEVSYQTPDGERKVDLRLKRQFMVEDGK
IKVVFDTDETVGDQMLLEVLDENSDTKMKSIVTTIQREQNKIIRDTSSDLLFVQGAAGSG
KTSAVLQRVAFLLYRYRGKLTSSQVILFSPNQLFNDYIDQVLPELGEQNMVQLTYYQHV
KRRLPKLHVETLQERFEEDQAPLKQKVKKLKDSLDFFKASKAYGAYLERKGMRFKDLK
LGGEVFFSKAHIKQIYYSFNENYHLRNRLEATKERLVKELNGQINRYAKSDQVQKAIQD
LSQEQLNALYGDHPRNFANGEKELNFLGRQYVMHAFKPLHQAIIKNRFLSINGQYVHFL
KSVPALVDLAKYDLTKEEWQLELAQVVADLKAKKLALSDVTAYLFLYDQISGKRGQTDI
KQVFIDEIQDYTAYQLAYLKEEFPRAKFTMLGDLNQAIFTHENSHSLLKELGTMFPQDK
TKVVQLTKSYRSTQQITDFTKEILLHGEAVTAFERAGEKPRLTTAKTEADLIADVVARLK
QNQADRLTTAVIGKSLAECEALTQALRAEGQAVTLIKTENQRLAAGTIVVPAYLAKGLEF
DAVIMWQANAKNYHADNDRQLVYTICSRAMHRLDIFVSGQVTPLFDRIDPKLYELITK
>DedA protein
MDKFFDLLFNFHDYLFPLVQSAGAWSYLVFFALIFAETGLIIFPYLPGGSLIFFASSIAAV
NGQVLNIYALLASFFISAYLGDSLNYFFGTKLEQLSFFKKHIPEAKLLTARRFFKRHGGK
TIIWGRYLPLIRNFIPLIAGSIGFNFKNFALYNLLAVSLWVFGTGLCGYFFGRIPLVQEHF
STIFLLIACLALIPTIFTALLHLYKQRKAFKRS
>Phosphoglycerate mutase family
MSFTVYLVRHGQTILNRYNRMQGWCDSDLTTLGISQAQQTGKRLAAIKFDHAFHSDTI
RATQTCQNILAQNHNLVPAHQESFYFREEGYGYYEGADSGNAWMIIGGPHGCKTFSEI
ISTYSIEDAKNFAKAADPFDDAEDNAEFWHRLDCGFDLLRSVATDGQNVLLVSHGTTIR
SIVARFAPEIDITVSPKNGSVTKLIVDQAEVKVDYFNQFADDFAY
>Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205)
MQLAKNIKLNIPIISAGMDTVTESAMAIAMARQGGLGVIHKNMTIERQADEVRKVKRSES
GVIIDPFFLTPDNSVSEAEALMRKYRISGVPIVESEASRKLCGIITNRDLRFVTDASERIG
DVMTKDKLVTAPEGTSLKEAEAILQKHKIEKLPMVDEAGRLTGLITIKDIEKVVEFPNAAK
DEHGRLLVAAAVGVTSDTYERATALLNAGADALVIDTAHGHSAGVIRKIKEIRETFPDAT
LIAGNVATAEATRALYDVGVDVVKVGIGPGSICTTRVVAGVGVPQLTAIYDAAGVAREY
GKTIIADGGIKYSGDIVKAIAAGGFAVMLGSMLAGTDEAPGETEIYQGRRFKTYRGMGS
MAAMDSSHGSSDRYFQSGVNEANKLVPEGIEGRVAYKGQVADIVYQMVGGLRAGMG
YCGAGDLNKLREEAQFVQITGAGLIESHPHDVQITKEAPNYSK
>contains helicase conserved C-terminal domain
MEQIKASLAASVLGSLVKQDDYEVNDYLAPKLLTNTAEDNLWGRLKAELQSCRSFKFA
VAFITLEMLVPLKVVLADLAKTGVTGQLVTSTYLGFNSPAVFRELLKLPNVTVKIASQEG
FHAKGYLFDHGDYQTAIIGSANFTRAALLKNAEWNLAVSSLVAGGFTREVKAEFEQAF
GSGQVLDEEWLKNYQQTYQAPKVVSTPTEKSSRLKPNSMQAAALKELAALRAKGAKK
SLVISATGTGKTYLGAFDVKQAQVKRFLFVVHREQILTAALKSFRRVLGGPASDYGILG
GGQKQLTARYLFATVQTLSKPEVLAQFKQTEFDYILIDEAHHVGAKTYQAILNYFQPEFL
LGMTATPERNDDLSVFANFDYQVAYEIRLQDALQAGMLTPFHYVGLVDYELAQTGETN
VKGSLQQVAVSDRLDYLLEEINYYQPSSQPIRGLMFVSRQDEARQLAQLLTRRGYPSL
ALTNQDSVATRLKAVSQLEAGQLTYLVTVDIFNEGIDIPSVNQVVFLRNTESSIVFAQQL
GRGLRKAPGKEAVLVLDFIGNYQNNYLIPQALSGDRSASKDKAYQTLFSQTNLGLATIN
FSKIAQEQIFQALAKVKLDAWQKLKAAYQSLSARLNRPALLMDFQRLGDFDAQVFARL
KLPNYGHFLQRMGQAVDLSTYEDQVLTFVTKELLLGLRPQELTLLKELVAKRRLAKVDF
KKQLQVKGWLADEGTLAGVENILGLGFFKIKAGSGLKSDQYGRLPLVEIKAGSYRLNA

QLQASLDTNRWFKLLFSDVLAAGLTLNQAYDNHQTFTRFQRYSRKDVCRLLNLAKDVS
APMYGYKVFDRVCPVFIIHQKLDQTKRAALYQNNFSHNDRIRWYTRSPRHLDSPEVKQ
LLAKRADGSFQVQIELFVKPSDSQESGFYYLGPATIVEDSVKEESLVQAGRKPKQVVG
MDLKLTQALSFSELQRL
>Protein tyrosine/serine phosphatase
MQRIFPLANAANFRELGGYPTEDGRQVKWRKLIRCGSLSYLTAAEEKQILAYGLKYDID
FRSDDETEIYPDTIKGQVIYHNAKVYPFMNSVFKNLGIVASVKLGELNFMEEAYVQMLA
DPHAQVAYREFFTYLLANDKDGESLAFHCAAGKDRTGVAAFLILNALGVKKQAVMADY
LLTNLAYSNDSVEQVNRLLTNTPQNQLADKLNTHLAVIPEGIEIMYKTCEVLAGSVENYL
VTKLGLRRADLASLKEIYLEEVD
>Magnesium and cobalt efflux protein CorC
MIIVVIVFFFAAFFVAAEFALVGVRKSALESALEAGKGNKKKLELALKMTNNLNEYLSTT
QVGITLAGIILGWIGADALAEMLKSVFETLNLNHATAVALSATLGVILLTYLEVVLTEIVPK
NISIDMPLQMVLFVAKPLHYFHIVFYPFVWLLNVSATGIVKLLGMKPAGEEDDVLSQNEI
LSISRAAVSGGALDKEDLVYMQRAFDFNDKVAKDIMVDRTSLEVIDITATVADVINAYLK
EGYTRFPVVANNDKDKILGYVYIYDMIRQAQVDATVRVSKVLRTIITVPEVTPIQDLLQQ
MIKKQTPIMVILDEYGGTSGIVTDRDIYEELFGTVKDEIDDVSGEDIIKQADGSYKVSGKT
TLYDFERYFRYRNEDFQESDSVTIAGYLLENYKVKLGTVITIGDFEIEVIDFKRNYIEWMK
VTKKVAADSKPKVD
>Methionine ABC transporter substrate-binding protein
MKKSRLFGFLTAGLAAAALFVGGVKTASADTTLTVGASANPHAIILKHVQKQLKKEGVN
LKVKVYDDYILPNKALANGDIDANYFQHTPFLDNWNKKNHGTLISAGNVHLEPIAVYSK
KYKSLKDLPKGAKIYVSSNTADYGRILKIFKDAGLITLKKGTKLASANFDDIKTNKKHLKF
VHTFEAKLMPKLYESNEAAATVINANYAVQAGLNPSKDSIARESDSSDYANLIAIRKDEK
NNKAIKKLVKALQSKSTQKWIKKHFKGAVLPAK
>Methionine ABC transporter permease protein
MNSNTGLASYFKFQNVNWPNLGAAAWETIWMTVVAVVIVAIVGLLLGLLLFETKDSKSP
LVKALNWVVAALVNVFRSIPFIILIVLLLNFTQVIVGSITGPKAAIPSLVFSAAPFYARLVEI
AFREVDGGVLEAATAMGATRWQIIYKVLIPESMPALVSGITVTAISLIGYTAMAGAIGAG
GLGQLAYTDGFQSYNNAITLAATVTIVIIVMVFQFLGDRVVKRIDKR
>Methionine ABC transporter ATP-binding protein
MIELKNLTKTFQTADGPVQAVADVNLTVNKGEVFGIVGFSGAGKSTLVRMLNGLETPT
SGEVIINGERIDNLTGTALRQKRRKIGMIFQHFNLLWSRTVLENIMFPLELTGTPKKERV
AKAKHLAELVGLGERIHAYPSQLSGGQKQRVGIARALANDPDILISDEATSALDPQTTD
EVLELLEDINKKLALTIVIITHEMHVIRRLADSVAVMEAGKVIEEGPVAQVFTNPQQELTK
RFVNAELDTSQTPDVQEISKELLQKYPDGLLVKLNFHGNQVGQPVVNRAIRKFPEIELSI
LEGSIHQVADKTATIGSLYIQLLGPKDQVAACLEFFKEVRVETVVINDEF
>Flavodoxin
MSKSSKTLIVYFSFDDGANREVERLRRKIEADIIKLEPVKPYPSEYTKRLAMIEREFQQG
KLPQIKTTLPNLQKYREIYLGYPPVFGNRLPSIIHSFFSSFDLRGKTIIPFSEIVEEPLRATL
CEMNELAQKYEVKLVNREV
>PTS system, sucrose-specific IIB component (EC 2.7.1.69) / PTS system, sucrose-specific IIC component (EC 2.7.1.69) / PTS system, sucrose-specific IIA component (EC 2.7.1.69)
MAMDYKKVASQVIEAVGRDNMVAAAHCATRLRLVLKDDQKIDQAALDNNADVKGTFK
TNGQYQIIIGPGDVNFVYDELIKQTGLSEASTEDLKQIAAGNQKFNPVMAFIKLLSDIFVP

IIPALVAGGLLMALNNFLTSAGLFGPKALVETYPAIKGLSSMIQVMSAAPFIFMPVLVGIS
AAKRFGANQFLGAAVGMIMTTPGLIEGKYWNIFGLHVSQTNYYYQVIPVLAAVYILSVLE
KWLHKKLPDAVDFTFTPLISLMVTGFLTFVIVGPVMRGLSDAITNGIVWLYDTLGFVGTG
VFGLLYSPIVLTGLHQAFPAIETQLITAFTQNHTGYGDFIFVVASMANVAQGAACFAVFM
LTKNKKMKGLSSSATVSALLGITEPALFGVNLKLRFPFFCALIGSGIASAIAGLTHVIAVSL
GSAAFLGFLSINAKSIPLYVMCELISFAVAFALTFFYGKTHADLVNPEVAATAASAAQAS
ATSSQVEAASEMVADVADVDNENIAAPVDGQVKDLGDVNDQVFSTKLMGDGAAIVPS
DGTVYSPVDGQVTVAYETKHAYGLKSKDGAEVLIHIGIDTVNLKGEGFESFVKQGQTV
KVGDKLGTVDLAKVKAEGYDTTVMVVVTNTNNYASVNRLNVTDVKHGDNVVAVTTK
>Sucrose-6-phosphate hydrolase (EC 3.2.1.B3)
MEWTREQRYLPYNKWDAETLLKLQAQADNSSYQLNYHIRPSSGLLNDPNGFSYYNNR
WHVFYQSYPFGAVHGLKSWVHLESCDLVHWQNKGLALGADTKYDSHGAYSGSATVI
DDKLFLMYTGNVRDQNWQRISYQNGATMDKNGIVTKLAAPLIYAPAHVTEHFRDPQLL
KHRDAYYVLLGAQDKTTLSGKVAVFKSKDLKNWEDLGYLDFSPETMGYMIECPNLVFI
DDKPVLIFCPQGLDKNVANYDNIYPNMYVAGDKLKFDAARMETKQVAPINLDDGFDVY
ATQAFNAPDGKAYAISWVGLPDMTYPTDEEGWANCFSQVKELSLDEQNRLVQRPVPA
MAGLRQAGQVIKAEKAINGRQILVDNASRSYELKLTIPKGQAGTLHLVANQDLSESLQL
KFATGDAAQLVVDRAHSGIQFASDYGFTRQIDLTANEDLELDIFVDNSLFEIFVNGGLKV
LTGRFFGSQANNQIAIQGTTKEELAFRGTYWRIDHI
>Sucrose operon repressor ScrR, LacI family
MRPKLTDVAKKAGVSATTVSRVINNYGYLSQQTKEKVYAAMKELNYQPNSLARSLHG
KSTKLIGLIFPSVTNPFFAELVEEIENKFFNAGYKVILCNSADNKEKERAYIRMLIANQVD
GIIAGAHNLGIEEYNKVGLPIVSFDRKLSDQIPIVSCDNYQGGKLATEELYHAGCRHIYFL
GSARRKGHPTDYRLNGYLDTITDLGLTQHVHSVAFSESATIKSVSVRKLLEHHHPDGIV
CTDDLTAILVLQMAQKLGIKVPTDLKVVGFDGTKFIQDYHPELSTIAQPIPDIAALLVNLL
QERISQPNKKLAQNQYILPVKFIRSNSTFSL
>FIG00743624: hypothetical protein
MITRQEINGKAPMNYNWFDVNELNDSDTRQLKRLFRLTPEIIDYVSDRHERPHYDFDIH
TQSELVVYDVPLWPDNVTAHFTTQPIIFLLQGTTVFTFHTDETAYVFEQFKQEKVKEIKN
ATEFIMTFLLHAAKYFNQALDKLNVERNQLDKNLNNHIVNTHLRHLAQIEKGLVYLSSSI
KTNLMMLESLRHTPLTLQITKEAQEKLDDILIETEQASRMVQISNEVTEKISTTSNNILNN
NLNDTMKFLTVWSLILTIPTIMTGFFGMNVKLPGAQFPLAWIAIIILNIVFMAWLYFYLKKN
HFL
>Peptide methionine sulfoxide reductase MsrB (EC 1.8.4.12)
MTEFDKEAALKRLTPEQFAVTQQAATEYPFSGQYDDFYEEGIYVDVVSGEALFSSKDK
YNAGCGWPSFTKPLTKLTYKRDRSHNMERTEVKSPKADSHLGHVFTDGPTDLGGLRY
CINSAALRFIPVSQLKEAGYAEYLAEFQN
>Integral membrane protein
MRTRSELKAHVKQTYKGNWGKAIGINILPIIGAIFSVVGVGALVAITVVALKTGVVDPNHL
SDSNGGTSSGMVVNIIGTMIATGIAYSLLDWLRTGKQPASAVKAMFAVFSKQYFIPVLV
LFIIRTIFEFFWTLLFIIPGIIKSYSYSQTYFIYKDCLASGRNEDMNYLDYVTESRKLMKGH
KFELFVLQLSFLGWAILSVLTLGIGFIWLIPYINGVMAEFYNDLAGDKFTNRQFTKNY
>5S RNA

>Large Subunit Ribosomal RNA; lsuRNA; LSU rRNA

>Small Subunit Ribosomal RNA; ssuRNA; SSU rRNA

>Putative regulator of the mannose operon, ManO
MVESLNSKVEIVVDGTSYAGFAQYGKIMIGNAGFEFYDDKKVSNFVQIPWEEIESVTVS
VLFGGKWIPRFSIKTKSDGSFIFAAKEVKAVLRAVRSHIGKEKMYRSLSFFQVIKRGTSA
LFKKKN
>PTS system, mannose-specific IID component
MEKKVQLTKKDRMQVAWRSQFLQGSWNYERMQNGGWCYSMIPAIKKLYPNKEDRV
AALKRHLEFFNTHPYVASPVVGVTLALEEERANGAEIDDVAISGVKVGMMGPLAGVGD
PVFWFTVRPILGALGASLALGGSILGPILFFVLWNLIRMGFIWYTQEFGYRVGTKITDDL
SGGLLQKVTRGASMLGMFIIGALVERWVSIKFTPVVSRIQQQKGAYIDWDHLPKGAAGI
KEAITQYSNGLSLDKYKVTTLQDNLDQLIPGLAALLLTLLCMWLLKKKVSPIVIILALFVVG
VVGHVIGLL
>PTS system, mannose-specific IIC component
MSAISMILVVIVAFLAGMEGILDEFQFHQPLVACTLIGLVTGNLTAGVILGGSLQMIALGW
ANIGAAVAPDAALASVASAIILVQGGQGVKGVGTAIAIAIPLAVAGLFLTMIVRTLSVAIIHL
MDAQAEKGNMRGIDALQILAICMQGLRIAIPSALLLVIPSSAVRAFLESMPAWLTDGMAI
GGGMVVAVGYAMVINMMASREVWPFFVIGFVLAALSKLTLIAIGALGVSLALMYLELEE
KAKGNGGNGGNSGSGDPLGDILDDY
>PTS system, mannose-specific IIB component (EC 2.7.1.69) / PTS system, mannose-specific IIA component (EC 2.7.1.69)
MVGIILASHGGFADGIYQSGEMIFGQQENVAHVILKPDEGPDDIRAKMEKAIASFDDDE
EVLFLVDLWGGTPFNQANNLVEQHKDKWAIVAGMNLPMVIEAYASRFSMNTAHEIAAH
IIETAKEGVKVMPESLQPKEAPKQAAAGNQPTGAIPEGTVIGDGKIKYVFARIDTRLLHG
QVATGWTKAMNPNRIIVVSDNVAKDNLRKNMIAQAAPSGVKANTVPIKKMIEVAKDPRF
GNTKAMLLFETPQDALAAIKGGVDIKELNVGSMAYSNGKVNVNQVLAMDQKDIDTFKE
LKKLGVKFDVRKVPSSSPENMDSLLKKAQDLVNESK
>PTS system, mannose-specific IIB component (EC 2.7.1.69) / PTS system, mannose-specific IIA component
MTMDVRLMRIDSRLLHGQVATNWAKTAKIDRILVVSDSVAKDELRKVIIMQAAPPGTKA
NVIPLAKMIRIYHDPRFYGLRVMILVENPQDAKRLIQGGIRVKSLNIGSLSFEVGRKMVT
DTIAVNQADIDAFTWIHNQGILLETRKVSTDSKRDLWKTLRERNLVE
>NtrC family Transcriptional regulator, ATPase domain
MQTRRERILTYIKEHSSQVSQGFTTAELVVALGLVRSNVSKELNVLVREGLLVKSKGRP
VRYQLQGQGKSPKQVVTPKQAEKQVGGQDVAPSDDVFEKIIGRKGSLKTQVEQAKAA
LLYPPHGLNVLITGPTGSGKTYFANAMYHFACSNGIIKAKQPFTTFNCADYAHNPQLLM
SHLFGYVKGSFTGANEDKDGLIQEADGGMLFLDEVHRLPPEGQEMIFYFMDHGTYNR
LGETAKTHHADVRIVCATTEDPESSLLSTFVRRIPITIQMPSFKQRTMRERLGLLRHLLII
EANRIQKNITLNENVVQALLGSVTYGNVGQLKSNIQLVCAQGFLNNMQNQTEIQITFDQ
LPPNIKDGLTLLANNRKELGELTQLLEPTMTIYPDNDTKVIDSTGGYELPYNLYEIIGDKA
AVLKNEGLDQEAINNFILTDINVHLKTFYRGAKIAQAEDNLNEIVDQEIIDITKAIMKELTTN
YDYALSSNFLYAMSLHISSFIKRVQLGRPLRAVSSDIVEMVKAYPLEQKMASRVKELLE
EHYHFPVPESEVYYLAILLVSLKSSPASGQVGIIVAAHGKSTASSMVQVVEQLLSADNL
RAFDMSLEMSPQTALAGIAQEVRKLDRGNGVLLLVDMGSLSTFSERLTKETGVQVRSL
DMVTTAMVLEAARKTTLIDRDLDAIYKELREFNGYSRQIEPPKKELSEKAHHLVLGDKR
PKVIVAICSTGEGTAKRIKKMLDRILIDNLVDDILVEPISVVGMEEKIAQLQQKYRVIAATG

ITQPKLDVPFVSLENLLQGGGEQFIKVIENIDEPLFEKTASLEESDTLTEEMASQYLEQY
YTFLNPVKVVKVLWDYCDYLDRRLDFEINNAMRISLIMHLGGAIERNLLNSSLTITKQEE
AEAKADSKYKLVVQANGIIAEALNIELSLAEVYYLIKLIDTQAANL
>hypothetical protein
MSKAGRLDIHCTDAYGNQFDIKLQMTNEHNVAQRAKYYHALMTNHMLTAGENYQSLK
ETYVIFLCPFNPLITEQQRAVGHYQMISEEGDYLNDGTHTLILNASGDWEGVSDELKGI
FQLVLRQPASFGQLGAQIMDKIDEIKHSKTGGRKYMELTGAFWQDARNEGRNEGRKE
GTDEAIIKLIPDLFALSNDRQAVIKILMTKMNLSQEQAEGYYQRAMVAEAKKVSR
>Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA (TC 3.A.1.5.1)
MKFKKVIATGASLVALGLALTACGSNSSKSGLADKQVLNWSYASELPSMDLSTATDTIS
FDQLNSTMEGIYRIGKNSKVEPGLATKTEVSKDGLTYTFTLRKNDKWSNGDPVTAQDF
VYSWRRTVDPKTGSQYAYLFDGIANANDIIAGKKAVDTLGIKAEGKYKLVVTLEKKLPYF
KLLMGFPVFFPQNQNVVEKYGKKYGTAAKYLVYNGPFKMEGWSGSNLSWKLVKNKN
YWDKKDVKLSQINFSVNKSTTTSYNLYQSKKLDYTPLSTEQAKQLKGKDGYQVLKEAR
TNYLEFNETNKVFANKKIRQALSYAVNRQVLADKVLGAGTLPSLGIVSRDLAFNKGKDF
AMAAKTTAGVTYNKAKAQKLLKEGLAEVGQSKLSFTLLGDDTDVSKQVTESLQSQIQQ
TLPDVDVSVSNVPFKTRLQRSEDGDFDVVVSAWRADFADPISFLDLFTSDNSYNNGK
WKNAEYDKLITASKTTDAGNVDKRWDDMVKASKILSEDQGVAPLYQLNVAYMLNPSV
KGVIQNTAGVTWSFKDAYIANK
>Ubiquinone biosynthesis monooxygenase UbiB
MTRPDSSNNLTRGSRRTRLLEIVRVMRNHNVITNFINQRNPREVRLAFQELGPTFIKAG
QLLSTRPDLISPAFIAEMRQLQDNVEVDDFASVKATFEEQTGKKLSEVFSFFDETPFAS
ASIGQTHRAILKDGTKVVVKIQHPEVARLIATDLALFRLALRMTKFTPDIGAINPREIFNEI
RTSLLNEINTEIEIENGQEFYHYNNNDGIIRVPKVYKQYSAQKVLVNSSMPGKSIKNYLA
QPISKDLAVAESQKAERKYLAQVLVNNFLKQVFEDNFFHADPHPGNILFYRLKEGDPN
YQENQAKEAFSYEFHGNKVVWAKRQPLPPYRLVYLDFGMMGRLTPSMIDGIAQIVLAL
NTKDIRQIGQAILAVCNQTGPVDSEDFYEELGLFLTPFMNMGLDQIDFPAMLYSVIGLC
RKNNLQMKAEVTLLVKAFGSLEGLVSQLDPDLSMMDVARPLGKAYLKRKFNLKTSLED
LSFDTLQSLKATSQLPTKASKFFDVVSSGQTRFSVRYKGQDKLLDRIDHLANRIIIALVLA
AIILASSLLVEGSANHPAIYNLGVTGYIVAIVLVALLILDDLHKRFKKRKRK
>Accessory gene regulator protein A
MLKVIVVDDQTEILEQVATGIEAYVKFENLDVNLALATGDPQAVLTYVRENPNDKYLFFL
DIDLATESYTGIDLASELRKLLPYEAIVFLTSHEELSLIVLEHRVLPLDYIVKQSDFTKVMK
AVKIDLQRTLELIGQESKVNKSYFTYVKRGRHYRIATTDIYYFEAMPGSPHLISLYANNK
RLNIRRSLSQIEEQLGADFYRCHRGYLINLARVVALDPVEKMVYFDEEQTIACPVAVRR
LRDLNRKLT
>Histidine kinase of the competence regulon ComD
MLAGLAGYIAVFGAYFYLLGELRLNKVLKIVASSLVSVLLFLSAFVDLGFIVLVLGQRWY
YRRVNKAHALLRRLALIVVIILFDAIGDVTTFTFFLIGKIANLSQLEVAIIGVPAQLLAFLVVI
WQLKKRQLFKTLAVHTDTAFKLALRLFIFYIFVHSIEVVLITLTDKWMLVIVSAIIAIFTIALL
VIVFYFYQDTVTKLENTALVAEQARQMKAFQLYLQEINDYYQEIANFRHDYKNLLLTLG
NRIAKSGDADLQSYYQQMVTYSDQQIRGVIKPRNLAQLEQIELQSLRSVILAKLTKANQ
LGIETQLNVTGWVQAGRVNELELARVVAILLDNAIEASSKQEQKWLGIGLDSYGEDGV
DILIENKFTGDKAPNLTAWFERGYTTKGSNHGRGLAIVNQLCDKNLIDLEVEYKRGLICF
ILGIEV

>ABC transporter, ATP-binding protein
MLTLDNVTKNFGTKQAVSDLSFTVKPGEILGLIGQNGAGKSTTFRMIMDFIQPTKGQIL
WQGKQFDRTKHQIGFMPEERGLYQKESIIDQLNYFADLHGRNRFDMRLELVKWMERL
NVVGKLTDKVESLSKGNAQKVQLIATMIFHPEILILDEPFSGLDPVNAQLLLDEIIRAKNE
GTAIIFSTHNMDNVAKVSDKLLMLKHGQTVLAGTLAEIKQTYPRLNLTVEGIEAEGFFSK
YHGVTELRKVGQQTFKMKLANEDLGRLIFQDVVKNRDYVPVFEQSAPSLDEIFKQLA
>hypothetical protein
MSNYQLMAIAKRIVSKQLKSFSYLSLVLLPMIVGEAAVYRNQEFLQTLTAKQLSLGLYQL
MPGVAAFLCAVYTGILATEVVADREAKLTEFLLAIVDAKTQLVGKILGTVILLVLHCLSYF
LVLLATVVIFKLRLAMVDVKYLVFLLLSLLLLLGSSLIWTVEFGIYIQEREQMALAMLTPVI
LVMTGEILATVYACTPHMFAGMVWFKVFLVVNGWVPALGTCLLPVAVSTQGLSYFWG
YFSLVVQVIILVIMFKKVLPKYQRGLLATKQRHPIIKAIFGK
>hypothetical protein
MRMKKYFKVARFEINHLMKRALFWTTFVWPLSMEGALSLLLPAKLNLSPEQLLTLVVIT
LLGFTAINALSIVNMFINSKSTRVGELVIGMSGYKVQYYGKLLAASLVYLLNLAASSGFLI
YLGQMEGFLQLNTWSKFELLNAGVVMLVVTIFGNILLSILAVYVNDEDKGQVIALGYLLF
VLVACGGSVVGLLGHSSVWLLIPIVNLLGQLSLTVWWEQLVIWGVNLALLVGGLQVGF
NYYQHNILNYARKVGGVSE
>hypothetical protein
MTFLLAEKILLIAGLMAIIFIGSELLTRVALLVKRLVQKRVRKQTEEELLYKAKVQASVRK
WLVTRKSYSWLLTGFYLVFLVIISWAFWYFSTLPFWGYLILVLFSLQQPLYLIPATISLQK
LPQYSPGLKKIEQQPLILRLTDYVCLCNLLPLGLLMLTATVYLDLNAISILAICEIIIGLIFSLL
LFRRVRTFWHYYRTVGAEYLVKYRHYLGANLVTWTLFYMVLGLIVYYLNV
>hypothetical protein
MAYISNSNLFWASLVLSIVILLGDVLVGRILIGIQRFSMKILPRSEEESIKQLAVGYANKKK
YSWRWLIVTTVAMTLLGLILCQTAAPPAWVYVFLLFFIIQTNFHLTPAKVAFHKLPNYVKI
IWNVRKQPFVFGYYNYCFLCNGVIIGATLIYATVQATLQVHPFNLFELVVYLAIGLAELFL
GLYFSLSFLHRLRLLVNALNKENPKYVIYWRRYLVARVINWGLLYIMLLIFLTYVPL
>hypothetical protein
MRLEQVVEIYLLSKMFLKFIGITLIGYFFFCVIVRVLMNNAKKTPVEISILDDSNEELAKTL
NGDKGFLRAKITAWLIKFVKNIQLYNSQLVSKKEMVTQNRLRKLYLLDITGHLIVGFSFS
SGLIELQNQLLAGIFFPLLYVVVGSLIDYKVIKIRAKLGSYLIDKGSFFLVPGLPLLLMINLL
WVGAALLISNLQAETLMIAGFDVGITDMFLGIATILLGLIGVLLQVKRLKAILKRLSLLGKQ
AEMSFKKSILLKASDYMIVYVILGIILFM
>ABC transporter, permease protein
MVGTLTYLAVTLIGNVMEILLRKWEFLKWNPLNFTNYGNQLVDPTFANITHLTTNQLLW
GSLAYTTVFLALGMWVFANKEV
>hypothetical protein
MAALLFFIKAAIRNSGYCEPVYYGANGYSCRYSNGKWDYKVTKTAAQAVGGVMAKG
WIGSLGGSYYQPGYRG
>hypothetical protein
MILNAHVTQDERAILVNFKNHVNDKNFANELMGLAAELRQLGIKNLAAQKKMSPETSAF
YKKIATYGQRELNWARGLAMTGVIF
>small heat shock protein

MANELNNRFNELMKGGDDFFSNLGRSFFNGFDDSFKEMKTDISESDTAYTVAVDLPG
VDKKDIDIDFKDHALTISAKRDSFSDESDAKGNLLSSERSYGRFTRQYQFPNVDKDKIS
AKYEAGVLTVTLPKTAEEISNSHKIEIE
>hypothetical protein
MYKDKRIIQRPYFTSRSWAALIFAALLLIRITSISLTALYHTQADFVYQLTFQSLWYLSSL
GLTLGLSFLFVYLFDIQLRFASLFRLYFVSRSPYLGACLLELLALLIKKQALTAATNLSIGA
GGIVLSCCYLAWLNRREAYLTTNQNFKLVIATGTFQLFSLLPEFLQVFS
>Major facilitator:Oxalate:Formate Antiporter
MKKTNRYVVATAGVVLHLMIGSVYAWSVYSNPIVAETGWSESAVSFAFSIAILFLGMSA
AFMGRLVEKLGPRVTGTISAICYGTGTILTGLAIHSHQLWLLYLAYGVIGGLGLGSGYVT
PVSTIIKWFPDKRGLATGLAIMGFGFAAMLTSPIAQRLIAGVGISNTFYILGLVYFTVMIIS
AQFIKKPQAGELPAAAVAKTNATSLTGGQQLTANQALKTKSFYFLWFMFFINITCGIGLV
SAASPMAQSMTAMTAGAAAVMVGVIGLFNGFGRLIWATLSDFIGRPLTFSLIFLVDIMM
LAAMLFFHVPFVFVVALCLLLSCYGAGFSVIPAYLSDVFGTKELGAIHGYILTAWGVAG
MVGPVLLSFTHELLGTYTVTLITFILIDLIAFIISFVIQKDFVKVSQGATEK
>Negative regulator of phenolic acid metabolism PadR
MPQKRILPYVILGILEKGDATGKQITDEFKYEIGEFWKSSHSQIYPELKHMLTDGWIDSY
QKEDNDKEIYYRPTKLGKEIIADWLATPVAKLPVSKDLFSLKVFFIKRADDPRLKTLIANQ
KKLVSQNLAHLKKREAELFDTAEKINHNYGHYLILQRAISRQENQLAWLDQLTVC
>unknown
MAPREIHFTFGPKEALKKLIQAHPDRKLLLFQAVTDKERYMLFDYSGQETIFSGGLSYQ
VVRQVEFDKDWDGFFEFRYLTLDEDEQKVFRAIMDKWVRKDGRPFGLNETVILQSEK
KNFEFLMINVWEAEADFVDWTNLKDNELQQFGNAGNDQALVVEYKRAK
>Fumarate reductase, flavoprotein subunit precursor (EC 1.3.99.1)
MKFVGIVGSIAEQSYNKMLLAYIAKHYQELAEIEILDIKDVPIFNESDDQTETPVIQNLVKK
IKAADGVILATPEHNHTTTAAMKNVLEWLSFKVHPFQDKPVLIVGASYFSQGSSRAQLS
LRQILDSPGVNALVMPGNEFLLGNVKEAFDAQGDLKDQGTVDFLGSVLAKFCEWVQV
LEVLSEKNKQTSWQDEDLSASKGTDTTIKGVDMRADDWVEQAAAKTKAASGKDYVKL
DRGVLTVDQLNWFLNTMPMELTYADDNNQFIYYNHTLPGKEMLAPRDPKQVGDPMD
AVHPPRAIKGVKHVISSLRNGNDLVEMPVPGNKLNERHIMHYYKAMRDSEGNYRGVN
EWVLDLWPIVESYLKQTGQKLVKDEAAPDVVSGASGNENQAPAVDSTTSASVDESSS
EDLGPRLDEGIDSVTGASE
>Fumarate reductase, flavoprotein subunit precursor (EC 1.3.99.1)
MNKYIGIAGTNSPRSTNKKLMEYIQSHFAGQAEIELMSIADLPVFYKVPGRVLPPRVEE
MDKLVDEADGVIIATPEYDHSVPAVLMNALEWMSYGGTHPFVGKPVMIVGASYGTLGA
SRAQAHLRQILDSPEMQARIMPSSEFLVDHSLQAFDEAGNLTDSEKVERLEQLFTDFE
TFVEISKHLNASLAANKQLAANYTWEDEEGAK
>COG1983: Putative stress-responsive transcriptional regulator
MKKKLYKSQDKVVSGVFGGIAEYFDWDKTWVRLAGIFLIIFPGSVIPGIFMYFIAAIIMPE
RPAANHDDETVIEGEFREK
>hypothetical protein
MKKMLKAGVLITVVGIVLLILSIMGGGVQSVEYVGWTPRIVKKEMTLKKQTVSEFENVK
VNTDNVGVKIIQGNGYHVDYKGRKSLAPKISVEGKTLVIKQKAHLRTGFIFDGVTIKSDR
HTSFTNTVTVTVPKDLDQISLTTSGNSDVTLSDLKLNSLALDVLDGDLHLTNTQVANSS
KVKLLDGDLYLNRVAFNNGTKMEIFDGDLSLVDTSLNNVQIENHDGDVSFNHLKVAGG

SYKGQDSDVNGTSLEVTAGYQFNLADGDVSITGAKADGYETTTTDGDNHLFNQSGSH
LTQGANSGNILKVTTTDGDLNIR
>integral membrane protein (putative)
MDKAHYFVELEGFLTALTAEERTDVIEFYSEYVDEAGLKSRAEIEGKLGTPRQLARKILA
DYSIKDFDDETGGQAEGQVAKTHSNLRMIWLILLVLITSPITIAILGVILAIFLAVVITVGAL
AFGFLAAFILTFAVALYTGIRLLFTSLMVGLFYIAVALLSLGLLGIIFPIVFWFFSQIIQGIAN
LARFLYGKFAKKKGVQVK
>Transcriptional regulator, PadR family
MKIQITADLLDGLVLALLDCKDYYGYSLTQDMQQAISISESTMYPVLRRLKKNEFLTTYD
KPYQGRNRRYYQITDAGREHLNRVRKMWVDYKGSLDEIFFGKKEEEISG
>FIG01280259: hypothetical protein
MKLTKTDHIALIVSDYYAAKEFYVDKLGFELLSQHDRPERGDIILNVGQGDLVLELFVKP
TAPKRPQLPQPENQGLRHLAFKVSDVAACLAEFDRLKIKHEPLRYDDYNGKKMAFFFD
PDGQPLEIHE
>Transcriptional regulator, MerR family
MTMFKLHVNKDFRVGIGDLSQMIGVSTRQLRYWESKGYIKSLAGEEGSTRKFSIGTCY
RAATIRAFLDEGYTLAKAVAKTEQVKNQVAIVRHFENEMLKDIKVTDEDKGYGEIDYGY
LANEPTKKVYGVIDENGYRMETREVNADETN
>Permeases of the major facilitator superfamily
MTKYKQKMKKLKKEGICLQSEKQALDINGKPYNRGLLVLVLMIGSFCTVLNGTILATALP
AIMKTFDINTSTVQWLTTGFMLINGVMIPVSAWMINKFSTRVMYIGAMTTFLLGTVLAYI
STNFGMLLAGRIIQGLGVGVTMPLLQTIMLSIFPPEKRGAAMGTVGLVIGLAPAIGPTLS
GWVVDNYTWRDLFGMIIPIVCFVLILSFFLMRNVIPNQDQKLDFISAVTSTIGFGSLLYGF
SNAGNDGWTDKFVLLELVIGFIFVALFIRRQLHMEEPFLDVRVFKYGEFTLAALLSGVA
NLAMFGIEMLLPLYIQNVRGESAFHSGLMLLPGAIAMGVMSPITGRIFDRFGARYMAIT
GMTLLTLGTIPFLFITENTSVAYIVIFYAVRMFGIAMTMMPVTTSGMNALPFNSISHGTAV
NNTFRQVMSSIGTAILTSVLTNVTTNSLPAKSMLKQTPLKYRSEAIDAVLVGYHSAFFVA
ACFALVALILTFFLKKGNNARQVNLNRKGATK
>hypothetical protein
MIILILIAFAVLFFMFNVLPITKGNHLWSLLFLAGFILSGAAIIANDVNHFGMQVQSQTKTY
HLASDSDKMGVLLYKALGNGQEKIYVYKTSAKQAKPKPTKTSDFSAKVTRTKEDSYLK
VTTKRYVYTNKWSKLLFGVFGNNHEVKHRYYHFYVNKDWYVLSTDQAARLSKLLKTK
QTEMKTKIASLVKAKVTQEIMANPSLAADSSKQQALTEAATKVATKQVLDQYVAQVK
>hypothetical protein
MTNTSATDIVPAQIIGLGTNSITLRTENNEIVKVKQPATAKKDKVFMQTMKDILKSGLWIP
VNKKLNQLLHYDWFDAPIESYSF
>Organic hydroperoxide resistance transcriptional regulator
MIEHTPGRLEDELCFAIYSAQKNYNKFYSEALRKFKLTYSQYITLLVLWEEKKPLMIKEL
GKRLGLDTGTLTPLLKRMEKDDWVKREKTNEDGRRVYISLTAKALENEQPVKQTVQS
CFYNMDMTKEEYVEDVARLKEITQTLEKNNNRLEKENF
>RNA-binding protein Jag
MTKFEGKTVQQAITLGLTTLALTEKQADITVLRQAKKGFLGLGKKLALVEVSAKKVPVA
KKAPVSEPVKTAPVTESDEAKREKRQAELEKALLAVGDYLAAVTQKMGIKATIDLTPSR
HTAYYKFETNQEGLLIGKRGRTLNALQLLAQDYLDKQVHQHVRVMLDVSDYRKQRQE
TLTHLAQNVAREAVLARSVMSLDPMPAYERKIIHAALAQDQRVQTYSKGSEPKRYVVIE
PV

>hypothetical protein
MQNLTLSKFIHLHCISKVLETQKISLKAEGEKANNIVGD
>oxidoreductase of aldo/keto reductase family, subgroup 1
MSLVDTYELSNGVKIPVVGFGTWQTPSGAVANFATKAALTSGYRHIDTAEMYGNEASI
GQAIKEAGISRKSLFITSKLNNHAHSYDKASAAIEQSLKDLQTDYLDLFLIHWPNPKPSR
DENWQAHLQDTWRALEDAYKAGKLRAIGVSNFKPHHFEILAQTQTVAPMVNQIRICPG
DHDEETIKYCRDHNILLEAYSPLGTGKIFSEPTIQKIAQELNRSVAQVCIRWSLQHGFLP
LPKSVHEDRIEDNTHVFDFELSQAQMDEIDALDGVVGYATDPDTASF
>Neutral endopeptidase O (EC 3.4.24.-)
MTQYPVNTAQIKDDLYLAVNGAWLKTAKIPADKPATGGFFDLADNIEKTLMADFEDFLN
GKTHTDDPYLREFLKYYRLASDFDQREKLGFKPAKPFLKRIEDLKNLADWEEKLPSLTL
AGYPSPFPLSVEADMKNTKAYALYADVASLILPDKTYYTKDNPQAPALLAIFQKMMTKL
FQLAGYDDDFASQTIAQALEFDRSLAPHQKDSTERADYTKQYNPYSLADFASFSSHLD
LAAYAKTLVKTDDLDKVIVTEPNYYQALDEILNEDTFPQFKSWLLVRSLRAVAHYLTNEA
RIVAGQYQRALSGSKEAMKAQKAAYYLAADQFDQVVGLYYAHKYFGPEAKADVHRMV
EQMISVYQKRLKANTWLSQATKDKAVLKLSTLGINVGYPDELHYLYKKFVVDENQDLLT
NDLNFTKITLSDHYGRLKKEVDRTRWAMPAHMVNAYYNPSFNVIVFPAAILQAPFYSLK
QKASENFGGIGAVIAHEISHAFDNNGAKFDEFGNLKNWWTDSDLKAFEDLAQAMISQF
DGLETEAGKVNGKLVVSENIADAGGLSCALEAAKEDPQADISAFFINWARIWGVKASLE
RQKLLLAIDVHAPHVLRANMQPKNLADFYTTFNIQPGDKMWLAPADRVNIW
>transposase
MNNSIKTILGIKDPYLKLDEKNFDNPIEDQPNQIIVHLIQTYPMHCPKCGHLMCKNGYKT
VNCLGPELHFKPTIWSIKKQKYICKASSSCPEVVTKLAAVEDIHYRNHISLAIKQRAMML
LTKNESQSDLAKELNVSDWTIRRVITNLDQFFKPNYHWLPRHIAFDDFKSGRFAPSGM
SMILMNIENKRTLDIILSRKNSYLRKYFLRYDRSARLAVQTVTVDLYTPYRHLIHELFPHA
LIIADHFHIVAQAYRAFNKIRIQVMNRAGAGTHKWRALKHFWKLLLTPANELKYDNYWS
RRNFSYAQLTDVEVIHRLLSFDNELKRAYEYYQNLILVIAHRSKKELKNLLAIKWTQLPQ
ALQKVQRTLRRHKQEIYNSFKYDTYTNGPVEGTNNKIKVIKRTAYGFRNFFNFRIRILLA
LPNTYIAINWRNKQTAHAKVQAQAA
>Asparagine synthetase [glutamine-hydrolyzing] (EC 6.3.5.4)
MCGFVGCLTTIKNTKDLNDQLTVSDMNSMIVHRGPDDSGYFEDEHITMGFRRLSIIDLE
GGHQPLSYDNDRYWMTFNGEIYNYLELRADLKKEGYEFKTDSDSEVIIAMYARYKEKC
VDFLRGMYAFVIWDRDTKTMFAARDHFGIKPLYYAEEGDKFYYASEGKAIYKVLKDKS
FDKAALQDYMTFQFVPEPKTLTKEIKAVLPGHYIIKKPGQAVVTKRYFKPEFNVIARPEE
EYVKRVRNVLLDSVEKHMRADVTVGSFLSGGIDSSIVVALARQLSPKLKTFSVGFERE
GYNEIDVAKETAAKLDVENFSTVITPQRFVEAFPCFIWNMDDPLADPAAVPQYFVAKEA
RKQVKVALTGEGADELFGGYTIYHEPYSLKMFNYTKPLNGALHALARIIPEGVKGKSFL
LRGTTPLEERYVGNAFIFNEKEKQQFMTQYDANQPFVDVTKPYYDIAKNYDPVTKMQF
VDINQWLNGDLLHNADRTTMAHSLELRTPFLDKEVFAVAREIPADLRIANNTTKYILRKA
VEGIVPDHVLHRRKLGFPVPLRFWLKDELYDWAKDQILASQTDKYLDKTYMLKLLDDH
RDGKADNSRKLWTVLAFMMWHKIYVENDYLPDAEEIQATAAKI
>sensor histidine kinase
MSLQRRIVTSYVIIIVLALLTLLGIGGAFMVAYYAKYDEYATDVRYFQIEKNGYKVKQLVT
NSQYSKANQAQIIRDLHAYHYRVRIFNQKTGKVIYTDKFAEDELYFVKSIRADDFSTKKA
KRYQWSNSAFVSRLYNYQGQKIGLVATKLTRYRPKKYNTHISKLDLRNRIIFTLWAGVLI
IVIVALTVLFFSKRTLRRIMAPLDKLIAASRRIEEGQEFKPVVYTGGDYEMEVLIKSFNDM

QTSIKEYMDKTAAYERARTKMISGISHDLNTPLTSIKGCLKGIIDGVARTPERQKKYIEIA
YNKAAEMEGLLQQLFYLSKLETGNMPANPIVLDLRTFVSDLVAKNRLAWEEEGAQLSL
RLGTETSQVKIDAEQLRRVLVNIIENAIKYNDKNTALVTIDLITSADEVKLCLGDNGPGVP
EDKLGEVFNEFYRGDESRNTKKKGSGLGLYIARTIINQAGGTISAKNEQGLKITLTLPRI
N
>DNA-binding response regulator
MKKILIIEDEEAIAMIEQDYLELSNFETEVVGDGASALQKLNDNSYDLILLDLMLPGGMS
GYDVCKKIRGQVDIPIIMVTAKTESLDKVLGLELGADDYLTKPFDPAELVARVKANLNQY
ERLKGQTAQAKEKLVIGNGIVMELGSYKVTKDGKELRFPNKEFELLRFLAQNPNYVYS
KEQLFEKIWGYDYIGDPATVAVHINRIREKIEQDSNNPKLIETVWGVGYRFNGGK
>FIG00752497: hypothetical protein
MEATTKTSAFANAGQDRTTVKEYAYNISAAVANAILVTLGMGLLMQTIAGFVHFQPLYQ
AGTLAQDLLAPALGLAVASQLRVTTLVMFSSMIAATVGANAVHFVTAANGVATVTATG
QTAVQAAGTGVFSTGQPISAVAAGLVAALVGKYLSGKTPLDMVLVPFGATMVGTLVGL
GLAAVVTPALLSVSAFIAESMKVNPFVGSFCISMVWALFLMTPASSAALAVALTLDPVS
AAAAAIGTTAQFVGFTAMSYKQNTIGANVAQFVVTPKLQFPNLIVNPWQLLPPLVSAMV
CAPLATVVFGFQAPYTVGGLGLNSFIAPIYFWGQGLGTFGVYVLCGLVLPALFSVVGFK
LMKKVGLVRDNELHLNIL
>Cyanate MFS transporter
MKTFKNIALVIIVAFNLRWGITSVSPVVDYLKLKLGLSNFQVSLLTSIPVFCMGIFAFFVGL
LQTKYGRKWSIFYLLLLMFIAAATRLGSHSFISLCLSAFIIGLTIAIIGPMLSGFIKEQFPHH
QGSMVGVYSLSMGVGATVASGAALSLSHRYSWQVGLGIWSVIALVSALVWQLGAAD
DKKAPKLPSQSSSPLKVSELLQTKQMWALILYFGFQSGIYYSAVAWTTSFFSSRNLPAA
KNTLYLTVMMVLITACSFIIPSLMDRFGSTKNWNIFCTVAALFSGLTFWLGSNDWAYLL
AILLFAIGSGGLFPISLLLPLQFSTQADEASLKTGLMQAGGYIMAGIIPLALGSLMDWTGN
YLYLPLNYVLSSVIMLVIGNRLLSEK
>Multidrug transporter
MTSKQIKWVTAGLLLASVIAGLDATIINTALPAIISDLNGIEYMGWIIAVFLLGMAVFTPLW
SKLGENWGNKKAFQLSVVLFLAGSLLEGLAPNIGFFIAARTLMGIGAGGMGALPYIIIGY
VFENINKRAHILGLISASFSAGSIIGPLLGGWIVDTLSWNWIFYINVPLGLVTISLIQVAYKE
EVRQTKQAFDYRGAWLMVLGLVAFLLGIQLLGMVSWYYCAGLVLVGSLLIWRLSKVEA
KQQTPLIPGHLFKNKALVVDFILFGIAWGASIAFNTYVPMWAQGLLAATALVGGLTQVP
GALTDFLGANLAPAFQRKFRDLTVVNLGLIGIMVSVLGLYFLDAKAAYGWLLFLSAFYG
FGVGLVFVVLQIRVQQDVDIKDMPAATSLSYLLRILAQTLMAAVYGVLLNFELARGVAQ
TKGKISLEMMNKLSNPVTNKELPLNLIPQMQLILHQGLHTIMLTATGLLLVGLVFSLWAQ
KKVFKN
>SSU ribosomal protein S18p @ SSU ribosomal protein S18p, zinc-independent
MAQQRRGGRRRRKVDFIAANHIEYIDYKDTDLLKRFISERGKILPRRVTGTSAKNQRRL
TVAIKRARIMGLLPFVAED
>Single-stranded DNA-binding protein
MINNVVLVGRLTRDPDLRFTPSGVATANFTLAVDRNFTNQQGQRDADFINCVVWRKA
AENFANYTHKGSLVGIEGRIQTRNYENQQGQRVYVTEVVVSNFSFLESKAQADAYRQ
QNGGGQGYAPQGQGYQANPYGTPAANQAQPGGFTNNNTSEAQPSADPFADNGDTI
DIADNDLPF
>SSU ribosomal protein S6p

MTNYEITYIINPSLDEAAKNDLVERFDNVLKNDGAEVVDSKDWSKRRFAYEIAGYTEGV
YHIVNIKAENADAINEFDRLAKINDGILRHMIVKREA
>DNA gyrase subunit A (EC 5.99.1.3)
MVDLPNRIKNVDLSKVMRTSFLDYAMSVIVARALPDVRDGLKPVHRRILYGMNELGVT
PDKPYKKSARIVGEVMGKFHPHGDSAIYESMVRMAQDFSYRYMLVDGHGNFGSVDG
DGAAAMRYTEARMSKITLEMLRDINKDTIDFQDNYDGTEREPKVLPARFPNLLVNGAT
GIAVGMTTNIPPHNLGEVISGLHILMKNPDATTADLMEAIPGPDFPTGGIVMGKSGIRKA
YETGKGQVTLRAKVDIEELKNGRQQIIVHELPYMVNKARLIERIAELAREKKIEGITDIKD
ESDREGMRITIDIRRDVSASVVLNNLYKLTLMQSSFGFNMLAIVDGTPKVLSLKQILQHY
LNHQESVIRRRTEFELKKAKARAHILEGLRIALDHIQAIIDIIRKSPSGDVAKDRLINEYALS
DKQAQAILDMRLVRLTGLERDKVEAEYQKTMEAIADYEDILAKPERIHQIIYEELLDIQAR
FGDERRTELMVGEVLSIEDEDLIEEESVAITVTHNGYIKRIPTSDFKQQNRGGRGVQGM
GVHDDDFIEQLITSSTHDELLFFTNLGKVYSMKGYEIPEYGRAAKGIPVINLLNIEGDEKI
TTVINVPKDSKVEQPDDLALFFTTKLGVVKRTSVAEFTNIRKNGLKALNLHEGDELINVV
AVKADANVIIGTHQGYSVSFAVENVRLMGRAAAGVRGIRLREADYVVGSDVLAADSKV
FVISEKGYGKQTLASEYPIKGRGGKGIKTANIGTKNGPLVGLTTVKGDEDIMVITDKGVM
IRFNIETVSQTGRATMGVHLINLDEDSVVSTMAKVEPEENSEISEEK
>DNA gyrase subunit B (EC 5.99.1.3)
MTEEERTTAQLEESKEEKAKEYDASQIQVLEGLEAVRKRPGMYIGSTSSQGLHHLVWE
IIDNGIDEALAGFASEINVTVEKDNSITVIDNGRGIPVDIQPKTGRPALETVFTVLHAGGK
FGGGGYKVSGGLHGVGASVVNALSTELDVEVYRGGKIYGMDFKLGKVAHPMEVKGT
CGSDQHGTKVHFLPDPDIFTETTIYDIKVLTTRIRELAFLNKGLKITIEDKRPEEPTYQEF
HYAGGIRHYVEFLNKGKEPLFEEPIYVEGEQKGITVEVALQYTNDFTSTLLTFTNNIHTY
EGGTHEAGFKTALTRVINDYARKVGLLKENDANLSGEDVREGLTAVISIKHPNPQFEGQ
TKTKLGNSDARAVTDKTFSETFNKFLLENPTVAKKIVEKGTVAAKARVAARRAREMTR
KKNGLEISNLPGKLADNTSKDPAISELFIVEGDSAGGSAKQGRSRLTQAILPIRGKILNVE
KASLDRILANEEIRSLFTALGTGFGKDFDVTKANYHKLIIMTDADVDGAHIRTLLLTLFYR
FMRPMIDAGYVYIAQPPLYQVRQGKMVRYIDSDAELNEVLGQLPPSPKPAIQRYKGLG
EMDAEQLWETTMDPENRRLLRVELSDAEEANAVFEMLMGDHVGPRRQFIEDNATFVE
NLDA
>DNA recombination and repair protein RecF
MYLKQITLTDFRNYQKQKLDFSPAVNVLIGKNAQGKTNLLEAIFVLAMAKSHRTSKDKE
LIRFKQEAARLEGTINRRLGDLKLELILSKKGKIAKVNHLERPKLSEYIGQLNVILFAPEDL
ALVKGAPQERRRFIDMEFSQIDPLYLHDLSQYKQVLRQRNKYLKALQLKQANDLVYLDI
LTEQLANLGGKVIAKRLIYLKRLETYAKDLHAQITQGKEQLTLRYKTNLKDLGTQTESNL
ATNLLAKFQQLRQKELSQGTSLAGPHRDDLSFYLNDNNVQTYGSQGQQRTTALSVKL
AEIELMHEQTGEYPVLLLDDVLSELDGQRQTHLLKTIQDKVQTFLTTPGLNDIAQQLIKE
PKIFHVNNGKISTDDENTLQVEIFYPHKDENDESTKN
>FIG002958: hypothetical protein
MQTEIKLKTPYITLGQLLKIADIISSGGQAKFFLQEYAGKILVNGQPENRRGKKLYPDDLV
EIADFGTYIMKA
>DNA polymerase III beta subunit (EC 2.7.7.7)
MKFSVQRTAFIKYLTDVQRAISSKVTLEILRNIKLDLFSDGLTLTGSNADISIETIISIADED
ANLQIFEEGSVALPARFFSEIVKKLPADTITIETDPNRFITKITSGSTEFSINGVDPDSFPH
LPEVGATQSLQIPSDILKQVINQTVMAVSKQESRPMLMGVHLIIQDGTLLAVATDSHRLS
QRKIQLDSAKDANYNIIVPGKSLDELAKMLTDSSEEIEIQIAENQILFILGHTSFYSRLLEG

NYPDTERLIPATSETQVEFNASELLKAIDRASLLSHAGRSNVVKLSIDVDKQTVMISSNS
PEVGMVEEALEFTNLNGNSLEISFNPDYMKEALRSFGQTQTTLSFTMALRPFTLVPTES
EGDFVQLITPMRTAN
>Chromosomal replication initiator protein DnaA
MPNLDALWQFICDNFHETLSLVSYDTWIKTAEPLQFTDNTLVIQVPSDFHRDYWKQRLI
TKVNQLVNDNFQTEIVVLILIPGEKNPLASVAPVVEKKAAQPEAETTGLLRNSPLNKAYT
FESFVVGDGNQMAHAAALAVSEELGTLYNPLYFYGGVGLGKTHLMHAIGNQLLKNQP
NAKVKYVSSETFTNDFVNSIKNKQQEEFRRAYRNVDLLLVDDVQFFAKKPGTIDEFFHT
FNDLYDDKKQIVLSSDRLPNMIENLGDRLISRFKWGLPVEITPPDYETRIAILETKATSEN
LTIPNDVLKFIAGRFETNVRELEGGLTRVKAYAQMKKAPITIDLADEALQNLTDMPAKRIL
DIHSIQTAVAKFYNVTVAEIVGKKRVKRIVVPRQVAMYLARELTDASLPKIGKAFGGKDH
TTVLHACDKIAEAILNDDKLQEEVATLKKRLDVETVN
>Ribonuclease P protein component (EC 3.1.26.5)
MRKSYRVKKEAEFQRVFTQGKSCANRQFVVYMIEKPEQVHFRVGISVGKKIGNAVAR
NWVKRRIRQSLTELKPQLKQDCDFIVIARPGVAWMEMAEVKDHLIHVLRLANVLVDED
K
>Inner membrane protein translocase component YidC, short form OxaI-like
MKKKKPVLASLAALSLILVLTGCSSQTAPITEHSTGFWDHYIIWNFVRAIEGLSSFLGHN
YGWGIVVFTVIVRIIILPLMVYQMKSMRKTSELQPKLKALQAKYPGKDPESLKMMREEQ
QRLYAEAGVNPVAGCLPLLIQMPILIALYQAIFRSETLKAGHFLWMELGDKDPYYILPILA
AVFTYATSKLSMMSQPEQNAMTSAMTYGMPIMIFIMALNFPAALSLYWVITNAFSVGQT
LLINNPFKINKERKEKEQAEKQQARKLAKAKRKALKSRKK
>Hypothetical similar to thiamin biosynthesis lipoprotein ApbE
MTALWTDTHPTFIPNLAARRLDKTYKALGTRIVLTAFGTATEADLDDAFTLIMRYEDLLT
VNRANSEVMAVNHAAGHHPVQVSTTVYKLTKLAVAMSKQNFGFNAAIGPLVKLWKIGF
ADAAVPKPEAIQANLALINPANIRLNDSDLSIFLTQAGMELDLGAIAKGYIADRIADLWAS
RGINAGIINLGGNLLMVGDSPLRSDKLWRIGIQDPFSKRGHSIGLVKMGPCSAVTSGIY
ERHLEVAGKSYHHILDPQTGYPYQTDLAGVTVFTTCSVQAEIETTRLFFAGKNVDLASL
NSPELLGAVFVTIDRGVKVVGLPTDSFQLLDQSYHLI
>Transcriptional regulator, GntR family domain / Aspartate aminotransferase (EC 2.6.1.1)
MVKSDFDFVSIQLKPEKDLPLYVQLYQQLADLIRAGKIVENYRLPAIRKLAQSLEVNPGT
VVNAYKELEKNGFIYSKLGSGSYVAKQEQLSESFNLPLGVTEVDDNQADYLDLTTVSL
SPDLVSVKDFQTAIDQVLKTAGSYAFSAQESQGYYPLRQALCSELAKSSIYTAPQNVQII
SGAQQGIDIIARTLLKHGDTVLVENPTYPGALAAFRACGAKIIDVKMTADGMDMEEAET
KIRLYRPKLIYTMPNVQNPTGITYSKEKKRKLLGLARHYGAYILEDDSLSGLYYQSENTK
PIKTYDQDERVIYILSLSKLFMPGLRLGYLLVPTELLARVRTIKHLADIATSGLIQRVYDYY
LRNGLWSKHLTTLKESYQQRFEYCQQLLEKQLGKQVSYQLPQGGLTFWLKLPAKVSA
KMVAQKAKKQGVLVTPGQVFYARQADDSHLRISFAGVDEQQAKEALIILGQIIKAE
>HNH endonuclease
MMSEVQKVNYIYVVDSNGTPLMPTSRLGMVRRWLKTGQARWFGNSRKIIQFVRPVTT
NTQELTLGVDAGFHLGLAVVGNQREYYASESLRKSEKDRITSRRELRRTRRNRLRHR
QARFDNRRRKDGWLAPSMQHRLDFTIKEIKCLYTFLPITKLVVEVTPFNNQKLVDSNIK
PWEYTQGKMHGFKTIKDYLLARDNYRDALDGKQYPASQLRVHHLVQRRDGGSNQPD
NLVLLSDINHSQANHHNGSLAKLRENRQKELDYRGAYFMSILATRLSDYFARYTTTQG
YLTANLRQKYQLEKSHLNDAFIIAGGTNTTLRTNNVYSRQKLRNNNRVLQKFYDAKYID

SRDGKQKTGKELSSGRTRRTRELNYANLRQFRQEKVKKGRVSIRRNHYQLRPHDVVL
NTKTNKVETVTGVQNNGTVIKFQTGKTCSLKNVVSLYHVNGMLEKKIEDI
>DNA topoisomerase III (EC 5.99.1.2)
MAKALIIAEKKSLADVYRQAIKRAGGKLPYEYDVLQFQGNVVELLEPEDIKPEWKTWKL
ADLPMLPQKFQYTPIKPKAGDDPKTAAYKRRGIAYFKEAKQKLLSGEYDLVVNGCDPD
REGQSIFDKFMLLMPAKIQRLPQLRLWVGDMQVETSLKYLQQPLNNKDPEYVNFGYE
AMGRGIMDWLFGLNASRAFSLVQNSGQAIHLGRGMTAILALVVKRQAEIENFVSQPYW
TVQAKLKTPELLEANQVTYPEVKLVNHFERQDAEAVREVISGQTAQLISSKQSPVRFRR
APKYFNTSDIIAYLSSATYHLSGASITKALQHLYELKITTYPRTSGREMTPSRAENLAAIIA
VAQSIPSLNEKLRRLEVSPQMLAKLKANRYYVKPLSKKTGHDCLTFTGKKFNYASLNAV
EQKVVAAIAFRLLLSLLPSEEHLDKELIFKAATYYFQAKATVLVNPGWTAYVPKFKGNYS
DLPQLVVNQAYPLAQVEILDHKTQAPKAYTTASLIKTMENIGRLYDDPALKAAIKDHGIG
TDATRQLLIDKNEKLGYWSIDKKSRLIQPTKLGRQIIQALGANPLTQVQLTALWEEKLAQ
VAAGQLSLGQFVDLAKEATVEIVAQIKAAPSTLAPRQTEKLTSPCPLDGGDLLDQGQVF
ICSNHDTGCPFIVYRNLLGAKLTAVDLEAILTKGKSNLKELYSSKTHKHFQGYFVWNQE
KGKVELAIQQAKPKLMGTCPLDGGKVYSLAGGKLYACEHHQSQKNAQGQWESSGCP
YYFSGRMYGKKLTQAVAKQIIETGSSPVLTFTSKKRKKFTAILKWNQAEAKLKLEFAEA
SSKAIGVCPLDGGQVSLHGKAYLCENHRAKRLADGSWQTTGCPLVIPQTFAKQKLSLK
TIKQLLKYKRTDLLTFTGQNQRDYQARLVWDEQAARVKFEF
>Glycerate kinase (EC 2.7.1.31)
MKVLIAADSFKGSASSKQVAAYLAAGLREAQADVEVSQVPIADGGEGTVESLVSACGG
KIYRTEVVGPLGESVDANWAMLSHNQAVIELAQAAGLPLVKGKLRPLAATTYGVGQLL
AQAIKHGAKRIYLGLGGSASTDGGVGLAQALGGHFLDAKGAEVKLGGGHLKEICQVDL
TALDKKLAGIEIIGLADVTNPLTGPLGAAAVFGPQKGASAAEVQFLDASLVHLAQVVAKV
HAGSYQASPGAGAAGGTGYGLLAFLGGRLQPGIETIMELIGLEKKLVDQDLVITGEGRL
DGQSLNGKAPLGIAKLAKKQGLPVIAVTGGVGAEIESVYKAGIDLVLSSTTRPMPVEEAI
KQAPSLIKAAGYTALKAYLLGR
>Mannitol operon activator, BglG family
MVKLTQRQEKLLTALAESGVGLSLKDLESLMKVSRRTIYREFADLKLYLQSQGVAITNE
NGLYRLTGQVAKLDGVFEDEVAKQELTTAKRESILACQLLLAGQATKIISLALNLKVSEG
TIQRDLTKLERSLSQYDLHVIKQKGVGVYIEGNEAKRRQVLTGILLNELNEYYFLRYVQK
GEEVAQFKQLFLDAINPQIFRQVYQGLSQTILPVVELNSDRQMIQLILIATISIKRAQLGK
HPHASKTKDNLRYLAKVYEFFATLDIGLRAKLTTDEALFIASQIHLRDYSEQQFDLNEDQ
EWPLTLKVRQFVLDVSKDFGWNFQRNPDFFKRLTRHILNLEKNSHDRLPNVQIETLQT
LTTKYQGLYNSIMKCWPQNFELRLVIPEYQLILLYFANEYTNRRYLQDLAALIVCENGIG
TANILASRLKKELPEIKRVSITHLAVLGQVDLQQYDLILSTLELPGFPREYQLVSPLLLND
EVKRLKDYLANYHRPDYGQVKTSLVKHKNALKKLSRMKAYLDLSLKLITQTQVHKVAN
QGSDLGASIEETLAHIPGEIINDKEKVTLALLRRINLAPIGIPDTNLALLHTSDAGVNQCYF
GIYELAHPIKMLGMDREEMQVSRYTLMVAPKEVTDLEQKVLGIISSSIVMNSENTNVFE
TGTSEAVQELLANQLLNEIPKEK
>Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] (EC 2.6.1.16)
MCGIVGVTGNPNAVEILLNGLEQLEYRGYDSAGIYVNDQKGDDYLVKEKGRIADLRAE
VSADVHGSTGIGHTRWATHGVPSKENAHPHVSNDKRFYLVHNGVIGNFTELKETYLQ
DVTLVSETDTEVIVQLVAKFAREGLSTKEAFKKVLGLVDESSSYAFLLLDREDPDTLYVA
KNKSPLLIGVGDGFNVVCSDALAMLKETHDFIELVDGEIVILKPTSITIEDPAGQVVDRKP
FHVDTDPAQAQKGAYPFYMLKEIDEQPGVMRNLIKEYIAEDGSIKIDADLLKAMADSDRI

YIVAAGTSYHAGLVGKELFEKLIQVPTEVHVASEFAYNPPLLSAKPFFIFLSQSGETADS
RQVLTQVNQAGYPSLTITNVMNSTLSREATYTLLLYAGPEIAVASTKAYTGQIALEAILA
KALGVYKQIATIKDFDVAAELTKAANGMQELVDEKGKIEKIAADMLKDTRNAFYIGRTM
DYAIVQEAALKLKEVSYIQTEGFASGELKHGTIALIQENTPVVAIITDKKTANHTRSNAQE
VKARGANVLYIVSRDLATPTDQIVLPEITPLLSALVAIVPTQLLAYYASIQRGYDVDKPRN
LAKSVTVE
>Integral membrane protein
MSLLILLPIGMGLLLGTQTAINSRLRAFVGSPYIASAISFTVGTLFLLSLTAIKEGLHLSISG
PWWAWLGGLLGTIALTTNILLFPKLGSVQTAVLPIFGQIFASLIIDQYGLFASVQTKLTWI
KLLGLLLVLAGVLLTVVEKNSSNQRPGHVKGLLAWQLAGILAGMLGACQIAINGHLGQI
QHSPLVAASYSFASGAIILWLVCLFKRHSFTSVKATFEQGPRYWWIWLGGFFGALYVF
GSAFLVTKIGTSQIVIFALFGQLSASALIEHFGLLESQPKPVHLQKICGLILLFIGVIITKLV
>periplasmic component of efflux system
MKKKKFIIIGVLAVVVLAVLGGIWSLTHSKEEKKANHYETMVAQKALPLTMSGNVSPKE
VRSFMVTKKNLAQMEVKDGQKVTNGQVLYTTYNQANASELDELKVTLAKYQRAKTTA
NKKLQDARNTLNKMQSSDEGYSEAKDAVASANDEVADATDSINATQRKIDQVSQEVS
PNSVAPFDGDVAVKYDDNGNAKVTVSSSALQMVTTVSEYDYAKVKAGDSVTVSAVAT
GKKQEATITLVDAHPTTTNATNGSKYKVYADLDAKNFIDGQTIKAKVAQAGIIIAKSCVR
QGKVYVVEDGKVTSQAITGKEVDGQYIVTAGLTAGQKVVSNPDGDLKEGEQLP
>ABC transporter, ATP-binding protein
MISLKKINKYYQQGDGRVHILHDISFDVNEGEFVGIMGQSGSGKSTLINIIGFLDREYEG
TYSFKGDDVSYLNNKRFTSLRNQNVGFVFQNFKLIQNLTVAENVALPLMYAGVKKKEID
QRVSKVLERVGLGGYEKYLPSRLSGGQQQRVSIARAIIGKPTFLIADEPTGALDSHTSA
EIMALFKQLNKENKTTIIIVTHDPKVGQQVDRLLTILDGKLISDTGVA
>ABC transporter, permease protein
MTYLDLIKVSYRSLIANKRRSALTMIGIVIGIAAVITIIALGNGVRNKMISEFKTSSSGEQKT
EISFFNAEKNSMQGFTQADLANIKQAFSREISNADFKLETENVATSGIQMGNAVVSGSV
SLLTSPLPSDMMKAGKNFTNSQLKLGQRVILVSEGYAKKVYKHKQNALGTVVIIGDKSY
QIQGLFKPQGYYKYGASFMLPKQAYLQGRKADRGSTLQLTIRNGQDANKVSERVVKY
LKKNGQLRKEGVYQYYNEGVILKSISKTMNMITYFVSAIAGISLFIAGIGVMNMMYISVSE
RTQEIGIRLAVGATQTNIMWQFLMEAIMLTVGGGLIGFALGWGLSTLISVFLPYGIHAVV
TVTDFLLAFSVSTIVGVIFGILPASQASKRNLIDILR
>PTS system, mannitol-specific IIB component (EC 2.7.1.69) / PTS system, mannitol-specific IIC component (EC 2.7.1.69)
MEAKKNTSFKAKVQKLGTNLSSMVMPNIAAIIAWGLITAIFMANGWFPNAKIALLISPML
SYLLPLLIAYVGGRLVYEERGAVVGAIAVMGVVVGAKVPMFLGAMIMGPLGGWCIKKF
DQLFQDKIPSGFEMIYNNFSSGILGMLLAILGTYVVNPLVVGGSTLMGKGVDFIISIHMLP
LANIFIEPAKVLFLNNAIGNGILVPLGIQQAASAGKSVLFLLESNPGPGLGILLAFMFFGK
GSAKASAPGAAIIHFFGGIHEIYFPYILMKPALIIAAIAGGVSGTFTFQLLGGGLKAAASPG
SIIAIALMTAKGAYFGTFAGVLVAAVVSFLIAAVILKHDKSMEDDDDLAAKQAQMQAMKA
ESKGKEMPTAMSAQESVESYRDVKKVIFACDAGMGSSAMGASLLRDKVKKAGITDVT
VTNTAINRLQDEAGLLVITQEELAERAAQRTPSAMHVAVDNFLSSPKYDEVVANLKAIS
QFEPTAKAEAKTETQASDPVANVDFAKVTEVDFISHDQHKGTATMATTLLKERLTKAN
KEVTVKNVKINDLDDDSSHLVVVTAQAAKNLRLRYTNVQIVVVDDLLTSDKYDEIVKKLN
>PTS system, mannitol-specific IIA component

MDLQKDMIVLDKHVATKNEAIELAGKLLVDSGCVEPEYIDSMIARNNDVSTYMGNFIAIP
HGTDDGKKYIKKTGISIVQIPMGVDFSSDDDANSEVVTVVFGIAGLNGEHLNILSQIALF
CSDVRNVAKLADAQSKDEIINLLKEVGE
>Mannitol-1-phosphate 5-dehydrogenase (EC 1.1.1.17)
MKAVHFGAGNIGRGFIGEVLASNGFEIDFVDVNETIIDALNERHGYTIELAAPGQEQITV
KNVSGINNGKHPEQVVAAIKDTNLVTTAIGPKILKFIAPLIAQGIEARKAAGNTAKLDVIAC
ENMIGGSQSLKKDVYSHLDAEGKEFADQYIGFPNAAVDRIVPLQKHDDPLFVSVEPFK
EWVVDEQQMANKDLKLKSVEYAPDLEPYIERKLFSVNTGHATVAYTGKSLGYETIGDAI
KDEKVLNQLKSVLKETGDLLIAKWGFDRAKHEAYQAKIVSRFENTYISDDIARVGRTPIR
KLGYDERFIRPIREAKERGLGYAHLLETVGLIYHFDEPQDEESQKLQAMLKEKDLETVV
KETTGLTDQDLVAQIVAAVK
>Acetate kinase (EC 2.7.2.1)
MEKVLAINSGSSSLKFKLFVMPAEEVIAAGKIERIGIKDSRVEIKYGLGQKYEKVEDIKDH
KVAVDLLMKLLIKLEIIRDYMEINGVGHRVVAGGEYFKHSEIVDDEVIRKIEEIQDLAPLH
NPANLIGIKAFKEILPYACSVVVFDTAFHQTMPEENYIYSVPYEWYEKLGIRRYGAHGTS
HRYVANEAQKLLNQPLKDLKLISCHLGAGASICAIKNGKSFNTSMGFTPLAGLTMATRS
GDVDASAIAYAMDHLGISSIDEMIEILNKKSGLLGISGVSSDMRDVNQAAREGNHRAAL
AGQIFERTVVDYIGQYIAEMGGVDGIIFTAGIGENSVSLRAHIVEKLAYLGIELDAGENSK
RGINNIISTPNSNVTVMRIMTDEEVMIARDVVALEAGQAD
>Lipid A export ATP-binding/permease protein MsbA
MEILRPHLQRYWQEIVGAVLAVIVAAVTALWQPHLLQAVLNAILTNDGVLVKRYGIELIV
VALIGILAGVANVFFAARLAQGVTSDLREETYRKIQTFSFGNIEHFSSGSLVVRLINDMN
QITNFVMTLVMQLIRVPVLFAGSFILGVVTIPRLWWVEVSLITIIVLVAYFIFKRMGALFER
VQTYMDKVNLQVKEELQGIRVVKSFNQEGKQQADFNQTSEQLNQVDLVIGYLFSASFP
AFTLVSYLGVALAIYLVAQTVTIHPQEIAALSAYITYLTILTMAIIIGGMTSMAFSRGLVSLG
RIKEVLVTKPDVEFEQSGFKGKLQGNVTFKEVSFAYPDAKEGDLSLADISFQVQAGQMI
GIVGATGAGKSTLAQLIARLYDPTTGTIEIDGHDLKTLSEETLRQTVSFVLQKALLFSGTI
ESNLKQGKYDATSEELERAIGIAQAAEFVNRYQEGLGHEIEERSANLSGGQKQRLSIAR
GVVGQPPILILDDSTSALDAKSEKLVQEALAKHLQNTTTFIIAEKIVSVLKADQILVLDQG
RLVAKGTHQELLASSPLYQEIYASQKAKERMVAVDD
>Lipid A export ATP-binding/permease protein MsbA
MTKLKTAVAYFSHYLKPFWKGILVVISLTLAATYFQVLAPVYMGKSITALTKLVTGKVAA
SQAWVTFYQAFGLMVLYYCLNALSICVAWMVTSRFNSLSNNQMRKGLYGKLQRLTVK
YFDTHQDGKMLALFNSDLDNIFNAMNEAIFEIISQSALFIGTIWMMMSINKTMTLITMAST
PLVLGLGWLIMRKAKKYLDLQQEEIASLNGYVNEQINGEQVIITNGLQRQSVAQFQKYN
DRVRKTMFKGQLYSGLLFPLMNGLSILNLAIVIAAGSLLVLSGQVETAAGLGLIVMFVQY
SQQYFQPLTQITAIYSLLQLAITGANRLAKVEAEVEEDRVEAGIKLTKLQDKVTLEDVHF
GYQPEREILHGINIEVKKGQAVALVGPTGSGKTTVMNLLNRFYDVTSGAVKFDGIDVRQ
LDLTTLRKQVGIVLQDSVLFTGTIATNIKFGKPNASDQEMIAAAKKANIHDFIMSLPKGYE
TEITEENNLFSTGQKQLVSIARTLLSDPDFLILDEATSNVDTVTEAKLQKAMDNVIAGRT
SFVIAHRLKTIINADKIVVLKAGTVIESGSHQELLAKQGFYYDLYRNQMAFE
>hypothetical protein
MKLIKAVVTLIRWTLKLILTGILLMLVAYFLASSNWLKADNKLALALNQGLVQVEAWLSK
AGVLPKLAMTERKASQADTDDNQAQGARWPQNQARVYLELTDPKLKRACLVAMSAW
NDTGAFKFVQTTDKDKAQIICRAISDSTNGAAGETETDLALPQKQLTHATVYLNKYYLF
NPSYSYSWQRIINTAEHELGHAIGLRHTNEISVMQPTGSNYSIQPADVTKVRELYSR

>putative esterase
MALLKVSFVSETLHRAVPLVVILPTDKIYKANQPKRSADKPYKTLYLLHGVLGNEEDWL
TGTRIQRWAEANDLAGFCCKVSDKLILR
>Mobile element protein
MTQFKGKQFQKDVIIVAVGYYLRYNLSYREVQEILYDRGIHVCHTTIYRWVQEYGKILY
QIWKKKNKQSFYSWKMDETYIKIKGKWHYLYRAIDVDGLTLDIWLRKKRDTQAAYAFLK
RLKNQFGEPKVLVTDKAPSIKSAFRKMDCI
>Undecaprenyl-diphosphatase (EC 3.6.1.27)
MALDFIEILKVIFLGIVEGITEWLPISSTGHMLLVDEFITLNMSEAFKEMFFVVIQLGAILAV
VVMFWNKMFPFQFKNKSQSIIKKDTFSLWFKVAVACVPSAIMGILFDDYLDAHLHTPVV
IAIMLILYGVLFIVIENRNKKRTATTSTLADISYKTALMIGVFQVLSLIPGTSRSGATIIGALLI
GVSRVAAAEFTFFLAVPTMLGASAFKLLKFGFDFTSAELLTLVIGMAVAFAVSVFVIKFL
MSYIKKHDFKVFGWYRIVLGILVLLITAI
>bacitracin ABC transporter, permease protein, putative
MLNLISCELSKLKRSKMVLISVAGVLSTPLLMLIEALQTHFDKPEIIFTLSDIYSDSVLYIML
LVNMMIYVAIAAYLYSREYTENTLKTILPIPISRTKLLIGKFCTLLLWIVMLTLVTWAGIFIV
CGLYHVVFTLEGYSLLVAISWLPKFLFGGILMFLTTSPFVFIAFKTKGFVAPVIASAVIVM
GSVALSNQELGALYPWTATFFLIDGRIESTGYPLALAIGIIILVSAVGFFMTFHHFKKEDL
K
>ABC transporter, ATP-binding protein
MIMEYVIETENLTKQYGETIVVNKINLHVPKGKIYGLLGRNGAGKTTAMKMMLQLAFPT
DGTVRLFGTNYKENIHTLYSKVGSIIETPGFYSNLTGYENLQILAKLRGGVSKSGVEKAL
EVVGLHKEKRKVFSDYSLGMKQRLGIAAAIMHEPELLILDEPINGLDPIGISEIRSFLSKL
SHENGTTIFISSHVLSEIEQIADVIGVMHEGHLVEEVNISELHKRNRKYTEFDVSDGKIAA
KILESSYHMTDFTVQDGTIRIYDFSQSVGEINREFARNGLLITRINDSEENLEDYFSKLIG
GGGIA
>Transcriptional regulator, XRE family
MEFNEKLQQLRTGKNLTQEQLAEQLYVSRTAISKWESGKGYPNMESLKCISKFFSVTI
DELLSGEELITLAETENRSNLKKIYNYIYGILDMMAVAFIFLPLYGNSVGGYVYAVNLLSF
TATTPFNLAVYWSAFAALIIIGIGKIISTHLDKEKWGGIATKCSLTITALAVCFFAAAREPYI
TVLVFLLLIGKIFVWIKQMGMK
>Mobile element protein
MTQFKGKQFQKDVIIVAVGYYLRYNLSYREVQEILYDRGIHVCHTTIYRWVQEYGKILY
QIWKKKNKQSFYSWKMDETYIKIKGKWHYLYRAIDVDGLTLDIWLRKKRDTQAAYAFLK
RLKNQFGEPKVLVTDKAPSIKSAFRKMDCI
>tributyrin esterase
MPAGENSFYVDNTFTGRAYSRFIAQELVEFTRKSFPLSPKKEDTFIGGLSMGGYGALY
NGLKYHDTFGAIVALSAALMVDESLADLPEQAEWFGATKAFWQGVFGPDLKAAGQSE
CNLKVLINNLLAQKVALPAIYLAIGDDDKLLAANVDFHEFLDQKQVAHVFKVAPGSHEW
DFWDRHLKRALDWLPLEQASAGINSGNVN
>GTPase and tRNA-U34 5-formylation enzyme TrmE
MATMTVTEYDTIAAISTPPGEGAISIVRLSGEEAVALAQKIFKGKDLSQVPSHTINYGHII
DPKTQAEIDEVMVSVMRAPKTFTREDVIEINCHGGIVATNKILQLLLANGARLAEPGEFT
KRAFLHGRIDLTQAESVMDLIRAKTDRSMKVALNQLDGNLSHLIRNLRQDILDVLAQVE
VNIDYPEYDDVETMTTKLLKEKALEVKARIHQLLETAQQGKVLREGLATAIVGRPNVGK
SSLLNYLLHEDKAIVTDIAGTTRDVIEEYVNVRGVPLKLIDTAGIRETTDKVEAIGVERSR

KAIEQADLVMLVLNSSEKLTAEDLELVAATNGKKRIVILNKMDLEPKLDVEVLKQHVDAA
EILTTSVLTNEGVDKLEERIAELFFGGIENSQTTVMVTNARHIGLLSQAEDALAAVLAGIE
QGMPVDLVQIDMTNAWDLLGEITGDSYQDELLNQLFSQFCLGK
>tRNA uridine 5-carboxymethylaminomethyl modification enzyme GidA
MKVKQYQAEDYDVIVVGAGHAGCEAALAAARMGNKTLLMTINLDMVAFMPCNPSVGG
PAKGIVVRELDALGGEMGRNIDKTYVQMRMLNTGKGPAVRALRAQADKHAYSAEMKH
TIEREPNLTLRQGIVDDLIIENQTVVGVVTNTGARYGAKTVVLTTGTAARGKIIIGELMYS
SGPNNSQPSMKLSGNLEKLGFVLERFKTGTPPRVDGNTIDYSVTEEQPGDKDPNHFS
FESKDEDYLPVKDQLSCWLTYTKEKTHEIIRENLERAPMFSGVIKGVGPRYCPSIEDKIV
RFADKDRHQLFIEPEGRNTDEYYIQGFSTSMPEEVQQEMVHSVPGLENARIMRPGYAI
EYDVVSPYQLRPTLETKLIKNLYTAGQTNGTSGYEEAAGQGILAGINAGRRALGKSEITL
KRSDAYIGVMVDDLVTKGTNEPYRLLTSRAEYRLLLRHDNADLRLTELGHEIGLISDER
YAKFLAKKAEIAAEKERLAQIRIKPNEAIDAYIEAHGDRPLKDGVLASEFLRRPYVSYEDL
LQFIPAPAKPLDRRVIEQVEIQYKYEGYIKKEDAKVEKLKRMEAKKIPARINYNAIEGLAT
EARQKLTKIQPETLAQASRISGVNPADISILAVYIEQGRIAKTQA
>Sensory box/GGDEF family protein
MFREFMKLLKPTEASIKGNLEEINARVGQILGPLTIGVGFVGIIYNINNHASVHSFVVNVL
VDCFLILTSVILLAYSNYRLSRNVKCWLKFNWVWFLYALASFLYGCYGAATYHGQGFF
TEMYIILTIVCACAFYIPPVLSFTYVVASFLLFIGFNRDFLASIRFIDVGLLAFAILFVNWIRY
LEMLNITEMNRDDQEKMAHLQVLSTNDELTNVKNRYAMRMDFPSYCQQGKIAVMITDI
DDFKFYNDMEGHSVGDKILINFAKCLVDYFGAEHVYRYGGDEFVVLYPNYQEEEFERL
TQSFQAAVDRLQFDNLKVPITCSGGYVYGTAKTPGDLRDFIKEADKCLYEAKRQGKNM
INGNAYLKEIRSNKTKLHYDELTNLMDIYNLRQVGEKILIDQNYRQKSLSFVFLDLDNFKL
INEKHGFKAGDKLLRFAADVIKDAFMQDFIGRASADHFVVISTLPLTEVKDRIKKLHEEIG
TYSQKEIIEIKAGIYYLAPDEEVGQVSTYFDKAKLACNSIKYKFNAIYRLYDQSIEQELLR
RKYIVDNLAKALEDGEIKRYYQPIIRTMTKKVYGLEVLARWHDDKVGAISPDEFIPILENV
RLIHKFDLYMLEGTLQDYQYLTREKGLFGLPFSINLSRLDFQLCDIFAEVDRLFRLYQVP
KEMARIEITESSLIYDSLKLHQTLDNFQKAGYRVWMDDYGSGYSSLSVLKDYSFDGIKI
DMGFLKSFEENHKAQLILASTLELAKSMNMQVVAEGVSSKEQYKFLAESGCDFIQGYY
TGKPQALTDYIGQLDNLETPADLDYYDEICRRNVANISPNLMANMRLRNFFKATPTAIIE
VEDKSRLYFLTENTGFRYFLKLIGYDGDDKNLSIEPADKQAFTDLFERAYLSGKIEYLAF
HRNNVLYQIYVQYISTNKRTQRSAYQALVYLPRAVISSE
>Integral membrane protein
MRKLGQLILLLFSCCLFLGWNNVKAADFSIDSYHVNIDILADGSANVNRTLTYDFDGEA
HGLLYQQAVRDQAQSLTDSKVFINNQPVTLNANQANNTYTLTREGNLYKFKVYHPVD
GDQLKITYKYHLNQVVKNYADTAELNWKIIGSNWDQDLANVKITINLPKRNVKSLQAWS
HGDLSGYTKVDKQRGQVTLSLASNPANSFVEAHLLFDPRVTFANTYRSSERVKAKIQA
QEAALAKKANAARKKQRLMHQIGALFLLIVGFGLNVFAVLRLSRRPKNTLFPASLKSAP
HNFEIPDLPVATAQALCFFTQPNSYAFSAHLLELEAQGQIAIEPLDVKKGLLKDQPNYLI
TLVNARLMHAPEEKLLQYLFKLNSKPAFSLKDLQNLSSKKKKRLAKLFEAWSKQAFEQ
ADIKKYQGFAHPAVSLSSAWVATNLCFFIAMMFANHWELLSLLGLVSLGLAGLDFYCY
KNHSPYLKEDVAKINQVRGFYTMLDDIGRFDLRELGERILWSEIMPYTVAFKLAPKVAK
ALHANFSEAELATVLGSYYPLFYHQSFAFATSFNNSFTAAIVTYSSSSTGSSGGFSGGS
SSGFGGSSGGGAF
>S-methylmethionine permease

MSQQNEHLARKMEARHIRMISLGGVIGTSLFLSSGYTIHEAGPLGTILAYLVGAILVYAV
MLCLGELSVAMPYTGAFHVYAKQFIGPATGFTVAVLYWLTWTIALGSEFTAAGLIMQT
WFPHSPVWLWSLLFIILIFLVNAFSVRLFAESEFWFAAIKVFAIVGFILLGGLAVLGILPIK
GYSHAPLFTNIFKDGLFPNGIGGVFTTMLTVNFAFSGTELIGITAGEAKNPDKTIPKAIHT
TLWRLVIFFIGSIVIMSCLIPYQQAGVTQSPFVHIFKLMGIPFAADLMNFVVLTAIVSAANS
GLYASTRMLWSLGNEGTLPKYFAKTTRAGSPLVALCFSMLGGILALLSSVFAADSVYLI
LVSISGLAVVLVWMAIALAEINFRKQFLAKGHSLAELKFKTPYFPILPYFAFGASLLSCLLI
YFDKNQRLALYCTVPFVLLCYLGYYLCKPKSR
>Homocysteine S-methyltransferase (EC 2.1.1.10)
MDFKTEITKGPLTLDGAMGTALELQDFDTNTDLWTAKALITDPQKVYRVHYDYFKAGA
RATITDTYQANVQAFMKKGYDETISREIIAKAVKLARKARDDFEVQTGRHNYVIAGIGPY
GAYLADGSEYRGDYELSLEAYQDFHQPRLEAVIRAGVDVIGIETQPKLSEVCALLEWLK
VRAPEQIAYVTFSLKDPATISEGTSLVEAVKKVSAYPNVVAVGANCIPLEQAQTALEVM
QSSTDLPLVIYPNSGAKYDPHTKTWTNPTGDLTLAKAASGWYQAGARLIGGCCTTTVK
DIGPLAQVVQNLG
>transposase
MNNSIKTILGIKDPYLKLDEKNFDNPIEDQPNQIIVHLIQTYPMHCPKCGHLMCKNGYKT
VNCLGPELHFKPTIWSIKKQKYICKASSSCPEVVTKLAAVEDIHYRNHISLAIKQRAMML
LTKNESQSDLAKELNVSDWTIRRVITNLDQFFKPNYHWLPRHIAFDDFKSGRFAPSGM
SMILMNIENKRTLDIILSRKNSYLRKYFLRYDRSARLAVQTVTVDLYTPYRHLIHELFPHA
LIIADHFHIVAQAYRAFNKIRIQVMNRAGAGTHKWRALKHFWKLLLTPANELKYDNYWS
RRNFSYAQLTDVEVIHRLLSFDNELKRAYEYYQNLILVIAHRSKKELKNLLAIKWTQLPQ
ALQKVQRTLRRHKQEIYNSFKYDTYTNGPVEGTNNKIKVIKRTAYGFRNFFNFRIRILLA
LPNTYIAINWRNKQTAHAKVQAQAA
>Mobile element protein
MFDQIKQLGLDGVGMLKRSSKVYYRYRGRLYSVKGFYERLRSEQRSPKATYQYSCVV
KSDSGIELRLVFVRNQRKANNYLVLATTKVSLRPAEIIQLYGRRWQIETYFKAAKQYLRF
DKTQVQNYDGLCGHSAIVMMTYDLLAWQERQEKDELTIGDLFYIMGEAMPDLSLTDVL
VWFIKLLNQLVESEPVAPIKQLNETVDKFISSLPQSIASQLQKA
>Nitroreductase family protein
MFGQSSEDVWDIVEERLKSEVANEKAFKKTQAKIASFREGFGTVLFFTDTDVVK
>Nitroreductase family protein
MQHYNPLIDELIASEFDLPSNWKLRAQMPFGSIEAPASSKDFIADEERFKVRFE
>Predicted cell-wall-anchored protein SasA (LPXTG motif)
MQVNKHNKLYKDKETKHEPKNFRLGKKGQELLAKVGLVHSASEAVSDKAKTMAQIED
QRIAAENAQISTNTKKKLKQSLAAGVALGGAVMTAQVAHADTTQGSSQNVTADTALAA
SKARVQAFIDQANGFMASSAYQNATPADQALYQNTVAQAKVDLASGSFNADYYNNVA
DTFQRVINQVERPSQIRFSHVASEEENTNATTTTQQVQAGKAELQTLVNQADSFMASP
AYQNATPADQSLYQTALSQAKAQLATDQADYTSALDNFQRVINQVEHPSHLFYSHTAS
EEEAPASEATTSQATTSEASATVASEAETSSQASTTAASPAQATAVLAANHEETAPLQ
SDFGKLVTEAKASGNLTAVSQTMAEFINTPQGETATAAATPSASQAKPAPATSQVSEK
DSASQAKEHLSYASVGVTGAVMTAAVGAARLKQAKNN
>hypothetical protein
MSLSESQRESLSLSVLAAKESQSEYEASQSTEHSLEESIQASESASLFTSESLLVEQSE
IMS
>Preprotein translocase secY subunit (TC 3.A.5.1.1)

MITGLKMKFRKSLLLRKGIVTILIMVVYLFGRNVALPHVVITNKGLDEMLQLASITTGGDM
ASNSVFSLGLTPWMTAVIIMGLFTTRQDSTFAKKSQKEQSMWQMLLALVLGVLQGLVK
IMETDFEVPGLGPKLATLVVLLGGSFFVVWLASMNTQFGVGGAGAIFMSNIVMANGM
MLGPGISYVLKHRQLLWPALIGGTLIMVLIIMVAVIFDRSEYRMPLKRVMVTNEYGNDIY
LPIKTAPAGAMPVMFGMALVTLPQYFCQLLLVIWPDSKILNWLTTNLVINSALGTVAYIIS
LLALSLAFTYVNISPEDTAEGLQKSGDYITGIAYGDETKKYIAHYMWHMGILGGLYMAVI
AGCPLILGLYYPELQRMVLLPGNLMLNTTFFLVIIDQIRTLRLANRYTKILD
>Accessory secretory protein Asp1
MYYFVTSWYGDQGKWASPNKYWYYEQRKMEFDDSVNQVKMFTLTNNHCELLVLNY
YPQLRNFMHRQGIYAVPYHNFFDELQGIDSNLIRKISFRDLNWPQGTDFRYTPFNVWA
FKDEEVLAYIEFTDEGNLNMIDWQLNGRTDHTYIFDDRGFLSSIRYYTADGNEWYQDY
FDLAGQKQFRQYFNGQGVDIFPGANCDFAKQHYQDMDEVLAERIQVFSQQHLSAKDT
VVVTANKRHNDLFVGRVPAKLALSFFNNRYDLTDEDQLTKLLDQVDLVVAASTKEAAQ
LKRLAPTTQVVEVTPFDTRLEIGKSQQTEYLKLLYWLGDQTDEEIVEALAALVALATKFP
KLALKFASYRSEEERIKTLVENYLAENDLEEEFCFGEPKFDPTDIDNLELEQKEVKVINF
LEVKSERDLVSELEFTRLIVDLGSEPDLFLQIAGISAGIPQINKVPSTYIEAGKNGLVVADI
AGLTQAVTYYFDGLKNWNRAMMYAVNKIMDYTSGKLVAKWKELLGD
>Accessory secretory protein Asp2
MIRMKVLQIGKHNWAEQYEIDESVEWDYCSANQVVATVTAYIAETIKEKGRLETYDCLL
LTDANFSLFALDKLAQIIDPYRVIYDRTLELSPELQTFLGYKCAWKLDVTKPAELIRQITY
YFFPGQSGIRTTPLQINISPNYFGPVNYRGSSQLDLEIDYGNTYQPLLSWGSGFHMDSL
KDLEFWLEFQKDDGCDLQLIFEEVGSIPHYSGKNRYVVTGSELKWPSLIRGHKYQSSL
MITLYAKGKGKITIGTLHSRWSRGPFGNYFLGGKRTVDPKREEVYSFFSPGNLKPPLC
VYFSGYRSLEGFEGYFMMRSLKHPFLLITDPRLEGGAFYIGTKEYEQQVKAAIEAALKR
LNFTNEQLILSGMSMGTFGALYYGAQLSPHAIILNKPLIHLGQMAAKIQTARPYVFDTSL
DLLLKMTARNDVAGADSLDQRIWATFSKGDFEKTTFAIAYMENDDYDLTAFNVIMELLA
SRKSQAKLISQGFLGRHNDDFIENAKWFKNQYKRILAEDFTRGAKQ
>Accessory secretory protein Asp3
MRNIIYWNNFKRTYAFGSQIEIQGEESVFFENHLLTSGQVLATWIMQANYQGERNEPQ
LPLLYPGHSYRLHLIAQTWPAASIYMRIIFKDINGHVLKRQVIRGFDGNFTYPKTAHEYQ
VELLAAGCYKLHFKRLEIEEYGLTVGDDDYLFSGVMNNNEQTAVLNVIFTEPLARLQAE
LTYEGLLDRVTNCLVINSTLSNAKLYLDDKVKAKLQELVKTYQHVNLISYGPVTNLTTWY
YGRYFGKKVTVYGTNDFYSLKRYQKLFREYLHEMFAVDTDLLVARLEKLPKKAARTDL
DLVKPLFNLETKLLAFPMFRKGEA
>Protein export cytoplasm protein SecA ATPase RNA helicase (TC 3.A.5.1.1)
MGLSMNVLALRKVKKILRKVNALKESVAQLSDEELQAKTPYFKQKIKEGVSLDKLLPEA
FAVMREADKRVLGLFPFDVQVMGGIVLHQGNVAEMKTGEGKTLTATLPLYLNALTGKG
TFLVTTNSYLAERDCEELKPVYQFMGLSCCFGAPEEKNLKPAVKRRIYDHDIVYTTNSA
LGFDYLIDNLAKDKESKYMRPFNYAIIDEADQVLLDTAQMPLIIAGAPRVQSNQYGTANT
FVTTLKKDEDYEFNEEETNVWLTEDGVKRAQSYYGIENIFTEEHHELLQHIILALRVNHL
LKRGDDYVVQDGKVELLDKNNGRVMEGNKLESGMHQAIEAKEEVKITPAMRAMASVT
YQNFFRMFPKIAGMTGTGKVAEEEFINTYYMKVVQIPTNRPVQRIDLPDRIYVTLPEKLL
ASLEVVKKIHATGQPLLIATANVEISEIYSELLLREKIPHNVLNANNVPKEAEIIKEAGQKD
AVTVATLMAGRGTDIKLGPGVKELGGLAVIGTEKLASKRDDLQLRGRSGRQGDPGMS
LFFTSLEDEVVIKHGLTWVHKYYDKNKDFDWDQPRLLTRRKFRRALENAQKASDNEG
QKGRETSLEFDESLRMQREIIYQQRNELINAQGGYDVEKIITDQIERFISTHPKLDAFTLS

HYIFSNLTYHYQGDITQVDLTNANAVKEHLLGIAREELALKKGQLANQAEVANFYRTAIL
RAIDACWIEEVDNLQQLRTVVSSRSLAQRQPMYEYHKEAFRSYGKMKADVYQKIVKNL
LLSSVVKTKKGNVIYFV
>hypothetical protein
MANKKEDDLLKQSAKQRLAKVRYDYQPPKKKKSRVRDYLNVVIIILVLSGLVISLLQLL
>Poly(glycerol-phosphate) alpha-glucosyltransferase GftA (EC 2.4.1.52)
MTVYNINLGIGWASSGVEYAQAYRASVLRKLGIKAKFIFMDYFNSENLADLTTNIGFKNE
EVLWLYTAFSDLALSQPSISVAQFKEQLGQEILHEQILPRQVVLQLANNLELRLHLSRLA
PHCVSWVEYLVAGKLQKKAYYTSQLSHEEVYGDKQVSYRQFYNQDGSVALTEMLTA
GKASLYRTQDQVFYSKEELIAAFMKKLGLTAQDIVILDRETGIGQVIFENVRPAKLGVVV
HAEHFSESSTNADYILWNNYYDYQFTNAKWVDFFVVATSVQQEILTEQLRRYNHHAAK
VYTIPVGSLAQLKEASEPRRMWSLVTASRLANEKHLDWLIEAVVKAKQVVTDLSLDIYG
EGNNRAKLAKEIEQADANAYIRLKGHQDLSHVYTKYEAYISASTSEGFGLTLLEALGSG
LPLIGLNVRYGNQTFIDDGQNGYLLNLASYTDKKAVVAALSQAIIKLFTQADLAAFSKHS
YEKAAAYLDSEIETKWEQLVSAEVGAGND
>GftB: Glycosyl transferase, family 8
MINLFERDTEDSRLLLTSLRRAGYDHPTVFLEETGYLPQGVTSPYAFFLAQAGAKLAGK
PLFFNQVPVPAYWEIKGSLSQAEVANHNQRVATISYDQTRGQERLVAKVAWQDANGK
VRLVDYYNQYGWCFKHSYYDETGRELRATYQTDQGVEVISEDKLTQTISLTYHKQQYF
FKNLVDFTIFFLKEAGYKLDQIWYNSLSTPMFIATRLGVTGVDILFWQEGKRDEVPDNM
DYILHSNPQRPTKICVTDKEAYAGFADLFSPADQDSLAYIGYLYNYRRENQARKQALIM
TNSDQIAHLQELVEQLPNLTFHIAALTEMSPKLMAFAAYDNVKLYPNVNPKQVSQLWQ
LADYYLDINYQDEILRATKQAFLNQMLILGFNETSHTNKFGLAENRFPVAAYQDLIAKLD
QLLAHDEEVSKEVFLQKQAAEQTSVAYYHEVLG
>Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12)
MTKLGWLLMAILCIEILCYYGGKKFKFISKLMALLSVGLIFAGLVIASQFVNYHNILADLLL
ATLWWTGLMIIVYFLTGIKKDRMQKPQGIYGQTEPPVAIILSVHDEDYEVVERAAAAVK
NLYYEDKSKLAYYLVDSSERAELKDLAADMEMTYCQKSQDFLADYQQFKGEYLLLLAA
DKVVQNNFLYRTLGYFDNPKVAALQVAFTHYNDYVLQYQPFAAYSATSQRRRDKFGA
AITTGECTMLRRQALASLQLKKPVQEIIELAPYFQRQGWLIRYVKESLANGLAPENMPA
LLRRFRYNMRLNRLGRRIVALRLGFWRQLLHLEWFTYVTFSWRKTLLWACPVIYFWLN
KQAFLHMHHLFLAALVPGGLAIILSLKLQGYSLRDTLGLLVYENALTPYILWDSLVGFIYK
GNRDYPAYHNNSNLKFRLNFKYAWIFLVELLLTLVTCYKAVSDGFQGYREYFLAAWLL
YTAFMNLATLILYAEMPRYRRAERFISDKTCQVYCQQYSLPAKIADWNEIGARVNLLAT
DTYNAADILRDMPGKLVVDNQLLNFKVIWASPTALGIVFDEIDSDQYSYLIDNTYAQPIE
NVKPSRYQEISRRDMRFEEMRPVTFKLSQLPERGQLLDISNSGCKLQARGDLQVEEK
RYILIQLNEHLWQPGKVRWLSKANDKTTLGVEFDK
>hypothetical protein
MNILFGIVFIVLGLFCLVVKPNHYVGIRLPWTLNDPKIWYQTNRLAGGLGIIIGLVI
>hypothetical protein
MDYQKLFEELKIRLDLAPDAKLQAAYQYGVQADELAQLVIKKIKTATTSAYDLYEKDEPL
PQAGGYDLILDSKDQPVALIQVDQVQIVPYLEVDEDHAFREGEGDRTLAYWRKAHDAF
FKEDYALAGKVFDPQTAKMVLEKFHLVVNLTENKLGPIYQINK
>glycosyl transferase, family 8
MKKESISFLLFVDGNEAYLPVWLQSIKEQDYPQTETIVVTKSPLAHLEALQKQYSFTVLL
TKGASRATQHELALAKASGTFVSFLYPTCGLVKDALKLVTPCLNEHPVDAVFSRFALLE

QGLFNFYKQYLGIEEVNRDNYYYYLKSNAEFRTLAGKIIRTSLAKNLDYRQSDQAILEQL
VAQEATVIWNSDQFFFWNLDLKPLPEFNGEAVTLHPRVLERLPLEPSDKISDTISILLCV
DDNYCQKIAPLLSSLGRTTKQQVDCYLVYYELSAASLAFVVKLGNYLSNVKLIPTKISEW
QHEQLQSFSMKNNKLPLAAYYRLLAAELLPNLNRVLYLDIDMLVNSSIEALWQTELGSN
VIGACRDLAFTDHKDSWSYQLLEEQGGNYFNSGLLLFNLEAMRKYKLVERFIEFIQATA
NVYLLGDQDAFNLFFINNTKLLPAKYNTILGYIDQSQIEQPAIIHYCGFEWLKPWRITAPV
AQVKLKWLNTYRQVNRRILGQVETKPLVSLIIPQLPNDSVAAFRKCESLLMQTYPQIELII
PASEQVDYLKANAGANIQINLAAGGASLETMIASAGGEYLFFLNGPHYFKEADALTKIM
ELALNYRSDLVLTHYLKSIKRGKDTILLGPDESKSFLDINLLSSEDIWREHVREFSSLNGL
LVNKHLVGELTKVQTVQALYQQIFNKAQAKFYLDDSLWIEVTN
>hypothetical protein
MTNEHNVAQRAKYYHALMTNHMLTAGEDYQNLKETYVIFLCPFNPLAIEQQRAVGHYR
MISEEGDYLNDGTHTLILNASGD
>hypothetical protein
MDKIDEIKHSKTGGRKYMELTGAFWQDARNEGRKEGRKKGKKAGIKIGERHGKKEGT
DEAIIKLIPDLFALSNNRQAVIKILMTKMGLSQEQAESYYQRAMAAEPKEK
>6-phospho-beta-glucosidase (EC 3.2.1.86)
MTKLQFPAGFLWGGAVAAHQLEGAWQAGGKGVSVADVMTAGANKVPREITAGVIPG
KNYPNHDAIDFYHHYREDIKLFAEMGFKCFRTSIAWTRIFPNGDELEPNEDGLKFYDDL
FDECHKYGIEPVITLAHFEMPYHLVEAYGGFANRKVIDFFVRFAVTCFKRYKDKVKYW
MTFNEIDNQSSFNNDFLMCTNSGLLLNNDKDNEAEMYQAAHYELVASALAVKEGKKIN
PDFKIGTMINFTPVYPASSKPADVLLAQRLDQRRNWFSDVQAWGYYPANVEAYIERKG
FRPDITAEDRIVLKEGTVDYIGFSYYQSMVVSADKVAPDDLADPTEAVVPNPYVEKSD
WDWEIDPVGLRYALNELTDRYHLPLFIVENGLGAVDKVEDDGSIHDPYRVDYLRKHLE
EVRKAVVLDGVDLMGYTPWGCIDLVSAGTGQMSKRYGFIYVDKDDEGHGTLARSRKD
SFYWYQKVIKTNGADLN
>hypothetical protein
MAKEDKERQAKIDALVKRLDNMSDKDWEAEHKKNVQVWIQNSLTKTVPLKKD
>Nucleotide sugar synthetase-like protein
MQVHITNLHGQSLKSTAQRAQNMVTEIATKQLDFNEMGLYFYPVKSDSPDKLSGRLD
GIISALNPGDVVILQTPTWNGTHYEHRLLDKIRGYSDVKIIAFIHDVEAWMFEGMEFVLD
EYIKVYNRCDLVISPTEKMTARLRKHGLTVKKVINQEFWDHPVAVDTTTFPPFNRVVNF
IGRGKKFDFVKNWSAADVKLILTDTDTGWGASKNIEFIGYQNDADLVSRMHRTGGFGL
LWSENQLWSEYMKLNVSYKLSTYLAAGIPVIVEPDTPQANTILAKKLGLVVNSLDEAVE
QVKQMTPADYQEMVVAVEHFAQLIRGGYFTKRLLTQAVFDLFYD
>Poly(glycerol-phosphate) alpha-glucosyltransferase GftA (EC 2.4.1.52)
MTIYNFNLGIGWANSGVEYAQSYRAKLLRELGQEAKFIFTDLILSENIEHLTANLGLKDE
EVIWLYSYFTDVKISPTTYTLADLKASLGQEVGGHCPAP
>hypothetical protein
MAKEDKELQAKIDALVKRLDNMSDKGWEAEHKKNVQV
>Nucleotide sugar synthetase-like protein
MNGQASTAQVAQNMVAKIAREQLGFTEMGIYSYNVASDSPDKLNGRLDGIIAGLSVDD
VVILQSPTWNGARFEHLLLDRIKAYRNVKVIVFIHDVVPWMNADAAGELTNYIKLYNRC
DLVISPNQKMTNRLRECGLMVTKVINQEFWDHPVAVDSTKVAPFAKRVTFIGLGSTKPL
IENWDSEQVKLVLTTDKQPMAADKNIEFMGYQYDYNLLGRLRQQGGFGLLWNALPIW

REYMQYNTSYKLSTYLAAGLPVITSPNIAQREVIEAKKLGLFVNSVDEAVRQIENMTSAE
YQEMRAGVEEFAHLIRNGYFTKRILTGAIFNLFY
>Poly(glycerol-phosphate) alpha-glucosyltransferase GftA (EC 2.4.1.52)
MTIYNFNLGIGWANSGVEYAQSYRAKLLRELGQEAKFIFTDLILSENIEHLTANLGLKDE
EVIWLYSYFTDVKISPTTYTLADLKASLGQEVVKEVRNGKAVYLYLAPKNSWVAAYLKN
SQSDLVERAEFVSNNCLYRKDYYTYTRLFSEYYTPRDNRAYVYQRRFFNEDGSVAYE
EIMDGDSPQMFRFPDKILYSKAEFIDYFVSKLAFKAGDILLLDRSKGLGQQVFRNKGEA
KLGVMVHADHFSEHSVDDDYILWNDYYDYQFTNSEAVDFFVVATASQRDLLLKQFAK
YEHKQPKVVVIPVGSVDKLRHPKEPRQPFALITASRLAPEKHVDWLVNACIRAHKEIPD
LTLDIYGEGSGKEQLEKLIANNQAQKYIRLKGHQDLKEQYVNYAAYITASTTEGFGLSLL
EAIASGLPLIGLDVRYGMQTFVDNGQNGYVCPWSDSMGAPRIVDNLAFAITELFKADLS
KFREHSYHKAQPFMEANVRKAWAELIATEGGLGHA
>GftB: Glycosyl transferase, family 8
MLNLFANYNEQAELLYHSLQAAGYSHPTVVLEDNGFLPAEVESPYKYFVEATQRSTTG
KPCFFNQVKVPALWEIKGDAKGAGVYYLDEKRANINYHYPKDRRLVESVTWLDRQGN
ARFVEHYNQNGWIFAKTSYNLLHEAITTAYFTQSGQEVIVENHITHDVTLNFQGKIYNFP
TKLDFVLYYLNFRHFELDRIFYNSLALPFMVVLRLPQTGSDILFWQEGLGDQVPGNMQ
LIFNNQAKRTKKIVVTNEEVYQKLRSLVGHNQRKLIKYLPYVYQVFRHNLLRPVSFTLTN
TDNLLNFEELLKAAPDITFNVAALTNMSNKLLDLLKYPNVILYPNANQARLNLLWQGADI
YLDLNVGEEVNNASRKAFENNMLILGFENSVHDDKLVDPESIYATPNYKALGKKLNEIA
NKPNLMKDLLLKQTKAARAANISKYKNILD
>Alcohol dehydrogenase (EC 1.1.1.1)
MKIKAAVVPEVGAQFEIRDNIDLHEVGPTDLQIHMVASGICHSDEAIRKGDASLGYPVIL
GHEGAGIVEKVGSEVKNFEVGDHVILSFYSDGTCDNCLKGMPTKCRSYAQYNLSGVR
ADGEDHFQENGQHISDMFNQSSFTTTTVVDQRNAVKIDKSFDLRKVGPLGCGYVTGS
GTVLNTLKPKPGDTIAVFGTGAVGLAAMMAGRISGCIKVIAIDIVPERLELAKELGATHTI
NSKQEDAVAKIKELTGGYGVDWTVDTTGLPAVIKNAISALAQGGTCAAIAVTPHLIELST
WNDLCVDDKKVVGVNMGDSIPQIDIPRLLEFYKLGWFDFDKTEKFYDFEEINQANADS
VSGKTIKPVLIIDKDYQPEA
>transposase
MVLTSGGIALTLDSLESTFYPFASLCCVFTEQKCYNISKESEVNLMEIQITIKVKLNLANA
EIASSFTNTMEQYRLACNYVSEYIFNHDFDMQQSRLSKVLYTNLRSLFMLKSQMAQSVI
RTVIARYKTVKTQMKRNPYKYQDINTGEWYREIRDLTWLQKPITFNRPQVDLQRNRDW
SYLSSGQLSINTLNGRVKVNPVCHGFDQYFDGTWKFGLAKLLKTGGKWYLHISATKTI
ADFDNQTVKHVVGIDRGLRFLATTYDEQGQTAFFDGKIVMRKRAKYQKLRAKLQAKGT
KSAKRRLKKLSGRENRWMTDVNHRLSKTLVQKFGSNTLFVLEDLTGVSFERTDLPKVL
RNQNKSWAFYQLAQFLTYKAHLNNSEVIEVSAQYTSQRCPKCGTIKKSNRNHDLHEY
HCSNCGYRSNDDRLGAMNIQLLGTQYISGQEQPKFELTTNA
>hypothetical protein
MDLRQNERANLKRCLGGLLYQSYTNLEILVLTKGTLTKQQAEFLAELQPYHRLEIVEAS
QLKEAVAKASGAYVYFLNINDLLLKEDVLAKFYQVANKGADLVASDYQRFDYQTNMIY
WVNKEEQMAFLGATNKQQYQRELGELTGWLVKTSLAKECLASGVSEQLAAKLLRQEA
KQVAVQLWLKTFNK
>Aspartate aminotransferase (EC 2.6.1.1)
MDLPAFGVEEWLNKWETKAKYDISQSSIKALTLEELIELDGTKVADFFVRYANTNLDYG
AIEGSAEFKQLVAGLYRHVEADNILQMNGATGGNFLALYALVNPGDQVISMLPSYQQL

YDIPKSLGAEVKFVRLKEEANWDYDLDQLAKLVTPQTKLITLNSANNPTGTWLKADQV
RRLAKIAASVDAYVLVDEVYEPLEARENFLSIVDAYDKGIAVSSLSKTYSLPGLRIGWTA
SNKEVADLFRKYRDYTTICGGVLNDALAVHALKNKDKILARNRKIVAENMQLVKDWVA
SEPKVALVEPESVSTAFIKLELKEDDYRFFTRLVKETGVLLVPGAAFGFPGHARLGYCC
QKEVLQAGLRRLSTFLHQQ
>2-hydroxy-3-oxopropionate reductase (EC 1.1.1.60)
MQIGFIGTGVMGNAISQNLIKAGHNLTVYNRTKAKTDNLVALGAKWAASPKEVATNSDL
VFTMVGMPSNVENVYFGDDGIFAGLTADKIAVDLTTSKPKLASRLYQLSLLIKNKVSYG
>Cadmium efflux system accessory protein
MEKSNHFLPTDEETERSVEIFKAFGDQTRYRILYLLFNKTLSVTEIATAIGLSQSAISHQL
KILRQTGLVKGIRAGQKVNYQLADEHIIMIFRQVKEHISED
>Lead, cadmium, zinc and mercury transporting ATPase (EC 3.6.3.3) (EC 3.6.3.5); Copper-translocating P-type ATPase (EC 3.6.3.4)
MVLTFKQLTKEERLDLSQIGLSLLLVVLARFYLGDTQLAVIVYIGAYLVAGFQIIVTAAKNI
WQGDWFDENFLMTLATLGALLIKQYPEACAVMIFYRLGEFFQERAVAKSQRAITSLLDL
RPDFARLAGQSEQIDPTQVKVGQVISVLPGERIPLDGIVTQGTAYLDTAALTGESKPRP
VTVGQEVLSGMVVLSSSIELRVTKSFGQSTVSKLLDLVQNSSQQKAQTENFIRRFAKR
YTPVVVLLALVLALLGPLVSGGPFQTWLYRACVFLVISCPCALVISIPLGFFGGIGAASR
AGVLIKGSNYLEALTKVTTIVYDKTGTLTEGKFKVSQVLPVEPDQAEKLITLAALAEKDS
PHPVAQAIVANFPGNIATYKVDQVEQLVGLGVKALYEGHTIFLGNARLMEKQKIAYQAV
AAPQATVCYLAYDGVYLGAILVADELKKTTKQALANLKQLGIKQQLMLTGDNQQVATK
VASELGIAVKSELLPQAKVAEVAKLKEQLSPTEKLAFVGDGLNDTPVLKQADIGIAMGAL
GSDAAIEAADIVLMNDELLTLTQAILIARKTKRIVWENISFALIIKALFLLLGLFGLTTMWVA
VFADVGVTLLAVANALRLVLK
>Nucleoside-binding protein
MKKIVTIIVGLVLLGVVFWFGVVYSNNRNHKQIPTVAMVTDGGGVEDGSFNAATWKGL
VEWGKEHHVAKGPNGYTYAQSLSDADFFPNMDKLVKKGYNTIIATGYRLTDAIKKASQ
KYPDTNFVMIDTVVKRPNVASVQFRDNEAAFLAGVAAAKTSKSKQVGFIGRTDSPVMA
VFYKGFVQGVHTVDPNIRINKKYVGTFSQPRIGQSLAAAMYDNGVDVIYTVAGGSGEG
AFTAAKAVRKNLGRTVWVIGVDQDQSKLGEYDGGNVTLASTIKRVDVAVKDMANKAM
TKKFPGGKTIYYDLSDRGVDIVNTSLPKSTWKLIQSYKQQIIAGQIKVAGRQS
>Exodeoxyribonuclease III (EC 3.1.11.2)
MRILTLNCHSLLEYEVNEQIQELVDFVVRESVTIIALQEASQSMKSPILSQTELLGSGFVP
PSESDVPVKVDNFAYLLVRALRRVKLNFAWSWTASHVGYDKLDEGIALLSLKPLQAVH
APVIVAKPEFASYQDYRRRKLLLAKVAGSGNLTIATGHFDGWQAEIFEAEWQQALSHF
DLKEPVYLMGDLNVDAEQNDCSYQKICQSFQDTYELAAAKGDGLTVRGAIDGWEDSQ
VGKRIDYVLTTASNEVLSSRVCFDGQDSARISDHAGIIVEVADTPSLATVKKRQHYTA
>4-methyl-5(B-hydroxyethyl)-thiazole monophosphate biosynthesis enzyme
MAKKIAVMIADGCEEIEALSPVDVLRRLGQDVDMVGLKDLAVTGAHQIKLTCDKVLDDS
LLDYDMVVLPGGYQGAQNLRNSDQLMALMQKRNQAGKWNAAMCAAPMAFARYGLL
NDSDYTCYPEMEKEFASDVTNGRHHEDMVVVDSTKHIITSRGPATALAYAYKIAEVLG
VDPAPQKEAMLYNFLATNI
>hypothetical protein
MNKKSTSRLITLSLIYLAYMLLFIDRSVLNISLAYIGKDFHLSPTALGSLASAFFFSYSLMQ
IPGGWLVDKLGSKKMVAFTLGLWSLLTIATGLAWSLLSLLIIRFAFGLAEGPYPAAALKQI
SETYPKSKRSQATSITLSSNYAGATIAPFLITFAIGQWSWRGAFFALGLLGLLILGSYCY

YERPFSQTVPTSSEPKQKLSLKQADAKIISFIIIGLSLNVVTKGLDTWMPTYLLAKRHLNL
KNIAWLVPLPTIAGGIAALISGFLLVHLFKNQAKLLIALSTLGTLIFMFGMYQATALSWVIV
CEVLTFFFKALAFTSCFAYIAQTPNLKNYGASIGLVNFGGQLGGFLAPIMIGAVVQLSGG
SFTAAFFCLVLWALLAFISSFFIKN
>hypothetical protein
MNLLEALEINQELLSKTEVRLTMPIKPKHLQPYGLLHGGLNAVLAESAASLGANLYCQP
DQFPVGLDITTHHLKAAQAGKTLTVIAKPLSTGRTIHVWQAKTYLNQTVTSFSTITLKINR
K
>Naphthoate synthase (EC 4.1.3.36)
MSNVTWEKIKDYSEIIFERTGKIAKITMNRPEKLNAFTPVTVQEMIDAFTICRDDSSIGVII
LTGAGDKAFSSGGDQSVRGNGGYVGPDKIARLNVLDLQHLIRIIPKPVIAMVKGWSVG
GGNILQLVCDLTLAADNAKFGQTGPKVGSFDAGYGSGYLARVIGHKRAKEVWFLNHF
YTAQEAYEMNWINKVVPLADLEKVTLEWCHEILKKSPTAIRFIKAAMNADTDGLAGLQQ
FAGDATMLFYTTDEGKEGRDAFNEKREPDFDQFPKFP
>O-succinylbenzoic acid--CoA ligase (EC 6.2.1.26)
MENWLKKRIALSPTKLAVVTPETKLTFRELGDLVSKQMYLLVKLVKKGTHVGVLTNNSL
TGYSLILALQQLGCVIVLLNKRLAAPELQYQLADAEVELLIYEDELALAGLTGLEMVSFS
ELANQTSEACQGQAEFKPDEVTTIMYTSGTSGQPKGVCQTYNNHLYSALGSALNLGL
TTEDNWLCVVPLFHISGLSIMMRSLIYGMTVTLLPKFDADEVMRYLVAADITAISVVPTM
LKQLLALGQSSLNPAFRFFLLGGGPIDKPSLEKCRAQGWPVIQSYGMTETASQVVALS
FNDAASHLGSAGKPHFFIQLKLASDGEIMLKGPNISPGYYRQALKNQRTFVDGWFKTG
DIGSLDQDGYLYVKGRKGEMIISGGENIFPREIEDCYLKLAAVAEIVVAAKEDAKWGQV
PVAFIKLKEDISREKLVTYGRKFLAHYKVPKEFYLVTDYPRTASGKIQRYKLLAQLAKLQ
KIV
>Cytochrome d ubiquinol oxidase subunit I (EC 1.10.3.-)
MSIVTLARFQFAMTTIFHFFFVPFTIGTALAVAIMELCYVKTNNEEYKKLTKFWGNIMLLS
FAVGVVTGIIQEFQFGMNWSDYSRFVGDIFGAPLAVEALLAFFMESTFLGLWMFTWDK
VGKKLHATFIWLVVAGSMLSAMWILAANSFMQHPVGYTIRHGRAEMTDFLALITNSQFF
YEYGHVITGAITMGGIIVAGLAAFRLLKKDKLSATTTNLYKKSVRLGLVISLIGSISVMGIG
DLQMKALLNDQPMKFAAMEGDYKDSGDPAAWTVVAWANEAKHEEVFGVKIPYMLSIL
SYGKPSGSVKGMDTVNKELTAKYGKRNYYPMVNTLFYSFRLMAAFGVLMFLVSALGL
FFTRKKNPILYEKKWMLWIMGLTTFAPFLANTFGWMITEFGRYPWTVYGLFTIQQSVSP
NVSVTSLLISNTVYFLLFTSLAITMIGLVVRELKKGPEYEEEQLRAVTNASVDPFDKGAF
>Cytochrome d ubiquinol oxidase subunit II (EC 1.10.3.-)
MSGLQFFWFLLIGVLFSGFFFLEGFDYGVGMAVQSLAHNEEEKDQIVATIGPVWDGNE
VWLITAGGAMFASFPYWYATLFSGYYLILLIILFGLIIRGVSFEFRHKVPVEKKQRWNWT
LSIGSCMVPFFLGVMFISMVQGMPIDANGNMSATFTDYINVFSVVGGVAVVLLTYLHGL
NYISLKTEGPVRERARNYAQFLYWVLYLGLVVFAVLLFFMTDFFKLHAATTLITLVVIVLL
TVIAHLAVFKRKEMTAFIASGLSLVGVVILLFQGLFPRVMISSLGSKYDLLIKDASSTPYT
LKLMSIIAITLVPFVLAYTVWAYYIFRKRIKQPVIVTGEK
>Transport ATP-binding protein CydD
MVDKAIMRLSGIQKLMGLLAGLDVLQAFFIIGQAYYLSSAITGLWSGESLTTQGSAIGLF
VLCYLGRHLINYFKDERLDSYAATQARLLRRQLLTKLFKLGPQIVQEEGTGNVITMALD
GISQVENYLHLILNKMVNMSIIPWIILAFVFYLDWQSGLVLLIVFPLIIIFMIILGYTAKAKAD
RQYESYQRLSNHFLDSLRGIDTLKFFGLSKRYAQSIYRTSESFRKATMSTLRIAILSTFA
LDFFTTLSIAVVAVLLGLRLIDGRIELFPALTALILAPEYFLPVRDFANDYHATLDGKNAM

QAIQAILTKEEIKKEQVKLPTWSEDSVLELNNVGMAYGQKKILAGTNLKFKGYQKVGIIG
MSGSGKSTLINLLSGFLPAKEGQIKLAGQEVANFDQADWKKQLLYIPQNPYVFEGSLR
ENLTFYTPNASEADILEAIKVVGLDELLAELPQGLETKIGNGARALSGGQAQRIALARAF
LDKQRRVMIFDEPTAHLDIETELELKEKMLPLMANRLVLFATHRLHWLNQMDLVVVLNE
GQVVEVGPYAELLAKKGYLYQLKNEMGGGVNE
>Transport ATP-binding protein CydC
MNKVALFKQLQGDKWVKPFLKQYKKTLYLALFLGFCTFFCGGALMFNSGYLISKSASH
PSNILLVYIPIVLTRAFGIGRPVFRYVERLTSHNWVLKMTSNLRVKLYRILEQDAIFFKRD
YSLGDVMGLLAEDINHLQNLYLRTIFPTVIAWLLYIFIVIALGFYSLWFALVMLFYFGVILIV
FPLWSVLVNGARQKQEKVLKNELYTNLTDNVLGISDWIFSQRGQEYVELHEQAQADLY
TVQAKLRKFNNLRGLLFELTYGGLALVTLLWASGQFVGDHGGAANWIAAFVLAIFPLVD
AFAGLPAASQETNVYADSLERLNKLPDQRVAKEVSELPQAPYDLVVKDLSFQYGKNQ
KEVLTDLNLMIKRGQKLAILGRSGSGKSTLANLLRGDLKPTKGQITLGGVDVSALNEQM
ADYIGVIQQAPYLFNTTLINNVRLGNEDASPEDVWAVLERVGLKKMVASLPKGLETMV
DEAGLRFSGGERHRLALARILLQDTPIVLLDEPTVGLDPVTEQALITTFMEQLQGKTLIW
ITHHLKGIEQVDKVIFIEDGQLEMEGSPEELAKTSPRYQKLKAIDDGRL
>Heptaprenyl diphosphate synthase component II (EC 2.5.1.30)
MSSYWKTYPVLEKQITQVKNLIEQRVQVRNPEISQAVAQLAQAGGKYLRPAFFFLFTNF
GPNQDEDRLTKIAASLEILHMATLIHDDIIDDSPLRRGTVSIQAKFGKDIAVYTGDLLFTIF
FDLTLETMNGSHYMQINAAAMKKILLGELDQMQLRYNQTQSIRDYLRSISGKTAELFRL
ASLEGAYFSGASPEVVRRAGRIGLNIGLAFQMLDDILDYSANQADFNKPVIEDLNTGVY
TLPLLLALKKEPAAFKPYLDKQNQMELADMQAVANLVRTLGGVDAAREVARKFTHKAL
VDIEGLPKSQTKKQLTKLTNHLLKRKI
>Putative prenyltransferase, contains 1,4-dihydroxy-2-naphthoate octaprenyltransferase domain
MTVPVFLELVEMKAKTASVLPFLIGLCFSYYYYGTVHVAYVLMYFVAMFLFNMFVDIWD
NYNDYHRAVDVHDYKQNTNIIGREHLSLVQVKTLMACFFISSLLIGLVLAWLVGWPLLIM
GGFCFLVGLLYSAGPRPLSGLPLGEFFSGFTMGFVISLICVYLNTYQDFAWTWANLGRI
LLIALPNTLWIANLMLANNLCDKEEDEKNHRYTLVHYIGLEGGLWLFTGANVVALLAIIG
EFLWGLAPATVLLSLLLIPFIYKQTKLLWQKQVKRETFICGVRILAVGSLTQVVTYLIGIIL
>NADH dehydrogenase (EC 1.6.99.3)
MKEVLVLGAGYAGLKTVRNLQKESGDFHITLVDMNDYHYEATELHEVAAGSQPKDKIT
YPIRDVLNDKKVTFLQDEVTKIDPANKLVSVKKNGSLHYDYVVVALGFCSETFGIKGAQ
ENALQMVDIDTAEAIHQHIIAMMKKYKETKDKNYLKLLICGAGFTGIELAGAFVDERKKY
AQIAGVDESEIEIICVEAATRILPMFDDEMAQYGVNLIKKLGVNLMTGCMIKEIRPGQVV
YAKGKDDPEEHVVEAATIVWTTGVSGSPVIGESGFEQRRGRVMVKPNLTAPDYDDVYI
IGDVSALMDSASNRPFPTTAQIATRMGAHVATNLAHQLKGEATEDFDFKSLGTVASVG
NTHAFGVVGTKKIKGYPASFVKKAIMNKSLVDIGGLKELLAKGRFDLYH
>lipoprotein precursor (putative)
MKKTLIVSLAALALMLAGCGNQTEKQAQASSDKVSTSHKKQAQTKKASSTSSSSSVSS
QSSSEASSSQAQSSATSSSAPDYSGQVRISATQNLQRLGQFNYLLKYNLGKEIVLPNN
PGYNDESKKLNVWYEGDGNNYSIHYSQAASGQDFNANSLRDVTDYAVFTKKTYGSAS
EAQAEINYQTAAMTQGLPKANLGYGIIGTIDSGAGQKYLHWNEGRWSLGVHAVAVND
EDPVPTARKVVELLERYYLPTPDKYGSGSFSVASSSGQQAPTLKWQAGNSVYSLEAK
DLTVLVRMASAVK
>Putrescine transport ATP-binding protein PotA (TC 3.A.1.11.1)

MKQAIISFNNVMKSYDDDVPVLKNISFDIEEGKFYTLLGPSGCGKTTILRMIAGFSEPTK
GDIYFEGKRINKVPANKRQVNTVFQDYALFPHMNVFDNIAFGLKMRHLPKEQIETKVKE
ALKMVRLDGYEKRGINEMSGGQKQRVAIARAIVNEPKILLLDEPLSALDHKLRTDMQYE
LRELQRRLGITFIFVTHDQEEALAMSDEIFVLNNGSIVQSGTPVDIYDEPINHYVADFIGE
SNIVDGTMIADELVEFVGHQFACVDGGMRPNEPVEVVLRPEDLTLTSVEAGKLKVKVD
TQLFRGVHYEIICYDEAGNEWMIHSTHKAVPGQEYGLNFESEDIHVMRFNESEEDFDA
RLESYED
>Spermidine Putrescine ABC transporter permease component PotB (TC 3.A.1.11.1)
MKKMQTLFAIPYFLWLVLFVFAPVALIVFYSFLDPQGHFTLTNYQTYFASGKYLLMTFN
SVWYAFLITIITLAISYPTAYFLTKLKHKDLWLLLVILPTWINLLLKAYAFIGIFGIHGGINRF
LGFFGLAPHQLLFTDASFLIVAMYIELPFMIIPIFNALEELNPSLLSASKDLGANSAQTFW
KVILPLSMNGVKSGIQAVFIPSLSLFMLTRLIGGNRVITLGTAIEEHFLVTQNWGMGSTIG
VVLIIAMFIVMLLTGEKKKGGK
>Spermidine Putrescine ABC transporter permease component potC (TC_3.A.1.11.1)
MKQLKWSNLYLTLVFIFLYLPIAYLIFYSFNAGDSMSSFKGFSLVHFQELFGDTRLIIILLN
TFMLAFLSALIATVIGIFGAMSIYYLKKRAAKTAMLSANNVLMVSPDVIIGASFLIFFTFLG
NLFSGFGLGLGFTTVLLSHIAFSIPIVVLMVLPKLQEMNDSMISAAYDLGANNWQILTNVI
LPYLTPGIIAGYFMAFTYSLDDFAVTFFVTGNGFTTLSVEIYSRARQGISLEINALSTLVFL
FSLVLVVGYYFISKDSRKRRGKKRRGALLK
>ABC transporter, periplasmic spermidine putrescine-binding protein PotD (TC 3.A.1.11.1)
MKKLRSLLIGIIVIIIVLVAGAYELNKSTGQAGAKIVNIYNWGDYIDPDLITKFEKETGYKVN
YETFDSNEAMYTKIQQGGTSYDIAVPSEYMVEKMRNDHLLVKLDKKKLPNLKYIGKEYL
NESFDPHNTYSVPYFWGTLGIIYNDKYYRASDVDTWAKLWNPRFKNQIMFIDGAREFI
GIGLSSLGYSLNSTNDKQVEQAYEKLKSLTPNAKAIVSDEIKMYMVNGESNLAVTYSGE
AVDMLSKNKHLHYVIPEEGTNLWFDNIVIPKGAKNLKGAYAFINFMLRPENAAQNAEYV
GYSTPDEGALKLLPKSVTGDKQFYPSKQQLAKMEVYKNLGPKWLERYNDLFLELKMY
RN
>Beta-lactamase (EC 3.5.2.6)
MKKFNLMLILVALLVSNPITTKAQTNPNTHAIKLAHKLIKQKHFAGTITLVKNGQIIDQISN
GYQNYQTRVKNTATTPYMIGSVSKAITAQLIALLAQNKKLALTDRLATYYPNLPGASEVT
LAELLNHRSGLDTSYFSLHKHYDSYQTYLTALVHQAKIDPQNINSWHYSSYNYNLLAGII
TQVSGQSYATFLKQAITTPLGIHLDLMQATALKSYSYLGGGFFRQTDYNPYNAYGAGQ
VAMNGRDLYYLLAAFLQSKLLAPNLTATLYHKPPTAAFPYYSAGFYHLQSGVYHLHGT
QLGDETSVTISPDGQQAVLVQSNVTAYTKGGVNYTFDLPLFAALNK
>Galactose-1-phosphate uridylyltransferase (EC 2.7.7.10)
MTPLEFIPAIARQKPTNQDACPFCNLASLEKIIDQTGDFLWLDNKYPTIKDTYQTLVIESQ
EHLGDISNYSYAYNRRLFNYLWDCFDKLQATRAYQSVVMFRNFGPRSGGSLRHPHTQ
IIGFDHVDAYEQLEEANFLGKELLAETTASPEITLSRQPLIGFSEFNVLVSKRQALAKLAD
GVKELVGYLMTAYFNGRCDSYNLFFYERPGKFICKVVPRFIASPYYIGYKLPQVDDATR
QAEVKAEFLAYRAQLRKDLLDEEA
>hypothetical protein
MKKRNLWLGLGWFVLALAIGAVLMYRFTHSGHMGIGPVRSFKLMALMVAALAAAIYQF
NK
>FIG00745420: hypothetical protein

MDVDYSYQRTKFSRLAFISTLLLGTFTMSISQSALSTAYPTLMSAFHLSAATIQWLTTGF
MLVMCISMPLSPWLLNNFSFKTMFLGALVLFDLGSLIIVLTPLKLQSFGFAIMMLGRILEA
FAVGILFPSYQTVLLTITPPAKRGSTMGLAGLVMGSALACGPIISGLILKFFKWQSLFVVF
ICLISLVIILSASGFIKDVMPRHPSKLDWYSVFFSLGLIGLVYVFNSLAKGTGNLKVNLGLL
VVSLVLLALFVYRQLQATSPLLDLRVLKSFNYDLALLLTTCSYIALIVVTIIFPLYYQGVLR
VSPFISGLALVPGAVFLSLLNFLAGKLADHLGFKPTILIGMIMISGGWYLTSNLLSFQSLT
LMIICAAIIEGGNAFVMMPAITLGSNFLSPKLVSHGTAVITTVRQVLGSAGVALATIILTQV
SDHQQLLGLTHLAATLTAFSYVCYTMLVLELCGLG
>putative alpha-dextrin endo-1, 6-alpha-glucosidase
MIVTKDFPFAPANSSRTLHIYLPDDYYQSQEAYPCMYYFDGQNLFFDYEASFGTSWGL
KDFLDVWPKKMIVVGIECSHQGNQRLAEYSPYSFFEEPFGQIEAYGKETLNWLVSELK
PYIDQKFRTYPFREATGIAGSSMGGLMAAYAGIAFNNYFSKMACLSSALSFNFEELMT
DLDKQELNADSRFYLSVGQLEAQANQNQLDPFQTPAVKRNLSLEAALQAKGAQTYFY
LQEYGQHNEASWQEQNQRYLEFLWF
>Trehalose operon transcriptional repressor
MPVQSKKEIIAQDIIQKIRFAQYKPGDYLPSEHRLCELYGTSRETVRKALAQLTSLGLIQK
IRGKGSLVLDVQNFTFPISGVVSFKDFDKQQGKSSDIEVLNLSERTSPSTFGGHNLNRR
APSYYVEHLRYIDNIPSVLNQEFLFKEAVPIITKADVINSLYSYLERTLGLNISYSTKEITVE
RVPERIAKILKLTDDHYAAVIRSMSYLDDTRLFQLTTSYQRPDKFKFIDFERRKNAARS
>hypothetical protein
MEIRKVTLFSPVHYAALALREEVLGSKPDLELEEKLTVFVAILNGQVIGTAAVQLYPLGIA
RIRQVATAKAARGQNVGTMLMDACEEYALTKGAHQVVLTGRQSASGFYYKRGYEQFL
LPFSSHGIDFLWMRKGLAVETEAEIA
>hypothetical protein
MFFLLLVVLWGVIQFQEVFTSNVTFILALNGYTFVGIIIYMAKLLFGSNLDEEWLHHENK
>Hypothetical secreted protein
MFKSLRRFTYSLLAILLASLSVPLLAKADGQSTTTHTLNSSYVVYGAGVASEHYADLDN
VLGTDSNFKKLTSTASDYRKYISDSTNTTDAAMISSVAIAPGDPGSGVKVNIKDYNGSN
NITTVTAQQYALVAQMAGVTDVTIVVTANRPVSGEAALTGVYKAFAADGATLNTTNTAT
ANQMLDATQDAINANSDDKTYPGKLMAAVGDTTKQIAQEKQDGNSLTTDQIAKLLEEA
LAKRGISEQTTTNQVNNIVNVLVNFQNSPISSSSKYVSNVGNTINNVKNSTGNLMNKAK
DWASSLDTKKASETAQNWFQKLIAWIQSLFN
>Membrane-bound protease, CAAX family
MKLKQFLLATISFGLLAVALFVPQISLRLATLPGKKALAWLAFFGLYALLICLYFWVLKKD
AFHKLSKNDWVAIIVGYFTLTMIKLVYADLSQWLYHQTDTKNDNLIRAAMGHNQTTALM
IVIIACIFAPLCEEMLFRGIFMKLFWPQKFFLPIIVSGLLFGLAHSNFNILGTLLYSALGWTL
AFVYKHTKNLSASLTLHVLNNAPVILVFFLGIH
>Phosphoesterase, DHH family protein
MKKFLPAFMKNEQLRLLAFLLLGLAVVAVITAFIVNGLWGLIVLVLSLIVGVLAYSVLLQA
NEDTSEYLADLSYRMKRGEQDAFIRMPLGVLFYNEQAEVEWVNPALQHFFGQEDVLG
RPLTEVDEDLAQIIAENQENHKAVEVEWRGAYFNLLVQSEIHAVYMMEITHYAKLQKRY
DNSRLVIGQIFLDNYDEITKSLNDKAISNLSNYMTNALSDWARQFKMFLKRVDEDHYFF
LGYTQALRELEAEKFKILDRVRESTSKRNSPLTLSMGIAYGGEDLAELSKLSQSNLDLA
LGRGGDQAVVREIDGQARFYGGKTNPMEKRTRVRARMISQALQELMNQSDQVFVMG
HQRPDMDSIGACLGLRRIAAMTKKECWIVLDQEKMHSDIERLLASLKQYPEIKGKIISPS
EALEKATKKSLLIMVDHSKPSISISAELYQRLQNRVMVIDHHRRGEEFPENPILVYIEPYA

SSTCELITEMFEYQTKDAEPINKIEATAMLMGIIIDTKSFSLRTGTRTFDAASYLRSVGAD
LQMTRHFMKENVDSYLQRNHLIDRVEFIGAGNAVVAGEEKKIYDPVTAAQAADSLLNIS
GVEASFVITKRAENTVGISARSNGETNVQVIMEKMGGGGHLSNAATQIEGQKIAQVKD
ELLAILNATLQNDTETE
>LSU ribosomal protein L9p
MKVIFTQDVRGKGKRGEVKELPDGYANFLIKSGKVKQANAQAMSQLAGQKRAEARHD
AEVLAEAKVLKEKLESGDFVVELKAKAGTDGRLFGSITSKQISQALEQQYQLKIDKRKM
DLKEPIRSMGYVNVPVKLHNEVTAKIRVHISEK
>Replicative DNA helicase (EC 3.6.1.-) [SA14-24]
MDNEVLEQNLPPQNIEAERAVLGAVFFNADALTDAMEYIEASDFYRRAHQLIFQAMVD
LNDVGEAIDVLTIQNKLAEKNQVEDVGGLDYLMELVDSVPTAANAGYYAKIVKDKAVGR
RLIAVASDIIGMAQEQNDDMTSILDQAEKKVMDVAENTNRAGFKKVASVLDEALNEVYR
LSTQQEDITGLATGYDALDKMTAGLQKDNLIIIGARPAVGKTAFVLNIAQNIAISTDKAVAI
FSLEMGAESLVNRMLCAEGSIRANNLRTGQLTDEEWDNLTMASGVLSRTNIYIDDTPGI
KVAEIRAKCRRLAKETGNLGLVVIDYLQLIEGSGRESRQQEVSEISRQLKKLAKELSVPV
IALSQLSRGLEQRQDKRPVLSDIRESGAIEQDADIVAFLYRDDYYEREGSDDEDTKASA
DEGANDAGEIEVIIEKNRAGARGTVKLLFLKAYNKFVNITYQQPEGV
>Adenylosuccinate synthetase (EC 6.3.4.4)
MSSVVVVGSQWGDEGKGKITDFLSRDAEVIARYQGGDNAGHTIVFNGETFKLRLIPSGI
FYSDKISVIGNGVVLNPKSLVAELDYLKEKGVPTDNLRISDRAHVILPYHIVLDGLQEKSK
KGNKIGTTNKGIGPAYMDKASRVGIRVADLLDKETFEAKLRTNLAEKNRLFEKFYDAEP
LKFEDIFEPYFEYGQRFKEYVTDTSVVLNDALDAGKRVLFEGAQGVMLDIDHGTYPFVT
SSNPVAGGVTIGSGVGPAKIDKVVGACKAYTSRVGDGPFPTELFDEIGDRIREVGHEY
GTVTKRPRRIGWFDSVVMRHSKRVSGLTNLCLNCLDVLTGLDEIKICTAYELNGEKIYH
YPASLKELAACKPVYETLPGWQEDITGCKTLEELPENARNYIHRIQDLVGVKVSTFSVG
PDREQTNVLDNVWARI
>hypothetical protein
MVISAETSTEDLIKQVMAERNCSYQEAIKLLLELARN
>Acyl-CoA hydrolase (EC 3.1.2.20)
MSSQILTCRQTLAVTSHRVLASDLNEHGTVYGGRLLEILDGTASISASRIGRASLVTASL
DELNFILPFKLDDSFCIETYVTGTGTRSIEVFAKIIGEHLNTGERYLGATAFLTFVVPDKG
VTLGQITPVSEEEQYICAGYPKRRQARMLAYQQQKAFQAQLSLKHPWEN
>5'-nucleotidase (EC 3.1.3.5)
MVKKTLKIAIVSICGLLLSLSSVQASSPKKVEVIFTHDLHSHLDSYRVASPKGIEDVGGVS
RLATVIKTQRQKHPDTLVVDAGDFSMGTLYQTLYERQATELNFLQELGFDATTLGNHE
FDYGDKALAKMLTKAKKLPLLEANIDWQKSASPKLEAEFKKYGLKKYTVVKKGNVKIALI
GVMGNDSIEDTPTTKLKFKDPIKSVKQTVAEIKRKEKVDMIVCLSHSGTSQTKTSEDEE
LAKKVPDLDLIISGHTHTVLNKPLKHGQTYIVSCGAYGRYTGLVKLKQKLNHRWKLVDY
RLISMTQAIPKDQQVEKQLRADTALIDQTYLSKLGYRKDQVIAYNPYNFESVDDHYFKP
GEHRLGDFMADSYLYALQQAKIKADVAIVPAGVIRDTLVKGPVTINDAFNLFSLGKGQD
KLPVYPLVKVYLTGKDLKKVAEIDASLANYLNSARLYMSGLGYEYNPHRALMNKVTAA
YLLDKQGHKRPIKDTKLYPVVADMYSAHMLGSLKKQSYGLLNVTPRKSNGQAVTDYQ
QMVVMKGQREYKVFEAFASYLQSFPKGSQGLATIPSYYNQLHKRKVQKASYNPLALL
SKPNRFTFALIGIVLGLGLIVIWVAKKLLRLKRAKFKRK
>nucleoside transport protein

MYLAINVIGLLIFLAIGWIFSKDRKAINWKAVLIMVVFNLVLAWFLTSFPIGRMIVKGAAAG
FDQLVQVAYEGIAFAFPSMVKVKQMDFFFSVLMPILLIVPLFDILTYIGILPWFIKLIGRGL
SWLTGQPKFESFFAVEMMFLGNTEALAVSSLQLKQMSKERNLTLAMMSMSCVTASIIG
SYTLMMPAQFILTAVPVNVVNALIVTAMLNPVSVPKEEDTIATMHSAVAEEAGDVEADG
KKEPFFSFLGDSILGAGKLILIICANVVAFVALAAMIDKLLGAIYPWLSLEHILGIIMFPFAW
LMGLPVDDAFKFAQYMGTKLVTNEFVVMGEVTKTIRTAAFSDHFRAVLTVFVTSFANF
STVGMIIGCFKGIVNKEKNSLISKNVGYMLISGILVSLLSAAMVGIFVW
>Mobile element protein
MLRVQKVRLYPNETMKQVLDDLCDYRRYCWNQGLALWNDMYDASLVLGDKKLRPSA
RKVRDELVATKEDWQYRLSARCLQLAISDLNKAWQNFFKKSLPDWGKPKFKSKKTAR
QGFKTDRAQIINGKLRLDKPREIKTWADISFKGAKSLVGDLKVVSIYRENGKYWASLPF
EVKVTKKTKTGQKTAVDVNVGHFDYTEGQVKTLPNNLKALYKRIKHYQRLLARKRVAN
GKKATQANNYVKTRAKLQRDYRRVANIQHDIVQKFTTMLVNNYDRIVSEDLAVKQMQ
MSHVASKGLHRSMFGYFKQVLKYKCEWYGKELILADRYYPSTQRCSQCGHVKTGAD
KVGLDGNKKHGTKHNEYVCYECGAVMERDKNAVRNLLALM
>tRNA-Lys-CTT

>tRNA-Lys-CTT

>Substrate-specific component PdxU of predicted pyridoxine ECF transporter
MRTNQNSLRAIVLTGLFAAIIYIGIWVLRIPIPAMVGRPFIHFGNTLTAVAILTLGLRNGML
AGIIGLGGFDILNGYAATSWLTMLEVVIVALVISGLFKAFNYQDSRRNIIILAIVAGVTKIFT
SYCTSIVEALMAGTVFKVALVASFFSLPATVINSISTAICTPLLYFFLKRILQAVTKR
>Pyridoxal kinase (EC 2.7.1.35)
MLDKVLISQDLSCYGQVSLGVALPLLGASGFAPTVLPTAVLSTHTGGFGANSYLDLADE
MLNIIHHWQQLDLKFTAIYLGYLGKRALSVWLNHYQTLAIPSTITLIDPVMGDNGRLYRG
FDDDYVKEMQTLIKTATIVTPNFTEAAFLLNQPQLATSPPTIELAEKSARTLAEKYSIPTV
VITGIDLAADQIGLVCYEQKQDRSQSFVQPKLAGNFFGTGDLFASSLLAGVLAGLATPV
ACKLAMDFISQAIANTPKNHDRRLGINYASSLDFLLNTLNKRRRVDAD
>Two-component response regulator SA14-24
MKKILVVDDEKPISDIIKFNLTKEGYDVYTAYDGQEALEQVEEVNPDLVILDLMLPKIDGL
EVAREVRKTHDMPIIMVTAKDSEIDKVLGLEMGADDYVTKPFSNRELVARVKANLRRQ
SAVAANSQAEEENANSDITIGDLTIHPEAYMVSKHGDKIELTHREFELLHYLAKHIGQVM
TREHLLQTVWGYDYFGDVRTVDVTVRRLREKIEDNPSRPVWLVTRRGVGYYLRNPEA
E
>Two-component sensor kinase SA14-24
MQQLKSDNIKAFKQRVELSTYVDNSLIASLNSSDTAAANQKIRKVLEDINNENISEIQVV
DVKGTIRGDSDVNNRSLVGQKTTDAEIKQVIYNSRSFTKTTYDKRDNRRYYIAITPLLSV
SGNTNTVVGAVYVKANMETVYDSINSVVVIFATASLIATAIGMALAIVISRAITRPIDEMKK
QTSRIARGDYSGQVRIYGQDELGQLAQAVNNLSVKIEEAQESSEAERRRLDSVLSHMS
DGVIATDRRGNITIINEMASEDLNVNPDEVIGRSILDVLDIRSQYTLREILENQDEIILDLSD
AEHEQILNAYFSLIQRESGFISGLVCVLHDVTEQQKIDRERRQFVSNVSHELRTPLTSM
RSYIEALNDGAWKDPEVAPQFLKVTQDETDRMIRMINDLLSLSRMDSGRLQLELELVN
LNELFNYILNRFDMMLDKDNNDTRDTKTKNYTIKRDFTKRDLWVEIDTDKFIQVVDNIM
NNAIKYSPDGGVITCRLLETHNHVILSITDQGLGIPKKDLNHVFDRFFRVDKARSRAQG
GTGLGLAISKEVIALHHGKIWVDSIEGKGSTFYISLPYEEYEGDAWDEA

>FIG011501: YycH protein
MRLNLRKYLLVLGLLVAVVISLVFSYLIWINPARSQANSSATSSTPNSALPDSIRKKVGDI
YLPTNVVFNDDAKQYRLTSDTLDLTNYTKKQLVKWQYEDIKRLPTKSASTYQRYLHTK
NALVLNYPVAVTARTVAETFNFKVSKYKSALISRIIIPLTKMDRIYLLDDKQEAVYMVGLK
KATKQKLIALAEDKDNLRSEVTFEKIGDKYTLFYQDGIKVAQYSYLLNKESTGLFVNRLL
GTNNTSAITAKEQDATTTYSDGAAMRLTIQKKTGAVTFTDYNSSNSRRNYSYENYLTA
GLTNLNNIGVALDDVRYDCYDSKRHLLTYRNYINGYPVIAAESAGAYSILLTNSGKKRLD
FSLYTLQIPVPSAGGALTLPNTRWVISQLLSRGYDQAKIGDIKVGYSWNQNQTSNQVV
DLKPTYFVKYDNQWSDYQSLLNS
>Distant homolog of hypothetical protein SA_21
MNFKKIIWIFFLAFIVLDIFLIYSYTQQNDYVDATTNTVNGNEASSILKSIHNDQITTGKISQ
AKGAGYYIASADDDYLRDKADQLNFVSWAYNNHRLTVTFATMVKLGTNPQKTMKALL
NKENQVINGKDYRYAASLSSKNTLVYTQLVHGKPVYSSAGQIRFTVKNGYVTGYTQGY
LSKVQTLREKKETISQQRALVWLYQYNKLPSNSTIEWSDLGYTKLLAVNNSTIYIPTWN
FYIKNNASGNYYYRRINAFTGTVMEED
>Beta-glucosidase (EC 3.2.1.21); 6-phospho-beta-glucosidase (EC 3.2.1.86)
MTVMRFPKQFNFGATILAKDDEGGFDQGNKGLTTYDVLTASENVAKLDLEEVLVAVEI
EGLASYPSRQGVDFYHRLDEDLQTLVNLGISQVLVSLSWARLFPKGEELSPNEVGLKF
YQILLTKIKALGLEPIVILGDGALPLNLSVKYNGWATRRLVADFNRYANACFDNFGQLVS
KWLTFIDLTAVLNAPYLAGGILSRADAEIAYQALVHQLLASSLASKNLHEQFQASQLGY
GISYQVDLTEGRAAFEQNFKQRLLADLVLRGEYSPLLAQLLEQAEIELDLDELDAETLAE
SKVDFIYLNWPETQHNLEVALLELTERYQLPIIVGDFSSKYRQNIQAVARALSQKATVVG
YFPQVFNGGLVAVASDLTRQKTAAYNDYQDIIFNNGLEI
>Zn-dependent hydrolase (beta-lactamase superfamily)
MASKDDKLKISVLASGSSGNVTYIETPKRKVLVDAGLSGKKIAGLMESIGRSLDEVDSLF
VTHEHSDHIQAVGVLARRYGLDVYANEKTWQAMGPKIGKIKAEQKHLFQPGELKAFDD
LDVVSYQVSHDAAQPQFYEFHHAGATFAIVTDTGYVSEEVAGVISDADAYLFECNHDIE
MLRVGTYPWSLKQRILSDQGHLSNEDGAKALMSVIGNHTKQIYLGHLSRENNMKPLAH
LTVANLLAEHDLGVGHDFELFDTDPAQATPLVSL
>Serine protease, DegP/HtrA, do-like (EC 3.4.21.-)
MAQKEEKQARLKSSFLKTALVAVVAGVLGGGLTIGAYNLYQNQAVSSSDDSQATTKVS
NVKVNVSTQATKAFNKIKRAVVTVEAYQKSDNSIDSIFGFSSGSSEETTSESEGSGVIY
KKNGNTAYVVTNNHVISGADSFYVMLYSGKRVKATVVGKDSVSDLAVLKIDAKNVEQT
ATFGNSDNIQVGETALAIGSPMGSEYATSLTQGIISAKKRTLDITNSSGVTTGSTTVIQT
DAAINSGNSGGPLVNLDGQVIGINSMKLSSNSSSSENASVEGMGFAIPSNEVVKIINQL
VTKGKVSRPALGISGIDLDYVSDSAKSDTLKLPDGVDSGVVVMKSDSDSPLKDVLSKY
DVIVSLGGKKVTGISSLKAALYAHKVGDSVTIKYYHDGQLKSQKVTLSLEASESTTTATS
SESEN
>LSU m3Psi1915 methyltransferase RlmH
MNIKIIGVGKLKEKYLKQGIAEYAKRLGKFCKFEIVEVPDEKAPEKLSAAEMEGVKEKEG
ERILAKIRDKEYVYALAIEGKERSSEEFAQEIAQLTTYGRSDITFVIGGSLGLSPAVLKRA
ETKISFGRFTLPHQLMRLVLSEQIYRAFMINNGSPYHK
>Transcriptional regulator, DeoR family
MLPAERLRLIKQLLTKKEELTTNEVALKLAVSKDTVRRDFDHLAQTGQALRIHGGLMTL
GDSGVPNFKERTKILSPLKHEMAKLAAKFVSHDGLYYFDTSTTLVQLAKLIDFKGTTIVT
NSLDNALALTNSRLMRVELLGGTLNKHNRFTSSMTSLGELANYQFDVAFIGTSRLSPA

GIFVVDQDDSYLPRQAIKQARKVVVVAEQYKFHTDRSSFKVCDLNEIDVLITDEQLTKT
QCKWLRSDTQILYAK
>Hydrolase (HAD superfamily)
MYKHLFSDMDNTLLRTDGSFSEENARVIKQTNLPFTLVSARAPMEMDFVLNKLKLSNP
QIAFNGALIYRQLGTQREVLYSQPLQPSDFTKLMAVIKAYPEISLTWYDANNWYAFKQD
EGTAYETSITHLEPTLVTEADLKSLTGPSYKFMLICFDLEVLGKLKAALLALGLANVSIVN
SNPTYIEITHKDATKANAIAHVQTAEKITKDQLMAFGDGENDLPMLTAVGTAVVVANAS
AKIKQVADFISKSNDEDGVAYAIKKLILDTEL
>YbbM seven transmembrane helix protein
MSQLAVNNVSLALAFALVLVAVYISNREKLGLTKDIFYSISRAIIQLIIVGYVLKYIFNVNNV
LLTAAMSFFIVLNAAYNAHKRNPNQSKHYLNSLLAIFISTYVTLGVLVLAGAIKFIPAQIVPI
SGMIASNSMVAIGLCYRNLHELFINQRQQVLEKLALGANRKQASLPLLHASIKTAMQPTI
DSAKTVGLVSLPGMMSGLIFAGVDPLHAIKYQIMVTFMLLSATSLGSIIAGYRTYVDYYN
DQLQLRI
>YbbL ABC transporter ATP-binding protein
MSILSLTDISYQVEQRLILNKVNLEVQAGDFLTLTGPSGGGKSTLLKIIASLISPSDGEIIF
QGQNQRDYEITTYRQLVSYCFQQPSLFGQTVADNLQFPFVIRNEASSQVQIEQALTKV
NLPTSYLNKNITELSGGERQRVALLRNTMFLPKILLLDEVTVGLDARSIEIVHQFIKQLW
QEAGLTIIQITHNQAEISQAAKVLWLEEGSLSDVTTSR
>Purine trans deoxyribosylase (Nucleoside deoxyribosyltransferase-I) (EC 2.4.2.6)
MNRIYLASPFFSDAQNERIAQVKAALDQNPTVSSVFEPAKHPYEAAEFGSLEWQKACF
MLDTSQIIKADAVVAIIDYKLEESDNEADSGTAFEIGYAYGNNIPVIVVQFDPNKEINLMIR
QAMTAYFDVNKDGLNALSNYDFNELMPSFNAERPVI
>Deoxyadenosine kinase (EC 2.7.1.76) / Deoxyguanosine kinase (EC 2.7.1.113)
MITLSGIIGSGKSSLTEILAQELGTKAYYEPVKDNPVLPIFYKGNEIAAKKRAQGDKDATN
PYAYLLQTFFLNRRFAMIKAAMAEDNNILDRSIYEDEIFMKMNTEMGNATEVEYDIYKSL
LNNMMEELPFAAHKKSPDLMITIKVSYETMIKRIEKRGRDYEQIKTDPSLVDYYHRLLKH
YDTWMKEYDASPLLIIDGDKYDFVANMADRVTVLEQIEHKLVELGNLSHDHLLNLKTKH
SELLA
>Universal stress protein family
MIKMSQEYKRILVAVDGSKGAEVALRTAVEVAKRNNTHLDILRVLDLNSLEYGGAGIAL
DGAKIYDIEQENEAYLLDLRDHLVNVDGLDEANVSVHLRFGNPKDVIVNDFQPEYHNDL
IIVGSTGKNLFQRIIVGSVASYVVRAAKCDVLMARTTEADLAK
>/ Glutamine transport system permease protein GlnP (TC 3.A.1.3.2)
MLKLKLVSCFSHLKSGLSPLKIVFSRTKTEYFLVLIRKIWYYEKVLIIYSQREELFMKKIYS
SLLLILGALCLFTLVHPDRVAAKTYVIGTDVTYPPFEFANDNNKYVGIDIDLMKTIAKREG
FQVEIKALGFNAAVQALESGQIDGVIAGMGITPERRKTFDFSDPYYKTGYVMAVAKNSS
ISSYKDLKGKRVAIKTGTAAANYAQSIQKKYGFKTVTFDDSDNMYNDVISGNSVACFDD
GPVLQYAVKTGTKLKLATKPADGGYYGFAVMKGQNQELLQKFNHGLKQLKADGTYQ
QIVNKYVGSASINKQKANKAAAHKESTILGLLKQNHGALLTGLEKTIYLTLVGIFFATIFGII
FGLLGVVPSKFCQGLSSTVIYIFRGLPLLVLALFIYNGVPALTGIKIPAFIAGIITLTLNEGA
YTAAFVKGGIESVDVGQMEAARSLGLPYGKAMKRVILPQAIKIMVPSFINQFIITLKDTSI
LSVIGILELTQTGKIIIARNLQGFRVWAIIGIIYLLVITILTLISKYIEKKLAR
>Branched-chain amino acid aminotransferase (EC 2.6.1.42)
MTKHQTVDLDWENLGFEYMDLPYRYRAFWKDGAWHKAGLTEDSTVAMNEGSPVLH
YGQAAFEGLKAYRCKDGSVNLFRPDQNAKRLQNSCKRLLMPEVPEEMFVDALKQVVK

ANEAYVPPYGTGATLYLRPYIIGVGENIGVHPASEYIFSVFAMPVGSYFKGGMVPTNFV
ASDYDRAAHKGTGASKVGGNYAASLYPGEKAHKAGFSDAIYLDPVEHRKIEEVGSANF
FGITKDGKFVTPKSPSILPSVTKYSLLYLAKERFGMEAVEGDVYVDELDNFAEAGACGT
AAVISPIGGFQVGDDFHVFYSTTEVGPVTKKLYDELVGIQFGDRPAPEGWIVKAVEA
>Transcriptional regulator, LysR family
MDIRLLTYYLKVAQVGNITKAAQQLHITQPTLSRQLTKLEEEVGAQLLTRSQHHLALTEA
GVTFERRAKQILSLLATTKKELQTQQNDLAGLLKIGCVESKLSLEVANWLQAFQAKHPQ
VQFEFYDADGDDLKEKLDTHYLDFAFLIEPVETAKYNFLTLPYYDQWGIIVNKSHPLAK
QKSLSLGDLKNEKLILSRRNIVKAEIDSWLAPLQALSPKFFINLPSNANFLVLNHNYCAL
GIEGILTLNAHPDLRFIPFAPSHHSGHVLAWRKNNHLGSLEQHFLAYIQALIK
>Transcriptional regulator
MAEFEQAVKENNLFGLGNLNTAYAQYFTGASYLQGLVQAQDEIDVNVANVNFEPGCR
NNWHVHHDGFQILLVTAGKGWYQEAGKPARLLSAGDVVVIHEGIKHWHGATKASWFS
HLAITKGTSEWLEAVTAEDYAKLG
>4-carboxymuconolactone decarboxylase (EC 4.1.1.44)
MVKKQTAGRKQLGDFAPKFAELNDDVLFGQVWAREDKLSAHERSVITITALIAGGNFE
QLKYHLQVGKDNGISKEEISEIITHLSFYVGWPKAWSTFNLAKEIWA
>permease of the major facilitator superfamily
MEQTKQVTTMRARVEVDHPWLAMTSMLIGAFVGMLSETSLNIALPSLKLALGVETSTL
QWLVTGYMLIIGIILPFSSLISKWFTTRQTVIFALADFLVGAVISATASGFGQLLLGRMIQG
IGTGLILPLMFTVAMQIFPPQKLGAAMGVSALVIMFAPAIGPTLTGLLLAKLSWNWIFWS
FAIILAIALVFAIFGLKNVAEITKPKVDFLSLVESVIGFGALVMGVSFASELGWSSPVVLGF
LIVGIVVLFLYSKRQLNLASPILNIKIFAKPAFRTGAVLVMLDFAIILSAMYLLPQYIQSGLA
LPVALTGMIMLPGGIINAVMSALAGRLYDSHGAKMPARIGFAIALVGAIMLALVKPTSSV
AYVICAHIIMMIGCPLAMSPAQTHALNALEGLESADGSTIMNTMQQIVGAIATALATSFL
MLGQSFYNGANKQAAFTNGVHFGIYFTIVLAVVGLCLAFKLKDPAKSKDVDLNSK
>hypothetical protein
MQKFKRFLIALGIASWLCGIIFLGLTFYVLKVLASTTLFFICLFFDFIAFTGIFTTITLGILAFL
KKRSKKD
>Pyruvate dehydrogenase E1 component alpha subunit (EC 1.2.4.1)
MTKGIDLKDYFATPETKTVKVLDGSGQVVNEDFFPDLSDDQLVELFKQMVWSRILNER
STKLNRQGRLGFFAPTAGEEASQIGSNFAMEKEDFLLPAYRDVPQLVQHGLPLYKAFL
WSRGHVEGNEYPADLQALPPQIIIGAQYVQAAGVALGLQKKGSKNVAYTYTGDGGTS
QGDFYEGINFAGAFKSPAVFIVQNNGYAISVPRATQTAAPTLSQKAVAAGIPGVQVDG
MDALAVYAVTKAAREWAAAGNGPVLIETLTYRFGPHTLSGDDPKRYRTSEEESEWQE
KDPLIRMRTFLDKKGLWSQEQEEAWTKEVNEQIDEAMAKTEAVPVQKVSDFLKNTFVE
TPAVIQEQIDVYAAKEGK
>Pyruvate dehydrogenase E1 component beta subunit (EC 1.2.4.1)
MAKTTMIKAITQALDQEIKRDDKVLVFGEDVGRNGGVFRATDGLQDANGKDRVFDTPL
AESGIGGLATGLALQGFRPVPEIQFFGFVFEVMDEIAGQISRIGHRMGGTRKMPITIRAP
FGGGVHTPELHSDSLEGLMAQTPGLKVVVPSGPYDAKGLLLSAIRSDDPVVFLEHMKL
YRSVKEEVPDEEYTIPLDKAAVKREGKDVSIITYGYMVQESLKAAADLAKEGIEAEVLDL
RTISPLDEEAILATVKKTGRAVIVQEAQRQAGVAAQVASVIAEKGILSLEAPIGRVSAPDT
PYPFSEAESIWLPNKADIIAKVKETVNF
>Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.12)

MSKYQFKLPDIGEGIAEGTIGEWHVKEGDKVEVDEDLVQIENDKSVEELPSPVAGTVTK
ILVAEGEVAEVGDPLIELEVAAGQGNVSDDTPAPATKSEAPAAGGGDVYQFKLPDIGE
GIAEGTVGEWHVKVGDTISADDDLVQIENDKSVEELPSPVAGTITKILVPEGEVAEVGD
VLVELAVAAGQGNASASAVAKATPAAAPTGAKPAATVGQAQDHSLPVLAMPSVRAYA
REQNVDLTKVNGSGNHGQVLRTDVDAFIANGGASPVEKVVPTTDLREEEVATSPAKV
KATVQALDAQWPEHREKMDQIRKATAKSVSRSVAEIPHVHIFDEVVVDKLWAHRKKYK
ELAASRDVKLTFLAYIVKALAVVMKEYPVFNSSVDMDNNEIVYKDFVNVGIATDTPRGL
FMPNIKNADRLSLFNIARAISENTAKAKDGKLTATDMRNGGMSITNIGSVGGGYFTPVIN
WPEVAILGIGRITQEPVVLDDKVQVARVMKLTLAFDHRVIDGATGQSAMNRLKELLADP
ELLLMEG
>Dihydrolipoamide dehydrogenase of pyruvate dehydrogenase complex (EC 1.8.1.4)
MVVGDFAIDLDTVVIGAGPGGYVAAIHAAELGQKVTVIEREFIGGVCLNVGCIPSKALIE
AAHHYQHAMDSQDMGLQVTAAKLDFTKTQEWKKSVVDRLTGGVASLFKKHHIDVIWG
SAYLKDAHSLRVIGEDKAQTYTFNNLIIATGSHPIEIPNFKFEGRVIDSTGALSLTEVPKE
LVVVGGGYIGSELASAYANLGAHVTILEGSNSILANYERDLVKVVEHHFSEQGVDIYTE
AMAKRAEQDGDKVEVTFEVKGEEKTITADYCIVSVGRRPNTKDMGLEQVGIKVDQRGL
IEVDNQSRTSVENIYAIGDVVPGFALAHKASYEGKVAAEAISGKATTVDYHAMPAVCYT
DTSVATTGLTFAEAKDQGLDVKKSQFPFAANGRAISMNSTDGFVRLVYLKDSGILVGA
QIVGVHASDLISELTLAIESGNTVEDIALTIHPHPSIAEAVMDVADVAIGFPTNI
>Lipoate-protein ligase A
MPKMRYIAMKTDDIRTNLATEQYLMNSGKLEPPFMLFYIEKPCIIVGRNQNTLEEINQKY
CQEHNITITRRLSGGGAMYQDLGNMCFSFVVQADKQKFGDFKSLVKPVVDALHEMGA
TGAEVTGRNDIVIDGKKFSGNAMYSRNGKTFSHGTLMFDVNTDSVADALNVPKDKIES
KGIKSVRSRVTNIKPYLKPEYQELDTYQFRDELIKKIWGVATIEEAKAYEYELTAEDQAGI
AEIEQKLYKNWDWVYGKSPEFSVQKRKHFDGGTIDARFQVEEGKIKELKIYGDFFGPG
DVTELEDALRGQEYTPDKMIAVLTKLDLGKYFVGIAQEDVIDLLAYQH
>putative secreted protein
MISWQLGIMQTCDNFATMPFRLGQLAQIVAYTPFLSGSKVRLVLSNLYGEADLCFDEIY
LSKEEKFHVKHQVTYLKQKQVVIRRGHKLQTDPLTLKVQAGEKLYVWMKASRKQTYA
DVASTYATDLINASYGQKFDYCPSLRVNEHQRKSWFSLEGVLVWTKHKPKVVEFAGD
SLMEMGFVSQALGKLFLEEYPQRVTYLNSAVSGSRLLYDCPTDSPLLATYGQSLLKRK
PKPNSVALSICFCGGNDLLLPAYSSQASKQVVTPATLVAGFRELLTSWPNPVVSSTILP
APKIQAQSEATRLAVNQELSSWSQIFDGAKLLVDGQGKLKPTYQLGDGLHINQAGGSL
LAQQLLAFLKTNSLL
>unknown
MYNLIPDTMIEVKTKLDWLLANGFDATEDSVIQDGILAGADLMFEDALEGPYWTVLWAV
VDEKLLVRGALGETVGYITPRAGFSKFSDDFKQNAPSTWFNISSQVEKLV
>ACT domain protein
MKAILTTVGQDQVGIIAGVSNYLAQKQINILDVSQTIMSGYFTMMMMVEVPSADLDFVQ
VATELKDLGTKLGVEINIRNADLYQAMHQL
>hypothetical protein
METAKILETIHMISDENLDVRTITMGISLLDCIDSDSSVACQKIYDKITSKAKDLVKVGQAI
EAEYGIPIINKRISVTPISLIAAASHDKDYLKYAHTLDRAAKAVGVNFIGGFSALVQKGYQ
TGDQTLIASLPQVLAETELVCASVNVGSTRSGINMDAVKQMGQVVVAASKKDVMTNA
KMVIFCNAVEDNPFMAGAFHGVGERDVVINVGVSGPGVVKTALEKVKGQPLDVVANTI
KETAFKVTRMGQLVGTVASERLGVPFGIVDLSLAPTPAVGDSVAQVLEEIGLEQVGTH

GTTAALAMLNDAVKKGGIMACSQVGGLSGAFIPVSEDAGMIDAVRAGCLNIEKLEAMT
AVCSVGLDMIAVPGDTPAETISAMIADEAAIGMINNKTTAVRVIPVPGAKVGDKVEFGGL
LGHAPVMAVNSASSAEMINRGGLIPAPIHSFKN
>Transcriptional repressor of the fructose operon, DeoR family
MLAEERHQIILNLLKTNSIVKLQDICEQTQCSESSARRDLQLLEEQGLLVRVHGGAKIKY
SLQRELDMTGKASRNTQEKAAIATAAQAHVQAEDVIYLDAGTSTLALISLLDSSLQLTVV
TNGVMHASLLADRGIRTILVGGELKNTTKAIVGVEAVEALKKYRFNKTFLGINGIHPKYG
YTTPDPDEAAVKAVAIEQSEQVYVLADSSKFDAVSFVKVAEIGQATIITSKLPQRIAGQY
SNQTTIQEVLT
>Tagatose-6-phosphate kinase (EC 2.7.1.144) / 1-phosphofructokinase (EC 2.7.1.56)
MIYTVTVNPSIDYIVQLKELTLGEVNRMDYDNKLPGGKGINVSRILKELGLDNVALGFLG
GFTGQFVKEELEKVGLKTNFTPVAEDTRINVKIKAQSETEINGRGPKISEAEVAAFKAQF
DKLTADDVVIFAGSLAPNLDDNFYFDLIELIRSKGAQFVIDTTGESLLKTLAQKPLVVKPN
NHELADLFDVEFTGMDDIVKYGRKLLDLGAQHVLISMAGDGGLMITKDKVYRSYAPKG
TVINSVGAGDSMIGGFVGTYAKTKDPLEAFRYGLACGSATAFSEDLADKAKIDEILPLIKI
EEY
>PTS system, fructose-specific IIA component (EC 2.7.1.69) / PTS system, fructose-specific IIB component (EC 2.7.1.69) / PTS system, fructose-specific IIC component (EC 2.7.1.69)
MDIRDLLMKDIMIMDLKATTKAEVIDEMVHNYFVKGIISDEELYKKDILAREAQSSTGMG
EGIAIPHAHDTAVKKPAVMFARSEKGIDYDSMDGKPAHLFFMIAAPEGGDNTHLQALAA
LSQVLMNPDVVNALKASTTADQVQDIFAKAVAEKEAQNKAEEEAEKTAANTNSGRPYI
VAVTACPNGIAHTYMAEENLKKAAKALGADIKVETNGSEGVKHRLTADEISRAVGVVIA
ADKKVEMGRFDGKPLVNKPVTAGIQHPEDLIQQVLDGNASVYHATGDDSSSSAESDG
TVWGNIYKQLMNGISHMLPFVIGGGILIAISFIVEQYMGGAKAPAFIFLNSAGSLAFAFMV
PVLAAYIAESIGDKPALMPGFVGGFMASIANSSLGGAYHVNFQANSTSPSGFLGGIAAG
FIAGYLVLGLKKLFAGLPRSVEGMKPMLLYPIFGLLFIALVMYYVINPIFSTINIWITHFLNG
MGTGNLVLLTLVLAGMMSIDMGGPFNKAAYVFASGAFANDPKSATAAILMAAVMVGG
MVPPFATAIGTAFFKNKYTEDERRAGVTNWVLGFSFITEGAIPFATADPGRVIPSCIVGS
AVSGAIVGAMRIGIPAPHGGFWVSPLATNVWGYFLAVIVGSIVAALIMSFWKKPVEK
>lipoprotein precursor (putative)
MQVFNKKLWILIVTIILVGMLGGCGSVTEKKTATELSVTKVSEAKVAYKKTLNLLSEGKT
SQAYEVLKPYQNSDRQSVKQLVVNLKNLNEVKKELAANNLKVVKNKLADLLEVTKPKE
FVKQVRLTNKEYNLISLANTYYQEIQRYYQAGKYNEAQGSLEALNGLESNYQVVNELQ
TKAKNYKDKIAVALAKQNTSSTTSSTTSSYVNAKNSKLVESQYSNQTGAALTSASSQA
VSSVASELTNNDIINQFQTASGISKQDGDQYFIQAQANKEYLIEIRTVSKTNPNVSVLKG
MYRFNSQTKVVQKLDSLTGEYKQISVN
>hypothetical protein
MVIYFDIFDFSKVLDGLRKEYRFEEINPEVLSGEKVGFILKFQENNQRFKLLGNELFISSS
YYLKNKGKVPDVEEFIKFEREFK
>Superfamily I DNA/RNA helicase protein
MRNLRWAKERNSENLDSYLGLRVDKNQVNLEIRKNKPDYNPAVFEQILAPHNYPLGRF
PSNTKYALSLMQQIVVNLTSNVDTKNIRSVNGPPGTGKTTLLKDVFADLVVKQAMKMS
ETALLSGKLGDNMGELAELPNGIARNNIVVASSNNGALENIVNELPLEAEIDGTLITELKK
VDYFMEIANEIAGDSGSNLSKWGLFSLAGGKYSNVKILIEALKKMLFYLESEYQPNFNV
YKEYRELHHSVELLRNEVDAKANEYRNLPVNIQKLNMLAQKYESFEEKFENEEAQKRY

LVQKKEADYRLLESQNKHKQAKLQLELLNLDSKITLNRETLNKITEQKPWFFEIRKKRFY
KNNLKACRHTLKQLEAQKVSLKTDYQYLAGLTTEAKIEFEQKQVEYSQWKKLAKDRQD
RLLREKSLQENKVETIKRGINTLRPLDLKVGYQQLQLMDPWFTEKYRRAQSRLFIHALG
VRKQFLYENRVQLRKAVKVWKNQNKYRLDSPKLINNAWNWINLAIPVISSTFASFGSM
FRNVDANTIGNLFIDEAGQATPQSAVGAILRSKRIMAVGDPEQIKPVITLDPLVLNMLKIR
FDLPEKYMTEDSSVQTLMDSASHYGFSKDNDTWIGISLRVHRRCAYPMFTIANDISYD
GLMI
>Mobile element protein
MVKKPKLRPRKVKQATQKRCFGKSIDERPVEANLRSEIGHFEIDTIQGKKMATDKVILVL
TDRKSRYNIVKVINSKTSQDVNKAILELKAEYGSKLFKTITADNGSEFSELNKVHSEVYY
AHPYSPRQRGSNENNNGFLRRVITKGKAISSFPEEEIYYVNDIMNEYPRRIFNGLSAEE
VFHTHYVLPSGPRVALSLTSGPLISNQPIFPYLSFP
>ABC transporter, permease protein
MKYIEKKMSAPLMSCILTILVLVSMYLIMGYAPFGSKTIAAMDANIQYLDFFAYLKNVLLG
KDGINYSYTSLLGGSNISLVSYYLMSPLNLLVIFFKKSQLGIFFNLIYLLKLALAATTMTYF
LKKVTEDKLSHAFSVILGICYALMQYNLAQGSNIIWLDGVYMLPLILLGVHYLIIGEKSILL
PVAVGCSIIFNWYTGGINCLFAGLWFIFELVLQTKLSSDELIKGIKKCIYFVLNMLLGVML
SAFIFIPAVMALGNGSRNQMDWAVLTNTFRGNVLSLVERYIIGGQSDATGVSLFSGSVV
LLGVVAYFFLKKVSLKRKITTLLFGIVVVLTLYWQPAYFVFSLLKQVDSYWCRFSYVAIA
FLIYVAAQFYVEYEKSSFYKYGLVGIASIYSVAEIVMHRHGLTWNVIVSIIFCIVIATLLTIP
KRNWTNSLALVLVIIELMVSAYPLLKFYSYADNQSFVNYELQAEKQVNELKQYDSSNYR
VTQTSPRGVVDDGIVANYNEGMAFNYWSISSYVSSQSRTQLDILNSLGYRSEQDRITV
VNSPIISSDALLGVKYIFSRYPVNGLTKINKLGTYNDKATYKNKYAFPLAFIGDGKAVNT
QGDWSTFVYQNRLFNNLYGKDIDLFQTVGYKQSVDGNNIKYKLTLPVGRYALYGNIIW
AKNMDAVLDVNNKTKIAYAKWLSPSVFYIPTNNSDKSATVTLTSSVNGLAIANQEFYAL
NLDKLKLVSEKANKNAVPVSFSKKSMAIKVTSTKKNQYLYLAFPKDKSWKITVNGKVVV
PETYANGLMVIPLGRGTNNLRLTYKVVHQKLGILITVIGILSVGMLFYINKRWR
>hypothetical protein
MNAANPKFAKYSNGPIEGINRKIKLLRRLSFGMPNFTNMKKRIILWLNFKPKNDAHKVL
GRARFSIFWQKSMK
>Mobile element protein
MNFKNYYNTLSSHQDEINQINHLYLNITSKGLIMDEMPHVEPYKKFDRALFVPAKLKFS
QRPTCIFCGFKAVEYHGYTPSTLKHIAADLSFPVFIELKKHRAKCLTCGHCFQAEDADL
APKYAQKTNQVTQKIISELLTLTTMKDITNHNNVSPSTVVNVLNESVTLSSFTLHDSLPE
HLCFDEVRTNDNQLSFICIDAVKHTLVALLPGKSSYDIEKFFNGFSLAQRARVKTVTTDL
NAYYASVAERLFPNTEIIIDRFHIIKMINQAFNHERVKKMKLFAKPSVEYNLFKCHWRLFL
LDPAKLDNEHPHYRRQLKRSMTDAQIVSEALELSEELLLSHNVIHKLHRAIICNDVLTLA
RCLRTFKQSFKQVSVQKAQ
>Bactoprenol glucosyl transferase (EC 2.4.1.-)
MTKTLSIIVPCYFEEESIPLFYAAVEKVKPQLPGIELEYWFINDGSTDNSLAEMKKLNEQ
DPAHVHYASFSRNFGKEAALYCGLKNATGDYVTVMDVDLQDPPEMLPEMFKYINEGY
DCVGTRRMNRKGEPPIRSFFAKLFYKLINKISDTEIVDGARDYRLMTRQMVDAILEMSE
YNRFSKGIFSWVGFKTKYLEYKNQERVAGETSWNFWKLFTYSIDGIVNFSQFPLDLAS
FFGFISSLISGLGIIFIIIRYLIHGNPTSGWASTVCIILFVGGIQLMFLGIIGKYIGKIFTEVKNR
PIYIVKEKK
>Predicted cell-wall-anchored protein SasA (LPXTG motif)

MNKVLQKVLLGMTLTAIFGGGATLAKADDVNTSSDSSGVNNVLANDSNSTQSTQASKT
SLADQISSTSQAAQVLTKLQAATSTSESSSTSYLNAETSSTNQTTTRNVEEPSSNTDSS
SSSSSVAEVATSSATASVGSDTASQASQTGKVLADGNANADTVKKAPRNNNTQAAKQ
AASTLATTMKGASQNNLTFYNDVDEAVVGFNGNNIYNRGYNYSYDNLDAVKYNGDTR
NDKAQVINDYANLTDSELKEVNFFAAQLLNTARQSVWAVNGQDSNTWQLMITKDSLG
LAVKVAQGYTRDRHSISTGTHDTAVLKQAYSSYGLHGYSEDYGSTALNYAATMTTNRK
TTMYDLKKAIFQTLTLMLYDDMPATSGGSNAGADGNGHAYSLAYGWDKYSFGANYFG
LAFDQMGQVHIVNFRPSMIWSDTRFDTNQLVYTKTLGNTSALTKNVPTAYSNYLKASV
FNNDSSISYVKEKVVGNTAYYFVTRNGQQYIYDSNLTKLKGQVKIGGTLYNLSNSTGIV
LKNNSEITATEVAQRLASNQNVAYFDSVSIDGANMIVSGWHAADASIIRPYSYIILFDKT
MNRELRRVAYTANYRGDVARVYPALYRSANSGFRVVFNLQGLNYAGHQLQLVARYS
DDVAGNGNTTDVWSSDYTFNTNAAYFDSVKIEGTNLLVRGWHIADAAANRKYAYIILFD
KSANRELRRIAYNPTNRSDVARVYHNVYNGANGGFDVSFDLRGLDFAGHQLQIVARY
TDDVAGNGNAMDVWSGDYTFNSNAGYLDSVKKVNDKTINVSGWHVADAASRQKYT
WIIVLDKTLGREIARTKYVQSTRNDVYTYYSGSYNARQSGFSNVNVQLSTNLSGLNNH
VFQVVARYSDNSYSGEGHYTDFYSKDVRFRNGMFYL
>hypothetical protein
MEETTYSLITIVNKEAVYKDFLDSLNTQTGVKYELITINNEHNQYSSARTAYNEAARRAR
GEYLIFLHPDIRFLASDSLAKICDYTKEIGDFGLVGIAGSPKELVKNERIILTNIIHGSNRET
VPNSKKITGPTEVQTLDEAFFIMKKSFWEKFPFPNKEGWHLYAVEQCLIANLNGYSNY
VVPANIWHTSDGKSEDYRYTLQVRKLIREYRPYVDNINTTVKRWRTRGLYANLYSLAYL
MKMYLKYLLNK
>Threonine synthase (EC 4.2.3.1)
MTLYRSTRDLQAKAITASQAVLQGLSPDGGLYVPVTMPQIDFKLEDALAFSYQELAYRI
LKPFFSDYSQAELTDCINSAYGPNFTDPLVAPVNYQADNAYLELFHGPTIAFKDIALQLL
PYLMTTAAKKNQLTNKIVILTATSGDTGKAAMEGFADVPNTEICVFYPKDGVSPIQEKQ
MLTQKGANTHVIGIKGNFDQAQTNVKKIFNNRDLAEQLATKGYQFSSANSINIGRLFPQ
VVYYFHAYLQLVKAGKVTLGAPVNFSVPTGNFGDILAGYYAKQLGLPIKKLICASNKNN
VLTDFFNEGTYDKNRPFYVTTSPSMDILVSSNLERLLFYISGEDAAKTAQLMANLNQTG
KYSLDEDMKANLADFAAGFADEASTEAEIARVYQASGYTLDPHTAVASAVAEQYKKAS
GDSTPMVVVATASPYKFPQAVLRALGASEIDSGLDAVNQLQALIKVPLPKTVADLFTAK
ELHSLVIEPSEMEATVIDILY
>Homoserine dehydrogenase (EC 1.1.1.3)
MEVINIGMLGLGTVGSGVVRMLTDHAEKISLITGRKLVVKTVVVHDLTKKRSVDMTGIN
VTDDVTAILNDPEIGLVVEVMGTVDIARDYIKRALKAGKHVVTANKDLIATHGPELTSLAL
EQGCDLYYEASVAGGIPILRTIVNSFAADKIVEVKGIVNGTTNYIMTQMNQAGISYEQAL
KQAQELGFAEADPTNDVEGIDAAYKMIILTQFAFGKDITLDDVKLEGISNVKAEDIAEAK
RLGYKLKLLGISKLVNKHLDVSVGPVLVPRNQPLAAVQNENNAVFVTGAAVGETMFYG
PGAGELPTANSVLSDIITVAKNIVMGTTGVRFNDYSADVKLAVPSEVVYPYYISLKVLDQ
PGQMLKLTEIMTKAQASFNLITQTQLDEDYARVTMVTHKMSEAQLNEIKAALAQLDTIE
VIATYKVLGG
>Homoserine kinase (EC 2.7.1.39)
MTVIKVPATSANLGPGFDSLGIALSMYLTLTVGPKSDEWIVEHDLGADIPTDEENMIIKT
ALSLAPNLEPHKVRVESKIPLARGLGSSSTAIVAGLALANHLGKLGLGKDELVQKATELE
GHPDNVAPAILGDFVAATYVNGHVTSVKGSFPKLSFLAYIPDYELATSQSREALPEQLD

FKDAVVAGSIGNALVASILTGNLAEIGPLMEADRYHEQARLKLVPHLEVLRKLGHEVGA
LATYLSGAGPTVMVVIDRNKVADFIKEAQALGLTGRFEQLEVERQGLMVSE
>CDP-glycerol:poly(glycerophosphate) glycerophosphotransferase (EC 2.7.8.12)
MPNIAIIGYNIFGIGGTTRSNLNLMHEFDHSEYQLTYYNYTEFNKRDVLSLKNKYPYTSK
VDFRFFLELYNLEPKQSYDYIFVTREDFFPLAKILRKAYPQAIIIGEIHTPLALLKRKLTGH
EYFSCLRVATESIKERLSSQYGYTRIYVQTVSLAHLNWNFKEQVTNSNLLVHSRFDDS
QKDISYTLRLLNQLVNNEGQKQVKLYLSGAGPDLGKYKKYIKNHQLKNNVTISQRKLPQ
NFTAISTSHYETLGYSIIESVAQGHRVLAYYGDDDVVYENLADVPLITWLSKDLKEDTPR
VLAALRARPTLAEFRESQAALEKLAGHYLEKFLANTQLYQGVKFDLAKITSADIPTCRN
QIEAILGSPLAPWYIKLYRYLKKTPLKIFLPK
>hypothetical protein
MEWLTLREVSEKSGISLNTLRQRFKRNYYFKRKNYVYDYGLDMWREVKLTKRDIQINP
KNKLKEKLISPRFADVIIQETVVKSEQAANEYFEKEIDKLEFPQTVKLTKLINRGILREDLI
KSSLKELRKAEQIRNDAIDKLEDLQGILEIASEYIDDIEFKSTSSSLKDLMSQKSSKIRDFN
KMKKDRFIDLKRDIEELKKESDYRRFH
>hypothetical protein
MKRRYNNLIVIFIFAFLSWMFSYYFIKIGGMVLDSDSSFHISRVEEIYLNLREGKVFTFIAT
RTFSNSGVGNFMFYPTLFLYPWALLRFVFNPVTSFYLWYGFFMFLTMVISYYSMLSYS
SSKIRSYAFALVYSLAAYHIHLGLWHYVLGEFIAYTFIPLAFLGIYHILFGDSKKWKILGCG
MALIMYAHLLSVYITALILVLITLGSIIYRKGIAKERLVSLVKSAVLALLLVSPQLLLFITDYK
GLGAPQKSFVYMSTVTEFIDSSLNNTSGTQTVGLVIIIVAFIGWYFVRNKSEEITAYVLGL
GFLLLATRLIPWDLLTKTFVFNILGQVQFPYRFNTYSSFFLAVTASLIIQEILEIGSRDYLK
LVNIIALTIVMVISYESAVTDVITRIDSNHIHLKKSNKELVTLPVNAYVDKDDYSNIFTYLVL
FGETDYYPKQSYASYNNINSSVNAQTIVKHEGYLNKRKISIKLVKTSANTIIYKVKTTKDG
KLDLPVVKYPHTKVILNGNATKFDTSDRGTIQVNVNKGINNIMIRYQPSKLFYAGIGIVVI
TWGILLVITISKVLLKIREHNIGGLRV
>UDP-galactopyranose mutase (EC 5.4.99.9)
MTKYLVVGAGLFGATFAYEAAKRGHEVKMIEKRDHIAGNIYTKEVNGIQVHQYGAHIFH
TSNKKIWDYVNQFADFNRYTNSPVANYNGEIYNLPFNMNTFNKLWGVVTPAEAQAKIE
EQRSILGDKRPENLEEQAISLIGTDIYEKLIKGYTEKQWGHKATDLPAFIIRRLPVRFTYD
NNYFNDDFQGIPVGGYTQIVEKMLAHENISVELNTDFFAKKEAYLNDYDKVVFTGMIDQ
FFDYKLGELEYRSLRFETEVLDEDNYQGNAVVNYTDAETPFTRIIEHKHFEFGKGEKGK
TVITREYPANWQRGDEPYYPVNNKENNALYTSYKELADREAKNVIFGGRLGQYRYYN
MDQVIMAALVTVEKEFEG
>Membrane protein involved in the export of O-antigen, teichoic acid lipoteichoic acids
MKVVRNYLYNAGYQILALITPLITAPYISRTLKPHGVGIYADTNAWIQWFVLIASIGIALYG
NREIAYVRDDKEKMSRTFWEIQIVKLVMTLVAYIALVIFLQIYTKYTWYIWVQSINILAATL
DISWLYMGLEDFKRTVIRNTGVKILSLVLILTFVKTLNDVALYIFLTGFSTLLGNATLWPR
LRVMLTKVKLSTLRPLRHIKPSLALFVPQIATQIYLILNKNMLGIIVNSTAVGYYNNSDSLV
KMVLALVTATGTVMLPHVASEFAKGNKEAINEMLYSSFDFVSFLSVAMAFGLAAVGLH
GVPYFYGKGFEPVGPAIMIESAVIVIIGWSNAIGIQYLLPTDRVKEFSTSVIIGAVVNFILN
FPLMELWGLMGAMVSTVLSEAMVTGYQLWTIRKDIEYKRLFTNTWKYLLGGIVMFIPV
SWLDRNLHTSILSILAEIILGIVIYMLLMLILKPTILGKASEMIRRKRGN
>Alpha-L-Rha alpha-1,3-L-rhamnosyltransferase (EC 2.4.1.-)
MISVVMATYNGEKYIEKQLESILNQSLIPDEVIIRDDGSSDATVQLVKDFIQKNKLSTWNI
EVNDRNLGYKKNFANLLSLAKGDYIFLSDQDDEWISDKIDKMVAIFKDNKAIQALNGGV

RLIDANSNKVEIVNRKNVYNANFYFSKKPVVRLNEVPLASLIISNVTPGCAMAITKELRD
TFISSYEGTFPHDWYLNMLAAYNGGCYFLNEEVISYRIHENNASSPTAATGILSKLQKF
NQEKKRRMERMHSQVAIVPNIENSLGKINSESSLAYNFVKARLAFYERSNLRNLIAMRK
YDDYKDRTVLKGEIWDLILAFKVDKLIYLFGNINGE
>hypothetical protein
MFYLAFITYLVGHFMSGTMYILVFGDNLQYRIWLVACLLALFKIILFNSYRTWQDASVVV
LIGLVIWASSLNTNTWDMIYYYVMIVAAQDMKFQNIVKVFLTTIITLVLITFISAKLGIIMGL
TNVRDGSTDLRYALGMVYPTDLAARLFYIELFYVVLRKFKLNLSEYVIWVAVIAFGYIVT
NTRLDALLMLMILIVSLLRKQVISVLKKLSVGGISVVVGIGIATIVGITYLYNPHNFILNKIDG
VLSGRLHFGHIAFERFNVQFFGQVIQQVGNGGLHGRVYDYFFIDCSYIRVLMMNGLVS
FLLLMILILGLIRYFMSVKSYTLVLALLFVILSSVIDQHLLEISFNCLFLAVFADITYFKNDEL
NSLYV
>Galactofuranose transferase
MSSDLVITSMDPGKNMGGSKAKLDIINFLEEVGYTSYIVNPYVSKLKKLYLAEFAIPKDF
KGKKYNNVIIQFPIPSEKLTKKLIKVIRENSPEAKIVFWIHDIQGLQSTDEEVTERELAIFEL
ADELIVHNQAMMGWLKNRGYSKKMVNLEIFDYANPQPVNDSEIYQQSLCFAGNLFKS
EFLRKLQLEHTIDVFGPNMFTDAPDCINYRGQYSPEELPKYLIQNFGLIWDGPSVDTCM
GTFGEYMRFNNPHKTSLYISSGLPIIIWKEAALAEFVQKNKIGIAIEKISDIDKALASLSQT
EYNEMKKNTLAIAQKLRTGHYTKAVVQKI
>Glycosyltransferase (EC 2.4.1.-)
MYYPLQVGSAEENFPGFLRDNTGDNIAEKNPNYCELTAQYWAAKNRNADIKGLVHYR
RLFSNGKRNFFSSVDKKFADVMTSDTLAKLMESYDLILPTKRNYYIETSWSQYAHVHY
EKDLVVTREVIAEKYPDYLPAFDKWVNRRAVHKFNMLIAKADIFDQYTEWLIDVLEEVE
RRIDISDYTPYEQRVYGFLSEILIDVWVEKQGIKYKEIPVMFMGNQHWVKKISKFLLRKV
RGRA
>Undecaprenyl-phosphate galactosephosphotransferase (EC 2.7.8.6)
MFFKNCNGAGYHLIKRIFDIVVSLIALVFVLPITLVVFLIDCFGENKGPVFYKQKRIGKNHT
PFYIYKYRSMVVDADKKLYDNKKLYELYVKNSYKLPPELDPRVTKFGHFIRKTSLDELP
QFFNILKGDMSLIGPRPVVEEELQEYGDRVDKFLSVTPGAMGYWQATGRSSIPYPERC
DVELYYVDNASLGFDIKILFKNIISIFKTEGAY
>Glycosyltransferase (EC 2.4.1.-)
MSRPLISIVIPVYNVEAYLERCLDSVLAQTYQNLEIILVDDGSTDSSTQIVDRYAHKDSRV
VVIHQTNGGLSNARNHGIDVANGEYLTFIDSDDYVATDYVEYLYQLLAKNNFQSKLALC
SLKTIFTQTGGFIDAGNGSEQTLSGEKAIEMMCYHDLVDTCAYAKLAKKELYDKVRFPE
GKLFEDIATTYKLFDQCETVECGFVAKYFYMIRGNSIVTSGFKPSKLDLLAMTDEMASY
VNQKYPSLKAATLRRQVYARFSTLNQMLTVTDYKAERDEILNFLRRNRKAIMSNPRTP
KRDKLGYLMLTFGFNFYKLAWSGYLKVKTRKH
>Alpha-L-Rha alpha-1,3-L-rhamnosyltransferase (EC 2.4.1.-)
MKISVVMSTYNGQKYILEQLDSLRNQTRQADEVLIFDDRSTDATATLVADYIKKYALNG
WQIKVNPQNKGWRRNFMEGMWSATGDIIFTCDQDDIWRSDKLAVMAKIMGEHPEIQL
LTSNYCEFFDNGKRRVGPNSNTKELKKILLKNNYLLVGQPGCTHCIRKSLIEVSKKYWQ
PEYPHDALLWRLAIFADGLYTYTDDLINWRKHATSAFAKESRDLKTVKEKKKWVATAQ
KFDKTLWRYIKEDVVSDTTKQERVLSRNDKWLAEREKFYQTKNILRGIWLARYWDCYP
RYRQYPAEWYLLFIKRK
>Glucose-1-phosphate thymidylyltransferase (EC 2.7.7.24)

MKGIILAGGSGTRLYPITRGISKQLIPVYDKPMVYYPLSTLMLAGIRDILVISTPEYTPLFE
ELLGDGSEFGINLSYKVQEQPNGLAEAFVLGKEFIGTDSVCLILGDNIYYGSGLSKLLQE
AAQKESGATVFGYHVNDPERFGVVEFDEEMHALSIEEKPAQPKSNYAVTGLYFYDND
VVDIAANLKPSPRGELEITDVNKEYLRRGKLDVKLMGRGYAWLDTGTHDSMLEAASFI
ATIQKRQNLKVACLEEIAYRMGWISKDKLVELAQPMKKNDYGQYLLRLAQED
>dTDP-4-dehydrorhamnose 3,5-epimerase (EC 5.1.3.13)
MGKLVVKETKLQDVKLITPAVFGDHRGFFTESYSDRDFKEAGIDFDFIQDNHSLSTQAG
VLRGLHFQRGKAAQTKLIRVATGAVLDVIVDLRKGSPTYKQWEGYILSESNHRQLLVPK
GFAHGFVTLTDNVNFLYKCDNYYDAAADGGISFKTPELNIDWPIDLDKAITSEKDAKQP
TLTEFEKDNPFVYGEI
>dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)
MFKNIIVTGGAGFIGSNFVHYVVNNHPEVHVTVLDKLTYAGNKENLAGLPADRVELVVG
DICDKELVDKLVQKADAVVHYAAESHNDNSLQNPEPFIQTNIVGTSVLISACTKYGVRF
HHVSTDEVYGDLPLREDLPGHGEGPGEKFTPETPYKPSSPYSSSKASSDLLVRAWVR
SFGLKATISNCSNNYGPYQHIEKFIPRQVTNILSGIRPKLYGTGKNVRDWIHTNDHSSAV
WKILTEGKLGETYLIGADGEKNNKEVLELILELMGQPADAYDQVKDRPGHDLRYAIDAS
KLRNELGWQPQYTDFKEGLKATIDWYTNHRDWWEAQKAAVEAKYAKNGQ
>dTDP-4-dehydrorhamnose reductase (EC 1.1.1.133)
MKILITGGNGQLGQEMQHLLNQHGITDYEAADVEKLDITDEASVEAYFAENKPEVVYHC
AAYTAVDKAEGEGKAVNEAVNATGTEIIAKACAKYDALLVYISTDYVFDGSRTEGEYMP
DDPTGPRNEYGRTKLLGEQAVKKYCQKYYIVRTAWVYGKYGHNFVYTMLNLAKTHDK
LTVVSDQVGRPTWTKTLAEFITYLVDNKIAYGTYHCSNDGTCSWYEFAKEILKDKDVEV
LPVSSDEYPTAAFRPHYSVMHLAKETGFEFPMWQDALHEFMAEIGEK
>Phosphoglycerol transferase and related proteins, alkaline phosphatase superfamily
MSHKLKEAISITLVLLLTGLVLFSQQYWQIDLSLEKSITNKYSYLSLFKGITFLGKPLFLIIL
GFLLGNETNLRYVRAIKFWLVMVVTSISLQVIILYFDNNFSISELYNSLLPLLRNTYPLAS
GIIVGLCFLKKVDEYSLKLSWRQLIIVLVVATGLPSIFGQDPFAMNDGNGLTFGLVMFLF
GAIVARKKVWLQTKRCLILATIMGASYLLINYWMPYISYDVHGNYSTAGRFLNSGALPA
VLFAIFMGMGILPFLGSLARTWKLRLADFAYLGLICATNSSLLLILKRYFEKNSTAPQVML
LIIAELGGLALGTLMLAVLARTAVNKRLEEQWLGYDLIDVLDSLKSKVLENKRYLLGILGA
VALSYSSFFFVTPNLELTIGKMTHDIGEYILFARPMMIWLNALIIIALWCIIFGLTNRYWVS
TLGVAISMLVLFVTENIKVGIRNEPILPSEIVMLKAINELLHMATTPVLVGGAVGLGLGISLI
VYMERKYPQPGIKWMTRLILVAVSCLFFSSSLFLNHNKGKMHTVSLALGNDPTFYNQL
LGAQQNGPLLQFLDNVDTQVMEKPADYSQAKMAELANKYNGVAQKLNQTRTNNWQD
QSIIFSLSESFADPNRVPGVSLAHNPIKQIQGIMAKNTSGLMLSSGYGGGTANIEYMSLT
GLTVANFSATLPTPYTQLVPFQKQAWSFNQLFAKSAAIHPYEGVFYSRPTTYKKFGFE
RFYHLGSKYKITDQETIDRSPYLDDKTAFANTLRVLKERGNGQFINLVSMQNHFPYDKG
YYDGTKKYRATGSAVTDETTADMIADFATGLSYTDKAVANFIKQIDKLNRPVTIVFYGD
HLPGIYSGNNMKKDGVVLHETDYFIYSNKYARKHGAKTKLAKATWLVAPNDFPAMVAE
QTNSKVNGYLALLTKVKDELPAMSTNPKLNATNNYNTMVSFVTKQGKILKPEQLTKKQ
QELYHDYQLVQYDMTAGKQYLLKNKDFSK
>hypothetical protein
MYLIIRGFVLLVTLLVLFLMTRSYFEDNKLMIAGIYLVVIFFASFLLYRLATFPNIFLDEGN
GMYDSWCLANYGMDSHLIKNPVYLEGHAGQGQSVLYAYIAGIFMKLLGYKLFIFRLPLV
LISISTLLILIYVALRRGSARIAFWITLVVASSPYILDVSRFGMDCNIAPFMVASGSAITYLG
ITQEKRLVRLIQLILGFLIIGLTAYSYNVGWIFLPVYLLTLAIFLLTTRKVKITEVILPVALLIIE

LIPIMIFAVRSNIPSLNKTVKILFWTSPRLQVGRADASFIDFKNHILVNIYHNVTEGLTMFI
NGTDGLSWNSVGNFGPYYMFTLPFFIWGLFIILKRRSIMDYFVIAQLVGMIPIMLIVTPNY
NHWIFLHFSVLMVIGIGIIDLVSNYNKFKYVLIIVYVISMVSYTKQYFTLERYTGWDIGAID
KIVSLNTSKYNKVYFASDDGAFLYVARVAAPVSPNEFQKTKDNPYSKVNLGMDNKYSN
FERLTPDSKISNNSLIILPREKVSEYGKKIKGRRNSNTFRLANVDYLVYE
>Choline binding protein A
MNTNRKKEVIKKAKKKKLQQLTLMAGLTIITSSAGVLGTKIKADQVDMASSTSATLVSQS
SSSEVVTSSSTSQVQSKQQDTGGDDTSKADSSTAEKVSESNQVEEKAETSASQEITK
TEVTSTEANVTSSTSSTSSTSSETQTDSNNTGAAESESTTNSGNEQVESQVTNNSTPS
ETTDVQTTTLDTSNITSQTSVANFAQVLAANTPDYFQIGDSRYTSVDAVDLASYQRWM
TQNDFNTLKALGVSTVIIKLTESTTYTNPEATKFINYARNAGLNIAIYHYVRFNNVSAAVG
EGNYLANSMRSHGLSGDTLIFADIEDEDTIVNSSNSNAIASLNAFWNTLNGRGFTNHAV
YVYQSYLYRDAVISTVGRDRTWLAQYPYTPTRGGYYETLHKNEGYGAWQFASTAYLP
GRAYMGSIDVSHDYNGLLTKRKEQNLAAFDDISLHGNTLHVSGWHAADISKAQGYASL
IVYDATTNREITRVSYTPNRRDDVYNVYSDIYNSRYSGFNVNIDLRPYNLANDTLKIVAR
YSNQPNGAGEYTDVWSDGKVLVNNVAHFDNVVITDTSLTASGWHASDNSYKQPNSFII
LYDITANKELARVKYTPTHRGDVQKVYGYIYNSANSGFSNIKFNINKKDLVGHTLQLIAR
FSDDKQGNGSYTDVWSDKYTFNQKIASFEQAEIVGTTLHAKGWYAADESLVKPYSYLI
LYDNTSKKELKRVRYTPQYRSDVYARHSEIYGSANSGFDVNFNLAGVNYAGHDLRLIA
RFSNDRAGNGSYTDFWSGTYKFNQNAAYFDKVEVKGTQLHVTGWHAADEAAVKPYS
YIILYDATSNREIKRVNYTATARPDVKNVYKNLYNSDNSGFDVNFNLAGVNYAGHDLRL
IARFSNDRAGNGNFTDVWSGTYKFNQNAAYFDKVEVKGTQLHVTGWHAADEAAVKP
YSYIILYDATSNREIKRVNYTATARPDVKNVYKNLYNSDNSGFDVNFNLAGVNYAGHDL
RLIARFSNDRAGNGNYTDVWSGTYKFNQNAAYFDKVEVKGTQLHVTGWHAADEAAV
KPYTYIILYDATSNREIKRINYTPTARPDVKNVYKNLYNSDNSGFDVNFNLAGLNLTGHK
IQLIARFSNDRAGNGNYTDVWSSLYKFNSMQRVLRA
>Lead, cadmium, zinc and mercury transporting ATPase (EC 3.6.3.3) (EC 3.6.3.5); Copper-translocating P-type ATPase (EC 3.6.3.4)
MRHLTKLWTILGIGLVALILQYGLNLQLYAQILVTVVGAVIALQMFIEMIGTLKSGSYGVD
LLAITAVVATLAVSEYWAAMVILVMLVGGDSLEDYAAKKANSELKALLDNSPQIAHKLVD
DDVVDLAVDEVLVGDKLLVRPGELVPVDSHIVKGESFFDESSLTGESRPLNKRVGDSL
MSGSLNGDASVTIQVDKLAIDSQYQQLVKLVKEAEKTPSKFVRLADRYAVPFTLAAYVI
AGIAWYLSKDPVRFAEVLVVASPCPLILAAPVAMVSGMSRASRNGIVVKTGDVLEKLAR
AKTGAFDKTGTITSGQLSVDKVVTLAGVKEEDLLAMIAGVETESNHILARSIVAYADDHG
IDGVATDSLSEVTGKGVSAMVAGQKVLVGKLSFVAPDSKIDKPQQTAIYLSVAGKLWG
YVLFIDHIRPEAAETMTTLQKLGVSHLMMLTGDQKAIATKVANQTGLTDVQADLLPQDK
IAALKAVPHSDHPVFMVGDGVNDAPSLATADVGIAMGAHGSSAASESADVVILKDDLS
KVAKAVIISQDTMRIARQSVLIGIFICLGLMLIASSGLIPAFIGAMFQEVVDTVSILWALKAR
YSHEQQVDVTKKHHYKAKLS
>hypothetical protein
MLVKDWVQDRVDYAVEQNTIAVTDRVTTEVTATMQEKFAQQLIKKAISTEQDYDSLRA
DLIDFFDENFADKMLTKYYK
>hypothetical protein
MFDIEMQTIQPKYLIQRILYYHSTMITERLYPRESYGEIPKTYVIFICLFDWYRLGNSFYE
VNLVPNGVN
>hypothetical protein

MKHYKNRRRQFLEEWEQSDISNDYIFSKVMAQPQLCSQLIKAPYPS
>Hypothetical protein DUF194, DegV family
MTATLKLLTDSSIQLTAEEIEKYHITIVPLSVEVDGKSYVDGEDISRQELVEHLRQGNIPK
TSQPAMGRFVEAYNELGKDGSEVLAIMMSDVLSGTYETAKTAATMTDTKVTVINSKST
DRGLAFQVLAAAKDIQAGKSLEEIIAHCQDIHKRTTIDVLIDNLDCLVAGGRVSKMAGML
TKLINLKIIVRLRDNSLDVYSKGRSKKLFLKHAKEIAERHLNNPIKALGLSNVGTDADFLQ
KLNDLILPADAAFSGLRELTSPVIMTHTGLNAVGVITLASKEEPADNQ
>Promiscuous sugar phosphatase YidA, haloacid dehalogenase-like phosphatase family
MYKLVTCDMDETLLSDDRIITPKTVAAIKAASAQGVYFVPNTGRNFLTIQDNLATLGLKQ
QKGQYVISFNGGAIVENAGPKVLTTQALDFEIAKKLWQLGMDRGYCVHVYTVDKLYIW
NPSQTDAAYLTGRVAGWELPQEAIDILATTPIVKVLYEVLDEQKRLEVKATVNAEIDASL
NITFSSDRYVEFNDQAATKGAAVLKLGEMLGIKREEIIAIGDNGNDLSMIEACGLGVSMA
NGRDFLKAKAQYVTQNDNNHDGVAEVIEKFILNPAD
>hypothetical protein
MRQGPHHSAQKSTKVKPDCVSASMLLADKFFAILIKPFSRNIDNYPNLP
>Thioredoxin
MMNPVLEALEKKYQGQIKFAKVNVDTNQALAQSYQIMSIPALILFKDGVAKEKVVGYKP
QEAMEKYLDSKLAEK
>transposase
MYLPNKGAIIISEESEVNLMETQVTVKVKLNLANAEIASSFTNTMEQYRLACNYASEYIF
NHDFDMQQSRLNKVLYTNLRSLFMLKSQMAQSVIRTVIARYKTVKTQMKRNPYKYQDI
NTGEWYREIRDLTWLQKPITFNRPQVDLQRNRDWSCLSSGQLSINTLNGRVKVNPVC
HGFDQYFDGTWKFGLAKLLKTGGKWYLHISATKTVANFDNQTVKHVVGIDRGLRFLAT
TYDEQGQTAFFDGKAMMRKRAKYQKLRAKLQAKGTKSAKRRLKKLSGRENRWMTDV
NHRLSKTLVQKFGSNTLFVLEDLTGVSFERTDLPKVLRNQNKSWAFYQLAQFLTYKAH
LNNSEVIEVSAQYTSQRCPKCGTIKKSNRNHDLHEYYCTNCDYRSNDDRLGAMNIQLL
GTQYISGQEQPKFELTTNA
>CrcB protein
MKLTNLLSVAFFAFIGGNLRYLIGQALGLVGTLSVNLGGCFLLAFLTYYLFDRRHVPSW
LTTGLGTGLIGAFTTFSTFSLELGKLLLAHAWGVASAYFVLSSVGGLTLVWLGYLLAQR
LRGLNR
>CrcB protein
MMLVGTGAALGACLRYLVTTLIKRYSHFNFPWATLLINLSGAFCLGFLTGSQVVNPFLG
TGLMGGFTTFSTFNTEVFALGDGGLKTLAVLYLVISYLGGLALALAGLFLGKTC
>DegV family protein
MQIALLTDSAAGLEPQLLKKYQVKVIPVPLEINGVRYYEDRDLNLEQYRDMLADKLVGSI
AQVRVAELVKNLEQLVAQGYDTVLGICPAGGLSTLATHLQACSGQIAGLTIHAFDTQTI
GRLQGQFVEASAEMIARGDDIQTILRKLGRMRRQLQTFIALTDLKSLQKQGYVSNGKS
FLSKKILPLQTLGEINEQGQFEVFNTKRRAKNAYAELQSWIYTNYDALEDRMQITIVGDY
DNEAFRDHWVPNLRQDFPKAKLEFAQMGVATALTFGLNSILVSWGPRWQEFLIE
>Homoserine O-succinyltransferase (EC 2.3.1.46)
MASSSPLKIGILNIMHDKVDTKKRFDYVLTRAQMPVELKYFYPKDHYQDRPVPPEVSA
MAAPLELEEVAKLDAFIVTGAPIEKIAFSEVTYMAELHELFAFLEAKQICQLYVCFGAMA
AANYFYGIEKHLLAQKLFGIYPQVIVNDSPLLAGLTPGFLAPHARYAELSLAQIKQTPEL
RLLATTAKDELFAMESRDGLQAYLFSHLEYEGDALLKEYQRELAAYPDQAQTLARPQN
YFAQPASMQGPLFKWQEAQQAFFANWLQLVARHVRVSVTNK

>Cysteine synthase (EC 2.5.1.47)
MDKIYDSVTELIGHTPIVKLNNLVPEDAADIYVKLEFFNPAGSVKDRIALAMIEDAEKAGK
LTPGGTIIEPTSGNTGIGLAAVGAAKGYQVIIVMPETMSLERRKLMQGYGAKLILTPGAD
GMKGSIAKASQLAKDNGYFMPMQFENPANPTVHEETTGQEILAAFDGKLDAFVAGVG
TGGTLTGVGHALKVADKNIQVYALEPAESPLLKEGKGGKHKIQGISAGFIPEVLDREIYD
DVLEVTSDDAIAMARQVAYKEGFLPGISAGANIYGAIEVAKKLGKGKKVVTLAPDGGDR
YLSTDLFAGYEG
>Xaa-Pro dipeptidyl-peptidase (EC 3.4.14.11)
MRFNQFGKVTVSYPQALKELQTIRFLTGVELNKITPNALWLHFLAQASPLAHSQSAKKA
RLNSLLANEYQTVLEYTTTNTVNAASFYAVALQLLGFEVGEDFDLADVFTALKRLNLAY
HPTINSKEDLLLAWYDLLETVGKNGRTLLDNLASAGYFTPFYELAATQKPLFFNGKSLPI
FDTSKLISEVVYVESDLDSDHDGRADLLKVEIIRPLDTNNGLKVPALYTASPYNQGTND
QLGKKLTHKVDVKLATKPVTNTSYEEIAYHAPKVGPIEKRPVAGHAQFAEESLERELSY
TFNDYMLARGFAAVYAAGIGTKDSDGLRTCGSPAETASTVAVIEWLAGSRKAFTNKTD
NIEIKAWWCNGAVAMTGKSYLGTLATAAATSGTPGLKTIISEAAISNWYDYYRDGGLVV
APGGFPGEDADVLIAECFSRKKEAADYLQIKDKFEADLKQVTVDQDRTSGNYNRFWD
ARNYLKDLHNIKCDIVMVHGLNDWNVKLRNVFNLYKQTENLKLKRKLILHQGQHIYINN
FQSLDFSDMMNLWLSHKLYGLDNQAPEILPDVLVQDNTLEGTWHQFADWDGAETELH
SLHFDQHQLIAKPNPAAGPASFTDHLPAEAFKRYSQNYHQWQADLLARKAPLTTSSLL
FESQPLTTALYLQGTPRLKLRVAASQDHGLLSFMLVDYGQAKRLGKTPDLLLSKGLNA
GYRFREDNLVEFKLKKASPYQLITKGHINLQNRTSLWQNDELKPDTFYDLNVVLQPMF
YKLPAGHKLGLIVYSSDMEMTVHANEELNYSLDLAACRLDFDASRIAD
>Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA (TC 3.A.1.5.1)
MIKKGWLALVSTALVGTLFLAACGQTKSQQASNKQELNVATTAEMATLNTVKYSDTSS
LEALQNTFEGLYRFNTKSKPVLAGAQSVKVSADKKTYTFTLRKDAKWSNGDPVVASD
YVYAWRNMVDPKQASPNSQRFLPIKNGLKTATGELPVDQLGVKALSKYQLEVQLEAPI
SYLPELLTGAPFYPQDQKFVEKTGSAYGTSAAKTVFNGPFVVKNWTGTNLKWNYVKN
KDYWDKKAVSLKQVNVQVVKDTSTAAKLYQTGKLDYTALSTDYIRQYKSTPGYHTKSI
PLIGYMAFNVHRPDTANVHFRKAIAQAYDKQLLVKKVLHAGSPLNGIVPANFAFNSQTD
TDYRKDAGDMLTYNLKAARKQWALAKKELGKDKLSVEILTSDTTESKQVGEYLQSQLE
ANLKGLTVTVRSIPLKSRLQATTNYDFDIVYGTWQPDYADPVNFISDGGQYHLDDDYH
NANYRNLLNQAATTYATDTNKRWATLIKAEKQLIKDDAFTAPVYQGAMAYLLKDKVHG
LNISPYGTMLFYRDVTIK
>Na+/H+ antiporter
MAYLGTLSLILITTALAGQLCRRFGIPAVIGQLLVGIILGPAVLGLIKTNDFVEIFAEIGVIIL
MFMAGLESDLGLLKKYVKPATSVAVMGVIVPVILVYGFGRLFEFNNEEAIFLGVTFAATS
VSISVEVLKELHKLDSKEGTTILGAAVIDDILAVLILGALVSVFSDVAQAEGGHVQTNLWA
GLIYQVLFFGVIYLLFKWIAPLLIKVAEKMTIPSAVILVSLVICLGMAYLAEMVGLSGIVGS
FFAGVAVAQTPAKNEIDENIEPIGYAVFIPVFFISIGLNMTFKGFFDDLVFIIGLTILALLTK
WVGCGLGAKLLGMDFSSGNIIGSGMVSRGEMALIIAQVGYEAHLLSSKYYSGVIIVIILTT
VIAPFMLKGAINRQRKLEAN
>Amino acid transporter
MGIRERIFQKESLSRYLDKDKKLVKTLNARDLIGMGIGAVIGTGIFILPGTVAAVHSGPAI
MLSFVLAALVCSTVAMCYAEFSSALPIAGSAYSFGNIIFGEVVGWFLGWALILEYMLAV
AAVSTSWAAYFNSLLAGFGIHMPKALTANFDPANGTYVNLVAILIILFIAFISSRGVRSSIR

LNNIMVLVKLAVIVLFLLVGVFYVKPANWSPFMPFGASGIFKGASLVFFAYLGFDVVAAS
AAEVKEPKKNMPRGILGTLIICTVLYILVSMVLTGMVSYTKLDVGDPVAFALNVVNQNW
VAGIISLGALAGMFTMMFSMIFSSSRLIYSIGRDGLLPKFLGQVNEKTHTPEHSMTVVVII
VALMGGFVSLTQLINLVNIGTLIAFTFVSLGVIPLRKRTDIPNEGGYKVPFYPVLPLVSVF
LCLFMLSQLSLETWIASLIWFAIGAVIYFTYGIKHSLLNKD
>Zinc uptake regulation protein ZUR
MLDEALYLLKSHNYKVTKQRKSLLECLYNKCRHHYVDVVEIDAELRKLYPGMSHNTIYR
NMREFAQIGIIEMKTKASGACVKYQCDFSNLHHHHFICRKCGLVKEVQMCPMTVFESQ
LPGCQIEEHRFELHGLCADCANKQAKNS
>extracellular protein
MAYRTAPFITPFAIMSAVLALGATNVAVNTVTHTSEGKVRSSAVSSVSQASSSIVSEKS
KPDDERESKSSMTSSEEQLRSSSSKTTTSQGDDDTPTTTDDDAASSSADKVSSSSST
AKASSSSKATDQTASSNQTTDTTNTNDGGTTTTTNDNTTTNNQ
>FIG00743716: hypothetical protein
MVSILSMINSYLSYINLNVKIKNQVYTTLATIGNFYLLYVSYRFFVNGFIVRGLLFILAFLLI
AYFAYLNILYYFTKDKKSRYDISPWIEKTFHLTPKDPLTEAAKKKASQVGFVQTNGIFKN
EDFLPAEVKFDTQAKENLEKLVKSLVQVGYLSLDYDGQSDEVLFKLAKQGQAQRALG
QPKALPYFELVPGEKKLLIYGGLNQMEKMVLGEITTVGLMPALEAYKTYHLYLATAMIT
GGPVKVAGRVDVIEKVEPYRLEIQVAYRKRSQAQSKSSVSVSAQATPTYRRRRH
>Glutamate transport ATP-binding protein
MAMIEFSHVDKYYGKFQALKDINLEISQGEVVVIIGPSGSGKSTLIRTINGLEEIESGRLV
VNGLNLADKKTNLNELRRNVGMVFQHFNLYANKTVLENIMLAPRLVLKRDEAENKRLA
MELLETVGLAEKANSYPRQLSGGQKQRVAIARSLAMKPKVLLFDEPTSALDPEMIDDV
LNVMQEIAKKGDMTMVVVTHEMGFAREVADRVIFFDQGQVVEDDSTEDFFKHPKEER
ARQFLSKIITHK
>Glutamine ABC transporter, periplasmic glutamine-binding protein (TC 3.A.1.3.2)
MRVNFKKLALMAAGLVCLFATSACSKSSQKNLYQADQQTKTITWGIKSDTRLFGLMNI
KNGKSEGFDIDLAKALTKKIYGQDAQAKFETTTITTKIPLLKNGNIDAVIAAMTITPERQK
YVDFSKPYFNAGQSLLVKEGSPVKSIADLNQKGRVVLAVKGSVAVANVKKYAPKAKVL
EYDDYGQTFTALKSGQGDAMTTDNGILYGIAKENPGYAVVGGTFTNEPYGIAVNKGQK
QLLDAINKALSELKADGTYDRIATKWFGDVAGFDLKGAKGE
>Glutamate transport membrane-spanning protein
MITLFLDNWQTYLTGFGNTILASIIALVFSLIIGSAFAIMELLPSRLLRVIAHIYIEVFRNIPLL
VVAMFFYIVVPQYVVKLDGFTAGTIGLTLYTSSFIAETVRAGILAVDPGQMEGARANGLT
YWQAMRYIVLPQAYKIVIPPLGNQFINLIKNSSVLAFLAGFDLMYQAQTVASATFKTFSA
YFIVGFFYLVLTLPLSYYMRYLEKKLA
>Glutamate transport membrane-spanning protein
MTLLAHAYSAENINFLFQGLWVTIKVSLISIILSFIFGLIFGIIRYVQFPYLSKIVGFIIDIIRNL
PLLLIFFFTYFGVPELGLRVNPVSASIMALVIFESMMLAEIVRSGILAVDPGQMEGARAN
GLTYWQAMQHIILPQALTKMIPPLVSQFISLIKDTSLATIIVLPDLLYHAQIIYSQNTTYMLP
MYLMIAIMYFVVCYLLSLVARYLEKRLA
>amino acid transport protein
MKQSEEQIGLNRSLGLISSLSLVIGTVIGSGIFFKQAAVLDYAHSTNLALVAWLVGGLITM
ASGLTIAEIGSQMPHTGGLYVYLEKIYGKFWGFLAGWMQIIVYGPTLIAALGSYLATLLV
AFFGLPKTSTPLLAVIVIALIAAFNLLSNRYGAAFQVITTVGKLLPIAAIIIFGLLFGKANAFN
QVATTSGSVTLNGFGMAILATLFAYDGWVLIANMGGEIKNPQKLLPKAIVFGIGLVLIIYM

LVTAGIFRVVPASLIHRLGDQAAAHFATIAFGQVGGKLLNIGIIISIAGCINGKVMTFPRIM
YAMAKNDELPFSKQLAYLNKRMRTPMVATLTVCSIAALMAVTVDADRLSELCIFTIYLFY
VLAFFGLFKLRRLKTASPFRVPLYPFTPIVAILGGVFVLVSELNSDLTGVLASIFFVALGIP
VYLIKCKK
>Acetyltransferase, GNAT family
MDIVVNPKINKSELSRFFTEMEQQADAEEVASFKKGLKKSYVIAAYEDKKLIGFVKAVS
DYATLVYIQDLLVLPSYQHLRVNKTLMHHLLNYFGTIEQVVVSPLIKADAKFYRQLGFSA
GTEQGLETYFVDRRAR
>GTP-binding protein HflX
MNTIIAGIENNQANFDYTMEELAALASADNLTVVKEIRQKLDQPVAATYFGKGKAEEIKT
ASEIYDAQLLVVNDELTPTQIRNLETITNLTILDRTALILEIFASRARTKEAKLQVQIAKLQY
QLPRLQTGKKDQLDQQTAGNAGGGYTNRGAGETKLELNRRVIQKRISNLRKELKELTK
NHEVQRQQRQKNQLKTVALVGYTNAGKSTTMNQILALTGANQSKQVFEKDMLFATLD
TAVRKISLEDKKEFLLSDTVGFVSKLPHQLVKSFRATLEEAAQADLLIQVIDGADEHAKE
MIATTEATLKEIGVTDIPMIYAYNKADKSQLAYPVAEGNQITYSARQVESIELLLELIKKEL
FKDYEEHTYLIPYDQGRYLDLLNQQASIPKQDYLETGYLVKAEVAPKLAGQLAQFLSE
>Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA (TC 3.A.1.5.1)
MKLSKIVKLGAVAALSATFLVACGQKSSSSSSQQVLNWDETAELPTMDLSTATDVVSF
DTLNNTMEGLYRLGKDSKLEPGLATKTEVSKDGLTYTFTLRKDSKWSNGDKVTANDF
VYSWQRTVDPKTASQYAYLFSGIKNADDIMNGKADVSTLGVKADGDYKLVVTLDKKIP
YFKLLMGFPTFFPQDKAVVEKYGSKYGTQSKYMVYNGPFKLTGWTGSDLNWTLAKN
NDYWDKKDVKLAKINFSVNKSNTTSYNLYQSGKLDATPLSSQQAKQLKGKEGYTIRRA
ASTFYLQFNQDKDPYFKNAKIREAISLVINRKQFVDKVLGDGSTVAKGLVSPGLATYNG
KDFSDQAYTAAAVSYNVAKAKKLWAEGLKELGVSELSFSVMSDDSDKAKQSTEFLQS
AIESNLKGAKVTATNVPFKTRLARSESGDFDIVLSAWGADFSDPISFLELFTSDNPQNN
GAWKNAEFDKLIEAAGNEDAGNTGKRWNDMVDAEKILMNDVGISPIYHQATAWMVNP
KVKGIVYNSAGANYNFKEAYIK
>Oligopeptide transport system permease protein OppB (TC 3.A.1.5.1)
MRKYLLKRVFYMLLTLFLVATITFFLMKLMPGTPYSNQEKMTAAQIQIMNEQYGLNKPV
WQQYLIYIADMLHGDLGTSFQYSNQPVTYLISSRIGPSAQLGLQAMIFGLVFGIIIGSIAAI
NHNTWIDTTATVVSIIGKSVPNFVLAVLLQYYIALKLGWFPIADWGGFSYTILPMIALGVG
PLAETARFIRTGMVDVLNSDYIELARAKGLSKFGIVFHHALRNSLIPLVTLIGPYTVALMT
GSMVIENIFSIPGIGEQFVKSILTNDYPTIMGVTMLYSFGLVVIILITDIVYGIIDPRIRLTGN
GD
>Oligopeptide transport system permease protein OppC (TC 3.A.1.5.1)
MEEKIKLSPDSFQPLPKEEELDNEKIAAPSLTFTQDAWRRLKKNKGAVISLFIILIMVIIAF
GSTPFIDKSTLVKSAPQYANLPARVPGLDFINGLNGKLKVGGQWVDQYAQNGASGKY
FYLGTDYLGRSLAQRIIYGTKISLIVALVATFFDLTIGVTYGIISGWKGGRVDNVMQRIIEII
SSVPNLVIVVLMLLVMKPGMTSIIAAIAISSWTTMARMIRAETLQLKTQEYVLAARSLGES
PIKIAWKHLVPNLSSIIIIQTMYTIPTAIFFEAFLSFIGIGISAPETSLGVLLNDGQKNFQFLP
YQMWYPAIVLCVLMISFNLLGDGLRDAFDPKTRK
>Oligopeptide transport system permease protein OppB (TC 3.A.1.5.1)
MGERILDVKNLEIDFHTYAGEVKAIRNVSFHLDKGETLAIVGESGSGKSVTTKSLMGLLA
NNARVKGGEILFHGTDLLKKTEKEMQSVRGKDIAMIFQDPMTSLDPTMKIGMQIAEPLI
KHKHVEKKKALAQALEILKLVGIKDAEKRINDYPHQFSGGQRQRIVIAIALVCYPEILIADE

PTTALDVTIQAQILDLMKELQAKIETSIIFITHDLGVVAGMADRVAVMYAGRIIEYGTVDEI
FYNPQHPYTWGLLNSMPTLDSTKLEAIPGTPPDLLDPPKGDAFAARNPYAMKIDAEKQ
PPFFKVSETHYAATWLLHPDAPKITPPDEIVRRQKIFAEMQAKKVD
>Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)
MPEKAKKILEVKHLKQYFGVGKPDMVKAVDDVTFDIYEGETFGLVGESGSGKTTTGRS
IIHLYNPTDGQIIFDGKDVSKLHTKADKLAFAREMQMIFQDPYASLNPRMTIEDIIAEGLDI
HHLVKNKEERRARVEELLETVGLNKDHASRYPHEFSGGQRQRIGIARALAVEPKFIIAD
EPISALDVSIQAQVVNLLKDLQKERGLTYLFVAHDLSMVKYISDRIGVMHYGRLLEVGP
ADDVYFKPLHDYTASLISAVPVPDPEFERNRKQIDYDPSIEEDKEGRPRKLHEITPNHFV
RCAEDEIEMYKQKAAAYEQK
>transposase
MEMQVIVKVKLNLANAEIASSFTNTMEQYRLACNYVSEYIFNHDFDMQQSRLNKVLYT
NLRSLFMLKSQMAQSVIRTVIARYKTVKTQMKHNPYKYQDINTGEWYREIRDLTWLQK
PITFNRPQADLQRNRDWSYLNSGQLSINTLDGRVKVNPICHGFDQYFDGTWSFGLAKL
LKTGGKWYLHISATKTVADFDNQTVKHVVGIDRGLRFLATTYDEQGQTAFFDGKTVMR
KRAKYQKLRAKLQAKGTKSAKRRLKKLSGRENRWMTDVNHRLSKTLVQKFGSNTLFV
LEDLTGVSFERIDLPKALRNQNKSWAFYQLAQFLTYKARLNNSEVIEVSAQYTSQRCP
KCGTIKKSNRNHDLHEYHCSNCDYRSNDDRLGAMNIQLLGTQYISGQEQPKFELTTNA
>transposase
MVLTAGGIALTLGSLEPTFYPFASLCCVFTEQKCYNY
>Phosphate regulon transcriptional regulatory protein PhoB (SphR)
MKILVVDDDKDIVELLSIYIKNEGYEVEKAYNGKEAITKLNTNPDIALMILDVMMPQMDGL
SVVREVRKDSQIPVLMLSAKTDDLDKIQGLIQGADDYVTKPFNPLEVMARVRSLLRRSQ
HEVKNDEPDQIEVGPLIIKKDSHEVTTVSGNSIQLTALEFGILHLLASHPNRVFSADEIFE
RVWQQESVVSAKTVMVHVSHLRDKIEEATDGEKVIQTVWGVGYKIEAR
>Osmosensitive K+ channel histidine kinase KdpD (EC 2.7.3.-)
MVKSAKFKLTGREKSELFGEGIVTIILLILLNMSIYVILSQLIATNPGLKNGIFQIKMSLKFG
PNDFQIWSWGWLFVVLMAIIDCFIVYWRLMRRYSQMQVRHIIAELHYIAAGHFDHRIPF
TLKGQTQRIVASVNALVDSTINSMEEERRIERSKDELITNVSHDIRTPLTSIIGYLGLIENR
QYHNEEEMLKYTHTAYIKAVQMKALVDDLFEYTKVRQTDRPLNIVSVDMGAMLEQLGA
SFELEAQKKGMVIHLDEPTTPLMIEADPELLARVFNNLLTNALKYGSDGKNIFINLTQISE
TEAEITIANDGAKIPTEALGQVFERFYRVEESRSRKTGGTGLGLAIAQGIIDLHHGSINVT
SNDKLTSFVMRLPLKHPKEEK
>D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
MEFFKGLKKTLVGLLAGVTLFATTAAGLGNLTTVSAASLPANSDSSLNLDAKAAIAVDA
TTGQVMYSKNAQEALPVASMSKLLTTYLVLKAIHSGKLSWNDKITPDEASQKVSQNTT
LSNVPLKADHSYTVKSLYEAMLIYSANGATMALAYKVGGSHQNFINMMRKQAKEFGIN
DASIYTANGLTNGEVGSAAYPGATKDDENKFSAQDMAIIAQKILTDYPEVLQTTSITKKE
FNNGTDKTMMENWNWMLKGLAKAYTALPVDGLKTGTSNTAGACFVATVHKDGHRIIT
VILGAPHSSDTDLSRFEQTQKLMSYVYNGFTYKTIKAQSDYKQTLPVHNGKDLTTKVVT
AKETHLWVKNTDTDAKITASVSGKKSLYEKGGLQAPLKKDQTVGTYNLKLKDQELYFL
DGTNALKVPAQTKVAVDKANIFVIGWRALTGLFK
>Phage infection protein
MRMIIEEFKNLKRHKILIATILGIMSLPFLYSLFFLKSVWDPYGSTSYLPVAVVNKDQATY
YQGKKFQVGHDLVSKLKDNDDLEWHFVSEKEANDGLSHKKYYMVITIPKNFSKNATTV
LNRHPHKMNLTYKTNDSLNYIGKVISETGVKQVNSEIRESVSTAYADTMFKVVKKVGK

GFSEAASGSKQLADGSKVLNDGLNTYTAGVGQVNNGVIQLAAGVVPLSDGVQKLTAG
STALATGIKTYTAGVSQVNNGVQTLNSSTGTLATGVNALAAGSNSLKDGIGQYTAGVS
QVNAGVGELQAKTPTLANGVGKLAAGSKEFNTKLAQYTNGVDKLNAGLSQVTANDLT
SLKQLPAAAAALYQLNYIMSASIDNGLKQIDQQSLNALLTGVSQVDTEKLATSLTVMNN
LMGQMESLQTTANQMQSGLPAMEQAISKNQQAIAASLGKIATYNGTNQKLADAVTGD
VNAVLASSNSTLSEADKAKLQDAIKNTATIKTNSTNSLTEVGVVGNNLTQLKGTLDKTQ
ASLVDLNTKMADLREQASSLDTTYLKNQLATLAKMQGALDKDKLAKLQTTLGQLQQLQ
AAATKASSIARRLNSSVNNGQDLSSSDLAKLMQTAQVSSDATTALVAQNTAKLAGPST
EKFTTLVAGLAELSSGAKELSANNGALLAGSNQIATGLGTLNDSIPALTNGVTKLANGT
NQLAAKNAELNSGASQLAGGLGTLNANVPALTSGVSQLASGTSQLVANNGQLNSGAS
QLAGGLGTLNAQIPTMTNGVNQLASGTGQLAANSSALTDGSDKIYKGNKKLAKSMKQ
GANTIKKVKTGKDNAKMFAAPSNLVEKHYSYVPNYGYALAPYMLSVALFVGAMVFNLV
YPIRRLESEDETAFEWFMAKVTIGSIVAIGNALVEALLMIAAGLHPDHLGAMLANGILFSL
AMMFLIMFFSVAFQNPGRFLCMILLVIQLGASGGSFPIEITKGLGGFFQAVNPYLPMTYS
IYGFRQALTSGLGASQMITSYSVSLLFLVISLGLLYVAMVVVRPNVEYVEEVDINNYTKQ
N
>hypothetical protein
MSEKNRTDGKEIEVSSLGKPFVKGSHVSFTFHRHHFSGVVEKQLRNSAILKFDQEFAK
TTTAMDMKQLIVISYSKMQLLPAK
>branched-chain amino acid transport protein
MKKEEPRWLEVLRVALPLCLSYVPIGLACGILLNKAGFNAILTLLVSLMVFSGGAQFLIA
SMLVTNAPIHTIVLMLFFLELRYALLGSSLSQYLRGTSKRFLYVFALHMNDENYAVNYLK
FATDKEWRPGDALMVEHYTLLFWSVANLIGSIIGGFIHINLEIVDFALTAMFLYMIVMQLK
NRLTILISFVSACFAIVAMCLLKSTMGLIVATLAASLIGFAMEDYLVKHGRERGYLLSKIK
RRKKRPVAVPESDSEGE
>FIG00742167: hypothetical protein
MGYSTLDHVLLVVFSMIAAFIPRVVPLHIFATRKIPNWFNEWMKYVPVSLFTALVVKDIFI
NTKTYTFVGLDNVAKILAAIIVMAIAYKTRSMGLSVVFGLIAVALLAMVVPV
>hypothetical protein
MKRWVWWSLVAVITAVIFIGFGYTKKVATQKDYEQLVARGQANIDRADYKQAKINFQD
ALKKKQNDKPAQIYLKQIATYQAGLKLLKQKQYQAARLNFQMVAATDGGSSTLVRRSA
NLQTELKEVIKELKIFKTAYDKAYKLSSRYQYSASNTKLAVILGYGSIDQDYYRTIRQKAK
KLQGYNNYVLRSLGYTVEVDDDSAETKVAPKNDKAISPERLAQAKKELARAGVNTKKL
SPTKLKRLIIKADREHKSVVKVLKEK
>FIG00744642: hypothetical protein
MASKNKTHLLKASCFLASFLILAVLVHTNNVLVTNFDTAIQTLLKPLTNPINTRIVEFITNL
GGPVMATIFSLLVIICVYLKYSTRDAIWALVTLASGNLLVFLLKNLMQRPRPSDKLVAIGG
FSFPSGHTFGTTIFILFIVYLLIVHFQSRNVRLSLTCLAAIWALLIALSRIYLHVHYPSDVLA
SFLLAGFTWQVARMFQPKSSWHTQS
>hypothetical protein
MQLLPKRNYAEMLIILLVTILYSLCPADLKQALGISESLYMQVMGLLLLAVMLLDVAVMS
YWRS
>hypothetical protein
MLSWVVTFIMVFSLAITFLGWGLKTTAFNVNFAKQEVQTSGLNKELATDLMLALDSYRS
DLGIEPGKLLTKQQVNEDIILAVENFYAGKETILDTKKIAEQVATNMDNYLQTGELGLGL
NANSERYLALKSTVTTGIQEKLAWNLGQNANINDVQTEFMKAKQIVDTLFKVGLTLSIIL

AVILVLLQGLRPFNWAYYLGLAGCWATVPFYLLLWLLVKYQIYSSFIPYSARTIAIPLEKI
AVAVTASLQATVFAGVVVALGLLILGIIGMRFQQKVGHYGIHMAKRR
>Similar to ribosomal large subunit pseudouridine synthase D, Bacillus subtilis YjbO type
MEFTWQNEGKQQTLKKFLAARGVSHRLMGKIRRQKGAVLVDGKKVPFGTVLHQNEQ
VTLVLPEEKPVIKIAASPKPIEVLYEDANWLLVAKPAGLTSVPGPSNREDTLVNRVKYYL
ERNGRKNVVPHIITRLDRYTSGVVLLAKHSFAQGLIDKQVSNHEFDKRYYAIVAGQVTK
QHGLIEAPLKRKEGENFQIVSPDGRQAVTEYWLKANLAGGSLVEVKLHTGRTHQIRVH
FQSKGHPLLGDELYGGPLDMGIERQALHAYHLAFNDPFTNKRLVVEAPLPSDMKSVLA
KLKVD
>Mn-dependent transcriptional regulator MntR
MTPKKEDYLKIIFELGGKKKKISNKQIALSLNVAAGSVTEMVKKLVASHLVEHEPYAGISL
TTEGAKAAEDLVRRHRLWETFLVNKLNYKLQDVHTDAEVLEHITSNNLMNHLEEFLGN
PKRCPHGGVIPRQDGQYEEDSHKVLASVLDGQEAVVDRFIDNHDLLEYLADVKLNIGD
KIKVTKHTPFEGPIELVLLASGEKMSIGYKAAHYIFVR
>Cold shock protein CspA
MEQGTVKWFNADKGFGFITRQDGSDVFVHFSAIQGEGFKTLDEGQSVTFDVEEGQR
GPQAVNVVKA
>Multidrug resistance protein B
MEKQVKELPLIWLLIGNLVESTGMSFIWPLTTIYMHDYLGKSLTVAGVVLFIGSMAMIAG
NYTGGYLYDHKDARHYLLLALTLSLFATGFLVFESGWPYYPLMIILNNFAIGVSGTIVNS
LATTIKRYDSRYVFNLMYFMANVGVVLGTLGVGFVINISVRLIFIINFVMFAVFFVIASLFY
KVPKNDKHHLKQAKQAGMKTRRYAVYIITAVLIAYFVIQLGYAQWQSNLSVYMEKLGIS
LSKYSLLWTINGVIIVCLQPIISWIDEHYNVDIFKKVYLGIGLFVVAYSSLIVAHAYWMFIM
SMVIVTVGEVCLFPAIPTVVNKLSSASEKGKYQGQIGIAASLGHAIGPLFGGVVIEHYSY
GILFTLMALGVFMSEGWLILCRRLKAKIGD
>FIG00745123: hypothetical protein
MSQKTDGLIIKTVMLAGRIMIESGADMARVDDTLKRIARNAGIKDPKIFETTTGITMSIPD
RHEAQVEPINKRTIDLEKVARVNDASRAFQNKQLSLPNFYLKLIAIDVSAPFFAFYWQL
MAAALVSATLVIMYGASWRDFLPTALVGALGYTCFYFINSRYKIRFVSEFLAALLIGFSA
VMLVRFNLGYSVDRIIIGALMPLVPGVPLTNAVRDLLAGHILSGMARAMEAILTACALGL
GIAFILKFM
>Integral membrane protein
MEWIMHAAFSYLSSLGFGVIINVPRRALNACGFIGMLGWLTYLVIKNLEAGAMLANLGA
ALAIGLCSIFMARWQKMPTILFNIPSLVPLVPGGQAYQAVRYFALNDNTVALRYVVQVA
MIAGAIAMGFFLAELLGQAYYRMKS
>Cell envelope-associated transcriptional attenuator LytR-CpsA-Psr, subfamily F2 (as in PMID19099556)
MENSNENLRRKKELRRPRPRKKQGCFTLGLLIFVVLALFAGSYLKKHLATAKQTAADIY
QPTNTKKARNVDAILKQNKPVSILLLGTDTGALGRTFQGRTDTMILAVLNPKQKTMTVV
SIPRDTEVAVFGYEQYFPSKINAAYAYGGAKTALKTVQKYLNVPIDYYATINMGGLENLI
NAVGGLDIKPLITFSYDNQSFTKDQTTHMNGKRALAYVRMRDDDPLGDYGRQARQRQ
VLTKLAFKSSKLINLLDDGFLKTIKSQISTDLTFDDLIYLGTDYRVATHHMKSYSMQGST
EVIDGQDFEVVDTKNKQAITDILRAALDLEPATTGSTLASDNTIDTQAKD
>hypothetical protein
MTKHDLQTEMNQVELLIGKIMRIGVLIAAFVMIIGLGLLLVTGKSGYPGTFYPTTLQTIIAG
VCAFKAAAWMMLGIFLLILTPVLRVVISIYAFAKEKDHLYVWITSIVLVILVISFIIGHH

>Arginine/ornithine antiporter ArcD
MLTTALTMLLVGIFAGTMGAILGIGGGMIITPIVTLALGLDIKYAIGASIIAVIATSSGATIAY
LKDDVLNLRVAMFLEIATTIGAIVGALLTGYFNGNILYFLFGALLVFSSWNMYRKLRLGK
EVLKRANPDKLATKFKLNSSYYDKNLNQQIDYQVENIPLGLSIMFGAGLASGMLGIGSG
AFKVMAMDNAMKMPLKPSSATSNLMMGVTAAASATVYFFNGAIKPEIAVPLALGILGG
AALGSRIMPHLPAKVIRMIFIPILLYLGLQMILKACGVMI
>Transcriptional regulator, LysR family
MKNTNLYLMIYKKSILGGGGLNLKQLKYFLAVAEEKQVTAAAKRLHIAQPPLSYQLKELE
KDLDVKLFKRTAKGVILTEAGKMLTDYARQIVDLTQVTEEKVRKANQGQIGTISLGLISS
SGGFMASVALADFQARYPDVSYEIYEDNTYGIIEKLHQGQLDLAIVRTPFNQRGLKLNY
LTKEKMVVVSANDYSPQLPTSFTLTNLEEHPLIIYRRFKNIFTDSFEHYGIKPNVLMECD
DARTALYWAEMKLGLALVPESIAKLARNCRIIPVDYPAWETQMALIWLKDKELRPILKEF
ISVIKEGWK
>hypothetical protein
MDYYCQRCGRIIGEAERVCPACGVSQDRFKADEEQKACKKCRQEIPVNSNYCYHCGA
DQAQYYYQSEPLKKETPQVKKEEQPDQDNIFKQIFEGDYDDLNASYQRILKAEEEQRA
ELKKLKGKVTRANENEKPGLIPSTKLYLKDWFNINKRMGRADFWWSFVGMIGLTFVFS
LVLSVVVTGLRSVEPNFANFFLNFMVRVFIGFILIANMSALIRRFRDAEIPVIAIVMVFIPV
VGEIAAYILATKKQVVTNDDYTFIDPVREVKKRRKARKK
>Acetyltransferase (EC 2.3.1.-)
MEIKTKRLLLRPYVIDDLTDVYEYASQKKVAQAAGFTAVSDIRGALLFLEVLQNQGALAI
YSRLHQKVIGNISFYPSFASSKFELGYALNSDYWQRGYMTEALQALCKTLGDKEVLAY
VYPNNEASLRLLQKCNFEAQGSWRKGAQLFKYIG
>Acetolactate synthase, catabolic (EC 2.2.1.6)
MSEEKKRFGADAIVDSLANHDVEYVFGIPGAKIDRVFERLEHPVNPKAPKLIVTRHEQN
AAFMAAGIGRITGKPGVVIATSGPGASNLATGLVTATAEGDPVLAIAGQVQRTDLLRLT
HQSMDNAALFAPITKYSAEVQEPENISEVIANAYQEAESSKQGASFVSVPQDVTDSEV
KTEVIAPRIPPKLGPASPIESTLLSQRIKEARLPVLLLGMRASDPDTTAAIRNLIAATHLPV
VETFQGAGIIPRNMEADFFGRVGLFRNQPGDKLLKQSDLVIAIGYDPIEYEPRNWNQD
GKAHIVVIDSMRSEIDKNFQPEKELVGDIAQTLDFLLPYMKGYKLPEDSQAYLDSIRKEE
ALRDEPPALEEGQVRSHPLSIIQALQDRVTDDMTVTVDVGSFYIWMARHFKSYEPRHL
LFSNGMQTLGVALPWAISAALVRPNTQIVSVSGDGGFLFSAQDLETAVRLNLNIVHIIW
NDGNYDMVKFQEEMKYGQSAAVKFGPVDFVKYAESFGAKGLRVNKAADLEKVLDEA
FAIDGPVIVDVPVDYSDNQELGKQILDDQLH
>Alpha-acetolactate decarboxylase (EC 4.1.1.5)
MTKDGQKVYQHGTLALLVPGLFDGTMKLAELLAHGDTGIGTATGLNGEMIVLDGTPYL
VQANGQVKTLSAETMIPFATVHFDQPQVEFAAQGLSMQALGKQILAKYPYHNIFFAVKI
VGKFKHMQTRAVEGQVKPYPTLVEVANKQAIFEEADSQGTVVGYFAPELFQGMAAAG
YHLHYLNDAKKMGGHILGFEIDQAQVYLQPFATIEQHLPVDNQEFMEHQADLESMGQ
QIKSAEEAD
>Lysine-specific permease
MAEKNNVKRALKTRHLSMIALGGSIGTGLFVASGSAISTAGPGGALVAYIGIGLMVYFL
MTSLGEMATYLPVTGSFATYSRRFVDPALGFAMGWNYWFNWAITLAVDISTAAIVLRY
WFPHVPSWVFSLGALIIIFLINALSVASFGETEYWMSLIKVVTVLIFLVVGFLTIVGIMGGH
ATYLENFAYKKAPFVGGVPSILTVFVVAGFSFQGTELIGITAGESKDPAKSIPKAIKQVF
WRILLFYILAIFVIACLIPYTSPNLLGSDASDITISPFTLVFKRAGLASAASIMNAVILTSVVS

AANSGMYASTRMLFALGVSGDAPKIFGKVNGRGIPMPALLGTTLVGLLTFLSSIFGEQI
YSFLVSASGLTGFIAWVGIALAHYRFRRAFKAQGKDLSQLRYKAKWFPVGPLLALIICLL
VIVGQDIKSFANLDWQAISVTYMSVPLFIILYLYYKLRYKTKLIPLTEIDLTRHELEHEQAD
K
>Seryl-tRNA synthetase (EC 6.1.1.11)
MLDMKDMRNNVEEYKRRLATRGVEEKEIDDLLAEDKKRRELLVEAENLKKHRNEVSE
AIANAKRNGESAEEQIVAMREVGERIKQLDEKLAAVEERVNDMAARLPNMPNPTIPVG
PDEDANVELYKVGTPRKFDFEPKAHWDIGEDLGILDFDRAAKISGARFVFYKGLGAKLE
RAVYNFMLDEHAKEGYTEVIPPYIVKAKAMFGTGQFPKFKDQVYQVNGEDMTLVPTAE
VPLTNYYREEVIPTEKLPVYFTALTPSFRSEAGSAGRDTRGLIRMHQFNKVEMVKYTKP
EDSYAELEKMTKNAGNIMEKLGLPYHVITLSTGDMGFSAAMTHDLEVWMPAQDTYREI
SSCSNCEAFQARRARIQYRDENGKLQYVHTLNGSGLAVGRTVAAILENYQNEDGTVT
VPEALRPYLHGLEVITKQD
>YheO-like PAS domain
MESNEIRTYIESFIPLVKFIASILGPNSEVVLNDVTDLEHSVIHIENANITHRKIGDPASNLA
LRTIKAGKKENRDFIANYRGRAGNTNLRSSTYFIRYQGQIVAMICVNTDQSSLNNLTKQ
LELVNQAFNLPTHSLEGYDDPNQSQEPVEHFNQSNDSFLKNVIADKCQELNVSVKYLR
KKDKLAIIQILYSDGYFLLKDSVAHVAAELSTSIPTVYRYLNQVKKLDN
>Transcriptional regulator pfoR
MSIVLGVLMLFAVITLMSLFTFKAPYGRKSVSALSGAACATFLPQAFLSYAIGEVFHVEIA
RQIGDVMGSMGGLAAGALVPLAFGISPVFSMIVGISLMKFKLLPAFIAAYLISFLIKEIEKR
VTGGLDLIVVVLVAPLLTNLVATLIQPGVMGILTVIGGTITAATKGNPYVMGAVLGAIIPLV
GMTPLSSMVLTSLIGLVGTPMAIGALGCTGNSFLNFSFFRKMKFGDDATTIAVTIEPLTQ
LDIISANPIPIFVTNLVAGAINGIIVTMFGLVVRVTGMATPWAGLIVVFGMNPLVRTLIAVIL
IFINSTIWAYVGAYFFNHADIKIHTIDEIREAKDEKEAGHVVHGHAAV
>L-serine dehydratase, beta subunit (EC 4.3.1.17)
MSGNYRSVFDIIGPVMIGPSSSHTAGAVAIGQAANKLFGGIPDEVTVHYYGSFAQTHKG
HGTDYAITAGVMGFETDDPRVPNSQEIAKKMGLEIEFIEESGDSPINHPNTAVLDMVGR
DKKMRVSGCSIGGGAIEIREIMLRGMEIRPAGPLPIILYIDSDKNKEKELALTKEIERVAPF
ENRQVLQGKECNIYEFNIKNYMRDKTKKSLASKYDGIICL
>L-serine dehydratase, alpha subunit (EC 4.3.1.17)
MESVKELVEKSMSQNIPISELMIQEECATSGASRQEVWDKMKNNLETMRAAVKRGLT
GDGVYSKTGLTGGEAVLIKKYRQEHKSLSGDAVMSAVQAAISTNEVNAAMGVVCATP
TAGSAGTLPGILFMLEERLNLSEEQMIRFLFTAGGFGMIIANQACIAGASGGCQAEVGS
ASAMGAAAAVEAAGGNAEQSAQALAMAISNLLGLVCDPIAGLVELPCVKRNAIGAGNA
LISADMALAGCVSKIPADEVISAMDKVGRNLPESLRETGIGGLAGTPTGQKIKMQIFGQ
KV
>Phosphoglycerate mutase family 5
MRHGKTEFNLAKRFQGGAVDSPLLPESIINAQKVGSYLADLKFAGAFASPQERAVVTA
QNIVNQWEVAPKIKKIADLREMNFGKWDGDLVAAHQGEELFEQFLKYPSQFEGEKLG
GENYYQFTHRVQAGLKAIFEEGFKAEENVLVAAHALVISFAVKSLLGKSFEEIRAAGLVA
NTSVTVLETNDFNEFKLLDWNNTNFLI
>Small Subunit Ribosomal RNA; ssuRNA; SSU rRNA

>tRNA-Ile-GAT

>tRNA-Ala-TGC

>Large Subunit Ribosomal RNA; lsuRNA; LSU rRNA

>5S RNA

>tRNA-Val-TAC

>tRNA-Lys-TTT

>tRNA-Thr-TGT

>tRNA-Gly-GCC

>tRNA-Leu-TAA

>tRNA-Arg-ACG

>tRNA-Pro-TGG

>tRNA-Pro-CGG

>Transcriptional regulator CtsR
MQSRNISDIIEQYLKDILADSQQIEIKRSEVAELFNCVPSQINYVIKTRFTIPNGYLVQSKR
GGGGYIRIAKVNLIDDTDILDELTAYIGTSINASRGLQVVQSLYEEKMISRREANLILATIE
RSTLAFADKKLEDAVRARILVGILERLRCAD
>ATP-dependent Clp protease, ATP-binding subunit ClpC / Negative regulator of genetic competence clcC/mecB
MDELYTPSANRVLILAQEQANYFKHQAIGTEHLLLALSLEENGIANKVLRQNVITATDIRE
EIERLTGYGNLRRQAPDNYLPYSPKAKAVLEKAKNEASILNAAKVGTEHILLALLDDETI
VSSRILASLNVDVKRIKQSIYHQLGVSPVQVRKATRANNMKNQEKTPTLDELARDLTAM
AANQEVDPVVGRDQEIKRVIQVLSRRTKNNPVLLGEPGVGKTAVAEGLAQKIVLEQVP
DNLLTKRIMMLDMGSLVAGTKFRGEFEDRLKKILDEIYQDGNVILFIDELHTLIGAGGAE
GAIDASNILKPALARGEVQVIGATTLDEYQKYVESDAALERRFARVNVEEPTKKVTLEIL
TGLRPRYEKFHKVKLTDEALKAAVDYSSRYITNRFLPDKAIDLMDESAARVRIAHSEKK
SKLSESKLKLDELSQAKDEALLREDFEEAAKLRNQELKLRQKLAEMKRREEEKAKKLE
FNIKVTAQDVAGVVADWTGVPVTQMTKTESERLINLEKVLHERVVGQAEAVSAVAKAI
RRARSGLSDPGRPIGSFMFLGPTGVGKTELAKALAQAMFGDEDSMIRIDMSEYMEKY
STSRLVGSPPGYVGYDEGGQLTEKVRNKPYSVVLLDEVEKAHPDVFNILLQVLDDGYL
TDSKGRRVDFRNTIIIMTSNLGATALRDDKAVGFGVAKPSDEYAAMSAKIREALKKSFR
PEFLNRIDEIVIFHPLVKEELHEIVKKMAKQIVDRLAQKDVKLKITPAAIDVIAKAGFDPEY
GARPIRRALQNKVEDKLSDALLMGQIVQGATVTLGASKGKITLNVKEVSPVSTK
>DNA-directed RNA polymerase beta subunit (EC 2.7.7.6)
MAGHLVKYGKHRVRRSYSRIKEVLDLPNLIEIQTDSYKWFLDEGLREMFDDIMPIEDFQ
GNLSLEFVDYQLLEPKYTVDEAREHVANYSAPLHVTLRLINHETDEIKSQDVFFGDFPL
MTKQGTFIINGAERVIVSQLVRSPGVYFNSELDKNGRENFGTTVIPNRGAWLEYETDAK

NVAYVRIDRTRKIPLTELVRALGYESDSDIVQILGTESDSLNLTLEKDVHKNTEDSRVEE
ALKDIYERLRPGEPKTADSSRSLLTARFFDPKRYDLAPVGRYKINKKLDLKTRLLNQRL
AETLVDPETGEIIVAKDEVIDKRVMEKLAPYLARDDFKMVTFTPSSEGVVQEDMTLQIIK
VYSQVDPERVVNMIGNGNISIDFKHITPADILASMNYFFNLQEGIGSTDDIDHLGNRRIR
SVGELLQNQFRIGLSRMERVVRERMSIQDTATVTPQQLINIRPVVASIKEFFGSSQLSQ
FMDQTNPLGELTYKRRLSALGPGGLTRDRAGYEVRDVHYTHYGRMCPIETPEGPNIG
LINSLSSYARVNRYGFIESPYRRVSWETHKVTDKIDYLTADEEDNYIVAQANSPLNDDG
SFVDDVVMARHKDENIEVSIDKVDYMDVSPKQVVAVATACIPFLENDDSNRALMGANM
QRQAVPLIQPHAPLVGTGIEYKAAHDSGDALISEHAGVVEYVDAREVRVRRDDGSLDK
YKLMKFRRSNGGKNYNQTPIVRVGDRVDADEVLADGPSMEQGELALGQNPLIAFMTW
DGYNFEDAIAINERLVKEDVYTSIHIEEHESEARDTKLGPEEITREIPNVGEDALKNLDEF
GIIRIGAEVKDGDILVGKVTPKGMTELSAEERLLHAIFGEKAREVRDTSLRVPHGGGGIV
QDVKIFTREAGDELSPGVNQMVRVYIAQKRKLQVGDKMAGRHGNKGTVSVVIPEEDM
PFMPDGTPIDIMLSPMGVPSRMNIGQVLDLHLGMAARKLGIHVASPVFDGARDEDIWG
ALKEAGLPSDGKTVLYDGRTGEAFDNRIAVGVMYYMKLAHMVDDKIHARSIGPYSLVT
QQPLGGKAQFGGQRFGEMEVWALEAYGAAYTLQEILTYKSDDVVGRVKTYEAIVKGE
QIPQPGVPESFRVLVKELQALGLDMKVLDADKNEIELRDKEDEDEEDGNKEKVLNLSKI
EKAKAEETKTAEAQAQNEK
>DNA-directed RNA polymerase beta' subunit (EC 2.7.7.6)
MIDVNKFDSMQIGLASSDKIRSWSYGEVRKPETINYRTLKPEKDGLFDERIFGPTKDWE
CACGRYKRIRYKGLICDKCHVEVTRSKVRRERMGHIELAAPVTHIWYFKGIPSRMGLVL
DMSPRALEEVIYFASYVVTDPGDTPLEKKQLMTEREYREKSAEYPGKFKAAMGAEAIR
TLLNDVDLEKECAELKEELKEATGQKRTRAVRRLDILEAFLKSGNDPAWMVMDAVPVI
PPDLRPMVQLEGGRFATSDLNDLYRRVINRNNRLQRLLDLHAPGIIVQNEKRMLQEAV
DALVDNGRRGRPVVGPGNRPLKSLSHMLKGKQGRFRQNLLGKRVDYSGRSVIDVGP
KLKLSQMGIPHEMALELFKPFMMKELVQREMASNIKNAKRKIDRRDDDIWDVLEEVIKE
HPVLLNRAPTLHRLGIQAFEPTLVDGKAMRLHPLACEAYNADFDGDQMAIHVPLSAEA
QAEARLLMLAAHHILAPKDGKPIISPSQDMTIGNYYITIEEANREGEGMIFKDVNEVRTA
YQNNYVHLHTRIGLQTDSLVKAGKPFTQWQKDRILVTTAGKAIFNEILPADFPYLNEPTS
ENLTGKVPDKYFLEPGQDIHEYLAQAPLVGPFKSGFLSDIIAQVYKDYKVTATAELLDR
MKDLGYYESTKSGLTVGIADVTDLKEKPAIVEQAHKEVANVAKQFRRGLITEDERYNNV
IAIWNKAKDDIQNKLVENMDPSNPIQMMSDSGARGNISNFTQLAGMRGLMAAPNGKT
MELPVIANFREGLSVLEMFISTHGARKGMTDTALKTADSGYLTRRLVDVAQDVIIREKD
CGTDRGLDVTAIREGNEMIEPLYDRILGRYTMKTVKNPATGEVIVPANVMIDEEMAQAV
VDAGVEMVTIRSAFTCNTRHGVCEHCYGRNMATGDEVEVGEAVGTVAAQSIGEPGTQ
LTMRTFHTGGVAGDDITQGLPRVQEIFEARNPKGRAQITEVTGIVETIEENPAERTKEVT
VKGETDTRVYSLPFTSVLKVKEGDAVHRGDALTVGSIDPKELILVRDVLSTENYILREVQ
KVYRMQGVEISDKHVEVMARQMLRKVRIMDPGDTDMLPGTLMDISDFKDRNTDTLISG
GIPATARPVLLGITKASLETNSFLSAASFQETTRVLTDAAIRGKNDPLVGLKENVIIGKIIP
AGTGMKKYRDIQPEEVGTVSKSVYSISDVEKSLKEKVESTSSED
>Leader peptidase (Prepilin peptidase) (EC 3.4.23.43) / N-methyltransferase (EC 2.1.1.-
)
MYLFAVFCYATVWASFSSCYAYRYAHNESIFYPGSYCDYCHHPLNFWQLIPILGFCLQ
RGRCYYCHHKISLHSTLVELTFGLYMLSLFASKAPYLWASLSIFCAWSLLLALSDYLTQ
SVPAFELYLGGLVLLTQKLSWPPLAPGCSLCFLVILVGATYLQLLGSGDLIYLGFLWASF
GLLFLLATTCLASCLALCYFSLKKDRPASLAFIPFLTTASFILLYFN

>SSU ribosomal protein S12p (S23e)
MPTINQLVRKGRKSKGSKSNAPALNFGYNSFKKVQTNNSAPQKRGVATRVGTMTPKK
PNSALRKYARVRLSNLIEVTAYIPGIGHNLQEHSVVLIRGGRVKDLPGVRYHIVRGALDT
AGVDGRMQSRSKYGTKKPKK
>SSU ribosomal protein S7p (S5e)
MPRKGAVAKREVLPDPIYNSKLVTRLINRLMLDGKRGTASKILYQAFDIIKEQTGNDPLE
VFEEAMKNIMPVLEVKARRVGGSNYQVPIEVRPDRRTTLGLRWLVNYSRLRGEHTMP
ERLAKEIMDAANNTGASVKKREDTHKMADANRAFAHYRW
>Translation elongation factor G
MANKREFPLEKTRNIGIVAHIDAGKTTTTERILYYTGKIHKIGETHDGASQMDWMEQEQ
ERGITITSAATTAQWKDHRINIIDTPGHVDFTVEVERSLRVLDGAVVVLDAQSGVEPQT
ENVWRQATTYGVPRIVFINKMDKIGANFDYSVSTIKDRLQANPLPVQMPIGAEDQFEGV
IDLIEMKADLYDEDELGSKWDVVDVPDEYKEEAESRRDEMIEALADVNEDIMEKYLDG
QEISNDEIRAAIRQATLNLDVYPVFAGSAFKNKGVQMMLDGVNDYLPSPVDVKPYKAT
NPDTDEEVDLIADDNKPFAGLAFKIATDPFVGRLTFFRVYTGTLEAGSYVLNATKGKRE
RVGRLLQMHSNNRTEIPEVFSGDIAAAIGLKNTTTGDSLTDPKTPLLLESMEFPEPVIQV
SVEPKSKADQDKMDVALQKLAEEDPTFKAETNPETGETLIAGMGELHLDIIVDRMKREF
NVEATVGAPQVAYRETFTKQVSAQGKFVRQSGGKGQYGDVWVEFTPNEEGKGFEFE
DAIVGGVVPREYIPSVEQGLKEAMQNGVLAGYPLIDVKAKLYDGSYHDVDSSEAAFKV
AASLALRNAAPKAGAVILEPVMKVEIIAPEDNLGDVMGHVTARRGRVEGMEARGNAQV
VNAFVPLAEMFGYATTLRSATQGRGTFTMTMDHYEPVPKSIQDEIIKKNGGNAE
>Nicotinamidase (EC 3.5.1.19)
MPKALLIIDYTNDFVATNGALTCGQVAQDLAPTIIDMADSFLAAGDYVILPTDCHHLNDP
YHPENKLFPPHNIEGTWGQEFYGPLNDWYQTHQDSENVYAFAKNRYSSFQNTNLDN
YLRSRQITDLTLSGVCTDICVLHTAVSAYNLNYQLTIPKKAVASFDANGHDWALNHFKT
CLGAKLI
>SSU ribosomal protein S10p (S20e)
MAKQKIRIRLKAYEHRILDQSADKIVETAKRTGAQISGPIPLPTERTLYTVIRSPHKYKDS
REQFEMRTHKRLIDIVNPTPKTVDSLMKLDLPSGVDIEIKL
>LSU ribosomal protein L3p (L3e)
MATKGILGKKVGMTQVFTENGELVPVTVVEVGTNVVLQVKTVENDGYEAIQLGFDDLR
EVLSNKPAKGHAAKANTAPKRFVREIRDVELGEYKVGDEVKADLFEAGDIVDVTGTSK
GHGFQGAIKFHGQSRGPMAHGSRYHRRPGSMGVIINRVMKGKLLPGRMGGNRVTIQ
NLEIVKADVENGVLMIKGNVPGANKSLVVVKSAVK
>LSU ribosomal protein L4p (L1e)
MPTVALFKQDGTQNGEVSLNDAVFGIEPNNNVVFDAVLMQRASLRQGTHAVKNRSAV
RGGGKKPWRQKGTGRARQGSIRSPQWVGGGTVFGPTPRSYAYKLPRKVRRLAIKSV
LSQKVLDESLVVVDALNFDAPKTKAFAEVLNNLDVKSKVLVVLEDDNKQAALAARNLA
NVTVIPAKGLNVLDVVNNDKLVITQGALSQVEEVLA
>LSU ribosomal protein L23p (L23Ae)
METRDVILRPVITEASMADMDNKRYTFDVDTRATKTQVKLAIEDIFDVKVAKVNIMNVK
GKKKRMGRYEGFTKKRRKAIVTLTAESKEIKFFED
>LSU ribosomal protein L2p (L8e)
MGIKKFKPTSNGRRNMTGSDFAEITKTSPEKSLLVSLNKTAGRNSYGRITVRHRGGGH
KRQYRLIDFKRVKDDVPAIVKAIEYDPNRSANIALIVYADGVKSYILAPKGLEVGQQVQS
GKNADIKVGNTLALADIPVGTVIHNIELKPGKGGQLVRSAGTSAQLLGKEGKYAIVRLAS

GETRMILVTCRATIGAVGNEQHELIKVGKAGRSRWAGKRPQSRGSVMNPNDHPHGG
GEGKAPVGRPSPMSPWGKKTTGLKTRSKKAKSNKFIIRKRKK
>SSU ribosomal protein S19p (S15e)
MSRSLKKGPFVDEHLMKKVEAQADKEKKSVIKTWSRRSTIFPSFIGYTIAVYDGRKHVP
VYIQEDMVGHKLGEFVPTRTFHGHVADDKKTGVRK
>LSU ribosomal protein L22p (L17e)
MAEQVTSAKATAKTVRIPARKARLVIDLIRGKSVADALGILKFTPRSGAQLIEKVLKSAIA
NAENNFDLDVEDLYVSEAFVNEGPTLKRFRPRAKGSASPINKRTSHITVVVSVK
>SSU ribosomal protein S3p (S3e)
MGQKVNPTGLRVGIIRDWDAKWYAEKDFASNLHEDLHIRNYIETKLADASVSTVEIERA
AKRVNISIHTAKPGMVIGKGGSEVEKLREELNNLTGKRVHINIVEIKKPDLEAKLVGESIA
QQLENRVAFRRAMKQAIQRTMRAGAKGIKVQVAGRLNGADMSRVERFSEGTVPLHTL
RADIDYAWVEAATTYGKLGVKVWIYRGEVLPTKKEK
>LSU ribosomal protein L16p (L10e)
MLVPKRTKHRREFRGKMRGEAKGGKEVTFGKYGLQAVGSHWITNRQIEAARIAMTHY
MKRGGKVWIKIFPHKSYTAKAIGVRMGSGKGAPEGWVAPVKRGKVMFEIDGVSEEVA
REALRLASHKLPVKTKFVKREEVGGESNEG
>LSU ribosomal protein L29p (L35e)
MKANEILEKKDKQVRDEIKTLSSTELVDLEKQFKEELFNLRFQLATGQLENTARLNKVR
KNIARIKTALRQAELNK
>SSU ribosomal protein S17p (S11e)
MSEGRNQRKVYQGRVVSDKMDKTITVVVETYKNHKVYGKRVKYSKKFKAHDENNQA
KVGDIVKIMETRPLSATKRFRLLEIVEKAVII
>LSU ribosomal protein L14p (L23e)
MIQQESRLKVADNSGAREILVIKILGGSRVKTANIGDIIVATVKQATPGGVVKKGDVVKA
VVVRTKSGAHRADGSYIKFDENAAVIIGEDKSPKGTRIFGPVARELRDGNFMKIVSLAP
EVL
>LSU ribosomal protein L24p (L26e)
MFVKKNDKVKVIAGKDKGKEGVIEKVFPAKDRVIVKGVNIVKKHQKPTNANPNGGIVEV
EAPIHVSNVMLIDPSNNEATRVGFKVEDGKKVRVSKKTGKTL
>LSU ribosomal protein L5p (L11e)
MVNRLKEKYDKEIVPSLMEKFNYSSIMQAPKVEKIVLNMGVGDAVSNAKNLDEAVDEL
TLISGQKPVITRAKKSIAGFRLREGMAIGAKVTLRGQRMYEFLDKLVSVSLPRVRDFHG
VSKRAFDGRGNYTLGVKEQLIFPEIDFDNVNKVRGLDIVIVTTANTDEESRELLTQLGM
PFAK
>SSU ribosomal protein S14p (S29e) @ SSU ribosomal protein S14p (S29e), zinc-dependent
MAKKSQIAKNKRPAKFSTQEYSRCERCGRPHSVYRKFKLCRVCLRDLAHKGQIPGMK
KASW
>SSU ribosomal protein S8p (S15Ae)
MVMTDPIADFLTRIRNANMAKHESVEAPASKIKRDIAEILKNEGFIRDVEYIDDDKQGVIR
VFLKYGKNNERVISGIRRISKPGLRSYVKADAVPKVLNGLGIAIISTSEGVVTDKVARAKK
IGGEVIAYVW
>LSU ribosomal protein L6p (L9e)
MSRIGYKEVDLPAGVEVKQDGNVVTVKGPKGELTREFSDKIKMNINGNVVTFERSAED
NKTKALHGTTRANFHNMVVGVTEGFTKELELRGVGYRAQMKGTTLVLNVGYSHPVEF

EKEDGIEFATPSATQVTVSGINKEIVGDVAARVRATRAPEPYKGKGIRYVGEFVRRKEG
KTGK
>LSU ribosomal protein L18p (L5e)
MISKPDKNKTRQKRHARVRGKISGTAECPRLNVYRSNKNIYAQVIDDVAGVTLVSASTL
DSEVSGNTKTEQAASVGAVVAKRAVEKGITEVVFDRGGYLYHGRVQALAEAARENGL
DF
>SSU ribosomal protein S5p (S2e)
MAFIDPAQLELEDRVVAINRITKVVKGGRRLRFAALVIVGDHNGHVGFGTGKAQEVPEA
IRKAVEAARKNLIEVPMVGTTLPHEVVGQYSGSRIMLKPAIAGSGVAAGGAVRAVMELA
GVSDVTSKSLGSNTPINVVRATMDGLANMKSAEAVAALRGVSVQHLAE
>LSU ribosomal protein L30p (L7e)
MANLKVTLVRSVIGRPQNQREIVKGLGLGKVNSSVVVPDNAAMRGVIRKINHLVDVELA
K
>LSU ribosomal protein L15p (L27Ae)
MKLHELHAAEGSRHVRNRVGRGTSSGNGKTAGRGQKGQKARSKVRVGFEGGQMPL
FRRMPKRGFKNINRKDYAIVNLETLNKFEDGAEVTPALLVESGLIKDEKDGIKVLGNGEL
KKKLTVKANKFSASAKAAIEAAGGQAEVI
>Preprotein translocase secY subunit (TC 3.A.5.1.1)
MLKTMKNALAVKEIRNKILFTLGVLIVFRLGTYITVPGINAQALSSVASSGLISILNTFSGG
GLTNYSILAMGVSPYITAQIVVQLLQMDIVPRFVEWSKQGEVGRRKLNQVTRYLTIVLAF
VQSIGITAGFNALSSLNLVTNPGIKTYISIGIILTGGTMFTTWLGDMITERGLGNGISMIIM
AGILARIPTGIYQIYTEQFGSASQIWQPILYVVILILVILIIVTFVTYVQQANYKIPIQYTRRLA
GASESSYLPLKVNVAGVIPVIFASSFIATPQTILLAFTANHSEDTWYQVMTNIFSTSTVAG
MTLYTVLIVLFTFFYAFVQVNPEKLSENLQKQGSYIPSVWPGLETQKYVSRLLIRLSSVG
SLFLGLVSLIPLVASAVWGLDESIGLGGTSLLIVVGVTLETMRQLEGMMMKREYVGFIR
E
>transport protein
MNNKHIMKGITIAAIASAIWGISGTTLQFISQNQNLPVAWYLSVRTLGAGTILLLISLLLYR
KQTFAIFKDKTALAWLLAYATLGLMANLLTFYLAIQTGNASTATVLQYLSPLFIVIGALVF
KRQTPMKTDLIAFIISLLGVVLVITKGDFTQLAISVESLLWGIGSGITAAFYIVLPRPLAKDY
SPIITLGWGTFIAGILFNLYNPVWVGGPKLTLPIILSVSTVIIFGTIIPFALLLYSARFAPSDVI
GIMDALQPVTTTILSVIFLKLSISVMEVIGICLVIIAIYVLQRGRRNLEAVNYREADF
>Adenylate kinase (EC 2.7.4.3)
MAMNLMLMGLPGAGKGTQAERIVDTYKIPHISTGDIFRAAMKNETPMGLEAKKYIDKG
ELVPDEVTNGIVKERLAQADTKDGFLLDGFPRNMAQAKALEEIGKELDRPLTGVINIHV
DPESLMERLTGRYICRACGATYHKVFNPTKVEGTCDRCGGHEFFQREDDKPETVKNR
LEVNIKMNTPLLDFYKEQNLLHEVNGNQEIDKVFADVKSVLDNL
>Translation initiation factor 1
MAKDDVIEIEGTIKETLPNAMFKVELENGHEILAHVSGKIRMHYIRILPGDKVTVEMSPY
DLTKGRITYRFK
>SSU ribosomal protein S13p (S18e)
MARIAGVDLPRNKRIVIGLTYIYGIGNTTAKKILAAAEVSEDVRVRDLTADQEDKIRAEVD
KVKVEGDLRREVSLNIKRLSEIGSYRGMRHRRHLPVRGQNTKNNARTRKGPAISIAGK
KK
>SSU ribosomal protein S11p (S14e)

MAVKKSRKRRVKKNIESGVAHIHSTFNNTLVMITDVHGNAVAWSSAGSLGFKGSRKST
PFAAQMAAEAAAKGAMDHGMKTVEVAVKGPGSGREAAIRALQTTGLEVTSIKDVTPV
PHNGCRPPKRRRV
>DNA-directed RNA polymerase alpha subunit (EC 2.7.7.6)
MIEFEKPKIHKIEETKDYGKFIIEPLERGYGTTLGNSLRRILLSSLPGAAITDVQIDGVLHE
FSTVEGVLEDVTAIILNLKKVALKIDSEDAQALEINVVGPMDVTAGDINGSSDVEILNPDL
HIATVAEGASFHVRMNADKGRGYVSASDNKAKNDEMPIGVLAIDSIYTPIERVNYQVEK
TRVGQKNDYDKLTLDVWTNGSITPSEAISLSAKILTEHLTLFVDLTDEAKNAEIMIEKEET
HKEKMLEMTIEELDLSVRSYNCLKRAGINTVQELTNKSEADMMKVRNLGRKSLEEVKA
KLADLGLSLRHDD
>LSU ribosomal protein L17p
MAYRKLGRTSAHRKSMLRNLTTDLLVNGKIVTTETRAKEVRKFADKMITLGKKGDLAA
RRRAAAFMMDVVADVKEDGDKVVVQTALQKLFDDVAPRYAERNGGYTRILKMEQRR
GDAAKMVVLELV
>ATPase component of general energizing module of ECF transporters
MTTMIKVKDLTYRYNPTDEQTALNKVSLEIEAGQWVAIIGHNGSGKSTLAKALDGLIEPE
AGTIEVAGLKLTAKTVWEVRNHIGMVFQNPDNQFVGADVEGDVAFGMENRGLARSEM
VRRVKQALAAVNMTAYAKHEPVRLSGGQKQRVAIAGVLALQPDVIIFDESTSMLDPEG
RQDIIDLMSKLNQEGFTIISITHDIDEASLAQRVIVLDDGQIIADQKPADLFAKGQALIELGL
DVPYPEKLKHELSKLGLRVPTKYLDEEGMVAWLWQLLSQA
>ATPase component of general energizing module of ECF transporters
MAMAITFTGLNYIYQANTPFEYHALKDINLTIPDDSYTAIIGHTGSGKSTLLQHLNGLLKP
TSGSLEVNGQLITPQTKNKQLGALRKQVGFVFQFPEAQLFEETVLKDIAFGPQNFGKS
QAEALEIAKQMAQLVNLPASLWEKSPFELSGGQMRRVAIAGILAMTPKTLILDEPTAGL
DPKGRLEMMEMFYRLHQEEKLNIILVTHNMDDVANYADNVVVLEGGQVLKTGSPQEV
FAEPAWLKKHHLGLPKTTAFAQKLVERGFKLDKLPLKEKQLAEKLVALIGGTGHE
>Transmembrane component of general energizing module of ECF transporters
MSKFLLGRYINGDSLVHQMDARAKLALSIYFIAIIFLANNWATYGLLLILVLACVLLSKIKL
GFFLRGVAAMIWLIIFTVLLQVFFTTGGHVYWHWGILNLTQVGLVRGAYVFCRFVLIIFM
STLLTLTTSPLEISDGLESLMKPLKYLKVPVYEIALMLSIALRFVPTLLDETQKIMNAQRS
RGVNFGEGKLLDQIKAVIPLLIPLFVSAINRAEDLATAMEARGYRGGEGRTKYRQQVW
QLKDSLALVVFGLVTVGLVFLRTW
>tRNA pseudouridine synthase A (EC 4.2.1.70)
MTKRYKITMAYDGTNFAGFQIQPNQRTVQGVLEKAINKMSKLDEYITVHGSGRTDAGV
HALAQVAHFDFPHEISAEGMLRGLNSMLPLDCEITACEIVPDSFHSRFTTHGKRYMYR
VSRGRYTNPFKRFYTGHYKYPLDIDRIKEAMPDVEGTHDFSSFVASGSQVKDHVRTIY
EATVREDKANDEIIFEFYGNGFLYNQVRIMVATLLEIGNGRRDVHDFLRLYEVKDRSQA
RETAPASGLYLKEVYYD
>hypothetical protein
MKAKLGSSHQQELTAALAEALADVRQATFILVQTTLGDLKQVATYLQTNFPGAASLATS
GQSYLATGAGQTLVAVAFEAGVAVETGVLTKLATAPLRDLYTLDQKLQAIKPNQQNTIC
LEYCTNDEEVTVSSMNVALEKYGVTLLGGSVFDEQKCTGEVLVDGQVYQDACGFALIK
KLDGKINTYSENIYEPVSNAIRHIATKVDLKTKTLVELDNQPAAEVYSQETGLPAQQIIEN
VLNQPLGRLVGDNYYIFSMHELAANGALVNYKRVNKNDTVAILQLADYQTINRQTLEKI
KRDNPRPELILSFNCIFRHQLFQKEGYLPTYLRQMGTLGPYLGYVCGGEQYGKQHVN
QTMICAVFE

>Methyl-accepting chemotaxis protein
MLFFKKKPAVTATESSEQAAVPRKIAAKLLPLKEIGRFALEQKSKLQQEESITINGIQEIQ
DSFEIVEEKYSGISQSVAGFKSEFKGINEVSNRLETITEQLKQTALDSQAGIKSVDSSSQ
SVLTTIDDMQVNFEAFQKSLAEIKKQVEQINYVATQTNLLALNASIEAARSGEAGRGFA
VVANQVNELSQEIKEMVTSIDSSMGELEQKNAELNQSLVDTRNAVDLSHETVNDTQAT
IATITTVADDVTKESEQLIQTFETCDTEIGLITQNINDSSHYFTNVAEKISDIKTKVSQKGF
MFEDMTNVLEQIQPLISEIESGEE
>LSU ribosomal protein L13p (L13Ae)
MRTTYIAKPGEVERKWYVVDATDIPLGRLSAVVASVLRGKNKPTFTPNVDTGDYVIVVN
AEKVALTGKKETDKIYYHHSNHPGGLKARTAGDYRANDPEKLIALSVKGMLPKGSLGR
KQAKKLHVYRGAEHNHAAQQPEVLDITNLI
>SSU ribosomal protein S9p (S16e)
MAQVQYSGTGRRKNSVARVRLVPGTGKIVMNGKPAEEYIPFANLREVMVQPFAVTET
KGAYDVLVNVNGGGFSGQAGATRHGIARALLDVDPDFRGVLKRAGLLTRDARMKERK
KPGLKKARKASQFSKR
>immunity protein PlnI, membrane-bound protease CAAX family
MIEPRMQETSAWLKLTFALVAYVIYVVVAEVTANLGCWKILIHASLVLLIVMASAWALKT
GYLLVPRSQTKAQMKTGRYFLIFVILYFGYFFQILSGVYSALGLEVGKLVRVTVVALGAG
FFEEFLVRGLILSAFAQLLEGKRGQISISVFSSALIFGLLHFCNLTYQTFPATLQQVIYAAV
IGVALGVVRVVTNGLVFPVILHSLIDFKPDVLEASELMNWSMLTIVLVPLLVVASITLINW
EFKRNYLLK
>Amino acid permease
MDKNKKIGLVSLIMMIFSSIFGFANMPVAFLQMGYASIIWYIFAAIFFFLPVAMMLAEYGA
TFKEEEGGIYSWLTHTLGEKWAFIGTFIWLSSWIIWLLSISSKVYIPFSAMLFGKDLTQT
WAFGPFNATQVIGLLGILWMLLVTFCASRGVNVISKVSSIGGTFVSAMVVVFLVATGIAF
VASHGQMQEALTVHSLVKAPNPAFQSPVATMSFVVYAIFAYAGLESLGGMIDSIDQPE
KTFPKGLLISSLLIAAAYSLMILLWGVSTNWQRDLSSDSVNLGNITYVMMQHLGVYLGQ
TLGLKASTSALIGNIFIRFVGLGMFLAYVGSFFILIYSPIKSFILGSKHLWPEKMTKLNKYG
IPEFAMWMQALLVCILIFLISFGGGQAKTFYTILTDMSNVSTSFPYLFLVGAFPIFKKKAF
DQ
>Mobile element protein
MRGRLTYGYRRVQARLDQLGIHLAGATIRHLMSELGVQVEIYNHHRNGKYSSYHGRV
GSVADNVLQQQFNAKDPYQVIHTDVTQVRLANQQWAYISAMTDEASKEVLAFQMSSH
PNQELIMNTLNELLDHLPDDAQPIIHSDQGWHYQLKYYTKKLKDHHFIQSMSRKGNCH
DNAPIESFFHLYKTELLQGLPPCNNITELRKLSLTYIDQFNNERISLKTKGMTPVGYRNH
ALAV
>Mobile element protein
MDQIRADQAELPKAQRYKIGDLLSRINLAKATYHDERKRIANPDDKYCEVKKVIRRICE
>hypothetical protein
MKKQTKKQKQKLVEKKEIDQLREELARTKQELYNTKMDRDILKKSLALFGPSKPGKKR
K
>TnpV
MDKYIYDDKNGLWYELQGDYYIPCLILPAEKEQPIGLWGQRHLRYLKEYRRATYITLFT
SGRLNNYLADIDRQAQERMERLTEQMKRVQGITEQLKAENALEWTQRMNNIRACAKEI
VEKEIIFA
>Transporter

METLKTIWDFFQNEVLGMNWLNRLIETFLNACGLDTTGKIGGSVQFFIYDTVKIMVLLG
VLILIISYIQSYFPPERAKKILGRFHGIWANIIAALLGTVTPFCSCSSIPLFIGFTSAGLPLGV
TVSFLISSPMVDLGSLVLLMSIFGWKVAVLYVIFGLVIAVAGGTLIEKLHLDNQVEEYIRN
GHSIDVPQEELHFKDRMKFAWEQVVETAKKVAPYVLIGVGIGAFIHNWIPEEFIVKALGE
NNPFGVIIATIVGIPMYADIFGTIPIAEALLAKGALLGVVLSFMMGVTTLSLPSMIMLRKAV
KPKLLGIFIGICAAGIIIVGYLFNAIQYLIV
>Redox-active disulfide protein 2
MALFNFGKKKEEEKKTPACACSCGCAASEAESITNDCCPDLKEGEICCIKVLAGCKSC
HEQFEYTKQAVKDMGLSVEVEYITDMQKVMEYGVMSMPALVVNEKVVAMGKVLKAA
DVVALLHKLGF
>Arsenate reductase (EC 1.20.4.1)
MDKKKVAFICVHNSCRSQIAEALGKHLAGDKFDFYSAGTETKSQINQDAVRLMKQLYGI
DMEQSQYSKTIDKIPTPDIAISMGCDVGCPYIGRAFDDNWGLPDPTGKSDECFIQVIERI
SQKIKDLKLEK
>hypothetical protein
MECVSALHHNKRKTKYHIHLIFSERRLLPEPDVKVASRSVFFDETGKRVRTKKEITGED
GQIRKGCTVIKKGEVYESHLFTTKDTRFKGEPFLREIKEVYTELINCHISDPEQHLKVFD
KNSVYLPAKKIGKNNPKEDEIKADNAARQEWNRTADMALLSGISEAKILEVKQTEIHEK
ASQSIKSKGWLPGLFRSIVNKAKDFLQNLIREHDMPPKPVLEIDMAEFRTMQKLMIKAQ
DKAKEIRHLQDTVLPKLKQQLADTKGIFKGKERKALTEQIQRTEKEIAEKLDKLPDVLKE
DGYPDVQAYMATYRKAEAVVEQFNRDLAAWEQQVREKQKPAKKEQAKPPERESVLK
RLRQLQAEGKQRNQPRQRKKSFDRDSR
>hypothetical protein
MEGGHALHTVPPAELLCKVREYLLRLVLRIDFGQGDNQFPCFDTFALCAAALKLLLAFP
CKVAPKGAVGGAVGGIEIFLPCMACNIRYSPFDIGQF
>hypothetical protein
MYGYRQEHHAEAGGMGKAARPQSGAQDHYPQGHAGTVSDGQSGKKSSEAR
>hypothetical protein
MKKETTFTAKQVGRRIKERRTELNITMPELGRRVGVNKSTIQRYEADGVDPKRTMVIN
GLAEALLTTPEWLTGLSDDKEYDTYTLCQRDIEEHIRKYLDTVSYTVKGEPHQQLLTTF
LGKMVDLYTVMTCYFADAMEEVDRVAEDKGLKESLGRYAIESGAIMEQVYRKKMEVPI
EDMKRFLDGILHIHDEGRTRMSMGALFGIVEEAEERLSEKENSVAP
>Integrase
MAKGSVRKKGKKWYGRFYIEDESGRKVQKEFAGTESKAETEAMLRKAIADYEEKQFV
GKAENITVGDMLDMWVEEELKPGNLSNGTVMSYQGTVNRIKQYPIGKRKLKTVTADHL
QAFIDFLSYGGTNPDGTTSKPMSKGYMLLFSAVLQNSFRFAVFPKKLITFNPMQYVKLR
GRKQETDIFSDSEEDTSSIPTITHEQFQKLEEFLKAKDNPALLPVQIAYYTGLRIGEVCG
LTWQDINLEEQYLTVRRSMRYNGTRHTTEVGTTKRSKVRTVDFCDTLAAILRAARTEQ
RKNRFRYGEFYHLNYYKEVKEKGRTYYEVYSLQRTEEVPEDYKEISFVCLRADGAYEA
PSTVGIMCRAAKKKVKGLEDFHFHTLRHTYTSNLLSGGAKPKDVQELLGHSDVSTTMN
IYAHSTREAKRTSARLLDKVVGGE
>hypothetical protein
MTKYSPEFKAQVVVEYMTDGLSSYTLEKKYKVSARQIRKWIQQYQLQGIKSFKQHHRK
RIFSAKFKWHVINYYQTHAESFAQVAAKFDLLPSQISI
>amino acid transport protein

MSNNYDALHQPFEVYKKHGWTQVITLISFLIILFGIIFTCVQPFLDGDYVTGFWTIIGPIFF
GGLAWLIYRQAEKKTK
>Galactofuranose transferase
MNNYLISITDKNKNDAGPKAKRDIEHFLAGKFHSVTLKYDLKKGWKSRLQKIYLAHVKV
PAFFKSHPDIDNLVLQYPLYSTYLLDVLLTNLKKYTQAKLYLVIHDLESLRLFKDKPNYV
QSEINFLNQADGLIVHNQKMISWLQANGCTSKLVDLEIFDYDNPQPLLTTDYYDASVCF
AGNLAKADFLTKLNLTKHQLILFGPNPSNNYPQNIAYQGQKTPEELPKFLQQSFGLVW
DGDSLSSCTGTFGEYMKYNNPHKVSLYLSSGLPVIIWKEAALADFVEKHHVGLTIADLN
DLDDLLDRLTPADYALIRQNVAAVANKMRQGYYIKRALKQLKI
>NADH:flavin oxidoreductase/NADH oxidase
MSELLKPAKVAGLDLKNRVVMSPMCMYEVAKEDGIATDFHFAHYGARALGGVSLIITE
STAVLPDGRITKNDLGLWNDQQAEKFSQLVSYLHGLGTKVGVQLNHAGRKAEDAKLK
WAPSAIAYNEIYGQPKEMSLAEIAEVKEAFVAAAKRAQAAGVDVIEIHGAHGYLLNQFL
SPKTNQRNDEYGGTLAKRYRLVDEIVRAVRAEFTGSLWLRISATDYAPEAENTMADW
KQVAGWLQAAGLDCLDVSTGGLLDVKPNIPLYDGYQVQFASALKAASDLSVAAVGLLD
NPGLCEYLVQTKQVDLILQARALLRNPNWVMKAATALHDHDYQAYNASYERARL
>FIG00742944: hypothetical protein
MMKSLSLPNFTKKKSKASVNEKLNVLRAGVLGSNDGILTVVGVLFSVAVATPNRFTIFIA
GLADLVACAFSMASGEYASVSTQKDTEVETVKAEQALLKNDFASELKVVQDYYVSQG
VSVQTAQEIAQNLLAKKPLATVLEVKYDLKLDDYQSPWSAAFSSLISASLGGILPLAAMS
LSPAAYMWPATIFAVCISVSLTGFLSAKLGDGLVKIAIIRNLLVGIITMIIHYSLGILL
>FIG00773410: hypothetical protein
MKKMSLAQKVNVMRAAVMGANDGILSVAGIVIGVAGASTSNFAIFISGIAGMLAGTVSM
AMGEYVSVNSQKDFQKKAIIQQQAALATNYDHEYDFVLQKYLAHGIAPRLAKRATEEM
MSKDPLTTTVRERFGFNPQEFTSPYEAALASMISFPTGSILPLAAITLFPEQIKVPATFLA
VIIALACTGYLAALLSKADRFHGVLRNVISGCITMLVTFGIGYLIGNY
>N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38)
MNTAILGVSGYSGSVLYKLLKEHPKVDKVNLYSRKPRRSLAEAVTAFKNQPESLQEYD
PQAIMAENDVLFLATPAGVSSQLASVFMEAKFPVIDLSGDLRLKKAASYEKWYQKPAA
ATKLLAQATYGLAEFEPVGTYIANPGCYATATLLALAPLVKEDLLQLDSIIVDAKSGLAG
AGKKLTESSHYAYINENAQAYKVNCHQHIPEIVQELQTWNDKLEAIQFTTTLIPITRGIMV
SAYAKVKPGFAQADLEAVFTKYYHDKKWVRFSGSDLPAIKDVIQSNYCDIGLSYNPVT
NTVLVVSVIDNLLKGASGQAVQNFNKLFGFEETLGLPVIPAWP

## 7.2 Supplementary data 2 – pVEcXLa sequence

>pVEcXLa
gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactgga
aagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggc
tcgtatgttgtgtggaattgtgagcggataacaatttcacacaggaaacagctatgaccatgattacgaattcatcgatatcta
gaggtacggatgccatatttaacccccgtagttataaatacggtggttataagatagtgttggatgccatactttaagtactgt
agttataactacataccgcaaccaccgtacttttacgtacgcttgtaaaaaacagggatgcaatcgccaaacattagtatta
actaagggtaaagcttttaacgtccttttataactatatacttaacaaccgtagttacaactacccttgtaagagtgcaagcaa
atattaccgagtatgtaaaaataagccgttttatcgtgatagacattttaggacaacacggctaaaccttggtaaatcttggta
tttttagagggcgtaactataattacacccttaccaccgtagttataacaacgcttgttacagtaatacttgtctagatataatat

acctaggctttgaattaggaccagctaccctatgaattagcttagtgaattgttgtaagactgtacgcccctcaacattaagca
tttaccgaggtgatgagtcatgacaacacgcgaaaggctagaggtgtttaaaattagcttggctttagttatcctagcattaa
gtcgcttaactctgagtatagcgcatgcttacaggataataaaaaaagctaaccacgattacaataaatcgtaattagctaa
tacacctaaagtcttaagcattcacgcttgagggtgaacaggttacgactgtttgccctcttttcaatatataataatacctgta
atagctttaaaagtaaagcgagatacaaactaaggggcgatgcaagatgccaaaagaaactaccgataaaagtaaaa
agcaagttgagtataataaacggtggtatgaaaaaaatagagaacacaagcgctatttaagcaagcgtagcacggcga
gaagtttttaagggtggcaactaaagaggatattcaagccctaattgaatttgctaacgaacaattaaaagaacgtttttaa
agtctagcatccaaactaggcttatttttatgtattgaatagaattagctcaagcaaagagctttttcagtgttgagagacgtatt
ttaagcctcgagctccacaaggtctttgtactaacctgtggttatgtataaaattgtagattttagggtaacaaaaaacaccgt
atttctacgatgtttttgcttaaatacttgtttttagttacagacaaacctgaagttaactatttatcaattcctgcaattcgtttacaa
aacggcaaatgtgaaatccgtcacatactgcgtgatgaacttgaattgccaaaggaagtataattttgttatcttctttataata
tttccccatagtaaaaataggaatcaaataatcatatccttttctgcaaattcagattaaagccatcgaaggttgaccacggta
tcatagatacattaaaaatgttttccggagcatttggctttccttccattctatgattgtttccataccgttgcgtatcactttcataat
ctgctaaaaatgatttaaagtcagacttacactcagtccaaaggctggaaaatgtttcagtatcattgtgaaatattgtatagc
ttggtatcatctcatatatccccaattcaccatcttgattgattgccgtcctaaactctgaatggcggtttacaatcattgcaa
tataataaagcattgcaggatatagtttcattccctttcctttatttgtgtgatatccactttaacggtcatgctgtatgtacaaggt
acacttgcaaagtagtggtcaaaatactcttttctgttccaactatttttatcaatttttcaaataccatctaagttccctctcaaatt
caagtttatcgctctaatgaacaaagatattataccacatttttgtgaatttttcaacttgcccacttcgactgcactcccgactta
ataacttcttgaacacttgccgaaaaagaactgcaggtcgacttaagctaaagccttaccttcagtcttactggtaacatcaa
cttgaacaccaacaaacttactaacccgaccatctggagtcttaattggagtaacagttaataaattccaaaatggagtacc
atccttccgataatttaataaccgaacactacaagcttcacccttcttaatagcatcccgaatcttttgaacttcctttggatcagt
accttcaccttgtaaaaaccgagcattgtgacctaaaacttcatctggaccataaccagtcatagcataaaaaccttcacta
gcataaactaatggacaatctggtaaagtagcatcagcaacaacaaaagtgtgccgtaaaccagccggccgctgcgtg
cacgggatcccattatcaaagtggcaactaacatactataaggatgctagaaaacggccttttaaaagagccttgtttaaa
cgtgatatacgcgcatctaagcaactttcttgtatccttgtaaactggtaaacccgaggtacttaaaagggtgtatacgtaaa
atagcaagcttacaaaaggcattgcactttgtgcagggcttaaaaaagcccccgactaaaagccgagagctagttttagg
tatgattacagaaaataaagaatacacttttattatagcatcccaaaataaaaacccaagctgtacagcctaggttttgtcat
atcatttagccacaaaatagagattattttatctagttgactagttgaccaacttgtcaacttttttagttgactagatgagacttatt
acagccaaagtaaaaaataaacccccttatttataggctttttcaaactctataatacgatgaaagtttgaaaaattattcccac
tcgatggttgccggtggcttggaagtgatgtcgtagacgacgcggttgacgtggtctacttcgttgacgattcgaactgagat
ggcttggagtacgtcccagtcaatcttagcgaagtcggcggtcattccgtcgatggaagtaacggcccggataccgatagt
gtagtcgtaagtccggccatcacccatgacccccacggagcggatacctggcaagacggtgaagtattgccagatttcctt
gtctaagccgtgcttggcaatttcttcgcgcaagatgtagtcggagtcacgcacgatttgaagtttttcatcggtaatttcaccg
ataacccggataccaagtccaggaccagggaatggttggcgccaaactaattcgtgtggcatccctagcttttcacctaatt
cacggacttcatccttgaataaggtgttaagtggttcgattaattcgaactgcatgtcttctggcaagcccccaacgttgtggtg
ggacttgatagtttgagcagtatctgttcctgattcgattacgtcagtgtaaagcgtaccttgggccaagaacttgatgccttta
agcttagtagcttcgtcatcgaaaacttggatgaactcatttccgataatcttgcgcttcttttctgggtcagagacgccagcaa
gtttgcctaagaagcggtccttagcgtcaaccttgatgatgtttaagccaaatttgccaccgaggctgtccataacttgttcgg
cttcacccttacggagcaacccgtggtcaacaaagatactggttaattggtcaccaatggccttgtgtaatagaaccccaa
caactgaggaatctaccccaccggaaagaccgaggagtacctttttgatcaccgaccgtttggcgaatcttggcaacttgc
atgtcgatgaagtcttccattgaccagttagcttgggcaccacacacatcgaaagcaaagtgacgtaagatgtcattaccgt
attcagtgtggcgaacttctgtgtggaattgcaacccataaagtttctttgctgggttagccatagcagcaattgggcagttag
cactagtagcactgacagtaaaaccagcaggtacttcagttactaagtccccgtgactcatccaaacagtttgagtggttgg
taagtccttaaacattacggcgtccttatcgactacttcaatattggccttaccgtattccttgttgtcagcactttcgacctttcca
cctaaatcgtgggccatcaattgcatcccgtagcaaatccctaagataggtagtcccaattcaaagattgctgggtcaacttt
aaatgcgtcttcgtcgtaaacactgttagggccacctgagaagatgatacccttagggttcatctgcttaatctcttgagctgta

agcttgtgagaaagaagttcagagtaaactcccatttcgcgaatacgccgagtaattaattggttatactgacttccaaagtc
caaaacgataatcttgtcgaaagcctgcatatcaacatttgccaagatattcacctcattaattaatttattatcctatattgtaa
cgacccaagggctttcggtcaatgattttcaccttaatccgatcgcccttcccaacagttgcgcagcctgaatggcgaatgct
agaatcgatacgattttgaagtggcaacagataaaaaaaagcagtttaaaattgttgctgaacttttaaaacaagcaaata
caatcattgtcgcaacagatagcgacagagaaggcgaaaacattgcctggtcgatcattcataaagcaaatgccttttcta
aagataaaacgtataaaagactatggatcaatagtttagaaaaagatgtgatccgtagcggttttcaaaatttgcaaccag
gaatgaattactatccctttatcaagaagcgcacaaaaagaaaaacgaaatgatacaccaatcagtgcaaaaaaagat
ataatgggagataagacggttcgtgttcgtgctgacttgcaccatatcataaaaatcgaaacagcaaagaatggcggaa
acgtaaaagaagttatggaaataagacttagaagcaaacttaagagtgtgttgatagtgcagtatcttaaaattttgtataat
aggaattgaagttaaattagatgctaaaaatttgtaattaagaaggagtgattacatgaacaaaaatataaaatattctcaa
aactttttaacgagtgaaaaagtactcaaccaaataataaaacaattgaatttaaaagaaaccgataccgtttacgaaatt
ggaacaggtaaagggcatttaacgacgaaactggctaaaataagtaaacaggtaacgtctattgaattagacagtcatct
attcaacttatcgtcagaaaaattaaaactgaatactcgtgtcactttaattcaccaagatattctacagtttcaattccctaaca
aacagaggtataaaattgttgggagtattccttaccatttaagcacacaaattattaaaaaagtggtttttgaaagccatgcgt
ctgacatctatctgattgttgaagaaggattctacaagcgtaccttggatattcaccgaacactagggttgctcttgcacactc
aagtctcgattcagcaattgcttaagctgccagcggaatgctttcatcctaaaccaaaagtaaacagtgtcttaataaaactt
acccgccataccacagatgttccagataaatattggaagctatatacgtactttgtttcaaaatgggtcaatcgagaatatcg
tcaactgtttactaaaaatcagtttcatcaagcaatgaaacacgccaaagtaaacaatttaagtaccgttacttatgagcaa
gtattgtctatttttaatagttatctattatttaacgggaggaaataattctatgagtcgcttttgtaaatttggaaagttacacgttac
taaagggaatgtagataaattattaggtatactactgacagcttccaaggagctaaagaggtccctagcgctcttatcatgg
ggaagctcggatcatatgcaagacaaaataaactcgcaacagcacttggagaaatgggacgaatcgagaaaaccctc
tttacgctggattacatatctaataaagccgtaaggagacgggttcaaaaaggtttaaataaaggagaagcaatcaatgc
attagctagaactatatttttttggacaacgtggagaatttagagaacgtgctctccaagaccagttacaaagagctagtgca
ctaaacataattattaacgctataagtgtgtggaacactgtatatatggaaaaagccgtagaagaattaaaagcaagagg
agaatttagagaagatttaatgccatatgcgtggccgttaggatgggaacatatcaattttcttggagaatacaaatttgaag
gattacatgacactgggcaaatgaatttacgtcctttacgtataaaagagccgttttattcttaatataacggctctttttatagaa
aaaatccttagcgtggtttttttccgaaatgctggcggtaccccaagaattagaaatgagtagatcaaattattcacgaatag
aatcaggaaaatcagatccaaccataaaaacactagaacaaattgcaaagttaactaactcaacgctagtagtggattta
atcccaaatgagccaacagaaccagagccagaaacagaatcagaacaagtaacattggatttagaaatggaagaag
aaaaaagcaatgacttcgtgtgaataatgcacgaaatcgttgcttattttttttaaaagcggtatactagatataacgaaaca
acgaactgaatagaaacgaaaaaagagccatgacacatttataaaatgtttgacgacattttataaatgcatagcccgat
aagattgccaaaccaacgcttatcagttagtcagatgaactcttccctcgtaagaagttatttaattaactttgtttgaagacgg
tatataaccgtactatcattatatagggaaatcagagagttttcaagtatctaagctactgaatttaagaattgttaagcaatca
atcggaaatcgtttgattgctttttttgtattcatttatagaaggtggagtttgtatgaatcatgatgaatgtaaaacttatataaaa
aatagtttattggagataagaaaattagcaaatatctatacactagaaacgtttaagaaagagttagaaaagagaaatatc
tacttagaaacaaaatcagataagtatttttcttcggagggggaagattatatatataagttaatagaaaataacaaaataat
ttattcgattagtggaaaaaaattgacttataaaggaaaaaaatctttttcaaaacatgcaatattgaaacagttgaatgaaa
aagcaaaccaagttaattaaacaacctattttataggatttataggaaaggagaacagctgaatgaatatccctttttgttgta
gaaactgtgcttcatgacggcttgttaaagtacaaatttaaaaatagtaaaattcgctcaatcactaccaagccaggtaaaa
gcaaaggggctatttttgcgtatcgctcaaaatcaagcatgattggcggtcgtggtgttgttctgacttccgaggaagcgattc
aagaaaatcaagatacatttacacattggacacccaacgtttatcgttatggaacgtatgcagacgaaaaccgttcataca
cgaaaggacattctgaaaacaatttaagacaaatcaatacctttctttattgattttgatattcacacggcaaaagaaactattt
cagcaagcgatattttaacaaccgctattgatttaggttttatgcctactatgattatcaaatctgataaaggttatcaagcatatt
ttgttttagaaacgccagtctatgtgacttcaaaatcagaatttaaatctgtcaaagcagccaaaataatttcgcaaaatatcc
gagaatattttggaaagtctttgccagttgatctaacgtgtaatcattttggtattgctcgcataccaagaacggacaatgtag
aatttttttgatcctaattaccgttattctttcaaagaatggcaagattggtctttcaaacaaacagataataagggctttactcgtt

caagtctaacggttttaagcggtacagaaggcaaaaaacaagtagatgaaccctggtttaatctcttattgcacgaaacga
aattttcaggagaaaagggtttaatagggcgtaataacgtcatgtttaccctctctttagcctactttagttcaggctattcaatc
gaaacgtgcgaatataatatgtttgagtttaataatcgattagatcaacccttagaagaaaaagaagtaatcaaaattgtta
gaagtgcctattcagaaaactatcaaggggctaatagggaatacattaccattctttgcaaagcttgggtatcaagtgattta
accagtaaagatttatttgtccgtcaagggtggtttaaattcaagaaaaaaagaagcgaacgtcaacgtgttcatttgtcag
aatggaaagaagatttaatggcttatattagcgaaaaaagcgatgtatacaagccttatttagtgacgaccaaaaaagag
attagagaagtgctaggcattcctgaacggacattagataaattgctgaaggtactgaaggcgaatcaggaaattttcttta
agattaaaccaggaagaaatggtggcattcaacttgctagtgttaaatcattgttgctatcgatcattaaagtaaaaaaaga
agaaaaagaaagctatataaaggcgctgacaaattcttttgacttagagcatacattcattcaagagactttaaacaagct
agcagaacgccctaaaacggacacacaactcgatttgtttagctatgatacaggctgaaaataaaacccgcactatgcc
attacatttatatctatgatacgtgtttgtttttctttgctgtttagcgaatgattagcagaaatatacagagtaagattttaattaatt
attagggggagaaggagagagtagcccgaaaacttttagttggcttggactgaacgaagtgagggaaaggctactaaa
acgtcgaggggcagtgagagcgaagcgaacacttgatttttttaattttctatcttttataggtcattagagtatacttatttgtcct
ataaactatttagcagcataatagatttattgaataggtcatttaagttgagcatattagaggaggaaaatcttggagaaatat
ttgaagaacccgattacatggattggattagttcttgtggttacgtggtttttaactaaaagtagtgaatttttgatttttggtgtgtgt
gtcttgttgttagtatttgctagtcaaagtgattaaatagaattccggatgagcattcatcaggcgggcaagaatgtgaataaa
ggccggataaaacttgtgcttatttttctttacggtctttaaaaaggccgtaatatccagctgaacggtctggttataggtacatt
gagcaactgactgaaatgcctcaaaatgttctttacgatgccattgggatatatcaacggtggtatatccagtgatttttttctcc
attttagcttccttagctcctgaaaatctcgataactcaaaaaatacgcccggtagtgatcttatttcattatggtgaaagttgga
acctcttacgtgccgatcaacgtctattttcgccaaaagttggcccagggcttcccggtatcaacagggacaccaggattt
atttattctgcgaagtgatcttccgtcacaggtatttattcggcgcaaagtgcgtcgggtgatgctgccaacttactgatttagtg
tatgatggtgtttttgaggtgctccagtggcttctgtttctatcagctgtccctcctgttcagctactgacggggtggtgcgtaacg
gcaaaagcaccgccggacatcagcgctagcggagtgtatactggcttactatgttggcactgatgagggtgtcagtgaag
tgcttcatgtggcaggagaaaaaaggctgcaccggtgcgtcagcagaatatgtgatacaggatatattccgcttcctcgctc
actgactcgctacgctcggtcgttcgactgcggcgagcggaaatggcttacgaacggggcggagatttcctggaagatgc
caggaagatacttaacagggaagtgagagggccgcggcaaagccgtttttccataggctccgcccccctgacaagcat
cacgaaatctgacgctcaaatcagtggtggcgaaacccgacaggactataaagataccaggcgtttccccctggcggct
ccctcgtgcgctctcctgttcctgcctttcggtttaccggtgtcattccgctgttatggccgcgtttgtctcattccacgcctgacac
tcagttccgggtaggcagttcgctccaagctggactgtatgcacgaaccccccgttcagtccgaccgctgcgccttatccgg
taactatcgtcttgagtccaacccggaaagacatgcaaaagcaccactggcagcagccactggtaattgatttagaggag
ttagtcttgaagtcatgcgccggttaaggctaaactgaaaggacaagttttggtgactgcgctcctccaagccagttacctcg
gttcaaagagttggtagctcagagaaccttcgaaaaaccgccctgcaaggcggtttttcgttttcagagcaagagattacg
cgcagaccaaaacgatctcaagaagatcatcttattaatcagataaaatatttctag

## 7.3 Supplementary data 3 - pVEcXLa-P$_{eft}$-*rnetB*

>pVEcXLa-P$_{eft}$-*rnetB*
gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactgga
aagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggc
tcgtatgttgtgtggaattgtgagcggataacaatttcacacaggaaacagctatgaccatgattacgaattcatcgatatcta
gaggtacggatgccatatttaacccccgtagttataaatacggtggttataagatagtgttggatgccatactttaagtactgt
agttataactacataccgcaaccaccgtacttttacgtacgcttgtaaaaaacagggatgcaatcgccaaacattagtatta
actaagggtaaagcttttaacgtccttttataactatatacttaacaaccgtagttacaactacccttgtaagagtgcaagcaa
atattaccgagtatgtaaaaataagccgttttatcgtgatagacattttaggacaacacggctaaaccttggtaaatcttggta
tttttagagggcgtaactataattacacccttaccaccgtagttataacaacgcttgttacagtaatacttgtctagatataatat

acctaggctttgaattaggaccagctaccctatgaattagcttagtgaattgttgtaagactgtacgcccctcaacattaagca
tttaccgaggtgatgagtcatgacaacacgcgaaaggctagaggtgtttaaaattagcttggctttagttatcctagcattaa
gtcgcttaactctgagtatagcgcatgcttacaggataataaaaaaagctaaccacgattacaataaatcgtaattagctaa
tacacctaaagtcttaagcattcacgcttgagggtgaacaggttacgactgtttgccctcttttcaatatataataatacctgta
atagctttaaaagtaaagcgagatacaaactaaggggcgatgcaagatgccaaaagaaactaccgataaaagtaaaa
agcaagttgagtataataaacggtggtatgaaaaaaatagagaacacaagcgctatttaagcaagcgtagcacggcga
gaagtttttaagggtggcaactaaagaggatattcaagccctaattgaatttgctaacgaacaattaaaagaacgtttttaa
agtctagcatccaaactaggcttatttttatgtattgaatagaattagctcaagcaaagagctttttcagtgttgagagacgtatt
ttaagcctcgagctccacaaggtctttgtactaacctgtggttatgtataaaattgtagattttagggtaacaaaaaacaccgt
atttctacgatgtttttgcttaaatacttgtttttagttacagacaaacctgaagttaactatttatcaattcctgcaattcgtttacaa
aacggcaaatgtgaaatccgtcacatactgcgtgatgaacttgaattgccaaaggaagtataattttgttatcttctttataata
tttccccatagtaaaaataggaatcaaataatcatatccttttctgcaaattcagattaaagccatcgaaggttgaccacggta
tcatagatacattaaaaatgttttccggagcatttggctttccttccattctatgattgtttccataccgttgcgtatcactttcataat
ctgctaaaaatgatttaaagtcagacttacactcagtccaaaggctggaaaatgtttcagtatcattgtgaaatattgtatagc
ttggtatcatctcatatatccccaattcaccatcttgattgattgccgtcctaaactctgaatggcggtttacaatcattgcaa
tataataaagcattgcaggatatagtttcattccctttcctttatttgtgtgatatccactttaacggtcatgctgtatgtacaaggt
acacttgcaaagtagtggtcaaaatactcttttctgttccaactatttttatcaatttttcaaataccatctaagttccctctcaaatt
caagtttatcgctctaatgaacaaagatattataccacatttttgtgaatttttcaacttgcccacttcgactgcactcccgactta
ataacttcttgaacacttgccgaaaaagaactgcaggtcgacttagtagtattcaattttatgatcttgccagttaatcttaaac
ataaattcgttatattcactagtactacttaacttattagtaccccgccattgagtagtttcccagtttaaaatgtaatcgttatcaa
accgttgatattcaacaataataacactttccttagcattctttggagcagttaaagctaaagccatatttggactaaaaccac
cactaattaaggtacttaaatcccgatcatcagtaaaattcttatcaccattgttgtataattgactcttcataaataattgattac
cgtaaatagcgtggtaactatcaatgttataaccatccttagtttcaacaaacttaatatcccaactagctaaattagcatcatc
cttccgttgaatagtccgaaaatctggttgttcataactaatagtattttgaacattataactagcattaataccagcaccagca
gtcttaccttcaacactaatattcccaccaatactatacccaatactattactaacatccttcttatcaatggtgttctttggaatac
tgtttgcaatcttaatgttgttgttaacatcagcactcttaacattaatccgataagtttctggccacttcatcttaccataatactta
ctaccaaaaatttgcttatcacttggaataaacccttctaaatttaataaagcagtcttcttatcactgtgtggatcttcaataaa
agtaccacttaaagtagccttccaacccttgtgactagcagtatcactactagtgtacttaatagcttccttaccgttttccttaat
aatttcaccacttaagttctttaattcaatcttgttaatatcgtttaattcactcatcggccgcatttaaaacgaacctcctgatatttt
gcaagaaataccttcaatatcaaaattgacgataatcctttaatttttattatactactttcctaagagaaaaccaagcctttctt
agctagcttagctcgttaagaaagaattccttgaactattatatagaggctaattaactttgtaaatacttaaatgaaggatcc
cattatcaaagtggcaactaacatactataaggatgctagaaaacggccttttaaaagagccttgtttaaacgtgatatacg
cgcatctaagcaactttcttgtatccttgtaaactggtaaacccgaggtacttaaaagggtgtatacgtaaaatagcaagctt
acaaaaggcattgcactttgtgcagggcttaaaaaagcccccgactaaaagccgagagctagttttaggtatgattacag
aaaataaagaatacacttttattatagcatcccaaaataaaaacccaagctgtacagcctaggttttgtcatatcatttagcc
acaaaatagagattattttatctagttgactagttgaccaacttgtcaacttttttagttgactagatgagacttattacagccaaa
gtaaaaaataaacccccttatttataggctttttcaaactctataatacgatgaaagtttgaaaaattattcccactcgatggttgc
cggtggcttggaagtgatgtcgtagacgacgcggttgacgtggtctacttcgttgacgattcgaactgagatggcttggagta
cgtcccagtcaatcttagcgaagtcggcggtcattccgtcgatggaagtaacggcccggataccgatagtgtagtcgtaag
tccggccatcacccatgacccccacggagcggatacctggcaagacggtgaagtattgccagatttccttgtctaagccgt
gcttggcaatttcttcgcgcaagatgtagtcggagtcacgcacgatttgaagtttttcatcggtaatttcaccgataacccggat
accaagtccaggaccagggaatggttggcgccaaactaattcgtgtggcatccctagcttttcacctaattcacggacttca
tccttgaataaggtgttaagtggttcgattaattcgaactgcatgtcttctggcaagcccccaacgttgtggtgggacttgatag
tttgagcagtatctgttcctgattcgattacgtcagtgtaaagcgtaccttgggccaagaacttgatgcctttaagcttagtagct
tcgtcatcgaaaacttggatgaactcatttccgataatcttgcgcttcttttctgggtcagagacgccagcaagtttgcctaaga
agcggtccttagcgtcaaccttgatgatgtttaagccaaatttgccaccgaggctgtccataacttgttcggcttcacccttacg

gagcaacccgtggtcaacaaagatactggttaattggtcaccaatggccttgtgtaatagaaccccaacaactgaggaat
ctaccccaccggaaagaccgaggagtacctttgatcaccgaccgtttggcgaatcttggcaacttgcatgtcgatgaagt
cttccattgaccagttagcttgggcaccacacacatcgaaagcaaagtgacgtaagatgtcattaccgtattcagtgtggcg
aacttctgtgtggaattgcaacccataaagtttctttgctgggttagccatagcagcaattgggcagttagcactagtagcact
gacagtaaaaccagcaggtacttcagttactaagtccccgtgactcatccaaacagtttgagtggttggtaagtccttaaac
attacggcgtccttatcgactacttcaatattggccttaccgtattccttgttgtcagcactttcgacctttccacctaaatcgtggg
ccatcaattgcatcccgtagcaaatccctaagataggtagtcccaattcaaagattgctgggtcaactttaaatgcgtcttcgt
cgtaaacactgttagggccacctgagaagatgatacccttagggttcatctgcttaatctcttgagctgtaagcttgtgagaa
agaagttcagagtaaactcccatttcgcgaatacgccgagtaattaattggttatactgacttccaaagtccaaaacgataa
tcttgtcgaaagcctgcatatcaacatttgccaagatattcacctcattaattaatttattatcctatattgtaacgacccaaggg
ctttcggtcaatgattttcaccttaatccgatcgcccttcccaacagttgcgcagcctgaatggcgaatgctagaatcgatacg
attttgaagtggcaacagataaaaaaaagcagtttaaaattgttgctgaacttttaaaacaagcaaatacaatcattgtcgc
aacagatagcgacagagaaggcgaaaacattgcctggtcgatcattcataaagcaaatgccttttctaaagataaaacgt
ataaaagactatggatcaatagtttagaaaaagatgtgatccgtagcggttttcaaaatttgcaaccaggaatgaattactat
cccttttatcaagaagcgcacaaaaagaaaaacgaaatgatacaccaatcagtgcaaaaaaagatataatgggagat
aagacggttcgtgttcgtgctgacttgcaccatatcataaaaatcgaaacagcaaagaatggcggaaacgtaaaagaag
ttatggaaataagacttagaagcaaacttaagagtgtgttgatagtgcagtatcttaaaattttgtataataggaattgaagtta
aattagatgctaaaaatttgtaattaagaaggagtgattacatgaacaaaaatataaaatattctcaaaacttttaacgagt
gaaaaagtactcaaccaaataataaaacaattgaatttaaaagaaaccgataccgtttacgaaattggaacaggtaaag
ggcatttaacgacgaaactggctaaaataagtaaacaggtaacgtctattgaattagacagtcatctattcaacttatcgtca
gaaaaattaaaactgaatactcgtgtcactttaattcaccaagatattctacagtttcaattccctaacaaacagaggtataa
aattgttgggagtattccttaccatttaagcacacaaattattaaaaaagtggtttttgaaagccatgcgtctgacatctatctga
ttgttgaagaaggattctacaagcgtaccttggatattcaccgaacactagggttgctcttgcacactcaagtctcgattcagc
aattgcttaagctgccagcggaatgctttcatcctaaaccaaaagtaaacagtgtcttaataaaacttacccgccataccac
agatgttccagataaatattggaagctatatacgtactttgtttcaaaatgggtcaatcgagaatatcgtcaactgtttactaaa
aatcagtttcatcaagcaatgaaacacgccaaagtaaacaatttaagtaccgttacttatgagcaagtattgtctatttttaata
gttatctattatttaacgggaggaaataattctatgagtcgcttttgtaaatttggaaagttacacgttactaaagggaatgtag
ataaattattaggtatactactgacagcttccaaggagctaaagaggtccctagcgctcttatcatggggaagctcggatca
tatgcaagacaaaataaactcgcaacagcacttggagaaatgggacgaatcgagaaaaccctctttacgctggattaca
tatctaataaagccgtaaggagacgggttcaaaaaggtttaaataaaggagaagcaatcaatgcattagctagaactata
ttttttggacaacgtggagaatttagagaacgtgctctccaagaccagttacaaagagctagtgcactaaacataattattaa
cgctataagtgtgtggaacactgtatatatggaaaaagccgtagaagaattaaaagcaagaggagaatttagagaagat
ttaatgccatatgcgtggccgttaggatgggaacatatcaattttcttggagaatacaaatttgaaggattacatgacactgg
gcaaatgaatttacgtcctttacgtataaaagagccgttttattcttaatataacggctctttttatagaaaaaatccttagcgtgg
tttttttccgaaatgctggcggtaccccaagaattagaaatgagtagatcaaattattcacgaatagaatcaggaaaatcag
atccaaccataaaaacactagaacaaattgcaaagttaactaactcaacgctagtagtggatttaatcccaaatgagcca
acagaaccagagccagaaacagaatcagaacaagtaacattggatttagaaatggaagaagaaaaaagcaatgact
tcgtgtgaataatgcacgaaatcgttgcttattttttttaaaagcggtatactagatataacgaaacaacgaactgaatagaa
acgaaaaaagagccatgacacatttataaaatgtttgacgacattttataaatgcatagcccgataagattgccaaaccaa
cgcttatcagttagtcagatgaactcttccctcgtaagaagttatttaattaactttgtttgaagacggtatataaccgtactatca
ttatatagggaaatcagagagttttcaagtatctaagctactgaatttaagaattgttaagcaatcaatcggaaatcgtttgatt
gctttttttgtattcatttatagaaggtggagtttgtatgaatcatgatgaatgtaaaacttatataaaaaatagtttattggagata
agaaaattagcaaatatctatacactagaaacgtttaagaaagagttagaaaagagaaatatctacttagaaacaaaatc
agataagtatttttcttcggagggggaagattatatatataagttaatagaaaataacaaaataatttattcgattagtggaaa
aaaattgacttataaaggaaaaaaatctttttcaaaacatgcaatattgaaacagttgaatgaaaaagcaaaccaagttaa
ttaaacaacctattttataggatttataggaaaggagaacagctgaatgaatatccctttgttgtagaaactgtgcttcatgac

ggcttgttaaagtacaaatttaaaaatagtaaaattcgctcaatcactaccaagccaggtaaaagcaaaggggctatttttg
cgtatcgctcaaaatcaagcatgattggcggtcgtggtgttgttctgacttccgaggaagcgattcaagaaaatcaagatac
atttacacattggacacccaacgtttatcgttatggaacgtatgcagacgaaaaccgttcatacacgaaaggacattctga
aaacaatttaagacaaatcaataccttctttattgattttgatattcacacggcaaaagaaactatttcagcaagcgatatttta
acaaccgctattgatttaggttttatgcctactatgattatcaaatctgataaaggttatcaagcatattttgttttagaaacgcca
gtctatgtgacttcaaaatcagaatttaaatctgtcaaagcagccaaaataatttcgcaaaatatccgagaatattttggaaa
gtctttgccagttgatctaacgtgtaatcattttggtattgctcgcataccaagaacggacaatgtagaatttttgatcctaatta
ccgttattctttcaaagaatggcaagattggtctttcaaacaaacagataataagggctttactcgttcaagtctaacggtttta
agcggtacagaaggcaaaaaacaagtagatgaaccctggtttaatctcttattgcacgaaacgaaattttcaggagaaa
agggtttaatagggcgtaataacgtcatgtttaccctctctttagcctactttagttcaggctattcaatcgaaacgtgcgaatat
aatatgtttgagtttaataatcgattagatcaacccttagaagaaaaagaagtaatcaaaattgttagaagtgcctattcaga
aaactatcaaggggctaatagggaatacattaccattctttgcaaagcttgggtatcaagtgatttaaccagtaaagatttatt
tgtccgtcaagggtggtttaaattcaagaaaaaaagaagcgaacgtcaacgtgttcatttgtcagaatggaaagaagattt
aatggcttatattagcgaaaaaagcgatgtatacaagccttatttagtgacgaccaaaaaagagattagagaagtgctag
gcattcctgaacggacattagataaattgctgaaggtactgaaggcgaatcaggaaattttctttaagattaaaccaggaa
gaaatggtggcattcaacttgctagtgttaaatcattgttgctatcgatcattaaagtaaaaaaagaagaaaaagaaagcta
tataaaggcgctgacaaattcttttgacttagagcatacattcattcaagagactttaaacaagctagcagaacgccctaaa
acggacacacaactcgatttgtttagctatgatacaggctgaaaataaaacccgcactatgccattacatttatatctatgat
acgtgtttgtttttctttgctgtttagcgaatgattagcagaaatatacagagtaagattttaattaattattagggggagaagga
gagagtagcccgaaaacttttagttggcttggactgaacgaagtgagggaaaggctactaaaacgtcgaggggcagtg
agagcgaagcgaacacttgatttttttaattttctatcttttataggtcattagagtatacttatttgtcctataaactatttagcagca
taatagatttattgaataggtcatttaagttgagcatattagaggaggaaaatcttggagaaatatttgaagaacccgattac
atggattggattagttcttgtggttacgtggtttttaactaaaagtagtgaatttttgatttttggtgtgtgtgtcttgttgttagtatttgct
agtcaaagtgattaaatagaattccggatgagcattcatcaggcgggcaagaatgtgaataaaggccggataaaacttgt
gcttatttttctttacggtctttaaaaaggccgtaatatccagctgaacggtctggttataggtacattgagcaactgactgaaat
gcctcaaaatgttctttacgatgccattgggatatatcaacggtggtatatccagtgatttttttctccattttagcttccttagctcct
gaaaatctcgataactcaaaaaatacgcccggtagtgatcttatttcattatggtgaaagttggaacctcttacgtgccgatc
aacgtctattttcgccaaaagttggcccagggcttcccggtatcaacagggacaccaggatttatttattctgcgaagtgat
cttccgtcacaggtatttattcggcgcaaagtgcgtcgggtgatgctgccaacttactgatttagtgtatgatggtgtttttgaggt
gctccagtggcttctgtttctatcagctgtccctcctgttcagctactgacggggtggtgcgtaacggcaaaagcaccgccgg
acatcagcgctagcggagtgtatactggcttactatgttggcactgatgagggtgtcagtgaagtgcttcatgtggcaggag
aaaaaaggctgcaccggtgcgtcagcagaatatgtgatacaggatatattccgcttcctcgctcactgactcgctacgctc
ggtcgttcgactgcggcgagcggaaatggcttacgaacggggcggagatttcctggaagatgccaggaagatacttaac
agggaagtgagagggccgcggcaaagccgtttttccataggctccgcccccctgacaagcatcacgaaatctgacgctc
aaatcagtggtggcgaaacccgacaggactataaagataccaggcgtttccccctggcggctccctcgtgcgctctcctgt
tcctgcctttcggtttaccggtgtcattccgctgttatggccgcgtttgtctcattccacgcctgacactcagttccgggtaggca
gttcgctccaagctggactgtatgcacgaaccccccgttcagtccgaccgctgcgccttatccggtaactatcgtcttgagtc
caacccggaaagacatgcaaaagcaccactggcagcagccactggtaattgatttagaggagttagtcttgaagtcatg
cgccggttaaggctaaactgaaaggacaagttttggtgactgcgctcctccaagccagttacctcggttcaaagagttggta
gctcagagaaccttcgaaaaaccgccctgcaaggcggttttttcgttttcagagcaagagattacgcgcagaccaaaacg
atctcaagaagatcatcttattaatcagataaaatatttctag

## 7.4 Supplementary data 4 – pVEcXLa-P$_{cwah}$-*rnetB*

>pVEcXLa-P$_{cwah}$-*rnetB*
gcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactgga
aagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggc
tcgtatgttgtgtggaattgtgagcggataacaatttcacacaggaaacagctatgaccatgattacgaattcatcgatatcta
gaggtacggatgccatatttaacccccgtagttataaatacggtggttataagatagtgttggatgccatactttaagtactgt
agttataactacataccgcaaccaccgtacttttacgtacgcttgtaaaaaacagggatgcaatcgccaaacattagtatta
actaagggtaaagcttttaacgtccttttataactatatacttaacaaccgtagttacaactacccttgtaagagtgcaagcaa
atattaccgagtatgtaaaaataagccgttttatcgtgatagacattttaggacaacacggctaaaccttggtaaatcttggta
tttttagagggcgtaactataattacacccttaccaccgtagttataacaacgcttgttacagtaatacttgtctagatataatat
acctaggctttgaattaggaccagctaccctatgaattagcttagtgaattgttgtaagactgtacgcccctcaacattaagca
tttaccgaggtgatgagtcatgacaacacgcgaaaggctagaggtgtttaaaattagcttggctttagttatcctagcattaa
gtcgcttaactctgagtatagcgcatgcttacaggataataaaaaaagctaaccacgattacaataaatcgtaattagctaa
tacacctaaagtcttaagcattcacgcttgagggtgaacaggttacgactgtttgccctcttttcaatatataataatacctgta
atagctttaaaagtaaagcgagatacaaactaaggggcgatgcaagatgccaaaagaaactaccgataaaagtaaaa
agcaagttgagtataataaacggtggtatgaaaaaaatagagaacacaagcgctatttaagcaagcgtagcacggcga
gaagtttttaagggtggcaactaaagaggatattcaagccctaattgaatttgctaacgaacaattaaaagaacgtttttaa
agtctagcatccaaactaggcttatttttatgtattgaatagaattagctcaagcaaagagctttttcagtgttgagagacgtatt
ttaagcctcgagctccacaaggtctttgtactaacctgtggttatgtataaaattgtagattttagggtaacaaaaaacaccgt
atttctacgatgtttttgcttaaatacttgtttttagttacagacaaacctgaagttaactatttatcaattcctgcaattcgtttacaa
aacggcaaatgtgaaatccgtcacatactgcgtgatgaacttgaattgccaaaggaagtataattttgttatcttctttataata
tttccccatagtaaaaataggaatcaaataatcatatccttttctgcaaattcagattaaagccatcgaaggttgaccacggta
tcatagatacattaaaaatgttttccggagcatttggctttccttccattctatgattgtttccataccgttgcgtatcactttcataat
ctgctaaaaatgatttaaagtcagacttacactcagtccaaaggctggaaaatgtttcagtatcattgtgaaatattgtatagc
ttggtatcatctcatcatatatccccaattcaccatcttgattgattgccgtcctaaactctgaatggcggtttacaatcattgcaa
tataataaagcattgcaggatatagtttcattccctttttcctttatttgtgtgatatccactttaacggtcatgctgtatgtacaaggt
acacttgcaaagtagtggtcaaaatactcttttctgttccaactatttttatcaatttttttcaaataccatctaagttccctctcaaatt
caagtttatcgctctaatgaacaaagatattataccacatttttgtgaatttttcaacttgcccacttcgactgcactcccgactta
ataacttcttgaacacttgccgaaaaagaactgcaggtcgacttagtagtattcaattttatgatcttgccagttaatcttaaac
ataaattcgttatattcactagtactacttaacttattagtaccccgccattgagtagtttcccagtttaaaatgtaatcgttatcaa
accgttgatattcaacaataataacactttccttagcattctttggagcagttaaagctaaagccatatttggactaaaaccac
cactaattaaggtacttaaatcccgatcatcagtaaaattcttatcaccattgttgtataattgactcttcataaataattgattac
cgtaaatagcgtggtaactatcaatgttataaccatccttagtttcaacaaacttaatatcccaactagctaaattagcatcatc
cttccgttgaatagtccgaaaatctggttgttcataactaatagtattttgaacattataactagcattaataccagcaccagca
gtcttaccttcaacactaatattcccaccaatactatacccaatactattactaacatccttcttatcaatggtgttctttggaatac
tgtttgcaatcttaatgttgttgttaacatcagcactcttaacattaatccgataagtttctggccacttcatcttaccataatactta
ctaccaaaaatttgcttatcacttggaataaacccttctaaatttaataaagcagtcttcttatcactgtgtggatcttcaataaa
agtaccacttaaagtagccttccaacccttgtgactagcagtatcactactagtgtacttaatagcttccttaccgttttccttaat
aatttcaccacttaagttctttaattcaatcttgttaatatcgtttaattcactcatcggccgatctttcactcccggatattaattttac
gactgtgattctattttaatcttttaatattgcattagggtacaaaaaaggttacaaaacttaattttttgtaacaaaagtccgata
aaagctaattataaagggattgttaacttttgtaagaatttgttaagtaattcttaagcaaagaaagcccgtggttgagctgtttt
tctgaagggtaagttgacttgtaattcttatctcaacggatcccattatcaaagtggcaactaacatactataaggatgctaga
aaacggccttttaaaagagccttgtttaaacgtgatatacgcgcatctaagcaactttcttgtatccttgtaaactggtaaaccc
gaggtacttaaaagggtgtatacgtaaaatagcaagcttacaaaaggcattgcactttgtgcagggcttaaaaaagcccc
cgactaaaagccgagagctagttttaggtatgattacagaaaataaagaatacacttttattatagcatcccaaaataaaa

acccaagctgtacagcctaggttttgtcatatcatttagccacaaaatagagattattttatctagttgactagttgaccaacttg
tcaactttttagttgactagatgagacttattacagccaaagtaaaaaataaacccccttatttataggctttttcaaactctataat
acgatgaaagtttgaaaaattattcccactcgatggttgccggtggcttggaagtgatgtcgtagacgacgcggttgacgtg
gtctacttcgttgacgattcgaactgagatggcttggagtacgtcccagtcaatcttagcgaagtcggcggtcattccgtcgat
ggaagtaacggcccggataccgatagtgtagtcgtaagtccggccatcacccatgacccccacggagcggatacctgg
caagacggtgaagtattgccagatttccttgtctaagccgtgcttggcaatttcttcgcgcaagatgtagtcggagtcacgca
cgatttgaagtttttcatcggtaatttcaccgataacccggataccaagtccaggaccagggaatggttggcgccaaactaa
ttcgtgtggcatccctagcttttcacctaattcacggacttcatccttgaataaggtgttaagtggttcgattaattcgaactgcat
gtcttctggcaagcccccaacgttgtggtgggacttgatagtttgagcagtatctgttcctgattcgattacgtcagtgtaaagc
gtaccttgggccaagaacttgatgcctttaagcttagtagcttcgtcatcgaaaacttggatgaactcatttccgataatcttgc
gcttcttttctgggtcagagacgccagcaagtttgcctaagaagcggtccttagcgtcaaccttgatgatgtttaagccaaattt
gccaccgaggctgtccataacttgttcggcttcacccttacggagcaacccgtggtcaacaaagatactggttaattggtca
ccaatggccttgtgtaatagaaccccaacaactgaggaatctaccccaccggaaagaccgaggagtacctttgatcacc
gaccgtttggcgaatcttggcaacttgcatgtcgatgaagtcttccattgaccagttagcttgggcaccacacacatcgaaa
gcaaagtgacgtaagatgtcattaccgtattcagtgtggcgaacttctgtgtggaattgcaacccataaagtttctttgctgggt
tagccatagcagcaattgggcagttagcactagtagcactgacagtaaaaccagcaggtacttcagttactaagtccccgt
gactcatccaaacagtttgagtggttggtaagtccttaaacattacggcgtccttatcgactacttcaatattggccttaccgtat
tccttgttgtcagcactttcgacctttccacctaaatcgtgggccatcaattgcatcccgtagcaaatccctaagataggtagtc
ccaattcaaagattgctgggtcaactttaaatgcgtcttcgtcgtaaacactgttagggccacctgagaagatgatacccctta
gggttcatctgcttaatctcttgagctgtaagcttgtgagaaagaagttcagagtaaactcccatttcgcgaatacgccgagt
aattaattggttatactgacttccaaagtccaaaacgataatcttgtcgaaagcctgcatatcaacatttgccaagatattcac
ctcattaattaatttattatcctatattgtaacgacccaagggctttcggtcaatgattttcaccttaatccgatcgcccttcccaac
agttgcgcagcctgaatggcgaatgctagaatcgatacgattttgaagtggcaacagataaaaaaaagcagtttaaaatt
gttgctgaacttttaaaacaagcaaatacaatcattgtcgcaacagatagcgacagagaaggcgaaaacattgcctggtc
gatcattcataaagcaaatgccttttctaaagataaaacgtataaaagactatggatcaatagtttagaaaaagatgtgatc
cgtagcggttttcaaaatttgcaaccaggaatgaattactatccctttttatcaagaagcgcacaaaaagaaaaacgaaatg
atacaccaatcagtgcaaaaaaagatataatgggagataagacggttcgtgttcgtgctgacttgcaccatatcataaaaa
tcgaaacagcaaagaatggcggaaacgtaaaagaagttatggaaataagacttagaagcaaacttaagagtgtgttga
tagtgcagtatcttaaaattttgtataataggaattgaagttaaattagatgctaaaaatttgtaattaagaaggagtgattacat
gaacaaaaatataaaatattctcaaaacttttttaacgagtgaaaaagtactcaaccaaataataaaacaattgaatttaaa
agaaaccgataccgtttacgaaattggaacaggtaaagggcatttaacgacgaaactggctaaaataagtaaacaggt
aacgtctattgaattagacagtcatctattcaacttatcgtcagaaaaattaaaactgaatactcgtgtcactttaattcaccaa
gatattctacagtttcaattccctaacaaacagaggtataaaattgttgggagtattccttaccatttaagcacacaaattatta
aaaaagtggtttttgaaagccatgcgtctgacatctatctgattgttgaagaaggattctacaagcgtaccttggatattcacc
gaacactagggttgctcttgcacactcaagtctcgattcagcaattgcttaagctgccagcggaatgctttcatcctaaacca
aaagtaaacagtgtcttaataaaacttacccgccataccacagatgttccagataaatattggaagctatatacgtactttgtt
tcaaaatgggtcaatcgagaatatcgtcaactgtttactaaaaatcagtttcatcaagcaatgaaacacgccaaagtaaac
aatttaagtaccgttacttatgagcaagtattgtctatttttaatagttatctattatttaacgggaggaaataattctatgagtcgct
tttgtaaatttggaaagttacacgttactaaagggaatgtagataaattattaggtatactactgacagcttccaaggagctaa
agaggtccctagcgctcttatcatggggaagctcggatcatatgcaagacaaaataaactcgcaacagcacttggagaa
atgggacgaatcgagaaaaccctctttacgctggattacatatctaataaagccgtaaggagacgggttcaaaaaggttta
aataaaggagaagcaatcaatgcattagctagaactatatttttggacaacgtggagaatttagagaacgtgctctccaa
gaccagttacaaagagctagtgcactaaacataattattaacgctataagtgtgtggaacactgtatatatggaaaaagcc
gtagaagaattaaaagcaagaggagaatttagagaagatttaatgccatatgcgtggccgttaggatgggaacatatca
attttcttggagaatacaaatttgaaggattacatgacactgggcaaatgaatttacgtcctttacgtataaaagagccgtttta
ttcttaatataacggctctttttatagaaaaaatccttagcgtggtttttttccgaaatgctggcggtaccccaagaattagaaat

gagtagatcaaattattcacgaatagaatcaggaaaatcagatccaaccataaaaacactagaacaaattgcaaagtta
actaactcaacgctagtagtggatttaatcccaaatgagccaacagaaccagagccagaaacagaatcagaacaagt
aacattggatttagaaatggaagaagaaaaaagcaatgacttcgtgtgaataatgcacgaaatcgttgcttattttttttaaa
agcggtatactagatataacgaaacaacgaactgaatagaaacgaaaaaagagccatgacacatttataaaatgtttga
cgacattttataaatgcatagcccgataagattgccaaaccaacgcttatcagttagtcagatgaactcttccctcgtaagaa
gttatttaattaactttgtttgaagacggtatataaccgtactatcattatatagggaaatcagagagttttcaagtatctaagcta
ctgaatttaagaattgttaagcaatcaatcggaaatcgtttgattgctttttttgtattcatttatagaaggtggagtttgtatgaatc
atgatgaatgtaaaacttatataaaaaatagtttattggagataagaaaattagcaaatatctatacactagaaacgtttaag
aaagagttagaaaagagaaatatctacttagaaacaaaatcagataagtatttttcttcggagggggaagattatatatata
agttaatagaaaataacaaaataatttattcgattagtggaaaaaaattgacttataaaggaaaaaaatctttttcaaaacat
gcaatattgaaacagttgaatgaaaaagcaaaccaagttaattaaacaacctattttataggatttataggaaaggagaac
agctgaatgaatatccctttttgttgtagaaactgtgcttcatgacggcttgttaaagtacaaatttaaaaatagtaaaattcgct
caatcactaccaagccaggtaaaagcaaaggggctatttttgcgtatcgctcaaaatcaagcatgattggcggtcgtggtg
ttgttctgacttccgaggaagcgattcaagaaaatcaagatacatttacacattggacacccaacgtttatcgttatggaacg
tatgcagacgaaaaccgttcatacacgaaaggacattctgaaaacaatttaagacaaatcaataccttctttattgattttgat
attcacacggcaaaagaaactatttcagcaagcgatattttaacaaccgctattgatttaggttttatgcctactatgattatca
aatctgataaaggttatcaagcatattttgttttagaaacgccagtctatgtgacttcaaaatcagaatttaaatctgtcaaagc
agccaaaataatttcgcaaaatatccgagaatattttggaaagtctttgccagttgatctaacgtgtaatcattttggtattgctc
gcataccaagaacggacaatgtagaatttttgatcctaattaccgttattctttcaaagaatggcaagattggtctttcaaaca
aacagataataagggctttactcgttcaagtctaacggttttaagcggtacagaaggcaaaaaacaagtagatgaaccct
ggtttaatctcttattgcacgaaacgaaattttcaggagaaaagggtttaatagggcgtaataacgtcatgtttaccctctctta
gcctactttagttcaggctattcaatcgaaacgtgcgaatataatatgtttgagtttaataatcgattagatcaacccttagaag
aaaaagaagtaatcaaaattgttagaagtgcctattcagaaaactatcaaggggctaatagggaatacattaccattctttg
caaagcttgggtatcaagtgatttaaccagtaaagatttatttgtccgtcaagggtggtttaaattcaagaaaaaaagaagc
gaacgtcaacgtgttcatttgtcagaatggaaagaagatttaatggcttatattagcgaaaaaagcgatgtatacaagcctt
atttagtgacgaccaaaaaagagattagagaagtgctaggcattcctgaacggacattagataaattgctgaaggtactg
aaggcgaatcaggaaattttctttaagattaaaccaggaagaaatggtggcattcaacttgctagtgttaaatcattgttgcta
tcgatcattaaagtaaaaaaagaagaaaaagaaagctatataaaggcgctgacaaattcttttgacttagagcatacattc
attcaagagactttaaacaagctagcagaacgccctaaaacggacacacaactcgatttgtttagctatgatacaggctg
aaaataaaacccgcactatgccattacatttatatctatgatacgtgtttgtttttttctttgctgtttagcgaatgattagcagaaat
atacagagtaagattttaattaattattagggggagaaggagagagtagcccgaaaacttttagttggcttggactgaacg
aagtgagggaaaggctactaaaacgtcgaggggcagtgagagcgaagcgaacacttgatttttttaattttctatcttttatag
gtcattagagtatacttatttgtcctataaactatttagcagcataatagatttattgaataggtcatttaagttgagcatattagag
gaggaaaatcttggagaaatatttgaagaacccgattacatggattggattagttcttgtggttacgtggtttttaactaaaagt
agtgaatttttgatttttggtgtgtgtgtcttgttgttagtatttgctagtcaaagtgattaaatagaattccggatgagcattcatca
ggcgggcaagaatgtgaataaaggccggataaaacttgtgcttatttttctttacggtctttaaaaaggccgtaatatccagct
gaacggtctggttataggtacattgagcaactgactgaaatgcctcaaaatgttctttacgatgccattgggatatatcaacg
gtggtatatccagtgatttttttctccattttagcttccttagctcctgaaaatctcgataactcaaaaaatacgcccggtagtgat
cttatttcattatggtgaaagttggaacctcttacgtgccgatcaacgtctattttcgccaaaagttggcccagggcttcccgg
tatcaacagggacaccaggatttatttattctgcgaagtgatcttccgtcacaggtatttattcggcgcaaagtgcgtcgggtg
atgctgccaacttactgatttagtgtatgatggtgtttttgaggtgctccagtggcttctgtttctatcagctgtccctcctgttcagct
actgacggggtggtgcgtaacggcaaaagcaccgccggacatcagcgctagcggagtgtatactggcttactatgttggc
actgatgagggtgtcagtgaagtgcttcatgtggcaggagaaaaaaggctgcaccggtgcgtcagcagaatatgtgata
caggatatattccgcttcctcgctcactgactcgctacgctcggtcgttcgactgcggcgagcggaaatggcttacgaacgg
ggcggagatttcctggaagatgccaggaagatacttaacagggaagtgagagggccgcggcaaagccgtttttccatag
gctccgcccccctgacaagcatcacgaaatctgacgctcaaatcagtggtggcgaaacccgacaggactataaagata

ccaggcgtttccccctggcggctccctcgtgcgctctcctgttcctgcctttcggtttaccggtgtcattccgctgttatggccgc
gtttgtctcattccacgcctgacactcagttccgggtaggcagttcgctccaagctggactgtatgcacgaacccccccgttca
gtccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggaaagacatgcaaaagcaccactggcagca
gccactggtaattgatttagaggagttagtcttgaagtcatgcgccggttaaggctaaactgaaaggacaagttttggtgact
gcgctcctccaagccagttacctcggttcaaagagttggtagctcagagaaccttcgaaaaaccgccctgcaaggcggttt
tttcgttttcagagcaagagattacgcgcagaccaaaacgatctcaagaagatcatcttattaatcagataaaatatttctag

# 8. References


ABARES (2013). "Agricultural commodity statistics 2013", Department of Agriculture**:** 1 - 254.

ABARES (2016). "Agricultural commodities Outlook: MARCH QUARTER 2016". D. o. A. a. W. Resources. http://www.agriculture.gov.au/, Australian Government**:** 1 - 286.

Abelson, P., M. P. Forbes and G. Hall (2006). "The annual cost of foodborne illness in Australia". A. G. D. o. H. a. Ageing. Australia**:** 1 - 108.

Aguilar, J. C. and E. G. Rodríguez (2007). "Vaccine adjuvants revisited." Vaccine **25**: 3752 - 3762.

Ahmed, B., M. Loos, D. Vanrompay and E. Cox (2014). "Oral immunization with *Lactococcus lactis*-expressing EspB induces protective immune responses against *Escherichia coli* O157:H7 in a murine model of colonization." Vaccine **32**(31): 3909 - 3916.

Aires, K. A., A. M. Cianciarullo, S. M. Carneiro, L. L. Villa, E. Boccardo, G. Pérez-Martinez, I. Perez-Arellano, M. L. S. Oliveira and P. L. Ho (2006). "Production of Human Papillomavirus Type 16 L1 Virus-Like Particles by Recombinant *Lactobacillus casei* Cells." Applied and Environmental Microbiology **71**(1): 745 - 752.

Alimolaeia, M., M. Golchina and M. Ezatkhah (2017). "Orally administered recombinant Lactobacillus casei vector vaccine expressing β-toxoid of *Clostridium perfringens* that induced protective immunity responses." Research in Veterinary Science **115**: 332 - 339.

Alpert, C.-A., A.-M. C.-L. Coq, C. Malleret and M. Zagorec (2003). "Characterization of a Theta-Type Plasmid from *Lactobacillus sakei*: a Potential Basis for Low-Copy-Number Vectors in lactobacilli." Applied and Environmental Microbiology **69**(9): 5574 - 5584.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic Local Alignment Search Tool." Journal of Molecular Biology **215**: 403 - 410.

An, H.-Y. and T. Miyamoto (2006). "Cloning and sequencing of plasmid pLC494 isolated from human intestinal *Lactobacillus casei*: Construction of an *Escherichia coli–Lactobacillus* shuttle vector." Plasmid **55**: 128 - 134.

Antikainen, J., V. Kupannen, K. Lahteenmaki and T. K. Korhonen (2007). "pH-Dependent Association of Enolase and Glyceraldehyde-3-Phosphate Dehydrogenase of *Lactobacillus crispatus* with the Cell Wall and Lipoteichoic Acids." Journal of Bacteriology **189**(12): 4539 - 4543.

Archer, G. S. and J. A. Mench (2014). "The effects of the duration and onset of light stimulation during incubation on the behavior, plasma melatonin levels, and productivity of broiler chickens." Journal of Animal Science **92**: 1753 - 1758.

Aronesty, E. (2011). "ea-utils : Command-line tools for processing biological sequencing data." from https://github.com/ExpressionAnalysis/ea-utils.

Auvray, F., M. Coddeville, R. C. Ordonez and P. Ritzenthaler (1999). "Unusual Structure of the attB Site of the Site-Specific Recombination System of *Lactobacillus delbrueckii* Bacteriophage mv4." Journal Of Bacteriology **181**(23): 7385 - 7389.

Aziz, R. K., D. Bartels, A. A. Best, M. DeJongh, T. Disz, R. Edwards, K. Formsma, S. Gerdes, E. Glass, M. Kubal, F. Meyer, G. Olsen, R. Olson, A. Osterman, R. Overbeek, L. McNeil, D. Paarmann, T. Paczian, B. Parrello, G. Pusch, C. Reich, R. Stevens, O. Vassieva, V. Vonstein, A. Wilke and O. Zagnitko (2008). "The RAST Server: rapid annotations using subsystems technology." BMC Bioinformatics **9**(75).

Baele, M., L. A. Devriese and F. Haesebrouck (2001). "*Lactobacillus agilis* is an important component of the pigeon crop Flora." Journal of Applied Microbiology **91**: 488 - 491.

Baldwin, S., R. J. Hughes, T. T. H. Van, R. J. Moore and D. Stanley (2018). "At-hatch administration of probiotic to chickens can introduce beneficial changes in gut microbiota." PLoS One **13**(3).

Bannam, T. L. and J. I. Rood (1992). "*Clostridium perfringens-Escherichia coli* Shuttle Vectors That Carry Single Antibiotic Resistance Determinants." Plasmid **229**: 233 - 235.

Bassford, P. J., T. J. Silhavy and J. R. Beckwith (1979). "Use of Gene Fusion to Study Secretion of Maltose-Binding Protein into *Escherichia coli* Periplasm." Journal Of Bacteriology **139**(1): 19 - 31.

Beery, J. T., M. B. Hugdahl and M. P. Doyle (1988). "Colonization of Gastrointestinal Tracts of Chicks by *Campylobacter jejuni*." Applied and Environmental Microbiology **54**(10): 2365 - 2370.

Berlin, K., S. Koren, C.-S. Chin, J. P. Drake, J. M. Landolin and A. M. Phillippy (2015). "Assembling large genomes with single-molecule sequencing and locality-sensitive hashing." Nature Biotechnology **33**(6): 623 - 634.

Bermúdez-Humarán, L. G. (2009). "*Lactococcus lactis* as a live vector for mucosal delivery of therapeutic proteins." Human Vaccines **5**(4): 264 - 267.

Bibby, S. Pers. comm. 11/8/14

Black, R. E., M. M. Levine, M. L. Clements, T. P. Hughes and M. J. Blaser (1988). "Experimental *Campylobacter jejuni* Infection in Humans." The Journal of Infectious Diseases **157**(3): 472 - 479.

Bochman, M. L., K. Paeschke and V. A. Zakian (2012). "DNA secondary structures: stability and function of G-quadruplex structures." Nature Reviews Genetics **13**(11): 770 - 780.

Bohez, L., R. Ducatelle, F. Pasmans, F. Haesebrouck and F. V. Immerseel (2007). "Long-term colonisation–inhibition studies to protect broilers against colonisation with *Salmonella* Enteritidis, using *Salmonella* Pathogenicity Island 1 and 2 mutants." Vaccine **25**: 4235 - 4243.

Boratyn, G. M., C. Camacho, P. S. Cooper, G. Coulouris, N. M. Amelia Fong, T. L. Madden, W. T. Matten, S. D. McGinnis, Y. Merezhuk, Y. Raytselis, E. W. Sayers, T. Tao, J. Ye and I. Zaretskaya (2013). "BLAST: a more efficient report with usability improvements." Nucleic Acids Research **41**: 1 - 5.

Branny, P., F. d. l. Torre and J.-R. Garel (1998). "An operon encoding three glycolytic enzymes in *Lactobacillus delbrueckii* subsp. *bulgaricus*: glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase and triosephosphate isomerase." Microbiology **144**: 905 - 914.

Brettin, T., J. J. Davis, T. Disz, R. A. Edwards, S. Gerdes, G. J. Olsen, R. Olson, R. Overbeek, B. Parrello, G. D. Pusch, M. Shukla, J. A. T. III, R. Stevens, V. Vonstein, A. R. Wattam and F. Xia (2015). "RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes." Scientific Reports **5**: 1 - 6.

Bron, P. A., M. G. Benchimol, J. Lambert, E. Palumbo, M. Deghorain, J. Delcour, W. M. d. Vos, M. Kleerebezem and P. Hols (2002). "Use of the alr Gene as a Food-Grade Selection Marker in Lactic Acid Bacteria." Applied and Environmental Microbiology **68**(11): 5663 - 5670.

Browning, D. F. and S. J. W. Busby (2004). "The regulation of bacterial transcription initiation." Nature Reviews Microbiology **2**(1): 57 - 65.

Bryant, J. A., L. E. Sellars, S. J. W. Busby and D. J. Lee (2014). "Chromosome position effects on gene expression in *Escherichia coli* K-12." Nucleic Acids Research **42**(18): 11383 - 11392.

Bujnakova, D., E. Strakova and V. Kmet (2014). "In vitro evaluation of the safety and probiotic properties of lactobacilli isolated from chicken and calves." Anaerobe **9**: 118 - 127.

Caldwell, D. Y., R. W. Moore, D. J. Caldwell and B. M. Hargis (2001). "Effect of photointensity, sound intensity, and ambient temperature on preening behavior and ingestion of spray-applied biologics." Journal Applied Poultry **10**: 99 - 106.

Canosi, U., G. Morelli and T. A. Trautner (1978). "The relationship between molecular structure and transformation efficiency of some *S. aureus* plasmids isolated from *B. subtilis*." Molecular and General Genetics **166**(3): 259 - 267.

Cardenas, L. and J. D. Clements (1993). "Stability, immunogenicity and expression of foreign antigens in bacterial vaccine vectors." Vaccine **11**(2): 126 - 135.

Catarame, T. M. G., K. A. O'Hanlon, D. A. Mcdowell, I. S. Blair and G. Duffy (2006). "Comparison of a real-time polymerase chain reaction assay with a culture method for the detection of *Salmonella* in retail meat samples." Journal of Food Safety **26**(1): 1 - 15.

Cawthraw, S., A. R. Ayling, A. P. Nuijten, T. Wassenaar and D. G. Newell (1994). "Isotype, Specificity, and Kinetics of Systemic and Mucosal Antibodies to *Campylobacter jejuni* Antigens, including Flagellin, during Experimental Oral Infections of Chickens." Avian Diseases **38**: 341 - 349.

Chang, S. and S. N. Cohen (1979). "High frequency transformation of *Bacillus subtilis* protoplasts by plasmid DNA." Molecular and General Genetics **168**(1): 111 - 115.

Chaudhari, N. M., V. K. Gupta and C. Dutta (2016). "BPGA- an ultra-fast pan-genome analysis pipeline." Scientific Reports **6**: 1 - 10.

Chen, C. Y., S. W. Chen and H. T. Wang (2017). "Effect of supplementation of yeast with bacteriocin and *Lactobacillus* culture on growth performance, cecal fermentation, microbiota composition, and blood characteristics in broiler chickens." Asian-Australasian Journal of Animal Sciences **30**(2): 211 - 220.

Chen, T., H. Siddiqui and I. Olsen (2017). "In silico Comparison of 19 *Porphyromonas gingivalis* Strains in Genomics, Phylogenetics, Phylogenomics and Functional Genomics." Frontiers in Infection Microbiology **7**(28): 1 - 24.

Chen, Y. S. and J. L. Steele (2005). "Analysis of promoter sequences from *Lactobacillus helveticus* CNRZ32 and their activity in other lactic acid bacteria." Journal of Applied Microbiology **98**: 64 - 72.

Chin, C.-S., D. H. Alexander, P. Marks, A. A. Klammer, J. Drake, C. Heiner, A. Clum, A. Copeland, J. Huddleston, E. E. Eichler, S. W. Turner and J. Korlach (2013). "Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data." Nature Methods **10**(6): 563 - 571.

Cho, S. H., P. C. Loewen and R. R. Marquardt (2004). "A plasmid DNA encoding chicken interleukin-6 and *Escherichia coli* K88 fimbrial protein FaeG stimulates the production of anti-K88 fimbrial antibodies in chickens." Poultry Science **83**(12): 1973 - 1978.

Conway, T., G. W. Sewell and L. O. Ingram (1987). "Glyceraldehyde-3-Phosphate Dehydrogenase Gene from *Zymomonas mobilis*: Cloning, Sequencing, and Identification of Promoter Region." Journal of Bacteriology **169**(12): 5653 - 5662.

Cooper, D. L., R. S. Lahue and P. Modrich (1993). "Methyl-directed Mismatch Repair is Bidirectional." The Journal of Biological Chemistry **268**(16): 11823 - 11829.

Corry, J. E. L. and H. I. Atabay (2001). "Poultry as a source of Campylobacter and related organisms." Journal of Applied Microbiology **90**: 96 - 114.

Costa, M. J. D., E. O. Oviedo-Rondón, M. J. Wineland, J. Wilson and E. Montiel (2014). "Effects of breeder feeding restriction programs and incubation temperatures on progeny footpad development." Poultry Science **93**: 1 - 10.

Costa, S. P. F. d., D. Mot, M. Bokori-Brown, C. G. Savva, A. K. Basak, F. V. Immerseel and R. W. Titball (2013). "Protection against avian necrotic enteritis after immunisation with NetB genetic or formaldehyde toxoids." Vaccine **31**: 4003 - 4008.

Crhanova, M., H. Hradecka, M. Faldynova, M. Matulova, H. Havlickova, F. Sisak and I. Rychlik (2011). "Immune Response of Chicken Gut to Natural Colonization by Gut Microflora and to Salmonella enterica Serovar Enteritidis Infection, ." Infection and Immunity **79**(7): 2755 – 2763.

Croix, M. D. S., I. Vacca, M. J. Kwun, J. D. Ralph, S. D. Bentley, R. Haigh, N. J. Croucher and M. R. Oggioni (2017). "Phase-variable methylation and epigenetic regulation by type I restriction–modification systems." FEMS Microbiology Reviews **41**: S3 - S15.

Cullender, T. C., B. Chassaing, A. Janzon, K. Kumar, C. Muller, E , J. Werner, J, L. Angenent, T, M. E. Bell, A. G. Hay, D. A. Peterson, J. Walter, M. Vijay-Kumar, A. T. Gewirtz and R. E. Ley (2013). "Innate and Adaptive Immunity Interact to Quench Microbiome Flagellar Motility in the Gut." Cell Host & Microbe **14**: 571 – 581.

Curtiss III, R. and J. O. Hassan (1996). "Nonrecombinant and recombinant avirulent *Salmonella* vaccines for poultry." Veterinary Immunology and Immunopathology **54**: 365 - 372.

Curtiss III, R. and J. O. Hassan (1996). "Nonrecombinant and recombinant avirulent *Salmonella* vaccines for poultry " Veterinary Immunology and Immunopathology **54**: 365 - 372.

Dalton, C. B., J. Gregory, M. D. Kirk, R. J. Stafford, R. Givney, E. Kraa and D. Gould (2004). "Foodbome disease outbreaks in Australia, 1995 to 2000." Communicable Diseases Intelligence Quarterly Report **28**(2): 211 - 224.

Danzeisen, J. L., H. B. Kim, R. E. Isaacson, Z. J. Tu and T. J. Johnson (2011). "Modulations of the Chicken Cecal Microbiome and Metagenome in Response to Anticoccidial and Growth Promoter Treatment." PLOS ONE **6**(11).

Dave, K. K. and N. S. Punekar (2015). "Expression of Lactate Dehydrogenase in *Aspergillus niger* for L-Lactic Acid Production." PLOS ONE **10**(12).

De Cort, W., S. Geeraerts, V. Balan, M. Elroy, F. Haesebrouck, R. Ducatelle and F. Van Immerseel (2013). "A *Salmonella* Enteritidis hilAssrAfliG deletion mutant is a safe live vaccine strain that confers protection against colonization by *Salmonella* enteritidis in broilers." Vaccine **31**: 5104 – 5110.

De Cort, W., F. Haesebrouck, R. Ducatelle and F. v. Immerseel (2015). "Administration of a *Salmonella* Enteritidis *hilAssrAfliG* strain by coarse spray to newly hatched broilers reduces colonization and shedding of a *Salmonella* Enteritidis challenge strain." Poultry Science **94**(1): 131 - 125.

de Zoete, M. R., J. P. M. van Putten and J. A. Wagenaar (2007). "Vaccination of chickens against *Campylobacter*." Vaccine **25**: 5548 - 5557.

Dhillon, A. S., A. H. L. Shivaprasad, B. D. Schaberg, A. F. Wier, A. S. Weber and D. Bandli (2006). "Campylobacter jejuni Infection in Broiler Chickens." Avian Diseases **50**: 55 - 58.

Dillard, J. P. and J. Yother (1991). "Analysis of *Streptococcus pneumoniae* Sequences Cloned into *Escherichia coli*: Effect of Promoter Strength and Transcription Terminators." Journal of Bacteriology **173**(16): 5105 - 5109.

DoAaF. (2012, 2012). "Pig Production: Performance standards." Retrieved 14/2/18, from https://web.archive.org/web/20160616102808/https://www.daf.qld.gov.au/animal-industries/pigs/managing-a-piggery/production-and-performance/performance-standards.

DoH. (2017). "Notification Rate of Campylobacteriosis*, received from State and Territory health authorities in the period of 1991 to 2016 and year-to-date notifications for 2017." National Notifiable Diseases Surveillance System Retrieved 19/4/17, from http://www9.health.gov.au/cda/source/rpt_3_sel.cfm.

Drissi, F., N. Labas, V. Merhej and D. Raoult (2015). "Draft Genome Sequence of the *Lactobacillus agilis* Strain Marseille." Genome Announcements **3**(4): e00840 -00815.

Dunne, C., L. Murphy, S. Flynn, L. O'Mahony, S. O'Halloran, M. Feeney, D. Morrissey, G. Thornton, G. Fitzgerald, C. Daly, B. Kiely, E. M. M. Quigley, G. C. O'Sullivan, F. Shanahan and J. K. Collins (1999). "Probiotics: from myth to reality. Demonstration of functionality in animal models of disease and in human clinical trials." Antonie van Leeuwenhoek **76**: 279 - 292.

EFSA (2016). "The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2015." EFSA Journal **14**(12).

EFSA ( 2010). "Analysis of the baseline survey on the prevalence of *Campylobacter* in broiler batches and of *Campylobacter* and *Salmonella* on broiler carcasses in the EU, 2008." EFSA Journal **8(03)**(1503): 1 - 100.

Field, D., B. Tiwari, T. Booth, S. Houten, D. Swan, N. Bertrand and M. Thurston (2006). "Open software for biologists: from famine to feast." Nature Biotechnology **24**: 801 - 803.

Finn, R. D., J. Clements, W. Arndt, B. L. Miller, T. J. Wheeler, F. Schreiber, A. Bateman and S. R. Eddy (2015). "HMMER web server: 2015 update." Nucleic Acids Research **43**(1): 30 - 38.

Fitzsimons, A., P. Hols, J. Jore, R. J. Leer, M. O'connell and J. Delcour (1994). "Development of an Amylolytic *Lactobacillus plantarum* Silage Strain Expressing the *Lactobacillus amylovorus* α-Amylase Gene." Applied and Environmental Microbiology **60**(10): 3529 - 3535.

FSANZ (2005). Scientific assessment of the public health and safety of poultry meat in Australia, Food Standards Australia New Zealand.

Garcia, P., K. L. Hopkins, V. Garcia, J. Beutlich, M. C. Mendoza, J. Threlfall, D. Mevius, R. Helmuth, M. R. Rodicio and B. Guerra (2014). "Diversity of Plasmids Encoding Virulence and Resistance Functions in *Salmonella enterica* subsp. enterica Serovar Typhimurium Monophasic Variant 4,[5],12:i:- Strains Circulating in Europe." PLoS One **9**(2).

Gasson, M. J. (1983). "Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic streptococci after protoplast-induced curing." Journal of Bacteriology **154**(1): 1 - 9.

Gervasi, T., N. Horn, U. Wegmann, G. Dugo, A. Narbad and M. J. Mayer (2014). "Expression and delivery of an endolysin to combat *Clostridium perfringens*." Applied Microbiology and Biotechnology **98**: 2495 - 2505.

Giannenas, I., E. Tsalie, E. Triantafillou, S. Hessenberger, K. Teichmann, M. Mohnl and D. Tontis (2014). "Assessment of probiotics supplementation via feed or water on the growth performance, intestinal morphology and microflora of chickens after experimental

infection with *Eimeria acervulina*, *Eimeria maxima* and *Eimeria tenella*." Avian Pathology **43**(3): 209 - 216.

Giovanardi, D., E. Campagnari, L. S. Ruffoni, P. Pesente, G. Ortali and V. Furlattini (2005). "Avian pathogenic Escherichia coli transmission from broiler breeders to their progeny in an integrated poultry production chain." Avian Pathology **34**(4): 313 - 318.

Glenting, J., H. C. Beck, A. Vranga, H. Riemanna, P. Ravna, A. M. Hansen, M. Antonssonc, S. Ahrné, H. Israelsena and S. Madsen (2013). "Anchorless surface associated glycolytic enzymes from *Lactobacillus plantarum* 299v bind to epithelial cells and extracellular matrix proteins." Microbiological Research **168**: 245 - 253.

Granato, D., G. E. Bergonzelli, R. D. Pridmore, L. Marvin, M. Rouvet and I. E. Corthesy-Theulaz (2004). "Cell Surface-Associated Elongation Factor Tu Mediates the Attachment of *Lactobacillus johnsonii* NCC533 (La1) to Human Intestinal Cells and Mucins." Infection and Immunity **72**(4): 2160 - 2169.

Grangette, C., H. Muller-Alouf, D. Goudercourt, M.-C. Geoffroy, M. Turneer and A. Mercenier (2001). "Mucosal Immune Responses and Protection against Tetanus Toxin after Intranasal Immunization with Recombinant *Lactobacillus plantarum*." Infection and Immunity **69**(3): 1547 - 1553.

Guiziou, S., V. Sauveplane, H. J. Chang, C. Clerté, N. Declerck, M. Jules and J. Bonnet (2016). "A part toolbox to tune genetic expression in *Bacillus subtilis*." Nucleic Acids Research **44**(15): 7495 - 7508.

Heel, A. J. v., A. d. Jong, M. Montalban-Lopez, J. Kok and O. P. Kuipers (2013). "BAGEL3: automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides." Nucleic Acids Research **41**(W448 - W453).

Heier, B. T., A. Lovland, K. B. Soleim, M. Kaldhusal and J. Jarp (2001). "A Field Study of Naturally Occurring Specific Antibodies against *Clostridium perfringens* Alpha Toxin in Norwegian Broiler Flocks." Avian Diseases **45**(3): 724 - 732.

Helmboldt, C. F. and E. S. Bryant (1971). "The Pathology of Necrotic Enteritis in Domestic Fowl." Avian Diseases **15**(4): 775 - 780.

Hiett, K. L., N. A. Cox and M. J. R. Jr (2014). "Polymerase chain reaction detection of naturally occurring *Campylobacter* in commercial broiler chicken embryos " Poultry Science **92**(4): 1134 - 1137.

Hijum, S. A. F. T. v. and A. L. Zomer. "Perform Promoter Prediction on FASTA Formatted DNA Sequences." Retrieved 1/8/14, from http://bioinformatics.biol.rug.nl/websoftware/ppp.

Hols, P., T. Ferain, D. Garmyn, N. Bernard and J. Delcou (1994). "Use of Homologous Expression-Secretion Signals and Vector-Free Stable Chromosomal Integration in Engineering of *Lactobacillus plantarum* for α-Amylase and Levanase Expression." Applied and Environmental Microbiology **60**(5): 1401 - 1413.

Hood, A. M., A. D. Pearson and M. Shahamat (1988). "The extent of surface contamination of retailed chickens with *Campylobacter jejuni* serogroups." Epidemiology and Infection **100**: 17 - 25.

Humphrey, S., G. Chaloner, K. Kemmett, N. Davidson, N. Williams, A. Kipar, T. Humphrey and P. Wigley (2014). "*Campylobacter jejuni* Is Not Merely a Commensal in Commercial Broiler Chickens and Affects Bird Welfare." mBio **5**(4): e01364 - 01314.

Inoue, H., H. Nojima and H. Okayama (1990). "High efficiency transformation of Escherichia coli with plasmids." Gene **96**(1): 23 - 28.

Invitrogen (2010). SuperScript™ II Reverse Transcriptase. Invitrogen.

Jacobs-Reitsma, W. F. (1997). "Aspects of epidemiology of campylobacter in poultry." Veterinary Quarterly **19**: 113 - 117.

Janes, M., B. Meyhack, W. Zimmermann and A. Hinnen (1990). "The influence of GAP promoter variants on hirudin production, average plasmid copy number and cell growth in *Saccharomyces cerevisiae*." Current Genetics **18**(2): 97 - 103.

Jang, S. I., H. S. Lillehoj, S.-H. Lee, K. W. Lee, E. P. Lillehoj, Y. H. Hong, D.-J. An, W. Jeong, J.-E. Chun, F. Bertrand, L. Dupuis, S. Deville and J. B. Arous (2012). "Vaccination with *Clostridium perfringens* recombinant proteins in combination with MontanideTM ISA 71 VG adjuvant increases protection against experimental necrotic enteritis in commercial broiler chickens." Vaccine **30**: 5401 - 5406.

Jee, P.-F., V. Tiong, M.-H. Shu, J.-J. Khoo, W. F. Wong, R. A. Rahim, S. AbuBakar and L.-Y. Chang (2017). "Oral immunization of a non-recombinant *Lactococcus lactis* surface displaying influenza hemagglutinin 1 (HA1) induces mucosal immunity in mice." PLOS ONE **12**(11): 1 - 17.

Jia, Y., S. Li, G. Allen, S. Feng and L. Xue (2012). "A novel glyceraldehyde-3-phosphate dehydrogenase (GAPDH) promoter for expressing transgenes in the halotolerant alga Dunaliella salina." Current Microbiology **64**(5): 506- 513.

Jiménez, B., F. Ugwu, R. Zhao, L. Ortí, T. Makhnevych, A. Pineda-Lucena and W. A. Houry (2012). "Structure of Minimal Tetratricopeptide Repeat Domain Protein Tah1 Reveals Mechanism of Its Interaction with Pih1 and Hsp90." Journal of Biological Chemistry **287**(8): 5698 - 5709.

Johansson, M.-L., L. G. Molin, L. B. Jeppsson, S. Nobaek, S. Ahrne and S. Bengmark (1993). "Administration of Different *Lactobacillus* Strains in Fermented Oatmeal Soup: In Vivo Colonization of Human Intestinal Mucosa and Effect on the Indigenous Flora." Applied and Environmental Microbiology **59**(1): 15 - 20.

Kajikawa, A., E. Midorikawa, K. Masuda, K. Kondo, T. Irisawa, S. Igimi and S. Okada (2016). "Characterization of flagellins isolated froma highly motile strain of *Lactobacillus agilis*." BMC Microbiology **16**(49).

Kaldhusdal, M., M. Hofshagen, A. Løvland, H. Langstrand and K. Redhead (1999). "Necrotic enteritis challenge models with broiler chickens raised on litter: evaluation of preconditions,*Clostridium perfringens* strains and outcome variables." FEMS Immunology and Medical Microbiology **24**: 337 - 343.

Kaldhusdal, M. and A. Lovland (2000). "The economical impact of *Clostridium perfringens* is greater than anticipated." World Poultry **16**(8): 50 - 51.

Kaldhusdal, M., C. Schneitz, M. Hofshagen and E. Skjerve (2001). "Reduced Incidence of *Clostridium perfringens*-Associated Lesions and Improved Performance in Broiler Chickens Treated with Normal Intestinal Bacteria from Adult Fowl." Avian Diseases **45**(1): 149 - 156.

Kamke, J., S. Kittelmann, P. Soni, Y. Li, M. Tavendale, S. Ganesh, P. H. Janssen, W. Shi, J. Froula, E. M. Rubin and G. T. Attwood (2016). "Rumen metagenome and metatranscriptome analyses of low methane yield sheep reveals a Sharpea-enriched microbiome characterised by lactic acid formation and utilisation." Microbiome **4**(56): 1 - 16.

Kanojia, G., R. t. Have, P. C. Soema, H. Frijlink, J.-P. Amorij and G. Kerstena (2017). "Developments in the formulation and delivery of spray dried vaccines." Human Vaccines & Immunotherapeutics **13**(10): 2364 - 2378.

Kaufmann, A. F. and J. C. Feeley (1968). "Culture Survey of *Salmonella* at a Broiler-Raising Plant." Public Health Reports **83**(5): 417 - 422.

Keyburn, A. L., J. D. Boyce, P. Vaz, T. L. Bannam, M. E. Ford, D. Parker, A. D. Rubbo, J. I. Rood and R. J. Moore (2008). "NetB, a New Toxin That Is Associated with Avian Necrotic Enteritis Caused by *Clostridium perfringens*." PLoS Pathogens **4**(2): 1 - 11.

Keyburn, A. L., J. D. Boyce, P. Vaz, T. L. Bannam, M. E. Ford, D. Parker, A. D. Rubbo, J. I. Rood and R. J. Moore (2008). "NetB, a New Toxin That Is Associated with Avian Necrotic Enteritis Caused by *Clostridium perfringens*." PLoS Pathogens **4**(2).

Keyburn, A. L., R. W. Portela, M. E. Ford, T. L. Bannam, X. X. Yan, J. I. Rood and R. J. Moore (2013). "Maternal immunization with vaccines containing recombinant NetB toxin

partially protects progeny chickens from necrotic enteritis." Veterinary Research **44(1)**(108): 1 - 7.

Keyburn, A. L., R. W. Portela, K. Sproat, M. E. Ford, T. L. Bannam, X. Yan, J. I. Rood and R. J. Moore (2013). "Vaccination with recombinant NetB toxin partially protects broiler chickens from necrotic enteritis." Veterinary Research **44**(54): 1 - 8.

Keyburn, A. L., S. A. Sheedy, M. E. Ford, M. M. Williamson, M. M. Awad, J. I. Rood and R. J. Moore (2006). "Alpha-Toxin of *Clostridium perfringens* Is Not an Essential Virulence Factor in Necrotic Enteritis in Chickens." Infection and Immunity **74**(11): 6496 - 6500.

Khosravinia, H. (2009). "Effect of the slow (K) or rapid (k+) feathering gene on carcass-related traits of broiler chickens selected for breast and thighs weight." Russian Journal of Genetics **45**(1): 98 - 104.

Kiatpapan, P., M. Yamashita, N. Kawaraichi, T. Yasuda and Y. Murooka (2001). "Heterologous expression of a gene encoding cholesterol oxidase in probiotic strains of *Lactobacillus plantarum* and *Propionibacterium freudenreichii* under the control of native promoters." Journal of Bioscience and Bioengineering **92**(5): 459 - 465.

Kim, S. F., S. J. Baek and M. Y. Pack (1991). "Cloning and Nucleotide Sequence of the *Lactobacillus casei* Lactate Dehydrogenase Gene." Applied and Environmental Microbiology **57**(8): 2413 - 2417.

Kimura, N., F. Mimura, S. Nishida and A. Kobayashi (1976). "Studies on the Relationship between Intestinal Flora and Cecal Coccidiosis in Chicken." Poultry Science **55**: 1375 - 1383.

Kinder, J. M., J. E. Then, P. M. Hansel, L. L. Molinero and H. A. Bruns (2014). "Long-Term Repeated Daily Use of Intragastric Gavage Hinders Induction of Oral Tolerance to Ovalbumin in Mice." Comparative Medicine **64**(5): 369 - 376.

Kinoshita, H., S. Ohuchi, K. Arakawa, M. Watanabe, H. Kitazawa and T. Saito (2016). "Isolation of lactic acid bacteria bound to the porcine intestinal mucosa and an analysis of their moonlighting adhesins." Bioscience of Microbiota, Food and Health **35**(4): 185 - 196.

Kirk, M., L. Ford, K. Glass and G. Hall (2014). "Foodborne Illness, Australia, Circa 2000 and Circa 2010." Emerging Infectious Diseases **20**(11).

Klaenhammer, T. R. (1993). "Genetics of bacteriocins produced by lactic acid bacteria." FEMS Microbiology Reviews **12**(1 - 3): 39 - 85.

Klebe, R. J., J. V. Harriss, Z. D. Sharp and M. G. Douglas (1983). "A general method for polyethylene-glycol-induced genetic transformation of bacteria and yeast." Gene **25**: 333 - 341.

Kobierecka, P. A., B. Olech, M. Książek, K. Derlatka, I. Adamska, P. M. Majewski, E. K. Jagusztyn-Krynicka and A. K. Wyszyńska (2016). "Cell Wall Anchoring of the *Campylobacter* Antigens to *Lactococcus lactis*." Frontiers in Microbiology **18**(7): 1 - 18.

Koren, S., B. P. Walenz, K. Berlin, J. R. Miller and A. M. Phillippy (2016). "Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation." bioRxiv: 1 - 33.

Kowalczyk, K., J. Daiss, J. Halpern and T. F. Roth (1985). "Quantitation of maternal-fetal IgG transport in the chicken." Immunology **54**: 755 - 762.

Krivonogov, S. V. (1984). "The recF-dependent endonuclease from *Escherichia coli* K12. Formation and resolution of pBR322 DNA multimers." Molecular and General Genetics **196**(1): 105 - 109.

Kubasova, T., J. Matiasovicova, I. Rychlik and H. Juricova (2014). "Complete sequence of multidrug resistance p9134 plasmid and its variants including natural recombinant with the virulence plasmid of *Salmonella* serovar Typhimurium." Plasmid **76**: 8 - 14.

Kulkarni, R. R., V. R. Parreira, Y.-F. Jiang and J. F. Prescott (2010). "A Live Oral Recombinant *Salmonella enterica* Serovar Typhimurium Vaccine Expressing *Clostridium perfringens* Antigens Confers Protection against Necrotic Enteritis in Broiler Chickens." Clinical and Vaccine Immunology **17**(2): 205 - 214.

Kulkarni, R. R., V. R. Parreira, S. Sharif and J. F. Prescott (2006). "*Clostridium perfringens* Antigens Recognized by Broiler Chickens Immune to Necrotic Enteritis." Clinical and Vaccine Immunology **13**(12): 1358 - 1362.

Lacey, J. A., P. A. Johanesen, D. Lyras and R. J. Moore (2016). "Genomic diversity of necrotic enteritis-associated strains of *Clostridium perfringens*: a review." Avian Pathology **45**(3): 302 - 307.

Lackenbach, F. I. (1912). "Rabies and the Pasteur Treatment." California State Journal of Medicine **10**(3): 123 - 125.

Lai, G.-H., Y.-C. Lin, Y.-L. Tsai, Y.-Y. Lien, M.-K. Lin, H.-J. Chen, W.-T. Chang, J. T. C. Tzen and M.-S. Lee (2014). "High yield production of pigeon circovirus capsid protein in the *E. coli* by evaluating the key parameters needed for protein expression." BMC Veterinary Research **10**(115).

Lan, P. T. N., L. T. Binh and Y. Benno (2003). "Impact of two probiotic *Lactobacillus* strains feeding on fecal lactobacilli and weight gains in chicken." Journal of General Applied Microbiology **49**: 29 - 36.

Łaniewski, P., M. Kuczkowski, K. Chrząstek, A. Woźniak, A. Wyszyńska, A. Wieliczko and E. K. Jagusztyn-Krynicka (2014). "Evaluation of the immunogenicity of *Campylobacter jejuni* CjaA protein delivered by *Salmonella enterica* sv. Typhimurium strain with regulated delayed attenuation in chickens." World Journal of Microbiology and Biotechnology **30**(1): 281 - 292.

Layton, S. L., D. R. Kapczynski, S. Higgins, J. Higgins, A. D. Wolfenden, K. A. Liljebjelke, W. G. Bottje, D. Swayne, L. R. Berghman, Y. M. Kwon, B. M. Hargis and K. Cole (2009). "Vaccination of chickens with recombinant *Salmonella* expressing M2e and CD154 epitopes increases protection and decreases viral shedding after low pathogenic avian influenza challenge." Poultry Science **88**(11): 2244 - 2252.

Layton, S. L., M. J. Morgan, K. Cole, Y. M. Kwon, D. J. Donoghue, B. M. Hargis and N. R. Pumford (2011). "Evaluation of Salmonella-Vectored *Campylobacter* Peptide Epitopes for Reduction of *Campylobacter jejuni* in Broiler Chickens." Clinical and Vaccine Immunology **18**(3): 449 - 454.

Lee, K. W., H. S. Lillehoj, M. S. Park, S. I. Jang, G. D. Ritter, Y. H. Hong, W. Jeong, H. Y. Jeoung, D. J. An and E. P. Lillehoj (2012). "*Clostridium perfringens* a-Toxin and NetB Toxin Antibodies and Their Possible Role in Protection Against Necrotic Enteritis and Gangrenous Dermatitis in Broiler Chickens." Avian Diseases **56**: 230 - 233.

Ley, D. D. (1971). Hybridization of DNA. Methods in Microbiology. J. R. Norris and D. W. Ribbons. Orlando, Florida, USA, Elsevier. **5A:** 311 - 322.

Li, B., V. Ruotti, R. M. Stewart, J. A. Thomson and C. N. Dewey (2010). "RNA-Seq gene expression estimation with read mapping uncertainty." Bioinformatics **26**(4): 493 - 500.

Li, H. (2013). "Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM." arXiv:1303.3997 [q-bio.GN].

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and G. P. D. P. Subgroup (2009). "The Sequence Alignment/Map format and SAMtools." Bioinformatics **25**(16): 2078 - 2079.

Lillehoj, H. S., S. I. Jang, A. Panebra, E. P. Lillehoj, L. Dupuis, J. B. Arous, S. K. Lee and S. T. Oh (2017). "In ovo vaccination using Eimeria profilin and *Clostridium perfringens* NetB proteins in Montanide IMS adjuvant increases protective immunity against experimentally-induced necrotic enteritis." Asian-Australasian Journal of Animal Sciences **30**(10): 1478 - 1485.

Lin, R., Y. Zhang, B. Long, Y. Li, Y. Wu, S. Duan, B. Zhu, X. Wu and H. Fan (2017). "Oral Immunization with Recombinant *Lactobacillus acidophilus* Expressing espA-Tir-M Confers Protection against Enterohemorrhagic *Escherichia coli* O157:H7 Challenge in Mice." Frontiers in Microbiology **8**: 1 - 9.

Lizier, M., P. G. Sarra, R. Cauda and F. Lucchini (2010). "Comparison of expression vectors in *Lactobacillus reuteri* strains." FEMS Microbiology Letters **308**: 8 - 15.

Lohans, C. T., M. J. v. Belkum, S. A. Cochrane, Z. Huang, C. S. Sit, L. M. McMullen and J. C. Vederas (2014). "Biochemical, Structural, and Genetic Characterization of Tridecaptin A1, an Antagonist of *Campylobacter jejuni*." ChemBioChem **15**(2): 243 - 249.

Lu, J., U. Idris, B. Harmon, C. Hofacre, J. J. Maurer and M. D. Lee (2003). "Diversity and Succession of the Intestinal Bacterial Community of the Maturing Broiler Chicken." Applied and Environmental Microbiology **69**(11): 6816 - 6824.

Lu, M. and U. Henning (1989). "The Immunity (imm) Gene of Escherichia coli Bacteriophage T4." Journal of Virology **63**(8): 3472 - 3478.

Lyras, D. and J. I. Rood (1998). "Conjugative Transfer of RP4-oriT Shuttle Vectors from *Escherichia coli* to *Clostridium perfringens*." Plasmid **39**(2): 160 -164.

Lyras, D., C. Storie, A. S. Huggins, P. K. Crellin, T. L. Bannam and J. I. Rood (1998). "Chloramphenicol Resistance in *Clostridium difficile* Is Encoded on Tn4453 Transposons That Are Closely Related to Tn4451 from Clostridium perfringens." Antimicrobial Agents and Chemotherapy **42**(7): 1563 - 1567.

Mabelebele, M., O. J. Alabi, J. W. Ng`ambi, D. Norris and M. M. Ginindza (2014). "Comparison of Gastrointestinal Tracts and pH Values of Digestive Organs of Ross 308 Broiler and Indigenous Venda Chickens Fed the Same Diet." Asian Journal of Animal and Veterinary Advances **9**: 71 - 76.

Mangell, P., P. Lennernas, M. Wang, C. Olsson, S. Ahrne, G. Molin, H. Thorlacius and B. Jeppsson (2006). "Adhesive capability of *Lactobacillus plantarum* 299v is important for preventing bacterial translocation in endotoxemic rats." APMIS **114**: 611 - 618.

Martin, J., W. Zhu, K. D. Passalacqua, N. Bergman and M. Borodovsky (2010). "*Bacillus anthracis* genome organization in light of whole transcriptome sequencing." BMC Bioinformatics **11 (Suppl 3)**(S10): 1 - 13.

Mason, C. K., M. A. Collins and K. Thompson (2005). "Modified electroporation protocol for lactobacilli isolated from the chicken crop facilitates transformation and the use of a genetic tool." Journal of Microbiological Methods **60**: 353 - 363.

Mcdevitt, R. M., J. D. Brooker, T. Acamovic and N. H. C. Sparks (2006). "Necrotic enteritis; a continuing challenge for the poultry industry." World's Poultry Science Journal **62**(221 - 247).

Meeusen, E. N. T., J. Walker, A. Peters, P.-P. Pastoret and G. Jungersen (2007). "Current Status of Veterinary Vaccines." Clinical Microbiology Reviews **20**(3): 489 - 510.

Meimandipour, A., M. Shuhaimi, M. Hair-Bejo, K. Azhar, B. M. Kabeir, B. Rasti and A. M. Yazid (2009). "In vitro fermentation of broiler cecal content: the role of lactobacilli and pH value on the composition of microbiota and end products fermentation." Letters in Applied Microbiology **49**: 415 - 420.

Mekonnen, M. M. and A. Y. Hoekstra (2012). "A Global Assessment of the Water Footprint of Farm Animal Products." Ecosystems **15**: 401 - 415.

Melo, R. T., P. C. Nalevaiko, E. P. Mendonça, L. W.Borges, B. B. Fonseca, M. E. Beletti and D. A. Rossi (2013). "*Campylobacter jejuni* strains isolated from chicken meat harbour several virulence factors and represent a potential risk to humans." Food Control **33**(1): 227 - 231.

Mergulhão, F. J. M., D. K. Summers and G. A. Monteiro (2005). "Recombinant protein secretion in *Escherichia coli*." Biotechnology Advances **23**(3): 177 - 202.

Meselson, M. and R. Yuan (1968). "DNA Restriction Enzyme from *E. coli*." Nature **217**: 1110 - 1114.

Messaoudi, S., G. Kergourlaya, M. Dalgalarrondo, Y. Choiset, M. Ferchichi, H. Prévost, M.-F. Pilet, J.-M. Chobert, M. Manai and X. Dousset (2012). "Purification and characterization of a new bacteriocin active against *Campylobacter* produced by *Lactobacillus salivarius* SMXD51." Food Microbiology **32**: 129 - 134.

Messaoudi, S., A. Madi, H. Prévost, M. Feuilloley, M. Manai, X. Dousset and N. Connil (2012). "*In vitro* evaluation of the probiotic potential of *Lactobacillus salivarius* SMXD51." Anaerobe **18**: 584 - 589.

Methner, U., P. A. Barrow, A. Berndt and G. Steinbach (1999). "Combination of vaccination and competitive exclusion to prevent *Salmonella* colonization in chickens: experimental studies." International Journal of Food Microbiology **49**: 35 - 42.

Millman, J. M., K. Waits, H. Grande, A. R. Marks, J. C. Marks, L. B. Price and B. A. Hungate (2013). "Prevalence of antibiotic-resistant *E. coli* in retail chicken: comparing conventional, organic, kosher, and raised without antibiotics." F1000Research **2**(155).

Milne, I., G. Stephen, M. Bayer, P. J.A.Cock, L. Pritchard, L. Cardle, P. D. Shawand and D. Marshall (2013). "Using Tablet for visual exploration of second-generation sequencing data." Briefings in Bioinformatics.

Mishra, N. and J. A. Smyth (2017). "Oral vaccination of broiler chickens against necrotic enteritis using a non-virulent NetB positive strain of *Clostridium perfringens* type A." Vaccine: 1 - 8.

Mitsuoka, T. (2002). "Research in Intestinal Flora and Functional Foods." Journal of Intestinal Microbiology **15**(2): 57 - 89.

Molin, L. G., B. Jeppsson, M.-L. Johansson, S. N. S. Ahrne, M. Stahl and S. Bengmark (1993). "Numerical taxonomy of *Lactobacillus* spp. associated with healthy and diseased mucosa of the human intestines." Journal of Applied Bacteriology **74**: 314 - 323.

Moric, I., M. Savic, T. Ilic-Tomic, S. Vojnovic, S. Bajkic and B. Vasiljevic (2010). "rRNA Methyltransferases And Their Role In Resistance To Antibiotics." Journal of Medical Biochemistry **29**: 165 - 174.

Morita, H., C. Shiratori, M. Murakami, H. Takami, Y. Kato, A. Endo, F. Nakajima, M. Takagi, H. Akita, S. Okada and T. Masaoka (2007). "*Lactobacillus hayakitensis* sp. nov., isolated from intestines of healthy thoroughbreds." International Journal of Systematic and Evolutionary Microbiology **57**: 2836 - 2839.

Morona, J. K., R. Morona and J. C. Paton (2006). "Attachment of capsular polysaccharide to the cell wall of *Streptococcus pneumoniae* type 2 is required for invasive disease." Proceedings of the National Academy of Sciences of the United States of America **97**(12): 5225 - 5243.

Mot, D., L. Timbermont, F. Haesebrouck, R. Ducatelle and F. V. Immerseel (2014). "Progress and problems in vaccination against necrotic enteritis in broiler chickens." Avian Pathology **43**(4): 290 - 300.

Mukherjee, A., K. B. Weyant, U. Agrawal, J. Walker, I. K. O. Cann and C. M. Schroeder (2015). "Engineering and Characterization of New LOV-Based Fluorescent Proteins from *Chlamydomonas reinhardtii* and *Vaucheria frigida*." Synthetic Biology **4**: 371 - 377.

Mulder, N.-D., C. Pan and W. Sawyer (2013). Rabobank Poultry Quarterly Q4 2013. Rabobank Poultry Quarterly Rabobank International.

Murphya, P., F. D. Bello, J. O'Doherty, E. K. Arendt, T. Sweeney and A. Coffey (2013). "The effects of liquid versus spray-dried *Laminaria digitata* extract on selected bacterial groups in the piglet gastrointestinal tract (GIT) microbiota." Anaerobe **21**: 1 - 8.

Murray, I. A., T. A. Clark, R. D. Morgan, M. Boitano, B. P. Anton, K. Luong, A. Fomenkov, S. W. Turner, J. Korlach and R. J. Roberts (2012). "The methylomes of six bacteria." Nucleic Acids Research **40**(22).

Myszewski, M. A. and N. J. Stern (1990). "Influence of *Campylobacter jejuni* Cecal Colonization on Immunoglobulin Response in Chickens." Avian Diseases **34**(3): 588 - 594.

Naito, T., K. Kusano and I. Kobayashi (1995). "Selfish Behavior of Restriction-Modification Systems." Science **267**: 897 - 899.

Nan, J., Y. Zhou, C. Yang, E. Brostromer, O. Kristensenc and X.-D. Su (2009). "Structure of a fatty acid-binding protein from *Bacillus subtilis* determined by sulfur-SAD phasing using in-house chromium radiation." Acta Crystallographica Section D **65**(5): 440 - 448.

Naqi, S. A., B. Marquez and N. Sahin (1983). "Maternal antibody and its effect on infectious bursal disease immunization." Avian Diseases **27**(3): 623 - 631.

Nariya, H., S. Miyata, E. Tamai, H. Sekiya, J. Maki and A. Okabe (2011). "Identification and characterization of a putative endolysin encoded by episomal phage phiSM101 of *Clostridium perfringens*." Applied Genetics and Molecular Biotechnology **90**: 1973 - 1979.

Netherwood, T., H. J. Gilbert, D. S. Parker and A. G. O'Donnell (1999). "Probiotics Shown To Change Bacterial Community Structure in the Avian Gastrointestinal Tract." Applied and Environmental Microbiology **65**(11): 5134 - 5138.

Newell, D. G. and J. A. Wagenaar (2000). Poultry infections and their control at the farm level. Campylobacter. I. Nachamkin and M. J. Blaser. Washington D.C., ASM Press**:** 497 - 509.

Niederstadt, L., O. Hohn, B. G. Dorner, R. Schade and N. Bannert (2012). "Stimulation of IgY responses in gene gun immunized laying hens by combined administration of vector DNA coding for the target antigen Botulinum toxin A1 and for avian cytokine adjuvants." Journal of Immunological Methods **382**: 58 - 67.

Nishimichi, N., M. Aosasa, T. Kawashima, H. Horiuchi, S. Furusawa and H. Matsuda (2005). "Biological activity of recombinant chicken interleukin-6 in chicken hybridoma cells." Veterinary Immunology and Immunopathology **106**: 97 - 105.

Novick, R. P. (1987). "Plasmid Incompatibility." Microbiological Reviews **51**(4): 381 - 395.

Nurmi, E. and M. Rantala (1973). "New aspects of *Salmonella* infection in broiler chickens." Nature **241**: 210 - 211.

O'Sullivan, D. J. and T. R. Klaenhammer (1993). "High- and low-copy-number *Lactococcus* shuttle cloning vectors with features for clone screening." Gene **137**: 227 - 231.

OECD/FAO (2017). "OECD-FAO Agricultural Outlook 2017-2026". Paris.

Onodera, O., A. D. Roses, S. Tsuji, J. M. Vance, W. J. Strittmatter and J. R. Burke (1996). "Toxicity of expanded polyglutamine-domain proteins in *Escherichia coli* " FEBS Letters **399**(1): 135 - 139

Overbeek, R., T. Begley, R. M. Butler, J. V. Choudhuri, H.-Y. Chuang, M. Cohoon, V. d. Crécy-Lagard, N. Diaz, T. Disz, R. Edwards, M. Fonstein, E. D. Frank, S. Gerdes, E. M. Glass, A. Goesmann, A. Hanson, D. Iwata-Reuyl, R. Jensen, N. Jamshidi, L. Krause, M.

Kubal, N. Larsen, B. Linke, A. C. McHardy, F. Meyer, H. Neuweger, G. Olsen, R. Olson, A. Osterman, V. Portnoy, G. D. Pusch, D. A. Rodionov, C. Rückert, J. Steiner, R. Stevens, I. Thiele, O. Vassieva, Y. Ye, O. Zagnitko and V. Vonstein (2005). "The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes." Nucleic Acids Research **33**(17): 5691 - 5702.

Overbeek, R., R. Olson, P. GD, G. Olsen, J. Davis, D. T, E. RA, G. S, P. B, M. Shukla, V. Vonstein, A. Wattam, F. Xia and R. Stevens (2014). "The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST)." Nucleic Acids Research **42**.

OzFoodNetWorkingGroup (2015). "Monitoring the incidence and causes of diseases potentially transmitted by food in Australia: Annual report of the OzFoodNet network." Communicable Diseases Intelligence **39**(2): 326 - 264.

PacificBiosciences. (2014). "MotifMaker." MotifMaker Retrieved 1/9/16, from https://github.com/PacificBiosciences/MotifMaker.

PacificBioSciences. (2015). "Pitchfork." Retrieved 10/6/16, from https://github.com/PacificBiosciences/pitchfork.

PacificBiosciences. (2016). "kineticsTools." Retrieved 1/9/16, from https://github.com/PacificBiosciences/kineticsTools.

Palombo, E. A., K. Yusoff, V. A. Stanisich, V. Krishnapillai and N. S. Willetts (1989). "Cloning and genetic analysis of tra cistrons of the Tra 2/Tra 3 region of plasmid RP1." Plasmid **22**(1): 59 - 69.

Palop, M. L., J. P. Smiths and B. t. Brink (1995). "Degradation of sinigrin by *Lactobacillus agilis* strain R16." International Journal of Food Microbiology **26**: 219 - 229.

Pawelec, D. P., D. Korsak, A. K. Wyszynska, E. Rozçynek, J. Popowski and E. K. Jagusztyn-Krynicka (2000). "Genetic diversity of the *Campylobacter* genes coding immunodominant proteins." FEMS Microbiology Letters **185**: 43 - 49.

Petersen, J., H. Brinkmann and R. Cerf (2002). "Origin, Evolution, and Metabolic Role of a Novel Glycolytic GAPDH Enzyme Recruited by Land Plant Plastids." Journal of Molecular Evolution **57**: 16 - 26.

Pike, J. E., R. A. Henry, P. M. J. Burgers, J. L. Campbell and R. A. Bambara (2010). "An Alternative Pathway for Okazaki Fragment Processing: Resolution Of Fold-Back Flaps By Pif1 Helicase." Journal of Biological Chemistry **285**(53): 41712 - 41712.

Pouwels, P. H. and R. J. Leer (1993). "Genetics of lactobacilli: plasmids and gene expression " Antonie van Leeuwenhoek **64**: 85 - 107.

Power, M. E., P. Guerry, W. D. Mccubbin, C. M. Kay and T. J. Trust (1994). "Structural and Antigenic Characteristics of *Campylobacter coli* FlaA Flagellin." Journal of Bacteriology **176**(11): 3303 - 3313.

Purves, J., A. Cockayne, P. C. E. Moody and J. A. Morrisse (2010). "Comparison of the Regulation, Metabolic Functions, and Roles in Virulence of the Glyceraldehyde-3-Phosphate Dehydrogenase Homologues gapA and gapB in *Staphylococcus aureus*." Infection and Immunity **78**(12): 5223 - 5232.

Ragione, R. M. L., M. J. Woodward, M. Kumar, J. Rodenberg, H. Fan, A. D. Wales and K. Karaca (2013). "Efficacy of a Live Attenuated *Escherichia coli* O78:K80 Vaccine in Chickens and Turkeys." Avian Diseases **57**: 273 - 279.

Raha, A. R., N. R. S. Varma, K. Yusoff, E. Ross and H. L. Foo (2005). "Cell surface display system for *Lactococcus lactis*: a novel development for oral vaccine." Applied Genetics And Molecular Biotechnology **68**: 75 - 81.

Rantala, M. and E. Nurmi (1973). "Prevention of the growth of *Salmonella infantis* in chicks by the flora of the alimentary tract of chickens." Poultry Science **14**: 627 - 630.

Raveh, A., M. Margaliot, E. D. Sontag and T. Tuller (2016). "A model for competition for ribosomes in the cell." Journal of The Royal Society Interface **13**(116): 1 - 17.

Rawal, P., V. B. R. Kummarasetti, J. Ravindran, N. Kumar, K. Halder, R. Sharma, M. Mukerji, S. K. Das and S. Chowdhury (2006). "Genome-wide prediction of G4 DNA as regulatory motifs: Role in *Escherichia coli* global regulation." Genome Research **16**(5): 644 - 655.

Reveneau, N., M.-C. Geoffroy, C. Locht, P. Chagnaud and A. Mercenier (2002). "Comparison of the immune responses induced by local immunizations with recombinant *Lactobacillus plantarum* producing tetanus toxin fragment C in different cellular locations." Vaccine **20**: 1769 – 1777.

Rice, B. E., D. M. Rollins, E. T. Mallinson, L. Carr and S. W. Joseph (1997). "*Campylobacter jejuni* in broiler chickens: colonization and humoral immunity following oral vaccination and experimental infection." Vaccine **15**(17): 1922 - 1932.

Riddell, C. and X.-M. Kong (1992). "The Influence of Diet on Necrotic Enteritis in Broiler Chickens." Avian Diseases **36**(6): 499 - 503.

Roberts, A. P., P. A. Johanesen, D. Lyras, P. Mullany and J. I. Rood (2001). "Comparison of Tn5397 from *Clostridium difficile*, Tn916 from *Enterococcus faecalis* and the CW459tet(M) element from *Clostridium perfringens* shows that they have similar conjugation regions but different insertion and excision modules." Microbiology **147**: 1243 - 1251.

Roberts, R. J. and S. E. Halford (1993). Type II Restriction Enzymes, Cold Spring Harbor Laboratory Press.

Robinson, D. A. (1981). "Infective dose of Campylobacter jejuni in milk." British Medical Journal **282**: 1584.

Rocap, G. "Genbank/EMBL to FASTA Conversion Tool." Retrieved 1/7/16, from http://rocaplab.ocean.washington.edu/tools/genbank_to_fasta.

Rosano, G. L. and E. A. Ceccarelli (2014). "Recombinant protein expression in *Escherichia coli*: advances and challenges." Frontiers in Microbiology **5**(172): 1- 17.

Rosenquist, H., N. L. Nielsen, H. M. Sommer, B. Nørrung and B. B. Christensen (2003). "Quantitative risk assessment of human campylobacteriosis associated with thermophilic Campylobacter species in chickens." International Journal of Food Microbiology **83**(1): 87 - 103.

Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream and B. Barrell (2000). "Artemis: sequence visualization and annotation." Bioinformatics **16**(10): 944 - 945.

Sadeyen, J.-R., Z. Wu, H. Davies, P. M. v. Diemen, A. Milicic, R. M. L. Ragione, P. Kaiser, M. P. Stevens and F. Dziva (2015). "Immune responses associated with homologous protection conferred by commercial vaccines for control of avian pathogenic *Escherichia coli* in turkeys." Veterinary Research **46**(1): 1 - 14.

Sadusk, J. F. and H. Kuhlenbeck (1946). "Dangers Associated with the Use of Living "Attenuated" Typhus Vaccine." American Journal of Public Health **3**: 1027 - 1030.

Saengkerdsub, S., P. Herrera, C. L. Woodward, R. C. Anderson, D. J. Nisbet and S. C. Ricke1 (2007). "Detection of methane and quantification of methanogenic archaea in faeces from young broiler chickens using real-time PCR." Letters in Applied Microbiology **45**: 629 - 634.

Sahin, O., N. Luo, S. Huang and Q. Zhang (2003). "Effect of *Campylobacter*-Specific Maternal Antibodies on *Campylobacter jejuni* Colonization in Young Chickens." Applied and Environmental Microbiology **69**(9): 5372 - 5379.

Sahin, O., Q. Zhang, J. C. Meitzler, B. S. Harr, T. Y. Morishita and R. Mohan (2001). "Prevalence, Antigenic Specificity, and Bactericidal Activity of Poultry Anti-Campylobacter Maternal Antibodies." Applied and Environmental Microbiology **67**(9): 3951 – 3957.

Saleh, N., S. I. Fathalla, R. Nabil and A. A. Mosaad (2011). "Clinicopathological and immunological studies on toxoids vaccine as a successful alternative in controlling clostridial infection in broilers." Anaerobe **17**: 426 - 430.

Salem, R. B., M. S. Abbassi, V. García, R. García-Fierro, J. Fernández, H. Kilani, I. Jaouani, M. Khayeche, L. Messadi and M. R. Rodicio (2017). "Antimicrobial drug resistance and genetic properties of *Salmonella enterica* serotype Enteritidis circulating in chicken farms in Tunisia." Journal of Infection and Public Health **10**(6): 855 - 860.

Schlegel, S., E. Rujas, A. J. Ytterberg, R. A. Zubarev, J. Luirink and J.-W. d. Gier (2013). "Optimizing heterologous protein production in the periplasm of *E. coli* by regulating gene expression levels." Microbial Cell Factories **12**(24): 1 - 12.

Schoeni, J. L. and A. C. L. Wong (1994). "Inhibition of *Campylobacter jejuni* Colonization in Chicks by Defined Competitive Exclusion Bacteria." Applied and Environmental Microbiology **60**(4): 1191 - 1197.

Sheedy, S. A. (2006). Development of a live bacterial vector to deliver prophylactic proteins to chickens. Doctor of Philosophy, Monash University.

Shike, D. W. (2013). Beef Cattle Feed Efficiency. Driftless Region Beef Conference.

Shin, M. S., S. K. Han, A. R. Ji, K. S. Kim and W. K. Lee (2008). "Isolation and characterization of bacteriocin-producing bacteria from the gastrointestinal tract of broiler chickens for probiotic use." Journal of Applied Microbiology **105**: 2203 - 2212.

Shine, J. and L. Dalgarno (1974). "The 3'-Terminal Sequence of *Escherichia coli* 16S Ribosomal RNA: Complementarity to Nonsense Triplets and Ribosome Binding Sites." Proceedings of the National Academy of Sciences **71**(4): 1342 - 1346.

Siegl, T., B. Tokovenko, M. Myronovskyi and A. Luzhetskyy (2013). "Design, construction and characterisation of a synthetic promoter library for fine-tuned gene expression in actinomycetes." Metabolic Engineering **19**: 98 - 106.

Silva, N. A. D. and J. E. Baile (1991). "Influence of plasmid origin and promoter strength in fermentations of recombinant yeast." Biotechnology and Bioengineering **37**(4): 318 - 324

Simon, D., A. Rouault and M.-C. Chopin (1985). "Protoplast transformation of group N streptococci with cryptic plasmids " FEMS Microbiology Letters **26**: 239 - 241

Simon, R., U. Priefer and A. Pühler (1983). "A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria." Nature Biotechnology **1**: 784 - 791.

Sitaraman, R. (2016). "The Role of DNA Restriction-Modification Systems in the Biology of *Bacillus anthracis*." Frontiers in Microbiology **7**(11).

Sletta, H., A. Nedal, T. E. V. Aune, H. Hellebust, S. Hakvåg, R. Aune, T. E. Ellingsen, S. Valla and T. Brautaset (2004). "Broad-Host-Range Plasmid pJB658 Can Be Used for

Industrial-Level Production of a Secreted Host-Toxic Single-Chain Antibody Fragment in *Escherichia coli*." Applied and Environmental Microbiology **70**(12): 7033 - 7039.

Song, L., H. Cui, L. Tang, X. Qiao, M. Liu, Y. Jiang, W. Cui and Y. Li (2014). "Construction of upp Deletion Mutant Strains of Lactobacillus casei and Lactococcus lactis based on Counterselective System Using Temperature-Sensitive Plasmid." Journal of Microbiological Methods **STILL IN PRINT**.

Spath, K., S. Heinl and R. Grabherr (2012). "Direct cloning in *Lactobacillus plantarum*: Electroporation with non-methylated plasmid DNA enhances transformation efficiency and makes shuttle vectors obsolete." Microbial Cell Factories **11**(141).

Sridhar, V. R., V. V. Smeianov and J. L. Steele (2006). "Construction and evaluation of food-grade vectors for *Lactococcus lactis* using aspartate aminotransferase and alpha-galactosidase as selectable markers." Journal of Applied Microbiology **101**(1): 161 - 171.

Stangegaard, M., I. H. Dufva and M. Dufva (2006). "Reverse transcription using random pentadecamer primers increases yield and quality of resulting cDNA." BioTechniques **40**(5): 649 - 656.

Stanley, D., M. S. Geier, R. J. Hughes, S. E. Denman and R. J. Moore (2013). "Highly variable microbiota development in the chicken gastrointestinal tract." PLOS ONE **8**(12): 1 - 7.

Stassi, D. L. and S. A. Lacks (1982). "Effect of strong promoters on the cloning in *Escherichia coli* of DNA fragments from *Streptococcus pneumoniae* " Gene **18**(3): 319 - 328.

Stephenson, D. P. (2010). Microbial ecology of *Lactobacillus* strains in the chicken gastrointestinal tract: identification of persistant strains for use as live vectors. Doctor of Philosophy, Australia National University.

Stephenson, D. P., R. J. Moore and G. E. Allison (2010). "*Lactobacillus* Strain Ecology and Persistence within Broiler Chickens Fed Different Diets: Identification of Persistent Strains." Applied and Environmental Microbiology **76**(19): 6494 - 6503.

Stephenson, D. P., R. J. Moore and G. E. Allison (2011). "Transformation of, and Heterologous Protein Expression in, *Lactobacillus agilis* and *Lactobacillus vaginalis* Isolates from the Chicken Gastrointestinal Tract." Applied and Environmental Microbiology **77**(1): 220 - 228.

Stern, N. J. (1992). Reservoirs for *C. jejuni* and approaches for intervention in poultry. *Campylobacter jejuni*: current status and future trends. I. Nachamkin, M. J. Blaser and L. S. Tompkins. Washington, D.C., American Society for Microbiology**:** 49–60.

Sun, Z., H. M. B. Harris, A. McCann, C. Guo, S. Argimon, W. Zhang, X. Yang, I. B. Jeffery, J. C. Cooney, T. F. Kagawa, W. Liu, Y. Song, E. Salvetti, A. Wrobel, P. Rasinkangas, J. Parkhill, M. C. Rea, O. O'Sullivan, J. Ritari, F. o. P. Douillard, R. P. Ross, R. Yang, A. E. Briner, G. E. Felis, W. M. d. Vos, R. Barrangou, T. R. Klaenhammer, P. W. Caufield, Y. Cui, H. Zhang and P. W. O'Toole (2015). "Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated genera." Nature Communications **29**(6).

Tanweer, A. J., N. Chand, U. Saddique, C. A. Bailey and R. U. Khan (2014). "Antiparasitic effect of wild rue (*Peganum harmala* L.) against experimentally induced coccidiosis in broiler chicks." Parasitology Research.

Tatusova, T., M. DiCuccio, A. Badretdin, V. Chetvernin, E. P. Nawrocki, L. Zaslavsky, A. Lomsadze, K. D. Pruitt, M. Borodovsky and J. Ostell (2016). "NCBI prokaryotic genome annotation pipeline." Nucleic Acids Research **44**(14): 6614 - 6624.

Teirlynck, E., M. D. E. Gussem, J. Dewulf, F. Haesebrouck, R. Ducatelle and F. V. Immerseel (2011). "Morphometric evaluation of "dysbacteriosis" in broilers." Avian Pathology **40**(2): 139 - 144.

Timbermont, L., F. Haesebrouck, R. Ducatelle and F. V. Immerseel (2011). "Necrotic enteritis in broilers: an updated review on the pathogenesis." Avian Pathology **40**(4): 341 - 347.

Totte, P., J. Nyanjui, A. Bensaid and D. McKeever (1999). "Bovine CD4+ T-cell lines reactive with soluble and membrane antigens of Cowdria ruminantium." Veterinary Immunology and Immunopathology **70**: 269 - 276.

UN (2013). "World Population Prospects: The 2012 Revision". P. D. Department of Economic and Social Affairs, United Nations.

USCB (2013). "Total midyear population for the world: 1950-2050".

van Asten, A. J. and J. E. van Dijk (2005). "Distribution of "classic" virulence factors among *Salmonella* spp." FEMS Immunology and Medical Microbiology **44**(3): 251 - 259.

van der Wielen, P. W. J. J., D. A. Keuzenkamp, L. J. A. Lipman, F. v. Knapen and S. Biesterveld (2002). "Spatial and Temporal Variation of the Intestinal Bacterial Community in Commercially Raised Broiler Chickens During Growth." Microbial Ecology **44**: 286 - 293.

Van Immerseel, F., U. Methner, I. Rychlik, B. Nagy, P. Velge, G. Martin, N. Foster, R. Ducatelle and P. A. Barrow (2005). "Vaccination and early protection against non-host-specific *Salmonella* serotypes in poultry: exploitation of innate immunity and microbial activity." Epidemiology and Infection **133**: 959 - 978.

Van Roekel, H. (1964). Proceedings of the National Conference on Salmonellosis. National Conference on Salmonellosis, Atlanta, U.S. Department of Health, Education and Welfare Public Health Service, Communicable Disease Center.

van Zyl, W. F., S. M. Deane and L. M. T. Dicks (2015). "Use of the mCherry Fluorescent Protein To Study Intestinal Colonization by *Enterococcus mundtii* ST4SA and *Lactobacillus plantarum* 423 in Mice." Applied and Environmental Microbiology **81**(17): 5993 - 6002.

Vezina, B. D. (2013). Construction of a shuttle plasmid for live vaccine vectoring. Bachelor of Health and Medical Science (Honours), Deakin University.

Vos, W. M. d., G. Venema, U. Canosi and T. A. Trautner (1981). "Plasmid transformation in *Bacillus subtilis*: Fate of plasmid DNA." Molecular and General Genetics **181**(4): 424 - 433.

Wade, B. and A. Keyburn. (2016). "The true cost of necrotic enteritis." Meat Retrieved 22/11/2017, from http://www.poultryworld.net/Meat/Articles/2015/10/The-true-cost-of-necrotic-enteritis-2699819W/#comments.

Wagenaar, J. A., M. A. P. Van Bergenm, M. A. Mueller, T. M. Wassenaar and R. M. Carlton (2005). "Phage therapy reduces *Campylobacter jejuni* colonization in broilers." Veterinary Microbiology **109**: 275 - 283.

Wagner, G. P., K. Kin and V. J. Lynch (2012). "Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples." Theory in Biosciences **131**(4): 281 - 285.

Wagner, G. P., K. Kin and V. J. Lynch (2013). "A model based criterion for gene expression calls using RNA-seq data." Theory in Biosciences **132**: 159 - 164.

Walk, C. L., T. T. Santos and M. R. Bedford (2014). "Influence of superdoses of a novel microbial phytase on growth performance, tibia ash, and gizzard phytate and inositol in young broilers." Poultry Science **93**(5): 1172 - 1177.

Walker, B. J., T. Abeel, T. Shea, M. Priest, A. Abouelliel, S. Sakthikumar, C. A. Cuomo, Q. Zeng, J. Wortman, S. K. Young and A. M. Earl (2014). "Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement." PLOS ONE **9**(11): 1 - 14.

Wang, Z., M. Gerstein and M. Snyder (2009). "RNA-Seq: a revolutionary tool for transcriptomics." Nature Reviews **10**: 57 - 63.

Wassenaar, T. M., B. A. M. van der Zeijst, R. Ayling and D. G. Newell (1993). "Colonization of chicks by motility mutants of Campylobacter jejuni demonstrates the importance of flagellin A expression

" Journal of General Microbiology **139**: 1171 - 1175.

Webste, L. T. (1922). "Experiments on normal and immune mice with a *Bacillus* of mouse typhoid." The Journal of Experimental Medicine **36**(1): 71 - 96.

Wegmann, U., N. Horn and S. R. Carding (2013). "Defining the *Bacteroides* Ribosomal Binding Site." Applied and Environmental Microbiology **79**(6): 1980 - 1989.

Weiss, N., U. Schillinger, M. Laternser and O. Kandler (1981). "*Lactobacillus sharpeae* sp.nov. and *Lactobacillus agilis* sp.nov., Two New Species of Homofermentative, meso-Diaminopimelic Acid-Containing lactobacilli Isolated from Sewage." Zentralblatt für Bakteriologie, Mikrobiologie und Hygiene **2**(3): 242 - 253.

Wheeler, D. L., T. Barrett, D. A. Benson, S. H. Bryant, K. Canese, V. Chetvernin, D. M. Church, M. DiCuccio, S. F. Ron Edgar, L. Y. Geer, Y. Kapustin, O. Khovayko, D. Landsman, D. J. Lipman, T. L. Madden, D. R. Maglott, J. Ostell, V. Miller, K. D. Pruitt, G. D. Schuler, E. Sequeira, S. T. Sherry, K. Sirotkin, A. Souvorov, G. Starchenko, R. L. Tatusov, T. A. Tatusova, L. Wagner and E. Yaschenko (2006). "Database resources of the National Center for Biotechnology Information." Nucleic Acids Research **35**(D5–D12): 1 - 8.

Wijewanta, E. A. and P. Seneviratna (1971). "Bacteriological studies of fatal *Clostridium perfringens* type-a infection in chickens." Avian Diseases **15**(4): 654 - 661.

Williams, R. B. (2005). "Intercurrent coccidiosis and necrotic enteritis of chickens: rational, integrated disease management by maintenance of gut integrity." Avian Pathology **34**(3): 159 - 180.

Wong, W. Y., P. Su, G. E. Allison, C.-Q. Liu and N. W. Dunn (2003). "A Potential Food-Grade Cloning Vector for *Streptococcus thermophilus* That Uses Cadmium Resistance as the Selectable Marker." Applied and Environmental Microbiology **69**(10): 5767 - 5771.

Xu, Q. Q., H. Yan, X. L. Liu, L. Lv, C. H. Yin and P. Wang (2014). "Growth performance and meat quality of broiler chickens supplemented with *Rhodopseudomonas palustris* in drinking water." British Poultry Science.

Xu, Y., X. Guan, Z. Liu, C. Tian and L. Cui (2015). "Immunogenicity in Swine of Orally Administered Recombinant *Lactobacillus plantarum* Expressing Classical Swine Fever Virus E2 Protein in Conjunction with Thymosin -1 as an Adjuvant." Applied and Environmental Microbiology **81**(11): 3745 - 3752.

Yan, X.-X., C. J. Porter, S. P. Hardy, D. Steer, A. I. Smith, N. S. Quinsey, V. Hughes, J. K. Cheung, A. L. Keyburn, M. Kaldhusdal, R. J. Moore, T. L. Bannam, J. C. Whisstock and J. I. Rooda (2013). "Structural and functional analysis of the pore-forming toxin NetB from *Clostridium perfringens*." mBio **4**(1): 1 - 9.

Yeoman, C. J., N. Chia, P. Jeraldo, M. Sipos, N. D. Goldenfeld and B. A. White (2012). "The microbiome of the chicken gastrointestinal tract." Animal Health Research Reviews **13**(1): 89 - 99.

Yuvaraj, S., S. Lahham, R. K. R. Marreddy, G. Dijkstra, W. A. M. Wolken, J. S. Lolkema, W. Helfrich, F.-E. Johansen, M. P. Peppelenbosch and N. A. Bos (2008). "Human scFv SIgA expressed on *Lactococcus lactis* as a vector for the treatment of mucosal disease." Molecular Nutrition & Food Research **52**: 913 - 920.

Zhao, K., M. Liu and R. R. Burgess (2007). "Adaptation in bacterial flagellar and motility systems: from regulon members to 'foraging'-like behavior in *E. coli*." Nucleic Acids Research **35**(13): 4441 - 4452.

Zhao, Y., L. Jiang, T. Liu, M. Wang, W. Cao, Y. Bao and J. Qin (2015). "Construction and immunogenicity of the recombinant *Lactobacillus acidophilus* pMG36e-E0-LA-5 of bovine viral diarrhea virus." Journal of Virological Methods **225**: 70 - 75.

Zhou, Y., Y. Liang, K. H. Lynch, J. J. Dennis and D. S. Wishart (2011). "PHAST: A Fast Phage Search Tool." Nucleic Acids Research **39**: W347 - W352.