Proteomic and Phosphoproteomic Studies of Monocyte Recruitment and Macrophage Polarisation

Cheng Huang

B.Sc., M.Med.

A thesis submitted for the degree of *Doctor of Philosophy* at
Monash University in *2017*
Department of Biochemistry and Molecular Biology

Abstract

Inflammation is the body's response to injury or infection. A hallmark of inflammation is the accumulation of monocytes, which subsequently differentiate into macrophages and remove pathogens and necrotic tissue by phagocytosis and proteolytic degradation. Macrophages may be polarised towards pro-inflammatory (M1) or tissue reparative (M2) phenotypes. The balance between these phenotypes can have a substantial influence on the outcome of inflammatory diseases such as atherosclerosis. Chapter 1 summarises current knowledge on the involvement of polarised macrophages in atherosclerosis and the importance of the chemokine and chemokine receptor families of proteins in promoting monocyte accumulation in inflamed tissues. In addition, the first chapter also reviews mass spectrometry-based proteomic methods, which are used throughout this thesis to study macrophage activation and monocyte recruitment processes.

Improved biomarkers of M1 and M2 macrophages would be beneficial for research, diagnosis and monitoring the effects of trial therapeutics in inflammatory diseases. In Chapter 3, to identify novel biomarkers, we have characterized the global proteomes of THP-1 macrophages polarised to M1 and M2 states, in comparison with unpolarised (M0) macrophages. M1 polarisation resulted in increased expression of numerous pro-inflammatory proteins, including the products of 31 genes under the transcriptional control of interferon regulatory factor 1 (IRF-1). In contrast, M2 polarisation identified proteins regulated by components of the transcription factor AP-1. Among the most highly upregulated proteins under M1 conditions were the three interferon-induced proteins with tetratricopeptide repeats (IFIT1, IFIT2 and IFIT3), which function in antiviral defence. Moreover, IFIT1 was strongly upregulated in M1 polarised human primary macrophages and was also expressed in a subset of macrophages in aortic sinus and brachiocephalic artery sections from atherosclerotic ApoE$^{-/-}$ mice. Based on these results, we propose that IFIT proteins may serve as useful markers of atherosclerosis and potentially other inflammatory diseases.

Monocyte chemoattractant proteins (MCP-1, MCP-2 and MCP-3) and chemokine receptor 2 (CCR2) are the major chemokines and receptor involved in recruitment of monocytes and macrophages. MCP-1 and CCR2 are considered to be potential therapeutic targets in a wide range of inflammatory diseases, especially in

atherosclerosis. However, a number of inhibitors targeting them have failed in clinical trials. Thus, there is motivation to identify alternative targets, including downstream signalling proteins that mediate the responses of cells to CCR2 activation. CCR2 is known to signal via G protein and β-arrestin-mediated pathways. However, the downstream signalling pathways have not been thoroughly explored.

In Chapter 4, we describe a large-scale proteomic study using data-independent acquisition (DIA) mass spectrometry to quantify the differences in protein and phosphopeptide levels between untreated and MCP-1- or MCP-3 -treated cell lines. To the best of our knowledge, this is the first time that DIA has been applied to globally quantify the phosphoproteome between multiple conditions. We successfully confirmed the involvement of known canonical pathways (MAPK, JAK/STAT and Akt/mTOR) as downstream effectors of MCP/CCR2 signalling and we also identified, mapped and manually curated additional signalling networks that have not been associated with the MCP/CCR2 signalling cascade. These include Rho guanine nucleotide exchange factors (ARHGEFs), nuclear pore complex (NPC) proteins and many actin cytoskeletal proteins.

We subsequently confirmed and extended the results of this study by analysing the dynamical changes underlying the identified phosphorylation events in Chapter 5. We were able to confirm well-established phosphorylation dynamics of the canonical networks and provided detailed and novel information on the phosphorylation kinetics of several proteins involved in the ARHGEFs, NPC and actin cytoskeleton networks.

Chapter 6 discusses a variety of novel hypotheses arising from the phosphoproteomic studies and proposes approaches to investigate these hypotheses. Together, the findings in this thesis will contribute to our understanding of monocyte recruitment, particularly the mechanisms of the MCP/CCR2 signalling network, and will lay a solid foundation for future work to more fully characterise this network and guide effective atherosclerosis therapies.

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: …Cheng Huang…

Date: ……04/12/2017…

Publications during Enrolment

Huang C., Lewis C., Borg N.A., Canals M., Diep H., Drummond G.R., Goode R.J., Schittenhelm R.B., Vinh A., Zhu M., Kemp-Harper B., Kleifeld O., Stone M.J., Proteomic identification of interferon-Induced proteins with tetratricopeptide repeats as markers of M1 macrophage polarization. J Proteome Res. 2018, in press.

Gulati T., **Huang C.,** Caramia F., Raghu D., Paul P.J., Goode R.J.A, Keam S.P., Williams S.G., Haupt S., Kleifeld O., Schittenhelm R.B., Gamell C., Haupt Y. Proteotranscriptomic measurements of E6-Associated Protein (E6AP) targets in DU145 prostate cancer cells. Mol Cell Proteomics. 2018, in press.

Keam S.P., Gulati T., Gamell C., Caramia F., **Huang C.,** Schittenhelm R.B., Kleifeld O., Neeson P.J., Haupt Y., Williams S.G., Exploring the oncoproteomic response of human prostate cancer to therapeutic radiation using data-independent acquisition (DIA) mass spectrometry. Prostate. 2018, in press.

Reichenbach A., Stark R., Mequinion M., Denis R.R.G., Goularte J.F., Clarke R.E., Lockie S.H., Lemus M.B., Kowalski G.M., Bruce C.R., **Huang C.,** Schittenhelm R.B., Mynatt R.L., Oldfield B.J., Watt M.J., Luquet S., Andrews Z.B. AgRP Neurons Require Carnitine Acetyltransferase to Regulate Metabolic Flexibility and Peripheral Nutrient Partitioning. Cell Rep. 2018, 22(7).

Chan F.L., Vinod B., Novy K., Schittenhelm R.B., **Huang C.**, Udugama M., Nunez-Iglesias J., Lin J.I., Hii L., Chan J., Pickett H.A., Daly R.J., Wong L.H., Aurora kinase B, a novel regulator of TERF1 binding and telomeric integrity. Nucleic Acids Res. 2017, 10.

Liu X., Nefzger C.M., Rossello F.J., Chen J., Knaupp A.S., Firas J., Ford E., Pflueger J., Paynter J.M., Chy H.S., O'Brien C.M., **Huang C.**, Mishra K., Hodgson-Garms M., Jansz N., Williams S.M., Blewitt M.E., Nilsson S.K., Schittenhelm R.B., Laslett A.L., Lister R., Polo J.M., Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. Nat. Methods 2017, 14(11).

Hayward J., Sanchez J., Perry A., **Huang C.**, Rodriguez Valle M., Canals M., Payne R.J., Stone M.J., Ticks from diverse genera encode chemokine-inhibitory evasin proteins. J. Biol. Chem. 2017, 292(38).

Huma Z.E., Sanchez J., Lim H.D., Bridgford J.L., **Huang C.**, Parker B.J., Pazhamalil J.G., Porebski B.T., Pfleger K.D.J., Lane J.R., Canals M., Stone M.J., Key determinants of selective binding and activation by the monocyte chemoattractant proteins at the chemokine receptor CCR2. Sci. Signal. 2017, 10 (480).

Stone M.J., Hayward J.A., **Huang C.**, Huma Z.E., Sanchez J., Mechanisms of regulation of the chemokine-receptor network. Int. J. Mol. Sci. 2017, 18.

Ludeman J.P., Nazari-Robati M., Wilkinson B.L., **Huang C.**, Payne R.J., Stone M.J., Phosphate modulates receptor sulfotyrosine recognition by the chemokine monocyte chemoattractant protein-1 (MCP-1/CCL2). Org. Biomol. Chem. 2015,13(7).

Thesis Including Published Works Declaration

I hereby declare that 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.

This thesis includes 1 original paper submitted in peer reviewed journals. The core theme of the thesis is chemokine receptor signal transduction investigation. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the student, working within the Department of Biochemistry and Molecular Biology under the supervision of Associate Professor Martin Stone.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

In the case of Chapter 3 my contribution to the work involved the following:

- Collection of mass spectrometric and Western blotting data
- Analysis of mass spectrometric and Western blotting data
- Involved in manuscript preparation, figure generation and writing of the first draft

Thesis Chapter	**Publication Title**	**Status**	**Extent of Contribution (%)**	**Co-author name, Nature, % of Co-author's contribution**	**Co-author(s), Monash student?**
3	Proteomic Identification of Interferon-induced Proteins with Tetratricopeptide Repeats (IFITs) as Markers of M1 Macrophage Polarisation	Submitted	50%	1. Caitlin Lewis, collected and analysed primary cell and tissue samples, performed some RT-PCR experiments, 30% 2. Natalie A. Borg, helped with discussion, 1% 3. Meritxell Canals, helped with discussion, 1% 4. Henry Diep, prepared arotic sections, 1%	Yes No No Yes

				5. Grant Drummond, help with discussion, 1%	No
				6. Robert J. Goode, help with PRM re-scheduling, 1%	No
				7. Antony Vinh, help with discussion, 1%	No
				8. Ralf B. Schittenhelm, supervise data collection and analysing, 1%	No
				9. Mingyu Zhu, prepared human primary cells, 1%	Yes
				10. Barbara Kemp-Harper, overall project direction, supervision of the data analysis, finalization of the manuscript, 5%	No
				11. Oded Kleifeld, overall project direction, supervision of the data analysis, finalization of the manuscript, 2%	No
				12. Martin J. Stone, overall project direction, supervision of the data analysis, finalization of the manuscript, 5%	No

I have not renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Student signature: [redacted] **Date: 04/12/2017**

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor signature: [redacted] **Date: 04/12/2017**

Acknowledgements

At the last stage of my PhD studies, when I was writing up this thesis, our little girl Bella was born - not the best timing and tiring, but she makes our life lovely. This actually made me realise that the long way of getting a PhD degree is like raising a child from being a baby to an adult. I would therefore like to divide it my professional development into three stages: newborn, youngster and adult.

I consider the first "newborn" stage – or more accurate the "gestation" stage – when my main supervisor, Associate Professor Martin Stone, recruited me to his lab. I would like to thank Martin for giving me this opportunity to obtain a PhD degree in his lab. But surely more than this, Martin was contributing to all the three stages of getting my PhD. Thank you for all your support, encouragement and questioning during this long journey. Without your fantastic contributions, this could never have happened.

During the second stage, from a newborn to a youngster, which I call the "learning" stage, I would like to thank my first secondary supervisor, Dr. Oded Kleifeld, who taught me most of my knowledge about mass spectrometry. The massive amount of information that I received from him was very useful for finishing up this thesis.

During the last "maturing" stage, from a youngster to an adult, I would like to express my gratitude to my second secondary supervisor, Dr. Ralf Schittenhelm, who brought me into a new "career". I started to work with people and to be responsible for other people's projects. All the skills that I have learnt and am continuing to learn will be helpful for my future. Furthermore, I would like to thank him for critical reading this thesis and for your generous support when my scholarship was finished.

Other than my supervisors, there are so many other people during this long journey who were so helpful in guiding and assisting me with both my work and my life.

I would like to thank Dr. Meritxell Canals in the Drug Discovery Biology, Monash Institute of Pharmaceutical Science, for her expertise and supervision for all cell-based work.

I would like to express my appreciation for the Monash Biomedical Proteomics Facility and the Monash Micro Imaging Platform. Without these facilities this thesis could have never been accomplished. In addition, I would like to thank Dr. Alex Flucher

for helping me out with cell imaging at the last stage of this study and I would like to thank the proteomics facility fellows Dr. David Steer, Robert Goode and Shane Reeve for their help and support throughout this thesis. In particular, I am very grateful to Iresha Hanchapola for assisting me with sample preparation when I was writing up this thesis and for giving me many essential "parenting" tips for Bella.

I would also like to acknowledge my PhD committee, Professor Roger Daly, Professor Tony Purcell, and Professor Mibel Aguilar for their guidance and support at all the milestones and their encouragement and motivation throughout my candidature, especially Roger, who supported me with a departmental scholarship to finish up my studies.

Of course, I would like to thank our lab fellows (Justin Ludeman, Zil E Huma, Julie Sanchez, Jenni Hayward and Sayeeda Chowdhury) and our little office 109 family (Nathan Habila, Melissa Honeydew, Hamed Karimian and Baydaa Hirmiz) for sharing the difficulties of PhD life. Life couldn't be better without sharing "stories" with you guys.

A huge thank also goes out to my parents and my brother for supporting my almost 22 years "student life". I can't image what would have happened without your endless support. Thank you for always being there and for supporting me, no matter what decisions I have made so far.

And lastly, I would like to dedicate this thesis to my wife Qin Liu and my daughter Jiayue (Bella) Huang. Thank you Qin for always backing me up to finish this study and my little cutie for making our life both "miserable" and "fun".

Big thanks also to all the other people I have forgotten to mention here. Without the support and help of so many people, I could never have accomplished this task.

Table of Contents

List of General Abbreviations

ANOVA	analysis of variance
BCA	bicinchoninic acid assay
C18	reverse phase C18
C8	reverse phase C8
CCR1	CC chemokine receptor 1
CCR2	CC chemokine receptor 2
CCR5	CC chemokine receptor 5
CCR7	CC chemokine receptor 7
CD14	monocyte differentiation antigen CD14
CD16	low affinity IgG receptor III
CD209	CD209 antigen
CXCL11	CXC chemokine 11
DDA	data dependent acquisition
DIA	data independent acquisition
FASP	filter-aided sample preparation
FDR	false discovery rate
GADPH	glyceraldehyde-3-Phosphate Dehydrogenase
GO-BP	gene ontology biological process
GO-MF	gene ontology molecular function
HEK293	human embryonic kidney 293 cell line
HPLC	high performance liquid chromatography
IFN-γ	interferon gamma
IL-1β	interleukin-1 beta
IL-4	interleukin-4
IL-6	interleukin-6
IRF	interferon regulator factor
LC	liquid chromatography
LC-MS/MS	liquid chromatography-tandem mass spectrometry
LFQ	label free quantification
LPS	lipopolysaccharide
M0	non-polarised macrophage
M1	polarised macrophage M1 phenotype
M2	polarised macrophage M2 phenotype
MCP-1 / CCL2	monocyte chemoattractant protein -1 / CC chemokine 2
MCP-2 / CCL8	monocyte chemoattractant protein -2 / CC chemokine 8
MCP-3 / CCL7	monocyte chemoattractant protein -3 / CC chemokine 7
MCP-4 / CCL13	monocyte chemoattractant protein -4 / CC chemokine 13
MRC1	mannose receptor, C type 1
MS/MS or MS2	tandem mass spectrometry
NCF-1	neutrophil cytosolic factor 1
NFκB2	nuclear factor-kappaB 2
PARP9	poly [ADP-ribose] polymerase 9

PBMC	peripheral blood mononuclear cell
PDBu	phorbol 12,13-dibutyrate
ERK/ MAPK	mitogen-activated protein kinase
pERK	phosphorylated ERK
ppm	parts per million
PRM	parallel reaction monitoring
RSLC nano HPLC	rapid separation LC nano high performance liquid chromatography
SDC	sodium deoxycholate
Z score	standard score
UHPLC	ultra high performance liquid chromatography
SOD2	superoxide dismutase 2
SORT1	sortilin
STAT1	signal transducer and activator of transcription 1
TGFB1	Transforming growth factor beta-1
TGM2	transglutaminase 2
THP-1	human monocytic cell line
TiO_2	titanium dioxide
TNF-α	tumor necrosis factor alpha
CK2	casein kinase 2
mTOR	serine/threonine-protein kinase mTOR
ErbB	epidermal growth factor receptor
KEGG	Kyoto Encyclopedia of Genes and Genomes
GEF	guanine nucleotide exchange factor
ARHGEF	Rho guanine nucleotide exchange factor
NPC	nuclear pore complex
v/v	volume to volume
w/v	weight to volume
RT	room temperature

Chapter 1. Introduction

1.1 Atherosclerosis

1.1.1 Introduction

According to the World Health Organization, cardiovascular disease is the number one cause of death globally. In 2008, an estimated 17.3 million people died from this disease, representing 30% of all global deaths and the number increase to 23.3 million by 2030 [1]. Atherosclerosis is the underlying cause of most cardiovascular diseases, including coronary heart disease, carotid artery disease, peripheral arterial disease and many cases of heart attack and stroke [2-4].

Atherosclerosis is a slow and progressive disease, which may start from childhood and is mainly localised to the intima of many medium-sized and large arteries, especially where the vessels divide [5]. It results from the accumulation in the arteries of plaque, which is made up of fat, cholesterol, calcium and other substances found in the blood. Atherosclerosis can be asymptomatic for decades, because arteries can enlarge at the positions of plaque, keeping the blood flow normal until enough plaque narrows or blocks the artery to induce symptoms [6]. Atherosclerosis has sometimes been considered to be a lipid storage disease associated with advanced modern lifestyles [2,7]. However, vascular calcification has recently been identified from computed tomography (CT) scans of a 5,300 year-old mummy [8] suggesting that this type of disease could date back to around 30th century BC. This finding has also been confirmed by additional CT evidence from other mummies [9-11]. Thus, the advance of science has proven that atherosclerosis is not only related to the modern dietary habits, but may be a more general maladaptive inflammatory response to infection [12-14].

1.1.2 Pathogenesis of Atherosclerosis

The vascular endothelium is a thin layer of cells in the artery, which provides a smooth and protective surface for the blood vessels, enabling the blood to flow fluently. Under normal circumstances, the leukocytes in the flowing blood do not adhere to the vascular endothelium and lipids such as cholesterol cannot penetrate this natural barrier [15]. However, when endothelial cells are damaged or become dysfunctional, lipids can penetrate the endothelium and accumulate in the smooth muscle cells. As a consequence, monocytes, which subsequently differentiate into macrophages, are recruited from the blood to clean up the accumulated lipids, but may become trapped in the damaged

endothelium area and provoke an inflammatory response, which can progress to atherosclerosis.

In detail, the pathogenesis of atherosclerosis can be separated into four steps [16]. First, monocytes are recruited to the site of damage in the endothelial cells, where they differentiate into macrophages [5]. Second, monocytes/macrophages engulf highly oxidized low density lipoproteins and become foam cells, which form the centre of atherosclerotic plaques [17]. Third, the trapped macrophages, foam cells and accumulated lipids form the necrotic core of plaques in the arteries [18]. Finally, the necrotic core or plaque can continue to grow and can form a thin fibrous cap. If the plaque is ruptured, it can initiate a blood clot, which can lodge in the heart or brain, causing a myocardial infarction (heart attack) or a thrombotic stroke [19].

1.1.3 Monocyte Recruitment and Macrophage Polarisation

Monocyte recruitment (Figure 1.1) is the first step in the development of atherosclerosis. Monocytes are a subtype of leukocytes (or white blood cells) that are produced in the bone marrow and circulate in the blood. They can travel across the endothelium to the inflamed or damaged tissue and then differentiate into macrophages, whose functions include to remove dying or dead cells or cellular debris, to phagocytize bacteria and damaged tissue, to regulate immune responses, to promote the development of inflammation and to promote tissue repair [20].

Macrophages can display diverse functional phenotypes according to the stimulus they encounter. This phenomenon is defined as macrophage polarisation. Polarised macrophages are broadly divided into two groups: the classically activated macrophage (M1) phenotype and the alternatively activated macrophage (M2) phenotype, mirroring the polarisation of helper T cells to Th1 and Th2 phenotypes, respectively. M1 macrophages, polarised by pro-inflammatory stimuli such as IFN-γ and TNF-α generate pro-inflammatory cytokines (e.g. IL-1β, IL-6, IL-12, IL-18), which are pro-oxidative, promote killing of microbial pathogens and intracellular parasites and cause tissue damage [21]. M2 macrophages, polarised by anti-inflammatory cytokines such as IL-4 and IL-13, are generally considered to have anti-inflammatory, anti-oxidative and tissue reparative properties [22]. The balance between M1 and M2 polarised macrophages is believed to be involved in various functions and diseases [23]. Monocyte recruitment and polarisation are illustrated in Figure 1.1.

Considering the importance of regulating the balance between M1 and M2 macrophages, there is much interest in identifying markers of these phenotypes and potential drug targets specific to these types of cells. To this end, Chapter 3 of this thesis describes the application of mass spectrometry-based proteomics to identify differences between the proteomes of non-polarised and M1 or M2 polarised macrophages.

Our understanding of the mechanism underlying monocyte recruitment has been advanced by the discovery and functional characterisation of chemokines and chemokine receptors [24,25]. It is now well recognised that monocyte rolling, firm adhesion and transmigration is controlled to a large extent by chemokines activating chemokine receptors [26]. Chemokines can present basal expression under normal condition and are immobilized on the endothelial or tissue cell surface through binding to glycosaminoglycans (GAGs). At the site of injury or infection, chemokines can cross the endothelium and activate their cognate chemokine receptors, which are expressed on the monocytes, leading to the trafficking of monocytes and production of inflammatory cytokines.

1.1.4 Determination of Polarised Macrophage Phenotypes

There are different ways to determine the phenotypes of polarised macrophages. Traditionally, M1 macrophages can be determined via detecting the increased expression of pro-inflammatory cytokines (e.g. IL-1β, TNF-α and IL-6), chemokines (e.g. CXCL9, CXCL10 and CXCL11) or their receptors (e.g. CCR7) with enzyme-linked immunosorbent assay (ELISA) or real-time polymerase chain reaction (RT-PCR) [27]. Similarly, M2 macrophages can be identified via the expression of pro-fibrotic (e.g. transforming growth factor-β (TGF-β) and insulin-like growth factor-1 (IGF-1) [28] and anti-inflammatory (e.g. IL-10) cytokines and scavenger receptors (e.g. mannose receptor C-type 1 (MRC-1) [29]. The macrophages can also be sorted by flow cytometry based on cell surface markers such as CD 80 and CD86 for M1 macrophages and CD36 and CD163 [27] for M2 macrophages.

Whilst such biomarkers are useful, current limitations include differences between the expression profiles of murine and human M1 and M2 macrophages [30,31] and differences between *in vitro* and *in vivo* macrophages [32,33]. As such, investigations would greatly benefit from an improved set of biomarkers for M1 and M2 macrophages.

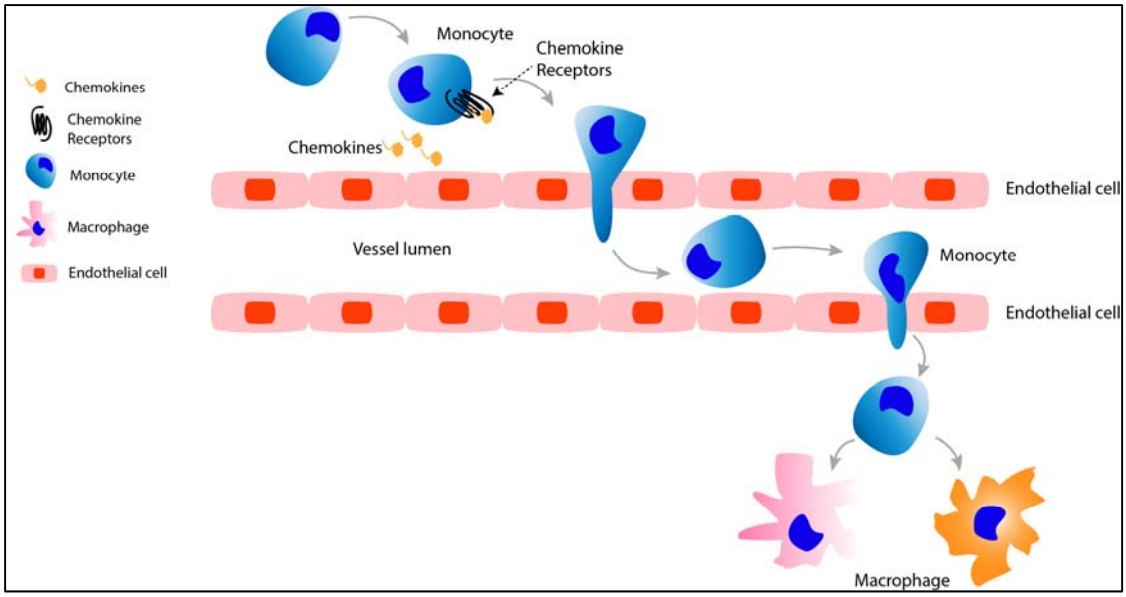


Figure 1.1. Monocyte recruitment and polarisation. The interaction between chemokines and chemokine receptors on the surface of the monocytes promotes transmigration of monocytes across endothelial cells from blood to the tissue, where they differentiate into macrophages and can be further polarised to M1 or M2 phenotypes.

Transcriptomics, a technique to measure the expression levels of all mRNA transcripts in a sample, was applied by Martinez et al. [31] in 2006 on human primary monocytes to profile mRNA expression of M1 and M2 macrophages. From this study [31], high levels of CCL8, CCL15, CCL19, CCL20 and CXCL13 transcripts were found in M1 cells and high levels of CCL13, CCL14, CCL17, CCL23, and CCL26 transcripts were found in M2 cells. Other than the expression of these classical chemokines, the mRNA expression of some G protein-coupled receptors (GPCRs) were also found to be up-regulated by M1 or M2 treatment. For example, M1 treatment can increase the expression of CCR7; whereas, GPR86, GPR105, P2Y8, P2Y11, and P2Y12 were increased in M2 macrophages. Interestingly, some solute carrier proteins also showed differential expression levels in M1 (e.g. SLC21A15 and SLC31A2) and M2 (SLC4A7 and SLC38A6) macrophages. All of these proteins could potentially serve as biomarkers. Indeed, several studies [34,35] have been published on transcriptional profiling of polarised macrophages by different treatments or cell types. However, it is important to note that the levels of mRNA do not always reflect the expression levels of the respective proteins. Therefore, it would be beneficial to identify protein markers that can be easily used in Western blot analyses to determine the phenotypes of polarised macrophages.

A few studies have explored the proteomes of polarized macrophages. An early study of primary human macrophages [36] identified a number of proteins that are specifically up- or down-regulated in macrophages treated with lipopolysaccharide (LPS) or IFN-γ compared to untreated controls, but did not investigate the more typical dual LPS/IFN-γ treatment or M2 polarized cells. Another study [37] focused on membrane associated proteins, identifying several that distinguish M1 or M2 polarized from non-polarized primary mouse macrophages. More recently, two studies [38,39] have used comparative proteomic methods (2D electrophoresis or SILAC) and the well-established THP-1 cell line to identify proteins expressed differentially between M1 and M2 macrophages, although neither of these studies included comparisons to non-polarized cells. In conclusion, all of these studies either were not able to identify a deep global proteome of the macrophages or neglected the control condition of non-polarised macrophages (M0). In Chapter 3 of this thesis, in particular in light of recent advances in label-free proteomics methodologies, we have performed a systematic global proteomics comparison of M1 polarized, M2 polarized and non-polarized THP-1 macrophages. The

study provides an expanded set of macrophage polarization markers that will serve as a valuable benchmark for future studies of inflammatory diseases and treatments.

1.2 Chemokines and Chemokine Receptors

1.2.1 Chemokines

Chemokines are a family of small cytokines, secreted in response to a variety of inflammatory stimuli [40,41]. Most chemokines consist of 70-80 amino acids, yielding molecular weights ranging from 8-10 kDa, and they have four highly conserved cysteine residues, which are important for structure and chemokine-mediated signalling [42-44]. Members of the chemokine family are divided into four groups (CC, XC, CXC and CX_3C, where X represents any amino acid), depending on the spacing of their first two cysteine residues. For example, in the two main subfamilies, CC and CXC chemokines, the first two cysteines are adjacent or separated by one amino acid, respectively. Table 1.1 lists known chemokines and their receptors grouped by chemokine families. Both systematic and common names are shown for chemokines in this table, but please note that the common names are used throughout the result chapters (Chapter3, 4 and 5).

The major role of chemokines is to act as chemoattractants to guide the migration of cells. Based on their functions, chemokines can be homeostatic, inflammatory or both. Homeostatic chemokines are constitutively expressed and important for many physiological processes, while the expression of inflammatory chemokines is induced or strongly upregulated by inflammatory stimuli [45].

Most of the chemokines are inflammatory chemokines that can be expressed by any type of cells upon stimulation. For example, by using Northern blot analysis, Ylä-Herttuala et al. [46] found that in normal human or rabbit arteries, mRNA expression of MCP-1 was not detectable; whereas, in atherosclerotic lesions, the expression of MCP-1 was remarkably increased. This was further confirmed by immunocytochemistry [46]. Similar observations can also be found in other diseases, such as CCL5/RANTES in breast cancer [47-49].

On the other hand, the homeostatic or constitutively expressed chemokines are restricted to specific cell types or tissues and their functions include T cell development, stem cell migration and lymphoid organogenesis [50]. For example, CCL25/TECK was found to be strongly expressed in the thymus, relatively weakly expressed in the small

intestine and liver, but absent in lung or colon, in line with its role in regulating T cell development [51]. Similarly, the homing chemokine CCL19 is selectively highly expressed by stromal cells and dendritic cells in the T cell zone of the lymph nodes and is involved in T cell homeostasis [52,53]. Interestingly, some chemokines, such as XCL1, XCL2, CCL17, CCL20, CCL21 and CCL22, can display either inflammatory or homeostatic function, depending on their cognate receptors, the tissue where they are expressed and stimulus they encounter [54-56].

Apart from regulating leukocyte trafficking, additional biological functions have been observed for some chemokines. These include direct antimicrobial activity, HIV inhibitory activity, angiogenic or angiostatic activity, tumour-promoting or tumour-inhibiting activity, apoptosis or mitogenic activity, modulating gene expression, T cell differentiation and phagocyte activation. However, a unifying aspect of most of these biological activities is that chemokines must bind to and activate their cognate receptors (chemokine receptors).

Table 1.1. Chemokine and chemokine receptor families[1]

Systematic name	Common name	Chemokine receptors
CC chemokine family		
CCL1	I309	CCR8
CCL2	MCP-1	CCR2, CCR5, ACKR1, ACKR2
CCL3	MIP-1α	CCR1, CCR5, ACKR2
CCL3L1	LD78β	CCR5, ACKR2
CCL3L3	MGC12815	CCR5, ACKR2
CCL4	MIP-1β	CCR1, CCR5, ACKR2
CCL4L1	LAG-1	CCR5
CCL4L2		CCR5
CCL5	RANTES	CCR1, CCR3, CCR5, ACKR1, ACKR2
CCL7	MCP-3	CCR1, CCR2, CCR3, CCR5, CXCR3, ACKR1, ACKR2
CCL8	MCP-2	CCR1, CCR3, CCR5, CCR8, ACKR2
CCL11	Eotaxin	CCR2, CCR3, CCR5, CXCR3, ACKR1, ACKR2
CCL13	MCP-4	CCR1, CCR2, CCR3, ACKR2,
CCL14	HCC-1	CCR1, CCR5, ACKR1, ACKR2
CCL15	HCC-2	CCR1, CCR3
CCL16	HCC-4	CCR2, CCR5
CCL17	TARC	CCR4, ACKR1, ACKR2
CCL18	MIP-4	Unknown
CCL19	MIP-3β	CCR7, ACKR4, CCRL2
CCL20	MIP-3α	CCR6
CCL21	SCYA21	CCR7, ACKR4
CCL22	MDC	CCR4, ACKR2
CCL23	MPIF-1	CCR1
CCL24	Eotaxin-2	CCR3
CCL25	TECK	CCR9, ACKR4
CCL26	Eotaxin-3	CCR2, CCR3
CCL27	ILC	CCR4, CCR10
CCL28	SCYA28	CCR3, CCR10
CXC chemokine family		
CXCL1	GROα	CXCR1, CXCR2
CXCL2	GROβ	CXCR2
CXCL3	GROγ	CXCR2
CXCL5	ENA-78	CXCR2, ACKR1
CXCL6	GCP-2	CXCR1, CXCR2, ACKR1
CXCL7	NAP-2	CXCR2
CXCL8	IL8	CXCR1, CXCR2, ACKR1
CXCL9	MIG	CCR3, CXCR3
CXCL10	IFI10	CCR3, CXCR3

CXCL11	I-TAC	CCR3, CXCR3, ACKR1, ACKR3
CXCL12	SDF-1	CXCR4, ACKR3
CXCL13	BCA-1	CXCR5
CXCL14	BRAK	Unknown
CXCL16	SCYB16	CXCR6
CXCL17	Dcip1	Unknown
CX3C chemokine family		
CX3CL1	ABCD-3	CX3CR1
C or XC chemokine family		
XCL1	LTN	XCR1
XCL2	SCM-1β	CXR1

[1] Data are taken from the International Union of Pharmacology (IUPHAR) database (http://www.iuphar-db.org)

1.2.2 Chemokine Receptors

Chemokine receptors are found on the surfaces of certain cells, most commonly leukocytes. They belong to the family of G protein-coupled receptors (GPCRs), also known as seven-transmembrane domain receptors, which is one of the largest families of proteins in mammalian genomes [57,58]; more than 1000 such receptors have been identified in the human genome.

Chemokine receptors are divided into four groups (CCR, XCR, CXCR and CX3CR), depending on the chemokine subclass specificity to the receptor [59]. For example, CCR2 ('R' represents receptor) can be activated by several CC chemokines (Table 1.1) [60]. To date, at least 45 human chemokines and 23 human chemokine receptors have been identified and interactions between them are highly promiscuous, often with one receptor being able to bind to multiple chemokines and some chemokines being able to bind to more than one receptor as indicated in Table 1.1.

Chemokine receptors are composed of about 340-370 amino acids that are divided into a short, acidic N-terminal region, seven helical transmembrane (TM) helices with three intracellular and three extracellular hydrophilic loops, and an intracellular C-terminus containing an eighth helix as well as several serine and threonine residues that act as phosphorylation sites during receptor regulation (Figure 1.2A and B). The sequences of chemokine receptors are highly conserved within the TM segments, whereas the amino and carboxyl termini and the third intracellular loop (between TM5 and TM6) show the highest sequence diversity [61]. Other structural features commonly found in chemokine receptors include cysteine residues that form two conserved disulfide bonds between the extracellular domains, a conserved 'DRY' (Figure 1.2 C) amino acid sequence motif, which has been implicated in G protein interaction, in the second intracellular loop, and an unusually short third intracellular loop which is enriched in positively charged residues [62].

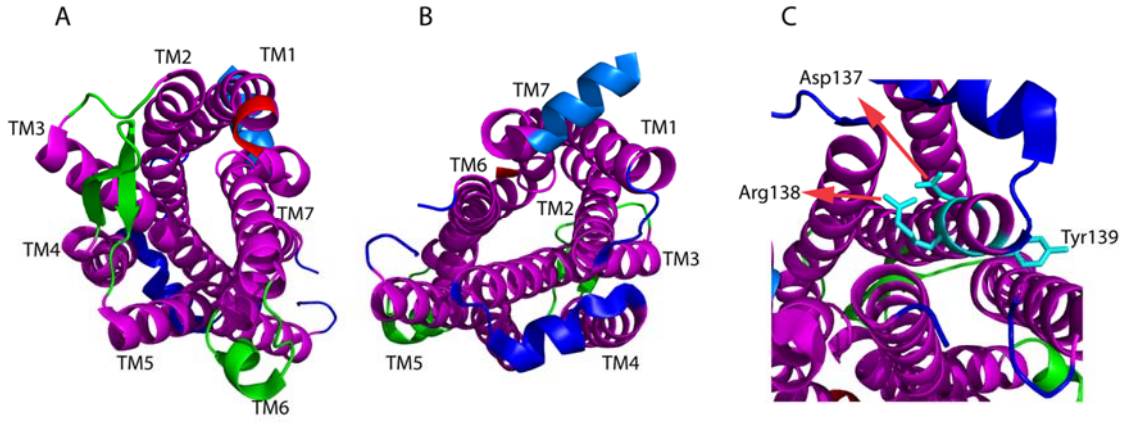


Figure 1.2. [63] Structure of CCR2 (PDB: 5T1A) represents common structure of chemokine receptors. A, Top view; B, Bottom view; C, DRY motif. Magenta, 7 transmembrane helices; red, N-terminus; green, extracellular loops; blue, intracellular loops; light blue, C-terminal helix 8; cyan sticks, Asp-137, Arg-138 and Tyr-139 forming the conserved DRY motif.

1.2.3 Signal Transduction in the Chemokine:Chemokine Receptor System

Chemokine receptors stimulated by chemokines share the common features of GPCR signalling pathways. Upon ligand binding, chemokine receptors associate with G proteins, exchanging GDP for GTP. In the active (GTP-bound) state, Gα proteins dissociate from the tight GβGγ complex and are able to activate signalling molecules that mediate changes in the cytoskeletal apparatus and transcription factors that regulate cell growth. Activated Gβ and Gγ subnits of G proteins trigger a series of reactions, commencing with the hydrolysis of phosphatidylinositol-4,5-bisphosphate (PIP_2), generating 1,4,5-inositol trisphophate (IP_3) and 1,2-diacylglycerol (DAG), an immediate consequence of which is the release of calcium from intracellular stores. This induction of calcium mobilization is one of the activities very often used to determine chemokine activity [64]. There is increasing evidence that chemokine receptors can also activate several different intracellular effectors downstream of $G\alpha_i$ coupling, including the low-molecular-weight proteins Ras and Rho, phospholipase-A2 (PLA2), phosphatidylinositol-3 kinase (PI3K), tyrosine kinases and the mitogen activated protein kinase (MAPK) pathway [65]. After binding to their specific receptors, chemokines may also induce receptor homodimerization and subsequently activate the receptor-associated Janus kinase (JAK) pathway, possibly by transphosphorylation on tyrosine residues. This may create SH2 docking sites, leading to the recruitment of Signal Transducers and Activators of Transcription factors (STAT) [66]. The canonical pathways of chemokine receptor activation are shown in Figure 1.3.

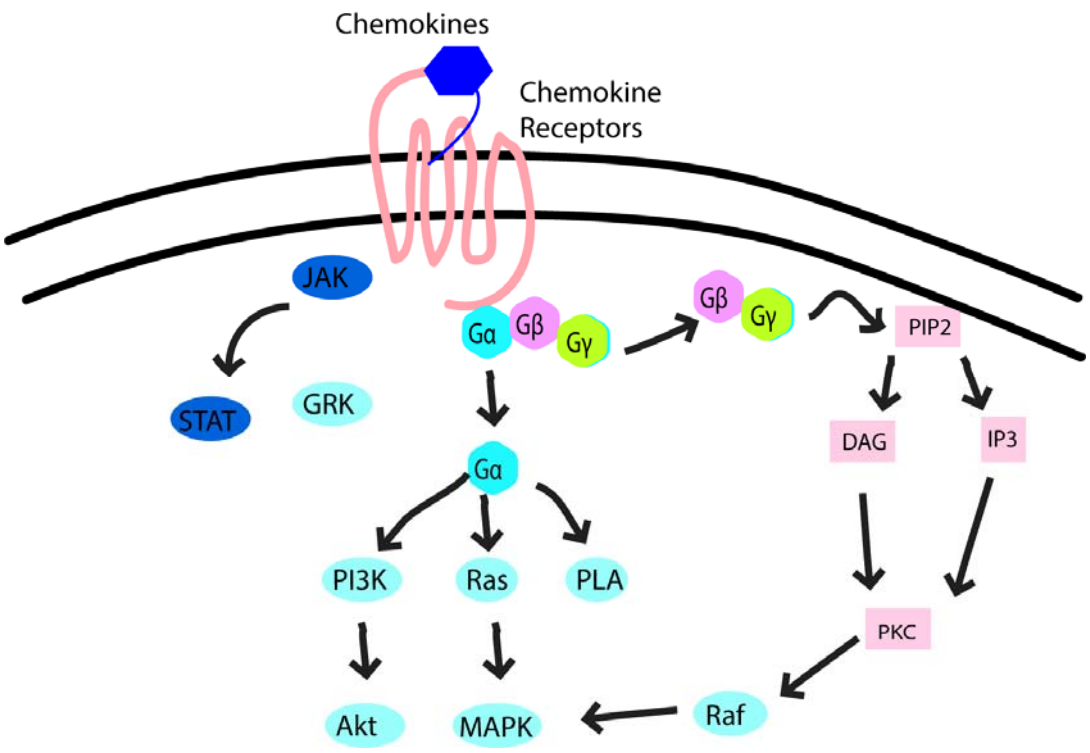


Figure 1.3. Generalised canonical pathways of chemokine receptor activation. Upon chemokine binding to the chemokine receptor, Gα proteins dissociate from the tight GβGγ complex and are able to activate downstream intracellular effectors, such as phosphatidylinositol-3 kinase (PI3K), small GTPase (Ras) and phospholipase-A2 (PLA2), which can further activate protein kinase B (Akt) and the mitogen activated protein kinase (MAPK) pathway. The activated Gβ and Gγ subunits of the G proteins trigger a series of reactions, commencing with the hydrolysis of phosphatidylinositol-4,5-bisphosphate (PIP_2) generating 1,4,5-inositol trisphophate (IP_3) and 1,2-diacylglycerol (DAG), which further activate Raf kinase that is connected to the MAPK pathway.

Although a few chemokine receptors interact specifically with a single ligand, most can be activated by several chemokines (Table 1.1). This phenomenon was previously considered to represent redundancy of the chemokine network [60,67,68]. However, there is mounting evidence that different ligands binding to the same receptor can produce distinct signalling responses [69]. For example, the cellular responses to three CXC chemokines (CXCL9, CXCL10, and CXCL11) were investigated using cells expressing their shared receptor, CXCR3. CXCL11 resulted in a transient increase of Ca^{2+}, while CXCL9 and CXCL10 were found to be ineffective at raising Ca^{2+}. Another signalling cascade, PI3K dependent phosphorylation of protein kinase B at Thr-308 and Ser-473 was stimulated by all these chemokines [70]. Similar observations have been discovered in CCR1, CCR2, CCR5, CCR7 and CXCR7 systems [70-75]. Rajagopal et al. [76] systematically studied ligand-biased signalling in the chemokine and chemokine receptors system. The abilities of different ligands binding to the same receptor to activate distinct signalling pathways was tested in CCR1, CCR5, CCR10, CXCR1, CXCR2 and CXCR3. 'Bias plots' were presented for qualitative analysis of ligand bias in CCR1, CCR10, CXCR2 and CXCR3. At CCR1, the ligands CCL15 and CCL23 demonstrated relatively decreased internalization compared to β-arrestin recruitment, while CCL3 and CCL5 showed the opposite trend. At CCR10, CCL27 was capable of β-arrestin recruitment and receptor internalization, while CCL28 was not as effective as CCL27. CXCL8 at CXCR1 and CXCL11 at CXCR3 displayed high levels of receptor internalization. These findings, combined with previous work, led the authors to suggest that biased agonism is likely to be an evolutionarily conserved biological mechanism for generating qualitatively distinct patterns of signalling via the same receptor in response to different endogenous chemokine ligands [76].

1.2.4 CC Chemokine Receptor 2 and Monocyte Chemoattractant Proteins

In 1994, Charo et al. [77] characterised two receptors, initially thought to be specific for monocyte chemotactic protein 1 (MCP-1/CCL2), and named them CCR2A and CCR2B. These two receptors are splice variants that differ in their carboxyl-terminal tails, which contain serine and threonine residues susceptible to phosphorylation [78]. Most subsequent research on these receptors has focused on the variant CCR2B, which will be referred to throughout this thesis as CCR2.

CCR2, one of the most extensively studied chemokine receptors, presents on the surface of several leukocyte subsets, such as monocytes, basophils, dendritic cells, natural killer cells and activated T lymphocytes. The functions of CCR2 include stimulation of chemotaxis, granule release, respiratory burst and cytokine release [79]. CCR2 shares the common structural features of chemokine receptors and other GPCRs. A high resolution crystal structure of CCR2 bound to two antagonists has recently been solved by Handel's group [63].

CCR2 can bind to a series of high-affinity ligands, including CCL2/MCP-1, CCL7/MCP-3, CCL8/MCP-2, CCL11/eotaxin-1, CCL12/MCP-5 (found in mice but not humans), CCL13/MCP-4 (found in humans but not mice), CCL16/HCC-4 and CCL26/eotaxin-3. The principal cognate ligands of CCR2 are generally considered to be the monocyte chemoattractant proteins (MCPs), which are primarily secreted by monocytes, macrophages and dendritic cells [80]. By activation of CCR2, MCPs can regulate the migration and infiltration of monocytes, memory T lymphocytes, and natural killer (NK) cells.

1.2.5 CCR2 and MCPs in Atherosclerosis

CCR2 and MCPs have been implicated in various diseases, including multiple sclerosis, rheumatoid arthritis, diabetes and especially atherosclerosis [81-83]. In atherosclerosis, monocytes are recruited from the blood into arterial walls where they differentiate into macrophages and subsequently polarise towards the M1 phenotype. M1 macrophages then contribute to plaque development and instability, resulting in major cardiovascular events (heart attack and stroke). CCR2 and MCPs are involved in guiding the transmigration of monocytes. Aiello et al. [84] transplanted bone marrow cells overexpressing MCP-1 into irradiated apoE-knockout mice, which can generate atheromatous lesions by themselves. Compared to the control mice, the MCP-1 transplanted mice presented increased oxidized lipid storage and monocytes/macrophages in lesion areas, indicating MCP-1 can promote atherosclerosis. Consistent with this conclusion, Gosling et al. [85] found that deletion of the MCP-1 gene in mice significantly reduced atherosclerotic lesions. Similarly, overexpression of MCP-3 in transgenic mice can also promote the development of atherosclerosis [86]. Although there is no direct evidence showing the functions of MCP-2, MCP-4 or MCP-5 in atherosclerosis, these MCPs also contribute to recruiting monocytes/macrophages

[87-89]. When investigating their main cognate receptor, CCR2, deletion of the CCR2 gene [90] or inhibition of CCR2 with antagonists [91] in mice both dramatically decreased atherosclerotic lesions and improved plaque stability. All these results indicate the importance of CCR2 and MCPs in the pathogenesis of atherosclerosis.

Despite numerous pharmacological and clinical studies targeting CCR2 or MCPs [91-94], currently there is no medicine on the market for treating atherosclerosis by inhibiting CCR2 or MCPs. A possible reason for this is that some of the same chemokines that activate CCR2 also activate other chemokine receptors. For this reason, rather than targeting individual chemokines or receptors, it may be more effective to target downstream signalling effectors that are activated by MCP/CCR2 signalling pathways as well as other chemokine/receptor combinations. To achieve this goal, it will be important to better understand MCP/CCR2 signal transduction.

MCP/CCR2 signalling pathways share the canonical chemokine and chemokine receptor signalling, as shown in Figure 1.3. In addition, it has also been reported that MCP-1 activation of CCR2 can induce fast signal transduction by increasing phosphorylation of tyrosine kinases, such as Lyn, paxillin and JAK2, leading to the phosphorylation of transcription factors STAT3 and STAT5 [95].

Considering the importance of signalling in chemokine/receptor function and the potential to target signalling effectors therapeutically, it is important to understand the pathways or networks of signal transduction as thoroughly as possible. Previous studies have used focused signalling assays, such as biosensors for second messengers, Western blots directed towards phosphorylated effector proteins, and transcriptional reporters, to identify a number of the effector molecules involved in these signalling pathways. While these methods are powerful for testing specific hypotheses, they are intrinsically biased towards observing pathways previously proposed to be important. It would therefore be valuable to complement such studies with a broader, unbiased approach to characterising the cellular signalling responses to chemokine receptor activation. One such approach is mass spectrometry-based proteomics.

To date, there have been no studies applying unbiased mass spectrometry-based proteomic technology to investigate the signalling transduction activated by MCPs/CCR2 or, indeed, any pro-inflammatory chemokine/receptor interactions. In Chapters 4 and 5 of this thesis, we describe the first such mass spectrometry-based

(phospho)proteomic study and the characterisation of an extensive signalling network resulting from MCP activation of CCR2.

1.3 Mass Spectrometry-based Proteomics

1.3.1 Approaches to Protein Sequence Determination and Identification

The term "proteome" was coined by Marc Wilkins in 1994 during his doctoral studies at Macquarie University, Australia and describes the entire set of proteins expressed by an organism [96,97]. A comprehensive study (often called a "large-scale" study) of a proteome is termed proteomics, which nearly always implies the use of a mass spectrometer and includes the analysis of protein quantities, expression levels, protein-protein interactions and/or post-translational modification [98,99].

Whole genomes could already be sequenced in the 1990s [100,101] and the human genome was fully assembled in 2003 [102]. However, it quickly became apparent that most biological questions cannot be answered solely by obtaining detailed information of the genome. Therefore, proteins started to become the focus of many research projects, as they are responsible for most functions in life such as catalysis, signal transduction and structural integrity. Unfortunately, studying a proteome is much more complicated that studying a genome. The human genome consists of a fixed set of 20,000 protein-coding genes [102], which are identical in each and every cell. In contrast, more than 290,000 non-redundant peptides from approximately 30,000 proteins have been identified in humans by mass spectrometry [103,104]. To complicate the matter, a proteome is quite plastic and differs dramatically between various cell types and tissues of the same organism and also changes under different conditions, stimuli or environmental factors. In addition, the vast majority of proteins are also modified after translation and it is estimated that more than 1,000,000 specific post-translational modifications are present in natural proteins [105]. The following sections will introduce various methods to study proteins including Edman degradation, peptide mass fingerprinting as well as tandem mass spectrometry used for top-down and bottom-up proteomics.

1.3.1.1 Edman Degradation

The first method of determining the sequence of a peptide or a protein was Edman degradation, developed by Pehr Edman in 1949 [106]. The uncharged amino terminus of

a purified protein can react with phenylisothiocyanate (PTC) under basic conditions to form a cyclic phenylthiocarbamyl derivative (PTC-protein). This amino acid can then be cleaved off by trifluoroacetic acid, generating a PTC-amino acid and a protein containing a new amino terminus. The PTC-amino acid can be identified by C18-HPLC with a pre-defined standard, while the protein with the new amino terminus can be used for another cycle of Edman degradation. This method is relatively simple and the automated Edman degradation instrument was developed in 1967 [107]. One-dimensional (1D) or two-dimensional (2D) gel electrophoresis coupled to Edman degradation enabled the identification of several proteins in a mixture [108,109]; however, some drawbacks limit its usage for deep proteome characterisation. First, it is relatively slow and even the fastest instrument takes approx. one hour to determine a single amino acid; second, the protein of interest has to be by far the most abundant species in the sample; third, 2-D gels suffer from lack of reproducibility; last, the amino terminus must not be modified for Edman degradation to work, but most proteins are known to be N-terminally acetylated.

1.3.1.2 Peptide Mass Fingerprinting

In the 1990s, a faster protein sequencing method called peptide mass fingerprinting (PMF) was established, which was enabled through the successful sequencing of genomes and the accompanied availability of protein databases. In brief, an unknown protein (or a simple protein mixture) was cleaved into smaller peptides by a protease and a mass spectrometer was used to measure the mass of these peptides. A protein database was "digested" *in silico* with the same protease and the measured masses of the unknown peptides were compared to the theoretical masses obtained from the protein database.

PMF relied heavily on the development and improvement of mass spectrometers. Fast atom bombardment (FAB) ionization was used until the end of the 1980s [110]; however, due to the lack of sensitivity and ionisation efficiency of large peptides, FAB was not implemented in many laboratories. It was not until 1991 that commercially available instruments coupled with soft ionization techniques, namely electrospray ionization (ESI) [111] and especially matrix assisted laser ionization (MALDI) [112], made PMF a routinely used method for protein identification. Although MALDI-TOF can provide mass accuracy of less than 50 ppm and requires only small amount of

proteins, a SDS-PAGE or 2D gel was still necessary to reduce the complexity of the sample [113]. Nevertheless, considering the low reproducibility of 2D gels mentioned above and also the expanding sizes of the protein databases, PMF was still not powerful or fast enough for global proteome identification.

1.3.1.3 Top-down Proteomics

Mass spectrometry-based proteomics is so far the most comprehensive, robust, and high-throughput approach for global proteome identification. It can be broadly divided into top-down proteomics and bottom-up proteomics.

Top-down proteomics is a technique to characterise - or even quantitate - intact proteins. In top-down proteomics, the intact protein mix is directly injected in and analysed by the mass spectrometer. The addition of a protease is avoided and the intact protein is directly fragmented by the mass spectrometer. Therefore, top-down proteomics can be used to detect protein proteoforms, including degradation products, mutations or post-translational modifications [114,115]. Due to significant advances in mass spectrometric instrumentation and protein separation procedures, such as the four-dimensional separation technique (gel-eluted liquid fraction entrapment electrophoresis), top-down proteomics is capable of identifying more than 2,000 proteoforms [116-119]. However, it is still impossible to achieve a deep proteome coverage using top-down proteomics due to the difficulties in fractionating, ionising and fragmenting intact proteins [3,120].

1.3.1.4 Bottom-up Proteomics

Bottom-up or shot-gun proteomics is based on enzymatic digestion of proteins. Proteins are cleaved into small peptides by a specific protease, usually trypsin, which can cleave the peptide backbone on the carboxyl side of lysine and arginine residues, thereby typically generating peptides between 500 and 3,000 Da. This mass range is ideal for mass spectrometric analysis and the presence of an arginine or lysine at the C-terminus helps peptide ionisation and fragmentation under acidic conditions [121,122]. The peptides are usually separated by reversed-phase chromatography according to their hydrophobicity and analysed on a tandem mass spectrometer, yielding the molecular (precursor) mass and masses of several fragments for each peptide. Peptide identification is achieved by comparing the observed mass-to-charge (m/z) ratios of the precursor and

fragment ions with theoretical m/z ratios obtained from *in-silico* digested protein databases [120]. Bottom-up proteomics is a very robust and advanced technique and it is currently the first choice for whole proteome characterisations.

1.3.2 Phosphoprotein Identification Methodologies

In both prokaryotic and eukaryotic organisms, phosphorylation is a key reversible post translational modification and is critical for the rapid transduction of messages from extracellular stimuli to elicit a cellular response. Thus, it is not surprising that an estimated 2 to 3% of the human genome is directly involved in phosphorylation [123,124] and an estimated 30 to 50% of proteins are thought to be phosphorylated. Protein kinases catalyse the transfer of a phosphate group from adenosine triphosphate (ATP) or GTP to the side chain of serine, threonine or tyrosine and sometimes, even to histidine or aspartic acid [125,126], while dephosphorylation is mainly regulated by phosphatases. Phosphorylation and dephosphorylation of proteins are essential for regulation of cellular signalling because they can lead to a series of changes in downstream signalling networks. Elucidating phosphorylation related signalling pathways and understanding how they are regulated is challenging but often rewarding [127].

1.3.2.1 Radiolabelling

Traditionally, protein phosphorylation was studied in one or only a few isolated events due to the lack of efficient isolation methods for analysing global phosphorylation events. The classical method of studying protein phosphorylation was through incorporation of radioactive ^{32}P-phosphate into proteins. When labelled ^{32}P-phosphate ATP is added exogenously into growing cells, protein kinases transfer the labelled orthophosphate to specific amino acids [128,129]. Proteins were extracted from cells, separated by SDS-PAGE and exposed to the film. This method is very sensitive and selective; however, some drawbacks of this method include the number of phosphorylation sites that can be analysed at a given time and the unpredictability of incubation time and incorporation efficiency.

1.3.2.2 Phospho-specific Antibodies

The rapid development of phospho-specific antibodies gave researchers alternative ways to identify and quantify protein phosphorylation. Nowadays, the

availability of hundreds of phospho-specific antibodies boosts the development of different methods such as antibody-based Western blotting, ELISA or immunoprecipitation [130]. The widespread use of antibodies for assessing protein phosphorylation not only eliminates the problems of dealing with radioisotopes, but also promotes the development of new anti-phospho antibodies. Antibody-based methods are sensitive and feasible, especially for tyrosine-phosphorylated proteins due to its relative low abundant. In addition, antibody-based ELISA, is not only highly sensitive because of utilizing two antibodies against target proteins, but also can be used for quantification [131]. Despite many advantages of antibody-based methods, however, the limitation of these methods is that they are only applicable to limited numbers of proteins and the costs of antibodies are very high. Phosphorylation of proteins in known to occur in so many proteins that antibody-based methodologies are not sufficient to study phosphoprotein changes on a global scale.

1.3.2.3 Phosphoproteomics

Mass spectrometry-based phosphoproteomics offers an unbiased and powerful option to analyse phosphorylation events on a global scale and it is capable of accurately identifying and quantifying thousands of phosphorylation sites in a single experiment [132,133]. More than 50,000 human phosphorylated peptides were detected using mass spectrometry-based phosphopteomics from Matthias Mann's group in 2014 [134]. More importantly, those phosphorylation events can be quantified across various conditions, which led to many important findings. For example, Francavilla et al. [135] used phosphoproteomics to study changes in the phosphoproteome of epithelial ovarian cancer (EOC), which led to the identification of cyclin-dependent kinase 7 as a potential therapeutic target for EOC. Kim et al. [136] investigated TANK-binding kinase 1 (TBK1) signalling in lung cancer cells by phosphoproteomics, which revealed its mitosis regulator function.

Phosphoproteomics is a specialized application of bottom-up proteomics. The proteins are digested with trypsin - or a combination of trypsin and endoproteinase Lys-C - into peptides. Instead of directly injecting the digested peptides into mass spectrometer, an additional enrichment step is used to increase the abundance of phosphopeptides. The enrichment is mostly performed using Immoblized Metal Affinity Chromatography (IMAC) or Titatnium dioxide (TiO_2), although there are some variants,

such as Sequential Elution from IMAC [137] or Ti^{4+} IMAC [138], which combine or modify IMAC and TiO_2 methods.

1.3.3 Sample Preparation and Processing in Proteomics and Phosphoproteomics

1.3.3.1 Protein Isolation, Digestion and Fractionation

In bottom-up proteomics, sample preparation starts with protein extraction. The proteins can be extracted from any type of sample including cells, tissues, blood, urine or even plants. The extracted proteins are then digested with a protease, normally with trypsin, generating tens of thousands of peptides and this peptide mixture is then loaded onto a High Performance Liquid Chromatography (HPLC) system [139] and sequenced by a mass spectrometer.

Comprehensive protein identification requires sophisticated protein separation techniques to reduce sample complexity. Although the latest mass spectrometers combine high resolution with fast scan speeds, they are still not advanced enough to characterize a full proteome in a single run due to the enormous complexity of a proteome. In order to reduce sample complexity and to identify even low abundant proteins, a sample can be pre-fractionated prior to MS analysis. Sample separation techniques have seen tremendous advances during the last decades ranging from gel-based separations to chromatographic-based fractionations [105]. Chromatographic-based separation techniques have been applied for peptide fractionation in numerous studies [140,141] and significantly increased the number of identified proteins.

1.3.3.2 Phosphopeptide Enrichment by Immobilized Metal Affinity Chromatography

Phosphorylated proteins are present at relatively low abundance in comparison to nonphosphorylated proteins, so it is necessary to enrich, in an unbiased manner, for phosphorylated peptides derived from proteolytic digestion of proteins on a large scale [125]. Immobilized metal affinity chromatography (IMAC) and metal oxide affinity chromatography (MOAC) are two of the most extensively used techniques [142,143].

Immobilized metal affinity chromatography (IMAC) enriches phosphopeptides by utilizing the affinity of positively charged metal ions to the negatively charged phosphate groups (Figure 1.4A). Conventionally, metal ions such as Fe^{3+}, Al^{3+}, Ga^{3+},

Zr^{4+} and Ti^{4+} are chelated onto a solid support (iminodiacetic acid or nitrilotriacetic acid). At low pH, the phosphate groups of the phosphopeptides are deprotonated but the carboxyl groups of peptides are protonated [144]. The negatively charged phosphate groups can bind to the positively charged metal ions. After removal of non-specifically bound peptides, the phosphopeptides can be eluted at high pH [130]. However, it is not always easy to protonate all the non-phospho peptides, especially those containing glutamate and aspartate, as this requires highly acidic pH, which may cause the metal ions to dissociate from their solid supports [144,145].

Recently, Zhou and Yu et al. [146-148] introduced new solid support-monodisperse microspheres possessing excellent stability, uniform size distribution and high column efficiency. Compared to Fe^{3+}-IMAC adsorbent, Ti^{4+}-IMAC microspheres showed high specificity and efficiency for enrichment of phosphopeptides. In 2010, Iliuk et al. [149] introduced a new metal ion-based reagent, termed polymer-based metal ion affinity capture (PolyMAC), which demonstrated outstanding reproducibility, exceptional selectivity, fast chelation times and high phosphopeptide recovery from complex mixtures. The PolyMAC strategy consists of two reaction steps: a homogeneous chelation between phosphopeptides and the PolyMAC-Ti reagent in solution phase and a heterogeneous reaction between the PolyMAC reagent and the solid phase support, which can be achieved by magnetic beads or hydrazide-agarose gels. This method is gaining in popularity and a commercial product is now available [150-153].

1.3.3.3 Phosphopeptide Enrichment by Metal Oxide Affinity Chromatography

Metal oxide affinity chromatography (MOAC), which uses a similar methodology to IMAC, incorporates metal oxides onto a solid support in place of metal ions (Figure 1.4B). The oxygen in the phosphate group of phosphorylated peptides can interact with the metal oxides via bidentate binding of the MOAC resin [154]. TiO_2, ZrO_2, Fe_3O_4, SnO_2, and CeO_2 can function as phosphopeptides enrichment metal oxides, although TiO_2 is most commonly used [143]. In 2004, Batalha et al. [155] reported the use of TiO_2 as a chromatographic medium for selectively enriching phosphopeptides combined with 2D LC-MS/MS, which enabled separation of phosphorylated peptides from nonphosphorylated peptides. Since then, numerous publications have appeared in which phosphopeptides have been enriched with TiO_2 and there are many commercial products available from different companies. In comparison to IMAC, metal oxide

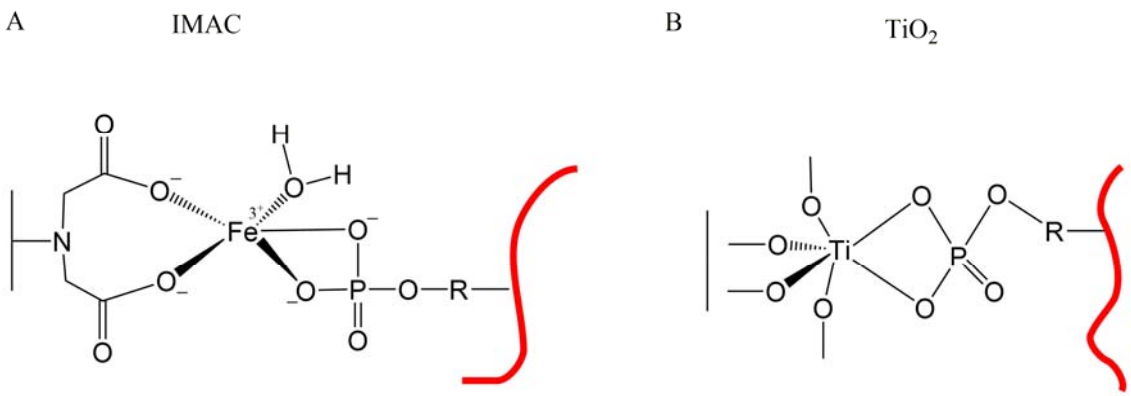


Figure 1.4. [157] **Comparison of the binding mechanisms of IMAC (A) and TiO_2 (B).** The red line indicates the peptide chain; R is the side chain. Both Fe^{3+}-IMA and TiO_2 are attached to certain matrices. In Fe^{3+}-IMAC, the positively charged Fe^{3+} ion binds to the negatively charged phosphate group. Similarly, in TiO_2 resin the oxygen atoms of the phosphate groups interact with the Ti^{4+} ion.

affinity chromatography performs better at low pH and does not present the drawback of metal ions leaching [145]. However, TiO_2 has a preference for mono-phosphorylated peptides, whereas IMAC is superior for multiply phosphorylated peptides, so some studies are also using combined TiO_2 and IMAC procedures [156].

1.3.4 Mass Spectrometry Instrumentation for Proteomics

The past 20 years have seen a tremendous increase in mass spectrometry-based proteomic studies due to significant advances in mass spectrometric instrumentation [158], ionisation techniques, mass analysers and fragmentation patterns. Several types of mass spectrometers are available in the market nowadays and the following section summarizes the most important ionization techniques, mass analysers and fragmentation methods.

1.3.4.1 Ionisation Techniques

One of the most important breakthroughs for the development of mass spectrometry-based proteomics was the establishment of soft ionization techniques: MALDI and ESI (Figure 1.5). Due to their tremendous contribution for analysing biological molecules John B. Fenn and Koichi Tanaka, who implemented these methods in modern biology, shared the Nobel Prize in Chemistry in 2002 [159].

1.3.4.1.1 Matrix-assisted Laser Desorption/Ionization

In matrix-assisted laser desorption/ionization (MALDI) mass spectrometry, a specific matrix consisting of a small crystallized UV-absorbing organic compound is mixed with the analyte of interest on a metal plate. The analyte recrystallizes with the matrix and is subsequently irradiated with a pulsed laser. The matrix absorbs the laser energy, promoting the desorption and ionization of the analyte of interest. The protonated analyte, which is usually singly charged by the addition of a proton ($[M+H]^+$), is then detected in a mass spectrometer [4,160]. The advantage of MALDI is that the resulting spectra are relatively simple and can be easily interpreted; however, lack of reproducibility can be a major issue with MALDI [161] [160].

1.3.4.1.2 Electrospray Ionization

Electrospray ionization (ESI) has become the most widely used ionization technique over the last decade. In contrast to MALDI, an electrospray source is directly

connected to a liquid chromatography (LC) system, which can provide a continuous flow and functions as a separation tool to reduce sample complexity. The eluent from the LC column is pushed through a voltage-containing needle tip and thus ionised. This ionization process can be briefly described as follows: high voltage is applied to the volatile solvent to create an aerosol through the glass or metal needle and the solvent then experiences a series of evaporation steps, transferring the charged ions from the solution into the gaseous phase, with the aid of high ESI-source temperature and vacuum [162] [4]. Currently, the ion evaporation model (IEM) and the charge residue model (CRM) are the two major theories for the production of gaseous phase ions. In IEM, the droplets from the spray continue evaporating until the electrical field energy reaches a certain point, which expels the ions from its surface into the gaseous phase; in contrast, CRM assumes that the droplet only contains a single analytical ion and the process is due to the self-formed Taylor Cone, which refers to the cone formed in electrospraying of mobile phase and repeated emission of charged droplets [163-165]. Figure 1.5 illustrates ionisation techniques of MALDI and ESI.

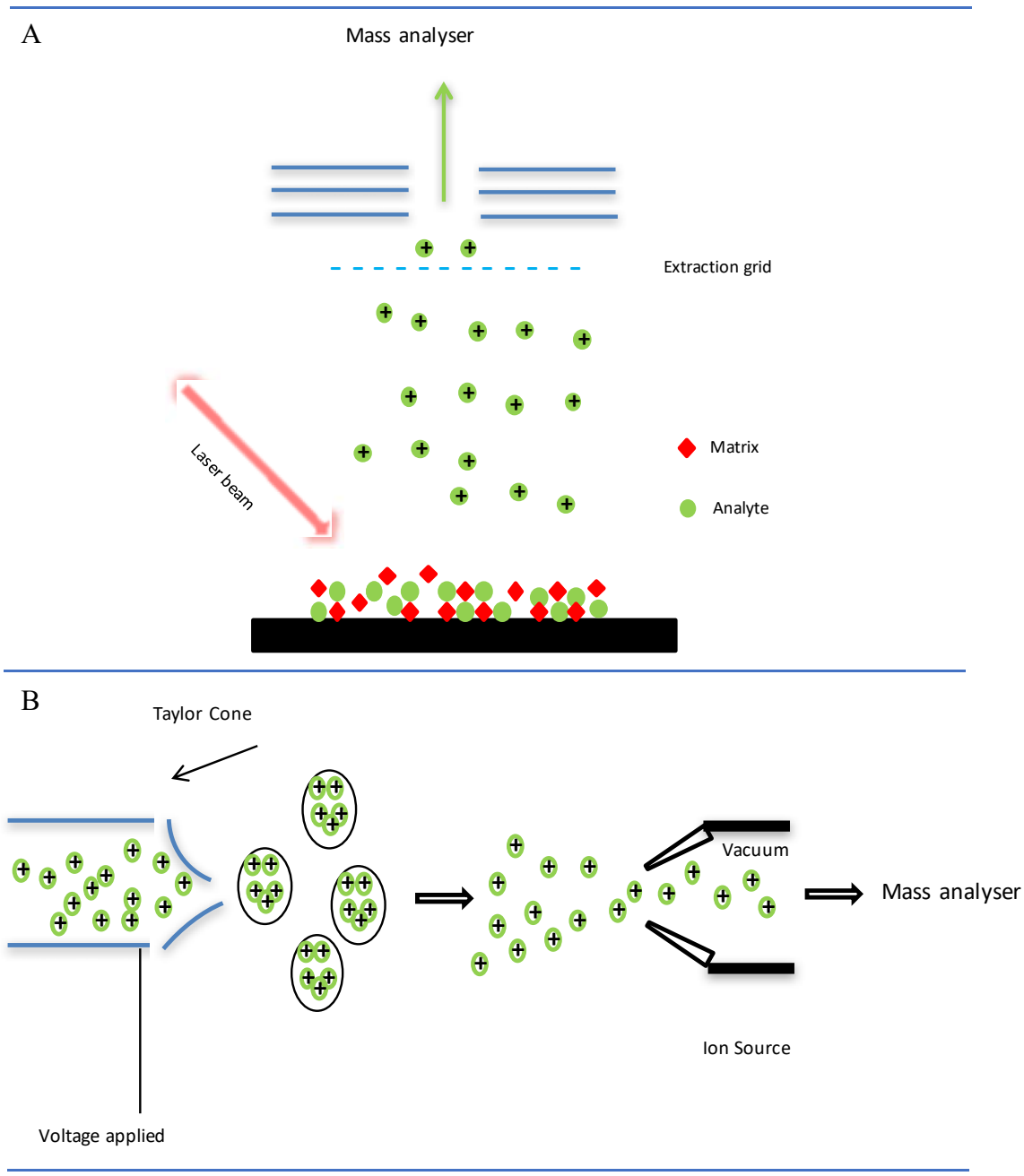


Figure 1.5. Schematic illustrations of soft ionization techniques. (A) MALDI. (B) ESI. A, a laser beam is used to protonate and ionise the analyte, which is co-crystallized with matrix molecules on a plate. The protonated analyte is then detected by mass spectrometry. B, the analyte is pushed through a thin 'needle-like' tip onto which a voltage is applied. The analyte becomes ionised and is then transferred and detected in a mass spectrometer.

1.3.4.2 Mass Analyser

After the ionization process, the ions from the peptides or proteins are resolved according to their mass to charge ratio (m/z) in the mass analyser. Currently, four of the most commonly used types of mass analysers are quadrupole (Q), time-of-flight (TOF), ion trap (IT) and Fourier transform (FT) mass analysers [166].

1.3.4.2.1 Quadrupole Mass Analyser

A quadrupole mass analyser, as implied by its name, contains four cylindrical rods, which are parallel to each other. It acts as a mass filter and only ions with the selected mass-to-charge ratio can pass through the pathway between the four rods (Figure 1.6). Each opposite rod forms a pair of electrodes, applied with a DC (constant) offset voltage and an AC (alternating) voltage. The DC voltage directs the ions towards one pair of the electrodes, while the AC voltage guides the ions through the rods. Ions that are either too large or small will be filtered out and only the ions with the right size can pass through the pathway between the electrodes and be analysed [167].

For tandem mass spectrometry, a quadrupole is often part of a triple quadrupole mass spectrometer [168,169]. The parent ion or precursor ion is selected in the first quadrupole (Q1) and fragmented in the centre or second quadrupole (Q2). The third quadrupole (Q3) can be employed to scan the entire m/z range or only certain product ions. Such QQQ instruments suffer from low resolution (500-1500 Full Width at Half Maximum (FWHM))) and low mass accuracy, but show outstanding ion transmission efficiency with over 90% transmission values and are often used for multiple reaction monitoring (MRM) or selected reaction monitoring (SRM) experiments [4]. A quadrupole is nearly always used as the first mass filter in modern hybrid instruments but may be combined with other mass analysers such as time-of-flight (TOF) or Orbitrap.

1.3.4.2.2 TOF Mass Analyser

The concept of the Time-Of-Flight mass analyser was first introduced by Stephens in 1946 [170] and the instrument was designed and named a ‘Velocitron’ in 1948 [171]. The ions in this instrument are accelerated with the same kinetic energy in an electric field towards to the detector (Figure 1.6). The time it takes for the ions to fly through the free-field path (a flight tube) depends on their velocity, which is determined by their mass and charge. Thus, small ions will travel faster than large ions in the same

charge state. Therefore, the ions are separated by their mass-to-charge ratio and the masses of the ions can be calculated based on known instrument parameters [172,173].

The simplest time-of-flight mass spectrometer is the linear mass analyser, which provides moderate mass resolution, but fast acquisition speed. However, due to limitation of the detection speed as well as space charging effects between ions, simple linear mass analyser are only rarely in use [174]. In order to improve the mass resolution and increase the limitation of detection, the ion mirror was proposed and built by the Russian scientists Alikhanov [175] and Mamyrin [176]. The ion mirror, also referred as a reflectron, has been applied to most modern TOF mass analysers. The reflectron is used to send the ions back through the flight tube multiple times, which can expand the time difference between ions, resulting in increasing the mass resolution from 100 to more than 10,000 FWHM [174,177]. Currently, the resolving power of TOF instruments can reach up to 40,000 FWHM, resulting from the improvement of a series of factors, such as ion production, ion sampling, ion detection and high-speed digitisation [178].

1.3.4.2.3 Ion trap Analyser

In ion trap mass analysers, the ions are first trapped in a ring electrode with two ellipsoid end caps and then expelled out of the ring electrode, according to their m/z ratio, by applying a resonant frequency voltage [173]. Ion trap mass analysers are robust, very simple and relatively inexpensive. However, the classical three-dimensional ion trap suffers from low mass accuracy due to poor ion trapping efficiencies [166] [179]. In the alternative two-dimensional or linear ion trap (LIT), ions are trapped in a two-dimensional radio frequency (RF) field by using a modified quadrupole. The trapped ions can be ejected in a direction perpendicular to the central axis of the trap via radial excitation [179]. In comparison to the traditional three-dimensional ion trap, the two-dimensional ion trap can provide high ion trapping capacity resulting in increased sensitivity and accuracy of the mass analyser [173,180]. Due to the property of trapping ions, ion traps can be combined with other mass analysers and function as ion storage devices or mass filters to isolate selected ions and they can also be used as stand-alone mass spectrometers [180]. The most commonly known and commercially available ion trap mass spectrometers are the LCQ and LTQ from Thermo Finnigan (now Thermo Fisher Scientific). The LCQ is a 3D ion trap mass analyser, while the LTQ is a 2D ion trap mass analyser [181]. Currently, stand-alone ion trap mass spectrometers are not the

main stream instruments for peptide or protein identification; however, they have been used in hybrid mass spectrometers such as the Thermo Fusion Tribrid mass spectrometer, which contains an ion trap as one of its mass analysers.

1.3.4.2.4 Fourier Transform Mass Analyser (FT-MS)

In Fourier transform mass analysers (FT-MS), the m/z ratio of ions is measured based on their excitation frequency, which can be measured very accurately. Therefore, FT-MS can offer much higher resolving power and mass accuracy than ion stability-based or ion transit time-based mass analysers [182-184]. The most widely known Fourier transform mass analysers are the Fourier transform ion cyclotron resonance (FT-ICR) and Orbitrap mass analysers.

The FT-ICR instrument was introduced by Comisarow and Marshall in 1974 [185]. Ions rotate in a cyclotron magnetic field and the cyclotron frequencies of ions reaching the detector are recorded. The detected signals are converted by Fourier transformation to the m/z ratios of ions, i.e. mass spectra [186]. So far, FT-ICR provides the highest mass resolution with up to 1,000,000 FWHM. However, due to the high complexity and cost of FT-ICR instruments, Orbitrap is the mainstream mass analyser in mass spectrometry-based proteomics.

The concept of “orbital trapping” was raised by Kingdon in 1923 [187]. In order to memorialise his contribution, an orbital trap is also called a Kingdon trap. In a Kingdon trap, a thin-wire central inner electrode is surrounded by a coaxial outer electrode. A voltage applied to the wire creates a radial logarithmic field. Ions with polarity opposite that of the central voltage can orbit the wire if given appropriate starting conditions. Endcaps are placed at the ends of the wire/cylinder to confine the ions axially. As a result, ions both orbit around the wire and oscillate down the length of the trap. However, due to the weak resonance of axial and radial ions, it didn’t produce mass spectra until the Orbitrap mass analyser was introduced by Makarov in 2000 [188,189]. An Orbitrap is a modified Kingdon trap mass analyser with special inner and outer electrodes [189]. The ions are trapped and rotating in a spindle-shaped electrostatic cell without a magnetic field and the orbital motion of ions is achieved through a combination of electrostatic attraction between the inner and outer electrodes, a centrifugal force and initial tangential velocity [4,190]. The axial field induces oscillation frequencies, which is proportional to m/z and that can be detected as an image current and then translated to accurate m/z and

mass spectra through Fourier transform [190]. As indicated before, ion frequency based Fourier transformation can give highly accurate m/z ratios. Therefore, Orbitrap is a high resolution mass analyser. The first commercialized Orbitrap mass analyser was introduced in 2005 by Thermo Electron, Bremen, Germany (now Thermo Fisher Scientific) [191]. It is a hybrid linear ion trap-Orbitrap mass spectrometer (LTQ Orbitrap Classic) with the resolution ranging from 60,000 to 100,000 FWMH at m/z 400 [189,190]. After a decade of development, the resolution of the latest Obritrap mass analyser (Orbitrap Fusion Lumos Tribrid) can reach 500,000 FWHM at m/z 200 with <3 ppm mass accuracy using external calibration and <1 ppm with internal calibration [192] [4].

1.3.4.3 Peptide Fragmentation Patterns in Mass Analyser

To obtain detailed information of the amino acid sequence of a peptide, the peptide needs to be fragmented in the collision cell of a mass spectrometer. The most common collision cells are based on collision-induced dissociation (CID), higher-energy collisional dissociation (HCD) and Electron-transfer dissociation (ETD) and electron-capture dissociation (ECD). Figure 1.7 illustrates typical fragment ions that are generated by CID (b and y ions), HCD (b and y ions) or ETD/ECD (c and y ions). Depending on where the peptide backbone breaks, the resulting fragment ion pairs are called a / x, b /y or c / z ions. A, b and c ions contain the N-terminus of the peptide, while x, y and z ions contain the C-terminus.

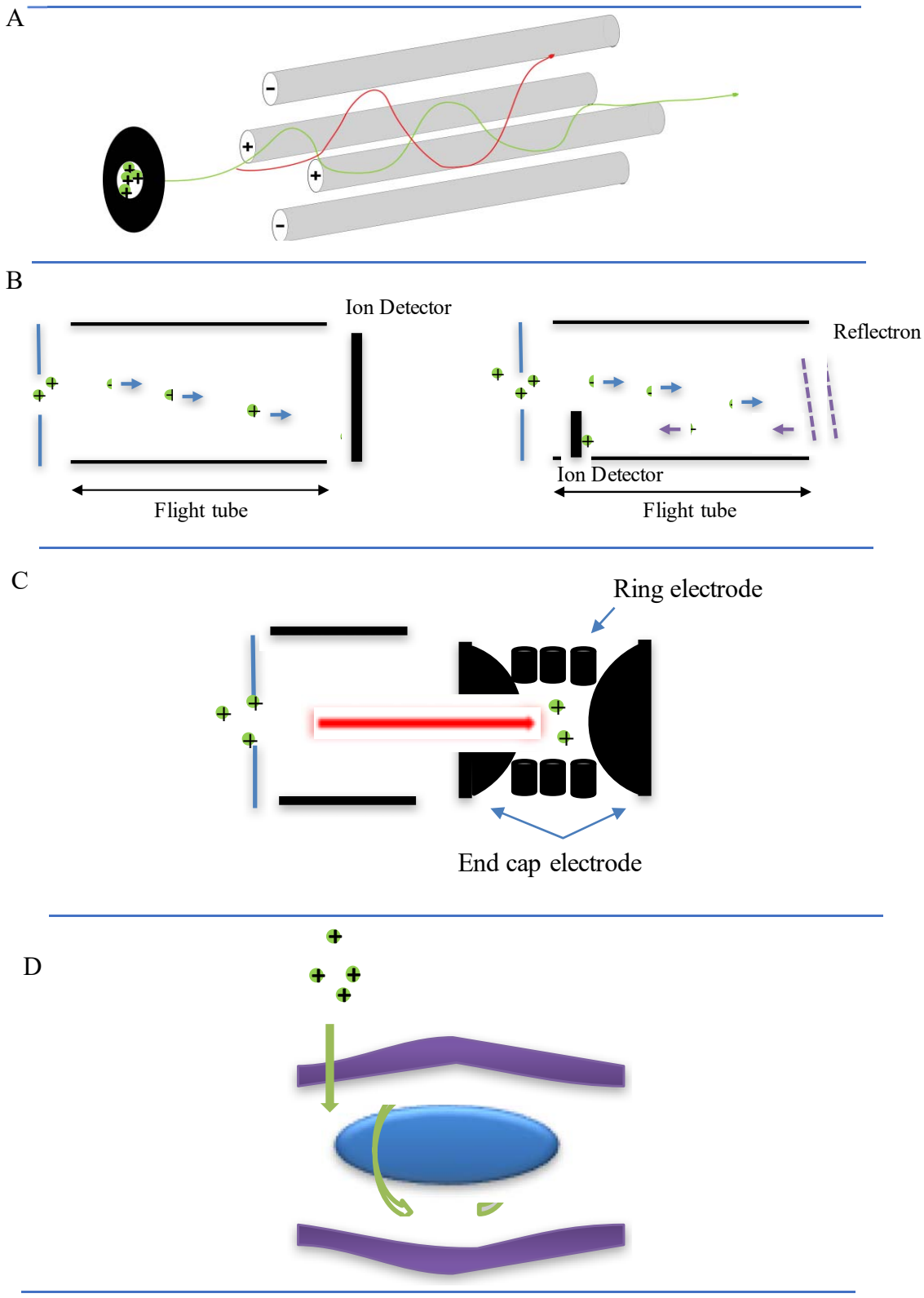


Figure 1.6. A schematic representation of 4 mass analysers. (A) quadrupole. (B) TOF. (C) ion trap. (D) Orbitrap. The green dots with a plus sign indicate positively charged ions, which are detected by different types of mass analysers Additional details are given in the text.

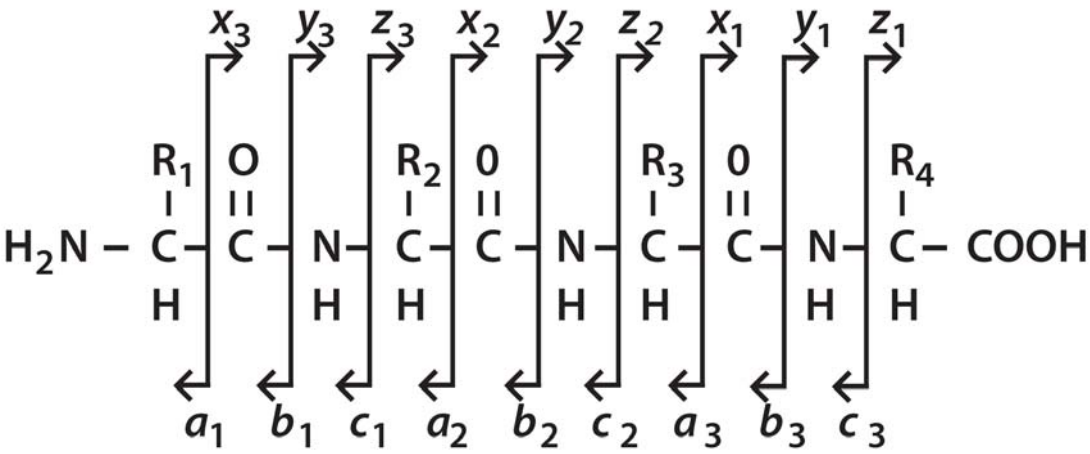


Figure 1.7. Different ion types generated by mass spectrometry from peptide backbone cleavages. a, b, c, x, y and z indicate six types of ions; R_1, R_2, R_3 or R_4 represents side chains.

1.3.4.3.1 Collision-Induced Dissociation

Collision-induced dissociation (CID), sometimes referred as collisionally activated dissociation (CAD), is the most widely used fragmentation method for proteome identification and quantification analysis. The ions are accelerated with high kinetic energy into the CID cell, where they collide with neutral molecules, often helium, nitrogen or argon. The vibrational energy of the impact is redistributed to the peptide/protein ion and can result in molecular dissociation reactions often occurring at the labile amide bonds along the peptide backbone, thereby generating b- and y-type fragment ions (Figure 1.7) or leading to losses of neutral atoms or small molecules [193]. CID can be used for detecting stable post-translational modifications, but it is not suitable for analysing labile modifications such as, phosphorylation or sulfation as those modifications can be lost in a charge separation process or by β-elimination events with a neutral loss of phosphoric acid or sulphur dioxide [194].

1.3.4.3.2 Higher-Energy Collisional Dissociation

Higher-energy collisional dissociation (HCD, formerly higher-energy C-trap dissociation) is a mass spectrometry fragmentation technique specific to the Orbitrap mass spectrometer in which fragmentation takes place external to the trap [195]. Compared to traditional ion trap-based CID, HCD uses higher activation energy and shorter activation times. HCD fragmentation combined with Orbitrap detection has no low-mass cut-off, relatively low intensities of neutral loss peaks, and increased yields of ion fragments resulting in higher quality MS/MS spectra [196], which make HCD better for identifying phosphorylation sites.

1.3.4.3.3 Electron-Transfer Dissociation and Electron-Capture Dissociation

Electron-transfer dissociation (ETD) and electron-capture dissociation (ECD) are complementary to CID and HCD. During the process of fragmentation, fluoranthene radical-anions are generated in a chemical ionization source and used as reagents to transfer an electron to a multiply charged peptide generated by electrospray ionization. This reaction is highly exothermic, reduces the peptide charge by one, and triggers fragmentation of the peptide backbone to produce a homologous series of complementary fragment ions of the c and z type (Figure 1.7), whereas CID and HCD always generate b and y ions [197]. Importantly, labile modifications can be retained by ETD and ECD.

There are some areas in which ETD and ECD are very beneficial, for example for the analysis of intact histones or histone tails essentially necessitates ETD or ECD [198,199]. As a downside, ETD or ECD does not result in identification of as many proteins as CID or HCD. In addition, ETD or ECD cells (as well as HCD cells) are more expensive than CID cells and need to be tuned carefully.

1.3.4.4 Tandem Mass Spectrometer

Modern mass spectrometers are capable of measuring m/z ratios at high resolution with accurate mass [200]. However, proteomic sample can be highly complex, especially when digested into small peptides for global proteome characterisation using bottom-up proteomics. A single mass analyser is not sufficient to provide unambiguous information for peptide identification. Therefore, tandem mass spectrometers, which usually contain two mass analysers, are designed to acquire the full mass (MS1) and the corresponding fragment masses (MS2) for unambiguous characterisation of peptides and proteins [4]. Currently, the most commonly used tandem mass spectrometers in global proteome identification are hybrid instruments, such as Quadrupole Time of Flight (Q-TOF) and Quadrupole Orbitrap (Q-OT).

1.3.4.4.1 Quadrupole Time of Flight

Q-TOF or triple-TOF instruments combine the merits of TOF and quadrupole analysers. TOF analysers can provide spectra at high resolution with accurate mass and quadrupoles offer highly efficient ion transmission. The simplest way to describe a Q-TOF is that the last quadrupole of a triple quadrupole instrument is replaced by a TOF mass analyser instrument. The structure of a triple-TOF instrument is shown in Figure 1.8A. Q3 is replaced by a TOF analyser and an additional Q0 is added before Q1. Only radio frequency (RF) voltages are applied to Q0 and Q2, which are more frequently replaced by hexapoles or octopoles. When the ions enter the mass spectrometer, Q0 functions as a collisional cooling and ion focusing component; Q1 acts as a mass filter and sends selected ions into Q2 for fragmentation. The ion modulator, which is located after Q2, accelerates the fragment ions in a perpendicular direction into the TOF analyser for spectral acquisition [200-202]. One of the latest instruments, a TripleTOF 6600 system (Sciex) can offer 35,000 resolution in MS mode, 30,000 resolution in MS/MS mode and up to 100 MS/MS scans per second.

1.3.4.4.2 Quadrupole Orbitrap

The concepts of quadrupole and Orbitrap have been described above. Generally, Q-OT instruments utilise the high ion transmission efficiency of a quadrupole and the high resolving power of an Orbitrap mass analyser. The first mass spectrometer that combined quadrupole and Orbitrap was called Q Exactive and launched by Thermo Fisher Scientific in 2011 [203]. Q Exactive mass spectrometers were used for the experiments described in this thesis. A brief history of the development of the Q Exactive instrumentation is given below.

As indicated above, the first tandem Orbitrap mass spectrometer was a hybrid linear ion trap-Orbitrap mass spectrometer (LTQ Orbitrap Classic; Figure 1.8 B). The high resolution of the Orbitrap mass analyser is exploited for the full MS scan, while high-speed ion trap scans are used for the acquisition of all subsequent MS2 scans. However, CID was the only fragmentation mechanism available in the LTQ Orbitrap Classic, which limited its usage to basic peptide identifications. The implementation of HCD (Higher-energy collisional dissociation) or ETD (Electron transfer dissociation) as alternative fragmentation mechanisms led to the development of the so-called LTQ Orbitrap XL in 2007, which allowed a faithful analysis of post-translational modifications (PTMs), such as phosphorylation and glycosylation. Both HCD and ETD are capable of keeping PTMs stable and thus, facilitate proper spectra assignment. The LTQ Orbitrap XL was further enhanced through the addition of new ion guides (S-lens) and improved dual-pressure ion trap mass analysers. These modified instruments appeared on the market in 2009 and 2011 and were named LTQ Orbitrap Velos and Elite, respectively [192,204,205].

The first stand-alone Orbitrap mass analyser, known as Exactive Classic, was developed in 2008, [206]. It consisted of an Orbitrap analyser, a HCD cell and a C-trap, which is a curved electrode used to capture the fragments from the HCD cell and then inject them simultaneously into the Orbitrap analyser. This C-trap was first developed in 2007 and is an essential part of all subsequent Orbitrap mass spectrometers [195]. The Exactive is capable of MS1 and MS2 scans without precursor selection, which is called all-ion fragmentation (AIF), but due to lack of ion selection modules, it is rarely used in proteomics [205].

The Q Exactive, which combined a quadrupole with an Orbitrap mass analyser, became commercially available in 2011. Various subsequent improvements led to the development of the Q Exactive Plus, Focus and HF [203,207]. All Q Exactive mass spectrometers are currently considered to be the best mass spectrometers for global proteomics including protein identification as well as protein quantification. Their design is shown in Figure 1.8 B. A S-lens ion guide (a stacked radio frequency device) is applied at the front end and increases ion transfer efficiency through ion focusing [208]; a bent flatapole is used for collisional cooling and for removal of solvent droplets and other neutral species [207]. This is followed by a quadrupole for ion selection and filtration. A C-trap is located after the quadrupole to store ions either from the quadrupole selection or the HCD cell fragmentation. The selected ions are then injected into the Orbitrap analyser. Due to this simple design, Q Exactive mass spectrometers require minimal maintenance, which make them the perfect choice for many laboratories around the world.

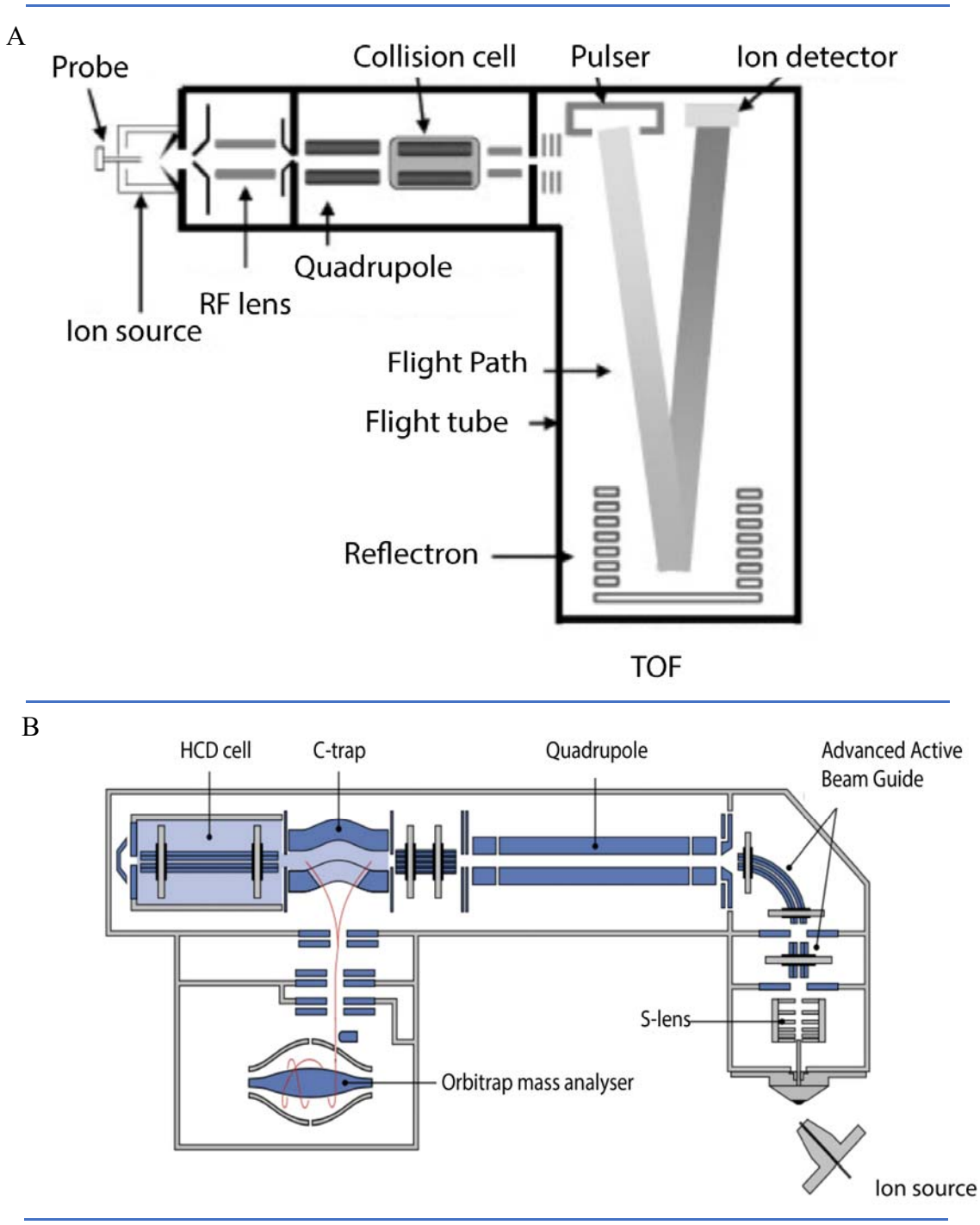


Figure 1.8. [209] **A schematic representation of Q-TOF (A) and Q-OT (B).** A: Ions are injected through ion source, focused to an ion beam in the RF lens, selected in the quadrupole, fragmented in collision cell and detected in the TOF mass analyser. B: Ions are injected through ion source, selected in S-lens and quadrupole, fragmented in HCD cells and detected in Orbitrap mass analyser.

1.3.5 Quantitative Proteomics and Phosphoproteomics

The recent developments of mass spectrometric instrumentation in combination with novel sample preparation methods have seen mass spectrometry as an extraordinary tool for the identification of proteins and phosphoproteins. However, for most biological questions, it is of utmost importance to be able to quantitate the proteomic changes across different conditions.

Stable isotope labelling is the most frequently used method for quantification of protein levels and protein phosphorylation levels by mass spectrometry. The proteins or peptides are incorporated with non-radioactive isotopes and after combing the samples from different conditions, the same peptides with different isotopes are analysed simultaneously and appear as pairs (doublets) in the mass spectra. The relative intensities of the doublet components directly reflect the relative quantities of the peptide in the two proteomes to be compared [133]. The most commonly used methods are metabolic incorporation of amino acids, chemical modification with isobaric tags, dimethyl labelling, and label-free quantitation [210].

1.3.5.1 Stable Isotope Labelling by Amino acids in Cell culture (SILAC)

Stable Isotope Labelling by Amino acids in Cell culture (SILAC), is a metabolic labelling strategy for incorporating “labels” into proteins through the addition of stable isotope-labelled amino acids in growth medium [211]. When labelled amino acids are supplied to cells in culture instead of the natural amino acid, they are incorporated into all newly synthesized proteins. After a number of cell divisions, each instance of this particular amino acid will be replaced by its isotope labelled amino acids. Since there is no chemical difference between the labelled and the natural amino acid, the cells behave exactly like the control cell population grown in the presence of the normal amino acid. Because of its high efficiency and reproducibility, SILAC has been used in many organisms such as yeast, bacteria, *Plasmodium falciparum*, plants and even in tissues using ‘spike-in’ SILAC [212] for quantifying proteome and phosphoproteome.

1.3.5.2 *In Vitro* Labelling Techniques

In contrast to SILAC, isobaric tagging and dimethyl labelling are *in vitro* labelling techniques. Instead of labelling the proteins *in vivo*, the proteins are extracted and digested into peptides separately, which are then labelled with different isotopes and

combined according to the experimental design. Isobaric Tags for Relative and Absolute Quantification (iTRAQ) is a set of reagents that widely used for isobaric tagging that is not only applicable for samples from cell culture or tissues but can also be used to differentially label and eventually mix up to 10 samples. Dimethyl labelling of peptides by formaldehyde is another relative robust, efficient and inexpensive method. It can be only used to compare up to three different samples, but it is a reasonable alternative because of its high efficiency and cheap price. Both of these two methods have already been successfully applied in studying proteome and global phosphorylation [210,213-215].

1.3.5.3 Label-free Quantification (LFQ)

Unlike protein labelling quantification methods, label-free quantification (LFQ) does not use isotopes for labelling the peptide or protein. This method assumes that identical peptides from different mass spectrometry experiments can be compared directly across any number of samples [216], but it requires rigorous attention to reproducibility and demands high standards for sample preparation, chromatographic separations and performance of the mass spectrometer. Despite these drawbacks, LFQ has been applied not only in global proteome quantification, but also in quantifying phosphoproteome [217,218].

1.3.6 Mass Spectrometric Acquisition Modes

Mass spectra graphically summarize all ions that hit the detector at a given point in time. They are typically XY graphs in which the m/z values of all ions are plotted on the X axis against their relative abundance on the Y axis. However, the mass spectrometric acquisition mode dictates which and how many individual ions are permitted through to the detector and thus, has a major impact on mass spectra and mass spectrometric data in general. The following mass spectrometric acquisition modes will be discussed in the following section: data-dependent acquisition, targeted acquisitions (MRM, SRM, PRM) and data-independent acquisition.

1.3.6.1 Data-Dependent Acquisition

Data-Dependent acquisition (DDA) is a well-established data acquisition mode used for protein identification and quantification. The concept is that, due to the limited scan speed of mass spectrometers, a pre-scan or survey scan (MS1 scan) determines the

intensity of all ions entering the mass analyser at a given time point. The 10-15 most abundant precursor ions are then selected for fragmentation and sequencing in subsequent MS2 scans. With the introduction of dynamic exclusions that the same ion will be excluded for a certain time to avoid repeated sequencing of the same precursor ion, which significantly increases the number of identified peptides [219]. Quantification is achieved by calculating the intensity of each corresponding peptide peak from the MS1 scans. The MS2 scans (together with the exact precursor masses from the MS1 scans) are used for characterisation.

Since the DDA acquisition mode is "data (intensity) dependent", it is biased towards abundant proteins while disfavouring low abundance proteins. Additionally, it suffers from a lack of reproducibility and less accurate quantification than some other methods. As indicated in some publications, the reproducibility is only 60-80%, even when the same sample is injected twice [220].Despite these drawbacks, DDA is still the method of choice in many labs around the world with thousands of proteins being detected in every run. A simple literature search using "data dependent acquisition" as query returned more than 10,000 publications (June 2017).

1.3.6.2 Targeted Acquisition- SRM, MRM and PRM

DDA is nearly always used for hypothesis generation, which results in a protein list that has to be further validated. Antibody-based validations such as Western blotting or fluorescence-activated cell sorting (FACS) is preferred in many laboratories. However, these validation methods rely on the availability and quality of specific antibodies [221]. Targeted mass spectrometry offers a better alternative. In contrast to DDA, which uses automatic precursor selection, targeted mass spectrometric approaches are designed to measure specific pre-defined precursor ions (as well as their fragments), which results in more reliable and reproducible quantification and which is not limited by the availability of commercial antibodies.

Two of the most commonly applied acquisition methods in the proteomics community are parallel reaction monitoring (PRM) and selected reaction monitoring (SRM), also known as multiple reaction monitoring (MRM). SRM/MRM is a method specific to triple-quadrupole based mass spectrometers. Q1 functions as a peptide precursor ion "sieve" (MS1 scan), and only a pre-defined precursor ion can pass through this filter. The precursor ions are then fragmented in Q2, the collision cell. The second

pre-defined transitions (fragmented/product ions) are filtered in Q3 and reach the detector for intensity determination (MS2 scan). In contrast, PRM is applied in Orbitrap-specific instruments. Similar to SRM/MRM, a quadrupole is used as a precursor ion filter; however, instead of analysing only one predefined fragment ion, full MS2 scans of all daughter ions are analysed in a Orbitrap mass analyser. Peptide identification by either SRM/MRM or PRM is based on a pre-acquired spectral library, which contains information such as m/z, retention time and MS2 fragments for each precursor ion.

In comparison to DDA, there are some major advantages of SRM/MRM or PRM. First, the targeted methods have high selectivity and specificity due to the targeted acquisition mode of the mass analyser, which increases the signal-to-noise ratio between analyte and background, this facilitating detection of low abundance proteins [222]. Second, the targeted methods have better mass accuracy and improved quantification. The quantification is based on profiled MS2 peak intensities, the data of which is less overlapped and therefore more accurate. Due to these benefits, these targeted proteomic methods have been applied in not only in targeted protein quantification [223] but also in targeted quantification of post-translational modified peptides, as is required for phosphoproteomics [224]. The research described in Chapter 3 of this thesis was achieved by combining DDA and PRM approaches.

1.3.6.3 Data-Independent Acquisition

DDA is the method of choice for maximizing protein characterisation, while SRM/MRM or PRM is ideal for reliably quantifying several proteins. The alternative approach of data-independent acquisition (DIA) tries to combine the advantages of DDA as well as of targeted methods such as MRM [225]. This method is made possible by the improved resolving power, acquisition speed and sensitivity of the latest mass spectrometers. A mass spectrometer operating in DIA mode can fragment all precursor ions in pre-defined m/z windows and acquire combined MS2 spectra for quantification [226]. DIA has the advantage that it can eliminate the effect of peptide/protein bias and lack of reproducibility encountered in DDA. Moreover, it can improve the quantification accuracy, a result of quantification based on MS2 spectra.

1.3.6.4 Comparison between DDA and DIA

As described in previous sections, there are advantages and disadvantages for both DDA and DIA. The following paragraph summarizes and compares DDA and DIA in greater detail. Firstly DDA is the 'traditional' way to identify proteins by mass spectrometry and benefits from the fact that each fragmentation spectrum (MS2) can be unambiguously assigned to its precursor ion scan (MS1), which is essential and required to accurately 'sequence' a peptide. In contrast, DIA requires the availability and usage of spectral libraries (DDA), although current efforts attempt to analyse DIA data without an underlying spectral library. Such spectral libraries are generated by DDA and contain the elution time and fragmentation pattern of all peptides. Secondly, the quantitative values are based on the MS1 features when using DDA, which is normally highly complex and noisy. Therefore, DDA quantification is typically more error-prone. In contrast, the quantitative values in DIA are based on all MS2 features of a peptide; since MS2 spectra are typically cleaner and without interference, this often results in better quantitative accuracy. Thirdly, DDA is biased towards highly abundant proteins, while low abundance proteins tend to be ignored. DIA in contrast can potentially quantify every peptide present in the sample as it does not suffer from any abundance biases. Lastly, DIA data can always be re-interrogated with a different spectral library and thus, can be regarded as a data repository.

1.3.7 Data Analysis

Different mass spectroemtric acquisition modes result in different data structures and thus, require different algorithms and software packages to extract meaningful information. Currently, there are several software packages available for analysing DDA-based data including MaxQuant [227], Mascot [228] and SEQUEST [229]. Those software packages use a "spectrum-centric" approach to perform protein quantification based on the MS1 peak intensities, while the MS2 spectra are only used for identification. A database containing the available protein sequences is digested *in silico* with the same protease that has been used in the sample preparation. A list of peptides is generated with theoretical m/z information, which is used to match the acquired MS1 spectrum. A peptide candidate list is first determined based on the acquired and theoretical m/z and then the MS2 spectrum is compared to the theoretical tandem mass spectrum (fragments) for reliably identification [230].

The spectral interpretation required for DIA can be quite challenging due to the highly complex MS2 spectra. In contrast to DDA, analysis of DIA data is not yet fully implemented in the general community, but the best approach is based on a peptide-centric analysis [230]. The workflow starts with the generation of a library, which contains all spectral information of each peptide such as m/z, fragment ion intensity pattern and retention time. Instead of searching each individual spectrum against the database, each peptide in the library is queried against the acquired DIA data to find the best matching peak within the data and the intensity from its MS2 spectrum is used for quantification [231]. There are currently only a handful of software packages available for DIA analysis. These include Skyline [232], Spectronaut (Biognosys) and OpenSWATH [233]. Spectronaut has been used in this thesis and a DIA workflow has been applied for the studies described in Chapters 4 and 5.

1.4 Thesis Overview

This thesis describes the application of state-of-the-art proteomics and phosphoproteomics approaches to better understand the recruitment of monocytes and subsequent polarisation of macrophages that are central to the pathogenesis of atherosclerosis.

In order to understand the physiological and pathological roles of macrophages and to monitor the disease progression and the effects of trial therapeutics, it is important to identify the phenotypes of macrophages present in specific inflamed tissues. In the study described in Chapter 3, we exploited recent advances in label-free proteomics to systematically compare the global proteomes of M1-polarised, M2-polarised and non-polarised THP-1 macrophages. We identified numerous proteins that are differentially regulated between the M1 and M2 phenotypes and validated these results by targeted proteomics and Western blotting analysis. Of note, a subset of these proteins can be used to identify the various polarisation phenotypes and act therefore as biomarkers for M1 and M2 macrophages. In addition, this study will also serve as a valuable benchmark for future studies of inflammatory diseases and treatments.

As discussed above, CCR2 is the major receptor involved in recruitment of monocytes, contributing to the pathogenesis of atherosclerosis and well as some other inflammatory diseases. MCPs are the main chemokines that activate CCR2 and induce recruitment of monocytes from the blood into arterial walls, where they differentiate into macrophages and contribute to development of atherosclerotic plaques [234]. Because a number of clinical trials targeting MCPs or CCR2 have been unsuccessful, there is a need to discover alternative targets such as downstream signalling effectors activated by MCP/CCR2 and other chemokine/receptor interactions. The studies described in Chapters 4 and 5 were undertaken to provide a more complete and unbiased understanding of MCP/CCR2 signalling.

The study in Chapter 4 was based on the observation that, relative to MCP-1, MCP-3 is a partial agonist of CCR2 as assessed by targeted signalling assays in our laboratory [235]. We hypothesised that MCP-1 and MCP-3 trigger distinct proteomic and phosphoproteomic changes in the downstream signalling transduction pathways that could potentially extend and explain the observed differences. Using DIA-MS, a cutting-edge mass spectrometric technology, we identified a number of differences in protein

phosphorylation induced by these two chemokines. Moreover, we were able to use the phosphoproteomic data to create a comprehensive map of the signal transduction network activated by MCP-1 and MCP-3 treatment, thus extending our knowledge of MCP/CCR2 signalling.

To further validate and extend the signal transduction network characterised in Chapter 4, we undertook the time course study described in Chapter 5, in which we determined the kinetics of phosphorylation at thousands of sites that had been observed in the previous chapter to be modulated upon MCP-1 treatment. This approach allowed us to both expand our network map and to predict, identify and confirm temporal linkages within the various signalling pathways.

In summary, this thesis describes the successful application of state-of-the-art proteomics and phosphoproteomics approaches to gain novel insights into the intricate mechanisms that lead to macrophage polarisation and to unravel signal transduction networks that are involved in monocyte recruitment. These studies will contribute to ongoing efforts to understanding the biological mechanisms leading to the development of atherosclerosis and will help guide future research towards effective atherosclerosis therapies.

Chapter 2. Materials and Methods

2.1 Materials

Roswell Park Memorial Institute (RPMI) medium and Dulbecco's Modified Eagle Medium (DMEM) were from Invitrogen. Blasticidin and hygromycin B were from Thermo Fisher Scientific. Tetracycline free foetal bovine serum (FBS) was from Clontech Laboratories. HisTrap HP nickel affinity column and HiLoad 16/60 Superdex 75 preparative grade size exclusion column (PSEC) were from GE Healthcare. Unless otherwise noted, all other chemicals/reagents were purchased from Sigma-Aldrich. All the buffers were prepared in MilliQ water, unless otherwise noted.

2.2 Buffers and Solutions

2.2.1 Cell Culture

2.2.1.1 RPMI1640 Cell Growth Medium

RPMI	500 mL
Heat-deactivated FBS	10% (v/v)

2.2.1.2 DMEM Cell Growth Medium

DMEM	500 mL
Hygromycin B	200 µg/mL
Blasticidin	5 µg/mL
FBS	10% (v/v)

2.2.1.3 Lysogeny Broth (LB) Medium

NaCl	10 g/L
Yeast extract	5 g/L
Tryptone	10 g/L

2.2.2 Protein Production

2.2.2.1 Lysis Buffer

Tris	20 mM
NaCl	500 mM
Imidazole	5 mM
Adjusted to pH 8.0	

2.2.2.2 Inclusion Body Washing Buffer

Tris	20 mM
NaCl	500 mM
Imidazole	5 mM

Triton X-100	0.5% (v/v)
Dithiothreitol	2 mM
Adjusted to pH 8.0	

2.2.2.3 IMAC Resin Equilibration Buffer

Tris	20 mM
Guanidinium chloride	6 M
Imidazole	20 mM
2-Mercaptoethanol	20 mM
Adjusted to pH 8.0	

2.2.2.4 IMAC Resin Elution Buffer

Tris	20 mM
Guanidinium chloride	6 M
Imidazole	200 mM
2-Mercaptoethanol	20 mM
Adjusted to pH 8.0	

2.2.2.5 Refolding Buffer

Tris	20 mM
NaCl	400 mM
L-Glutathione reduced	2 mM
L-Glutathione oxidized	0.5 mM
Adjusted to pH 8.0	

2.2.2.6 HisTrap HP, Buffer A

Tris	20 mM
Imidazole	20 mM
NaCl	500 mM
Adjusted to pH 8.0	

2.2.2.7 HisTrap HP, Buffer B

Tris	20 mM
Imidazole	200 mM
NaCl	500 mM
Adjusted to pH 8.0	

2.2.2.8 Thrombin Cleavage Buffer

Tris	20 mM

NaCl	400 mM
$CaCl_2$	2.5 mM
Adjusted to pH 8.0	

2.2.2.9 Preparative Size Exclusion Chromatography Buffer

HEPES	10 mM
NaCl	150 mM
Adjusted to pH 8.0	

2.2.3 SDS-PAGE and Western Blotting

2.2.3.1 4X SDS-PAGE Loading Buffer

Glycerol	40% (v/v)
Tris-HCl (pH 6.8)	200 mM
SDS	8% (w/v)
Bromophenol blue	0.4% (w/v)
Dithiothreitol	100 mM

2.2.3.2 SDS-PAGE Running Buffer

Glycine	192 mM
SDS	0.1% (w/v)
Tris-HCl (pH 8.8)	24.8 mM
Adjusted to pH 8.0	

2.2.3.3 Separating Gel

Bis-acrylamide	12% (w/v)
Tris-HCl (pH 8.8)	375 mM
SDS	0.1% (w/v)
Ammonium persulfate	0.05% w/v)
TEMED	0.1% (v/v)
Adjusted to pH 8.8	

2.2.3.4 Stacking Gel

Bis-acrylamide	5% (w/v)
Tris-HCl (pH 6.8)	125 mM
SDS	0.1% (w/v)
Ammonium persulfate	0.05 (w/v)
TEMED	0.1% (v/v)
Adjusted to pH 6.8	

2.2.3.5 Western Blotting Transfer Buffer

Glycine	192 mM
Tris	25 mM
Methanol	20% (v/v)
Adjusted to pH 8.3	

2.2.3.6 Tris-Buffered Saline and Tween 20 (TBST)

Tris	25 mM
NaCl	150 mM
Tween-20	0.1% (v/v)
Adjusted to pH 8.5	

2.2.3.7 Blocking Solution

Bovine serum albumin in TBST	5% (w/v)

2.2.3.8 Phosphate Buffered Saline (PBS) Stock Solution

NaCl	150 mM
KCl	2.7 mM
Na_2HPO_4	1.4 mM
KH_2PO_4	4.3 mM
Adjusted to pH 7.3	

2.2.3.9 PBS-EDTA (Versene)

NaCl	150 mM
KCl	2.7 mM
Na_2HPO_4	1.4 mM
KH_2PO_4	4.3 mM
EDTA	0.53 mM
Adjusted to pH 7.3	

2.2.4 Affinity-Purification Mass Spectrometry (AP-MS) Buffer

2.2.4.1 AP-MS Cell Lysis Buffer

Tris	50 mM
EDTA	1 mM
$MgCl_2$	1 mM
NaCl	150 mM
NP-40	1% (v/v)
Sodium deoxycholate	0.5% (w/v)

Adjusted to pH 7.4

2.2.4.2 AP-MS Washing Buffer

Tris	50 mM
EDTA	1 mM
$MgCl_2$	1 mM
NaCl	150 mM

Adjusted to pH 7.4

2.2.5 Mass Spectrometry Sample Preparation Buffer

2.2.5.1 Cell Lysis Buffer

HEPES	100 mM
Sodium deoxycholate (SDC)	1% (w/v)
PhosSTOP (phosphatase inhibitors) (Roche)	1 tablet per 10 mL

Adjusted to pH 8.0

2.2.5.2 Water-saturated Ethyl Acetate

Ethyl acetate	80 mL
MilliQ water	20 mL

2.2.5.3 TiO_2 Binding Buffer

Lactic acid	200 mg/mL
Acetonitrile	50% (v/v)
Trifluoroacetic acid (TFA)	5% (v/v)

2.2.5.4 TiO_2 Washing Buffer

Acetonitrile	50% (v/v)
TFA	5% (v/v)

2.2.5.5 TiO_2 Elution Buffer

Ammonium Hydroxide	5% (v/v)
Acetonitrile	50% (v/v)

2.2.5.6 Reverse Phase C18 High pH Fractionation Buffer A

Ammonium hydroxide	10 mM

2.2.5.7 Reverse Phase C18 High pH Fractionation Buffer B

Ammonium hydroxide	10 mM
Acetonitrile	80% (v/v)

2.2.5.8 Mass Spectrometry Sample Resuspension Buffer

Acetonitrile	2% (v/v)

Formic acid	0.1% (v/v)
Indexed retention time (iRT) peptide (Biognosys)	10 pmole

2.2.5.9 StageTip (C18) Desalting Buffer A

Formic acid	0.1% (v/v)

2.2.5.10 StageTip (C18) Desalting Buffer B

Acetonitrile	60% (v/v)
Formic acid	0.1% (v/v)

2.2.5.11 LC-MS/MS Buffer A

Formic acid	0.1% (v/v)

2.2.5.12 LC-MS/MS Buffer B

Acetonitrile	80% (v/v)
Formic acid	0.1% (v/v)

2.2.6 Cell Imaging Buffer

2.2.6.1 Fixing Buffer

Paraformaldehyde in PBS	4% (v/v)

2.2.6.2 Imaging Blocking Buffer

Bovine serum albumin in PBS	2.5% (w/v)
Triton X-100	0.05% (v/v)

2.3 General Methods

2.3.1 Cell Culture

2.3.1.1 HEK293 Cells

The FlpIn-CCR2-HEK293 cell line was generated in Dr. Meritxell Canals' laboratory at the Monash Institute of Pharmaceutical Science (MIPS) as described. Briefly, the c-Myc-His-CCR2 construct (in pcDNA/FRT/TO) was transfected into human embryonic kidney (HEK) FlpIn TRex 293 cells using Lipofectamine (Invitrogen). Successfully transfected cells were selected based on their hygromycin B resistance. The stably transfected FlpIn-c-Myc-His-CCR2-HEK293 cell line (called FlpIn-CCR2-HEK293 for the rest of the thesis) was maintained in DMEM Cell Growth medium.

2.3.1.2 THP-1 Cells

THP-1 monocytes were maintained in RPMI 1640 Cell Growth Medium. Both cell cultures were passaged every 3-4 days and fresh cells were thawed every 3-4 months

from liquid nitrogen to substitute the old batch. All cell culture work was performed under sterilized conditions in a biosafety cabinet.

2.3.1.3 Storage of Cells in liquid Nitrogen

For long-term storage of cell lines, the cells were harvested during the exponential growth phase, centrifuged at 800*g* for 5 min at room temperature and reconstituted in FBS to a final concentration of 10^7 cells/mL. An equal volume of FBS containing 20% (v/v) DMSO was added resulting in a final concentration of 5 x 10^6 cells/mL. Aliquots of 1 mL cells were stored in cryovials in a -80 °C freezer for 14 h and then transferred into a liquid nitrogen container for long-term storage.

2.3.1.4 Reconstitution of Cells After Storage in Liquid Nitrogen

Cryovials containing frozen cells from liquid nitrogen were warmed up quickly in a 37 °C water bath and resuspended in 5 mL of pre-warmed (37 °C) DMEM or RPMI 1640 Cell Growth Medium. Cells were spun down at 800*g*, 5 min, room temperature and the supernatant was discarded to remove DMSO. The cell pellet was resuspended in DMEM cell growth medium, transferred to a T75 or T175 flask and incubated at 37 °C, 5% CO_2.

2.3.1.5 Determination of Cell Density

In order to determine cell density, 10 µL of appropriately diluted cell mixture was loaded into the chamber of a cell counting hemocytometer (Neubauer Chamber, Celeromics) and the cells were counted under a TMSmicroscope (Nikon). The following equation was used to calculate the cell concentration:

$$\text{Cell concentration (cells / mL)} = \frac{\text{Number of cells} * 10{,}000}{4 * \text{dilution factor}}$$

The dilution factor was empirically determined. Usually, a 50 times dilution was required for a T75 flask, while a 100 times dilution was necessary for a T175 flask.

2.3.2 Recombinant Chemokine Production

All chemokines were expressed in BL21 (DE3) *E.coli* cells. Briefly, the *E.coli* cells containing pET11a-MCP-1, MCP-2 or MCP-3 plasmids were freshly spread on an LB plate containing 50 µg/mL ampicillin. Ten individual colonies were picked to test the

chemokine expression levels in 50 mL cultures and the highest expressing bacterial colony was chosen for a 4 L scale expression. The *E. coli* cells were inoculated at a ratio of 1:100 and incubated at 37 °C for 2-3 h until they reached an OD of 0.6-0.8. After induction with 1 mM IPTG, cells were grown overnight at 37 °C then harvested via centrifugation (Sorvall Evolution RC-SLC 6000 rotor, 5,000 rpm, 15 minutes, 4 °C) and resuspended in 50 mL Lysis Buffer. After addition of chicken egg white lysozyme (10 mg), the resuspended cell pellets were incubated at RT for 30 min and then sonicated 5 times with an amplitude of 10 for 30 s (with 1 min incubation on ice between the cycles) (on ice). The inclusion bodies were obtained by centrifugation (Sorvall Evolution SS-34 rotor, 15,000 rpm, 15 minutes, 4 °C) and resuspended in 40 mL of Lysis Buffer, containing 0.1 mg of DNase I from bovine pancreas. After incubation at RT for 30 min on a shaker, the inclusion bodies were centrifuged and resuspended in Inclusion Body Washing Buffer. The inclusion bodies were again centrifuged and washed with Lysis Buffer to remove any residual Triton X-100. The inclusion bodies were dissolved in 20 mL of IMAC Resin Equilibration Buffer by vigorously mixing and were stored at 4 °C overnight. After centrifugation, the supernatant was retained for further purification. Five millilitres of resuspended IMAC resin was washed with IMAC Resin Equilibration Buffer, added into the supernatant and the mixture was incubated at RT for 30 min. The IMAC resin was washed with IMAC Resin Equilibration Buffer and eluted with 10 mL of IMAC Resin Elution Buffer. The eluate was added dropwise to 2 litres of Refolding Buffer while rapidly stirring at RT. The refolded protein was filtered and loaded onto a 5 mL HisTrap HP column (GE Healthcare Life Sciences) pre-equilibrated with HisTrap HP Buffer A and eluted using a stepwise isocratic gradient of HisTrap HP Buffer B. The elution solution from the HisTrap HP column was dialysed overnight against thrombin cleavage buffer and cleaved with thrombin at a ratio of 10 units of thrombin per mg protein at 37 °C for 16 h. The reaction was quenched by adding a final of 200 µM of phenylmethylsulfonyl fluoride (PMSF). The cleaved protein was passed through a 5 mL HisTrap HP column to remove the His tag and any uncleaved protein. The flow-through, which contained the cleaved protein, was concentrated down and further purified by HiLoad Superdex 75 preparative size exclusion chromatography (GE Healthcare Life Sciences). The purified protein was analysed by SDS-PAGE to ensure a purity of at least ~95% and MALDI-TOF was used to verify its molecular weight.

2.3.3 Cell Treatments

2.3.3.1 THP-1 Cells

The suspended THP-1 monocytic cells were spun down at 1000*g* for 10 min and the cell pellets were resuspended in 10 mL of RPMI 1640 Cell Growth Medium. One million cells were seeded per well of a 6-well plate. The cells in each well were treated with phorbol 12,13-dibutyrate (PDBu, 100 nM) for 24 h at 37 °C, 5% CO_2 to differentiate them into macrophages. The macrophages were subsequently treated for 48 h with: IFN-γ (5 ng/ml) and LPS (10 ng/ml) for M1 polarisation; IL-4 (25 ng/ml) for M2 polarisation; or left untreated (M0). The cells were finally washed with ice-cold PBS, harvested by centrifugation and stored at -80 °C.

2.3.3.2 FlpIn-CCR2-HEK293 Cells

Around 10 million cells were plated in 15 cm dishes (BD Biosciences) and incubated at 37 °C, 5% CO_2 for 24 h. The DMEM Cell Growth Medium was replaced with FBS-free DMEM medium containing 10 µg/mL tetracycline to induce the expression of CCR2 for 16 h. The cells were treated with chemokines (the concentration and incubation time varied and is specified in the main text), washed with ice-cold PBS 3 times quickly and lysed with 1mL of Cell Lysis Buffer. The lysate was harvested with a cell scraper, the DNA was sheared 3 times with a probe sonicator (Soniprep 150, MSE) at amplitude 10 for 30 s and the lysate was stored at -80 °C.

2.3.4 Mass Spectrometry Sample Preparation

2.3.4.1 Affinity Purification

Two to three million cells were seeded in 100 mm dishes (BD Biosciences) and incubated for 48 h. After 24 h, tetracycline (final concentration 10 µg/mL) was added to half of the dishes to induce expression of CCR2, while the other half were used as controls. After 48 h, the cells were washed three times with ice-cold PBS, detached by scraping the dish with a cell scraper and centrifuged at 1,200*g* for 5 min. The cell pellet was reconstituted in 1 mL ice-cold AP-MS Lysis Buffer and incubated on ice for 10 min. To aid protein solubilization, the lysate was then sonicated twice at amplitude 2 (MSE Soni prep 150 plus) for 30 s with a 1 min break on ice in between and incubated at 4 °C on a spinning rotator for 30 min before it was passed 10-times through a 21-gauge needle.

The cell debris was spun down at 500*g* for 3 min at 4 °C and 300 µL of the supernatant was added to 30 µL of PBS-washed anti-cMyc agarose affinity gel (A7470, Sigma Aldrich). The sample mixture was incubated on a spinning rotator for 4 h in a cold room and spun down at 500*g* for 3 min at 4 °C. The agarose gel, which should contain the c-Myc tagged CCR2 protein, was washed 3-times with AP-MS Washing Buffer and then eluted with 50 µL of 4X SDS-PAGE Loading Buffer at 37 °C for 30 min. The sample was loaded onto a NuPAGE 4-12% SDS-PAGE gel (Invitrogen), which ran for only 3 min. The gel was stained with Instant blue (Expedeon) for 1 h and only one gel band, which contained all proteins, was excised from the gel [236]. The band was cut into approximately 2 mm cubes, which were destained using 100mM Ammonium bicarbonate (ABC) followed by 50mM ammonium bicarbonate / 50% acetonitrile. The gel cubes were subsequently reduced with 5 mM dithiothreitol (DTT) at 37 °C for 20 min and alkylated with 20 mM chloroacetamide (CAA) at room temperature in the dark for 20 min. The reduced and alkylated gel cubes were washed with 50 mM ABC / 50% acetonitrile and dehydrated completely with 100% acetonitrile. The dehydrated gel cubes were rehydrated by the addition of 12.5 ng/µL trypsin digestion solution in 100 mM ammonium bicarbonate / 5% acetonitrile and incubated at 37 °C overnight. The digestion reaction was quenched by adding 50% acetonitrile / 5% formic acid and the solution, which contained the digested peptides, was extracted with 100% acetonitrile, dried in a Speed Vac and resuspended in 20 µL of Mass Spectrometry Sample Resuspension Buffer.

2.3.4.2 Mass Spectrometry Sample Lysis

The frozen cells were lysed in 1 mL of Cell Lysis Buffer, heated at 95 °C for 5 min and sonicated at amplitude 10 for 30 s (Soniprep 150, MSE). Please note, we used sodium deoxycholate as the cell lysis detergent in this thesis and we assessed the efficiency of this detergent for protein extraction in Appendix I.

2.3.4.3 Determination of Protein Concentration Using a Bicinchoninic Acid (BCA) Assay

The protein concentration was determined with a Bicinchoninic acid (BCA) assay (Pierce™ BCA Protein Assay Kit, Thermo Fisher Scientific) according to the user manual. In brief, 5 µL of either pre-diluted BSA standards (0.125 µg/mL, 0.25 µg/mL, 0.5 µg/mL, 0.75 µg/mL, 1 µg/mL and 2 µg/mL; Pierce™ Bovine Serum Albumin Standard Pre-Diluted Set, Thermo Fisher Scientific), blanks or cell lysates (diluted if

necessary) were added in triplicate into a 96-well plate. Two hundred µL of working reagent (50 parts of Reagent A and 1 part of Reagent B as indicated in the manual) was added into each well. After incubating the plate at 37 °C for 30 min, the colorimetric absorbance was measured with a plate reader (FluoSTAR Optima, BMG Labtech) at 492 nm and the protein concentrations of the lysates were determined based on the BSA linear standard curve.

2.3.4.4 Protein Digestion

The required amount of protein was denatured and alkylated with 10 mM tris(2-carboxyethyl)phosphine and 40 mM CAA at 95 °C for 5 min. The pH of the lysate was adjusted to 7-8 and trypsin (Trypsin Gold, Promega) was added at a ratio of 1:100 (trypsin: protein (w/w)) and incubated at 37 °C while shaking on a Thermomixer (Eppendorf) for 16 h.

2.3.4.5 Sodium Deoxycholate Removal

The digested cell lysate was acidified with 100% formic acid to a final concentration of 2% and an equal volume of Water-Saturated Ethyl acetate was added. The solution was mixed by vortexing for 1 min and centrifuged at 13,000*g* for 5 min (5415D, Eppendorf) resulting in 2 distinct phases. The aqueous bottom layer, which contained the peptides, was separated from the organic top layer, which contained the SDC, and transferred into a new microcentrifuge tube. This ethyl acetate mediated removal step of SDC was repeated and the volume of the final aqueous layer was reduced in a SpeedVac concentrator (Centrivap, Labconco) to less than 500 µL for HPLC fractionation or for C18 desalting.

2.3.4.6 Reverse Phase (C18) High pH Fractionation

Reverse Phase C18 High pH Fractionation Buffers A and B, which are used for fractionation, are listed in section 2.2.5.6 and 2.2.5.7. The peptides were fractionated with a ZORBAX 300 Extend-C18 4.6 x 250 mm 5 µm (or 4.6 x 50 mm 3.5 µm) column (Agilent Technologies) connected to a HPLC 1100 (Agilent Technologies). The gradient was run as follows: 0% Buffer B-2.5% Buffer B in 1 min, 2.5% Buffer B in 6 min, 2.5% Buffer B - 50% Buffer B in 72 min, and then ramped to 100% B in 0.1 min and held for 5 min to fully clean the column. Then, the column was equilibrated with Buffer A for 10 min. Sixty-six (66) peptide-containing fractions were collected between 7 and 72 min

and pooled in a non-contiguous manner as previously described to obtain a total of 6 fractions [237]. We performed a preliminary comparison between reverse phase (C18) and strong cation exchange (SCX) fractionation, which is described in Appendix I.

2.3.4.7 Desalting using Oasis HLB Cartridges

The HLB cartridge (Waters) was stacked on top of a vacuum extraction manifold (Waters) and the pressure of the manifold was increased to a level of 5'' Hg under vacuum. The column was then washed and equilibrated with 6 mL of methanol and StageTip (C18) Desalting Buffer A. The sample was load onto the cartridge (6 mL at a time) and washed 3 times with StageTip (C18) Desalting Buffer A. The peptides were eluted with 6mL of StageTip (C18) Desalting Buffer B and the eluate was dried in a benchtop lyophilizer (FreeZone Freeze dry, Labconco).

2.3.4.8 Phosphopeptide Enrichment

The dried peptides were resuspended in TiO_2 Binding Buffer (300 µL). In parallel, TiO_2 beads (GL Sciences) were weighed out at a ratio of 4:1 (beads: peptides w/w) and washed with TiO_2 Binding Buffer (1 mL) through incubation at room temperature on a rotating vortexer for 10 min. The beads were pelleted by centrifugation at 400*g* for 3 min and the supernatant was discarded. The washed beads were resuspended in a sample-dependent volume of TiO_2 Binding Buffer (typically, 600-1000 µL). The peptides were added to the beads and the mixture was incubated on the vortexer (Vortex Genie 2, Scientific Industries) for 30 min at maximum speed at room temperature. During this incubation, C8 (Empore SPE extraction disk, 3M) StageTips were prepared by punching layers of a C8 membrane with an 18-gauge blunt syringe needle and placing them in 200 µL pipette tip. These C8 StageTips were equilibrated with 40 µL methanol, TiO_2 Washing Buffer and TiO_2 Binding Buffer by centrifugation at 3,000*g* for 2 min at room temperature. The beads containing the phosphopeptides were loaded into the C8 StageTips and washed with TiO_2 Binding buffer and TiO_2 Washing Buffer by centrifugation at 3000 g for 2 min. Each washing step was repeated 3 times. The C8 StageTips (containing the beads) were transferred to new microcentrifuge tubes and eluted twice with 200 µL of TiO_2 Elution Buffer by centrifugation at 1,000g for 10 min. The combined elution sample was acidified with 400 µL of 20% TFA, dried down in a SpeedVac concentrator (Centrivap, Labconco) and desalted with C18 StageTips using the protocol outlined below. The desalted phosphopeptides were dried in a SpeedVac

concentrator and reconstituted in Mass Spectrometry Sample Resuspension Buffer. We performed a preliminary study to test the feasibility of SDC-based phosphopeptide label free quantification, which is described in Appendix I.

2.3.4.9 Desalting Using StageTips (C18)

In order to desalt and clean the peptide mixture prior to mass spectrometry, all peptide solutions were desalted with homemade StageTips (C18) [238]. A C18 StageTip was prepared by punching 1-4 layers of a C18 membrane (Empore SPE extraction disk, 3M) with an 18-gauge blunt syringe needle and stacking them in a 200 µL pipette tip.

The pH of the peptide mixture was lowered to 2-3 by addition of 10% formic acid. The C18 StageTip was activated with 40 µL of methanol and equilibrated with StageTip (C18) Desalting Buffer A by centrifugation at 500*g* for 1-5 min. After the StageTip was fully equilibrated, the peptide mixture was loaded and the StageTip was centrifuged at 500*g* for 1-5 min. The StageTip (containing bound peptides) was washed with 40 µL of StageTip (C18) Buffer A, transferred to a new microcentrifuge tube and eluted with 40 µL of StageTip (C18) Buffer B. The desalted peptide (or phosphopeptide) eluate was dried in a SpeedVac concentrator and reconstituted in Mass Spectrometry Sample Resuspension Buffer.

2.3.5 Mass spectrometry Data Acquisition

2.3.5.1 DDA Analyses

All samples/peptides were analysed by LC-MS/MS using a Q Exactive Plus mass spectrometer (ThermoFisher Scientific) coupled online to a RSLC nano HPLC (Ultimate 3000, UHPLC ThermoFisher Scientific). Samples were loaded onto a 100 µm, 2 cm nanoviper Pepmap100 trap column, eluted and separated on a RSLC nano column 75 µm x 50 cm, Pepmap100 C18 analytical column (ThermoFisher Scientific). The detailed gradients used for protein identification and phosphopeptide identification are listed in Table 2.1 and Table 2.1, respectively.

The eluent was nebulised and ionised using a nano electrospray source (ThermoFisher Scientific) with a distal coated fused silica emitter (New Objective). The capillary voltage was set at 1.7 kV. The Q Exactive mass spectrometer was operated in the data dependent acquisition mode to automatically switch between full MS scans and

subsequent MS/MS acquisitions. Survey full scan MS spectra (m/z 375–1800) were acquired in the Orbitrap with 70,000 resolution (at m/z 200) after accumulation of ions to a 3 x 10^6 target value with a maximum injection time of 30 ms. Dynamic exclusion was set to 20 s. The 10 most intense multiply charged ions ($z \geq 2$) were sequentially isolated and fragmented in the collision cell by higher-energy collisional dissociation (HCD) with a fixed injection time of 60 ms, 17,500 resolution and automatic gain control (AGC) target of 5 x 10^4.

2.3.5.2 DIA Analyses

The LC methods were the same as for the DDA analyses. The following settings were applied to the mass spectrometer operated in data-independent mode: MS1 scan range from 370 to 2,000 m/z with a resolution of 70,000 (at m/z 200) using an AGC target of 1 x 10^6 and a maximum ion injection time of 50 ms; MS2 scan with a resolution of 17,500 (at m/z 200) using an AGC target of 2 x 10^5 with automatic injection time and a loop count of 25. In order to minimize the number of co-fragmented peptides and phosphopeptides and to increase the quantitative accuracy, we injected each sample 2 or 3 times using either 2 (376-675 m/z and 675-975 m/z) or 3 (376-675 m/z, 675-975 m/z and 975-1275 m/z) distinct m/z windows. The inclusion lists can be found in Table 2.3.

Table 2.1 The LC gradient used for proteome identification and quantification

Retention time (min)	Flow (μL/min)	%B
0	0.25	2.5
5	0.25	5
6	0.25	12.5
114	0.25	32.5
120	0.25	42.5
125	0.25	99
132	0.25	99
133	0.25	2.5
155	0.25	2.5

Table 2.2 The LC gradient used for phosphoproteome identification and quantification

Retention time (min)	Flow (μL/min)	%B
0	0.25	2.5
5	0.25	5
6	0.25	12.5
84	0.25	32.5
90	0.25	42.5
95	0.25	99
102	0.25	99
103	0.25	2.5
125	0.25	2.5

Table 2.3 DIA m/z inclusion list

Mass (375-675 m/z)	Mass (675-975 m/z)	Mass (975-1275 m/z)
381.4233	681.5597	681.5597
393.4287	693.5651	693.5651
405.4342	705.5706	705.5706
417.4396	717.5761	717.5761
429.4451	729.5815	729.5815
441.4505	741.587	741.587
453.456	753.5924	753.5924
465.4615	765.5979	765.5979
477.4669	777.6033	777.6033
489.4724	789.6088	789.6088
501.4778	801.6143	801.6143
513.4833	813.6197	813.6197
525.4887	825.6252	825.6252
537.4942	837.6306	837.6306
549.4997	849.6361	849.6361
561.5051	861.6415	861.6415
573.5106	873.647	873.647
585.516	885.6525	885.6525
597.5215	897.6579	897.6579
609.5269	909.6634	909.6634
621.5324	921.6688	921.6688
633.5379	933.6743	933.6743
645.5433	945.6797	945.6797
657.5488	957.6852	957.6852
669.5542	969.6907	969.6907

2.3.6 Mass Spectrometric Data Analysis

2.3.6.1 DDA Data Analyses

All generated files were analysed with MaxQuant (version 1.5.2.8) and its implemented Andromeda search engine [239] to obtain protein identifications as well as their label-free quantitation (LFQ) intensities. Database searching was performed with the following parameters: cysteine carbamidomethylation as a fixed modification; methionine oxidation N-terminal acetylation and phosphorylation (serine, threonine and tyrosine) as variable modifications; up to 2 missed cleavages permitted; mass tolerance of 20 ppm; 1% protein false discovery rate (FDR) for protein and peptide identification; and minimum 2 peptides for pair-wise comparison in each protein for label-free quantitation. The human protein sequence database was downloaded from Uniprot/SwissProt, containing 20,244 proteins.

The MaxQuant result output was further processed with Perseus (Version 1.5.0.31), a module from the MaxQuant suite. After removing reversed and known contaminant proteins, the LFQ values were $\log_2$ transformed and the reproducibility across the biological replicates was evaluated by a Pearson's correlation analysis. The replicates were grouped accordingly and all proteins were removed that had only one "valid value" in each group. The missing values were replaced by imputation and multiple-samples ANOVA tests (FDR < 5%) were performed to obtain a list of significantly regulated proteins, which were combined and normalized according to their Z-scores. Hierarchical clustering was used to assess the similarity between each replicate and condition using the Euclidean for distance calculation and averaged for linkage in row and column trees with a maximum of 300 clusters [240]. TM4:MeV (http://www.tm4.org/mev.html) [241] was used to group the proteins into different clusters based on k-Means/Medians clustering with a predefined number based on figure of merits analyses. The proteins from each cluster were submitted into FunRich 3.0 [242] for biological pathway and transcription factor enrichment analysis.

2.3.6.2 DIA Data Analyses

Spectral libraries were generated in Spectronaut (Biognosys) based on MaxQuant results using default parameters, which includes that the MaxQuant score type was set to 1% (PSM FDR) and protein inference was performed. Spectral libraries for

phosphopeptides and global peptides were generated identically, but separately. The multiple acquired DIA files for each sample were merged into the htrms format using the HTRMS Converter (Biognosys) and then loaded into Spectronaut (Biognosys) for generation of normalized phosphopeptide or protein intensity. All the parameters for phosphopeptide or protein quantification were kept as default. In addition, for global proteome quantification only non-modified peptides and the sum of peptide or precursor peak area was used for quantification. In contrast, only modified peptides were used for phosphopeptide quantification.

For DIA protein quantifications, a report containing normalized peak areas was exported as an Excel spreadsheet. Proteins with a q-value less than 0.01 and which were quantified in at least 2 out of 3 biological repeats in at least one condition were used for further analysis. The reproducibility was evaluated by Pearson's correlation analyses, which were visualized by heat maps generated with ComplexHeatmap [243] in R. The missing values were replaced based on the Gaussian distribution from global quantified proteins in Perseus 1.5.5.3 [244]. To identify proteins that were significantly regulated, two-sample student's t tests were used and cutoffs were applied as follows: p-value ≤ 0.01; fold change ≥1.5 or ≤ -1.5. These significantly regulated proteins were exported to an Excel spreadsheet and volcano plots were generated with ggplot2 [245] in R.

For DIA phosphopeptide quantifications, a report containing normalized peak areas for modified peptides was exported to an Excel spreadsheet. Only peptides falling below a q-value of 0.01 were retained for further processing and unphosphorylated peptides were filtered out. All phosphorylated peptides including their quantitative values were imported into Perseus with a unique identifier for each phosphorylation event and only phosphopeptides quantified in at least 2 out 3 biological repeats in at least one condition were used for analysis. The reproducibility of DIA phosphopeptide quantification was evaluated by Pearson's correlation analyses, which were visualized by heat maps generated with ComplexHeatmap [243] in R. The missing values were replaced based on the Gaussian distribution from phosphoproteome quantification in Perseus 1.5.5.3 [244]. To identify phosphopeptides that were significantly regulated, two-sample student's t tests were used and cutoffs were applied as follows: p-value ≤ 0.01; fold change ≥1.5 or ≤ -1.5. These significantly regulated phosphopeptides were exported to an Excel spreadsheet and volcano plots were generated with ggplot2 [245] in R.

The significantly changed phosphopeptides, which were derived from proteins that changed significantly we excluded for further analyses. All phosphosites and phosphorylated residues were extracted from a human database with a Python script (Appendix V). The localization probability of each significantly regulated phosphopeptide was retrieved from MaxQuant output and only phosphopeptides with a localization probability > 0.75 were kept and used for clustering analysis in MeV [241]. Z score transformed $\log_2$ intensities were grouped based on k-means clustering. The number of clusters was determined by figure of merits analyses and empirically assigned to the most distinctive clusters [246]. The phosphopeptides from each cluster were submitted to Motif-X [247] to extract overexpressed motifs based on significance at 0.000001 and fold change of at least 3, while the IPI human proteome was used as the background gene list. Each of the cluster was further analysed in NetworKin 3.0 [248] for prediction of upstream kinases. A minimum NetworKin score of 1, max score difference of 4 and 10 predictions were applied as the filter. To reveal enriched kinase activity in each cluster, the enrichment was performed based on the number of predicted sites by a specific kinase to the total number of sites in each cluster [246]. Heat maps comparing motif and kinase activity were generated in R with the ComplexHeatmap package [243]. Categorical annotation was queried from ClueGO [249] in Cytoscape [250] with the form of Gene Ontology molecular function and based on p value < 0.05. In order to connect the regulated phosphorylation events to signaling cascades, each of the clusters was submitted to DAVID [251,252] for KEGG pathway mapping. The network was manually curated and mapped in Cytoscape [250].

2.3.7 SDS-PAGE

SDS-PAGE was carried out to separate proteins prior to Western blotting. Protein samples were quantified using the BCA kits as described in section 2.3.4.3 and 10-50 µg protein from each sample was diluted in SDS-PAGE Loading Buffer, boiled for 5 min and cooled to room temperature before loading onto the gel. SDS-PAGE was performed using a NuPage Novex Gel System with 4-12% Bis-Tris Protein Gels (Life Technologies). Electrophoresis was carried out at 200 V until the Bromophenol Blue running front reached the bottom of the gel (approximately 1 h).

2.3.8 Western Blotting

Proteins separated by SDS-PAGE were transferred onto PVDF membranes (GE Healthcare Life Sciences) using a Mini Trans-Blot Cell system (Bio-Rad Laboratories) filled with Transfer Buffer for 1 h at 100 volts. The membrane was washed three times with TBST Buffer for 5 min and blocked for 1 h at room temperature with Blocking Solution. The blocked membrane was incubated overnight at 4 °C with the primary antibody and next day with a horseradish peroxidase-conjugated secondary antibody for 1 h at room temperature. The PVDF membrane was developed with Western Lightning Plus ECL reagent (PerkinElmer) and documented with a ChemiDoc and ImageLab software (Bio-Rad Laboratories). The following antibodies were used: rabbit anti-SOD2 (AB10364, Merck Millipore); mouse anti-p47-phox (NCF-1) (SC-17845, Santa Cruz Biotech); mouse anti-NFκB p100/p52 (AB31474, Abcam); mouse anti-STAT1 (610115, BD Biosciences); rabbit anti-GADPH (2118, Cell Signaling Technology); rabbit anti-phospho-p44/42 MAPK (ERK1/2) (Thr202/Tyr204) (9101, Cell Signaling Technology); rabbit anti-p44/42 MAPK (ERK1/2) (9102, Cell Signaling Technology); rabbit anti-phospho-c-Jun (Ser63) (2361, Cell Signaling Technology); rabbit anti-c-Jun (Ser63) (9165, Cell Signaling Technology); rabbit anti-phospho-CK2 (Ser209) (AB12861, Abcam); rabbit anti-CK2 (AB10466, Abcam); goat anti-mouse IgG (M8642, Sigma Aldrich) and goat anti-rabbit IgG (AB6721, Abcam) as the secondary antibodies.

2.3.9 Confocal Microscopy

Eighty thousand cells were grown on a poly-D-lysine coated coverslip in each well of a 24 wells plate for 24 h using DMEM Cell Growth Medium. CCR2 receptor expression was induced by addition of tetracycline to a final concentration of 10 µg/mL and incubation for another 12 h, then the DMEM Cell Growth Medium was replaced by serum-free DMEM and incubation continued for another 12 h, giving a total cell growth time of 48 h. Serum starved cells were treated with 300 nM MCP-1 for 60 min, while non-treated cells were used as vehicle control. After the treatment, MCP-1 was removed by washing with PBS and cells were fixed immediately with 500 µL of Fixing Buffer for 15 min at room temperature. The fixed cells were permeabilised with 500 µL of 0.25% Triton X-100 for 5 min at room temperature and blocked with Imaging Blocking Buffer for 15 min. To make the staining solution, DAPI (1:2,000) and Alexa Fluor 647 Phalloidin (1:50) (Thermo Fisher Scientific) were added to fresh Blocking buffer. The

staining was performed by flipping the coverslip upside down onto 20 μL of staining solution and incubation at room temperature for 20 min. Excess staining solution was washed away three times with PBS and the coverslip was mounted in mounting oil by leaving at room temperature overnight. The images were taken with 63x oil objectives on a Leica SP8 inverted confocal microscope and processed in ImageJ.

Chapter 3. Proteomic Study of Macrophage Polarisation

3.1 Preface to Chapter 3

Macrophages have been implicated as the main cell type that contribute to the pathogenesis of atherosclerosis. At the site of inflammation, they can be polarised to two major phenotypes: pro-inflammatory (M1) and tissue reparative (M2) phenotypes. In order to understand the physiological and pathological roles of macrophages and to monitor the disease progression and the effects of trial therapeutics, it is important to identify the phenotypes of macrophages present in specific inflamed tissues.

Traditionally, M1 and M2 macrophages are distinguished based on the expression level of some cell surface markers and cytokines. Whilst such biomarkers are useful, current limitations include differences between the expression profiles of murine and human M1 and M2 macrophages [30,31] and differences between *in vitro* and *in vivo* macrophages [32,33]. As such, investigations would greatly benefit from an improved set of biomarkers for M1 and M2 macrophages.

This chapter comprises a manuscript that has recently been submitted for publication. As described in the declaration, I performed the proteomics study described in this manuscript, including: cell growth and treatments; sample processing; data collection; data analysis; and network mapping. Other aspects of the study were performed by the other authors.

Please note that all supplemental figures and tables belonging to this chapter can be found either in Appendix II or requested from the candidate (Cheng Huang) or supervisors (Martin J. Stone or Ralf B. Schittenhelm).

Journal of proteome research

Article

Cite This: J. Proteome Res. XXXX, XXX, XXX−XXX

pubs.acs.org/jpr

Proteomic Identification of Interferon-Induced Proteins with Tetratricopeptide Repeats as Markers of M1 Macrophage Polarization

Cheng Huang,[†,∥] Caitlin Lewis,[‡,∥] Natalie A. Borg,[†] Meritxell Canals,[⊥] Henry Diep,[‡] Grant R. Drummond,[¶] Robert J. Goode,[§] Ralf B. Schittenhelm,[§] Antony Vinh,[¶] Mingyu Zhu,[‡] Barbara Kemp-Harper,[*,‡] Oded Kleifeld,[*,#] and Martin J. Stone[*,†]

[†]Infection and Immunity Program, Monash Biomedicine Discovery Institute & Department of Biochemistry and Molecular Biology, [‡]Cardiovascular Disease Program, Biomedicine Discovery Institute & Department of Pharmacology, and [§]Monash Biomedical Proteomics Facility, Monash Biomedicine Discovery Institute & Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia

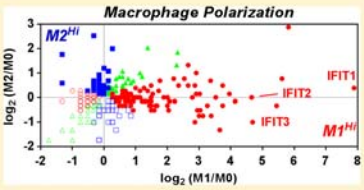
[⊥]Drug Discovery Biology, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Victoria 3052, Australia

[¶]Department of Physiology, Anatomy & Microbiology, School of Life Sciences, La Trobe University, Bundoora, Victoria 3086, Australia

[#]Faculty of Biology, Technion-Israel Institute of Technology, Technion City, Haifa 3200003, Israel

Supporting Information

ABSTRACT: Macrophages, which accumulate in tissues during inflammation, may be polarized toward pro-inflammatory (M1) or tissue reparative (M2) phenotypes. The balance between these phenotypes can have a substantial influence on the outcome of inflammatory diseases such as atherosclerosis. Improved biomarkers of M1 and M2 macrophages would be beneficial for research, diagnosis, and monitoring the effects of trial therapeutics in such diseases. To identify novel biomarkers, we have characterized the global proteomes of THP-1 macrophages polarized to M1 and M2 states in comparison with unpolarized (M0) macrophages. M1 polarization resulted in increased expression of numerous pro-inflammatory proteins including the products of 31 genes under the transcriptional control of interferon regulatory factor 1 (IRF-1). In contrast, M2 polarization identified proteins regulated by components of the transcription factor AP-1. Among the most highly upregulated proteins under M1 conditions were the three interferon-induced proteins with tetratricopeptide repeats (IFITs: IFIT1, IFIT2, and IFIT3), which function in antiviral defense. Moreover, IFIT1, IFIT2, and IFIT3 mRNA were strongly upregulated in M1 polarized human primary macrophages and IFIT1 was also expressed in a subset of macrophages in aortic sinus and brachiocephalic artery sections from atherosclerotic ApoE$^{-/-}$ mice. On the basis of these results, we propose that IFITs may serve as useful markers of atherosclerosis and potentially other inflammatory diseases.



KEYWORDS: *macrophage, polarization, atherosclerosis, biomarker, proteomics, interferon-induced protein with tetratricopeptide repeats, IFIT1, IFIT2, IFIT3, interferon-stimulated gene*

INTRODUCTION

A common feature of inflammatory diseases is the accumulation of macrophages, derived from circulating monocytes, in the inflamed tissue. Rather than being a uniform cellular population, macrophages may be "polarized" in response to the local cytokine environment toward M1 or M2 phenotypes, which represent the two extremes of a spectrum of possible polarization states.[1] M1 macrophages, polarized by pro-inflammatory stimuli such as interferon-γ (IFN-γ) and tumor necrosis factor-α (TNF-α), generate pro-inflammatory cytokines (e.g., interleukin-1β (IL-1β), IL-6, IL-12, IL-18), are pro-oxidative, and cause tissue damage. In contrast, M2 macrophages, polarized by anti-inflammatory cytokines such as IL-4 and IL-13, are generally considered to have anti-inflammatory, antioxidative, and tissue reparative properties.[2] Thus, the balance between M1 and M2 macrophages plays a key role in determining whether the inflammation is exacerbated or readily resolved.

The balance between M1 and M2 macrophage populations is believed to have a significant influence on the outcome of vascular inflammation and the progression of atherosclerosis. M1 macrophages are most prevalent in the rupture prone "shoulder" regions of atherosclerotic plaques and are key contributors to plaque instability, whereas the fibrous caps of atherosclerotic plaques have a more even balance of M1 and

Received: November 19, 2017
Published: March 6, 2018

DOI: 10.1021/acs.jproteome.7b00828
J. Proteome Res. XXXX, XXX, XXX−XXX

M2 macrophages, consistent with the tissue reparative properties of M2 macrophages.[3] While M1 macrophages are generally considered detrimental in the context of atherosclerosis, it is important to recognize that M2 macrophages may also contribute to atherogenesis as they are associated with early plaque development, are highly plastic and can change phenotype in response to the local microenvironment, and some M2 macrophage subsets exhibit inflammatory properties.[4] Nevertheless, in the context of unstable atherosclerotic plaques, it is generally thought that it would be advantageous to maintain or enhance M2 macrophages while limiting M1 macrophages.

To understand the physiological and pathological roles of macrophages and to monitor the disease progression and the effects of trial therapeutics, it is important to identify the phenotypes of macrophages present in specific inflamed tissues. Traditionally, M1 macrophages have been identified via the expression of pro-inflammatory cytokines (e.g., IL-1β, TNF-α, and IL-6), chemokines (e.g., CXCL9, CXCL10, and CXCL11) and their receptors (e.g., CCR7) together with inducible nitric oxide synthase (iNOS) and the costimulatory molecules CD80 and CD86.[5] Conversely, M2 macrophages are identified via the expression of pro-fibrotic (e.g., transforming growth factor-β (TGF-β) and insulin-like growth factor-1 (IGF-1))[6] and anti-inflammatory (IL-10) cytokines and scavenger receptors (mannose receptor C-type 1 (MRC-1), CD36)[7] and the anti-inflammatory cell surface marker, CD163.[5] While such biomarkers are useful, current limitations include differences between the expression profiles of murine and human M1 and M2 macrophages[8,9] and differences between *in vitro* and *in vivo* macrophages.[10,11] As such, investigations would greatly benefit from an improved set of biomarkers for M1 and M2 macrophages.

Several previous studies have explored the proteomes of polarized macrophages. An early study of primary human macrophages[12] identified a number of proteins that are specifically up- or down-regulated in macrophages treated with lipopolysaccharide (LPS) or IFN-γ compared to untreated controls, but did not investigate the more typical dual LPS/IFN-γ treatment or M2 polarized cells. Another study[13] focused on membrane associated proteins, identifying several that distinguish M1 or M2 polarized from nonpolarized primary mouse macrophages; in particular, intercellular adhesion molecule 1 (ICAM-1) and integrin components ITGB1 and ITGAL were upregulated in M1 polarized macrophages, whereas transferrin receptor (TFRC), disintegrin and metalloproteinase domain-containing protein ADAM10 and integrin component ITGAX were upregulated in M2 polarized macrophages. More recently, two studies[14,15] have used comparative proteomic methods (2D electrophoresis or SILAC) and the well-established THP-1 cell line to identify proteins expressed differentially between M1 and M2 macrophages, although neither of these studies included comparisons to nonpolarized cells.

In light of recent advances in label-free proteomics methodologies, we have now performed a systematic global proteomics comparison of M1 polarized, M2 polarized and nonpolarized THP-1 macrophages. From these data, we have identified proteins that are either up- or down-regulated in the M1 or the M2 phenotypes and we have further validated these proteins using targeted proteomics and, for a subset of proteins, Western blot analysis. Importantly, these observations were confirmed in human primary macrophages and *in vivo*, utilizing a murine model of atherosclerosis. The study provides an expanded set of macrophage polarization markers that will serve as a valuable benchmark for future studies of inflammatory diseases and treatments.

■ EXPERIMENTAL PROCEDURES

Materials

RPMI 1640 media and heat-inactivated fetal bovine serum (FBS) were purchased from Life Technologies. Phorbol 12,13-dibutyrate, was purchased from Calbiochem. Human pan monocyte isolation kit and M-CSF were purchased from Miltenyi Biotec and Ficoll-Paque PLUS gradient from GE Healthcare. Formic acid, sodium deoxycholate (SDC), HEPES, β-casein from bovine milk, ethyl acetate, chloroacetamide (CAA), LPS, IFN- γ, IL-4, phosphate buffered saline (PBS), and Tris (2-carboxyethyl)phosphine hydrochloride solution (TCEP) were purchased from Sigma-Aldrich. Sequencing grade modified trypsin was purchased from Promega. The RNeasy Mini kit and RNase-free DNase were purchased from Qiagen. The following antibodies were used: rabbit anti-SOD2 (Merck Millipore); mouse anti-p47-phox (neutrophil cytosolic factor 1, NCF1) (Santa Cruz Biotech); mouse anti-NFkB p100/p52 (Abcam); mouse anti-STAT1 (BD Biosciences); rabbit anti-IFIT1 (OriGene); mouse anti-iNOS (Abcam); anti-F4/80 (BioRad), rabbit anti-GADPH (Cell Signaling Technology); goat antimouse IgG (Sigma-Aldrich) and goat antirabbit IgG (Cell Signaling Technology), goat antirat Alexa Fluor 594 IgG and chicken antirabbit Alexa Fluor 647 IgG (Life Technologies) as the secondary antibodies. The Mouse on Mouse Immunodetection kit and Vectashield mounting medium with diamidino-2-phenylindole (DAPI) were purchased from Vector Laboratories. The High Capacity cDNA Reverse Transcription kit and PCR primers for IFIT1, IFIT2, IFIT3, ICAM1, CXCL11, CCR7, IL-1β, MRC-1, β-actin, and 18s were all purchased from Applied Biosystems.

Isolation of Primary Human Monocytes

Primary human monocytes were isolated from healthy blood donor buffy coats (Australian Red Cross Blood Service, Melbourne, Australia). Buffy coats were mixed with PBS supplemented with 0.5% FBS and 2 mM ethylenediaminetetraacetic acid (EDTA; Sigma-Aldrich) and layered onto Ficoll-Paque PLUS for density gradient centrifugation (400g, 40 min, acceleration = 1, deceleration = 0). The peripheral blood mononuclear cell (PBMCs) layer was collected and monocytes isolated using the human pan monocyte isolation kit according to the manufacturer's instructions. The purity of the monocyte population was confirmed to be at least 85% as determined by flow cytometry using $CD14^+/CD16^+$ expression.

Cell Treatment Conditions

Human monocytic cells (THP-1) were cultured in high glucose RPMI 1640 medium, supplemented with 10% heat-inactivated FBS. Monocytes were grown in T75 tissue culture flasks in a humidified incubator (Sanyo MCO-18AIC CO_2 incubator, 5% CO_2, 37 °C) and passaged every 3−4 days. For cell treatment, monocytes were seeded into six-well plates at a density of 1×10^6 cells/well and differentiated to macrophages with phorbol 12,13-dibutyrate (PDBu; 10 nM) for 24 h. The successfully differentiated macrophages were subsequently treated with IFN-γ (5 ng/mL) and LPS (10 ng/mL) for M1 phenotype, IL-4 (25 ng/mL) for M2 phenotype, or left untreated (M0) as control for 48 h (37 °C, 5% CO_2). After the treatment: for real-

DOI: 10.1021/acs.jproteome.7b00828
J. Proteome Res. XXXX, XXX, XXX−XXX

time PCR analysis and Western blotting, the cells were lysed directly on the plate; for secretome analysis, the media were collected, and for analysis of the global cellular proteome, the cells were trypsinized, washed with ice-cold PBS, and stored at −80 °C for later use. Isolated donor blood-derived primary monocytes (1×10^6 cells/mL) were differentiated into macrophages by culturing for 7 days in RPMI 1640 Glutamax medium, supplemented with 10% FBS, 1× antibiotic/antimyotic (Gibco Life Technologies, USA), 1 mM sodium pyruvate (Sigma-Aldrich), 1× nonessential amino acids (NEAA; Gibco Life Technologies), and 50 ng/mL M-CSF. Following differentiation, the primary human macrophages were either left untreated (M0), treated with 100 ng/mL LPS and 20 ng/mL IFN-γ (M1), or treated with 25 ng/mL IL-4 (M2). Cells were polarized for either 6 h for real-time PCR or 24 h for Western blotting. All the experiments were performed at least three times independently.

RNA Extraction and Real-Time PCR

Total RNA was extracted from macrophages using the RNeasy Mini Kit according to the manufacturer's instructions. RNase-free DNase was used to remove any contaminating DNA. The amount of RNA in each sample was quantified using a Nanodrop 1000D spectrophotometer (ThermoScientific), which measures absorbance at 260 and 280 nm. An A_{260}/A_{280} ratio of 2 or more was considered sufficiently pure. One microgram of RNA from each sample was reverse transcribed into cDNA using the High Capacity cDNA Reverse Transcription Kit with the reaction run in a thermal cycler (BioRad MyCycler, BioRad Laboratories). The resultant cDNA was used as a template for real-time PCR with Taqman primers and probes for IFIT1, IFIT2, IFIT3, ICAM1, CXCL11, CCR7, IL-1β, and MRC-1. β-actin and 18S were used as housekeeping genes. Real-time PCR was run in triplicate on the CFX96 Touch Real-Time PCR Detection Machine (BioRad Laboratories). Gene expression was quantified relative to the average M0 value using the comparative cycle threshold (Ct) method with the formula: Fold change = $2^{-\Delta\Delta Ct}$.[16]

Sample Preparation for Mass Spectrometric Analyses

To prepare samples for analysis of global proteomes, THP-1 cells were thawed and then lysed in 100 mM HEPES, 1% sodium deoxycholate, pH 8.2. The sample was sonicated and the protein concentration was determined with a BCA kit (Thermo Fisher Scientific). Two-hundred micrograms of proteins from the cell lysate was reduced with 10 mM TCEP, alkylated with 40 mM CAA, and then digested overnight with sequencing grade trypsin (Promega) with the ratio of 1:100 (trypsin to total protein, w/w) at 37 °C. Phase transfer was used to remove sodium deoxycholate.[17] For better coverage of the proteome, off-line basic pH reverse-phase fractionation was performed using an Agilent Zorbax 300 Extend-C18 5 μM column (4.6 × 250 mm^2) on a HPLC 1100 (Agilent Technologies). The gradient was run as follows: 100% buffer A (10 mM ammonium hydroxide) for 10 min, 2.5% buffer B (10 mM ammonium hydroxide, 80% acetonitrile) for 4 min, 2.5% buffer B to 40% buffer B for 60 min, 55% buffer B for 8 min, 100% buffer B for 8 min, 0% buffer B for 10 min. The fractions were collected from 10 to 82 min. Seventy-two fractions were collected and pooled in a noncontiguous manner as previously described to obtain a total of six fractions.[18] Fractions were then dried completely in a lyophilizer (Labconco) and dissolved in buffer A (0.1% formic acid, 2% acetonitrile).

To prepare samples for protein quantification by parallel reaction monitoring (PRM), the cells were identically processed, but the fractionation step was omitted. Instead, the tryptic peptides were directly desalted with C18 Stage-Tips,[19] dried in a Speed Vac and dissolved in buffer A (0.1% formic acid, 2% acetonitrile).

To prepare samples for the analysis of secreted proteins, 8 mL of media from the THP-1 cell cultures was concentrated with 3 kDa-cutoff Amicon Ultra-4 Centrifugal Filter Devices (Millipore) and the buffer was exchanged to 50 mM Tris, 150 mM NaCl, pH 8.0 by centrifuging multiple times for 30−40 min, 4000g, 4 °C. The protein concentration was determined using a Bradford assay (Expedeon). Two milligrams of protein was reduced with 10 mM TCEP at 50 °C for 20 min and alkylated by incubation with 20 mM CAA for 20 min in the dark. The protein mixture was loaded into a 50 kDa-cutoff Amicon Ultra-0.5 Centrifugal Filter Device (Millipore) and centrifuged at 14 000g for 15 min. The concentration of the filtrate was determined using a NanoDrop spectrophotometer. Trypsin was added at a ratio of 1:100 (w/w) and the sample incubated overnight at 37 °C. The sample was acidified with formic acid and desalted using C18 Omix tips (Agilent). The eluted peptides were dried in a Speed Vac (Labconco) and dissolved in buffer A.

LC−MS/MS for Proteome Analysis

For analysis of global cellular proteomes and secretomes, the peptide samples were analyzed by LC−MS/MS using a Q Exactive and Q Exactive Plus mass spectrometer (Thermo Fisher Scientific) respectively, coupled online to a RSLC nano HPLC (Ultimate 3000, UHPLC Thermo Fisher Scientific). Samples were loaded onto a 100 μm, 2 cm nanoviper Pepmap100 trap column and the peptides were separated on a RSLC nanocolumn 75 μm × 50 cm, Pepmap100 C18 analytical column (Thermo Fisher Scientific) using the following gradient of buffer A (0.1% formic acid, 2% acetonitrile) and buffer B (0.1% formic acid in 80% acetonitrile): 2.5% buffer B for 5 min; 2.5% to 12.5% buffer B over 1 min; 12.5% to 32.5% buffer B over 108 min; 32.5% to 42.5% buffer B over 6 min; 42.5% to 99% buffer B over 5 min; 99% buffer B for 7 min; 99% to 2.5% buffer B over 1 min; re-equilibration at 2.5% buffer B for 20 min. The eluent was nebulized and ionized using a Thermo nano Flex electrospray source with a distal coated fused silica emitter (New Objective). The capillary voltage was set at 1.7 kV. The Q Exactive instrument was operated in the data dependent acquisition mode to automatically switch between full scan MS and MS/MS acquisition. Each survey full scan (m/z 375−1800) was acquired in the Orbitrap with 70 000 resolution (at m/z 200) after accumulation of ions to a 3×10^6 target value with maximum injection time of 30 ms. Dynamic exclusion was set to 20 s. The 10 most intense multiply charged ions ($z \geq 2$) were sequentially isolated and fragmented in the collision cell by higher-energy collisional dissociation (HCD) with a fixed injection time of 60 ms, 17 500 resolution and automatic gain control (AGC) target of 5×10^4.

For analysis of targeted proteomics by PRM, peptide samples were subjected to LC−MS/MS using the same LC gradient as used for the analysis of global cellular proteomes and secretomes (above). The 224 precursor ions with scheduled retention times were set to target the MS2 scans. The inclusion list can be found in Table S1. A survey full scan and data-independent acquisition (DIA) mode were applied to generate

DOI: 10.1021/acs.jproteome.7b00828
J. Proteome Res. XXXX, XXX, XXX−XXX

targeted data. The survey full scan MS spectra (m/z 375–1800) were acquired in the Orbitrap with 70 000 resolution (at 200m/z), after accumulation of ions to a 1×10^6 target value with maximum injection time of 50 ms. The DIA scan was acquired with 17 500 resolution, AGC target 2×10^5, loop counts of 20 and an isolation window of 2 m/z.

Raw DDA data files have been deposited at PRIDE[20] with the data set identifier PXD008204 (password: 3fFbrLyP).

Global Proteomics Data Analysis and Statistics

For identification of the global cellular proteome, all generated files were submitted to MaxQuant (version 1.5.2.8), coupled with the Andromeda search engine,[21] to generate a list of proteins with label-free quantitation (LFQ) intensities. Database searching was performed with the following parameters: cysteine carbamidomethylation as a fixed modification; methionine oxidation and N-terminal acetylation as variable modifications; up to two missed cleavages permitted; mass tolerance of 20 ppm; 1% protein false discovery rate (FDR) for protein and peptide identification; and minimum two peptides for pairwise comparison in each protein. The human protein sequence database was downloaded from Uniprot in July 2014.

The lists of proteins with LFQ values were processed with Perseus (Version 1.5.0.31), a module from the MaxQuant suite. After removing the reversed and known contaminating proteins, the LFQ values were $\log_2$ transformed and the reproducibility across the biological replicates was evaluated by Pearson's correlation analysis. The replicates were grouped into M0, M1, and M2 with categorical annotation rows. At least two valid values out of the three repeats in at least one treatment condition was set to filter out the proteins without LFQ intensity. Missing values after filtering were replaced by imputation. Multiple-samples Test based on ANOVA (FDR = 5%) was performed to obtain identify the proteins with significantly changed expression levels. Only the proteins with ANOVA significant were used for the subsequent analyses.

For identification of the secretome under each treatment condition, LFQ values were generated in the same manner as for the global cellular proteome, except we used updated versions of MaxQuant (1.5.3.30), Perseus (1.5.4.0), and the protein sequence database (May 2016). The proteins were submitted to ProteinSide (http://www.proteinside.org/) for signal peptide prediction. The secreted proteins were imported back into Perseus for statistical analysis. The statistical analysis was the same as that used for the global cellular proteome.

Targeted Proteomics Data Analysis

A spectral library was built in Skyline 3.1,[22] based on the global proteomics MS/MS data from MaxQuant generated from global proteome identification. All precursors from 18 selected proteins were refined in Skyline, making sure that at least two proteotypic peptides were used for quantification of each protein. In total, a transition list of 224 precursor ions with scheduled retention times was generated in and exported from Skyline. All PRM data analysis and data integration was performed in Skyline. The mass spectrometric raw files were imported into Skyline under the targeted acquisition mode and each peak was manually examined to ensure accurate quantification. A dot product cutoff of 0.9 was chosen (1 indicates a perfect match to the spectral library) and the sum of the peak area for each protein was exported to Excel and normalized to the corresponding peak area in untreated cells (M0). Statistical analysis was based on a one way ANOVA, Dunnett's post hoc test in GraphPad Prism 6.0.

Western Blot

Equal amounts of protein from M0, M1, or M2 treated THP-1 or primary macrophages were separated with a NuPAGE 4–12% SDS-PAGE (Invitrogen) and electro-transferred to a polyvinylidene difluoride (PVDF) membrane. The PVDF membrane was blocked with 5% skim milk in Trisec-Buffered Saline (TBS, 200 mM Tris, 150 mM NaCl, pH 7.5) with 0.1% tween-20 for 1 h and subsequently probed with primary antibodies overnight at 4 °C. Secondary antibodies bound with horseradish peroxidase were incubated at room temperature for 1 h, detected by enhanced chemiluminescence (Western Lightning Plus ECL, PerkinElmer), documented and quantitated with the ChemiDoc MP Imager and ImageLab software (Bio-Rad). All the band intensities generated from ImageLab software were first normalized to the housekeeping gene GADPH and then expressed relative to the untreated (M0) conditions. Statistical significance was based on one way ANOVA, Dunnett's post hoc test in GraphPad 6.0.

Functional Annotation

Data for proteins showing significant differences between treatment conditions in the global cellular proteome and secretome were combined and submitted to Perseus for Z-score normalization, with matrix access by rows. Hierarchical clustering was used to assess the similarity between each replicate and treatment condition using the Euclidean for distance calculation and average for linkage in row and column trees with a maximum of 300 clusters.[23] Subsequently, TM4:MeV[24] was used to group the proteins into different clusters based on k-Means/Medians Clustering[25] with the setting of six clusters. The proteins from different clusters were submitted into FunRich 3.0[26] for biological pathway and transcription factor enrichment analysis.

Experimental Animals

All procedures were approved by the Monash Animal Research Platform (MARP) Animal Ethics Committee and conducted in compliance with the National Health and Medical Research Council of Australia's (NHMRC) Guidelines for the Ethical and Humane Use of Animals in research. ApoE$^{-/-}$ mice were obtained from the Animal Resources Centre (Canning Vale, WA, Australia). Mice were bred on the C57BL6/J background and only male mice were used in this study.

ApoE$^{-/-}$ mice were weaned at 3 weeks of ages, housed in groups of up to four littermates, and given access to normal chow and water *ad libitum*. At 5 weeks of age, they were placed on a high fat diet (HFD; 22% fat, 0.15% cholesterol; SFOO-219, Specialty Feeds, WA, Australia) and maintained on this diet for a further 14 weeks. At this time point, mice were killed by an overdose of isoflurane inhalation (Baxter Healthcare, Australia) and hearts and brachiocephalic arteries were frozen in Optimal Cutting Temperature compound (OCT; Sakura Finetek, USA) for immunohistochemistry.

Immunohistochemistry

Frozen mouse hearts in OCT were cut into 10 μm sections through the aortic sinus and mounted on poly-L-lysine coated glass slides. Brachiocephalic arteries in OCT were also cut into 10 μm sections and mounted on poly-L-lysine coated glass slides. Sections were fixed in 4% paraformaldehyde (PFA; Sigma-Aldrich, USA) for 15 min and washed with PBS. Fixed sections were blocked with 2% bovine serum albumin (BSA; Sigma) in PBS with 0.2% Triton X-100 (TX; Sigma) and Mouse on Mouse (MOM) Ig blocking reagent (Vector

DOI: 10.1021/acs.jproteome.7b00828
J. Proteome Res. XXXX, XXX, XXX−XXX

Journal of Proteome Research Article

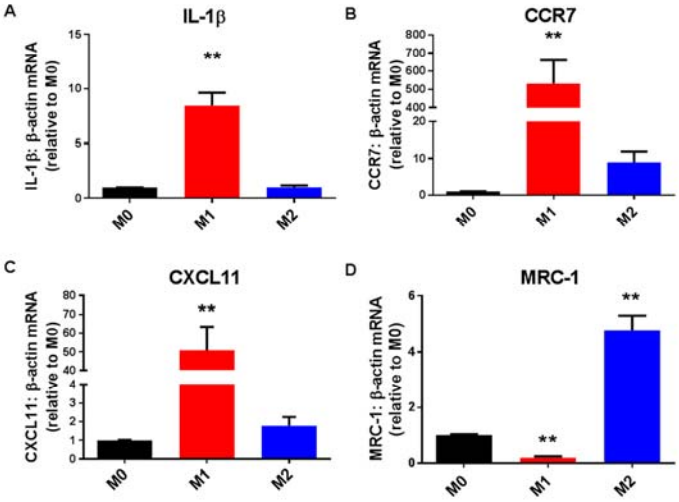


Figure 1. M1 and M2 marker mRNA expression in THP-1 macrophages. PDBu-differentiated THP-1 macrophages were left untreated (M0) or treated with IFN-γ/LPS (5 ng/mL IFN-γ + 10 ng/mL LPS; M1 conditions) or IL-4 (25 ng/mL; M2 conditions) for 48 h. mRNA expression of (A–C) M1 markers (A) IL-1β, (B) CCR7, and (C) CXCL11 and (D) M2 marker MRC-1 were determined by qRT-PCR and expressed relative to untreated macrophages (M0). All data are presented as mean ± SEM, $n = 8$. $**P < 0.01$ vs M0 (1-way ANOVA followed by Dunnett's post hoc test).

Laboratories, USA) for 1 h and subsequently incubated with primary antibodies against rabbit anti-IFIT1 (1:100, OriGene) and rat anti-F4/80 (cell surface macrophage marker, 1:100, Bio-Rad Laboratories) overnight at room temperature. Sections were then washed and incubated with the donkey antirabbit-488 (1:250, Life Technologies, USA) and goat antirat-564 (1:1000, Life Technologies, USA) secondary antibodies in the dark, at room temperature for 2 h. Finally, slides were washed, Vectashield mounting medium with diamidino-2-phenylindole (DAPI) added (Vector Laboratories, USA), and slides coverslipped. Appropriate antibody controls were performed by staining sections in the absence of primary or secondary antibodies and imaged on the same day under identical settings. Sections were imaged using a Nikon C1 upright confocal fluorescence microscope equipped with 20× or 60× objectives and run on NIS Elements Software (Nikon, Tokyo, Japan). Representative images were chosen from $n = 6$ (brachiocephalic arteries) or $n = 8$ (aortic Sinus) ApoE$^{-/-}$ mice, three sections per animal.

■ RESULTS

M1 and M2 Polarization of THP-1 Macrophages

THP-1 cells are an immortalized human monocytic cell line commonly used as a model for monocytes and macrophages. In this study THP-1 cells were differentiated to a macrophage state by treatment with PDBu, then polarized toward either an M1 or M2 phenotype by treatment with IFN-γ and LPS or with IL-4, respectively; untreated control cells (differentiated but not polarized) are labeled M0. To validate these polarization conditions, we used qRT-PCR to determine the mRNA levels for several recognized polarization markers (Figure 1). Treatment with IFN-γ/LPS lead to a significant increase in the M1 markers IL-1β (8-fold), CCR7 (500-fold), and CXCL11 (50-fold). While IL-4 treatment had no effect on M1 marker expression, the M2 marker MRC-1 was increased five-fold. In addition, M1 polarization was associated with an ~80% reduction in MRC-1 mRNA expression. Taken together, these results confirm polarization to M1 and M2 macrophage phenotypes by IFN-γ/LPS and IL-4 treatments, respectively.

Global Proteomics of Differentially Polarized THP-1 Macrophages

We used global label-free quantification (LFQ) proteomics[27] to compare cellular protein expression levels in unpolarized (M0), M1-polarized, and M2-polarized THP-1 macrophages. Using a FDR cutoff of 0.01, we identified 6656 proteins and we were able to quantify 6131 proteins in at least two out of three independent repeats of each treatment (Table S2). The excellent reproducibility of the LFQ methods is illustrated by the coefficient of variation (CV) curves (Figure S1); the median CV values were 14.7% for M0, 15.1% for M1, and 16.3% for M2 conditions, respectively.

Statistical analysis (ANOVA, permutation-based FDR < 0.05) identified 280 proteins (Table S3) that show significantly different expression levels between any two of the three treatment conditions. Hierarchical clustering, based on the expression levels of these proteins, indicated that the three repeats from each treatment are clustered together and that the expression profiles for the M0 and M2 conditions are more similar to each other than to the profile for M1 conditions (Figure 2A). This is confirmed by a principal component analysis (Figure S2) and is also consistent with previous reports that IL-4 treated macrophages showed less substantial changes in transcriptional and protein expression levels in comparison to LPS and IFN-γ treated cells.[9,14]

DOI: 10.1021/acs.jproteome.7b00828
J. Proteome Res. XXXX, XXX, XXX−XXX

Journal of Proteome Research Article

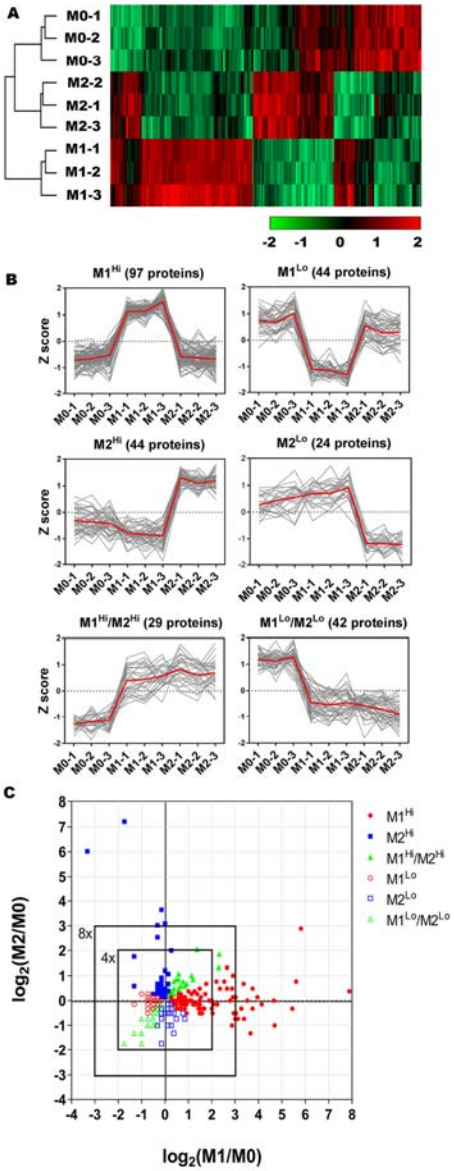


Figure 2. Cluster analysis of proteomics data. (A) Hierarchical clustering of the protein intensities determined for three repeats under each of three conditions; labels indicate the condition followed by the repeat number (e.g., M1−2 is the second repeat under M1 conditions). The significantly up- or down-regulated proteins were clustered in Perseus. The color bar indicates the Z-score of intensities. (B) Z-score graphs for each of the six clusters. Each graph shows a set of thin gray lines, each connecting the Z-scores for one protein in the cluster, and a thick red line connecting the average Z-score for proteins in the cluster. Z-scores are presented for three repeats under each of three conditions, labeled as in panel A. (C) Scatter plot showing the changes in protein intensities under M1 and M2 conditions, expressed relative to the basal (M0) conditions, for the clusters $M1^{Hi}$ (filled red circles), $M1^{Lo}$ (open red circles), $M2^{Hi}$ (filled blue squares), $M2^{Lo}$ (open blue squares), $M1^{Hi}/M2^{Hi}$ (filled green triangles), and $M1^{Lo}/M2^{Lo}$ (open green triangles). The black boxes indicate four-fold and eight-fold changes.

Cluster Analysis of Significantly Regulated Proteins

The 280 differentially expressed proteins were subjected to cluster analysis, yielding the six clusters shown in Figure 2B and C and summarized in Table 1. Among the 280 differentially expressed proteins, 97 were classified as upregulated in M1 conditions (cluster $M1^{Hi}$), 44 as upregulated in M2 conditions (cluster $M2^{Hi}$), and 29 additional proteins as upregulated in both M1 and M2 conditions (cluster $M1^{Hi}/M2^{Hi}$). In addition, 44 and 24 proteins were downregulated under M1 and M2 conditions, respectively, and 42 additional proteins were downregulated in both polarization conditions; these three clusters are labeled $M1^{Lo}$, $M2^{Lo}$, and $M1^{Lo}/M2^{Lo}$, respectively. Interestingly, none of the clusters involved predominant upregulation of proteins under one polarization condition and downregulation of the same proteins under the alternative polarization condition, and a figure of merit analysis (Figure S3) indicated that extension to a larger number of clusters was not justified.

The proteins in each cluster were analyzed to identify significantly enriched biological processes and transcription factors whose target gene products were significantly enriched in each cluster (Table 1). Cluster $M1^{Hi}$ was significantly enriched ($p < 0.01$; at least 4 proteins per process) for 24 biological processes, many of which are related to interferon or cytokine signaling, as expected from the M1 treatment conditions. Consistent with this finding, transcription factor enrichment analysis strongly indicated enhanced activity of interferon regulatory factor 1 (IRF1) in the $M1^{Hi}$ group; the mechanistic significance of this finding is discussed below. Smaller numbers of significantly enriched processes were found for other clusters. In particular, cluster $M2^{Hi}$ was enriched in processes related to signal transduction, platelet degranulation and aggregation, oxidation−reduction, cell−cell adhesion, and the innate immune response. Among the 12 transcription factors whose gene products were significantly enriched in the $M2^{Hi}$ cluster, the top five (FOS, FOSB, JUN, JUNB, and JUND) are all components of the activator protein 1 (AP-1) transcription factor, which is responsible for a variety of cellular functions, including cell proliferation, differentiation and apoptosis.[28]

Identification of Secreted Proteins from Differentially Polarized THP-1 Macrophages

Previous studies indicate that M1 and M2 macrophages differ substantially in the cytokines that they secrete. Therefore, to complement the cellular proteomics data, we analyzed the secreted proteins (secretome) from M1 and M2 polarized in comparison to unpolarized THP-1 macrophages. Although the use of serum-free media may have allowed detection of a larger

DOI: 10.1021/acs.jproteome.7b00828
J. Proteome Res. XXXX, XXX, XXX−XXX

Table 1. Summary of Protein Clusters Identified

cluster	number of proteins	biological pathways	transcription factors
$M1^{Hi}$	97	type I IFN-mediated signaling; defense response to virus; response to virus; negative regulation of viral genome replication; IFN-γ-mediated signaling; innate immune response; negative regulation of type I IFN production; inflammatory response; positive regulation of NF-κB activity; apoptotic process; response to cytokine stimulus; I-κB kinase/NF-κB cascade; leukocyte migration; positive regulation of protein phosphorylation; negative regulation of apoptotic process; cellular response to LPS; positive regulation of angiogenesis; positive regulation of transcription from RNA polymerase II promoter; cell adhesion; angiogenesis; negative regulation of neuron apoptotic process; T cell receptor signaling; aging; protein homooligomerization	IRF-1
$M2^{Hi}$	44	platelet aggregation; platelet degranulation; cell−cell adhesion; signal transduction; oxidation−reduction process; innate immune response	FOS; FOSB; JUN; JUNB; JUND; SRF; PPARA; BACH2; BACH1; SP1; RUNX2; STAT3
$M1^{Hi}/M2^{Hi}$	29	tRNA aminoacylation for protein translation	ZEB1; SP1; ELF2; TGIF1; TEAD1; SP4; MEIS2; FOXD1; HOXB3; EGR1; NFYA; IRF1; ELF1
$M1^{Lo}$	44	positive regulation of GTPase activity	
$M2^{Lo}$	24	mitosis	MYC; ESRRA
$M1^{Lo}/M2^{Lo}$	42		FOXK1; FOXJ2

number of secreted proteins, we chose to retain the same growth conditions as in the above analysis of cellular proteins because growth in serum-free media could also have altered the secretome. To reduce interference from serum proteins, we focused our analysis on proteins with molecular masses below ~30 kDa, which were separated by ultrafiltration from higher molecular weight proteins, including the major serum proteins albumin and immunoglobulins.

Using the same global LFQ workflow as for the cellular proteins above, we identified 168 secreted proteins from the media (Table S4), of which 40 were derived from precursor proteins containing signal peptides, as determined using Proteinside (http://www.proteinside.org/), and an additional 76 were determined to be potentially secreted proteins without signal peptides (Table S5). Six of these proteins (Table S6) were observed in only one treatment condition and, strikingly, these included four members of the chemokine family (CCL3, CCL4, CCL5, and CXCL10), all of which were detected in media from M1 polarized cells but not in either unpolarized or M2 polarized cells. Consistent with our observations, Cassol et al. have reported previously that M1 polarization of human monocyte-derived macrophages can significantly increase the expression of these chemokines.[29] In contrast to this clear upregulation of chemokines in M1 polarized cells, we did not identify any chemokines or other cytokines upregulated in the secretome of M2 polarized cells.

Validation of Significantly Regulated Proteins

To verify the changes in the protein expression observed above, 18 of the cellular proteins that showed differential expression in the global proteomics data set were arbitrarily selected for analysis by parallel reaction monitoring (PRM), a targeted, quantitative MS/MS method.[30] These included 11 proteins from the $M1^{HE}$ cluster, six proteins from $M2^{HE}$ cluster, and one protein from the $M1^{HE}/M2^{HE}$ cluster. The changes in protein levels for the 18 selected proteins under M1 and M2 polarization conditions are listed in Table 2 for the PRM data set in comparison with the global proteomics data set. Among the 11 proteins from the $M1^{HE}$ cluster, all showed significant increases under M1 conditions and none showed significant changes under M2 conditions in the PRM data set. Similarly, among the six proteins from the $M2^{HE}$ cluster, all showed significant increases under M2 conditions and none showed significant changes under M1 conditions in the PRM data set. Finally, the protein from the $M1^{HE}/M2^{HE}$ cluster displayed significant increases under both M1 and M2 conditions in the PRM data set. Thus, the PRM data set was in excellent agreement with the global LFQ proteomics data set.

Considering that the major changes in protein expression levels observed were induced by M1 polarization, we used Western blots to further evaluate changes for four of the proteins found to be significantly increased under M1 conditions in the global proteomics and PRM data sets (NCF1, NFκB2, SOD2, STAT1); these proteins have previously been reported to be overexpressed in M1 macrophages.[31−33] As shown in Figure 3 and Table 2, all four proteins showed significant increases by Western blot under M1 conditions but not under M2 conditions, confirming the changes observed by proteomics methods.

DOI: 10.1021/acs.jproteome.7b00828
J. Proteome Res. XXXX, XXX, XXX−XXX

Journal of Proteome Research Article

Table 2. Comparison of Changes in Protein Levels Determined by Global Proteomics, Targeted Proteomics, and Western blot

	global proteomics (DDA) fold change		targeted proteomics (PRM) fold change		Western blot fold change	
gene name	M1	M2	M1	M2	M1	M2
DTX3L	2.7 ± 0.4[c]	0.9 ± 0.1	5.4 ± 0.8[b]	1.3 ± 0.1		
IFIT3	25.8 ± 3.6[d]	0.5 ± 0.1	39.1 ± 4.2[c]	0.9 ± 0.1		
NCF1	6.3 ± 0.3[d]	1.4 ± 0.2	6.0 ± 1.0[b]	1.7 ± 0.4	5.5 ± 0.4[c]	1.7 ± 0.4
NFKB2	6.9 ± 0.7[d]	1.0 ± 0.1	10.5 ± 0.7[b]	1.2 ± 0.2	2.2 ± 0.4[a]	0.9 ± 0.1
OAS3	10.4 ± 2.3[d]	0.6 ± 0.2	15.7 ± 2.4[d]	1.0 ± 0.1		
OASL	17.4 ± 3.3[d]	0.9 ± 0.2	12.1 ± 1.0[c]	1.6 ± 0.6		
PARP9	11.5 ± 5.4[b]	1.0 ± 0.5	15.1 ± 2.5[d]	1.2 ± 0.1		
RELB	4.7 ± 0.8[c]	0.9 ± 0.1	13.6 ± 3.1[b]	2.5 ± 0.1		
SOD2	7.5 ± 0.8[d]	1.1 ± 0.1	6.9 ± 0.9[c]	1.3 ± 0.2	2.2 ± 0.2[b]	1.0 ± 0.2
STAT1	3.2 ± 0.1[d]	0.8 ± 0.1	4.5 ± 0.7[c]	1.1 ± 0.1	1.8 ± 0.1[b]	0.9 ± 0.1
EIF2AK2	2.8 ± 0.4[d]	1.0 ± 0.1	3.0 ± 0.5[b]	1.1 ± 0.1		
TGM2	0.3 ± 0.1	149 ± 68[c]	0.8 ± 0.2	32 ± 15[b]		
CD209	0.8 ± 0.1	8.1 ±1.2[c]	1.3 ± 0.2	5.7 ± 0.1[d]		
F13A1	0.1 ± 0.1	65 ± 22[d]	2.6 ± 1.1	666 ± 65[d]		
DOK2	0.9 ± 0.1	1.7 ± 0.2[b]	0.9 ± 0.2	1.5 ± 0.1[a]		
LSP1	0.4 ± 0.1	3.4 ± 0.4[d]	0.6 ± 0.1	2.1 ± 0.2[c]		
SORT1	0.9 ± 0.1	12.5 ± 2.1[d]	0.7 ± 0.1	17.0 ± 0.7[d]		
TGFB1	1.5 ± 0.1[a]	1.8 ± 0.2[b]	1.8 ± 0.3[a]	1.8 ± 0.1[a]		

[a] $p < 0.05$. [b] $p < 0.01$. [c] $p < 0.001$. [d] $p < 0.0001$.

Enhanced Expression of Interferon-Induced Protein with Tetratricopeptide Repeats 1 (IFIT1) in Primary Human Macrophages

The THP-1 macrophage proteomics data revealed large and significant increases, under M1 conditions, in the expression levels of three closely related proteins known as interferon-induced proteins with tetratricopeptide repeats 1, 2, and 3 (IFIT1, IFIT2, and IFIT3);[34−36] levels of these proteins were increased by factors of 236 ± 101, 25 ± 17, and 26 ± 4, respectively, in M1-polarized THP-1 macrophages relative to vehicle-treated controls. Since THP-1 cells are an immortalized cell line, it is important to determine whether the expression changes we observed in THP-1 macrophages are relevant to macrophages derived from primary monocytes. To this end, we initially used macrophages differentiated *in vitro* (using macrophage-colony stimulating factor (M-CSF)) from primary human monocytes, isolated from the buffy coat of donor blood, to determine the mRNA levels for IFIT1, IFIT2, and IFIT3 as well as the established M1 polarization marker ICAM-1. M1 polarization was associated with an increase in mRNA expression of IFIT1, IFIT2, IFIT3, and ICAM-1 by 250-, 80-, 100-, and 20-fold, respectively (Figure 4A). Moreover, Western blotting confirmed a robust increase in IFIT1 protein expression in M1 polarized human primary macrophages (Figure 4B,C).

Expression of IFIT1 in Atherosclerotic Plaque Macrophages of ApoE$^{-/-}$ Mice

Considering that macrophages are abundant in atherosclerotic plaques and M1 macrophages are particularly implicated in atherosclerosis,[37] we used the ApoE$^{-/-}$ mouse model of atherosclerosis to assess the potential of IFIT1 as a M1 macrophage marker *in vivo*. The expression of IFIT1 in aortic sinus and brachiocephalic artery sections from atherosclerotic ApoE$^{-/-}$ mice was assessed by immunohistochemistry. Immunofluorescent staining revealed a subpopulation of atherosclerotic plaque macrophages (F4/80 positive cells) expressing IFIT1 in both the aortic sinus and brachiocephalic arteries (Figures 5 and 6). IFIT1 expression was also evident in nonmacrophage cells within the plaque and in the media and adventitia of the vessels. Despite this diffuse expression of IFIT1 there was a substantial population of F4/80 positive cells that did not appear to express IFIT1, likely to represent macrophage phenotypes other than the pro-inflammatory M1 subset, suggesting that IFIT1 could potentially be used as a marker to distinguish between macrophage subsets.

■ DISCUSSION

The balance between pro-inflammatory (M1-like) and anti-inflammatory (M2-like) polarized macrophages is thought to be critical in controlling the progression and clinical outcome of macrophage-mediated inflammatory diseases such as atherosclerosis. To understand the biochemical factors contributing to the functions of polarized macrophages and to identify macrophages with distinct phenotypes, it is important to identify proteins with differential expression in differently polarized (and nonpolarized) macrophages. In this study, we applied quantitative proteomics to identify differences in protein expression between THP-1 macrophages polarized to M1 and M2 states, in comparison with nonpolarized THP-1 macrophages. We found numerous significant differences between the differently treated cells and identified a number of proteins whose expression is significantly enhanced under either M1 or M2 conditions as well as a small number exhibiting significantly reduced expression levels (Figure 2). Here we focus on the proteins with expression enhancements of ≥4-fold (Table 3), most of which were observed under M1 conditions.

Among the most upregulated proteins in M1 conditions, there are some that have previously been identified as M1 markers, including the adhesion molecule ICAM1, two components of the noncanonical NFκB pathway (NFκB2 and RELB), and NCF1 (Figure 3, Table 2), a subunit of NADPH oxidase, which promotes formation of reactive oxygen species (ROS), a characteristic feature of M1 macrophages (Figure 3 and 4 and Table 2).[38] Some other M1 markers, such as CD38,[39] NFκB1,[1] and CD14,[40] were also significantly upregulated in M1 conditions, albeit more weakly. Similarly,

DOI: 10.1021/acs.jproteome.7b00828
J. Proteome Res. XXXX, XXX, XXX−XXX

Journal of Proteome Research Article

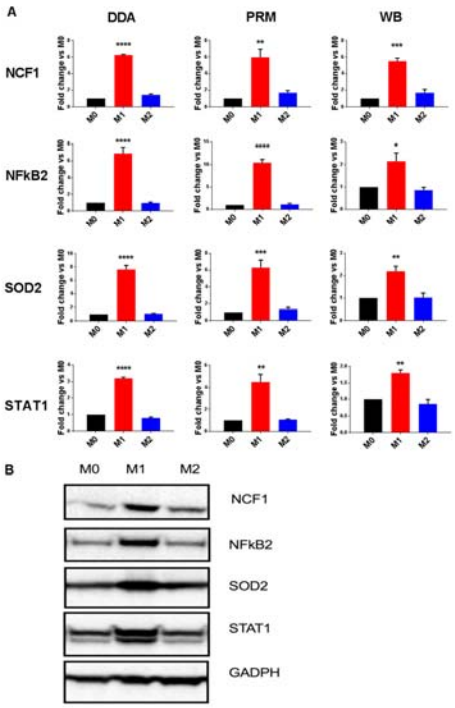


Figure 3. Validation of selected M1 markers by targeted proteomics and Western blot in THP-1 macrophages. (A) PDBu-differentiated THP-1 macrophages were left untreated (M0, black) or treated for 48 h with IFN-γ/LPS (5 ng/mL IFN-γ + 10 ng/mL LPS; M1, red) or IL-4 (25 ng/mL; M2, blue). The four proteins listed (left side) were independently quantified by DDA (left panels), PRM (center panels), and Western blot (right panels); quantities are shown relative to basal (M0) control. Data are presented as mean $\pm$ SEM, $n = 3$. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$ vs M0 (1-way ANOVA followed by Dunnett's post hoc test). (B) Western blot from a single representative experiment; glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as a loading control.

under M2 polarization conditions, we observed strong induction of CD209 (Table 2), a C-type lectin receptor expressed on macrophages, which is induced by IL-4[41] and has been used previously as a M2 polarization marker.[42] In addition, TGM2, a potential M2 marker,[43,44] was also strongly upregulated under M2 conditions. The observation of these proteins in the current study is reassuring and helps to validate THP-1 cells as an appropriate model for macrophage polarization. Furthermore, several of the upregulated proteins identified in this study have also been observed in previous proteomics studies of polarized THP-1 or primary macrophages, including ICAM1,[13] ACSL1,[13] C3,[14] and SOD2.[14] Nevertheless, as discussed below, the current study also identified a number of proteins that have not been found previously to be polarization markers, including IFIT1, IFIT2, IFIT3, OASL, OAS2, MX1, and MX2.

Perhaps our most striking observation was that the three related proteins IFIT1, IFIT2, and IFIT3 were all enhanced by at least 25-fold under M1 treatment conditions (Table 3). The enhanced expression in M1 macrophages was validated at the mRNA level for IFIT1, IFIT2, and IFIT3 in macrophages obtained by differentiation and subsequent polarization of primary human monocytes *in vitro*. IFIT1 protein upregulation was also confirmed in these cells. Furthermore, IFIT1 staining was clearly evident in the atherosclerotic plaques of ApoE$^{-/-}$ mice. For a population of cells, this staining colocalized with the macrophage marker F4/80; the identities of other IFIT1-positive cells remain to be determined. As not all F4/80 positive cells were IFIT1 positive, the double positive cells are likely to represent the M1 macrophage population in plaques. IFIT proteins are antiviral proteins whose expression is strongly induced by type I interferons (IFN-α and IFN-β).[35] Although they are structurally related, each being composed of several tetratricopeptide repeat (TPR) domains, they have different mechanisms of antiviral action.[34−36] IFIT1 and IFIT2 inhibit translation of viral mRNAs by binding to structural features present in viral mRNA, but not cellular mRNA, and also by binding elongation initiation factor 3 (eIF3).[35,36] IFIT1 has also been reported to inhibit viral replication by binding to the E1 helicase of human papilloma virus.[45] By contrast, IFIT3 acts as an adaptor protein, bridging the mitochondrial antiviral signaling (MAVS) complex to TANK-binding kinase (TBK1) and thereby causing derepression of the transcription factor nuclear factor kappa B (NFκB) in response to viral infection.[34,46] Considering previous reports that IFIT gene expression is only weakly induced by IFN-γ,[47] our observation of robust induction in THP-1 macrophages polarized by treatment with IFN-γ and LPS is novel and suggests that IFIT proteins may be useful markers for M1 polarization. To our knowledge, these proteins have not previously been characterized as M1 polarization markers, although Jablonski et al. reported 3.5- and 2-fold induction of IFIT-1 and IFIT-2 mRNA, respectively, in M1-polarized murine macrophages and Sudan et al. observed strong induction of IFIT1 mRNA in macrophages treated with IFN-β or IFN-γ.[39,48] Interestingly, two studies have shown reductions in IFIT mRNA in mouse macrophages or THP-1 monocytes treated with reputedly anti-inflammatory food extracts (from pistachio nuts and cranberries, respectively), prompting the authors of one study to suggest that IFIT-2 may be a biomarker for inflammatory cardiovascular disease.[49,50] The validation of IFIT proteins as cardiovascular disease markers will require future comparisons of tissue samples from patients and healthy controls.

Consistent with the elevation of IFIT proteins under M1 polarization conditions, we also observed significantly higher levels of several other proteins, or groups of proteins, involved in virus detection and MAVS-mediated antiviral signaling. In particular, we observed 7- and 8-fold enhancements of the two retinoic acid-inducible gene 1 (RIG-I)-like receptors, RIG-I and melanoma differentiation associated protein 5 (MDA5), which sense viral RNA and, upon activation, associate with MAVS on the surface of mitochondria. In addition, we found 8- and 7-fold enhancements, respectively, of interferon-stimulated gene 15 (ISG15), which is an antiviral ubiquitin-like protein, and Ubch8, the ubiquitin-conjugating (E2) enzyme that catalyzes conjugation of ISG15 to target proteins, including both RIG-I and IFIT1.[51,52] Under M1 conditions, there was also a large (18-fold) increase in the level of the 2′-5′-oligoadenylate synthetase (OAS)-like protein (OASL), which can bind to viral

DOI: 10.1021/acs.jproteome.7b00828
J. Proteome Res. XXXX, XXX, XXX−XXX

Journal of Proteome Research Article

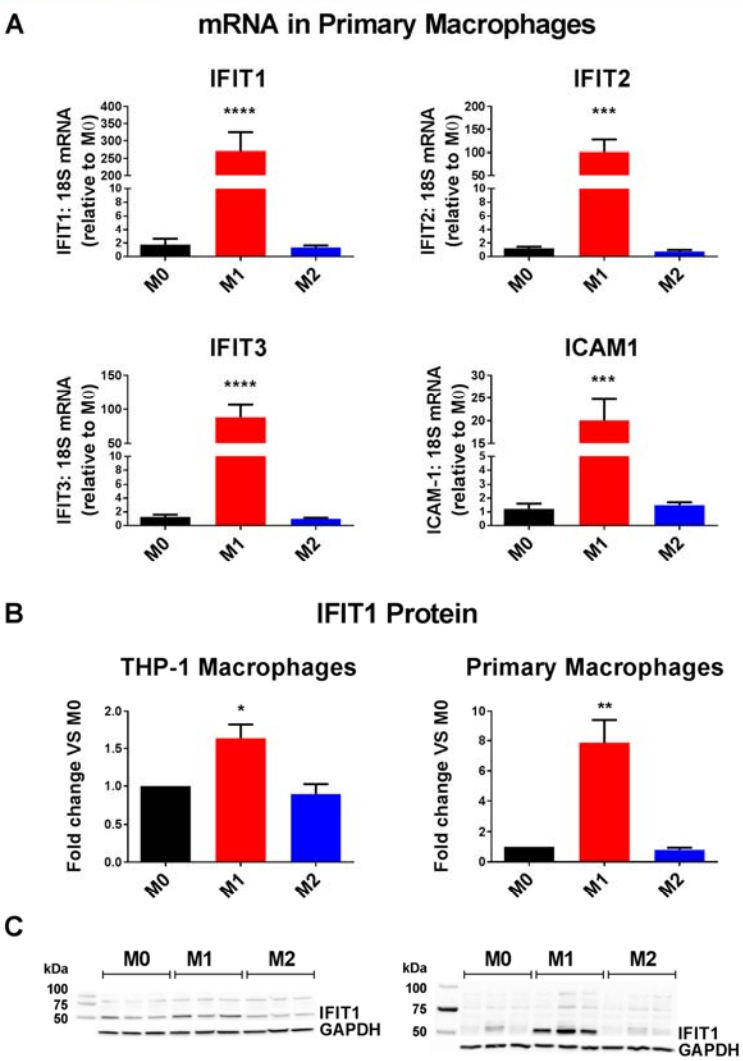


Figure 4. Validation of IFIT1, IFIT2, IFIT3, and ICAM1 expression in human primary macrophages. Monocytes from donor buffy coats were differentiated to macrophages and subsequently left untreated (M0, black), treated with IFN-γ/LPS (20 ng/mL IFN-γ + 100 ng/mL LPS; M1, red), or treated with IL-4 (25 ng/mL; M2, blue). (A) Following 6 h of polarization, mRNA expression of IFIT1, IFIT2, IFIT3, and ICAM1 was determined by qRT-PCR and expressed relative to untreated macrophages (M0). (B) Protein levels of IFIT1 in THP-1 and primary macrophages were assessed, via Western blotting, following 48 and 24 h of polarization, respectively, and expressed relative to untreated macrophages (M0). All data are presented as mean ± SEM, $n = 5-7$. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$ vs M0 (1-way ANOVA followed by Dunnett's post hoc test). (C) Western blot of 3 independent experiments; glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as a loading control.

double-stranded RNA (dsRNA) and mediate RIG-I activation and consequent MAVS formation and signaling.[53,54] In addition, the related OAS1, OAS2, and OAS3 were all increased 7- to 10-fold under M1 conditions. These proteins

DOI: 10.1021/acs.jproteome.7b00828
J. Proteome Res. XXXX, XXX, XXX−XXX

Journal of Proteome Research Article

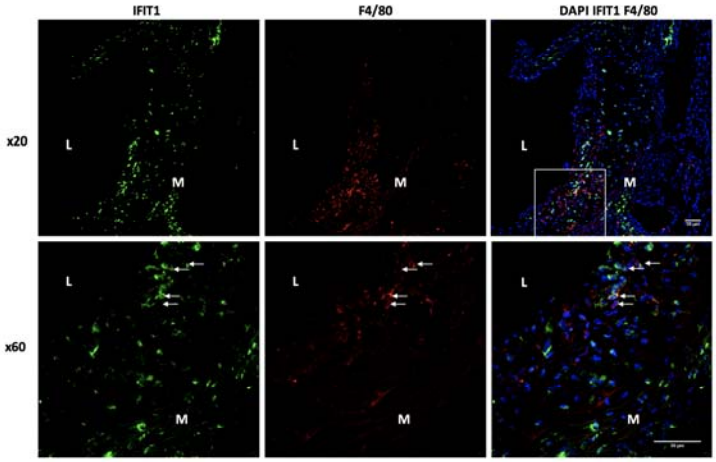


Figure 5. Expression and colocalization of IFIT1 with macrophage marker F4/80 in the aortic sinus of ApoE$^{-/-}$ mice. Immunohistochemical staining for IFIT1 (green), F4/80 (macrophage marker, red), and DAPI (nucleated cells, blue) in the aortic sinus of ApoE$^{-/-}$ mice fed a high fat diet for 14 weeks. 'L' indicates vessel lumen and 'M' indicates media layer, arrows indicated cells in which colocalization was observed. Images were taken at a magnification of 20× (top panel) and 60× (bottom panel, field of view indicated with white box on 20× image). Scale bar = 50 μm. Representative images from n = 8 are shown.

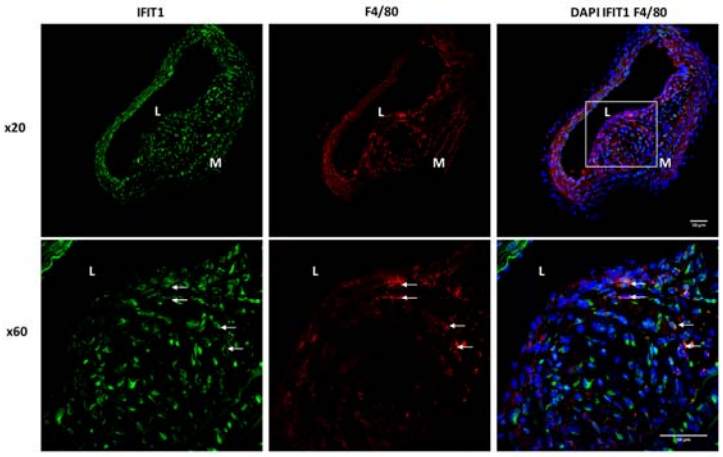


Figure 6. Expression and colocalization of IFIT1 with macrophage marker F4/80 in the brachiocephalic arteries of ApoE$^{-/-}$ mice. Immunohistochemical staining for IFIT1 (green), F4/80 (macrophage marker, red), and DAPI (nucleated cells, blue) in brachiocephalic arteries of ApoE$^{-/-}$ mice fed a high fat diet for 14 weeks. 'L' indicates vessel lumen and 'M' indicates media layer, arrows indicated cells in which colocalization was observed. Images were taken at a magnification of 20× (top panel) and 60× (bottom panel, field of view indicated with white box on 20× image). Scale bar = 50 μm. Representative images from n = 6 are shown.

inhibit viral translation by sensing viral dsRNA and, in response, synthesizing oligoadenylate dimers, which activate RNase L, leading to degradation of viral RNA and RIG-I/MAVS activation by the degradation products.[55] Finally, under M1 polarization conditions, we also observed substantial (13- to 15-fold) increases in the levels of the two myxovirus resistance proteins MX1 and MX2, which bind to components of viral capsids, thereby sequestering them or preventing their translocation into the nucleus.[55] Notably, MX1 is also a target of ISG15 conjugation.

All of the viral defense proteins discussed above have previously been identified as products of interferon stimulated

DOI: 10.1021/acs.jproteome.7b00828
J. Proteome Res. XXXX, XXX, XXX−XXX

Table 3. Proteins Whose Expression Was Enhanced by ≥4-Fold under Polarization Conditions

gene name	M1/M0	M2/M0	cluster
IFIT1	236 ± 101	1.3 ± 0.7	$M1^{Hi}$
ICAM1	57 ± 6	7.4 ± 0.6	$M1^{Hi}$
AMPD3	49 ± 18	1.7 ± 0.7	$M1^{Hi}$
IL411	434 ± 14	0.8 ± 0.2	$M1^{Hi}$
IFIT3	25.8 ± 3.6	0.5 ± 0.1	$M1^{Hi}$
IFIT2	25 ± 17	1.0 ± 0.7	$M1^{Hi}$
OASL	17.6 ± 3.3	0.9 ± 0.2	$M1^{Hi}$
MX1	15.5 ± 6.8	1.1 ± 0.3	$M1^{Hi}$
SIGLEC1	13.4 ± 5.3	0.8 ± 0.2	$M1^{Hi}$
C3	13.4 ± 1.7	1.4 ± 0.4	$M1^{Hi}$
MX2	12.6 ± 2.5	0.4 ± 0.1	$M1^{Hi}$
PARP9	11.5 ± 5.4	1.0 ± 0.5	$M1^{Hi}$
OAS3	10.4 ± 2.3	0.6 ± 0.2	$M1^{Hi}$
MARCKS	10.3 ± 3.1	1.6 ± 0.3	$M1^{Hi}$
HELZ2	9.6 ± 2.3	0.8 ± 0.2	$M1^{Hi}$
CD14	8.7 ± 0.6	0.62 ± 0.02	$M1^{Hi}$
OAS2	8.2 ± 2.3	0.7 ± 0.2	$M1^{Hi}$
MMP9	8.1 ± 2.2	1.1 ± 0.3	$M1^{Hi}$
DDX58	8.0 ± 2.2	0.6 ± 0.1	$M1^{Hi}$
ISG15	7.7 ± 2.3	0.7 ± 0.2	$M1^{Hi}$
EPB41L3	7.5 ± 0.8	2.0 ± 0.2	$M1^{Hi}$
SOD2	7.5 ± 0.8	1.1 ± 0.1	$M1^{Hi}$
OAS1	7.5 ± 2.7	1.1 ± 0.2	$M1^{Hi}$
UBE2L6	7.3 ± 1.4	1.4 ± 0.3	$M1^{Hi}$
IFIH1	7.3 ± 2.6	0.5 ± 0.3	$M1^{Hi}$
NFKB2	6.9 ± 0.7	1.0 ± 0.1	$M1^{Hi}$
HMOX1	6.3 ± 1.0	2.5 ± 0.3	$M1^{Hi}$
NCF1	6.3 ± 0.3	1.4 ± 0.2	$M1^{Hi}$
GBP1	5.8 ± 1.6	1.7 ± 0.5	$M1^{Hi}$
MGLL	5.1 ± 0.9	1.4 ± 0.1	$M1^{Hi}$
RELB	4.7 ± 0.8	0.9 ± 0.1	$M1^{Hi}$
CYP27A1	4.6 ± 0.7	1.7 ± 0.2	$M1^{Hi}$
APOE	4.2 ± 0.8	0.9 ± 0.2	$M1^{Hi}$
SAMD9L	4.1 ± 0.8	0.7 ± 0.2	$M1^{Hi}$
PLXND1	4.1 ± 1.1	0.7 ± 0.1	$M1^{Hi}$
ADA	4.0 ± 0.5	1.0 ± 0.2	$M1^{Hi}$
GBP4	4.0 ± 0.9	0.8 ± 0.2	$M1^{Hi}$
TGM2	0.3 ± 0.1	149 ± 68	$M2^{Hi}$
F13A1	0.1 ± 0.1	64.9 ± 22	$M2^{Hi}$
SORT1	0.9 ± 0.1	12.5 ± 2.1	$M2^{Hi}$
FCER2	1.0 ± 0.1	8.5 ± 2.5	$M2^{Hi}$
CD209	0.8 ± 0.1	8.1 ± 1.2	$M2^{Hi}$
GAS6	0.8 ± 0.1	5.8 ± 0.4	$M2^{Hi}$
HLA-DRB1	1.2 ± 0.2	4.0 ± 0.8	$M2^{Hi}$
FABP4	5.0 ± 1.7	3.6 ± 1.2	$M1^{Hi}/M2^{Hi}$
TLR8	4.9 ± 0.9	2.5 ± 0.4	$M1^{Hi}/M2^{Hi}$
COL6A1	2.6 ± 0.3	4.1 ± 0.2	$M1^{Hi}/M2^{Hi}$

genes (ISGs).[56] ISGs are a large family (hundreds) of proteins, mostly involved in priming cells for defense against viruses and bacteria, whose expression is upregulated by either type I interferons (including IFN-α and -β), type II interferons (IFN-γ), or type III interferons (IFN-λ1 to IFN-λ4).[55] Considering that M1 polarization was induced in the current study by treatment of the cells with IFN-γ and LPS, it is possible to rationalize the upregulation of these ISG proteins in terms of known aspects of interferon and LPS signaling. In one likely mechanism, presented in Figure 7, LPS signals via toll-like receptor 4 (TLR4), leading to activation of the transcription factor nuclear factor kappa B (NFκB) and consequent upregulation of numerous inflammatory genes (IGs). In parallel, IFN-γ signals via the IFN-γ receptor, thereby activating the Janus kinase and signal transducer and activator of transcription (JAK-STAT) pathway, driving the expression of hundreds of IFN-γ responsive genes, including transcription factors of the interferon regulatory factor 1 (IRF-1) family.[57] Importantly, transcription of IRF-1 can be upregulated synergistically in response to STAT-1 and NFκB[58] and this mechanism has been proposed to induce expression of IFN-β.[57] In turn, IFN-β is expected to be secreted and to act in an autocrine manner to further stimulate the JAK-STAT pathway via the type I IFN receptor,[55] thereby upregulating expression of additional ISGs. In support of this proposed mechanism, transcription factor enrichment analysis of the 97 proteins in the $M1^{Hi}$ cluster indicated that this cluster is highly enriched with the products of genes previously determined to be regulated by IRF-1 (31 genes; $p = 4 \times 10^{-17}$, Bonferroni correction) including IFIT1−3, RIG-I (gene DDX58), ISG15, Ubch8 (gene UBE2L6), OAS1, OAS3, and MX1. Moreover, expression of CCL5, one of four chemokines whose secretion levels were upregulated in M1 polarization conditions, has also been reported to be mediated by IRF-1.[59] In summary, the proposed mechanism accounts for many of our observations of enhanced protein expression under M1 conditions.

While the relatively simple scheme in Figure 7 may account for many of the protein enhancements observed under M1 polarization conditions, we also observed enhancements of several proteins that suggest the existence of feedback mechanisms that could further amplify the proposed pathways. These include STAT1 (a known IRF-1-regulated gene); STAT2, the canonical NFκB transcription factor NFκB1; two components of the noncanonical NFκB pathway, NFκB2 and RelB; and CD14, an LPS pattern-recognition coreceptor.[60] Although the mechanisms leading to induction of these factors remain to be determined, they are likely to amplify the initial responses to IFN-γ and LPS as well as the induced response to IFN-β.

The addition of IL-4 (M2 stimulus) to macrophages resulted in less pronounced protein changes than the addition of M1 stimuli. Nonetheless, the upregulation of several known M2 markers was observed with increases in the cell surface marker CD209,[9] the transglutaminase TGM2,[61] and proteins consistent with the pro-fibrotic functions of M2 macrophages such as TGF-β1 and COL6A1.[62] However, subsequent pathway analyses did not produce conclusive outcomes to explain the M2 induced effects from a mechanistic point of view. AP-1 transcription factor, the major enriched transcription factor in M2 macrophages, has been mainly implicated in regulating M1 polarized macrophages.[63] However, Fontana et al.[64,65] found that JUNB, a member of the AP-1 family, was involved in regulating both M1 and M2 polarized macrophages and further investigation performed by Yang et al.[66] identified that JUN transcription factor was significantly overexpressed in mouse M2 polarized macrophages. This observation is consistent with our findings, but it also highlights the need for additional studies to unravel the mechanisms underlying M2 polarization.

■ CONCLUSION

In conclusion, proteomic analysis has shown that M1 polarized macrophages express high levels of numerous antiviral proteins, predominantly under the control of the transcription factor IRF-1, whereas M2 polarization results in more modest changes

DOI: 10.1021/acs.jproteome.7b00828
J. Proteome Res. XXXX, XXX, XXX−XXX

Journal of Proteome Research Article

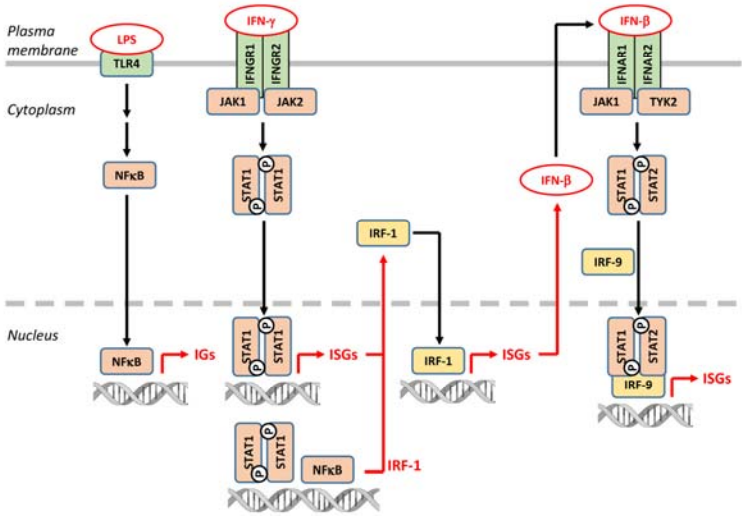


Figure 7. Proposed mechanism for induction of interferon-stimulated genes (ISGs) under M1 polarization conditions. LPS activates TLR4, stimulating expression of NFκB-regulated inflammatory genes (IGs). IFN-γ activates the JAK-STAT pathway, thereby inducing expression of ISGs as well as synergizing with NFκB to upregulate IRF-1. IRF-1 induces expression of numerous ISGs, including IFN-β, which, in turn, activates additional ISG expression via the JAK-STAT pathway and interferon regulatory factor 9 (IRF-9). Signal transduction and transport pathways are indicated by black arrows, whereas transcription and translation are indicated by red arrows.

in expression of proteins related to pro-fibrotic functions. This study has identified IFIT proteins as highly upregulated by M1 polarization. Moreover, the high expression level of IFIT1 in a subset of macrophages within atherosclerotic plaques suggests that the IFIT proteins may serve as useful markers (in combination with other proteins) of M1 macrophages in experimental animals or in human pathology applications. Future studies will be needed to determine whether the IFITs and other proteins overexpressed under M1 conditions also contribute to the pathology of atherosclerosis or other macrophage-associated diseases.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.7b00828.

CV histograms for protein intensities under each treatment condition; PCA, generated in Perseus, of significantly differentially expressed proteins; FOM for clustering analysis obtained using TM4 MeV (PDF)

Precursor/inclusion list used for Skyline; identified and quantified cellular proteins from global proteome study; cellular proteins with significant changes in expression levels upon M1 or M2 polarization; intensities of proteins identified in secretome study; proteins identified in secretome study and predicted to be secreted or potentially secreted; secreted proteins observed under only single treatment condition (XLSX)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: martin.stone@monash.edu. Phone +61-3-9902-9246.
*E-mail: barbara.kemp@monash.edu. Phone: 61-3-9905-4674.
*E-mail: okleifeld@technion.ac.il. Phone: 972-77-8871453.

ORCID

Martin J. Stone: 0000-0002-6468-4427

Author Contributions

‖These authors have contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge the facilities and the scientific and technical assistance of the Monash Biomedical Proteomics Facility and of Monash Micro Imaging, Monash University, Victoria, Australia. We also thank Dr. Ashish Banerjee for sharing the anti-NF-κB2 antibody and Prof. David Jans for sharing the anti-STAT1 antibody. N.A.B. was funded by an ARC Future Fellowship (110100223). This project was financially supported by Australian Research Council Discovery Grant No. DP130101984, ANZ Trustees Grant No. 12-3831, and Monash University Platform Access Grant No. PAG15-0091.

■ ABBREVIATIONS

M0, nonpolarized macrophages; M1, macrophages polarized to M1 phenotype; M2, macrophages polarized to M2 phenotype; ApoE$^{-/-}$, ApoE knockout; SDC, sodium deoxycholate; TCEP,

DOI: 10.1021/acs.jproteome.7b00828
J. Proteome Res. XXXX, XXX, XXX−XXX

Tris (2-carboxyethyl)phosphine hydrochloride solution; Ct, cycle threshold; DDA, data-dependent acquisition; LC−MS/MS, liquid chromatography−tandem mass spectrometry; LFQ, label-free quantification; PRM, parallel reaction monitoring; FDR, false discovery rate; ANOVA, analysis of variance; HFD, high fat diet; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology; CS, charge state; NCE, normalized collision energy

REFERENCES

(1) Wang, N.; Liang, H.; Zen, K. Molecular mechanisms that influence the macrophage m1-m2 polarization balance. *Front. Immunol.* **2014**, *5*, 614.

(2) Mantovani, A.; Garlanda, C.; Locati, M. Macrophage diversity and polarization in atherosclerosis: a question of balance. *Arterioscler., Thromb., Vasc. Biol.* **2009**, *29* (10), 1419−23.

(3) Stoger, J. L.; Gijbels, M. J.; van der Velden, S.; Manca, M.; van der Loos, C. M.; Biessen, E. A.; Daemen, M. J.; Lutgens, E.; de Winther, M. P. Distribution of macrophage polarization markers in human atherosclerosis. *Atherosclerosis* **2012**, *225* (2), 461−8.

(4) Anderson, C. F.; Gerber, J. S.; Mosser, D. M. Modulating macrophage function with IgG immune complexes. *J. Endotoxin Res.* **2002**, *8* (6), 477−81.

(5) Colin, S.; Chinetti-Gbaguidi, G.; Staels, B. Macrophage phenotypes in atherosclerosis. *Immunol. Rev.* **2014**, *262* (1), 153−66.

(6) Gordon, S.; Martinez, F. O. Alternative Activation of Macrophages: Mechanism and Functions. *Immunity* **2010**, *32* (5), 593−604.

(7) Stein, M.; Keshav, S.; Harris, N.; Gordon, S. Interleukin 4 potently enhances murine macrophage mannose receptor activity: a marker of alternative immunologic macrophage activation. *J. Exp. Med.* **1992**, *176* (1), 287−92.

(8) Raes, G.; Van den Bergh, R.; De Baetselier, P.; Ghassabeh, G. H.; Scotton, C.; Locati, M.; Mantovani, A.; Sozzani, S. Arginase-1 and Ym1 are markers for murine, but not human, alternatively activated myeloid cells. *J. Immunol.* **2005**, *174* (11), 6561.

(9) Martinez, F. O.; Gordon, S.; Locati, M.; Mantovani, A. Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: new molecules and patterns of gene expression. *J. Immunol.* **2006**, *177* (10), 7303−11.

(10) Sica, A.; Mantovani, A. Macrophage plasticity and polarization: in vivo veritas. *J. Clin. Invest.* **2012**, *122* (3), 787−95.

(11) Murray, P. J.; Allen, J. E.; Biswas, S. K.; Fisher, E. A.; Gilroy, D. W.; Goerdt, S.; Gordon, S.; Hamilton, J. A.; Ivashkiv, L. B.; Lawrence, T.; Locati, M.; Mantovani, A.; Martinez, F. O.; Mege, J. L.; Mosser, D. M.; Natoli, G.; Saeij, J. P.; Schultze, J. L.; Shirey, K. A.; Sica, A.; Suttles, J.; Udalova, I.; van Ginderachter, J. A.; Vogel, S. N.; Wynn, T. A. Macrophage activation and polarization: nomenclature and experimental guidelines. *Immunity* **2014**, *41* (1), 14−20.

(12) Brown, J.; Wallet, M. A.; Krastins, B.; Sarracino, D.; Goodenow, M. M. Proteome bioprofiles distinguish between M1 priming and activation states in human macrophages. *J. Leukocyte Biol.* **2010**, *87* (4), 655−62.

(13) Becker, L.; Liu, N. C.; Averill, M. M.; Yuan, W.; Pamir, N.; Peng, Y.; Irwin, A. D.; Fu, X.; Bornfeldt, K. E.; Heinecke, J. W. Unique proteomic signatures distinguish macrophages and dendritic cells. *PLoS One* **2012**, *7* (3), e33297.

(14) Meijer, K.; Weening, D.; de Vries, M. P.; Priebe, M. G.; Vonk, R. J.; Roelofsen, H. Quantitative proteomics analyses of activation states of human THP-1 macrophages. *J. Proteomics* **2015**, *128*, 164−72.

(15) Zhang, F.; Liu, H.; Jiang, G.; Wang, H.; Wang, X.; Wang, H.; Fang, R.; Cai, S.; Du, J. Changes in the proteomic profile during the differential polarization status of the human monocyte-derived macrophage THP-1 cell line. *Proteomics* **2015**, *15* (4), 773−86.

(16) Schmittgen, T. D.; Livak, K. J. Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.* **2008**, *3* (6), 1101−8.

(17) Masuda, T.; Tomita, M.; Ishihama, Y. Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis. *J. Proteome Res.* **2008**, *7* (2), 731−40.

(18) Batth, T. S.; Francavilla, C.; Olsen, J. V. Off-line high-pH reversed-phase fractionation for in-depth phosphoproteomics. *J. Proteome Res.* **2014**, *13* (12), 6176−86.

(19) Rappsilber, J.; Mann, M.; Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2007**, *2* (8), 1896−906.

(20) Vizcaino, J. A.; Csordas, A.; del-Toro, N.; Dianes, J. A.; Griss, J.; Lavidas, I.; Mayer, G.; Perez-Riverol, Y.; Reisinger, F.; Ternent, T.; Xu, Q. W.; Wang, R.; Hermjakob, H. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **2016**, *44* (D1), D447−56.

(21) Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26* (12), 1367−1372.

(22) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **2010**, *26* (7), 966−8.

(23) Devaux, S.; Cizkova, D.; Quanico, J.; Franck, J.; Nataf, S.; Pays, L.; Hauberg-Lotte, L.; Maass, P.; Kobarg, J. H.; Kobeissy, F.; Meriaux, C.; Wisztorski, M.; Slovinska, L.; Blasko, J.; Cigankova, V.; Fournier, I.; Salzet, M. Proteomic Analysis of the Spatio-temporal Based Molecular Kinetics of Acute Spinal Cord Injury Identifies a Time- and Segment-specific Window for Effective Tissue Repair. *Mol. Cell. Proteomics* **2016**, *15* (8), 2641−70.

(24) Howe, E. A.; Sinha, R.; Schlauch, D.; Quackenbush, J. RNA-Seq analysis in MeV. *Bioinformatics* **2011**, *27* (22), 3209−10.

(25) Soukas, A.; Cohen, P.; Socci, N. D.; Friedman, J. M. Leptin-specific patterns of gene expression in white adipose tissue. *Genes Dev.* **2000**, *14* (8), 963−80.

(26) Pathan, M.; Keerthikumar, S.; Ang, C. S.; Gangoda, L.; Quek, C. Y.; Williamson, N. A.; Mouradov, D.; Sieber, O. M.; Simpson, R. J.; Salim, A.; Bacic, A.; Hill, A. F.; Stroud, D. A.; Ryan, M. T.; Agbinya, J. I.; Mariadason, J. M.; Burgess, A. W.; Mathivanan, S. FunRich: An open access standalone functional enrichment and interaction network analysis tool. *Proteomics* **2015**, *15* (15), 2597−601.

(27) Cox, J.; Hein, M. Y.; Luber, C. A.; Paron, I.; Nagaraj, N.; Mann, M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **2014**, *13* (9), 2513−26.

(28) Karin, M.; Liu, Z.; Zandi, E. AP-1 function and regulation. *Curr. Opin. Cell Biol.* **1997**, *9* (2), 240−6.

(29) Cassol, E.; Cassetta, L.; Rizzi, C.; Alfano, M.; Poli, G. M1 and M2a polarization of human monocyte-derived macrophages inhibits HIV-1 replication by distinct mechanisms. *J. Immunol.* **2009**, *182* (10), 6237−46.

(30) Peterson, A. C.; Russell, J. D.; Bailey, D. J.; Westphall, M. S.; Coon, J. J. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol. Cell. Proteomics* **2012**, *11* (11), 1475−88.

(31) Biswas, S. K.; Mantovani, A. Macrophage plasticity and interaction with lymphocyte subsets: cancer as a paradigm. *Nat. Immunol.* **2010**, *11* (10), 889−96.

(32) Padgett, L. E.; Burg, A. R.; Lei, W.; Tse, H. M. Loss of NADPH oxidase-derived superoxide skews macrophage phenotypes to delay type 1 diabetes. *Diabetes* **2015**, *64* (3), 937−46.

(33) Leopold Wager, C. M.; Hole, C. R.; Wozniak, K. L.; Olszewski, M. A.; Mueller, M.; Wormley, F. L., Jr. STAT1 signaling within macrophages is required for antifungal activity against Cryptococcus neoformans. *Infect. Immun.* **2015**, *83* (12), 4513−27.

(34) Zhou, X.; Michal, J. J.; Zhang, L.; Ding, B.; Lunney, J. K.; Liu, B.; Jiang, Z. Interferon induced IFIT family genes in host antiviral defense. *Int. J. Biol. Sci.* **2013**, *9* (2), 200−8.

(35) Fensterl, V.; Sen, G. C. Interferon-induced Ifit proteins: their role in viral pathogenesis. *J. Virol.* **2015**, *89* (5), 2462−8.

(36) Vladimer, G. I.; Gorna, M. W.; Superti-Furga, G. IFITs: Emerging Roles as Key Anti-Viral Proteins. *Front. Immunol.* **2014**, *5*, 94.

DOI: 10.1021/acs.jproteome.7b00828
J. Proteome Res. XXXX, XXX, XXX−XXX

(37) Chinetti-Gbaguidi, G.; Colin, S.; Staels, B. Macrophage subsets in atherosclerosis. *Nat. Rev. Cardiol.* **2015**, *12* (1), 10−7.
(38) Tan, H.-Y.; Wang, N.; Li, S.; Hong, M.; Wang, X.; Feng, Y. The Reactive Oxygen Species in Macrophage Polarization: Reflecting Its Dual Role in Progression and Treatment of Human Diseases. *Oxid. Med. Cell. Longevity* **2016**, *2016*, 16.
(39) Jablonski, K. A.; Amici, S. A.; Webb, L. M.; Ruiz-Rosado, J. D. D.; Popovich, P. G.; Partida-Sanchez, S.; Guerau-de-Arellano, M. Novel Markers to Delineate Murine M1 and M2Macrophages. *PLoS One* **2015**, *10* (12), e0145342.
(40) Italiani, P.; Boraschi, D. From Monocytes to M1/M2Macrophages: Phenotypical vs. Functional Differentiation. *Front. Immunol.* **2014**, *5*, 514.
(41) Relloso, M.; Puig-Kroger, A.; Pello, O. M.; Rodriguez-Fernandez, J. L.; de la Rosa, G.; Longo, N.; Navarro, J.; Munoz-Fernandez, M. A.; Sanchez-Mateos, P.; Corbi, A. L. DC-SIGN (CD209) expression is IL-4 dependent and is negatively regulated by IFN, TGF-beta, and anti-inflammatory agents. *J. Immunol.* **2002**, *168* (6), 2634−43.
(42) Fujisaka, S.; Usui, I.; Bukhari, A.; Ikutani, M.; Oya, T.; Kanatani, Y.; Tsuneyama, K.; Nagai, Y.; Takatsu, K.; Urakaze, M.; Kobayashi, M.; Tobe, K. Regulatory Mechanisms for Adipose Tissue M1 and M2Macrophages in Diet-Induced Obese Mice. *Diabetes* **2009**, *58* (11), 2574−2582.
(43) Martinez, F. O.; Helming, L.; Milde, R.; Varin, A.; Melgert, B. N.; Draijer, C.; Thomas, B.; Fabbri, M.; Crawshaw, A.; Ho, L. P.; Ten Hacken, N. H.; Cobos Jimenez, V.; Kootstra, N. A.; Hamann, J.; Greaves, D. R.; Locati, M.; Mantovani, A.; Gordon, S. Genetic programs expressed in resting and IL-4 alternatively activated mouse and human macrophages: similarities and differences. *Blood* **2013**, *121* (9), e57−69.
(44) Arnold, C. E.; Gordon, P.; Barker, R. N.; Wilson, H. M. The activation status of human macrophages presenting antigen determines the efficiency of Th17 responses. *Immunobiology* **2015**, *220* (1), 10−9.
(45) Terenzi, F.; Saikia, P.; Sen, G. C. Interferon-inducible protein, P56, inhibits HPV DNA replication by binding to the viral protein E1. *EMBO J.* **2008**, *27* (24), 3311−21.
(46) Liu, X. Y.; Chen, W.; Wei, B.; Shan, Y. F.; Wang, C. IFN-induced TPR protein IFIT3 potentiates antiviral signaling by bridging MAVS and TBK1. *J. Immunol.* **2011**, *187* (5), 2559−68.
(47) Der, S. D.; Zhou, A.; Williams, B. R.; Silverman, R. H. Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (26), 15623−8.
(48) Sudan, B.; Wacker, M. A.; Wilson, M. E.; Graff, J. W. A Systematic Approach to Identify Markers of Distinctly Activated Human Macrophages. *Front. Immunol.* **2015**, *6*, 253.
(49) Hannon, D. B.; Thompson, J. T.; Khoo, C.; Juturu, V.; Vanden Heuvel, J. P. Effects of cranberry extracts on gene expression in THP-1 cells. *Food Sci. Nutr.* **2017**, *5* (1), 148−159.
(50) Zhang, J.; Kris-Etherton, P. M.; Thompson, J. T.; Vanden Heuvel, J. P. Effect of pistachio oil on gene expression of IFN-induced protein with tetratricopeptide repeats 2: a biomarker of inflammatory response. *Mol. Nutr. Food Res.* **2010**, *54* (Suppl 1), S83−92.
(51) Morales, D. J.; Lenschow, D. J. The antiviral activities of ISG15. *J. Mol. Biol.* **2013**, *425* (24), 4995−5008.
(52) Zhang, D.; Zhang, D. E. Interferon-stimulated gene 15 and the protein ISGylation system. *J. Interferon Cytokine Res.* **2011**, *31* (1), 119−30.
(53) Ibsen, M. S.; Gad, H. H.; Andersen, L. L.; Hornung, V.; Julkunen, I.; Sarkar, S. N.; Hartmann, R. Structural and functional analysis reveals that human OASL binds dsRNA to enhance RIG-I signaling. *Nucleic Acids Res.* **2015**, *43* (10), 5236−48.
(54) Zhu, J.; Zhang, Y.; Ghosh, A.; Cuevas, R. A.; Forero, A.; Dhar, J.; Ibsen, M. S.; Schmid-Burgk, J. L.; Schmidt, T.; Ganapathiraju, M. K.; Fujita, T.; Hartmann, R.; Barik, S.; Hornung, V.; Coyne, C. B.; Sarkar, S. N. Antiviral activity of human OASL protein is mediated by enhancing signaling of the RIG-I RNA sensor. *Immunity* **2014**, *40* (6), 936−48.
(55) Schneider, W. M.; Chevillotte, M. D.; Rice, C. M. Interferon-stimulated genes: a complex web of host defenses. *Annu. Rev. Immunol.* **2014**, *32*, 513−45.
(56) Rusinova, I.; Forster, S.; Yu, S.; Kannan, A.; Masse, M.; Cumming, H.; Chapman, R.; Hertzog, P. J. Interferome v2.0: an updated database of annotated interferon-regulated genes. *Nucleic Acids Res.* **2012**, *41*, D1040−6.
(57) Boehm, U.; Klamp, T.; Groot, M.; Howard, J. C. Cellular responses to interferon-gamma. *Annu. Rev. Immunol.* **1997**, *15*, 749−95.
(58) Pine, R. Convergence of TNFalpha and IFNgamma signalling pathways through synergistic induction of IRF-1/ISGF-2 is mediated by a composite GAS/kappaB promoter element. *Nucleic Acids Res.* **1997**, *25* (21), 4346−54.
(59) Liu, H.; Ning, H.; Men, H.; Hou, R.; Fu, M.; Zhang, H.; Liu, J. Regulation of CCL5 expression in smooth muscle cells following arterial injury. *PLoS One* **2012**, *7* (1), e30873.
(60) Kitchens, R. L. Role of CD14 in cellular recognition of bacterial lipopolysaccharides. *Chem. Immunol.* **1999**, *74*, 61−82.
(61) Beljaars, L.; Schippers, M.; Reker-Smit, C.; Martinez, F. O.; Helming, L.; Poelstra, K.; Melgert, B. N. Hepatic Localization of Macrophage Phenotypes during Fibrogenesis and Resolution of Fibrosis in Mice and Humans. *Front. Immunol.* **2014**, *5*, 430.
(62) Wynn, T. A.; Vannella, K. M. Macrophages in Tissue Repair, Regeneration, and Fibrosis. *Immunity* **2016**, *44* (3), 450−462.
(63) Juhas, U.; Ryba-Stanislawowska, M.; Szargiej, P.; Mysliwska, J. Different pathways of macrophage activation and polarization. *Postepy Hig. Med. Dosw.* **2015**, *69*, 496−502.
(64) Fontana, M. F.; Baccarella, A.; Kellar, D.; Oniskey, T. K.; Terinate, P.; Rosenberg, S. D.; Huang, E. J.; Herbert, D. R.; Kim, C. C. Myeloid expression of the AP-1 transcription factor JUNB modulates outcomes of type 1 and type 2 parasitic infections. *Parasite Immunol.* **2015**, *37* (9), 470−8.
(65) Fontana, M. F.; Baccarella, A.; Pancholi, N.; Pufall, M. A.; Herbert, D. R.; Kim, C. C. JUNB is a key transcriptional modulator of macrophage activation. *J. Immunol.* **2015**, *194* (1), 177−86.
(66) Yang, Y.; Qin, J.; Lan, L.; Li, N.; Wang, C.; He, P.; Liu, F.; Ni, H.; Wang, Y. M-CSF cooperating with NFkappaB induces macrophage transformation from M1 to M2 by upregulating c-Jun. *Cancer Biol. Ther.* **2014**, *15* (1), 99−107.

DOI: 10.1021/acs.jproteome.7b00828
J. Proteome Res. XXXX, XXX, XXX−XXX

Chapter 4. A Global Phosphoproteomic Study of Chemokine Signal Transduction Mediated by the Chemokine Receptor CCR2

4.1 Introduction

Chemokine receptors (and to a lesser extent chemokines themselves) have been the targets of numerous drug discovery programs and clinical trials. Two receptor antagonists have reached the market. Maraviroc, an antagonist of the chemokine receptor CCR5, is used as an antiviral agent in HIV infection to block viral entry into macrophages [253]. Plerixafor, an inhibitor of receptor CXCR4, blocks homing of hematopoietic stem cells to the bone marrow, thereby mobilising these cells to the bloodstream, allowing them to be collected for later transplantation [254]. However, a number of clinical trials have also been unsuccessful due to side effects (on- or off-target) or lack of efficacy. The clinical trial failures highlight the importance of better understanding the intricacies of the chemokine-receptor network.

CCR2 is the major chemokine receptor on monocytes and macrophages, which play key roles in numerous inflammatory diseases. The primary cognate ligands of CCR2 are considered to be the monocyte chemoattractant proteins (MCPs), although several other chemokines also activate CCR2. Genetic deletion of CCR2 or silencing of MCP-1 protects mice from developing atherosclerosis [90,255,256]. Moreover, in a phase II clinical trial, inhibition of CCR2 significantly reduced levels of C-reactive protein, an atherosclerosis biomarker [93]. More than 15 clinical trials have targeted CCR2 for a broad range of indications [94], including diabetes and cancer. However, no CCR2 inhibitor has reached the market to date. Therefore, there is a need to discover alternative targets such as downstream signalling effectors that are activated by MCP/CCR2 as well as possibly by other chemokine/receptor pairs.

To gain comprehensive and unbiased insights into MCP/CCR2 signal transduction networks, we analysed phosphorylation and dephosphorylation events in MCP-treated, CCR2-expressing cells using high-resolution mass spectrometry, which is often referred as phosphoproteomics. The dynamic regulation of protein phosphorylation and dephosphorylation is crucial in signal transduction [257] and mass spectrometry-based proteomics and phosphoproteomics provides an unbiased way of unveiling the network [258]. For instance, Robles et al [259] characterised a highly complex circadian signal transduction network with mass spectrometry-based phosphoproteomics. Similarly, Biggelaar et al. [260] performed a phosphoproteomic study and gained an in-depth understanding about thrombin signalling. However, so far, there are not many

global unbiased studies characterising downstream signalling resulting from chemokine receptor activation. The only three publications describing phosphoproteomics of chemokine receptor signalling all focus on the homeostatic receptor CXCR4. Yi et al [261] performed a systematic quantitative phosphoproteomic study on breast cancer stem cells and identified many previously unknown phosphorylation sites. Based on these data, a CXCL12/CXCR4 signalling network was reconstructed, which is involved a wide range of biological processes and lay the groundwork for future signalling studies and therapeutic target investigations. Wojcechowoskyj et al. [262] used a SILAC-based approach to investigate the effect of CXCL12/CXCR4 signalling on HIV infection and found a list of proteins that could be potential targets for developing HIV therapeutics. Finally, O'Hayre et al. [263] studied CXCL12/CXCR4 signal transduction in chronic lymphocytic leukemia and found with programmed cell death protein 4 and heat shock protein 27 two major players of potential therapeutic interest. All of these studies provide a precedent to demonstrate that proteomic studies of chemokine receptor signalling can lead to discovery of new pathways and mechanisms. Therefore, considering such studies had not yet been performed for any pro-inflammatory chemokine receptors and the importance of CCR2 in atherosclerosis and other diseases, it was important to characterise CCR2-related signal transduction.

An additional motivation to characterise CCR2 signalling pathways more thoroughly came from a study that our laboratory recently published in *Science Signaling* [235]. In this study, we showed that the ligands MCP-2 and MCP-3, are partial agonists of CCR2, i.e. they evoke a submaximal response of the same pathways relative to the full agonist MCP-1, and we identified structural features of both the chemokines and the receptor that mediate partial versus full agonism. Partial agonism is the ability of some agonists to evoke a lower response than others upon activation of the same receptor. This is a broadly recognised pharmacological phenomenon and can be most simply explained by different agonists binding with different strengths to the activated versus inactivated states of the receptor. To further elaborate the observed effects from partial agonism at a global level and to investigate the mechanisms that could potentially explain partial agonism, a global quantitative proteomic or phosphoproteomic study provides a good option.

As described in Chapter 1, quantification of changes in peptide or protein levels is benefiting from the astounding power of modern mass spectrometers and high-

throughput protein identification or quantification can be easily accomplished by tandem mass spectrometry [264]. Shotgun or bottom-up proteomics is a well-known tandem mass spectrometry technique in which the mass spectrometer operates in "discovery mode" to sequence as many peptides as possible. This is also termed data-dependent acquisition (DDA). In DDA, a fast MS1 scan will determine the intensity (peak area) of all peptides and the most abundant peptides will be subjected to fragmentation and sequencing. However, this procedure is known to introduce an inherent bias towards the more intense proteins, while ignoring less abundant ones. Although this issue can be reduced by extensive pre-fractionation techniques, quantification on MS1 signal intensities is relatively inaccurate and irreproducible. In addition, in order to quantitate on fractionated samples, relatively complicated computation algorithms are typically applied, which are known to be error-prone.

Recently, due to the improvement of scan speeds of mass spectrometers, data-independent acquisition (DIA) has provided a better option. In DIA, as implied by its name, the quantification of proteins is not based on the initial mass spectral (MS1) intensity. Instead, all available precursor ions in a certain m/z range are co-fragmented and MS/MS (MS2) spectra are recorded for quantification. Thus, the DIA method reduces bias towards highly abundant proteins and improves reproducibility and accuracy of protein quantification.

In this chapter, we applied the DIA methodology to quantify the levels of proteins and phosphopeptides in CCR2-expressing cells treated with vehicle (as a negative control), the full agonist MCP-1 or the partial agonist MCP-3. By quantitating phosphorylation changes induced by MCP-1 or MCP-3, we have: (1) characterised the common downstream signal transduction pathways resulting from CCR2 activation by these chemokines; and (2) compared the differences between the phosphorylation changes induced by MCP-1 and MCP-3 to gain insights into the outcomes of full versus partial agonism.

4.2 Experimental Design

4.2.1 Rationale of Using the HEK293 Cell Line

As discussed in Chapter 1, CCR2 is a seven-transmembrane receptor mainly expressed in monocytic cells. In order to provide a fully physiologically relevant study of CCR2 signalling, it would be ideal to use human primary cells or cell lines that

endogenously express CCR2 such as THP-1 cells (human monocytes) or RAW 264.7 cells (mouse monocytes). However, these cell lines also express other chemokine receptors such as CCR1, CCR3 and CCR5. Considering the promiscuity of chemokine-receptor interactions, the presence of these additional receptors would greatly complicate the interpretation of signalling data. In the future, CCR2-selective experiments could potentially be performed with these cell lines by knocking out other receptors with the CRISPR/Cas9 genome editing technique or by using a combination of inhibitors targeting different receptors. However, for the purpose of this thesis, to investigate and compare the signal transduction induced via the receptor CCR2 and simplify the already complex chemokine-receptor interaction system, we chose to express CCR2 in the model cell line HEK293 that does not endogenously express any of the chemokine receptors. This cell line has been used to exogenously express and study chemokine receptors in many studies, including in our lab [235,265]. Specifically, in the study described here, we used the stable FlpIn-CCR2-HEK293 cell line, in which expression of CCR2 is induced by the addition of tetracycline to the culture media. We confirmed the expression of CCR2 in this system by both Western blotting and mass spectrometry (Section 4.3).

4.2.2 Data-Independent Acquisition Workflow

Due to the highly complex MS2 spectra in DIA data, accurate assignment of spectra to peptides is very challenging. This is an intrinsic feature of the DIA acquisition mode due to the wide range m/z window, resulting in co-fragmentation of all the available precursor ions. To overcome this challenge, it is necessary to generate a spectral library [266], which contains the unique information for each precursor/peptide, such as retention time, charge state and fragment ions ranking. An unknown DIA spectrum is then assigned to a pre-defined precursor by comparing with the spectral library to determine the best match [266].

We used the workflow shown in Figure 4.1, which represents a well-accepted DIA workflow but incorporates phosphopeptide identification and quantification. The

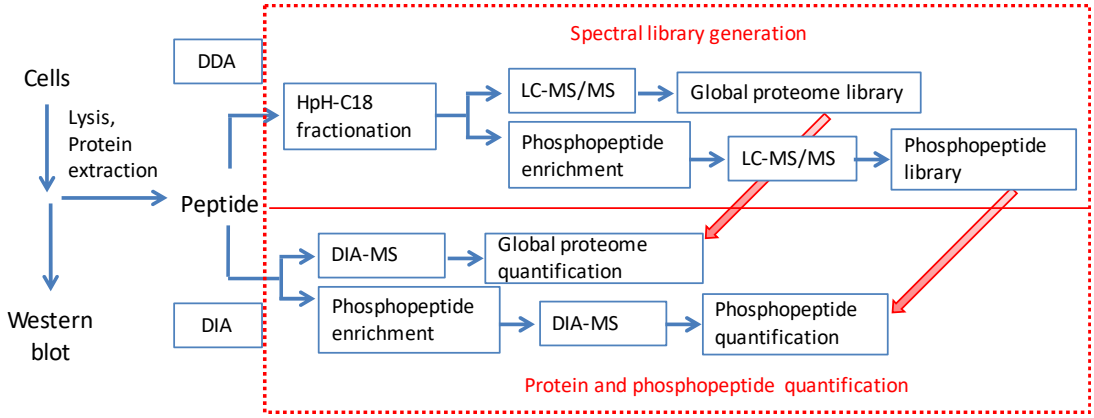


Figure 4.1. The workflow of phosphopeptide/protein spectral library generation and quantification.

top half presents the procedures for spectral library generation. Briefly, the protein was extracted from lysed cells and digested into peptides. Then it was fractionated with off-line basic pH reversed phase C18 (HpH) column, a small aliquot was injected directly into the mass spectrometer for protein identification, and the remainder was enriched with TiO_2 beads and injected into the mass spectrometer for phosphopeptide identification. Thus, spectral libraries containing data for proteins or phosphopeptides were created to enable quantification. For quantification by DIA-MS, shown in the bottom half of the figure, a small aliquot of the digested peptide was directly detected with the mass spectrometer and the rest was used for phosphopeptide enrichment and detection. Identification was achieved by using the spectral information from the spectral libraries and quantification was achieved from the MS2 intensities of the identified peptides and phosphopeptides.

4.3 Confirmation of CCR2 Expression

In order to confirm the expression of CCR2 in FlpIn-CCR2-HEK293 cells, we performed an immunoprecipitation study; untransfected HEK293 (without the cMyc-His-CCR2 construct) cells were used as a negative control. Both FlpIn-CCR2-HEK293 and untransfected cells were grown in the presence or absence of tetracycline, lysed in buffer containing detergent to help solubilise the membrane receptor CCR2, and immunoprecipitated using an anti-cMyc antibody, directed against the affinity tag of the CCR2 construct. The affinity purified sample was separated into two parts. One was used for Western blotting, while the other was used for mass spectrometric analysis.

The Western blot result is shown in Figure 4.2. Irrespective of whether the cells were treated with or without tetracycline, none of the untransfected HEK293 cells expressed CCR2 protein. In contrast, in FlpIn-CCR2-HEK293 cells, there are 3 bands (A-C) detected ranging from around 40 kDa to 100 kDa, when induced with tetracycline (lanes 7 and 8), but no band was observed in uninduced cells. The molecular weight of band C is slightly higher than the theoretical molecular weight (41 kDa) and there are 2 additional bands with even higher molecular weight. We suspect that band A represents CCR2 with a low level of post-translational glycosylation, band B has a higher level of glycosylation and band C may be a dimeric form, which has been observed previously for CXCR4 [267]. The Western blot results clearly show that only the transfected, tetracycline-treated cell line expressed CCR2.

To determine whether we could observe the expressed CCR2 by mass spectrometry, the affinity-purified samples were resolved by SDS-PAGE, excised, subjected to in-gel digestion and analysed by mass spectrometry. As shown in Figure 4.3, two of the CCR2 peptides (peptide 315-321 and 338-360) were observed in this MS experiment with most of their MS2 fragments (b and y ions) identified. They are only present in the FlpIn-CCR2-HEK293 cells when stimulated with tetracycline, but not in the empty HEK293 cells or in the FlpIn-CCR2-HEK293 cells without tetracycline treatment. Notably, both the identified peptides are from the C-terminus of CCR2. This was not surprising because CCR2 is a 7-transmembrane protein with seven highly hydrophobic transmembrane regions, which are known to be difficult to detect by mass spectrometry. Furthermore, we were not able to identify any glycosylation sites in the CCR2 protein as this modification occurs primarily in the N-terminal region [268], which was not observed in this study. Taken together, the results of the Western blotting and mass spectrometric analyses indicate that the CCR2 receptor is successfully expressed in the FlpIn-CCR2-HEK293 cells and the cell line can be applied for further analysis.

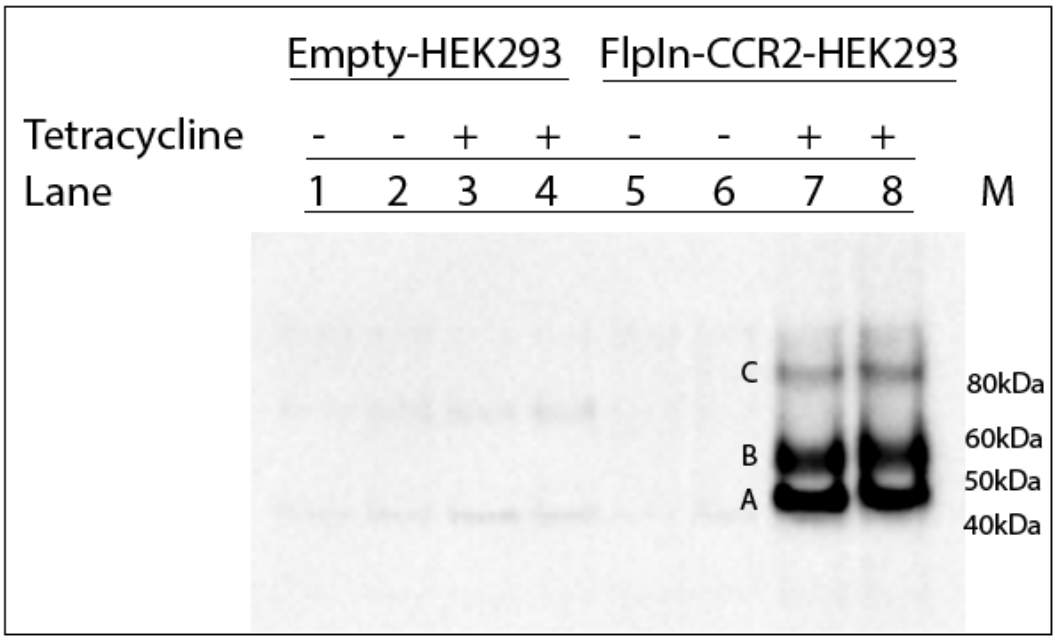


Figure 4.2. Western blot analysis of duplicate affinity immunoprecipitation experiments. Samples in lanes 1-4 are from empty HEK293 cells treated (or not treated) with tetracycline. Samples in lanes 5-8 are from FlpIn-CCR2-HEK293 cells treated (or not treated) with tetracycline. M indicates the positions of bands in the protein standard.

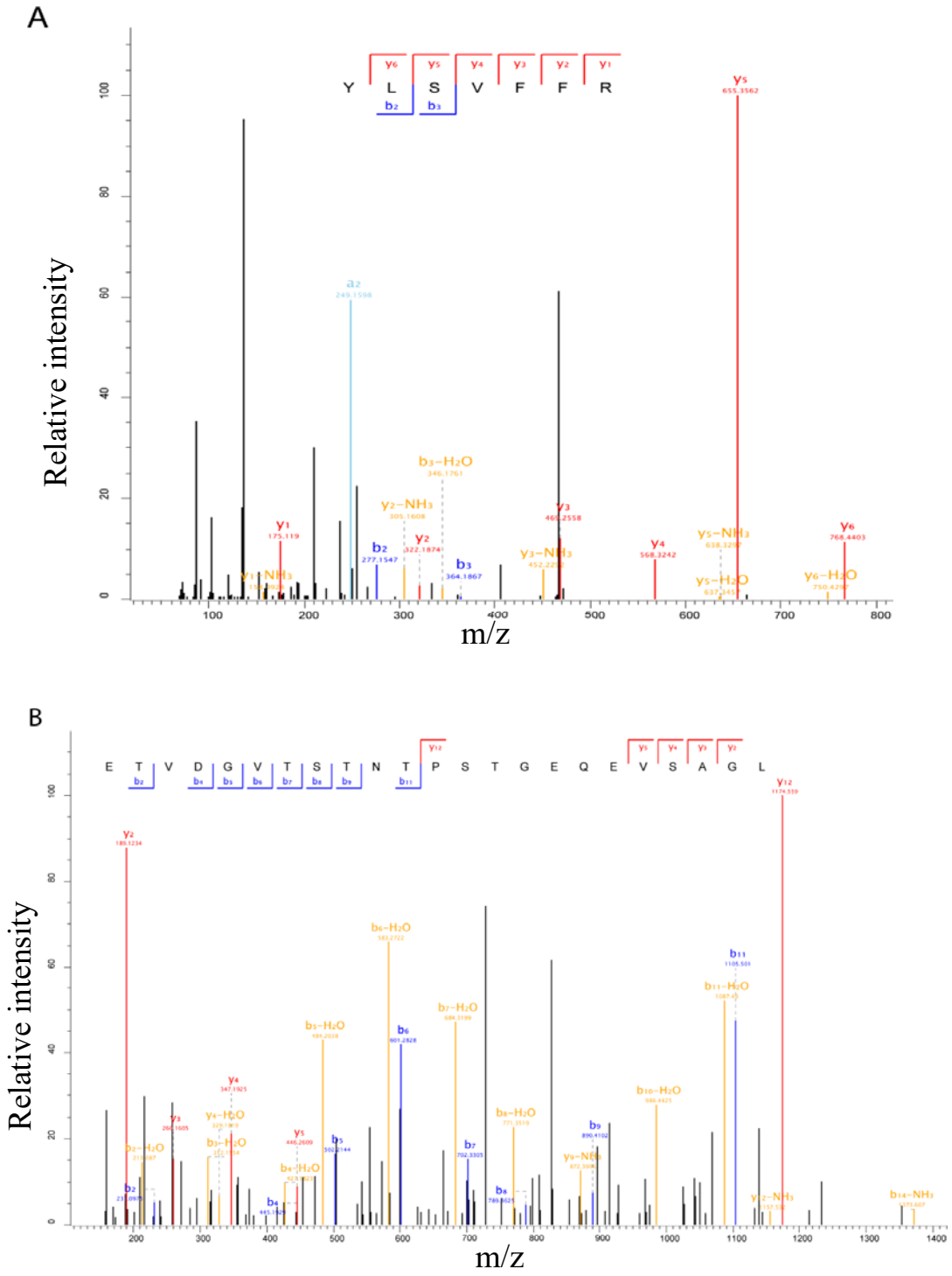


Figure 4.3. MS2 spectra of (A) peptide 315-321 (m/z 466.26 at charge 2) and (B) peptide 338-360 (m/z 1140.53 at charge 2) from CCR2. Coloured peaks represent b (blue), b-H_20 (orange) and y (red) fragment ions. The identified peptide and fragmentation pattern are shown schematically at the top of the figure.

4.4 Global Quantitative Proteomics and Phosphoproteomics

4.4.1 Generation of Spectral Libraries

Since precursor/peptide identification in DIA depends heavily on the spectral library, the library should fulfil stringent criteria: a) it should contain all precursors/fragments from all different conditions; b) the spectra should be of high quality; and c) the mass spectrometer has to be operated under the same settings (in particular the collision energy has to be constant) and preferably the same instrument should be used for the whole experiment. These conditions are required to avoid sample bias and incorrect quantification.

In order to generate a spectral library containing as many high-quality spectra for each precursor/peptide as possible, the digested peptides from each condition are fractionated to maximise the number and confidence of MS2 spectra (Figure 4.1). It has been reported that almost a full core human proteome can be identified through extensive fractionation [104,269]. Our aim was to achieve as deep a human proteome as possible to avoid protein identification bias. In this study, we applied off-line basic pH reversed phase C18 fractionation technology to fractionate one replicate for each condition, i.e. treatment with each chemokine and vehicle-treated control; meanwhile, a concatenated way of fraction combing approach was used as it limits the overlap of peptides between fractions and provides better sequence coverage for the protein [141]. A small part of each fraction was retained for global proteome identification, while the rest was enriched with TiO_2 beads for phosphopeptide identification.

Our global proteome spectral library contained 132,430 unique precursor ions from 100,204 peptides, of which 95,574 are proteotypic peptides. These peptides are derived from 8,058 proteins, which is about 80% of the human core proteome [104,269]. Our phosphoproteome spectral library consisted of 19,069 unique stripped peptide sequences from 34,599 precursors. More importantly, 29,269 were unique phosphopeptide isoforms, derived from 5,118 proteins.

4.4.2 In-depth Quantification of the Proteome and Phosphoproteome

For DIA quantification of the proteome and phosphoproteome, the sample preparation technique was similar to that of the spectral library generation, except that the samples were not fractionated. After digestion of the proteins with trypsin, the generated peptide mixture was directly desalted, a small percentage (10%) of each

sample was used for global proteome quantification, and the rest was enriched with TiO_2 beads for phosphopeptide quantification.

Using a DIA approach as shown in the bottom half of Figure 4.1, we managed to quantify 7,143 proteins across the various conditions. More than 6,000 of these proteins were quantified in all replicates. In addition, we quantified 14,145 phosphopeptides, among which 10,311 were singly phosphorylated, 3,464 doubly phosphorylated and 370 multiply phosphorylated. Not surprisingly, 86% of the identified phosphorylated sites were phosphoserines, 13.7% phosphothreonines and 0.3% phosphotyrosines, which is consistent with the phospho-amino acid distribution observed by different detection methods and other global phosphoproteomic studies [270,271]. Among these 14,145 quantified phosphopeptides, more than 12,000 phosphopeptides were quantified all replicates.

4.4.3 Reproducibility and Accuracy of DIA-MS Data

To determine whether the DIA-MS quantification was reproducible and accurate across the biological replicates, we performed Pearson correlation (R) and coefficient of variation (CV) analyses. As shown in Figure 4.4 and Appendix III, Table S4.1, Table S4.2, Pearson correlation coefficients of higher than 0.9 were observed across all biological replicates for both the proteome (Figure 4.4A and C) and the phosphoproteome (Figure 4.4B and D) quantification, indicating high quantitative reproducibility. Furthermore, analysis of the CVs confirmed a high quantitative accuracy of our DIA-MS with median CVs smaller than 15% (Figure 4.5A) or 16% (Figure 4.5B) for proteome and phosphoproteome quantification, respectively.

To the best of our knowledge, the study described here is the first study in the field of chemokine receptor signalling that exploits ‘data-independent acquisition mass spectrometry’ to quantify phosphoproteomes at a global level. From this study, we identified more than 7,000 proteins and 14,000 phosphopeptides in a single repeat without any fractionation steps required. More importantly, we quantified more than 6,000 proteins and 10,000 phosphopeptides in a single experiment with high reproducibility and accuracy, which has not been reported in any publication based on DDA quantification. It would be possible to further increase the coverage of the phosphoproteome by fractionation of phosphopeptides prior to MS analysis, as demonstrated by Sharma et al [134]. However, fractionation would require a larger

amount of material and more instrument time. As described below, the extent of the phosphoproteome observed in this study enabled us to identify a major signalling network. Therefore, this study helps to establish DIA-MS as an excellent approach to characterise protein phosphorylation and signalling networks.

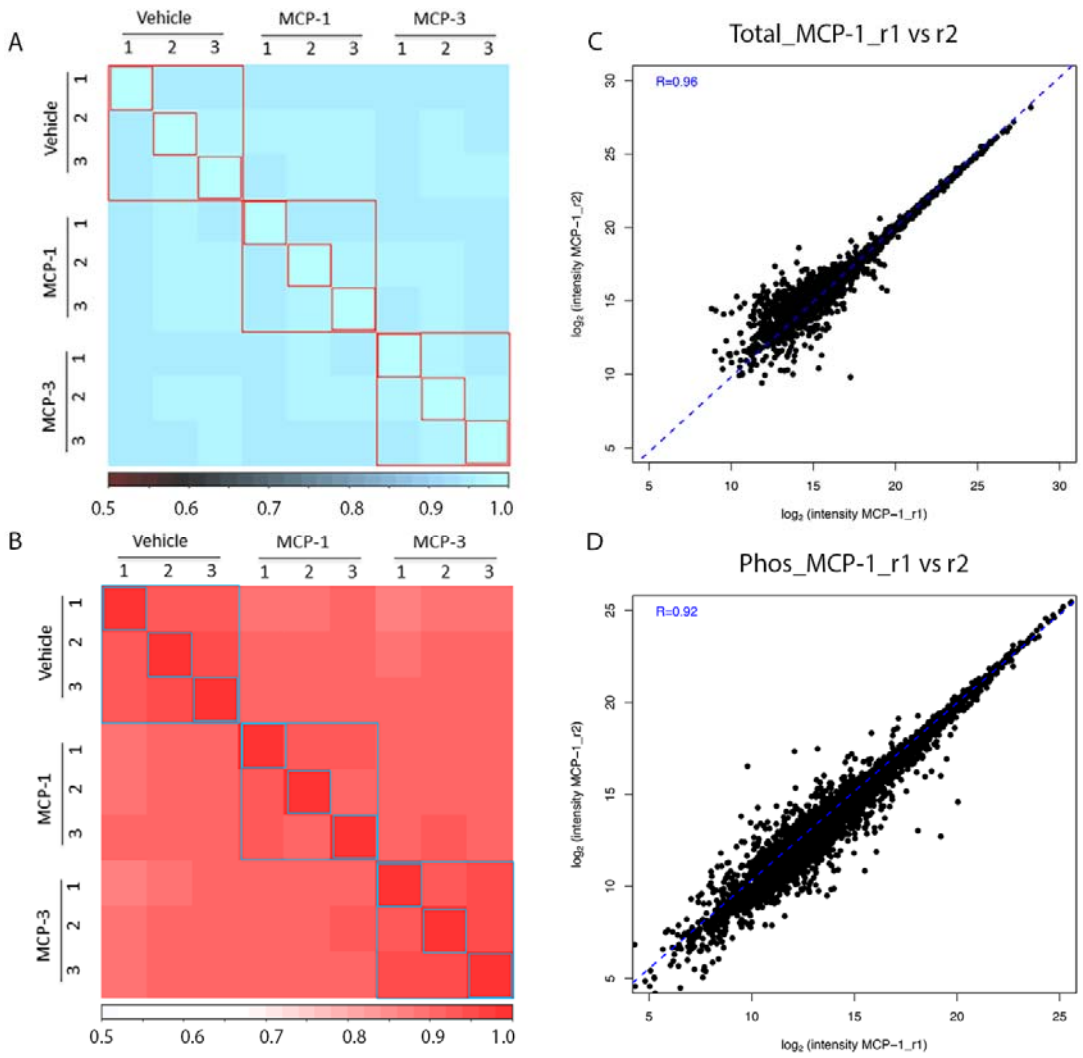
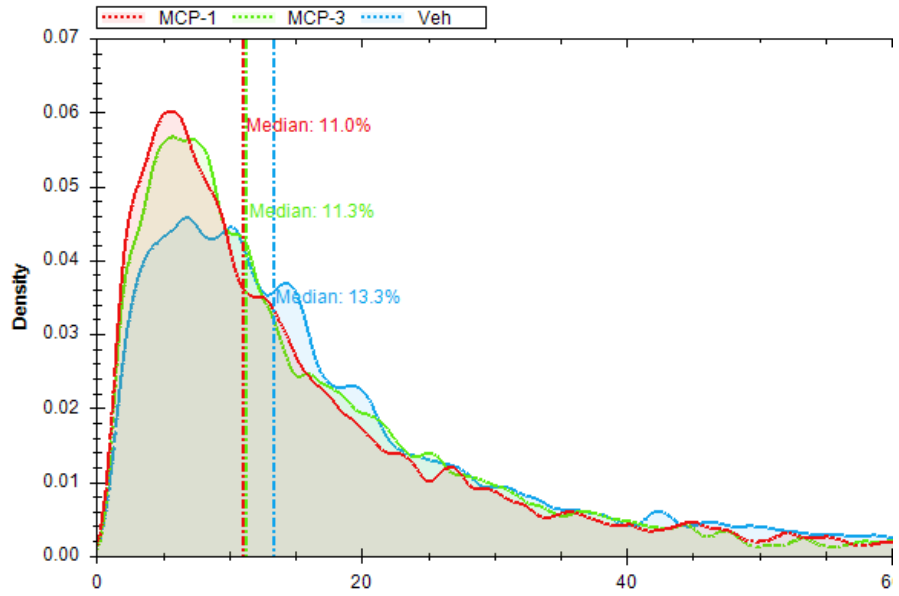


Figure 4.4. Reproducibility of proteome quantification shown by analysis of Pearson correlation coefficients. The heat maps (A and B) and scatter plots (C and D) are shown to present the Pearson correlation coefficients (R) between protein (A and C) or phosophopeptide (B and D) intensities in independent experiments. A and C: the big square indicates 3 independent repeats for the same treatment and the small red square is the same sample (R=1); the colour scale is shown for R values of 0.5-1.0. C and D: scatter plot between MCP-1 repeat 1 and repeat 2 for total proteins and phosphopeptides; the Pearson correlation is indicated on the top right with blue value (R); each black dot indicates a single protein or phosphopeptide; x and y axis is $\log_2$ transformed intensity.

A



B

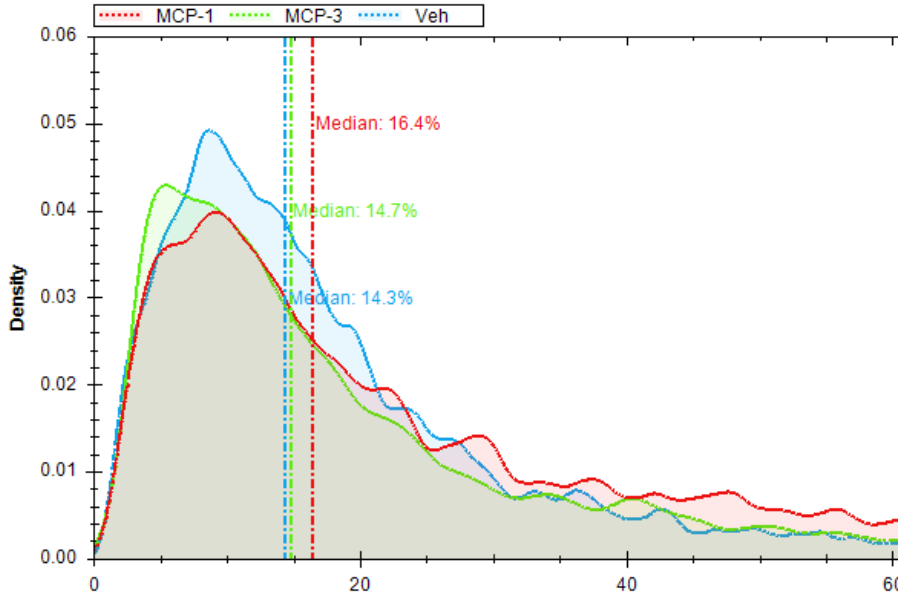


Figure 4.5. Coefficient of variation (CV) plots for protein and phosphopeptide quantification. Shown are the distributions of CV values for (A) proteome quantification and (B) phosphoproteome quantification. The median CV values for MCP-1, MCP-3 and vehicle treatment are indicated by red, green and blue dotted lines and labels, respectively.

4.5 Changes in Protein and Phosphopeptide Levels upon Chemokine Treatment

The protein and phosphopeptide intensities determined by DIA-MS were used to identify proteins and phosphopeptides that increased or decreased significantly upon treatment with MCP-1 or MCP-3, in comparison to vehicle-treated control cells. These changes are represented by the volcano plots shown for MCP-1 treatment (Figure 4.6) and MCP-3 treatment (Appendix III, Figure S4.1). In this section, we focus on the data for MCP-1. As can be seen from the few black dots (proteins) in the pink (significantly increased) and light blue (significantly decreased) areas in Figure 4.6A, only a few proteins were significantly up- or down-regulated. In contrast, substantially more changes can be observed in phosphopeptide levels (Figure 4.6B). This indicates that the changed abundance of phosphopeptides is due to phosphorylation or dephosphorylation of the peptide instead of changes in the abundance of proteins. The result is consistent with many other reports that a relatively short treatment of cells would not cause significant changes in protein expression level but can induce robust phosphorylation or dephosphorylation [260,272].

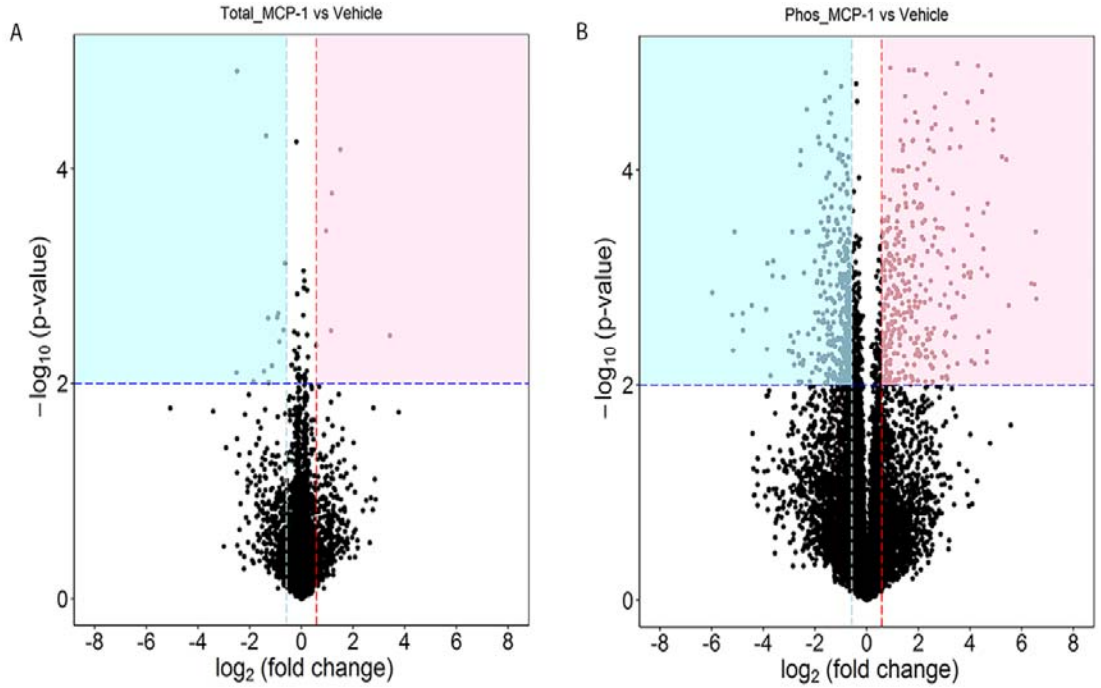


Figure 4.6. Volcano plots showing the effects of MCP-1 treatment; A: comparison between vehicle-treated and MCP-1-activated cells in protein quantification; B: comparison between vehicle-treated and MCP-1-activated cells in phosphopeptide quantification. The x-axis is $\log_2$-transformed fold change and the y-axis is $-\log_{10}$ (p value); each black dot indicates one protein or phosphopeptide. The horizontal blue dash line indicates p = 0.01; the vertical grey dashed line (left) indicates a 1.5-fold decrease and the vertical red dash line (right) indicates a 1.5-fold increase. The blue and pink areas highlight proteins or phosphopeptides that are significantly decreased or increased, respectively.

Based on the criteria of p value ≤ 0.01 and fold change (increase or decrease) ≥1.5, we identified 49 proteins and 1,420 phosphopeptides (spread across 832 phosphoproteins) with significantly increased or decreased abundance in either one or both chemokine treatment(s), relative to vehicle control. Only two of the significantly changed phosphopeptides were derived from significantly changed proteins and were excluded from subsequent analyses.

Although modern proteomic methods enable confident peptide identification, it is generally not possible to unambiguously assign every single phosphate group to a specific residue. This is because many peptides contain multiple potentially phosphorylatable residues (serine, threonine and tyrosine) and the decisive fragment ion may be missing from the mass spectra. The solution to this problem, generally agreed by the phosphoproteomic community, is to calculate the localization probability of phosphorylation on each residue. There are different computing algorithms, but a localization probability higher than 0.75 (1 means unambiguous assignment) is the well-accepted threshold and has been applied in most phosphoproteomic studies. As indicated before, DIA-MS is used for quantification, instead of identification, so we incorporated the localization probability from the spectral library to the significantly up or down regulated phosphopeptides. Out of the 1,418 phosphopeptides, 1,168 phosphopeptides had a localization probability higher than 0.75 and only these were retained for further analysis.

4.6 MCP/CCR2 Network Characterization

4.6.1 An Overrepresented Network in MCP-Activated CCR2 Cells

In order to extend our understanding of CCR2 signalling pathways, we submitted the genes, whose phosphosites were significantly regulated by both MCP-1 and MCP-3, to DAVID [251,252] to identify overrepresented KEGG pathways; KEGG is a database containing a collection of manually assigned pathways. Table 4.1 lists the most significantly enriched pathways (FDR < 5%) in addition to the MAPK signalling pathway, which is included with a FDR of 9.8% and 18 genes identified. Not surprisingly, the majority of these pathways have been previously implicated in MCP signalling or other chemokine signalling. For example, Jimenez-Sainz et al [273] used a combination of pharmacological inhibitors and mutations of downstream signalling proteins to characterise a MAPK signalling cascade in MCP-1 activated THP-1 and

HEK293 cells. Phosphoproteins within this network are also enriched in our phosphoproteomic study but our study extends previous knowledge of this pathway by identifying many additional modified phosphosites, as shown in Appendix III, Table S4.3. In addition, MCP-1 activation on CCR2 has been linked to mTOR and ErbB signalling pathway from previous studies [274,275]. The most significantly enriched pathways, adherens junction and tight junction, are highly connected to regulation of the actin cytoskeleton, the function of which is involved in cell attachment and movement. This indicates that many of the phosphorylation changes observed in this study are related to cell movement, which makes perfect sense considering that the major role of chemokines and their receptors is guiding cell transmigration.

Table 4.1. Enriched KEGG pathways in MCP-1 and MCP-3 activated CCR2 cells.

Term	No. of genes	FDR (Benjamin-Hochberg corrected)
Adherens junction	16	3.40E-06
Tight junction	20	4.10E-05
mTOR signalling pathway	11	2.30E-03
ErbB signalling pathway	13	3.20E-03
Insulin signalling pathway	15	0.01
AMPK signalling pathway	13	0.028
Regulation of actin cytoskeleton	18	0.03
MAPK signalling pathway	18	0.098

4.6.2 MCP/CCR2 Network Mapping

Based on the identified core pathways mentioned above, we used our comprehensive dataset to expand and manually curate these pathways and to develop an extended MCP/CCR2 network, which shown in Figure 4.7. The fold change of all mapped proteins and phosphosites is listed in Appendix III, Table S4.3.

The manually curated network indicates activation of several canonical pathways, including JNK, PI3k/Akt and MAPK/ERK. Consistent with these results, others have observed phosphorylation of some proteins within these pathways. For example, activation of JNK and MAPK/ERK pathways through phosphorylation has been detected in MCP-1 activated vascular endothelial cells with Western blotting [276] and this was confirmed by many other studies [277,278]. The activation of PI3K/Akt pathway by MCP-1 has been shown by Arefieva et al. [279], Kleibeulker et al. [280] and Jin et al. [281], with a maximum response achieved with 3-10 minutes' treatment, the activation of which was confirmed by a series of subsequent studies with PI3K inhibitor, LY294002. Previous findings add confidence to our manually curated network, but more information is provided from this study, including detailed phosphosites and quantitate changes in the levels of phosphorylation at these sites, all from a single experiment.

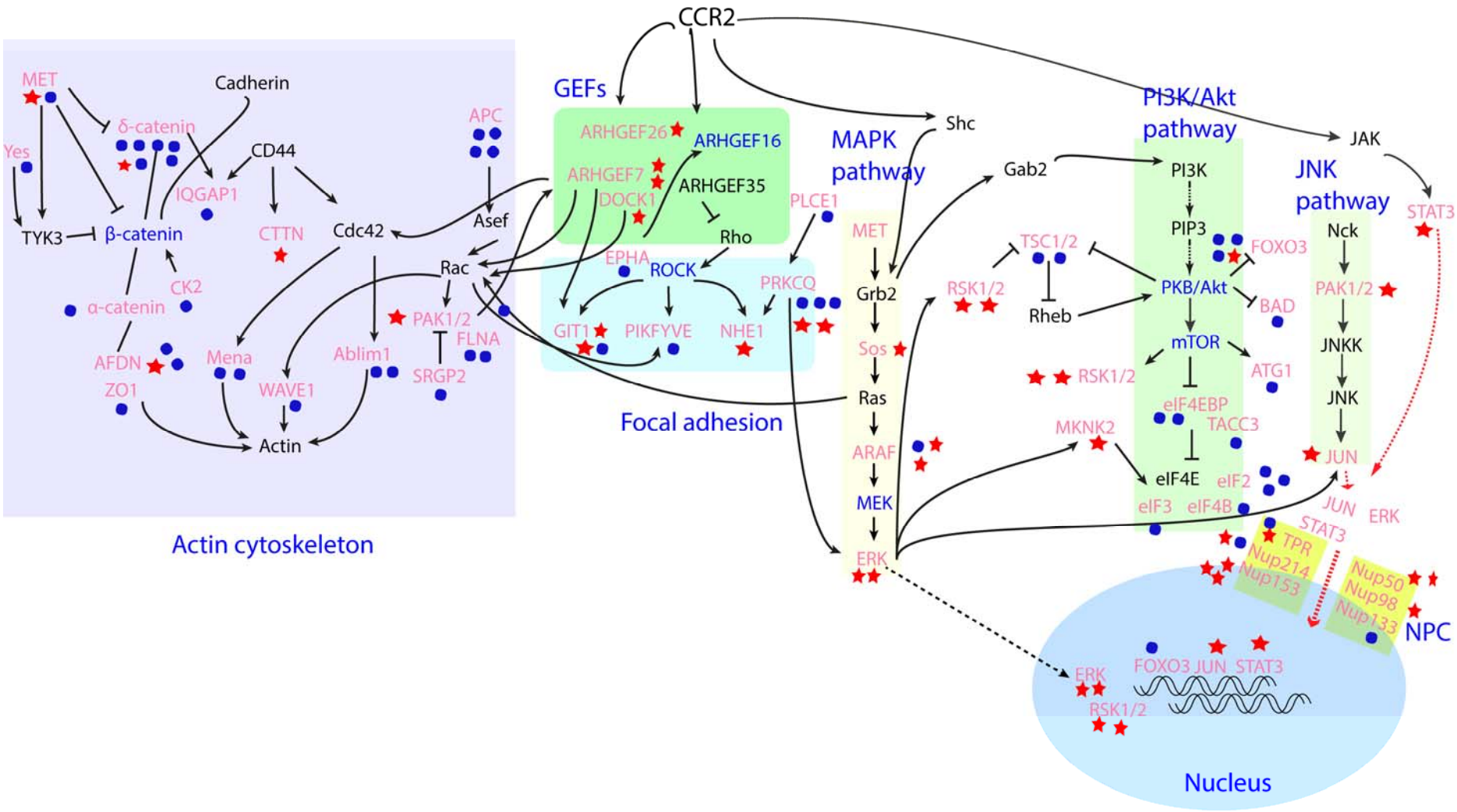

CCR2
GEFs
MAPK pathway
PI3K/Akt pathway
JNK pathway
Actin cytoskeleton
Focal adhesion
Nucleus
NPC
MET
Cadherin
δ-catenin
Yes
CD44
IQGAP1
TYK3
β-catenin
CTTN
Cdc42
APC
Asef
Rac
CK2
α-catenin
PAK1/2
FLNA
AFDN
Mena
Ablim1
SRGP2
ZO1
WAVE1
Actin
ARHGEF26
ARHGEF16
ARHGEF7
ARHGEF35
DOCK1
Rho
EPHA
ROCK
PRKCQ
GIT1
PIKFYVE
NHE1
PLCE1
Shc
Gab2
MET
Grb2
Sos
Ras
ARAF
MEK
ERK
PI3K
PIP3
TSC1/2
FOXO3
RSK1/2
Rheb
PKB/Akt
BAD
mTOR
ATG1
RSK1/2
eIF4EBP
TACC3
MKNK2
eIF4E
eIF2
eIF3
eIF4B
JAK
STAT3
Nck
PAK1/2
JNKK
JNK
JUN
JUN
ERK
STAT3
TPR
Nup214
Nup153
Nup50
Nup98
Nup133
ERK
FOXO3
JUN
STAT3
RSK1/2

Figure 4.7. A manually curated MCP/CCR2 signalling pathway. Each pathway is shaded and coloured. A protein name in pink indicates that one or more phosphorylation site(s) belonging to this protein have been found to be significantly changed in our data set. A protein name in blue indicates that the protein has been identified in our study, but not shown as significantly changes. A protein name in black indicates that the protein has not been identified in our study. Each blue circle indicates a particular phosphorylation site whose phosphorylation level has been observed to be decreased. Each red star indicates a particular phosphorylation site whose phosphorylation level has been observed to be increased.

In addition to these canonical pathways, we observed changes in several other (but connected) signalling networks. For example, we detected phosphorylation of six nucleoporins, which are parts of the nuclear pore complex (NPC; light yellow boxes in Fig 4.7), the function of which is to serve as a gateway for transporting molecules, such as DNA, mRNA or proteins, in and out of the nucleus [282]. This is the first time that phosphorylation of NPCs has been identified in the MCP/CCR2 signalling network. As descripted in Chapter 1, activation of CCR2 signalling can increase the activity of several transcription factors including JUN, FOXO3 and STAT3, which leads to their translocation between the cytoplasm and the nucleus. The connection between CCR2 signalling and NPC is unknown, but it is known that increased transcriptional activity can inhibit cell mitosis and promote protein translation [283]; therefore, we suspect that activation of CCR2 signalling can inhibit cell mitosis and promote protein biosynthesis, which is needed for cell migration, the major function of CCR2.

Moreover, we have mapped many phosphosites/proteins that are in the actin cytoskeletal organisation pathway. For example, Rho guanine nucleotide exchange factors (ARHGEFs or GEFs, green box in Figure 4.7) belong to the guanine nucleotide exchange factors (GEF) family, which can act as key controllers, limiting the rates of downstream signal transduction [284]. The GEF family includes more than 150 protein members that bind to GTPases, facilitating or promoting exchange of GDP for GTP to switch on downstream signal transduction. They are involved in a wide range of biological functions, such as cell proliferation, differentiation migration and cytoskeletal function.

ARHGEFs, the specific group of GEFs identified here, are mainly responsible for regulation of cell migration and cytoskeletal organisation and can activate Rho GTPases, such as CDC42 and Rac [285]. The mechanism of GEF regulation of GTPase is not fully understood, but it has been proposed that phosphorylation of auto-inhibited GEF can activate GEF, which can then bind to GTPases and propagate signal transduction [286]. Our data indicated increased phosphorylation of ARHGEF7 and ARKGEF26, which are known to regulate the activity of cdc42 and Rac [287], thereby triggering signalling events in actin-related processes (light purple box Figure 4.7). Moreover, although we did not observe phosphorylation changes of cdc42 and Rac directly, we observed such changes for several of their downstream targets in the actin cytoskeleton network. Thus,

it appears likely that ARHGEFs represent an important link between CCR2 activation and cytoskeletal reorganisation.

We also observed phosphorylation changes of some other signalling proteins, such as WAVE1 and CTTN, whose phosphorylation changes have been previously observed in actin cytoskeletal organisation. WAVE1, Wiskott-Aldrich syndrome protein family member 1, is a downstream effector of Rac. Kim et al. identified that decreased phosphorylation at Ser-310 of WAVE1 was accompanied by increased actin polymerisation *in vitro* [288]. A further WAVE1 knockdown study confirmed the result with observation of more filopodia in primary hippocampal neurons. Decreased phosphorylation on the same site in this protein is also detected in our phosphoproteomic data in both MCP-1 and MCP-3 activated CCR2 cells (Figure 4.8 A), which implies actin may also polymerise in response to these chemokines.

Cortactin (CTTN) is a Src kinase substrate or nucleation promoting factor, which is involved in actin cytoskeletal organisation, formation of lamellipodia and cell migration [289]. In contrast to decreased phosphorylation of WAVE1 at Ser-310, Janjanam et al. found that increased phosphorylation at Ser-405 and Ser-418 of CTTN promotes actin polymerisation [290]. They observed that CTTN extracted from MCP-1-treated human aortic smooth muscle cell (HASMC) caused actin polymerisation; a mutagenesis study further confirmed the requirement of phosphorylated Ser-405 and Ser-418 in HASMC migration [290]. The observed increase in phosphorylation of CTTN, as shown in Figure 4.8 B, adds confidence to the hypothesis that MCP activation of CCR2 may induce actin polymerisation in the HEK293 cells used in this study; this hypothesis is further explored in the next section.

Other than above mentioned proteins, we observed phosphorylation changes for several other proteins/phosphosites that have previously been connected to chemokine/receptor signalling pathways. Yes kinase has been related to CXCL12 induction [291], phosphorylated Ser-2152 in FLNA has been connected to CCR2 recycling [292] and PAK has been reported to be regulated by CCL5 [293]. The consistency of our proposed network with these previous findings add confidence to our manually curated network. Appendix III, Table S4.3 summarises the protein or phosphorylation residues found in this study that have been connected to MCPs, CCR2, other chemokines or other chemokine receptors. It is clearly shown that most of the

proteins and phosphosites were not previously known to be involved. Thus, this first report of a comprehensive phosphorylation and signalling network induced by MCP activation of CCR2 greatly extends our knowledge and understanding of MCP/CCR2 signalling mechanisms.

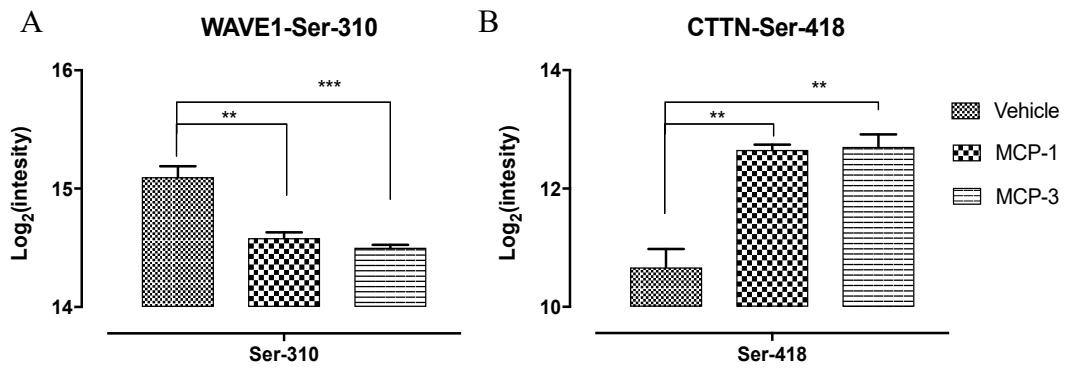


Figure 4.8. Dephosphorylation of WAVE1 and phosphorylation of CTTN upon chemokine treatment. Shown are the abundance of phosphorylated peptides in phosphoproteomic data for cells treated with vehicle, MCP-1 or MCP-3. (A) Levels of phosphorylated Ser-310 in WAVE1. (B) Levels of phosphorylated Ser-418 in CTTN. All data are presented as mean ± SEM, n=3. * $P < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.00001$ vs time 0min (one-way ANOVA followed by Dunnett's post hoc test)

4.6.3 Validating the Characterised MCP/CCR2 Network

To further increase the confidence in our phosphoproteomic data, we compared the phosphorylation observed by mass spectrometric analysis to Western blot results for the mitogen-activated protein kinases or extracellular signal-regulated kinases (MAPK/ERK). Phosphorylation of MAPK/ERK is a well-established downstream consequence of chemokine receptor activation, typically detected by Western blots or AlphaScreen analyses [273,276]. As clearly shown in Figure 4.9A, our phosphoproteomic dataset showed a significant increase in the abundance of phosphorylation in MCP-1 and MCP-3 treated cells, relative to vehicle treatment, both for singly phosphorylated and doubly-phosphorylated forms of ERK-1/2. The increased ERK phosphorylation from mass spectrometric analysis is confirmed by three independent Western blot analyses (Figure 4.9B.); a representative blot is shown in Figure 4.9C.

Additionally, to validate the phosphorylation changes in the actin cytoskeletal organisation pathway, we used both Western blotting and confocal microscopy. In Western blotting, due to the limited availability of phosphorylation-specific antibodies, we were only able to test for three of the phosphorylation sites identified in the proteomic data set: CK2-Ser-209, PAK2-Ser-197 and FLNA-Ser-2152. Among these, only anti-CK2-Ser-209 worked well, while the other two were not specific in protein binding. CK2 is casein kinase 2, which is a serine/threonine kinase and can phosphorylate a large number of proteins, regulating a wide range of cellular functions, including cell growth, proliferation and apoptosis [294]. The regulation of CK2 can be through phosphorylation and dephosphorylation. CK2 can be activated by decreased phosphorylation [295], which is detected in our phosphoproteomic data and supported by quantitative Western blotting (Figure 4.10). These results indicate the activation of CK2 in this pathway.

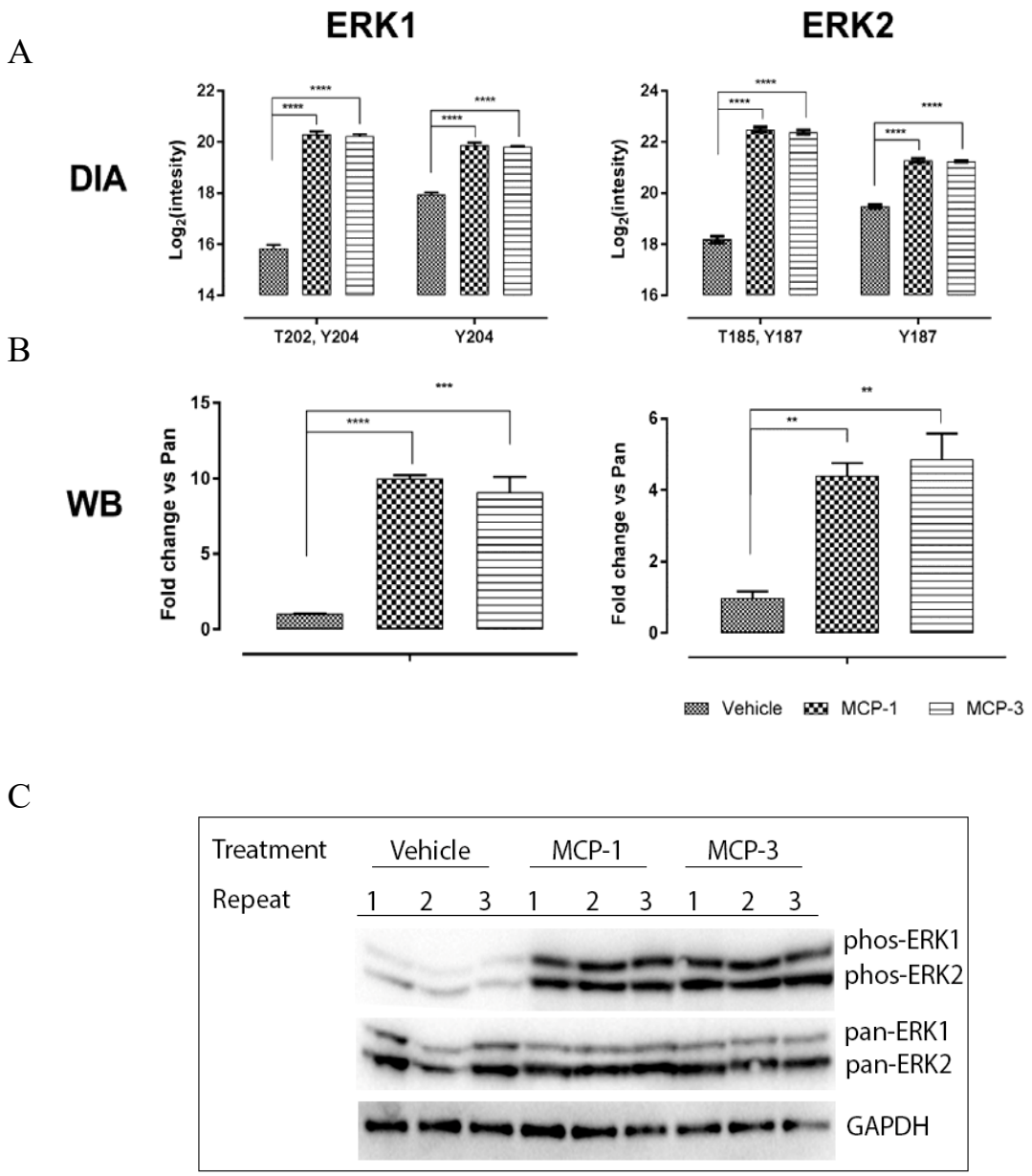


Figure 4.9. Comparison of DIA-MS and Western blotting quantification for phosphorylation of ERK1 and ERK2. (A) Quantification by DIA-MS, showing single phosphorylation (Tyr-204 (Y204) on ERK1; Tyr-187 (Y187) on ERK1) and double phosphorylation (Thr-202 (T202) and Tyr-204 (Y204) on ERK1; Thr-185 (T185) and Tyr-187 (Y187) on ERK2), (B, C) Quantification by Western blot; the phosphorylation-specific antibody (phos-ERK1/2) detects both singly and doubly phosphorylated forms. All data are presented as mean ± SEM, n=3. * $P < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.00001$ vs time 0min (one-way ANOVA followed by Dunnett's post hoc test)

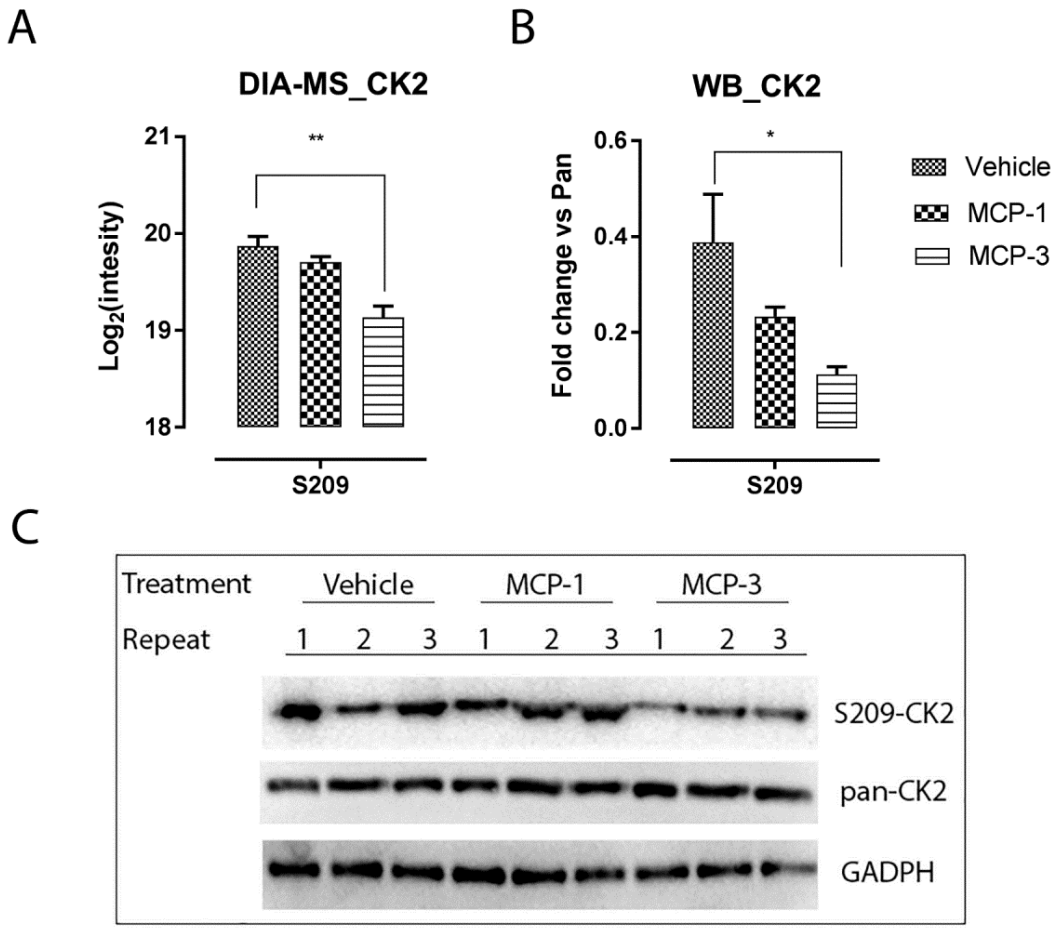


Figure 4.10. Dephosphorylation of CK2 upon chemokine treatment. (A) DIA-MS and (B) Western blotting quantification. (C) Western blot data. All data are presented as mean ± SEM, n=3. * $P < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.00001$ vs time 0min (one-way ANOVA followed by Dunnett's post hoc test)

The outcome of activating the actin cytoskeletal organisation pathway is to induce actin polymerisation, leading to cellular shape changes or movement. Therefore, to provide a validation of our phosphoproteomic data at the cellular level, we stained MCP-1-treated cells with phalloidin, a reagent specific for filamentous (polymerised) actin (F-actin) and examined them by confocal microscopy. Considering that the kinetics of actin polymerisation are slower than the initial phosphorylation changes [296-298], it is not surprising that short treatments (<30 min) did not induce noticeable changes in cellular morphology (data not shown). However, in cells treated for one hour we observed more lamellipodia in the cells treated with MCP-1 than in the vehicle-treated control, as indicated by the arrows in Figure 4.11. Lamellipodia are membrane protrusions at the edges of the cells. They are made up of actin filaments and function as part of the driving force for cell migration [299]. This exactly matches the biological consequence of the actin related pathway and the functions of chemokines and their receptors. Due to limited time, this experiment has only been performed in a qualitative manner at present but the data support our findings that chemokine activation of CCR2 stimulates components of the actin cytoskeletal organisation pathway.

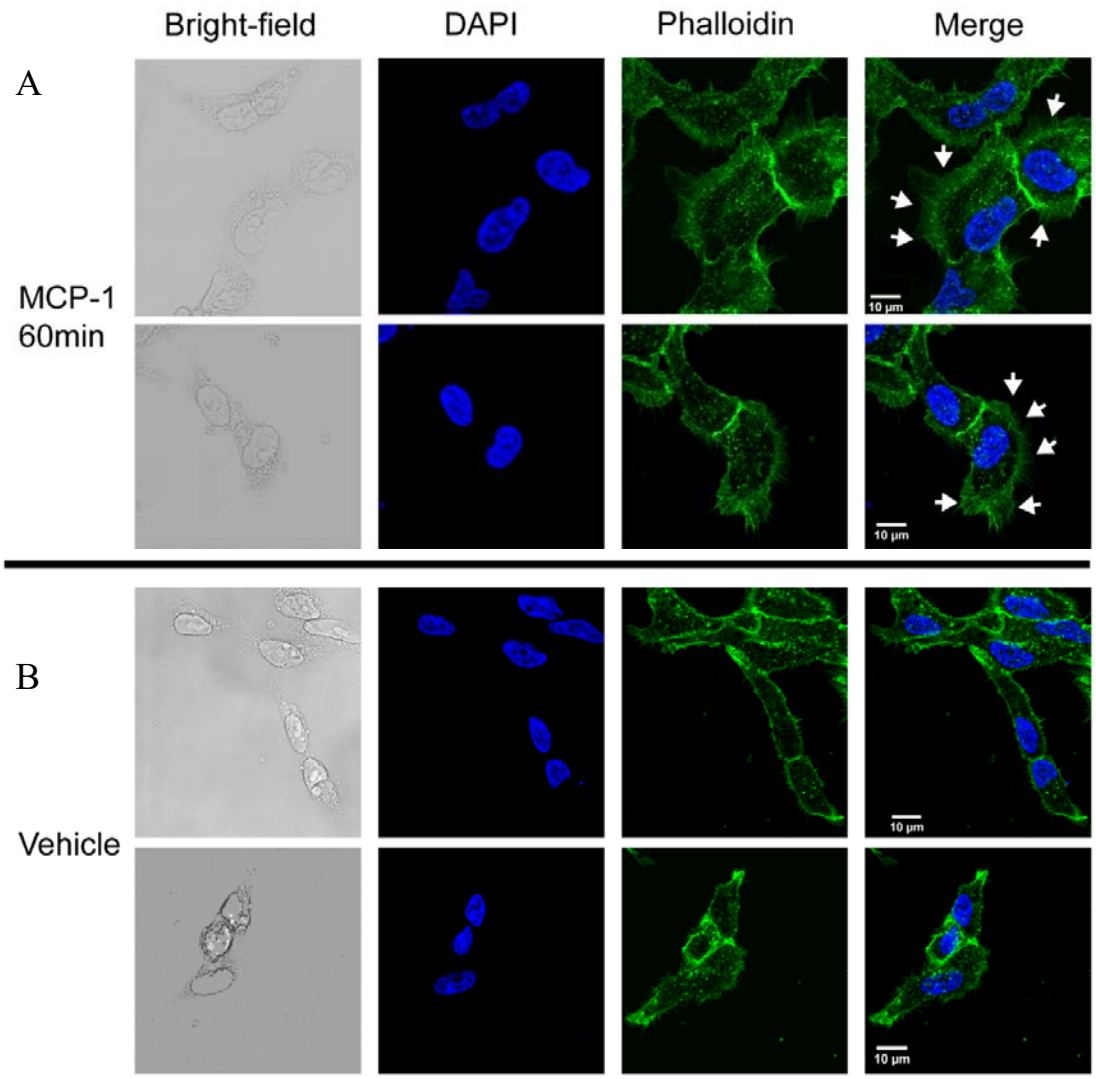


Figure 4.11. **Evaluation of the effects of MCP-1 on the actin cytoskeleton.** FlpIn-CCR2-HEK293 cells were treated with (A) MCP-1 for 1 h or (B) vehicle then immediately fixed and stained with DAPI (blue, to stain nuclei) and phalloidin-Alexa Fluor 647 (green, to stain F-actin) and imaged by bright field (left panels) or confocal microscopy. The images were merged in ImageJ. A duplicate experiment is shown here. Arrows indicate lamellipodia.

4.7 Comparison between MCP-1 and MCP-3 Activated Pathways

As discussed in the introduction, we recently showed that in contrast to MCP-1, MCP-3 can induce only partial activation of the same pathway upon activation of CCR2. Therefore, MCP-1 was defined as a full agonist, while MCP-3 was identified as a partial agonist. To gain a more comprehensive understanding of this phenomenon, in this section we compared the phosphorylation changes induced by MCP-1 or MCP-3 that lead to changes in cellular pathways.

4.7.1 Cluster Analysis of Significantly Regulated Phosphopeptides

Firstly, to separate the significantly regulated phosphopeptides that were exclusively induced by MCP-1 or MCP-3, we allocated the 1,168 significantly regulated phosphopeptides from section 4.5 into six distinct clusters based on either up or down-regulation in one or more conditions. Figure of merit analyses (Appendix III, Figure S4.2) indicated that there was no benefit in assigning more than 6 clusters. As shown in Figure 4.12, these 6 distinct clusters are: Cluster 1, MCP-1_Increase (only increased in MCP-1 activated cells); Cluster 2, MCP-3_Increase (only increased in MCP-3 activated cells); Cluster 3, MCP-1_MCP-3_Increase (increased in both MCP-1 and MCP-3 activated cells); Cluster 4, MCP-1_Decrease (only decreased in MCP-1 activated cells); Cluster 5, MCP-3_Decrease (only decreased in MCP-3 activated cells); and Cluster 6:, MCP-1_MCP-3_Decrease (decreased in both MCP-1 and MCP-3 activated cells). The number of significantly regulated phosphopeptides is as indicated at the top of each panel of the figure.

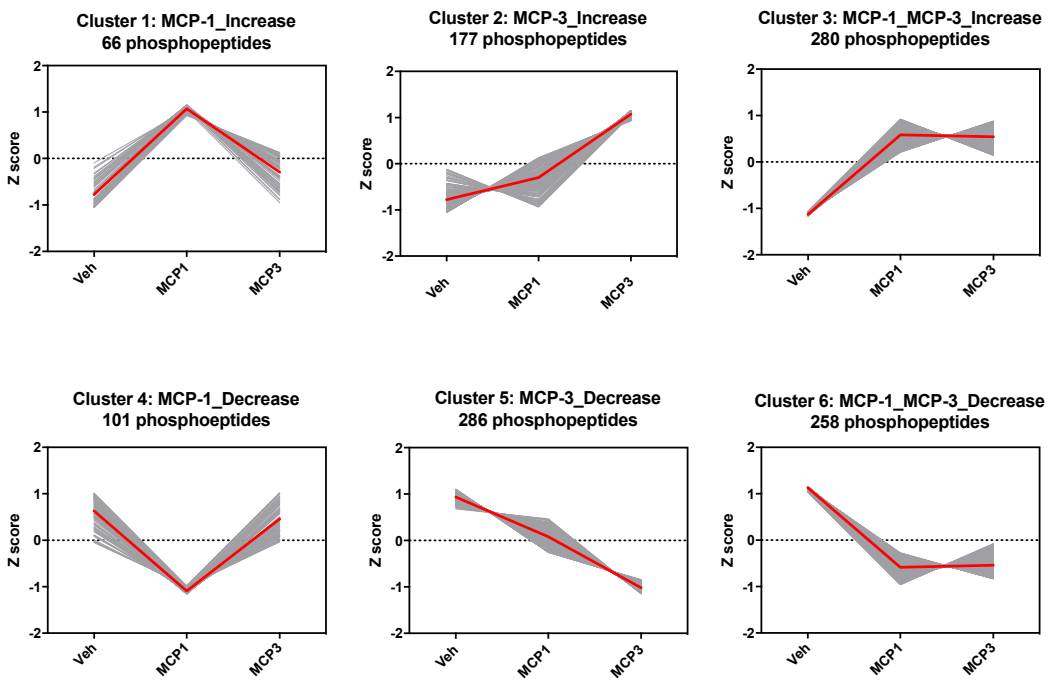


Figure 4.12. Clustering analysis of significantly regulated phosphopeptides. The 1,168 phosphopeptides were grouped into 6 clusters using the program TM4:MeV (http://www.tm4.org/mev.html). The diagrams show the z-score transformed value of the intensity of each individual phosphopeptide (grey lines). The thick red line indicates the average value within a cluster and the number of phosphopeptides within the cluster is given under the cluster name at the top of each panel.

4.7.2 Comparison of Overrepresented Phosphorylation Motifs and Upstream Kinases

4.7.2.1 Overrepresented Phosphorylation Motifs

Peptide (de)phosphorylation is always controlled by upstream kinases or phosphatases, which can recognize a certain pattern of amino acid residues. This pattern is defined as a phosphorylation motif. All the phosphopeptides from each cluster were submitted to MotifX [247,300] to look for overrepresented sequence motifs and enriched motifs (relative to random expectation) were plotted based on motif score.

As shown in Figure 4.13A, all the clusters consistently show enrichment in the **s**P motif, a proline-directed kinases motif; thus, there is no substantial difference in this motif between treatment conditions. When comparing Cluster 1 and 2 or 4 and 5, there are few differences observed, although **s**xxxxxxP and Rxx**s** motifs are slightly increased in Cluster 2 and Rxx**s** and **s**D are slightly enriched in Cluster 5. On the other hand, for Cluster 3 (phosphopeptides increased by both MCP-1 and MCP-3) a distinct Px**s**P motif is present, complemented by **t**P and Rxx**s** motifs, whereas for Cluster 6 (phosphopeptides decreased by both MCP-1 and MCP-3) the **s**PxxR motif is clearly overrepresented, complemented by **t**P and **s**xxxL motifs. Thus, the overall conclusion from this motif analysis is that there are only minor differences between the motifs phosphorylated upon treatment with MCP-1 and MCP-3. Instead, the two chemokines both caused net phosphorylation of Px**s**P sites and net dephosphorylation of **s**PxxR sites.

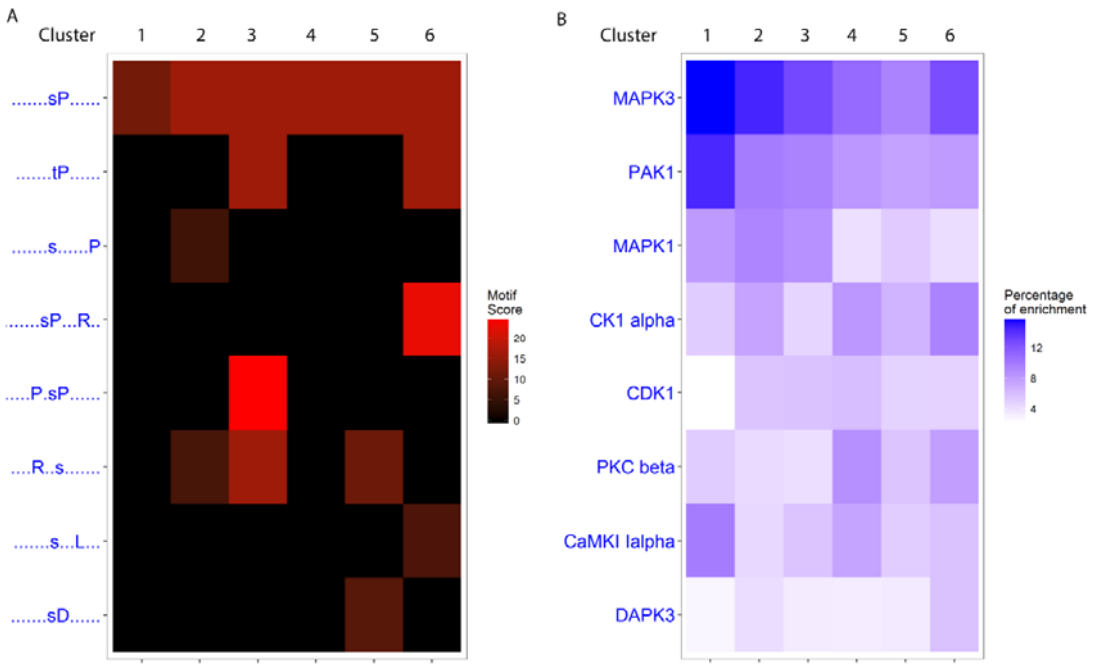


Figure 4.13. Enriched phosphorylation motifs and predicted upstream kinases for each cluster. (A) Phosphorylation motifs obtained by submitting the phosphorylation sites identified in each cluster to MotifX and (B) predicted upstream kinases obtained by NetworKin. Amino acids are denoted by s (phosphoserine), t (phosphothreonine), P (proline), K (lysine), R(arginine), D (aspartic acid) and L (leucine). Kinase abbreviations are: MAPK1/3, Mitogen-activated protein kinase 3; PAK1, Serine/threonine-protein kinase PAK 1; CK1 alpha, Casein kinase I isoform alpha; CDK1, Cyclin-dependent kinase 1; PKC beta, Protein kinase C beta type; CAMKII alpha, Calcium/calmodulin-dependent protein kinase type II subunit alpha; DAPK3, Death-associated protein kinase 3

4.7.2.2 Overrepresented Upstream Kinases

To predict the upstream kinases that are responsible for the changes of abundance of phosphopeptides and retrieve more information regarding MCP-1 and MCP-3 activated signal transduction, all the phosphosites within each cluster were submitted to NetworKIN 3.0 [301].

As presented in Figure 4.13B, similar percentages of MAPK3 substrates were enriched in Cluster 1 and 2. The substrates of most other kinase had only low levels of enrichment in all clusters. However, PAK1 and CaMKIIα substrates were substantially more highly enriched in the cluster of MCP-1 increased phosphorylation (Cluster 1) than in other clusters. PAK1 is a serine/threonine-protein kinase, the activity of which is regulated through binding to CDC42, Rac1 [302] and CaMKII. Notably, the function of CaMKII is to transduce Ca^{2+} signals and CaMKII is activated by flux of free Ca^{2+}, a well-established consequence of chemokine receptor activation [303]. The activation of PAK1 and CaMKII by MCP-1 are supported by many studies showing that MCP-1 can activate CDC42 and Rac1 [304,305]. Moreover, the increased activation of these kinases by MCP-1, in comparison to MCP-1, is consistent with our previous observation that MCP-1 is a stronger (higher efficacy) agonist than MCP-3 for CCR2 activation.

4.7.3 Gene Ontology and Cellular Pathway Comparison

To obtain further insights into the biological pathways activated by each individual chemokine, we submitted the corresponding proteins in Clusters 1, 2, 4 and 5 to DAVID [251,252] for gene ontology and cellular KEGG pathway analysis. Table 4.2 summarizes the most significantly enriched pathways from each cluster. For clusters 1 and 2, there are no significantly enriched pathways; the FDR is much higher than 5% for the most significant pathways. The same two pathways are significantly overrepresented in clusters 4 and 5. Moreover, a similar finding is obtained when analysing the molecular functions of their involved genes (Table 4.3 and 4.4). In general, there are no substantial differences in the percentage of genes classified into any of the molecular function between each pair of comparison (Cluster 1 vs Cluster 2 and Cluster 4 vs Cluster 5). In summary, these results indicate that MCP-1 and MCP-3 largely activate the same biological pathways.

4.7.4 Expression Level Comparison of Phosphopeptides Mapped in the Network

Lastly, we compared the level of the phosphopeptides that are mapped in our manually curated network (Figure 4.7) after MCP-1 versus MCP-3 treatment. The correlation between phosphopeptide intensity changes observed for MCP-1 and MCP-3 is presented in Figure 4.13. Although there is a small number of outliers, the overall pattern between MCP-1 and MCP-3 is quite similar with a squared correlation coefficient of 0.72. This correlation supports the above conclusions that the two chemokines have very similar effects on the phosphorylation and dephosphorylation of cellular proteins.

Table 4.2. Enriched KEGG pathway from each cluster

Clusters	KEGG term	No. of gene	FDR
1	Axon guidance	3	1
2	RNA transport	5	0.71
4	Adherens junction	5	0.04
	RNA transport	6	0.075
5	RNA transport	10	0.015
	Adherens junction	7	0.008

Table 4.3. Enriched molecular function within clusters 1 and 2 [1]

	% of genes in:	
GO_Molecular function	Cluster 1	Cluster 2
protein binding	66.1	74.3
ATP binding	19.4	15.4
poly(A) RNA binding	17.7	24.3
GTPase activator activity	9.7	5.9
protein serine/threonine kinase activity	9.7	7.4
protein kinase activity	8.1	6.6

[1]percentage of genes in the cluster that have been annotated as having the listed molecular function

Table 4.4. Enriched molecular function within clusters 3 and 4 [1]

	% of genes in:	
GO_Molecular function	Cluster 4	Cluster 5
protein binding	67.7	70.4
ATP binding	17.2	12.8
poly(A) RNA binding	16.1	17.7
cell-cell adhesion	11.8	9.3

[1] percentage of genes in the cluster that have been annotated as having the listed molecular function

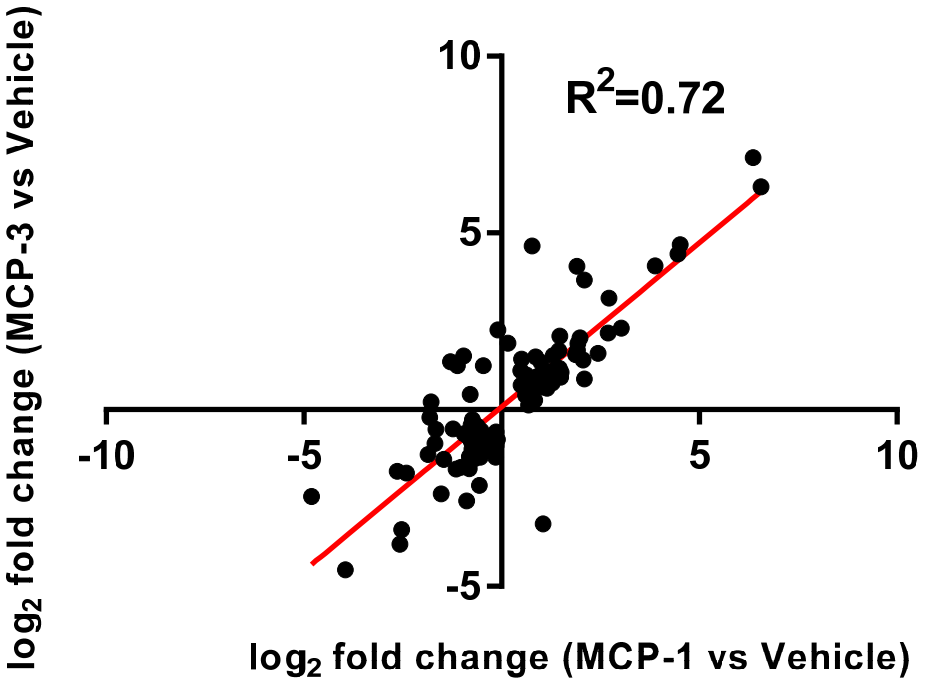


Figure 4.13. Correlation of phosphorylation changes induced by MCP-1 and MCP-3. Shown is a scatter plot correlating the log_2-transformed phosphorylation changes induced by MCP-1 and MCP-3 (each relative to vehicle control) for all 57 proteins that are mapped in the MCP/CCR2 network (Fig 4.7). The red line is the best fit and the squared correlation coefficient ($R^{2)}$ is shown.

4.7.5 MCP-1 vs MCP-3: A Summary

In conclusion, we have performed a systematic comparison between MCP-1 and MCP-3 activated signal transduction cascades by analysing their accompanied phosphorylation motifs, predicted upstream kinases, gene ontology-molecular functions, KEGG pathways and phosphopeptide levels. Although some different phosphosites were identified in each treatment and although small differences were observed when analysing the predicted upstream kinases, no major differences have emerged from all other analyses. We therefore conclude that the difference between MCP-1 and MCP-3 activated signal transduction is difficult to rationalise at the phosphopeptide level.

The current phosphoproteomic experiments were performed using 300 nM concentrations of each chemokine. Under these conditions, we have previously observed stronger activation of the β-arrestin2 pathway, but no differences in the activation of other pathways studied, such as ERK1/2 phosphorylation. In the phosphoproteomic experiments we did not observe β-arrestin2 but we did observe similar levels of ERK1/2 phosphorylation induced by both chemokines. Thus, the current results are not inconsistent with our previous signalling study but overall, they suggest that there are minimal differences between the global signal transduction networks activated by MCP-1 and MCP-3.

4.8 Conclusion

In this chapter, we report the first global phosphoproteomic study to quantify the changes in protein phosphorylation initiated by activation of a pro-inflammatory chemokine receptor – CCR2. By comparing the phosphorylation changes induced by MCP-1 and MCP-3, we found that these two chemokines cause very similar changes in cellular protein phosphorylation and signalling networks. Using the global proteomic and phosphoproteomic data, we developed a comprehensive map of the MCP/CCR2 network. In addition to enhancing our knowledge of the phosphorylation changes in some previously identified pathways, we, for the first time, characterised a highly connected actin polymerisation pathway activated by MCP/CCR2 activation, phosphorylation changes of ARHGEFs that control signal transduction and translocation of transcriptional factors by changing the phosphorylation states of NPC. The cellular outcome of actin polymerisation was further validated by confocal microscope. Future work will be needed to characterise the novel aspects of the proposed network using cells endogenously expressing CCR2, including monocytes and macrophages. This newly mapped signalling network may provide new therapeutic targets for chemokine-mediated inflammation, thus potentially transferring the spotlight from targeting chemokines or receptors to targeting downstream signalling effectors.

Chapter 5. Global Analysis of Protein Phosphorylation Kinetics in MCP-1 Activated Cells

5.1 Introduction

In Chapter 4, we have described the application of DIA-MS to reproducibly and accurately quantify phosphorylation sites involved in signal transduction activated by either MCP-1 or MCP-3, relative to vehicle control. Notwithstanding some differences in the phosphorylation network between MCP-1 and MCP-3 treated cells, the overall systemic cellular responses were observed to be similar for both chemokine treatments. We therefore created a manually curated network map of pathways and signal transduction cascades affected by activation of CCR2. In addition to known pathways such as the MAPK pathway, the JAK/STAT3 pathway and the PI3K/Akt pathway, we also identified novel phosphorylation events and protein networks that were previously not connected to the MCP/CCR2 pathway such as Rho guanine nucleotide exchange factors (ARHGEF) or proteins involved in the formation and function of nuclear pore complexes. More importantly, we could show for the first time that the MCP/CCR2 pathway directly acts on the actin cytoskeleton, which was further validated by confocal microscopy.

To further characterise our manually curated network map and, more importantly, to study the kinetics of the phosphorylation events and pathways identified in Chapter 4, we performed a time course experiment in which FlpIn-CCR2-HEK293 cells were treated with MCP-1 for 0, 1.5, 3, 6, 12, 25 and 60 minutes. After harvesting the cells and extracting proteins, we used DIA-MS to analyse changes in phosphorylation of many proteins at each time point.

5.2 Global Quantitative Proteomics and Phosphoproteomics

5.2.1 Generation of Spectral Libraries

To ensure that all possible phosphorylation events and proteins within this 60 minutes' time course study are present in the spectral library, we did not rely on the spectral library from Chapter 4. Instead we fractionated one biological repeat from each time point using high pH reverse phase C18 column chromatography to generate a new proteome and phosphoproteome spectral library. Similar to Chapter 4, the individual fractions were acquired by DDA and analysed with MaxQuant 1.5.5.1 [306] (see Figure 4.1, Chapter 4, for workflow).

The established global proteome library contained 152,441 unique precursors from 115,093 peptides, of which 109,852 are proteotyptic peptides. These peptides are derived from 8,164 proteins, which represent about 80% of the human core proteome [104,269]. In contrast, the established phosphoproteome library contained 22,170 unique peptides derived from 45,054 precursors. More importantly, the new library contained 36,982 modified peptides from 5,417 proteins. These spectral libraries are comparable to the ones from Chapter 4 in terms of the numbers of identified proteins and phosphopeptides. Indeed, comparative analysis revealed an overlap of 91% and 72 % between the proteins and phosphopeptides, respectively.

5.2.2 In-depth Quantification of the Proteome and Phosphoproteome

The quantification of proteins and phosphopeptides was accomplished by DIA-MS as described in detail in Chapter 4. Each time point was analysed using 3 independent biological replicates and the mass spectrometric raw files were analysed in Spectronaut 9.0 (peak picking and quantification). Perseus 1.5.5.3 and several in-house generated R scripts were used to visualise and further mine the data.

Using this strategy, we managed to quantify a total of 7,433 proteins and 19,331 phosphopeptides across all samples considering a q-value cutoff of 1%. Of note, more than 6,500 proteins were quantified in each and every sample. In addition, 16,590 phosphopeptides fell below the predefined q-value threshold of 1% in 2 (out of 3) biological repeats for each time point. Among these 16,590 phosphopeptides, 77.7% are singly phosphorylated, 20.4% are doubly phosphorylated and 1.9% are triply phosphorylated. As expected, and similar to the results from Chapter 4, the distribution of phosphoserine, phosphothreonine and phosphotyrosine was observed to be 84.5%, 14.9% and 0.6%, respectively.

To assess the quality of these data, we determined the Pearson correlation coefficient (R) between each sample (Figure 5.1) and calculated the coefficients of variation (CV) for each time point (Figure 5.2) The Pearson correlation coefficients were observed to be higher than 0.9 between each replicate suggesting very low variability between the biological replicates. This is further confirmed by the CVs, which were found to be around 11% and 22% for the proteome and phosphoproteome quantification, respectively, illustrating a very good quantitative accuracy as well as a very high reproducibility of these DIA-data.

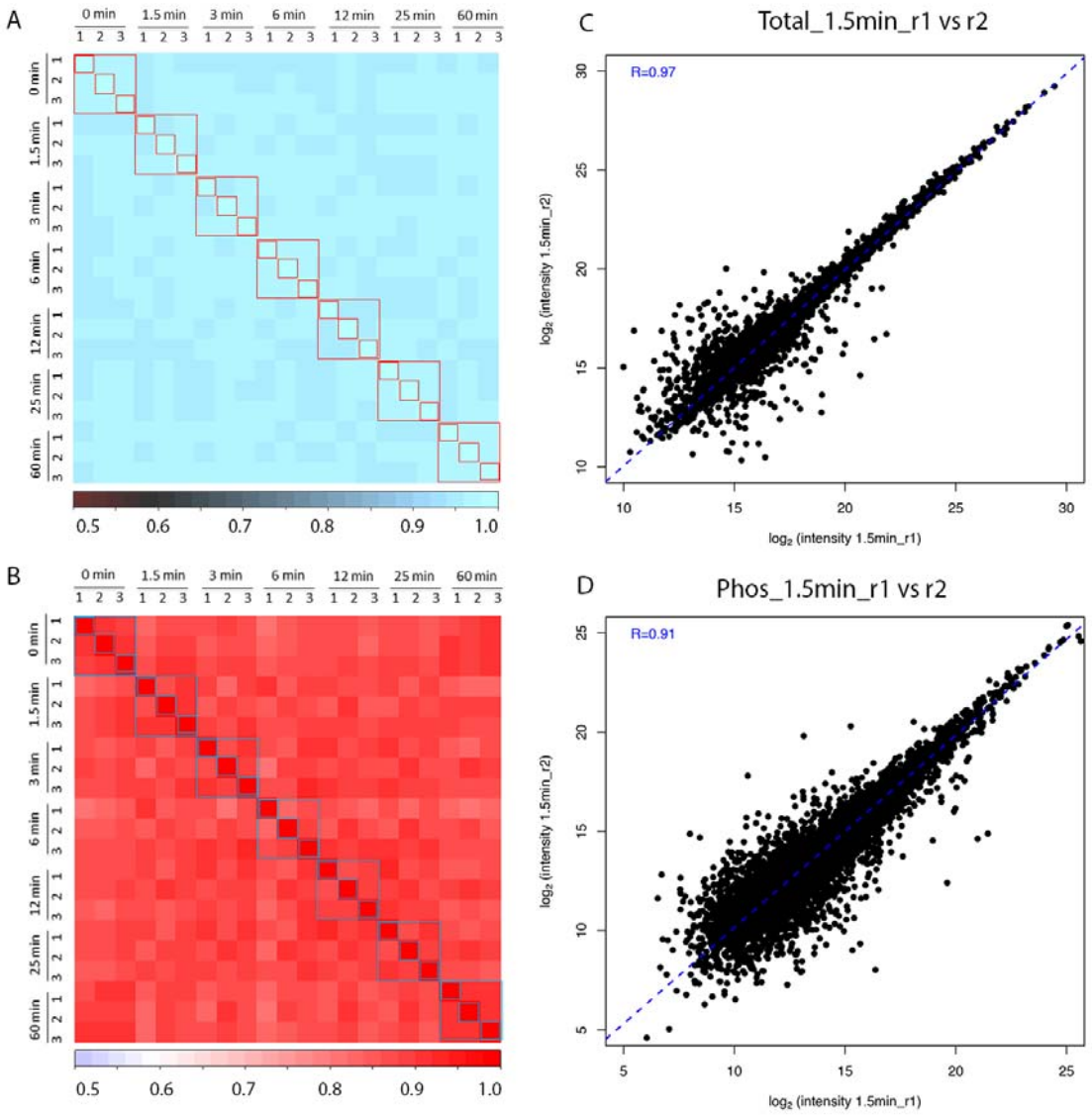


Figure 5.1. The heat map (A and B) and scatter plot (C and D) representing Pearson correlations (R). The heat maps (A and B) and scatter plots (C and D) represent the Pearson correlation coefficients (R) between protein (A and C) or phosophopeptide (B and D) intensities across independent experiments. A and C: the big square indicates 3 independent repeats for the same treatment and the small red square is the same sample (R=1); the colour scale is shown for R values of 0.5-1.0. C and D: scatter plot between 1.5 minutes' treatment repeat 1 and repeat 2 for total proteins and phosphopeptides; the Pearson correlation is indicated on the top right with blue value (R); each black dot indicates a single protein or phosphopeptide; x and y axis is $\log_2$ transformed intensity.

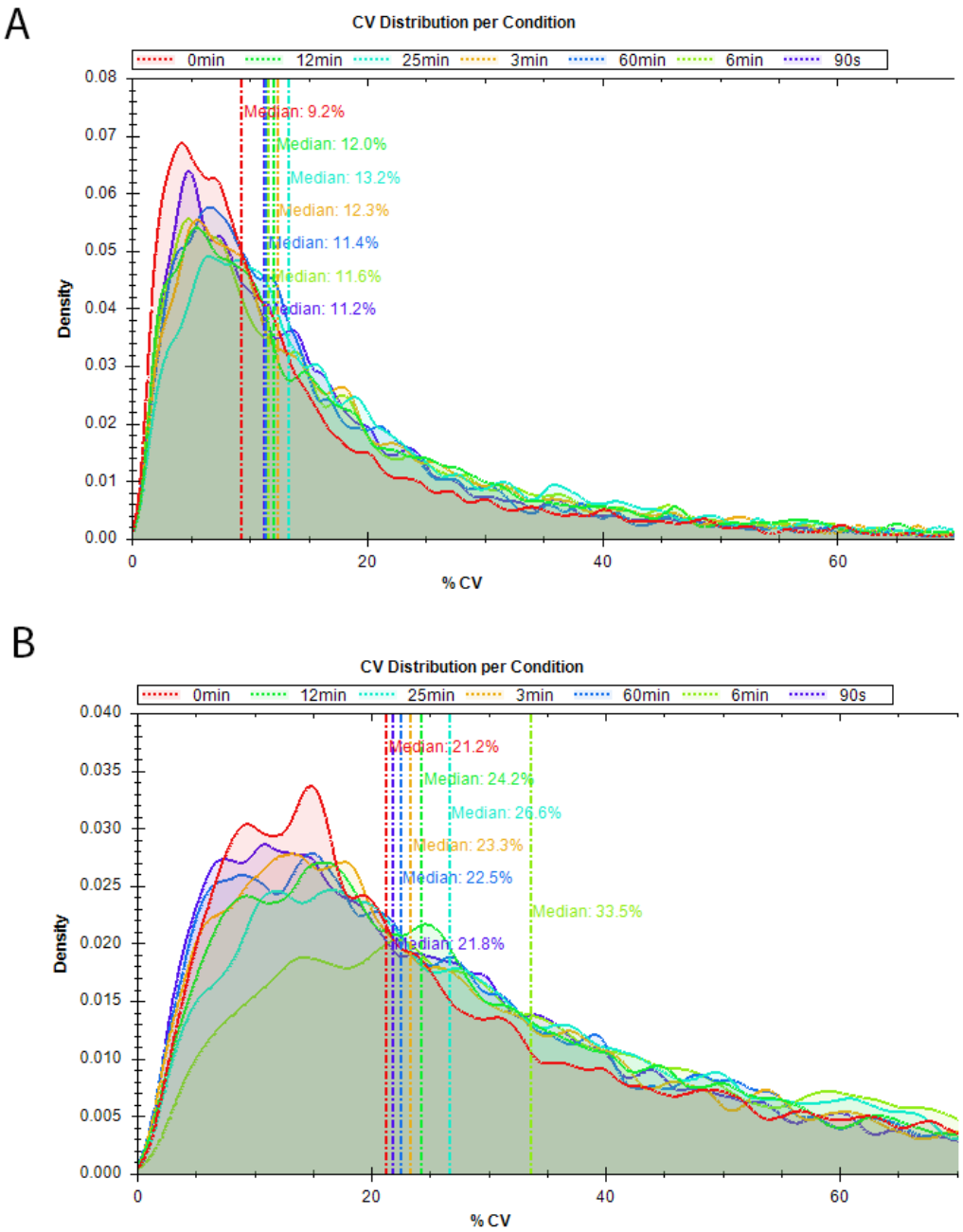


Figure 5.2. Coefficient of variation (CV) plots for protein and phosphopeptide quantification. Shown are the distributions of CV values for (A) proteome quantification and (B) phosphoproteome quantification. The median CV values for each time point is as indicated on top of each figure.

Two-sample T tests were performed in Perseus to calculate quantitative p values and fold changes of each protein and phosphopeptide between each time point ($t > 0$) and the reference time point ($t = 0$), which was used as control. Volcano plots were used to visualise these results (Figure 5.3 and Appendix IV, Figure S5.1 to S5.5), in which significantly up or down regulated proteins/phosphopeptides are located in the left and right upper quadrants. As expected, very few proteins (Figure 5.3A) were observed to be differentially regulated after 1.5 min in comparison to untreated cells. In contrast, hundreds of phosphopeptides (Figure 5.3B) were found to be significantly up- or down-regulated already after 1.5 min of MCP-1 treatment.

To restrict all subsequent analyses to the most statistically regulated phosphopeptides, we applied 3 stringency criteria: (1) we only kept proteins and phosphopeptides with a p value of less than 1% and a fold change of ≥1.5 or ≤ 0.67 [corresponding to $-\log_{10}$ (p-value) ≥ 2 and $\log_2$ (fold change) ≥ 0.58 or ≤ -0.58 for one or more time point]; (2) we excluded all phosphopeptides, which were derived from proteins that changed significantly between the different time points (3) we only kept phosphopeptides with a localisation probability (for the position of phosphorylation) higher than 0.75. Overall, 1,716 phosphopeptides were kept for further analysis after applying these stringency criteria.

5.3 Data Comparison to Chapter 4

The time course experiment described in the current chapter includes a 3-minute time point. Of note, the time point chosen for the static experiment in Chapter 4 was also 3 minutes. This circumstance provides an ideal opportunity to assess the accuracy, robustness and reproducibility of our experimental set-up and workflow as both 3 minutes' treatments with MCP-1 should result in identical cellular responses.

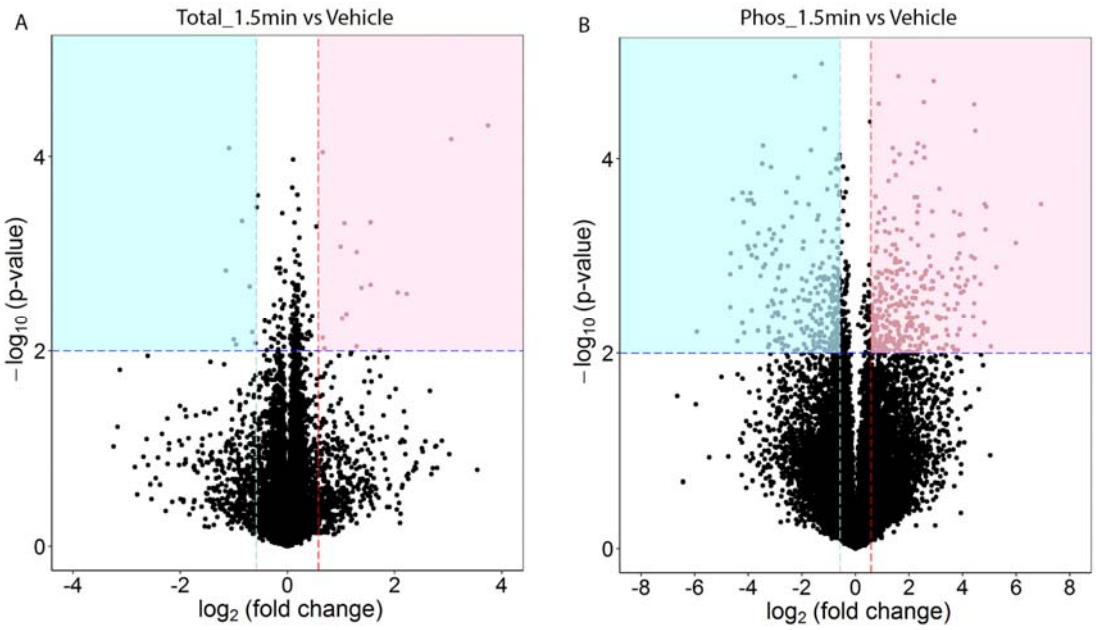


Figure 5.3. Volcano plots showing the effects of 1.5 minutes' treatment; A: comparison between vehicle-treated and MCP-1-activated cells (1.5 minutes) in protein quantification; B: comparison between vehicle-treated and MCP-1-activated cells (1.5 minutes) in phosphopeptide quantification. The x-axis shows the $\log_2$-transformed fold change and the y-axis the corresponding $-\log_{10}$ (p value); each black dot indicates a protein or phosphopeptide. The horizontal blue dash line indicates a p value of 0.01; the vertical grey dashed line (left) indicates a 1.5-fold decrease and the vertical red dash line (right) indicates a 1.5-fold increase. The blue and pink areas highlight proteins or phosphopeptides that are significantly decreased or increased, respectively.

First, we determined the overlay between the two experiments and observed that approximately 5,000 (out of around 16,000) phosphopeptides had been quantified in both experiments. The relatively low number of shared phosphopeptides is possibly due to the problem of missing cleavage, which happens quite often when detecting phosphopeptides [194]. Second, we averaged the intensity of each of these 5000 phosphopeptides across the 3 biological repeats within each experiment and calculated a Pearson correlation coefficient to determine the similarity of these independent experiments. As shown in Appendix IV, Figure S5.6A, the resulting R value is higher than 0.6 suggesting an acceptable correlation between these completely independent experiments. Of note, most of the data points (black dot) are clustered around the central diagonal with only relatively few outliers that affect the R value substantially (Appendix IV, Figure S5.6B). Third, instead of analysing the Pearson correlation of all identified phosphopeptides, we focused on only the significantly regulated phosphopeptides identified in both experiments. As shown in Figure 5.4A, we observed an R^2 value of higher than 0.8 confirming that both independent experiments resulted in very similar cellular responses. As examples, we specifically show the significantly increased phosphorylation of ERK1/2 at Thr-202/Tyr-204 (Figure 5.4B) and Thr-185/Tyr-187 (Figure 5.4C).

In summary, we compared the results of two completely independent experiments in which FlpIn-CCR2-HEK293 cells were treated with MCP-1 for 3 minutes. Despite some differences, good correlations have been observed between these 2 experiments, especially if only significantly regulated phosphopeptides are considered.

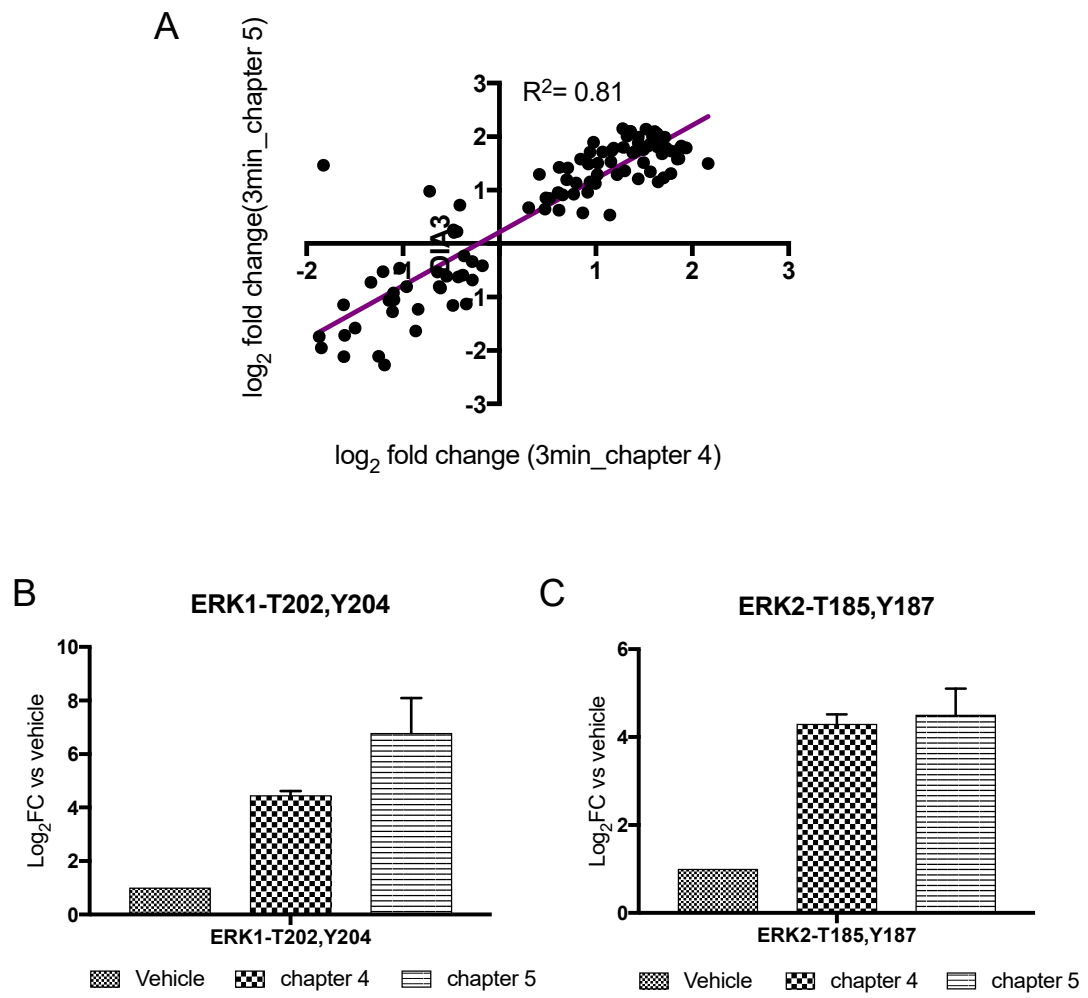


Figure 5.4. The expression level of significantly regulated phosphopeptides that identified in both experiments is plotted in (A) and coefficient of determination (R^2) is used to present the correlation. The log_2 transformed fold change, normalised to vehicle treatment, of ERK1-Thr-202 (T202)/Tyr-204 (Y204)and ERK2-Thr-185 (T185)/Tyr-187 (Y187) is shown in (B) and (C).

5.4 Expanding the MCPs/CCR2 Network

In Chapter 4, we used our mass spectrometric data to create a map of protein networks that are affected by MCP-1 (and MCP-3) treatment. With these time course data in hand, we sought to expand this network map. To accomplish that, proteins from the 1,716 significantly regulated phosphopeptides were submitted to STRING 10.5 to retrieve a list of their known and documented protein interaction partners [307]. We added a protein to the network map only if: (1) the interaction confidence score with an already mapped protein was higher than 0.4; (2) this new interaction had been confirmed by experimental evidence reported in the literature; and (3) the coefficient of determination (R^2) of phosphorylation dynamics was higher than 0.5 between the two interacting proteins. The newly added protein nodes are shown in magenta in Figure 5.5; most are located in the nucleus or in the actin cytoskeleton pathway.

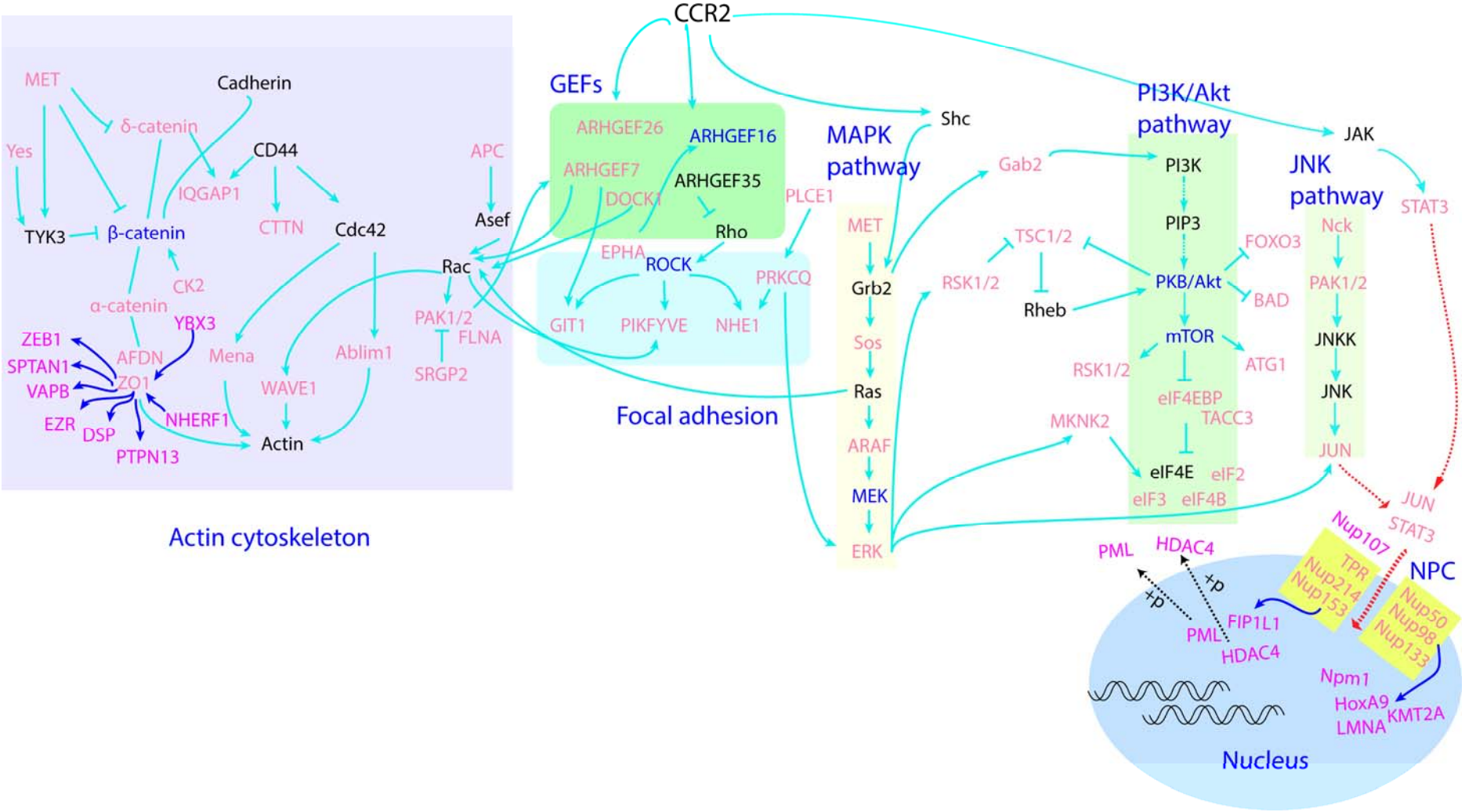

CCR2
GEFs
ARHGEF26
ARHGEF16
ARHGEF7
ARHGEF35
DOCK1
Rho
EPHA
ROCK
PRKCQ
GIT1
PIKFYVE
NHE1
Focal adhesion
PLCE1
MAPK pathway
Shc
MET
Grb2
Sos
Ras
ARAF
MEK
ERK
PI3K/Akt pathway
Gab2
PI3K
PIP3
TSC1/2
FOXO3
RSK1/2
PKB/Akt
BAD
Rheb
mTOR
ATG1
RSK1/2
eIF4EBP
TACC3
MKNK2
eIF4E
eIF2
eIF3
eIF4B
JNK pathway
JAK
STAT3
Nck
PAK1/2
JNKK
JNK
JUN
JUN
STAT3
Nup107
TPR
Nup214
Nup153
NPC
Nup50
Nup98
Nup133
PML
HDAC4
+p
+p
PML
HDAC4
FIP1L1
Npm1
HoxA9
LMNA
KMT2A
Nucleus
MET
Cadherin
δ-catenin
Yes
CD44
APC
IQGAP1
Asef
TYK3
β-catenin
CTTN
Cdc42
Rac
CK2
α-catenin
YBX3
PAK1/2
FLNA
ZEB1
SPTAN1
VAPB
AFDN
Mena
Ablim1
ZO1
WAVE1
SRGP2
EZR
DSP
NHERF1
PTPN13
Actin
Actin cytoskeleton

Figure 5.5. A summarised MCP/CCR2 signalling network. This new network is based on the re-constructed network from Chapter 4. A protein name in pink indicates that one or more phosphorylation site(s) belonging to this protein have been found to be significantly changed in our data set. A protein name in blue indicates that the protein has been identified in our study, but not shown as significantly changes. A protein name in black indicates that the protein has not been identified in our study. The newly added proteins from this chapter are indicated with magenta colour.

As discussed in Chapter 4, GEF, NPC and actin cytoskeleton are the three main newly mapped sub-networks in MCP/CCR2 signalling. Therefore, we mainly focused on these 3 sub-networks. Using our approach, we were unable to identify additional GEFs suggesting that the GEFs characterised in Chapter 4 are possibly the major GEFs affected by MCP-1 in HEK293 cells. In contrast, 8 additional proteins were added into the NPC network (magenta nodes in Figure 5.6). One protein (Nup107) represents an actual nuclear pore protein, while the other seven are nuclear proteins. Phosphorylation of two of these nuclear proteins (PML and HDAC4) has been shown to trigger their export from the nucleus to the cytosol [308-310], while the others are localised in the nucleus, functioning as transcription factors or activators.

Moreover, 8 additional proteins were added into actin cytoskeleton network or more specifically connected to the ZO1 protein of the actin cytoskeleton network (Figure 5.7). ZO1 (or TJP1) acts as a scaffold protein in tight junctions, localises at the apical end of cell membrane and plays a vital role in forming cell adhesion complexes and cell-cell connections [311]. Therefore, it is not surprising that ZO1 is an important protein for regulating cell migration [312]. Six of the new proteins that interact with ZO1 (SPTAN1, VAPB, EZR, DSP, PTPN13 ad NHERF1) are components of adherens junctions or the cytoskeleton, while YBX3 and ZEB1 can function as a transcription repressor or an activator, respectively [313,314].

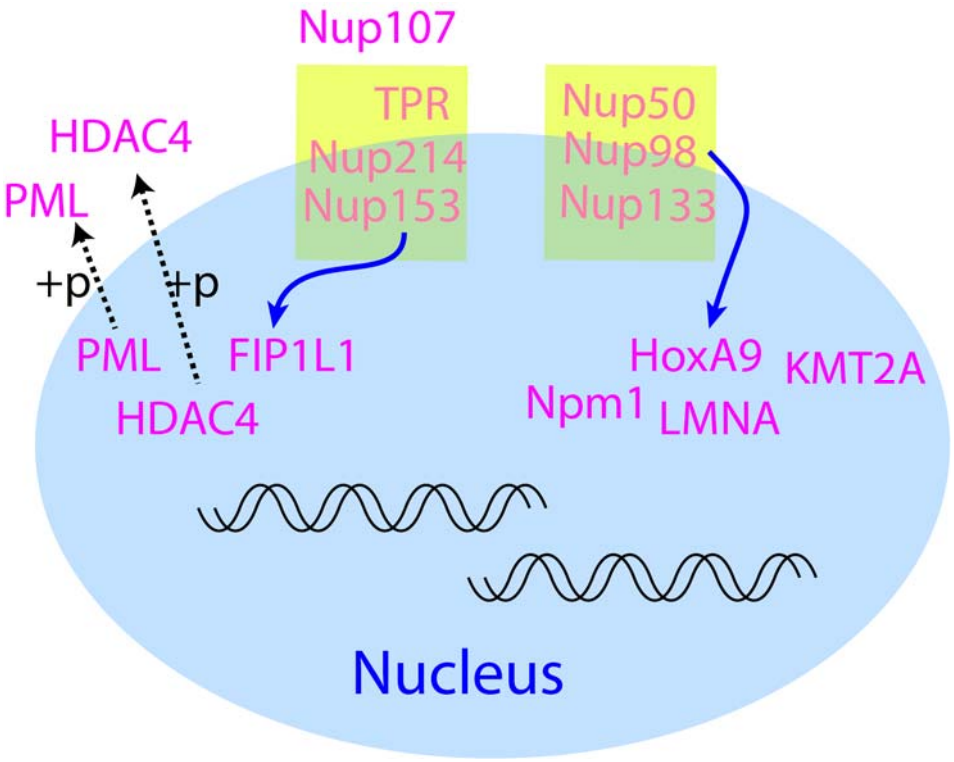


Figure 5.6. Expanded network in NPC. The newly added protein nodes are coloured with magenta and new connections is indicated by blue arrow.

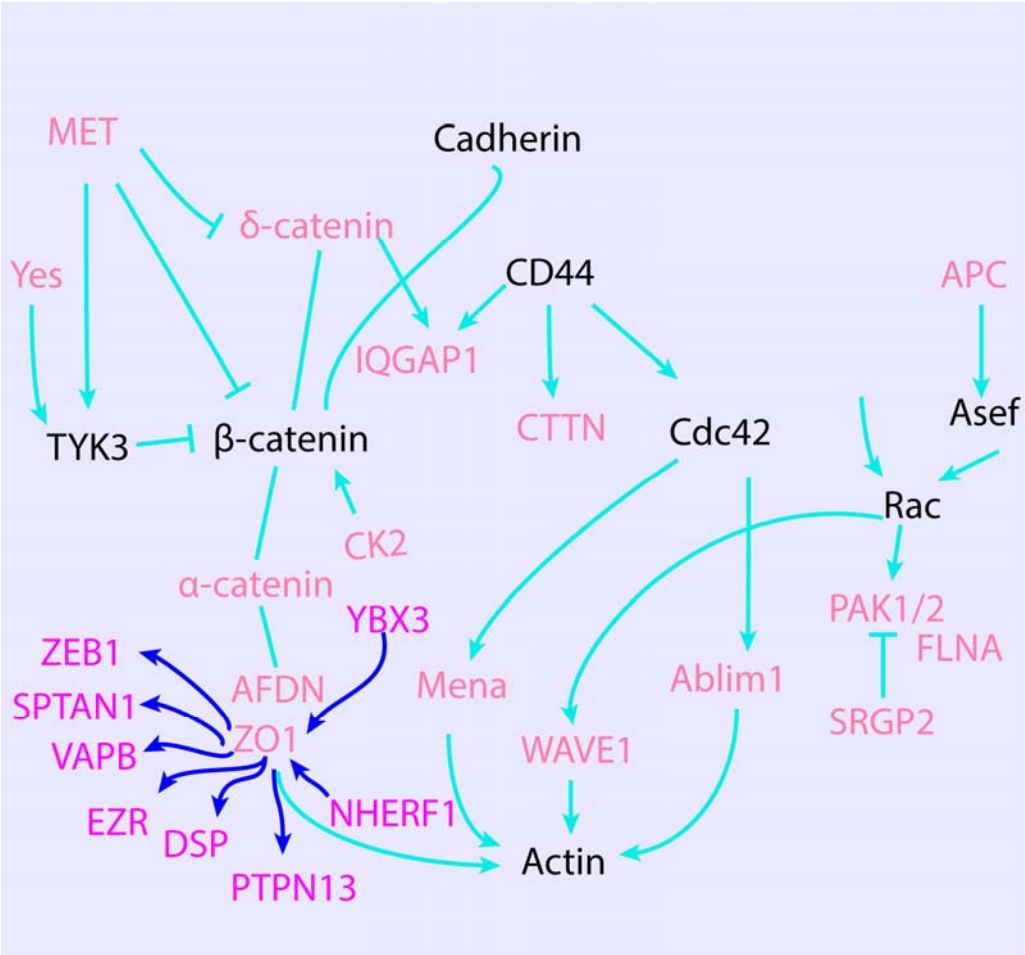


Figure 5.7. Expanded network in actin cytoskeleton pathway. The newly added protein nodes are coloured with magenta and new connections is indicated by blue arrow.

5.5 Phosphorylation Dynamics

Signal transduction is dynamically regulated by protein phosphorylation and dephosphorylation. In order to understand the dynamics of phosphorylation events in the MCP/CCR2 signalling pathways, we investigated the phosphorylation dynamics of selected proteins within our expanded network map. In addition to the phosphorylation dynamics in canonical pathways, we focussed on three sub-networks that have been newly characterised in this thesis. Of note, these three sub-networks directly or indirectly affect cell migration, which is one of the most important biological responses following MCP-1 treatment. The three sub-networks are: (1) the actin cytoskeleton pathway, which directly controls cell migration; (2) guanine nucleotide exchange factors, which are responsible for relaying the signal from CCR2 to the actin-related pathway; and (3) nuclear pore complex proteins and nuclear proteins, which control the transcriptional machinery that is required to produce the critical mass of proteins for cell migration.

5.5.1 Phosphorylation Dynamics in Canonical Pathways

Before analysing the above three pathways, we first examined the phosphorylation cascade in the canonical pathways. Numerous studies have confirmed that MCP-1 treatment leads to phosphorylation of ERK1/2, which in turn phosphorylates c-Jun [315,316]. This regulation cascade is also supported by our dynamic phosphoproteomic data. As shown in Figure 5.8A and B, a maximal phosphorylation response of ERK1/2 at Thr-185/Tyr-187 and Thr-202/Tyr-204 upon MCP-1 treatment was detected between 1.5 and 3 minutes, but basal phosphorylation levels were rapidly reached after 6 minutes. Both the increase and decrease in ERK phosphorylation were confirmed by Western blotting (Figure 5.8D, E and G). As expected for a substrate of ERK1/2, c-Jun was maximally phosphorylated after 6 to 12 minutes, after the maximal phosphorylation of ERK1/2, and the phosphorylation was then gradually decreased after a prolonged

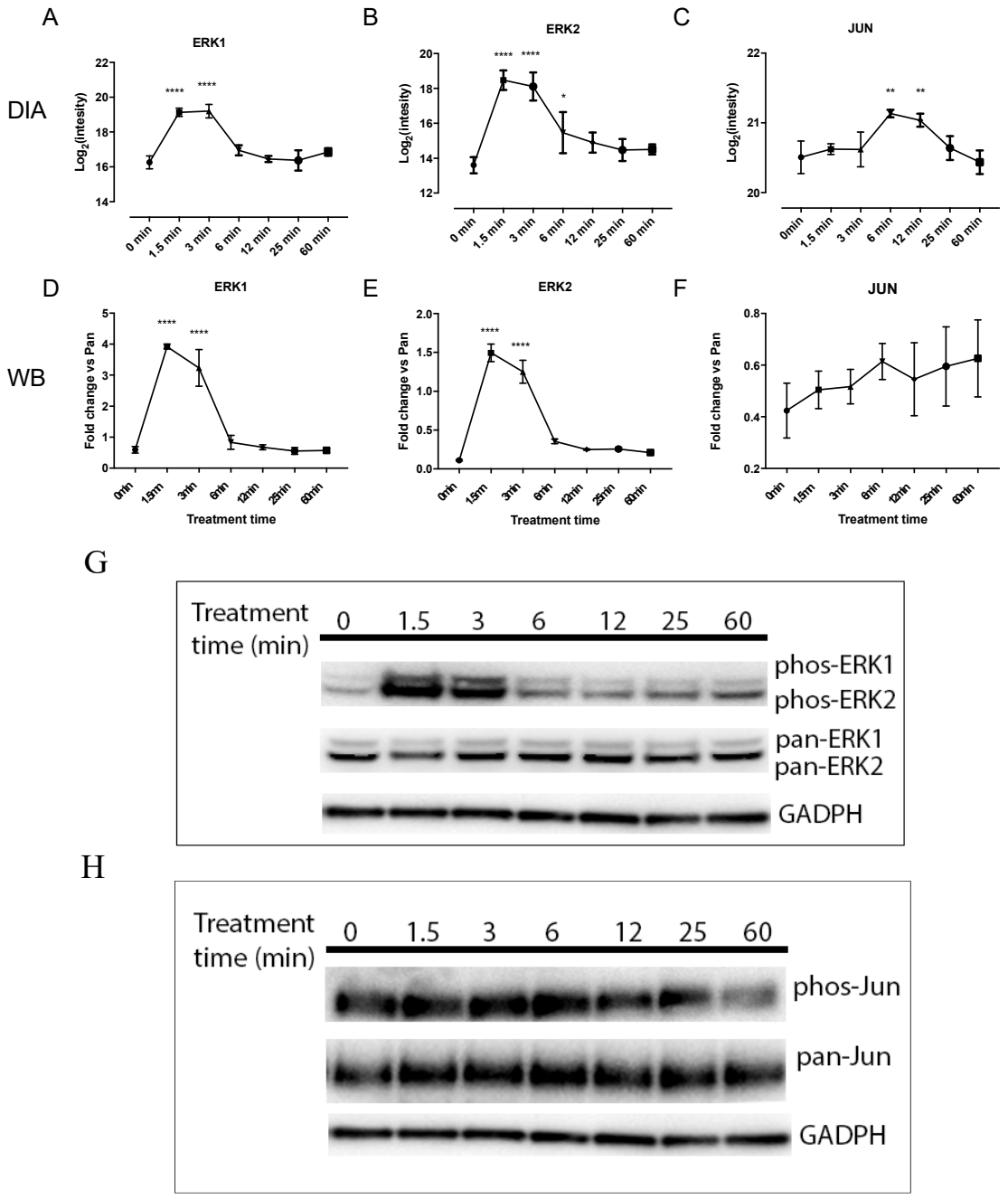


Figure 5.8. Phosphorylation dynamics of ERK1/2 and c-Jun. (A-C) DIA-MS quantification results for (A) ERK1 at Thr-185/Tyr-187, (B) ERK2 at Thr-202/Tyr-204 and (C) c-Jun at Ser-63. (D-H) Western blot quantification results for (D,G) ERK1, (E,G) ERK2 and (F,H) c-Jun at Ser-63. All data are presented as mean ± SEM, n=3. * $P < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.00001$ vs time 0min (one-way ANOVA followed by Dunnett's post hoc test)

treatment time (Figure 5.8C). Western blotting was consistent with the increase in phosphorylation in the first 6 minutes (albeit below statistical significance), but the errors at later time points were too high to enable confirmation of the subsequent decrease (Figure 5.8F and H).

MAP kinase-interacting serine/threonine-protein kinase 2 (MKNK2) is another substrate of ERK1/2 [317] as shown in Figure 5.5. In agreement with ERK1/2 phosphorylation kinetics (Figure 5.8), an increased phosphorylation level at Ser-453 was observed for MKNK2 after 3 minutes, which then gradually decreased to basal levels over an extended MCP-1 treatment time (Figure 5.9A).

In addition to the ERK1/2 node of the network, many other canonical nodes show a phosphorylation pattern that is consistent with previous publications. For example, signal transducer and activator of transcription 3 (STAT3) is a transcription factor that can autophosphorylate itself or can be phosphorylated in-trans by the Janus kinases (JAK) after activation of CCR2 signalling [318]. Moreover, phosphorylation of its serine residue 727 has been shown to increase the transcriptional activity of STAT3 [319]. In agreement with this, we observed an increased phosphorylation of Ser-727 (Figure 5.9B), which peaked after 6 minutes of MCP-1 treatment and then gradually decreased. This observation not only confirms and extends the notion that the transcriptional activity of STAT3 is modulated by CCR2 signalling but also supports our hypothesis (generated in Chapter 4) that activation of CCR2 signalling can increase transcriptional activity, which subsequently facilitates cell migration by promoting protein biosynthesis.

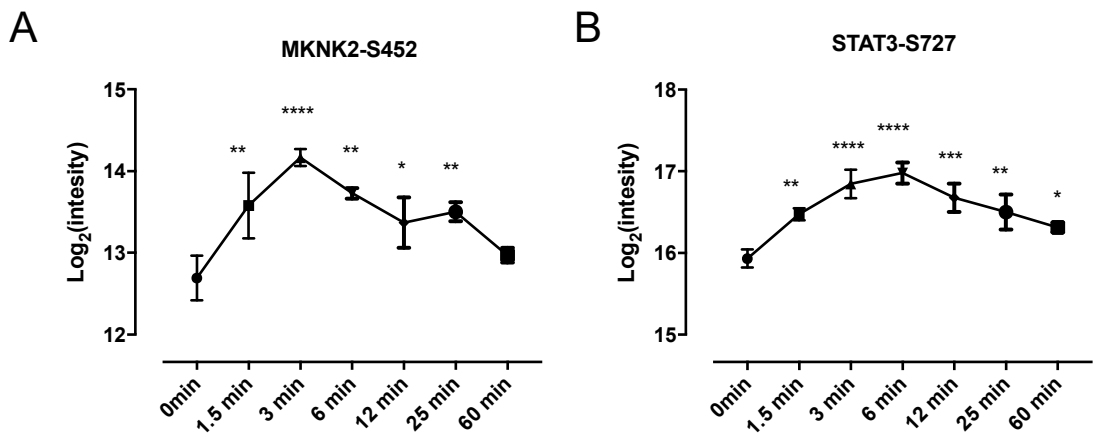


Figure 5.9. The phosphorylation kinetics of (A) MKNK2 at Ser-452 (S452) and (B) STAT3 at Ser-727 (S727). All data are presented as mean ± SEM, n=3. * $P < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.00001$ vs time 0min (one-way ANOVA followed by Dunnett's post hoc test)

5.5.2 Phosphorylation Dynamics of the Actin Cytoskeleton Pathway to Induce Cell Migration

In general, cell movement is one of the major biological responses to chemokine signalling as attracted cells have to migrate to the source of the chemokine. A series of coordinated mechanisms are required for this cell movement. First, actin has to polymerise to forms protrusions. Second, cell contraction leads to disassembly of adhesion molecules and movement. Finally, the new protrusion area is stabilised by focal adhesion, leading to cell reattachment [320]. In the following sections these three basic steps of cell movement will be connected to and partially explained by the underlying phosphorylation changes of selected proteins observed in this time course study. This discussion combines current knowledge from the literature with the results of in-depth data mining of our phosphoproteomic data set. While some parts are obvious or supported by existing evidence, others are speculative and will need to be tested by additional experiments; this is also true for sections 5.5.3 and 5.5.4.

As mentioned earlier in Chapter 4 and as expected, phosphorylation changes of actin-related proteins are directly involved cell migration. For example, dephosphorylation of WAVE1 at Ser-310 results in actin polymerisation, which is the first step of cell movement signalling [288]. As clearly shown in Figure 5.10A, decreased phosphorylation of WAVE1 was detected until 25 minutes of MCP-1 treatment, returning to basal levels after 60 minutes. This observation suggests that actin depolymerises and cells reattach again between approximately 25 and 60 minutes after receiving the MCP-1 stimulus.

Consistent with this notion is the observed (de-)phosphorylation pattern of casein kinase 2 (CK2), which is known to regulate a broad spectrum of pathways [321,322]. Dephosphorylation of CK2 can increase its activity [295], which results in inhibition of cell mitosis [323] and stabilisation of the E-cadherin/β-catenin complex, the function of which is to promote and strengthen cell adhesion [324]. A distinct drop in the Ser-209 phosphorylation levels is observed after 60 minutes (Figure 5.10B) further suggesting that cell reattachment, which is the last step during cell migration, occurs after approximately 60 minutes.

In addition to WAVE1 and CK2, we managed to add several new proteins to the actin cytoskeleton node of our network (Figure 5.5 and 5.7). Interestingly, all of these

newly added proteins are connected to Zonula occludens-1 (ZO1). ZO1, localised at the leading edge of cell membranes, is one of the tight junction proteins, which connects tight junctions to the actin cytoskeleton [325]. Dephosphorylation of ZO1 can result in disassembly of tight junctions [326], which is the second step in cell migration. As expected, we identified in our time course data three phosphopeptides derived from ZO1, two of which (Figure 5.11A and B) showed a rapid (~3-12 minute) dephosphorylation profile after activation of the CCR2 signalling cascade. The phosphorylation levels of ZO-1 recovered after 25 minutes implying that tight junctions are formed again after 25-60 minutes, which is in perfect agreement with our previous data and further supports the involvement of these proteins in MCP/CCR2-initiated cell migration. Another example that supports this finding is ezrin (EZR). EZR, together with radixin and moesin, can form a complex that connects the actin cytoskeleton to the plasma membrane [327]. Phosphorylation of EZR results in a loss of tight junctions and a subsequent improved cell spreading ability [328,329]. As shown in Figure 5.11C, we expected and observed a rapid increased of phosphorylated EZR, which then gradually decreases back to basal levels, again by 25 minutes.

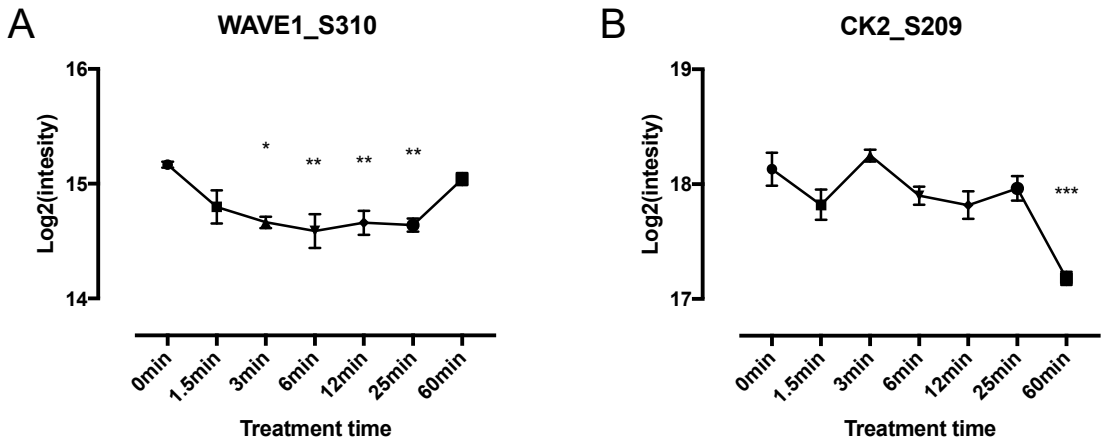


Figure 5.10. The phosphorylation kinetics of WAVE1 at Ser-310 (S310) (A) and CK2 at Ser-209 (S209) (B). All data are presented as mean ± SEM, n=3. * $P < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.00001$ vs time 0min (one-way ANOVA followed by Dunnett's post hoc test)

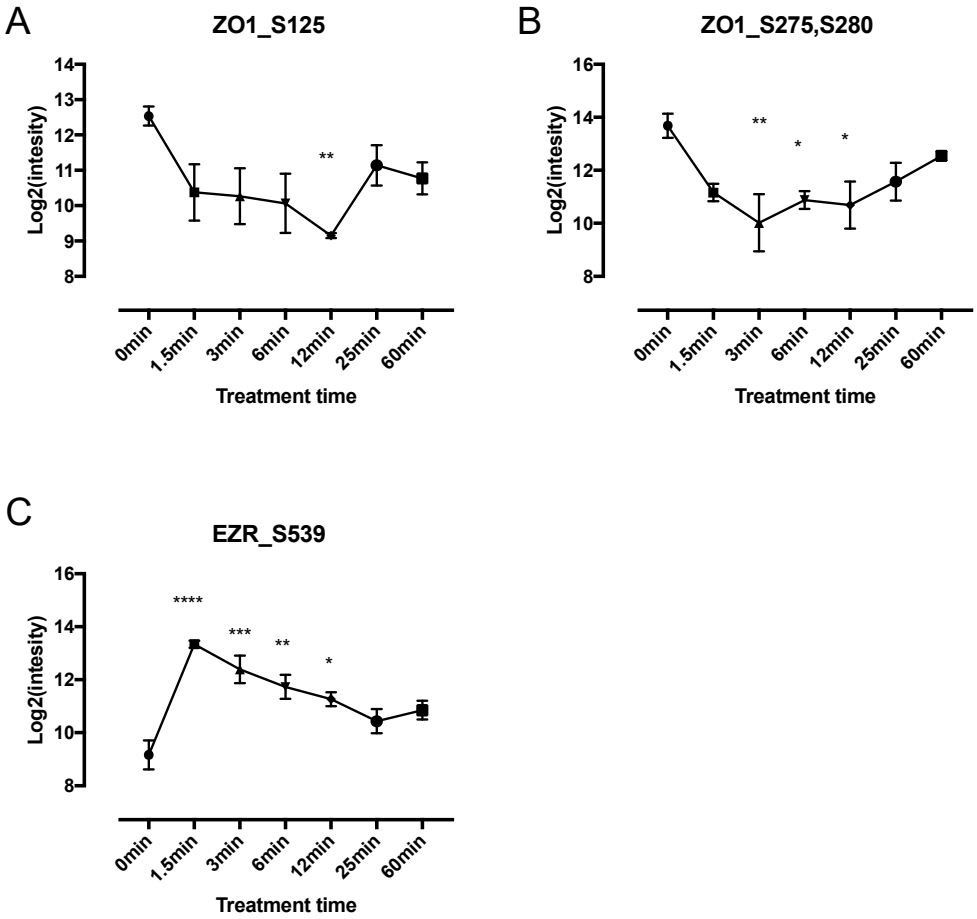


Figure 5.11. The phosphorylation kinetics of Z01 at Ser-125 (S125) (A), Ser-275/Ser-280 (S225, S280) (B) and EZR at Ser-539 (S539)(C). All data are presented as mean ± SEM, n=3. * $P < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.00001$ vs time 0min (one-way ANOVA followed by Dunnett's post hoc test)

In conclusion, cell movement as a result of activating MCP/CCR2 signalling can be explained, in part, through altered phosphorylation levels of actin-related proteins as shown here for WAVE1, CK2, ZO-1 and EZR. Our time course series is in good agreement with the notion that rapid phosphorylation changes (occurring in less than ~12 minutes) induce actin polymerisation to form cellular protrusions and to detach cells by resolving tight junctions and adhesions. These effects revert back after prolonged (~25 minute) CCR2 activation resulting in cell reattachment and actin depolymerisation.

5.5.3 Phosphorylation Dynamics of Guanine Nucleotide Exchange Factors

As discussed in Chapter 4, ARHGEFs, which belong to the Rho guanine nucleotide exchange factors gene family, act as the main "switch" of controlling cytoskeletal organization and cell migration [285]. The direct regulation of ARHGEFs by GPCRs or chemokine receptors is not fully understood, but it has been assumed that, upon being activated by their ligand, GPCRs can be phosphorylated, resulting in the translocation and activation of ARHGEFs [330-332]. In addition, well-controlled phosphorylation events are considered to be one of the major mechanisms to activate ARHGEFs [286].

In our phosphoproteomic time course data, as shown in Figure 5.12, ARHGEF26 is quickly phosphorylated at serine residues 127 and 222 upon treatment of cells with MCP-1 then the phosphorylation levels at both sites gradually decrease after 6 minutes. In contrast, ARHGEF7 (Ser-176) appears to be maximally phosphorylated at 1.5 minutes, then immediately dephosphorylated again. Although the phosphorylation patterns of ARHGEF26 and ARHGEF7 are slightly different, they both reach a basal level after 25 minutes, which may suggest that ARHGEF-mediated signalling is turned off after 25 minutes.

As discussed in Chapter 4, ARHGEFs are mainly responsible for regulation of cell migration and cytoskeletal organisation [285]. Based on these known connections and the relatively rapid phosphorylation profile of ARHGEFs, we hypothesise that the rapid phosphorylation and dephosphorylation of ARHGEF26 and 7 is one of the mechanisms to transduce the initial CCR2 signal leading to phosphorylation cascades involved in cytoskeletal actin and tight junction rearrangements through proteins such as WAVE1, CK2, ZO-1 and EZR. We speculate that phosphorylated (and active) ARHGEF26 and 7 contribute to cell migration and their dephosphorylation turns off the

signal required for cell migration. In support of this hypothesis, ARHGEF1 has been found to be phosphorylated in MCP-1 activated human aortic smooth muscle cells, leading to cell migration; in contrast, down-regulation of ARHGEF1 or inhibiting the necessary phosphorylation event compromised cell migration [333]. Interestingly, the expression of ARHGEFs is heavily dependent on the exact cell type [334,335], so the involvement of different ARHGEF subtypes (e.g. ARHGEF1 in smooth muscle cells vs. ARHGEF26 and 7 in HEK cells) is a potential mechanism to regulate cell type-specific responses to the same chemokine.

5.5.4 Phosphorylation Dynamics of Nuclear Pore Complex and Nuclear Proteins

Nuclear pore complex (NPC) proteins are involved in actively shuttling (de)phosphorylated transcription factors (including the well-known modulators ERK, Jun or STAT3) between the cytoplasm and the nucleus resulting in the activation or repression of transcriptional processes [336]. Of note, increased transcriptional activity can lead to increased protein biosynthesis, which has been shown to contribute to cell migration [283].

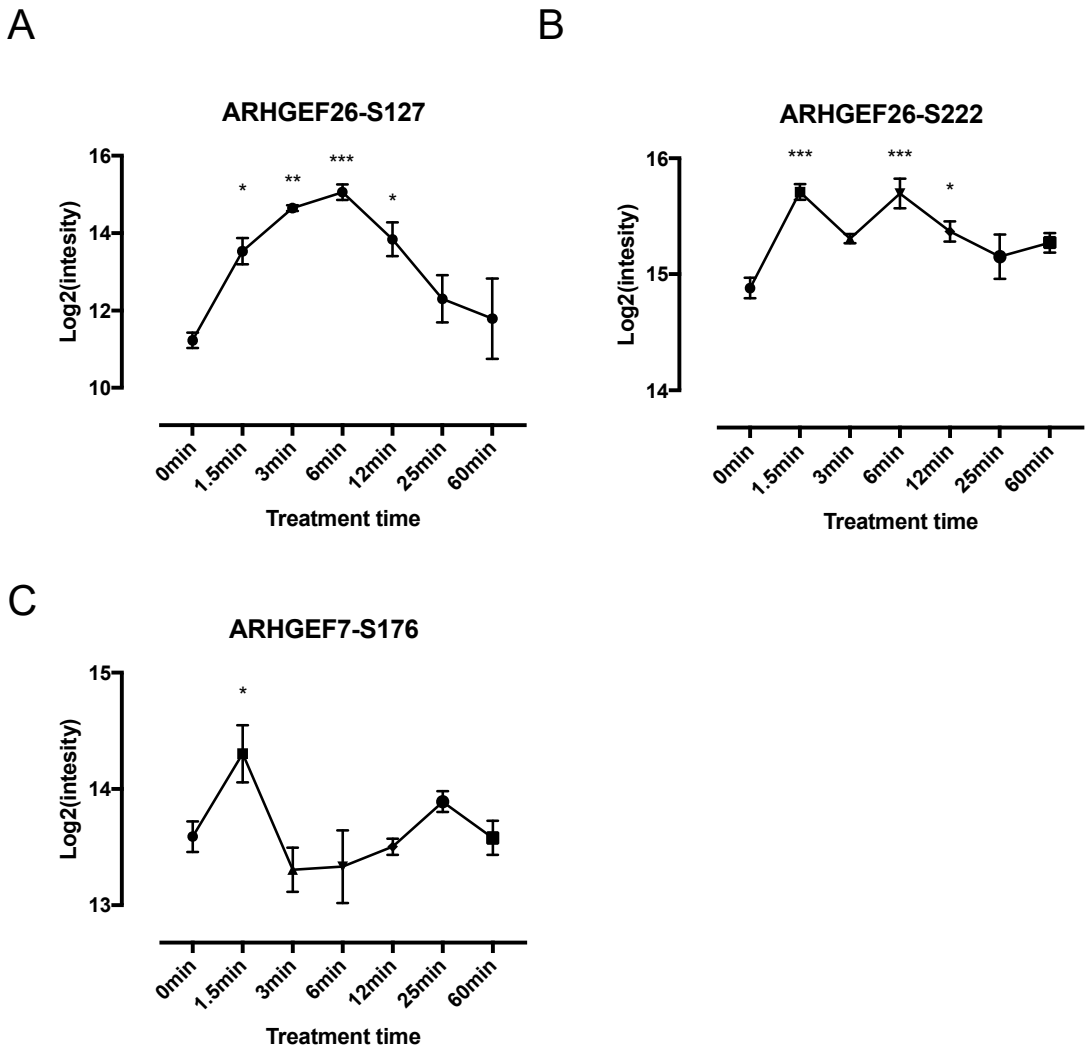


Figure 5.12. The phosphorylation kinetics of ARHGEF26 at Ser-127 (S127) (A) and Ser-222 (S222) (B) and 7 at Ser-176 (S176) (C). All data are presented as mean ± SEM, n=3. * $P < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.00001$ vs time 0 (one-way ANOVA followed by Dunnett's post hoc test)

The NPC itself can be controlled through increased or decreased phosphorylation [337,338]. In contrast to the phosphorylation pattern of the GEFs, which were observed to be maximally (de)phosphorylated after 1.5 to 6 minutes, most proteins involved in the NPC showed a maximal phosphorylation change after 25 min despite being altered already after 1.5 minutes, (Appendix IV, Figure S5.7). As an example, Ser-125 of NPM1 (Figure 5.13A) is gradually dephosphorylated after over the first ~6 minutes but retains its maximal dephosphorylation level over the range 6 to 25 minutes then returns to the basal level.

To date, no direct connection between the NPC and CCR2 signalling has been described, although Negi et al. found that increased phosphorylation of Ser-125 can increase the mobility of NPM1 into the nucleus resulting in cell mitosis [339]. Considering our observation that numerous NPC proteins undergo phosphorylation changes in response to MCP-1, we propose that activation of CCR2 signalling leads to the (de)phosphorylation of NPC proteins, which subsequently suppresses cell division, activates transcription and results in protein production. All of these effects contribute heavily to cell migration as protein production is required for cell movement [283]. In addition, as the phosphorylation events did not recover until 25 minutes of treatment, the process of transcription activation and protein biosynthesis appears to be a fairly long-lived process.

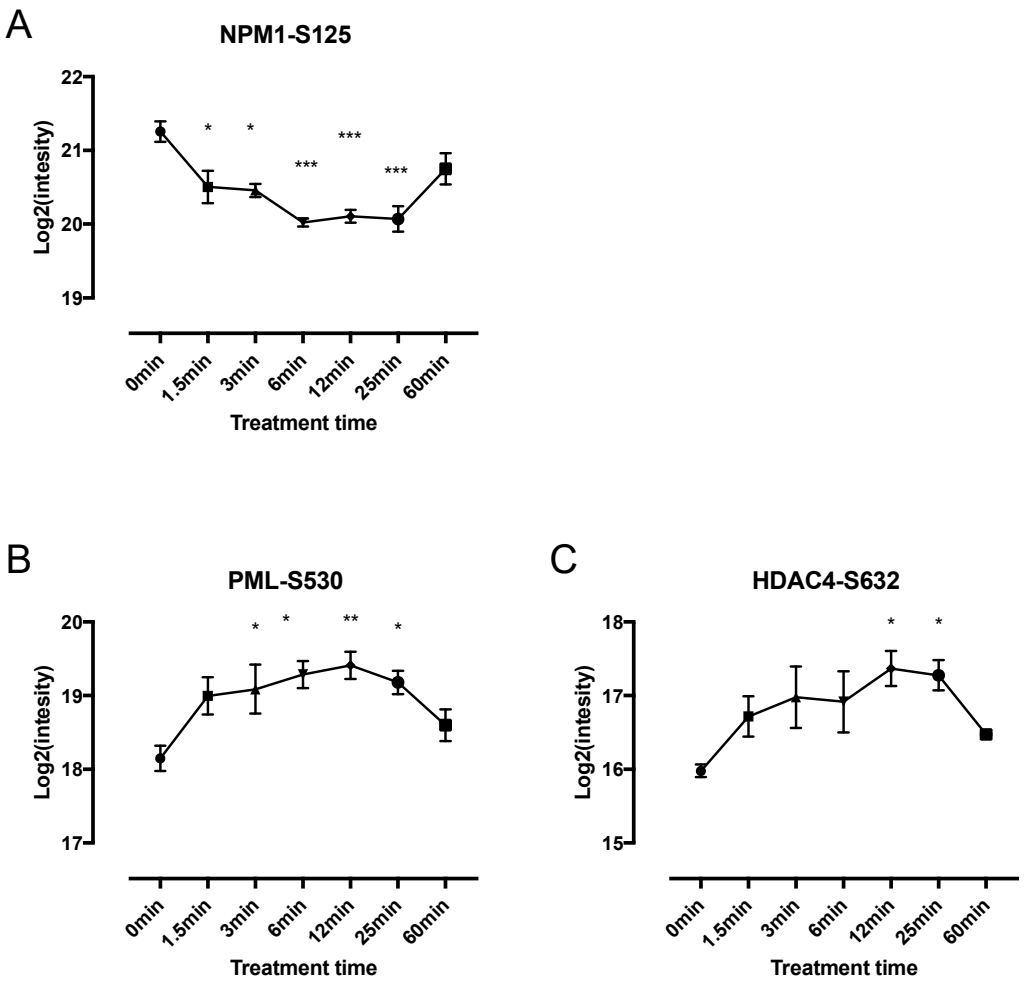


Figure 5.13. The phosphorylation kinetics of NPM1 at Ser-125 (S125) (A), PML at Ser-530 (S530) (B) and HDAC4 at Ser-632 (S632) (C). All data are presented as mean ± SEM, n=3. * $P < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.00001$ vs time 0min (one-way ANOVA followed by Dunnett's post hoc test)

This notion is further supported by the observed phosphorylation dynamics of other nuclear proteins such as promyelocytic leukemia protein (PML) and histone deacetylase 4 (HDAC4). PML is a hyper-phosphorylated protein and can act as a suppressor of transcription [340]. Phosphorylation of PML at Ser-530 has been found to be involved in its degradation [340-342]. In our phosphoproteomic data an increased phosphorylation level of PML at Ser-530 up to 25 minutes post-treatment was also observed, as shown in Figure 5.13B, implying that PML is degraded resulting in activation of transcription. In addition, HDAC4 can also function as a transcription repressor in the nucleus, while phosphorylation of HDAC4 can cause it to be exported out of nucleus to alleviate the inhibition of transcription [309,310]. As shown in Figure 5.13C, HDAC4 displayed a quite similar phosphorylation pattern to that of PML, consistent with increased transcriptional activity after MCP-1 treatment.

Taken together, we have very good evidence that CCR2 signalling acts on the phosphorylation status of several NPC proteins and nuclear proteins, which in turn promote activation of transcription through various mechanisms. An increased transcriptional activity is essential for cell migration as it provides the required amount of proteins, which are collectively involved in the cell migration machinery.

5.6 Summary

In this chapter, we used DIA-MS to study the phosphorylation kinetics of various regulated phosphosites and proteins in response to MCP-1-activated CCR2 signalling. We established that our DIA time course data are of high quality and shows excellent reproducibility and accuracy across the different time points and biological replicates. As expected, well-documented phosphorylation events of canonical pathways such as the phosphorylation of ERK1/2 or c-Jun were observed, which further confirms the validity of our time course data.

In addition, we examined the phosphorylation dynamics of three main sub-networks (actin cytoskeleton, GEF, NPC and nuclear proteins), which are proposed here to be directly or indirectly involved in cell migration. In light of previous studies describing the roles of these proteins in other signalling responses, our observations reinforce the involvement of numerous actin cytoskeleton proteins in responding to chemokine activation. Moreover, the data strongly suggest that phosphorylation changes of GEFs and nuclear pore (or other nuclear) proteins may play important roles in

regulating the activation of actin cytoskeleton proteins in response to MCP-1 (and perhaps also other chemokines).

In Figure 5.14, we present a schematic diagram illustrating the main steps in chemokine-initiated cell migration and showing how the timing of these event might be related to the dynamics of phosphorylation in canonical pathways and of GEF, NPC proteins and actin cytoskeleton proteins. From left to right, CCR2 is activated by MCP chemokines (Step 1), resulting in immediate activation of signalling through increased or decreased phosphorylation of proteins in the canonical pathways (Step 2). Phosphorylation changes of GEFs also occurs as part of this initial response and is proposed to play an important role in controlling "on and off" MCP/CCR2 signalling. Activation of GEFs gives rise to subsequent phosphorylation changes in actin cytoskeleton proteins, which occurs at a slower time scale than observed for GEFs. Similarly, activation of the canonical (e.g. Jak-STAT) pathways gives rise to downstream phosphorylation changes of nuclear pore proteins (also on a slower time scale). We propose that the changes in NPC phosphorylation also contribute to regulation of cell motility by controlling the biosynthesis of proteins involved in cytoskeleton rearrangements and adhesion. However, this remains to be tested and the detailed mechanisms will need to be determined. In any case, the phosphorylation changes of actin cytoskeleton proteins enable the cell to start detaching, losing tight junction (Step 2) and to form lamellipodia, which guides cell movement (Step 3). Presumably, the timing of these steps is regulated by dynamically changing the phosphorylation status of actin-related protein and tight junction or adhesion proteins. With increased treatment time, phosphorylation gradually returns to basal levels, resulting in the cell returning to its normal shape and reattachment (Step 4). The scheme presented in Figure 5.14 raises several novel questions that will need to be addressed in future research. These questions, and approaches to answering them, are discussed in the final chapter of this thesis.

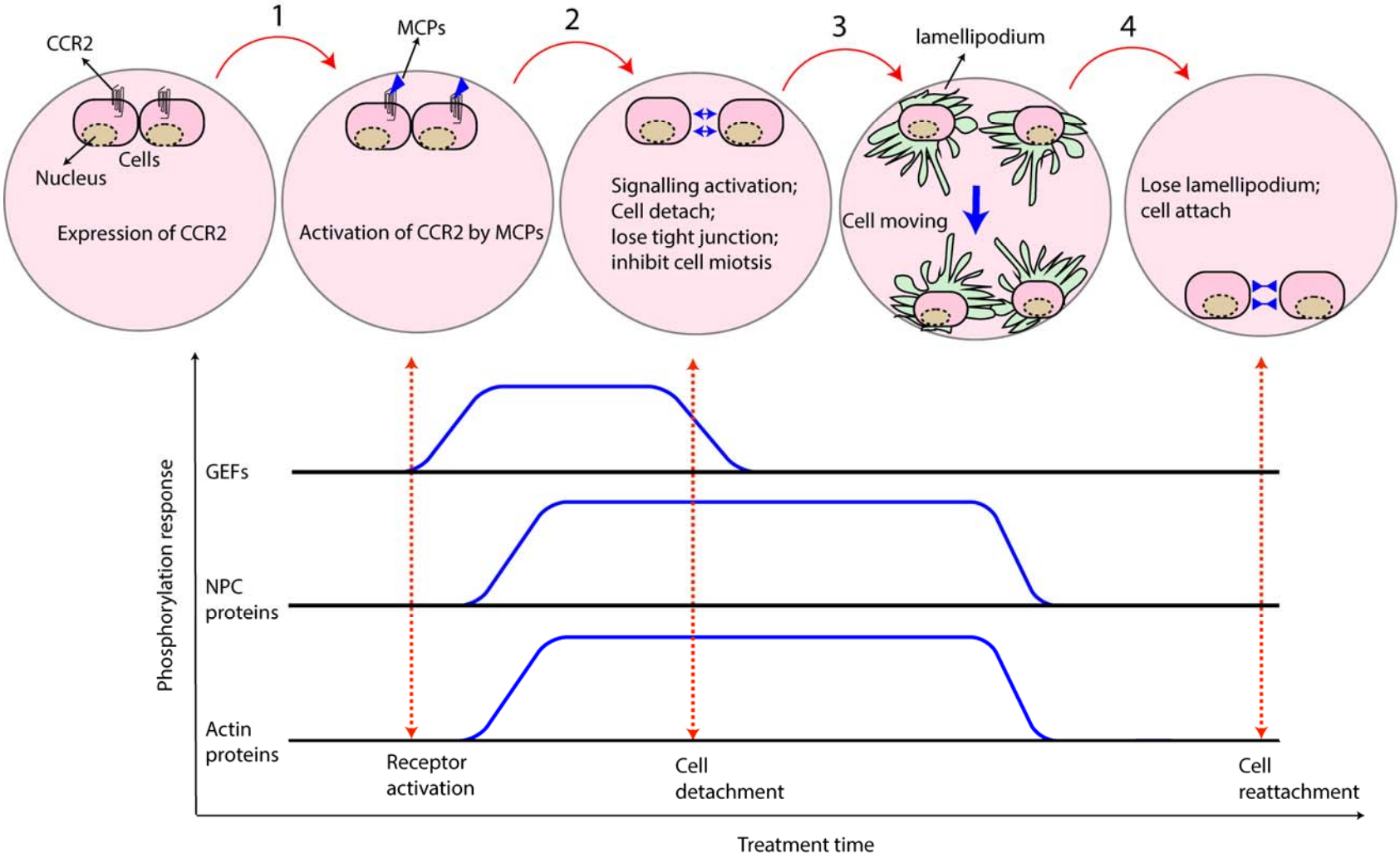


Figure 5.14. A schematic diagram to show proposed mechanism of activation of CCR2 induced cell movement.

Chapter 6. General Discussion

6.1 Thesis summary

Atherosclerosis is a potentially devastating condition and one third of all global deaths are attributed to this disease [343]. It is caused by the accumulation of plaques in arteries and can be asymptomatic for years before the development of severe medical complications such as heart attack or stroke [344]. The plaque is made up of atheroma, cholesterol and calcium. Atheroma in turn is mainly composed of polarised macrophages, which are the core "players" in forming atheromas and thus atherosclerosis [345].

Macrophages are broadly classified into two subtypes based on the stimuli they receive. The subtypes are called M1 and M2 macrophages. The function of M1 macrophages is to promote inflammation, whereas M2 macrophages attenuate inflammation processes. In the first result chapter, Chapter 3, which was submitted for publication recently, we performed a systematic proteomics comparison between non-polarised and M1 or M2 polarised macrophages using label-free quantification. In this study, we identified and quantified more than 6,000 proteins. More importantly, 280 proteins were observed to be significantly different between the various treatment conditions. Out of these 280 proteins, 141 proteins were significantly up- or down-regulated in M1 macrophages, 68 in M2 macrophages and 71 in both M1 and M2 macrophages.

Notably, the closely related proteins IFIT1, IFIT2 and IFIT3 (interferon-induced protein with tetratricopeptide repeats 1, 2 and 3) were observed to be significantly upregulated specifically in the M1 phenotype. The high expression of IFIT1 in M1 macrophages was further validated and confirmed in primary human macrophages as well as in aortic sinus and brachiocephalic artery sections from high fat diet fed ApoE$^{-/-}$ atherosclerotic mice. This implies that IFIT1 (and, by extension, possibly all IFIT proteins) may serve as a new biomarker for M1 macrophages. Moreover, a large number of other antiviral proteins that have previously been identified as products of interferon-stimulated genes (ISGs) were found to be significantly over-expressed in M1 macrophages. After examining the functions of these proteins, a detailed mechanism that can induce the expression of these proteins was proposed.

In contrast, the addition of M2 stimuli caused a relatively weak response in M2 macrophages, although some classical biomarkers were identified. This highlights the need for additional studies to unravel the mechanisms underlying M2 polarization.

As discussed at the beginning of this thesis, chemokines and their receptors are the key players in recruiting monocytes and macrophages to the sites of infection. The monocyte chemoattractant proteins (MCPs) and CC chemokine receptor 2 (CCR2) are the most important chemokine/receptor pair to regulate these processes. However, targeting MCPs or CCR2 has not yet been a successful therapeutic strategy so there is motivation to intervene at alternative points in the biochemical processes that give rise to monocyte recruitment. To better characterise the signalling events that link CCR2 activation to chemotaxis, we performed a large-scale mass spectrometry-based proteomics study to quantify the differences in protein and phosphopeptide levels between untreated cells and cells treated with MCP-1 or MCP-3 (Chapter 4). To the best of our knowledge, this was the first time that the data-independent acquisition (DIA) method was applied to globally quantify the phosphoproteome between multiple conditions. This novel workflow for phosphoproteome and proteome quantification showed excellent reproducibility and accuracy. More importantly, we were not only able to successfully confirm the involvement of known canonical pathways as downstream effectors of MCP/CCR2 signalling (including MAPK, JAK/STAT and Akt/mTOR pathways) but we have also identified, mapped and manually curated additional signalling networks that had not previously been associated with the MCP/CCR2 signalling cascade. In particular, we have described Rho guanine nucleotide exchange factors (ARHGEFs), nuclear pore complex (NPC) proteins and actin cytoskeleton as distinct targets of MCP/CCR2 signalling.

We subsequently confirmed and extended the results of this CCR2 signalling study by analysing the dynamical changes underlying the identified phosphorylation events (Chapter 5). We were able to confirm well-established phosphorylation dynamics of the canonical networks and we provided detailed and novel information on the phosphorylation kinetics of several proteins involved in the ARHGEFs, NPC and actin cytoskeleton networks. Of note, we were also able to link the observed phosphorylation changes to biological effects of the MCP/CCR2 signalling cascade (in particular cell migration) and we further used these datasets to propose theories and hypotheses that might help to understand the MCP/CCR2 signalling cascade in a broader context.

The following sections describe some of the hypotheses that can be proposed by considering the results of our proteomic studies of CCR2 signalling in the context of the

available literature. Potential experimental approaches are suggested to test these hypotheses.

6.2 MCP/CCR2 Signalling: Connecting the Pieces

6.2.1 Connecting ARHGEF7 and ARHGEF26 to MCP/CCR2 signalling

It is known that activation of chemokine receptors by chemokines can lead to phosphorylation and translocation of guanine nucleotide factors (GEFs), which can accelerate the exchange of GDP to GTP on G proteins [330,331], resulting in the activation of signal transduction. In addition to activation of classical G protein Gα, GEFs can also activate other G proteins, such as Rac1 or CDC42, the activation of which is involved in the cell migration pathway [346].

ARHGEFs are members of the GEF family. A direct connection between ARHGEFs and Rac1 or CDC42 was not established in our study, as we did not identify and quantify Rac1 or CDC42. However, it has been recently reported that addition of MCP-1 leads to the phosphorylation of p115RhoGEF (ARHGEF1), which in turn can further activate Rac1 resulting in the migration of vascular smooth muscle cells [333]. The importance of Rac1 in cell migration has also been confirmed by another study, which showed that the formation of F-actin fibres, an essential step for cell migration, is impaired in Rac1 mutant cells following MCP-1 treatment [347]. In addition to Rac1, CDC42 is another G protein that has been implicated as a downstream effector protein of ARHGEFs in chemokine induced cell migration [348,349]. It is therefore highly conceivable that Rac1 and/or Cdc42 are the missing link proteins from ARHGEF26/ARHGEF7 (the 2 GEFs, which have been identified in this study) to the actin cytoskeleton signalling network (illustrated in Figure 6.1).

To verify the importance of these two linker proteins (Rac1 and CDC42) in MCP/CCR2 signalling, the following experiment could be performed. Rac1 inhibitors, NSC23766 [350] and EHT1864 [351], or CDC42 inhibitor, ML141 [352], could be used either to partially inhibit cell migration with a single inhibitor or completely inhibit cell migration with two combined inhibitors. Another option is to use siRNA or CRISPR/CAS9 to knock down or knock out the expression of these two genes one by one or perform a double knock out study. We predict that a single inhibition or knock down/out would not completely inhibit cell migration but would reduce cell mobility; whereas a double inhibition or double knock out study would largely abrogate cell

movement. This proposed experiment would confirm the importance of these two proteins in MCP/CCR2 signalling, but would not indicate how they are activated.

To determine whether the activity of Rac1 or CDC42 is dependent on ARHGEF7 and ARHGEF26 *in vivo*, we propose using the Rac1 Pull-down Activation Assay Biochem Kit (Cytoskeleton Inc.), which is designed to determine the active level of Rac1 and CDC42 in a cell-derived sample. The experiment could be performed as follows. First, the FlpIn-CCR2-HEK293 cells would be either treated with or without MCP-1 and the ratio of non-activated (expected to be higher in vehicle-treated cells) vs activated (expected to be higher in MCP-1-activated cells) Rac1/CDC42 would be determined. Second, this experiment could be repeated with knock down/out or inhibition of either ARHGEF7 or ARHGEF26. We predict the activated Rac1/CDC42 in MCP-1-activated cells would decrease. To further confirm that the activation of Rac1/CDC42 is dependent on ARHGEF7 and ARHGEF26, a recovery study could be added to the second proposed experiment, in which co-transfecting ARHGEF7 and ARHGEF26 into the knock down/out cell line would recover the level of active Rac1/CDC42 in MCP-1-activated cells, while co-transfecting null or mutant ARHGEF7 and ARHGEF26 would not introduce any difference. From this study, we would confirm the activation of Rac1/CDC42 is dependent on ARHGEF7 and ARHGEF26.

Finally, to confirm the direct activation of Rac1/CDC42 by ARHGEF7 and ARHGEF26, another commercial kit (RhoGEF Exchange Assay Kit (Cytoskeleton Inc.)) could potentially be used. This kit uses a fluorescent nucleotide analogue to determine the rate of exchange of nucleotide uptake by GTPases (e.g. Rac1 and CDC42) *in vitro* and it is supplied with purified Rac1 and CDC42. By addition of ARHGEF7 and ARHGEF26 to this kit, one could readily determine whether these two GEFs activate Rac1 and CDC42. From these three complementary experimental approaches, we would be able to ascertain the function of Rac1 and CDC42 in MCP/CCR2 signalling and to determine whether ARHGEF7 and ARHGEF26 is connected to the actin cytoskeleton pathway through Rac1/CDC42.

6.2.2 Connecting NPC to MCP/CCR2 Signalling

One of the most interesting and novel findings in this study is the observed phosphorylation changes of several nuclear pore complex (NPC) proteins in response to MCP-1. The downstream effects of these phosphorylation changes are well described in

the literature to be inhibition of cell mitosis and increased protein biosynthesis, which are essential processes for cell migration [283], However, the molecular mechanisms and kinases that mediate phosphorylation of these NPC proteins, are unknown.

The NPC itself is made up of about 30 nucleoporins (Nups), which can form a giant rigid protein structure that serves as the gateway to shuttle macromolecules between the nucleus and the cytoplasm [353]. There is limited information about the mechanisms of Nup regulation, but phosphorylation events have been proposed to play a pivotal function [354]. In particular, the direct phosphorylation of Nup50 and Nup153 by ERK has been observed to disrupt the interaction between the NPC and importins, which are cargo transporting proteins involved in the translocation of macromolecules into the nucleus [355]. This leads us to hypothesise that some transcription factors (possibly ERK, Jun and/or STAT) that are activated by CCR2 signalling cause the (de)phosphorylation of Nups to promote their own translocation into the nucleus, which is required for them to exert their transcriptional roles [356,357].

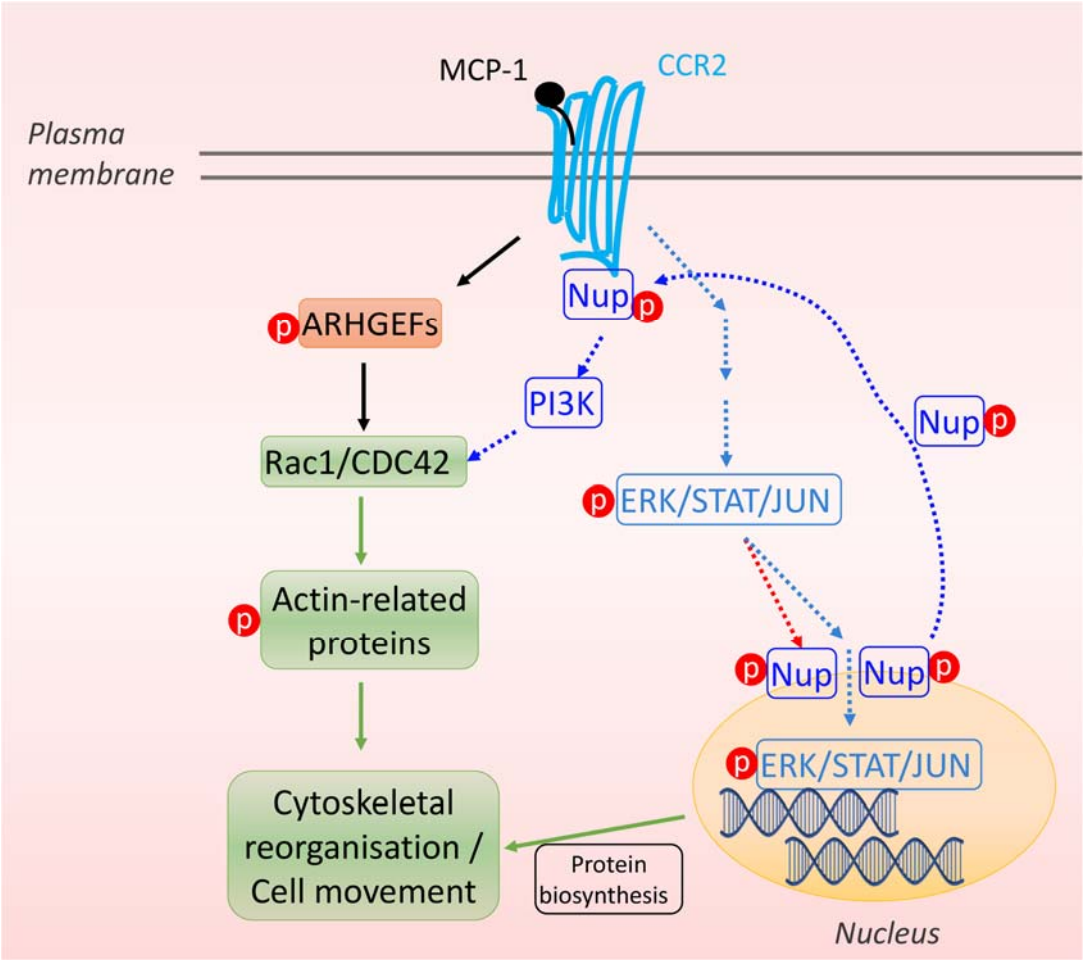


Figure 6.1. A proposed signal transduction network involved in cytoskeletal reorganisation and cell movement following CCR2 activation. Activation of CCR2 by MCP-1 was observed to induce phosphorylation changes (indicated by a red circle containing a white P) of ARHGEFs (*this study*), which possibly act through Rac1 or CDC42 [333] to phosphorylate or dephosphorylate many actin-related proteins (*as observed in this study*) resulting in cytoskeletal reorganisation and cell movement. In addition, activation of CCR2 can phosphorylate many transcriptional factors such as ERK, Stat or Jun (*this study and literature*), which in turn phosphorylate Nups (*observed in this study for CCR2 and literature for other chemokines* [355]). The phosphorylated Nups contribute to protein biosynthesis [283], which is essential for cell migration processes. Alternatively, Nups can relocate to the plasma membrane to interact with CCR2, resulting in activation of Rac1 via PI3K [305].

To test whether the phosphorylated transcriptional factors can induce the phosphorylation of Nups, we propose taking advantage of the powerful phosphoproteomics methods described in this thesis. We already have data in our peptide and phosphopeptide libraries that enable us to determine the relative levels of phosphorylation of many Nups. We could therefore use targeted proteomics, guided by these libraries, to determine the levels of Nup phosphorylation in cell lysates treated exogenously with inactive (non-phosphorylated) or active (phosphorylated) forms of each transcription factor. To enhance the sensitivity of this experiment, it may be advantageous to first enrich the cell lysates with proteins that are potential targets for the transcription factor of interest, which could be achieved using antibodies selective for the known phosphorylation motifs of those transcription factors. For example, Phospho-MAPK Substrates Motif [PXpTP] (14378, Cell Signaling Technology) and Phospho-MAPK/CDK Substrates (PXS*P or S*PXR/K) (2325, Cell Signaling Technology) antibodies can be used to pulldown the substrates of activated ERK. This experiment would demonstrate that the transcription factors are able to phosphorylate the Nups *in vitro* and may indicate the selectivity of different transcription factors for different Nups. To show that these reactions also occur in live cells would require approaches in which the specific transcription factors are inhibited or knocked down, which may be quite challenging considering their central roles in cell viability.

To our knowledge Nups have not previously been observed to be phosphorylated through CCR2 signalling. However, FROUNT, also called Nup85 (not observed in this study), has been mentioned several times to be involved in MCP/CCR2 signalling and CCR2-related cell migration [305,358,359]. FROUNT was first identified by Terashima et al. to bind to the C-terminal region of CCR2 [305]. Additional cell imaging studies confirmed that FROUNT relocated to the membrane of migrating cells after activation of CCR2 by MCP-1 and over-expression of FROUNT promoted the formation of lamellipodia, whereas depletion of FROUNT inhibited macrophage infiltration. Belema-Bedada et al. identified that CCR2-related homing and migration of stem cells was dependant on FROUNT [358]. Toda et al. found that FROUNT can also mediate CCR5-related cell migration. In a step towards understanding how FROUNT mediates CCR2-dependent cell migration, the formation of a complex between CCR2 and FROUNT was found to promote activation of phosphoinositide 3-kinase (PI3K), which can further active Rac1 and contribute to cell migration signalling [305,358]. Based on these

previous observations relating to FROUNT, we propose here that the (de)phosphorylation of other Nups in response to CCR2 activation may also give rise to feedback regulation of the CCR2 signalling network, as indicated by the blue dashed line in Figure 6.1.

Suggested Nups and CCR2 interactions could be tested by affinity-purification mass spectrometry (AP-MS) experiments or in combination with microscopic analyses to track their intracellular distribution. For instance, after the receptor CCR2 being activated by MCP-1, a crosslinking-based AP-MS experiment could be used to determine the binding partners of CCR2. This may have to be performed using a relatively long treatment time or as a time-dependent experiment. After the binding partner of CCR2 is determined, we propose that both Nups and CCR2 could be tagged with different fluorescence labels and the interaction could be visualised and confirmed with live cell imaging analyses or perhaps by a fluorescent resonance energy transfer approach.

Taken together, our datasets in combination with these studies prompted us to hypothesise the following mechanism (Figure 6.1). After activation of CCR2, phosphorylated transcriptional factors such as ERK can dynamically change the phosphorylation levels of several Nups to promote the translocation of these same transcription factors into the nucleus. In addition to their roles as nuclear pore complex components, some of Nups relocate to the membrane to interact with the receptor CCR2 to enhance the signalling cascade, a positive feedback mechanism. Moreover, the CCR2/Nups complex (or complexes) can further activate PI3K, which in turn activates Rac1 and Cdc42, thereby controlling actin cytoskeleton rearrangement and cell migration.

6.3 Targeting Signalling Proteins in MCP/CCR2 Network in Atherosclerosis

The first step during the development of atherosclerosis involves the recruitment of monocytes/macrophages to the site of inflammation. This process is guided and controlled by chemokines and their receptors with MCP chemokines and CCR2 being the most important [360,361]. It has been described that ARHGEFs play an essential role in mediating the signal transduction events by acting as a “switch” [284,362]. Indeed, we observed increased phosphorylation levels of ARHGEF26 and ARHGEF7 after 1.5 minutes of MCP1 treatment, which returned to basal levels after 25 minutes, indicating that the pathway is being turned on then off. We therefore propose that ARHGEFs could

be potentially used as therapeutic targets during atherosclerosis. Conventional chemokine and receptor antagonists, which have been tested so far, target only one chemokine/receptor pair but have failed in clinical trials due to their lack of efficacy. We suspect that the inactivation of only one chemokine or receptor is not sufficient to evoke a complete response as other chemokines and receptors might be able to compensate, which results in only a partially inhibited signal. This problem may be solved by targeting the ARHGEFs instead of the chemokines and receptors, as inhibition of ARHGEFs efficiently terminates GPCR signalling by severely reducing the exchange rate of GDP to GTP. In support of this approach, van Buul et al. [363] investigated the function of ARHGEF26 (also called SGEF) *in vitro*. They observed that overexpression of ARHGEF26 in endothelial cells could induce membrane ruffles, a feature of cell migration. In contrast, silencing of ARHGEF26 can block the trans-migration activity of endothelial cells. To explore the function of ARHGEF26 *in vivo* and to extend its role in atherosclerosis, Samson et al. [364] generated an ARHGEF26 knockout mouse line. When fed with high fat and high cholesterol diet, ARHGEF26 knockout mice showed decreased ability to develop atherosclerosis, indicating a critical function of ARHGEF26 in atherosclerosis and its potential as a therapeutic target. Although this previous study supports the targeting of ARHGEFs in atherosclerosis, it should be noted that different ARHGEFs could be involved in responses to different chemokines/receptors or could be redundant in their responses to the same chemokine/receptor. Thus, this approach could encounter similar lack of efficacy as targeting the chemokines or receptors themselves. Moreover, a side effect of targeting and inhibiting ARHGEFs, which act downstream of numerous receptors, might be that other signalling responses that are indispensable in maintaining normal cellular functions may also be inhibited. Therefore, before pursuing such an approach it will be essential to thoroughly characterise the effects of blocking ARHGEF activity, either individually or in combinations.

In this thesis, we curated an actin regulation network that influences cell migration through the activation of CCR2. Cell migration is an essential step in the development of atherosclerosis in that monocytes/macrophages have to transmigrate from the blood to the site of the atherosclerotic plaque. Thus, it may be also possible to target proteins in this actin-dependent pathway to inhibit or weaken cell migration. In this regard, one of the most interesting proteins that has been described in this thesis is ZO1, a tight junction protein, which contains multiple phosphorylation sites that appear

to be regulated by CCR2. Strikingly, we also observed changes in the phosphorylation states of six downstream targets of ZO1 in response to MCP-1/CCR2 activation. Therefore, ZO1 could be another potential therapeutic target in atherosclerosis and indeed, the importance of ZO1 in atherosclerosis has been implicated in other studies [365,366]. However, Katsuno et al. [367] observed that ZO1 is an essential protein, which leads to lethality in mice when fully knocked out. Thus, instead of completely inhibiting ZO1, it may be preferable to target upstream kinases involved in phosphorylation of specific phosphorylation sites within ZO1. Alternatively, blocking some of the downstream targets of ZO1 may be a more effective way to regulate cell migration activity without drastic side effects. Once again, more research is needed to understand the roles of individual proteins and phosphorylation events in this network and then to select the most appropriate targets for inhibition.

In addition to all these targeted approaches, other omics techniques such as genomics or transcriptomics could also be exploited to study the effect of MCP-1 on cells and gain more insight into the cellular responses by integrative, cross-omics comparisons.

6.4 Conclusion

Our global quantitative proteomics and phosphoproteomics approach provided a high-quality dataset to further understand CCR2-mediated signalling pathways, which play a critical role in atherosclerosis. By mining this data, we confirmed known cellular effects of CCR2 activation, but more importantly, we generated several new hypotheses based on our findings. In summary, we proposed the following hypotheses: (1) specific phosphorylation changes of actin-related proteins, such as ZO1, control cytoskeleton reorganisation and cell migration; (2) phosphorylation changes of ARHGEFs control the signalling cascade by acting as “switches” between the on and off state; (3) Rac1 and CDC42 are essential links between phosphorylation of the ARHGEFs identified here and phosphorylation of actin cytoskeleton proteins; (4) phosphorylation changes of NPC proteins contribute to an increase in protein biosynthesis and to inhibition of cell mitosis, which together promote cell movement; and (5) Nups are involved in a feedback regulation of CCR2.

For reasons of practicality, the phosphoproteomics study describe here was performed using the HEK293 cell line, which is an embryonic kidney cell line. To ensure

that the observed changes are not only cell-line specific responses, follow-up studies, using either proteomics or the other approaches discussed above, should be performed with different cell lines such as macrophage cell lines or ideally primary human cells.

In summary, this thesis contributes significantly to our understanding of the MCP/CCR2 signalling network. It is evident that this signalling cascade is highly complex and a plethora of additional studies still have to be conducted to fully unravel its complexities. Nevertheless, this study provides a solid basis for the generation and testing of new hypotheses that may ultimately contribute to the development of new therapeutic approaches towards atherosclerosis and other macrophage-driven inflammatory diseases.

References

1. Mathers, C.D. and D. Loncar, Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med., 2006. 3(11): e442.
2. Hansson, G.K., A.K. Robertson, and C. Soderberg-Naucler, Inflammation and atherosclerosis. Annu. Rev. Pathol., 2006. 1: 297-329.
3. Catherman, A.D., O.S. Skinner, and N.L. Kelleher, Top Down proteomics: Facts and perspectives. Biochem. Biophys. Res. Commun., 2014. 445(4): 683-693.
4. ILenz, C. and H. Dihazi, Introduction to Proteomics Technologies, in Statistical Analysis in Proteomics, K. Jung, Editor. 2016, Springer New York: New York, NY. 3-27.
5. Frostegard, J., Immunity, atherosclerosis and cardiovascular disease. BMC Med., 2013. 11: 117.
6. Ross, R., The pathogenesis of atherosclerosis: a perspective for the 1990s. Nature, 1993. 362(6423): 801-9.
7. David, A.R., A. Kershaw, and A. Heagerty, Atherosclerosis and diet in ancient Egypt. Lancet, 2010. 375(9716): 718-9.
8. Murphy, W.A., Jr., et al., The iceman: discovery and imaging. Radiology, 2003. 226(3): 614-29.
9. Allam, A.H., et al., Computed tomographic assessment of atherosclerosis in ancient Egyptian mummies. JAMA, 2009. 302(19): 2091-4.
10. Allam, A.H., et al., Atherosclerosis in ancient Egyptian mummies: the Horus study. JACC Cardiovasc. Imaging, 2011. 4(4): 315-27.
11. Thompson, R.C., et al., Atherosclerosis across 4000 years of human history: the Horus study of four ancient populations. Lancet, 2013. 381(9873): 1211-22.
12. Tabas, I., Macrophage death and defective inflammation resolution in atherosclerosis. Nat. Rev. Immunol., 2010. 10(1): 36-46.
13. Poole, J.C. and H.W. Florey, Changes in the endothelium of the aorta and the behaviour of macrophages in experimental atheroma of rabbits. J. Pathol. Bacteriol., 1958. 75(2): 245-51.
14. Ross, R. and L. Harker, Hyperlipidemia and atherosclerosis. Science, 1976. 193(4258): 1094-100.
15. Libby, P., Inflammation in atherosclerosis. Nature, 2002. 420(6917): 868-74.
16. Hopkins, P.N., Molecular biology of atherosclerosis. Physiol. Rev., 2013. 93(3): 1317-542.
17. Lusis, A.J., Atherosclerosis. Nature, 2000. 407(6801): 233-41.
18. Hansson, G.K. and P. Libby, The immune response in atherosclerosis: a double-edged sword. Nat. Rev. Immunol., 2006. 6(7): 508-19.
19. Bentzon, J.F., et al., Mechanisms of plaque formation and rupture. Circ. Res., 2014. 114(12): 1852-1866.
20. Woollard, K.J. and F. Geissmann, Monocytes in atherosclerosis: subsets and functions. Nat. Rev. Cardiol., 2010. 7(2): 77-86.
21. Mosser, D.M. and J.P. Edwards, Exploring the full spectrum of macrophage activation. Nat. Rev. Immunol., 2008. 8(12): 958-69.
22. Mantovani, A., C. Garlanda, and M. Locati, Macrophage diversity and polarization in atherosclerosis: a question of balance. Arterioscler. Thromb. Vasc. Biol., 2009. 29(10): 1419-23.
23. Wang, N., H. Liang, and K. Zen, Molecular mechanisms that influence the macrophage M1-M2 polarization balance. Front. Immunol., 2014. 5: 614.

24. Stone, M., et al., Mechanisms of regulation of the chemokine-receptor network. Int. J. Mol. Sci., 2017. 18(2): 342.
25. Moser, B., et al., Chemokines: multiple levels of leukocyte migration control. Trends Immunol., 2004. 25(2): 75-84.
26. Weber, C. and H. Noels, Atherosclerosis: current pathogenesis and therapeutic options. Nat. Med., 2011. 17(11): 1410-22.
27. Colin, S., G. Chinetti-Gbaguidi, and B. Staels, Macrophage phenotypes in atherosclerosis. Immunol. Rev., 2014. 262(1): 153-66.
28. Gordon, S. and F.O. Martinez, Alternative activation of macrophages: mechanism and functions. Immunity, 2010. 32(5): 593-604.
29. Stein, M., et al., Interleukin 4 potently enhances murine macrophage mannose receptor activity: a marker of alternative immunologic macrophage activation. J. Exp. Med., 1992. 176(1): 287-92.
30. Raes, G., et al., Arginase-1 and Ym1 are markers for murine, but not human, alternatively activated myeloid cells. J. Immunol., 2005. 174(11): 6561; author reply 6561-2.
31. Martinez, F.O., et al., Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: new molecules and patterns of gene expression. J. Immunol., 2006. 177(10): 7303-11.
32. Sica, A. and A. Mantovani, Macrophage plasticity and polarization: in vivo veritas. J. Clin. Invest., 2012. 122(3): 787-95.
33. Murray, P.J., et al., Macrophage activation and polarization: nomenclature and experimental guidelines. Immunity, 2014. 41(1): 14-20.
34. Becker, M., et al., Integratedt transcriptomics establish macrophage polarization signatures and have potential applications for clinical health and disease. Sci. Rep., 2015. 5: 13351.
35. Beyer, M., et al., High-resolution transcriptome of human macrophages. PLoS One, 2012. 7(9): e45466.
36. Brown, J., et al., Proteome bioprofiles distinguish between M1 priming and activation states in human macrophages. J. Leukoc. Biol., 2010. 87(4): 655-62.
37. Becker, L., et al., Unique proteomic signatures distinguish macrophages and dendritic cells. PLoS One, 2012. 7(3): e33297.
38. Zhang, F., et al., Changes in the proteomic profile during the differential polarization status of the human monocyte-derived macrophage THP-1 cell line. Proteomics, 2015. 15(4): 773-86.
39. Meijer, K., et al., Quantitative proteomics analyses of activation states of human THP-1 macrophages. J. Proteomics, 2015. 128: 164-72.
40. Baggiolini, M., B. Dewald, and B. Moser, Human chemokines: an update. Annu. Rev. Immunol., 1997. 15: 675-705.
41. Baggiolini, M., B. Dewald, and B. Moser, Interleukin-8 and related chemotactic cytokines--CXC and CC chemokines. Adv. Immunol., 1994. 55: 97-179.
42. Domanska, U.M., et al., The chemokine network, a newly discovered target in high grade gliomas. Crit. Rev. Oncol. Hematol., 2011. 79(2): 154-63.
43. O'Hayre, M., et al., Chemokines and cancer: migration, intracellular signalling and intercellular communication in the microenvironment. Biochem. J., 2008. 409(3): 635-49.
44. Bajetto, A., et al., Characterization of chemokines and their receptors in the central nervous system: physiopathological implications. J. Neurochem., 2002. 82(6): 1311-29.

45. Raman, D., T. Sobolik-Delmaire, and A. Richmond, Chemokines in health and disease. Exp. Cell Res., 2011. 317(5): 575-89.
46. Yla-Herttuala, S., et al., Expression of monocyte chemoattractant protein 1 in macrophage-rich areas of human and rabbit atherosclerotic lesions. Proc. Natl. Acad. Sci. U. S. A., 1991. 88(12): 5252-6.
47. Soria, G. and A. Ben-Baruch, The inflammatory chemokines CCL2 and CCL5 in breast cancer. Cancer Lett., 2008. 267(2): 271-85.
48. Niwa, Y., et al., Correlation of tissue and plasma RANTES levels with disease course in patients with breast or cervical cancer. Clin. Cancer Res., 2001. 7(2): 285-9.
49. Ueno, T., et al., Significance of macrophage chemoattractant protein-1 in macrophage recruitment, angiogenesis, and survival in human breast cancer. Clin. Cancer Res., 2000. 6(8): 3282-9.
50. Zlotnik, A., A.M. Burkhardt, and B. Homey, Homeostatic chemokine receptors and organ-specific metastasis. Nat. Rev. Immunol., 2011. 11(9): 597-606.
51. Vicari, A.P., et al., TECK: a novel CC chemokine specifically expressed by thymic dendritic cells and potentially involved in T cell development. Immunity, 1997. 7(2): 291-301.
52. Takada, K. and S.C. Jameson, Naive T cell homeostasis: from awareness of space to a sense of place. Nat. Rev. Immunol., 2009. 9(12): 823-32.
53. Link, A., et al., Fibroblastic reticular cells in lymph nodes regulate the homeostasis of naive T cells. Nat. Immunol., 2007. 8(11): 1255-65.
54. Johnson, L.A. and D.G. Jackson, Inflammation-induced secretion of CCL21 in lymphatic endothelium is a key regulator of integrin-mediated dendritic cell transmigration. Int. Immunol., 2010. 22(10): 839-49.
55. Richter, J.R., et al., Macrophage-derived chemokine (CCL22) is a novel mediator of lung inflammation following hemorrhage and resuscitation. Shock, 2014. 42(6): 525-31.
56. Damas, J.K., et al., Homeostatic chemokines CCL19 and CCL21 promote inflammation in human immunodeficiency virus-infected patients with ongoing viral replication. Clin. Exp. Immunol., 2009. 157(3): 400-7.
57. Fredriksson, R., et al., The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol. Pharmacol., 2003. 63(6): 1256-72.
58. Lander, E.S., et al., Initial sequencing and analysis of the human genome. Nature, 2001. 409(6822): 860-921.
59. Murphy, P.M., et al., International union of pharmacology. XXII. Nomenclature for chemokine receptors. Pharmacol. Rev., 2000. 52(1): 145-76.
60. Allen, S.J., S.E. Crown, and T.M. Handel, Chemokine: receptor structure, interactions, and antagonism. Annu. Rev. Immunol., 2007. 25: 787-820.
61. Kobilka, B.K., G protein coupled receptor structure and activation. Biochim. Biophys. Acta, 2007. 1768(4): 794-807.
62. Oppermann, M., Chemokine receptor CCR5: insights into structure, function, and regulation. Cell. Signal., 2004. 16(11): 1201-10.
63. Zheng, Y., et al., Structure of CC chemokine receptor 2 with orthosteric and allosteric antagonists. Nature, 2016. 540(7633): 458-461.
64. Patel, J., K.M. Channon, and E. McNeill, The downstream regulation of chemokine receptor signalling: implications for atherosclerosis. Mediators Inflamm., 2013. 2013: 12.

65. Murdoch, C. and A. Finn, Chemokine receptors and their role in inflammation and infectious diseases. Blood, 2000. 95(10): 3032-43.
66. Mellado, M., et al., Chemokine control of HIV-1 infection. Nature, 1999. 400(6746): 723-4.
67. Scholten, D.J., et al., Pharmacological modulation of chemokine receptor function. Br. J. Pharmacol., 2012. 165(6): 1617-1643.
68. Proudfoot, A.E., Chemokine receptors: multifaceted therapeutic targets. Nat. Rev. Immunol., 2002. 2(2): 106-15.
69. O'Hayre, M., et al., Emerging concepts and approaches for chemokine-receptor drug discovery. Expert Opin Drug Discov, 2010. 5(11): 1109-22.
70. Kouroumalis, A., et al., The chemokines CXCL9, CXCL10, and CXCL11 differentially stimulate G alpha i-independent signaling and actin responses in human intestinal myofibroblasts. J. Immunol., 2005. 175(8): 5403-11.
71. Berchiche, Y.A., et al., Different effects of the different natural CC chemokine receptor 2b ligands on beta-arrestin recruitment, Galphai signaling, and receptor internalization. Mol. Pharmacol., 2011. 79(3): 488-98.
72. Rajagopal, S., et al., Beta-arrestin- but not G protein-mediated signaling by the "decoy" receptor CXCR7. Proc. Natl. Acad. Sci. U. S. A., 2010. 107(2): 628-32.
73. Leach, K., S.J. Charlton, and P.G. Strange, Analysis of second messenger pathways stimulated by different chemokines acting at the chemokine receptor CCR5. Biochem. Pharmacol., 2007. 74(6): 881-90.
74. Tian, Y., et al., Differential chemokine activation of CC chemokine receptor 1-regulated pathways: ligand selective activation of Galpha 14-coupled pathways. Eur. J. Immunol., 2004. 34(3): 785-795.
75. Oppermann, M., et al., Differential effects of CC chemokines on CC chemokine receptor 5 (CCR5) phosphorylation and identification of phosphorylation sites on the CCR5 carboxyl terminus. J. Biol. Chem., 1999. 274(13): 8875-85.
76. Rajagopal, S., et al., Biased agonism as a mechanism for differential signaling by chemokine receptors. J. Biol. Chem., 2013. 288(49): 35039-48.
77. Charo, I.F., et al., Molecular cloning and functional expression of two monocyte chemoattractant protein 1 receptors reveals alternative splicing of the carboxyl-terminal tails. Proc. Natl. Acad. Sci. U. S. A., 1994. 91(7): 2752-6.
78. Yamasaki, R., et al., Role of CCR2 in immunobiology and neurobiology. Clin. Exp. Neuroimmuno., 2012. 3(1): 16-29.
79. Feria, M. and F. Díaz-González, The CCR2 receptor as a therapeutic target. Expert Opin. Ther. Pat., 2006. 16(1): 49-57.
80. Yadav, A., V. Saini, and S. Arora, MCP-1: chemoattractant with a role beyond immunity: a review. Clin. Chim. Acta, 2010. 411(21-22): 1570-9.
81. Jones, K.L., J.J. Maguire, and A.P. Davenport, Chemokine receptor CCR5: from AIDS to atherosclerosis. Br. J. Pharmacol., 2011. 162(7): 1453-69.
82. Szczucinski, A. and J. Losy, Chemokines and chemokine receptors in multiple sclerosis. Potential targets for new therapies. Acta Neurol. Scand., 2007. 115(3): 137-46.
83. Solomon, M., B. Balasa, and N. Sarvetnick, CCR2 and CCR5 chemokine receptors differentially influence the development of autoimmune diabetes in the NOD mouse. Autoimmunity, 2010. 43(2): 156-63.
84. Aiello, R.J., et al., Monocyte chemoattractant protein-1 accelerates atherosclerosis in apolipoprotein E-deficient mice. Arterioscler. Thromb. Vasc. Biol., 1999. 19(6): 1518-25.

85. Gosling, J., et al., MCP-1 deficiency reduces susceptibility to atherosclerosis in mice that overexpress human apolipoprotein B. J. Clin. Invest., 1999. 103(6): 773-8.
86. An, S.J., et al., Functions of monocyte chemotactic protein-3 in transgenic mice fed a high-fat, high-cholesterol diet. J. Microbiol. Biotechnol., 2013. 23(3): 405-13.
87. Breland, U.M., et al., Raised MCP-4 levels in symptomatic carotid atherosclerosis: an inflammatory link between platelet and monocyte activation. Cardiovasc. Res., 2010. 86(2): 265-73.
88. Van Damme, J., et al., Structural and functional identification of two human, tumor-derived monocyte chemotactic proteins (MCP-2 and MCP-3) belonging to the chemokine family. J. Exp. Med., 1992. 176(1): 59-65.
89. Lutgens, E., et al., Gene profiling in atherosclerosis reveals a key role for small inducible cytokines: validation using a novel monocyte chemoattractant protein monoclonal antibody. Circulation, 2005. 111(25): 3443-52.
90. Boring, L., et al., Decreased lesion formation in CCR2-/- mice reveals a role for chemokines in the initiation of atherosclerosis. Nature, 1998. 394(6696): 894-7.
91. Bot, I., et al., A novel CCR2 antagonist inhibits atherogenesis in apoE deficient mice by achieving high receptor occupancy. Sci. Rep., 2017. 7(1): 52.
92. Gong, J.H., et al., An antagonist of monocyte chemoattractant protein 1 (MCP-1) inhibits arthritis in the MRL-lpr mouse model. J. Exp. Med., 1997. 186(1): 131-7.
93. Gilbert, J., et al., Effect of CC chemokine receptor 2 CCR2 blockade on serum C-reactive protein in individuals at atherosclerotic risk and with a single nucleotide polymorphism of the monocyte chemoattractant protein-1 promoter region. Am. J. Cardiol., 2011. 107(6): 906-11.
94. Struthers, M. and A. Pasternak, CCR2 antagonists. Curr. Top. Med. Chem., 2010. 10(13): 1278-98.
95. Biswas, S.K. and A. Sodhi, Tyrosine phosphorylation-mediated signal transduction in MCP-1-induced macrophage activation: role for receptor dimerization, focal adhesion protein complex and JAK/STAT pathway. Int. Immunopharmacol., 2002. 2(8): 1095-1107.
96. Wilkins, M.R., et al., Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. Biotechnol. Genet. Eng. Rev., 1996. 13: 19-50.
97. Wasinger, V.C., et al., Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium. Electrophoresis, 1995. 16(7): 1090-4.
98. James, P., Protein identification in the post-genome era: the rapid rise of proteomics. Q. Rev. Biophys., 1997. 30(4): 279-331.
99. Tyers, M. and M. Mann, From genomics to proteomics. Nature, 2003. 422(6928): 193-197.
100. Goffeau, A., et al., Life with 6000 genes. Science, 1996. 274(5287): 546, 563-7.
101. Fleischmann, R.D., et al., Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science, 1995. 269(5223): 496-512.
102. International Human Genome Sequencing, C., Finishing the euchromatic sequence of the human genome. Nature, 2004. 431(7011): 931-45.
103. Kim, M.-S., et al., A draft map of the human proteome. Nature, 2014. 509(7502): 575-581.
104. Wilhelm, M., et al., Mass-spectrometry-based draft of the human proteome. Nature, 2014. 509(7502): 582-587.

105. Mayne, J., et al., Bottom-Up Proteomics (2013–2015): Keeping up in the Era of Systems Biology. Anal. Chem., 2016. 88(1): 95-121.
106. Edman, P., A method for the determination of amino acid sequence in peptides. Arch Biochem, 1949. 22(3): 475.
107. Edman, P. and G. Begg, A protein sequenator. Eur. J. Biochem., 1967. 1(1): 80-91.
108. O'Farrell, P.H., High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem., 1975. 250(10): 4007-21.
109. Aebersold, R.H., et al., Electroblotting onto activated glass. High efficiency preparation of proteins from analytical sodium dodecyl sulfate-polyacrylamide gels for direct sequence analysis. J. Biol. Chem., 1986. 261(9): 4229-38.
110. Barber, M., et al., Fast atom bombardment of solids (F.A.B.): a new ion source for mass spectrometry. J. Chem. Soc. Chem. Commun., 1981(7): 325-327.
111. Fenn, J., et al., Electrospray ionization for mass spectrometry of large biomolecules. Science, 1989. 246(4926): 64-71.
112. Tanaka, K., et al., Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom., 1988. 2(8): 151-153.
113. Henzel, W.J., C. Watanabe, and J.T. Stults, Protein identification: the origins of peptide mass fingerprinting. J. Am. Soc. Mass Spectrom., 2003. 14(9): 931-942.
114. Smith, L.M., N.L. Kelleher, and P. Consortium for Top Down, Proteoform: a single term describing protein complexity. Nat. Methods, 2013. 10(3): 186-7.
115. Siuti, N. and N.L. Kelleher, Decoding protein modifications using top-down mass spectrometry. Nat. Methods, 2007. 4(10): 817-21.
116. Tran, J.C., et al., Mapping intact protein isoforms in discovery mode using top-down proteomics. Nature, 2011. 480(7376): 254-258.
117. Ntai, I., et al., Integrated Bottom-Up and Top-Down Proteomics of Patient-Derived Breast Tumor Xenografts. Mol. Cell. Proteomics, 2016. 15(1): 45-56.
118. Lee, A.E., et al., Preparing to read the ubiquitin code: characterization of ubiquitin trimers by top-down mass spectrometry. J. Mass Spectrom., 2016. 51(4): 315-321.
119. Ntai, I., et al., A Method for Label-Free, Differential Top-Down Proteomics, in Quantitative Proteomics by Mass Spectrometry, S. Sechi, Editor. 2016, Springer New York: New York, NY. 121-133.
120. Zhang, Y., et al., Protein Analysis by Shotgun/Bottom-up Proteomics. Chem. Rev., 2013. 113(4): 2343-2394.
121. Gillet, L.C., A. Leitner, and R. Aebersold, Mass spectrometry applied to bottom-up proteomics: entering the high-throughput era for typothesis testing. Annu. Rev. Anal. Chem. (Palo Alto Calif.), 2016. 9(1): null.
122. Olsen, J.V., S.-E. Ong, and M. Mann, Trypsin Cleaves Exclusively C-terminal to Arginine and Lysine Residues. Mol. Cell. Proteomics, 2004. 3(6): 608-614.
123. Hubbard, M.J. and P. Cohen, On target with a new mechanism for the regulation of protein phosphorylation. Trends in biochemical sciences, 1993. 18(5): 172-7.
124. Manning, G., et al., The protein kinase complement of the human genome. Science, 2002. 298(5600): 1912-34.
125. Engholm-Keller, K. and M.R. Larsen, Technologies and challenges in large-scale phosphoproteomics. Proteomics, 2013. 13(6): 910-31.
126. Ciesla, J., T. Fraczyk, and W. Rode, Phosphorylation of basic amino acid residues in proteins: important but easily missed. Acta Biochim. Pol., 2011. 58(2): 137-48.

127. Iliuk, A.B., J.V. Arrington, and W.A. Tao, Analytical challenges translating mass spectrometry-based phosphoproteomics from discovery to clinical applications. Electrophoresis, 2014. 35(24): 3430-40.
128. Ubersax, J.A. and J.E. Ferrell, Jr., Mechanisms of specificity in protein phosphorylation. Nat. Rev. Mol. Cell Biol., 2007. 8(7): 530-41.
129. Peck, S.C., Analysis of protein phosphorylation: methods and strategies for studying kinases and substrates. Plant J., 2006. 45(4): 512-22.
130. de Graauw, M., P. Hensbergen, and B. van de Water, Phospho-proteomic analysis of cellular signaling. Electrophoresis, 2006. 27(13): 2676-86.
131. Bulayeva, N.N., B. Gametchu, and C.S. Watson, Quantitative measurement of estrogen-induced ERK 1 and 2 activation via multiple membrane-initiated signaling pathways. Steroids, 2004. 69(3): 181-92.
132. Huang, P.H. and F.M. White, Phosphoproteomics: unraveling the signaling web. Mol. Cell, 2008. 31(6): 777-81.
133. Macek, B., M. Mann, and J.V. Olsen, Global and site-specific quantitative phosphoproteomics: principles and applications. Annu. Rev. Pharmacol. Toxicol., 2009. 49: 199-221.
134. Sharma, K., et al., Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. Cell Rep., 2014. 8(5): 1583-94.
135. Francavilla, C., et al., Phosphoproteomics of primary cells reveals druggable kinase signatures in ovarian cancer. Cell Rep., 2017. 18(13): 3242-3256.
136. Kim, J.Y., et al., Dissection of TBK1 signaling via phosphoproteomics in lung cancer cells. Proc. Natl. Acad. Sci. U. S. A., 2013. 110(30): 12414-9.
137. Thingholm, T.E. and M.R. Larsen, Sequential elution from IMAC (SIMAC): an efficient method for enrichment and separation of mono- and Mmlti-phosphorylated peptides. Methods Mol. Biol., 2016. 1355: 147-60.
138. de Graaf, E.L., et al., Single-step enrichment by Ti4+-IMAC and label-free quantitation enables in-depth monitoring of phosphorylation dynamics with high reproducibility and temporal resolution. Mol. Cell. Proteomics, 2014. 13(9): 2426-34.
139. Camerini, S. and P. Mauri, The role of protein and peptide separation before mass spectrometry analysis in clinical proteomics. J. Chromatogr. A, 2015. 1381: 1-12.
140. Chan, K.C. and H.J. Issaq, Fractionation of peptides by strong cation-exchange liquid chromatography. Methods Mol. Biol., 2013. 1002: 311-5.
141. Yang, F., et al., High-pH reversed-phase chromatography with fraction concatenation for 2D proteomic analysis. Expert Rev Proteomics, 2012. 9(2): 129-34.
142. Villen, J. and S.P. Gygi, The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. Nat. Protoc., 2008. 3(10): 1630-8.
143. Thingholm, T.E., et al., Highly selective enrichment of phosphorylated peptides using titanium dioxide. Nat. Protoc., 2006. 1(4): 1929-35.
144. Yang, C., X. Zhong, and L. Li, Recent advances in enrichment and separation strategies for mass spectrometry-based phosphoproteomics. Electrophoresis, 2014.
145. Negroni, L., et al., Comparison of IMAC and MOAC for phosphopeptide enrichment by column chromatography. J. Chromatogr. B Analyt. Technol. Biomed. Life Sci., 2012. 891-892: 109-12.

146. Zhou, H., et al., Robust phosphoproteome enrichment using monodisperse microsphere-based immobilized titanium (IV) ion affinity chromatography. Nat. Protoc., 2013. 8(3): 461-80.
147. Zhou, H., et al., Enhancing the identification of phosphopeptides from putative basophilic kinase substrates using Ti (IV) based IMAC enrichment. Mol. Cell. Proteomics, 2011. 10(10): M110 006452.
148. Yu, Z., et al., Preparation of monodisperse immobilized Ti(4+) affinity chromatography microspheres for specific enrichment of phosphopeptides. Anal. Chim. Acta, 2009. 636(1): 34-41.
149. Iliuk, A.B., et al., In-depth analyses of kinase-dependent tyrosine phosphoproteomes based on metal ion-functionalized soluble nanopolymers. Mol. Cell. Proteomics, 2010. 9(10): 2162-72.
150. Radu, M., et al., ArhGAP15, a Rac-specific GTPase-activating protein, plays a dual role in inhibiting small GTPase signaling. J. Biol. Chem., 2013. 288(29): 21117-25.
151. Walls, C.D., et al., Phosphatase of regenerating liver 3 (PRL3) provokes a tyrosine phosphoproteome to drive prometastatic signal transduction. Mol. Cell. Proteomics, 2013. 12(12): 3759-77.
152. Xue, L., et al., Sensitive kinase assay linked with phosphoproteomics for identifying direct kinase substrates. Proc. Natl. Acad. Sci. U. S. A., 2012. 109(15): 5615-20.
153. Yu, S., et al., Syk inhibits the activity of protein kinase A by phosphorylating tyrosine 330 of the catalytic subunit. J. Biol. Chem., 2013. 288(15): 10870-81.
154. Aryal, U.K. and A.R.S. Ross, Enrichment and analysis of phosphopeptides under different experimental conditions using titanium dioxide affinity chromatography and mass spectrometry. Rapid Commun. Mass Spectrom., 2010. 24(2): 219-231.
155. Pinkse, M.W., et al., Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. Anal. Chem., 2004. 76(14): 3935-43.
156. Thingholm, T.E., et al., SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides. Mol. Cell. Proteomics, 2008. 7(4): 661-71.
157. Rogers, L.D. and L.J. Foster, Phosphoproteomics--finally fulfilling the promise? Mol. Biosyst., 2009. 5(10): 1122-9.
158. Yates, J.R., C.I. Ruse, and A. Nakorchevsky, Proteomics by mass spectrometry: approaches, advances, and applications. Annu Rev Biomed Eng, 2009. 11: 49-79.
159. Markides, K. and A. Gräslund, Advanced information on the Nobel Prize in Chemistry 2002. The Royal Swedish Academy of Sciences, 2002: 1–13.
160. Singhal, N., et al., MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. Front. Microbiol., 2015. 6: 791.
161. Hines, H.B., Microbial Proteomics Using Mass Spectrometry, in Microbial Systems Biology: Methods and Protocols, A. Navid, Editor. 2012, Humana Press: Totowa, NJ. 159-186.
162. Ho, C.S., et al., Electrospray ionisation mass spectrometry: principles and clinical applications. Clin. Biochem. Rev., 2003. 24(1): 3-12.
163. Thomson, B.A. and J.V. Iribarne, Field induced ion evaporation from liquid surfaces at atmospheric pressure. J. Chem. Phys., 1979. 71(11): 4451-4463.

164. Iribarne, J.V. and B.A. Thomson, On the evaporation of small ions from charged droplets. J. Chem. Phys., 1976. 64(6): 2287-2294.
165. Wilm, M., Principles of electrospray ionization. Mol. Cell. Proteomics, 2011.
166. Aebersold, R. and M. Mann, Mass spectrometry-based proteomics. Nature, 2003. 422(6928): 198-207.
167. Hill, L., https://www.jic.ac.uk/services/metabolomics/topics/lcms/single1.htm.
168. Yost, R.A. and C.G. Enke, Selected ion fragmentation with a tandem quadrupole mass spectrometer. J. Am. Chem. Soc., 1978. 100(7): 2274-2275.
169. Morrison, J.D., Personal reminiscences of forty years of mass spectrometry in Australia. Org. Mass Spectrom., 1991. 26(4): 183-194.
170. Stephens, W.E., A pulsed mass spectrometer with time dispersion. Phys. Rev., 1946. 69(11-12): 691.
171. Cameron, A.E. and D.F. Eggers, An ion "Velocitron". Rev. Sci. Instrum., 1948. 19(9): 605-607.
172. Wiley, W.C. and I.H. McLaren, Time - of - flight mass spectrometer with improved resolution. Rev. Sci. Instrum., 1955. 26(12): 1150-1157.
173. Edmond, d.H. and S. Vincent, Mass spectrometry principles and applications. 2007: 1-502.
174. Doroshenko, V.M. and R.J. Cotter, Ideal velocity focusing in a reflectron time-of-flight mass spectrometer. J. Am. Soc. Mass Spectrom., 1999. 10(10): 992-999.
175. Alikhanov, S.G., A new impulse technique for ion mass measurements Sov. Phys. JETP 1957. 4: 452-4533.
176. Mamyrin, B.A., et al., The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution Sov. Phys. JETP, 1973. 37: 45-48.
177. Mamyrin, B.A., Time-of-flight mass spectrometry (concepts, achievements, and prospects). Int. J. Mass Spectrom., 2001. 206(3): 251-266.
178. Fjeldsted, J.C., Chapter 2 - Advances in Time-of-Flight Mass Spectrometry, in Comprehensive Analytical Chemistry, P.E. Sandra Pérez and B. Damià, Editors. 2016, Elsevier. 19-49.
179. Hager, J.W., A new linear ion trap mass spectrometer. Rapid Commun. Mass Spectrom., 2002. 16(6): 512-526.
180. Douglas, D.J., A.J. Frank, and D. Mao, Linear ion traps in mass spectrometry. Mass Spectrom. Rev., 2005. 24(1): 1-29.
181. Blackler, A.R., et al., Quantitative comparison of proteomic data quality between a 2D and 3D quadrupole ion trap. Anal. Chem., 2006. 78(4): 1337-1344.
182. Marshall, A.G. and C.L. Hendrickson, High-resolution mass spectrometers. Annu. Rev. Anal. Chem. (Palo Alto Calif.), 2008. 1: 579-99.
183. Scigelova, M., et al., Fourier transform mass spectrometry. Mol. Cell. Proteomics, 2011. 10(7): M111.009431.
184. Nagornov, K.O., et al., High-resolution fourier transform ion cyclotron resonance mass spectrometry with increased throughput for biomolecular analysis. Anal. Chem., 2014. 86(18): 9020-9028.
185. Comisarow, M.B. and A.G. Marshall, Fourier transform ion cyclotron resonance spectroscopy. Chem. Phys. Lett., 1974. 25(2): 282-283.
186. Nikolaev, E.N., Y.I. Kostyukevich, and G.N. Vladimirov, Fourier transform ion cyclotron resonance (FT ICR) mass spectrometry: Theory and simulations. Mass Spectrom. Rev., 2016. 35(2): 219-258.
187. Kingdon, K.H., A method for the neutralization of electron space charge by positive ionization at very low gas pressures. Phys. Rev., 1923. 21(4): 408-418.

188. Makarov, A., Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. Anal. Chem., 2000. 72(6): 1156-1162.
189. Hu, Q., et al., The Orbitrap: a new mass spectrometer. J. Mass Spectrom., 2005. 40(4): 430-443.
190. Scigelova, M. and A. Makarov, Orbitrap mass analyzer – overview and applications in proteomics. Proteomics, 2006. 6(S2): 16-21.
191. Makarov, A., et al., Performance evaluation of a hybrid linear ion trap/Orbitrap mass spectrometer. Anal. Chem., 2006. 78(7): 2113-2120.
192. Martins, C.P.B., et al., Chapter 1 - Orbitrap Mass Spectrometry: Evolution and Applicability, in Comprehensive Analytical Chemistry, P.E. Sandra Pérez and B. Damià, Editors. 2016, Elsevier. 3-18.
193. Huang, T.Y. and S.A. McLuckey, Gas-phase chemistry of multiply charged bioions in analytical mass spectrometry. Annu. Rev. Anal. Chem. (Palo Alto Calif.), 2010. 3: 365-85.
194. Molina, H., et al., Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. Proc. Natl. Acad. Sci. U. S. A., 2007. 104(7): 2199-204.
195. Olsen, J.V., et al., Higher-energy C-trap dissociation for peptide modification analysis. Nat. Methods, 2007. 4(9): 709-12.
196. Jedrychowski, M.P., et al., Evaluation of HCD- and CID-type fragmentation within their respective detection platforms for murine phosphoproteomics. Mol. Cell. Proteomics, 2011. 10(12): M111 009910.
197. Chi, A., et al., Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry. Proc. Natl. Acad. Sci. U. S. A., 2007. 104(7): 2193-8.
198. Garcia, B.A., et al., Pervasive combinatorial modification of histone H3 in human cells. Nat. Methods, 2007. 4(6): 487-9.
199. Phanstiel, D., et al., Mass spectrometry identifies and quantifies 74 unique histone H4 isoforms in differentiating human embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A., 2008. 105(11): 4093-8.
200. Ladumor, M.K., et al., Chapter 7 - High-Resolution Mass Spectrometry in Metabolite Identification, in Comprehensive Analytical Chemistry, P.E. Sandra Pérez and B. Damià, Editors. 2016, Elsevier. 199-229.
201. Chernushevich, I.V., A.V. Loboda, and B.A. Thomson, An introduction to quadrupole–time-of-flight mass spectrometry. J. Mass Spectrom., 2001. 36(8): 849-865.
202. Andrews, G.L., et al., Performance characteristics of a new hybrid quadrupole time-of-flight tandem mass spectrometer (TripleTOF 5600). Anal. Chem., 2011. 83(13): 5442-5446.
203. Michalski, A., et al., Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. Mol. Cell. Proteomics, 2011. 10(9): M111 011015.
204. Zubarev, R.A. and A. Makarov, Orbitrap mass spectrometry. Anal. Chem., 2013. 85(11): 5288-5296.
205. Eliuk, S. and A. Makarov, Evolution of Orbitrap mass spectrometry instrumentation. Annu. Rev. Anal. Chem. (Palo Alto Calif.), 2015. 8: 61-80.
206. Andreas, W., et al., Development and applications of a new benchtop Orbitrap mass spectrometer. 2008.

207. Scheltema, R.A., et al., The Q Exactive HF, a Benchtop mass spectrometer with a pre-filter, high-performance quadrupole and an ultra-high-field Orbitrap analyzer. Mol. Cell. Proteomics, 2014. 13(12): 3698-708.
208. Scientific, T., Improved sensitivity through enhanced ion transmission using an S-Lens on the LTQ Velos linear ion trap.
209. Lacorte, S. and A.R. Fernandez-Alba, Time of flight mass spectrometry applied to the liquid chromatographic analysis of pesticides in water and food. Mass Spectrom. Rev., 2006. 25(6): 866-80.
210. Nilsson, C.L., Advances in quantitative phosphoproteomics. Anal. Chem., 2012. 84(2): 735-46.
211. Ong, S.E. and M. Mann, A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). Nat. Protoc., 2006. 1(6): 2650-60.
212. Monetti, M., et al., Large-scale phosphosite quantification in tissues by a spike-in SILAC method. Nat. Methods, 2011. 8(8): 655-8.
213. Jones, A.M. and T.S. Nuhse, Phosphoproteomics using iTRAQ. Methods Mol. Biol., 2011. 779: 287-302.
214. Unwin, R.D., J.R. Griffiths, and A.D. Whetton, Simultaneous analysis of relative protein expression levels across multiple samples using iTRAQ isobaric tags with 2D nano LC-MS/MS. Nat. Protoc., 2010. 5(9): 1574-82.
215. Wu, C.J., et al., Quantitative phosphoproteomics studies using stable isotope dimethyl labeling coupled with IMAC-HILIC-nanoLC-MS/MS for estrogen-induced transcriptional regulation. J. Proteome Res., 2011. 10(3): 1088-97.
216. Wong, J.W. and G. Cagney, An overview of label-free quantitation methods in proteomics by mass spectrometry. Methods Mol. Biol., 2010. 604: 273-83.
217. Soderblom, E.J., et al., Quantitative label-free phosphoproteomics strategy for multifaceted experimental designs. Anal. Chem., 2011. 83(10): 3758-64.
218. Montoya, A., et al., Characterization of a TiO(2) enrichment method for label-free quantitative phosphoproteomics. Methods, 2011. 54(4): 370-8.
219. Gatlin, C.L., et al., Automated identification of amino acid sequence variations in proteins by HPLC/Microspray tandem mass spectrometry. Anal. Chem., 2000. 72(4): 757-763.
220. Bateman, N.W., et al., Maximizing peptide identification events in proteomic workflows using data-dependent acquisition (DDA). Mol. Cell. Proteomics, 2014. 13(1): 329-38.
221. Marx, V., Finding the right antibody for the job. Nat. Methods, 2013. 10(8): 703-7.
222. Bauer, M., et al., Assessment of current mass spectrometric workflows for the quantification of low abundant proteins and phosphorylation sites. Data Brief, 2015. 5: 297-304.
223. Ronsein, G.E., et al., Parallel reaction monitoring (PRM) and selected reaction monitoring (SRM) exhibit comparable linearity, dynamic range and precision for targeted quantitative HDL proteomics. J. Proteomics, 2015. 113: 388-99.
224. Lawrence, R.T., et al., Plug-and-play analysis of the human phosphoproteome by targeted high-resolution mass spectrometry. Nat. Methods, 2016. 13(5): 431-4.
225. Hu, A., W.S. Noble, and A. Wolf-Yadlin, Technical advances in proteomics: new developments in data-independent acquisition. F1000Res, 2016. 5.
226. Gillet, L.C., et al., Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Mol. Cell. Proteomics, 2012. 11(6): O111 016717.

227. Cox, J. and M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol., 2008. 26(12): 1367-72.
228. Perkins, D.N., et al., Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis, 1999. 20(18): 3551-67.
229. Eng, J.K., A.L. McCormack, and J.R. Yates, An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J. Am. Soc. Mass Spectrom., 1994. 5(11): 976-989.
230. Ting, Y.S., et al., Peptide-Centric proteome analysis: an alternative strategy for the analysis of tandem mass spectrometry data. Mol. Cell. Proteomics, 2015. 14(9): 2301-2307.
231. Tsou, C.-C., et al., DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. Nat. Methods, 2015. 12(3): 258-264.
232. MacLean, B., et al., Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics, 2010. 26(7): 966-968.
233. Rost, H.L., et al., OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. Nat. Biotechnol., 2014. 32(3): 219-23.
234. Colin, S., G. Chinetti-Gbaguidi, and B. Staels, Macrophage phenotypes in atherosclerosis. Immunol. Rev., 2014. 262(1): 153-66.
235. Huma, Z.E., et al., Key determinants of selective binding and activation by the monocyte chemoattractant proteins at the chemokine receptor CCR2. Sci. Signal., 2017. 10(480).
236. Shevchenko, A., et al., In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat. Protocols, 2007. 1(6): 2856-2860.
237. Batth, T.S., C. Francavilla, and J.V. Olsen, Off-line high-pH reversed-phase fractionation for in-depth phosphoproteomics. J. Proteome Res., 2014. 13(12): 6176-86.
238. Rappsilber, J., M. Mann, and Y. Ishihama, Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc., 2007. 2(8): 1896-906.
239. Cox, J. and M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol., 2008. 26(12): 1367-72.
240. Devaux, S., et al., Proteomic analysis of the spatio-temporal based molecular kinetics of acute spinal cord injury identifies a time- and segment-specific window for effective tissue repair. Mol. Cell. Proteomics, 2016.
241. Saeed, A.I., et al., TM4: a free, open-source system for microarray data management and analysis. Biotechniques, 2003. 34(2): 374-8.
242. Pathan, M., et al., FunRich: An open access standalone functional enrichment and interaction network analysis tool. Proteomics, 2015. 15(15): 2597-601.
243. Gu, Z., R. Eils, and M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics, 2016. 32(18): 2847-9.
244. Tyanova, S., et al., The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods, 2016. 13(9): 731-40.
245. Wickham, H., ggplot2: Elegant Graphics for Data Analysis. 2009: Springer-Verlag New York.

246. de Graaf, E.L., et al., Single-step enrichment by Ti4+-IMAC and label-free quantitation enables in-depth monitoring of phosphorylation dynamics with high reproducibility and temporal resolution. Mol. Cell. Proteomics, 2014. 13(9): 2426-2434.
247. Schwartz, D. and S.P. Gygi, An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. Nat. Biotechnol., 2005. 23(11): 1391-8.
248. Linding, R., et al., NetworKIN: a resource for exploring cellular phosphorylation networks. Nucleic Acids Res., 2008. 36(Database issue): D695-9.
249. Bindea, G., et al., ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics, 2009. 25(8): 1091-3.
250. Shannon, P., et al., Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res., 2003. 13(11): 2498-504.
251. Huang da, W., B.T. Sherman, and R.A. Lempicki, Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res., 2009. 37(1): 1-13.
252. Huang da, W., B.T. Sherman, and R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc., 2009. 4(1): 44-57.
253. Saita, Y., et al., Structural basis for the interaction of CCR5 with a small molecule, functionally selective CCR5 agonist. J. Immunol., 2006. 177(5): 3116-22.
254. Uy, G.L., M.P. Rettig, and A.F. Cashen, Plerixafor, a CXCR4 antagonist for the mobilization of hematopoietic stem cells. Expert Opin. Biol. Ther., 2008. 8(11): 1797-804.
255. Dawson, T.C., et al., Absence of CC chemokine receptor-2 reduces atherosclerosis in apolipoprotein E-deficient mice. Atherosclerosis, 1999. 143(1): 205-11.
256. Gu, L., et al., Absence of monocyte chemoattractant protein-1 reduces atherosclerosis in low density lipoprotein receptor-deficient mice. Mol. Cell, 1998. 2(2): 275-81.
257. Nestler EJ, G.P., Protein Phosphorylation is of Fundamental Importance in Biological Regulation, in Basic Neurochemistry: Molecular, Cellular and Medical Aspects, A.B. Siegel GJ, Albers RW, Editor. 1999: Philadelphia: Lippincott-Raven.
258. Choudhary, C. and M. Mann, Decoding signalling networks by mass spectrometry-based proteomics. Nat. Rev. Mol. Cell Biol., 2010. 11(6): 427-39.
259. Robles, M.S., S.J. Humphrey, and M. Mann, Phosphorylation is a central mechanism for circadian control of metabolism and physiology. Cell Metab., 2017. 25(1): 118-127.
260. van den Biggelaar, M., et al., Quantitative phosphoproteomics unveils temporal dynamics of thrombin signaling in human endothelial cells. Blood, 2014. 123(12): e22-36.
261. Yi, T., et al., Quantitative phosphoproteomic analysis reveals system-wide signaling pathways downstream of SDF-1/CXCR4 in breast cancer stem cells. Proc. Natl. Acad. Sci. U. S. A., 2014. 111(21): E2182-90.
262. Wojcechowskyj, J.A., et al., Quantitative phosphoproteomics of CXCL12 (SDF-1) signaling. PLoS One, 2011. 6(9): e24918.

263. O'Hayre, M., et al., Elucidating the CXCL12/CXCR4 signaling network in chronic lymphocytic leukemia through phosphoproteomics analysis. PLoS One, 2010. 5(7): e11716.
264. Parker, B.L., et al., Targeted phosphoproteomics of insulin signaling using data-independent acquisition mass spectrometry. Sci. Signal., 2015. 8(380): rs6.
265. Hayward, J., et al., Ticks from diverse genera encode chemokine-inhibitory evasin proteins. J. Biol. Chem., 2017. 292(38): 15670-15680.
266. Lam, H., et al., Development and validation of a spectral library searching method for peptide identification from MS/MS. Proteomics, 2007. 7(5): 655-67.
267. Chabot, D.J., et al., N-linked glycosylation of CXCR4 masks coreceptor function for CCR5-dependent human immunodeficiency virus type 1 isolates. J. Virol., 2000. 74(9): 4404-13.
268. Preobrazhensky, A.A., et al., Monocyte chemotactic protein-1 receptor CCR2B is a glycoprotein that has tyrosine sulfation in a conserved extracellular N-terminal region. J. Immunol., 2000. 165(9): 5295-303.
269. Schirle, M., M.A. Heurtier, and B. Kuster, Profiling core proteomes of human cell lines by one-dimensional PAGE and liquid chromatography-tandem mass spectrometry. Mol. Cell. Proteomics, 2003. 2(12): 1297-305.
270. Hunter, T. and B.M. Sefton, Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. Proc. Natl. Acad. Sci. U. S. A., 1980. 77(3): 1311-5.
271. Olsen, J.V., et al., Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell, 2006. 127(3): 635-48.
272. Humphrey, S.J., S.B. Azimifar, and M. Mann, High-throughput phosphoproteomics reveals in vivo insulin signaling dynamics. Nat. Biotechnol., 2015. 33(9): 990-5.
273. Jimenez-Sainz, M.C., et al., Signaling pathways for monocyte chemoattractant protein 1-mediated extracellular signal-regulated kinase activation. Mol. Pharmacol., 2003. 64(3): 773-82.
274. Chen, X., et al., CCL2/CCR2 regulates the tumor microenvironment in HER-2/neu-driven mammary carcinomas in mice. PLoS One, 2016. 11(11): e0165595.
275. Ji, W.T., et al., Monocyte chemotactic protein 1 (MCP-1) modulates pro-survival signaling to promote progression of head and neck squamous cell carcinoma. PLoS One, 2014. 9(2): e88952.
276. Werle, M., et al., MCP-1 induces activation of MAP-kinases ERK, JNK and p38 MAPK in human endothelial cells. Cardiovasc. Res., 2002. 56(2): 284-92.
277. Cambien, B., et al., Signal transduction involved in MCP-1-mediated monocytic transendothelial migration. Blood, 2001. 97(2): 359-66.
278. Cai, K., et al., MCP-1 upregulates amylin expression in murine pancreatic beta cells through ERK/JNK-AP1 and NF-kappaB related signaling pathways independent of CCR2. PLoS One, 2011. 6(5): e19559.
279. Arefieva, T.I., et al., MCP-1-stimulated chemotaxis of monocytic and endothelial cells is dependent on activation of different signaling cascades. Cytokine, 2005. 31(6): 439-46.
280. Kleibeuker, W., et al., Physiological changes in GRK2 regulate CCL2-induced signaling to ERK1/2 and Akt but not to MEK1/2 and calcium. J. Neurochem., 2008. 104(4): 979-92.
281. Jin, D., et al., MCP-1 stimulates spinal microglia via PI3K/Akt pathway in bone cancer pain. Brain Res., 2015. 1599: 158-67.

282. Dallapiccola, B., et al., Identification of the origin of a 22p+ chromosome by triplex dosage effect of LDH B, GAPHD, TPI and ENO2. Ann. Genet., 1980. 23(2): 111-3.
283. Olson, E.N. and A. Nordheim, Linking actin dynamics and gene transcription to drive cellular motile functions. Nat. Rev. Mol. Cell Biol., 2010. 11(5): 353-65.
284. Rossman, K.L., C.J. Der, and J. Sondek, GEF means go: turning on RHO GTPases with guanine nucleotide-exchange factors. Nat. Rev. Mol. Cell Biol., 2005. 6(2): 167-80.
285. Bos, J.L., H. Rehmann, and A. Wittinghofer, GEFs and GAPs: critical elements in the control of small G proteins. Cell, 2007. 129(5): 865-77.
286. Schmidt, A. and A. Hall, Guanine nucleotide exchange factors for Rho GTPases: turning on the switch. Genes Dev., 2002. 16(13): 1587-609.
287. Chahdi, A. and J.P. Raufman, The Cdc42/Rac nucleotide exchange factor protein beta1Pix (Pak-interacting exchange factor) modulates beta-catenin transcriptional activity in colon cancer cells: evidence for direct interaction of beta1PIX with beta-catenin. J. Biol. Chem., 2013. 288(47): 34019-29.
288. Kim, Y., et al., Phosphorylation of WAVE1 regulates actin polymerization and dendritic spine morphology. Nature, 2006. 442(7104): 814-7.
289. von Holleben, M., et al., Immunoinhibitory adapter protein Src homology domain 3 lymphocyte protein 2 (SLy2) regulates actin dynamics and B cell spreading. J. Biol. Chem., 2011. 286(15): 13489-501.
290. Janjanam, J., et al., PLCbeta3 mediates cortactin interaction with WAVE2 in MCP1-induced actin polymerization and cell migration. Mol. Biol. Cell, 2015. 26(25): 4589-606.
291. Varrin-Doyer, M., et al., Phosphorylation of collapsin response mediator protein 2 on Tyr-479 regulates CXCL12-induced T lymphocyte migration. J. Biol. Chem., 2009. 284(19): 13265-76.
292. Pons, M., et al., Phosphorylation of filamin A regulates chemokine receptor CCR2 recycling. J. Cell Sci., 2017. 130(2): 490-501.
293. Weiss-Haljiti, C., et al., Involvement of phosphoinositide 3-kinase gamma, Rac, and PAK signaling in chemokine-induced macrophage migration. J. Biol. Chem., 2004. 279(41): 43273-84.
294. Hanif, I.M., et al., Casein kinase II: an attractive target for anti-cancer drug design. Int. J. Biochem. Cell Biol., 2010. 42(10): 1602-5.
295. Agostinis, P., et al., Regulation of casein kinase 2 by phosphorylation/dephosphorylation. Biochem. J., 1987. 248(3): 785-9.
296. Wang, T., et al., Identification and functional characterisation of 5-HT4 receptor in sea cucumber Apostichopus japonicus (Selenka). Sci. Rep., 2017. 7: 40247.
297. Li, S., et al., MCP-1-induced ERK/GSK-3beta/Snail signaling facilitates the epithelial-mesenchymal transition and promotes the migration of MCF-7 human breast carcinoma cells. Cell. Mol. Immunol., 2017. 14(7): 621-630.
298. Lehto, M., et al., The R-Ras interaction partner ORP3 regulates cell adhesion. J. Cell Sci., 2008. 121(Pt 5): 695-705.
299. Ballestrem, C., et al., Actin-dependent lamellipodia formation and microtubule-dependent tail retraction control-directed cell migration. Mol. Biol. Cell, 2000. 11(9): 2999-3012.
300. Chou, M.F. and D. Schwartz, Biological sequence motif discovery using motif-x. Curr. Ptotoc. Bioinformatics, 2011. Chapter 13: Unit 13 15-24.
301. Horn, H., et al., KinomeXplorer: an integrated platform for kinome biology studies. Nat. Methods, 2014. 11(6): 603-4.

302. Manser, E., et al., Expression of constitutively active alpha-PAK reveals effects of the kinase on actin and focal complexes. Mol. Cell. Biol., 1997. 17(3): 1129-43.
303. Soderling, T.R., B. Chang, and D. Brickey, Cellular signaling through multifunctional Ca2+/calmodulin-dependent protein kinase II. J. Biol. Chem., 2001. 276(6): 3719-22.
304. Tanaka, T., et al., Monocyte chemoattractant protein-1/CC chemokine ligand 2 enhances apoptotic cell removal by macrophages through Rac1 activation. Biochem. Biophys. Res. Commun., 2010. 399(4): 677-82.
305. Terashima, Y., et al., Pivotal function for cytoplasmic protein FROUNT in CCR2-mediated monocyte chemotaxis. Nat. Immunol., 2005. 6(8): 827-35.
306. Tyanova, S., T. Temu, and J. Cox, The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat. Protoc., 2016. 11(12): 2301-2319.
307. Szklarczyk, D., et al., STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res., 2015. 43(Database issue): D447-52.
308. Shandilya, J., et al., Phosphorylation of multifunctional nucleolar protein nucleophosmin (NPM1) by aurora kinase B is critical for mitotic progression. FEBS Lett., 2014. 588(14): 2198-205.
309. Wang, Z., G. Qin, and T.C. Zhao, HDAC4: mechanism of regulation and biological functions. Epigenomics, 2014. 6(1): 139-50.
310. Nakagawa, Y., et al., Class II HDACs mediate CaMK-dependent signaling to NRSF in ventricular myocytes. J. Mol. Cell. Cardiol., 2006. 41(6): 1010-22.
311. Shen, L., Tight junctions on the move: molecular mechanisms for epithelial barrier regulation. Ann. N. Y. Acad. Sci., 2012. 1258: 9-18.
312. Huo, L., et al., Cdc42-dependent formation of the ZO-1/MRCKbeta complex at the leading edge controls cell migration. EMBO J., 2011. 30(4): 665-78.
313. Coles, L.S., et al., Cold shock domain proteins repress transcription from the GM-CSF promoter. Nucleic Acids Res., 1996. 24(12): 2311-7.
314. Liu, Y., et al., The ZEB1 transcription factor acts in a negative feedback loop with miR200 downstream of Ras and Rb1 to regulate Bmi1 expression. J. Biol. Chem., 2014. 289(7): 4116-25.
315. Deng, Z., et al., Radiation-induced c-Jun activation depends on MEK1-ERK1/2 signaling pathway in microglial cells. PLoS One, 2012. 7(5): e36739.
316. Kayahara, M., X. Wang, and C. Tournier, Selective regulation of c-jun gene expression by mitogen-activated protein kinases via the 12-o-tetradecanoylphorbol-13-acetate- responsive element and myocyte enhancer factor 2 binding sites. Mol. Cell. Biol., 2005. 25(9): 3784-92.
317. Waskiewicz, A.J., et al., Mitogen-activated protein kinases activate the serine/threonine kinases Mnk1 and Mnk2. EMBO J., 1997. 16(8): 1909-20.
318. Biswas, S.K. and A. Sodhi, Tyrosine phosphorylation-mediated signal transduction in MCP-1-induced macrophage activation: role for receptor dimerization, focal adhesion protein complex and JAK/STAT pathway. Int. Immunopharmacol., 2002. 2(8): 1095-107.
319. Schuringa, J.-J., et al., Ser727-dependent transcriptional activation by association of p300 with STAT3 upon IL-6 stimulation. FEBS Lett., 2001. 495(1): 71-76.
320. Goicoechea, S.M., S. Awadia, and R. Garcia-Mata, I'm coming to GEF you: regulation of RhoGEFs during cell migration. Cell Adh. Migr., 2014. 8(6): 535-49.

321. Faust, M. and M. Montenarh, Subcellular localization of protein kinase CK2. A key to its function? Cell Tissue Res., 2000. 301(3): 329-40.
322. Lin, C.H., C.H. Shih, and B.C. Chen, Thrombin-induced IL-8/CXCL8 release is mediated by CK2, MSK1, and NF-kappaB pathways in human lung epithelial cells. Eur. J. Pharmacol., 2015. 767: 135-43.
323. Litchfield, D.W., et al., Phosphorylation of casein kinase II by p34cdc2 in vitro and at mitosis. J. Biol. Chem., 1992. 267(20): 13943-51.
324. Bek, S. and R. Kemler, Protein kinase CKII regulates the interaction of beta-catenin with alpha-catenin and its protein stability. J. Cell Sci., 2002. 115(Pt 24): 4743-53.
325. McNeil, E., C.T. Capaldo, and I.G. Macara, Zonula occludens-1 function in the assembly of tight junctions in Madin-Darby canine kidney epithelial cells. Mol. Biol. Cell, 2006. 17(4): 1922-32.
326. Nunbhakdi-Craig, V., et al., Protein phosphatase 2A associates with and regulates atypical PKC and the epithelial tight junction complex. J. Cell Biol., 2002. 158(5): 967-78.
327. Gautreau, A., et al., Ezrin, a plasma membrane-microfilament linker, signals cell survival through the phosphatidylinositol 3-kinase/Akt pathway. Proc. Natl. Acad. Sci. U. S. A., 1999. 96(13): 7300-5.
328. Parnell, E., et al., Phosphorylation of ezrin on Thr567 is required for the synergistic activation of cell spreading by EPAC1 and protein kinase A in HEK293T cells. Biochim. Biophys. Acta, 2015. 1853(7): 1749-58.
329. Simonovic, I., et al., Enteropathogenic Escherichia coli activates ezrin, which participates in disruption of tight junction barrier function. Infect. Immun., 2001. 69(9): 5679-88.
330. Kranenburg, O., et al., Activation of RhoA by lysophosphatidic acid and Galpha12/13 subunits in neuronal cells: induction of neurite retraction. Mol. Biol. Cell, 1999. 10(6): 1851-7.
331. Chikumi, H., S. Fukuhara, and J.S. Gutkind, Regulation of G protein-linked guanine nucleotide exchange factors for Rho, PDZ-RhoGEF, and LARG by tyrosine phosphorylation: evidence of a role for focal adhesion kinase. J. Biol. Chem., 2002. 277(14): 12463-73.
332. Sprang, S., GEFs: master regulators of G-protein activation. Trends Biochem. Sci., 2001. 26(4): 266-7.
333. Singh, N.K., J. Janjanam, and G.N. Rao, p115 RhoGEF activates the Rac1 GTPase signaling cascade in MCP1 chemokine-induced vascular smooth muscle cell migration and proliferation. J. Biol. Chem., 2017. 292(34): 14080-14091.
334. Hernandez-Garcia, R., et al., Endothelial RhoGEFs: A systematic analysis of their expression profiles in VEGF-stimulated and tumor endothelial cells. Vascul. Pharmacol., 2015. 74: 60-72.
335. Brazier, H., et al., Expression profile of RhoGTPases and RhoGEFs during RANKL-stimulated osteoclastogenesis: identification of essential genes in osteoclasts. J. Bone Miner. Res., 2006. 21(9): 1387-98.
336. Kosako, H. and N. Imamoto, Phosphorylation of nucleoporins: signal transduction-mediated regulation of their interaction with nuclear transport receptors. Nucleus (Calcutta), 2010. 1(4): 309-13.
337. Kabachinski, G. and T.U. Schwartz, The nuclear pore complex--structure and function at a glance. J. Cell Sci., 2015. 128(3): 423-9.

338. Glavy, J.S., et al., Cell-cycle-dependent phosphorylation of the nuclear pore Nup107-160 subcomplex. Proc. Natl. Acad. Sci. U. S. A., 2007. 104(10): 3811-6.
339. Negi, S.S. and M.O. Olson, Effects of interphase and mitotic phosphorylation on the mobility and location of nucleolar protein B23. J. Cell Sci., 2006. 119(Pt 17): 3676-85.
340. Hayakawa, F. and M.L. Privalsky, Phosphorylation of PML by mitogen-activated protein kinases plays a key role in arsenic trioxide-mediated apoptosis. Cancer Cell, 2004. 5(4): 389-401.
341. Cheng, X. and H.Y. Kao, Post-translational modifications of PML: consequences and implications. Front. Oncol., 2012. 2: 210.
342. Lim, J.H., et al., Mitogen-activated protein kinase extracellular signal-regulated kinase 2 phosphorylates and promotes Pin1 protein-dependent promyelocytic leukemia protein turnover. J. Biol. Chem., 2011. 286(52): 44403-11.
343. Barquera, S., et al., Global overview of the epidemiology of atherosclerotic cardiovascular disease. Arch. Med. Res., 2015. 46(5): 328-38.
344. Insull, W., Jr., The pathology of atherosclerosis: plaque development and plaque responses to medical treatment. Am. J. Med., 2009. 122(1 Suppl): S3-S14.
345. Davies, M.J., et al., Risk of thrombosis in human atherosclerotic plaques: role of extracellular lipid, macrophage, and smooth muscle cell content. Br. Heart J., 1993. 69(5): 377-81.
346. Itoh, R.E., et al., Phosphorylation and activation of the Rac1 and Cdc42 GEF Asef in A431 cells stimulated by EGF. J. Cell Sci., 2008. 121(Pt 16): 2635-42.
347. Kundumani-Sridharan, V., et al., Nuclear factor of activated T cells c1 mediates p21-activated kinase 1 activation in the modulation of chemokine-induced human aortic smooth muscle cell F-actin stress fiber formation, migration, and proliferation and injury-induced vascular wall remodeling. J. Biol. Chem., 2013. 288(30): 22150-62.
348. Takesono, A., et al., Requirement for Tec kinases in chemokine-induced migration and activation of Cdc42 and Rac. Curr. Biol., 2004. 14(10): 917-22.
349. Weber, K.S., et al., Chemokine-induced monocyte transmigration requires cdc42-mediated cytoskeletal changes. Eur. J. Immunol., 1998. 28(7): 2245-51.
350. Zins, K., et al., A Rac1/Cdc42 GTPase-specific small molecule inhibitor suppresses growth of primary human prostate cancer xenografts and prolongs survival in mice. PLoS One, 2013. 8(9): e74924.
351. Onesto, C., et al., Characterization of EHT 1864, a novel small molecule inhibitor of Rac family small GTPases. Methods Enzymol., 2008. 439: 111-29.
352. Surviladze, Z., et al., A Potent and Selective Inhibitor of Cdc42 GTPase, in Probe Reports from the NIH Molecular Libraries Program. 2010: Bethesda (MD).
353. Ibarra, A. and M.W. Hetzer, Nuclear pore proteins and the control of genome functions. Genes Dev., 2015. 29(4): 337-49.
354. Nardozzi, J.D., K. Lott, and G. Cingolani, Phosphorylation meets nuclear import: a review. Cell Commun. Signal., 2010. 8: 32.
355. Kosako, H., et al., Phosphoproteomics reveals new ERK MAP kinase targets and links ERK to nucleoporin-mediated nuclear transport. Nat. Struct. Mol. Biol., 2009. 16(10): 1026-35.
356. Whitmarsh, A.J., Regulation of gene transcription by mitogen-activated protein kinase signaling pathways. Biochim. Biophys. Acta, 2007. 1773(8): 1285-98.
357. McBride, K.M., et al., Regulated nuclear import of the STAT1 transcription factor by direct binding of importin-alpha. EMBO J., 2002. 21(7): 1754-63.

358. Belema-Bedada, F., et al., Efficient homing of multipotent adult mesenchymal stem cells depends on FROUNT-mediated clustering of CCR2. Cell Stem Cell, 2008. 2(6): 566-75.
359. Satoh, M., et al., A novel activator of C-C chemokine, FROUNT, is expressed with C-C chemokine receptor 2 and its ligand in failing human heart. J. Card. Fail., 2007. 13(2): 114-9.
360. Shi, C. and E.G. Pamer, Monocyte recruitment during infection and inflammation. Nat. Rev. Immunol., 2011. 11(11): 762-74.
361. Combadiere, C., et al., Combined inhibition of CCL2, CX3CR1, and CCR5 abrogates Ly6C(hi) and Ly6C(lo) monocytosis and almost abolishes atherosclerosis in hypercholesterolemic mice. Circulation, 2008. 117(13): 1649-57.
362. Lopez-Lago, M., et al., Tyrosine phosphorylation mediates both activation and downmodulation of the biological activity of Vav. Mol. Cell. Biol., 2000. 20(5): 1678-91.
363. van Buul, J.D., et al., RhoG regulates endothelial apical cup assembly downstream from ICAM1 engagement and is involved in leukocyte trans-endothelial migration. J. Cell Biol., 2007. 178(7): 1279-93.
364. Samson, T., et al., The guanine-nucleotide exchange factor SGEF plays a crucial role in the formation of atherosclerosis. PLoS One, 2013. 8(1): e55202.
365. Li, J., et al., Akkermansia muciniphila protects against atherosclerosis by preventing metabolic endotoxemia-Induced inflammation in Apoe-/- mice. Circulation, 2016. 133(24): 2434-46.
366. Fan, J., et al., Role of endothelin-1 in atherosclerosis. Ann. N. Y. Acad. Sci., 2000. 902: 84-93; discussion 93-4.
367. Katsuno, T., et al., Deficiency of zonula occludens-1 causes embryonic lethal phenotype associated with defected yolk sac angiogenesis and apoptosis of embryonic cells. Mol. Biol. Cell, 2008. 19(6): 2465-75.
368. Hommerson, P., et al., Comparison of electrospray ionization and atmospheric pressure photoionization for coupling of micellar electrokinetic chromatography with ion trap mass spectrometry. J. Chromatogr. A, 2008. 1204(2): 197-203.
369. Nakayama, H., et al., Capillary column high-performance liquid chromatographic-electrospray ionization triple-stage quadrupole mass spectrometric analysis of proteins separated by two-dimensional polyacrylamide gel electrophoresis. Application to cerebellar protein mapping. J. Chromatogr. A, 1996. 730(1-2): 279-87.
370. le Maire, M., et al., Electrospray ionization mass spectrometry on hydrophobic peptides electroeluted from sodium dodecyl sulfate-polyacrylamide gel electrophoresis application to the topology of the sarcoplasmic reticulum Ca2+ ATPase. Anal. Biochem., 1993. 214(1): 50-7.
371. Liebler, D.C. and A.J. Ham, Spin filter-based sample preparation for shotgun proteomics. Nat. Methods, 2009. 6(11): 785.
372. Manza, L.L., et al., Sample preparation and digestion for proteomic analyses using spin filters. Proteomics, 2005. 5(7): 1742-5.
373. Zhou, J., et al., Evaluation of the application of sodium deoxycholate to proteomic analysis of rat hippocampal plasma membrane. J. Proteome Res., 2006. 5(10): 2547-53.
374. Zhang, F., et al., Changes in the proteomic profile during the differential polarization status of the human monocyte-derived macrophage THP-1 cell line. Proteomics, 2015. 15(4): 773-86.

375. Proc, J.L., et al., A quantitative study of the effects of chaotropic agents, surfactants, and solvents on the digestion efficiency of human plasma proteins by trypsin. J. Proteome Res., 2010. 9(10): 5422-37.
376. Masuda, T., et al., Unbiased quantitation of Escherichia coli membrane proteome using phase transfer surfactants. Mol. Cell. Proteomics, 2009. 8(12): 2770-7.
377. Kulak, N.A., et al., Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. Nat. Methods, 2014. 11(3): 319-24.
378. Wisniewski, J.R., et al., Universal sample preparation method for proteome analysis. Nat. Methods, 2009. 6(5): 359-62.
379. Masuda, T., M. Tomita, and Y. Ishihama, Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis. J. Proteome Res., 2008. 7(2): 731-40.
380. Batth, T.S., C. Francavilla, and J.V. Olsen, Off-line high-pH reversed-phase fractionation for in-depth phosphoproteomics. J. Proteome Res., 2014. 13(12): 6176-6186.
381. Pujol, F., P. Kitabgi, and H. Boudin, The chemokine SDF-1 differentially regulates axonal elongation and branching in hippocampal neurons. J. Cell Sci., 2005. 118(Pt 5): 1071-80.
382. Bamidele, A.O., et al., IQGAP1 promotes CXCR4 chemokine receptor function and trafficking via EEA-1+ endosomes. J. Cell Biol., 2015. 210(2): 257-72.
383. Li, H., et al., Association between Galphai2 and ELMO1/Dock180 connects chemokine signalling with Rac activation and metastasis. Nat Commun, 2013. 4: 1706.

Appendices

Appendix I Implementation and Evaluation of Methodology for Global Proteomic Analyses

1. Implementation and Evaluation of Methodology for Global Proteomic Analyses

1.1. Comparison of Detergents for Extraction and Processing of Cellular Proteins

1.1.1. Introduction to Protein Extraction Methods

A key step in preparing samples for bottom-up proteomics is always to extract proteins from the biological material (e.g. cultured cells, bacteria, yeast, tissues or plants) in a manner that enables them to be efficiently digested with a protease. Urea and sodium dodecyl sulfate (SDS) are two of the most frequently used reagents for extracting, solubilizing and denaturing proteins.

Urea, typically used at a concentration of 8 M, helps to solubilize proteins by disrupting their non-covalent bonds and also makes the buried peptide bonds more accessible to proteases. However, isocyanic acid, which is a degradation product from urea, can carbamylate (add a $C(O)NH_2$ group to) the amino terminus of proteins and peptides, which change their physiochemical properties. In addition, it may even modify the side chains of lysine and arginine residues resulting in decreased digestion efficiencies when using trypsin or LysC proteases. More importantly, most proteases lose their enzymatic activity in high concentrations of urea, which can only be avoided by diluting the sample to a larger volume, which is often not desired.

SDS is also an efficient anionic detergent used to solubilize proteins. However, even a small amount of SDS in the sample can inhibit the ionization of peptides in the electrospray source of a mass spectrometer [368], which is detrimental for peptide and protein identification. Various methods have been used in the past to remove SDS such as the use of reverse phase columns [369] or separating the proteins by acetone precipitation [370]. However, none of the method can completely deplete SDS in the sample.

Filter Aided Sample Preparation (FASP) is nowadays commonly used to efficiently deplete SDS [371,372]. The protein-SDS solution is loaded into a centrifugal

filter, with 10kDa or 30kDa cut off, and SDS is replaced with urea after 3-4 centrifugation cycles. All subsequent steps, including depletion of detergent, buffer exchange, protein reduction, alkylation, and digestion, are performed in the same centrifugal filter.

In preliminary FASP experiments, we found that the quality of the centrifugal filters varied from batch to batch, which affects the protein recovery. In addition, filters would occasionally dislodge during a centrifugation step, which not only causes the loss of very precious samples but also reduces the reproducibility of experiments. Moreover, FASP columns have a capacity of only 200-400 µg of protein, which is not sufficient for some proteomic experiments (such as phosphoproteomic) that require large amounts of starting material (several milligrams of protein).

The above-mentioned drawbacks of SDS, urea and FASP prompted us to look for alternatives. Sodium deoxycholate (SDC) is an anionic detergent, which has been widely used to extract cell membrane or plasma membrane proteins. [373,374] More importantly, it was reported that the presence of SDC can enhance the digestion activity of trypsin and that SDC can be completely removed by phase transfer with a water-immiscible organic solvent at low pH [375,376]. Furthermore, Kulak et al. [377] performed a whole sample preparation procedure (from cell lysis to protein digestion and finally peptide fractionation) in a single pipette tip using SDC, which showed excellent reproducibility and a better yield of membrane and nuclear proteins in comparison to SDS-containing samples. To test the implementation of this sample preparation technique in our laboratory, we systematically evaluated SDC-based protein extraction in comparison to the SDS-FASP method.

1.1.2. Experimental Design

HEK293 cells were lysed in HEPES buffer containing either 4% SDS or 1% SDC and the protein concentrations were determined. Equal amounts of protein were processed as shown schematically in Figure 1.1 and analysed by LC-MS/MS. Briefly, the SDS-extracted sample was processed by FASP, which included depletion of detergent, buffer exchange, protein reduction, alkylation and trypsin digestion on the centrifugal filter, whereas the SDC-extracted sample was digested in solution and depleted of detergent.

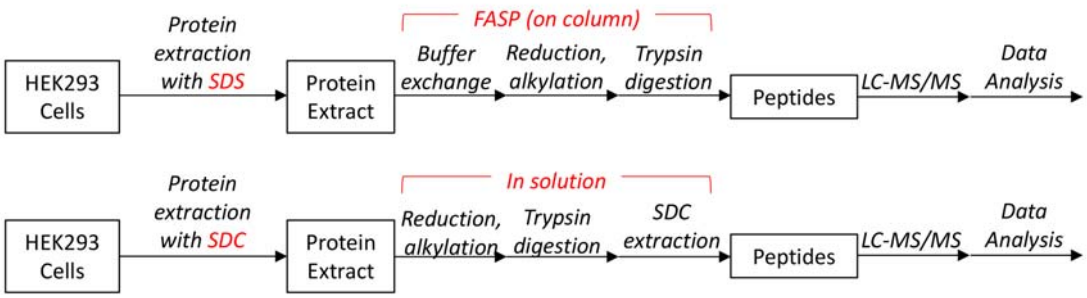


Figure 1.1. The experimental design of SDS and SDC comparison.

1.1.3. Results

The number of identified proteins and peptides from both sample preparations are shown in Figure 1.2. Starting with the same amount of protein in three independently repeated experiments and only considering protein and peptide identifications falling below a false-discovery rate (FDR) of 1%, we identified 37,594 ± 716 peptides resulting in 3,890 ± 34 protein identifications in the samples extracted with SDC. In comparison, the FASP-based sample preparation resulted in the identification of 33,699 ± 1,212 peptides derived from 3,684 ± 97 proteins. Although these differences are not statistically significant, we consistently identified 5-10% more peptides and proteins in the SDC-based protocol compared to the SDS-based FASP protocol. Thus, in our hands the SDC-based protocol was at least as effective, and possibly slightly more effective, than the SDS-FASP protocol.

In order to determine if the two protocols show a bias towards proteins in certain molecular weight ranges or towards certain types of proteins, we subjected the identified proteins to analysis with the program FunRich [242]. For this analysis, the subcellular distribution of identified proteins as well as their molecular weights (MW) were retrieved from the Uniprot database. As shown in Figure 1.3, there is no significant difference between the two methods with respect to these analysed features. We suspect that the observed differences are due to intrinsic properties of a mass spectrometer operated in data-dependent acquisition mode (a detailed introduction can be found in Chapter 1). In brief, during a DDA run, the intensities of all precursor ions dictate which peptides are sequenced, which inherently introduces reproducibility issues, especially for low abundant proteins [222].

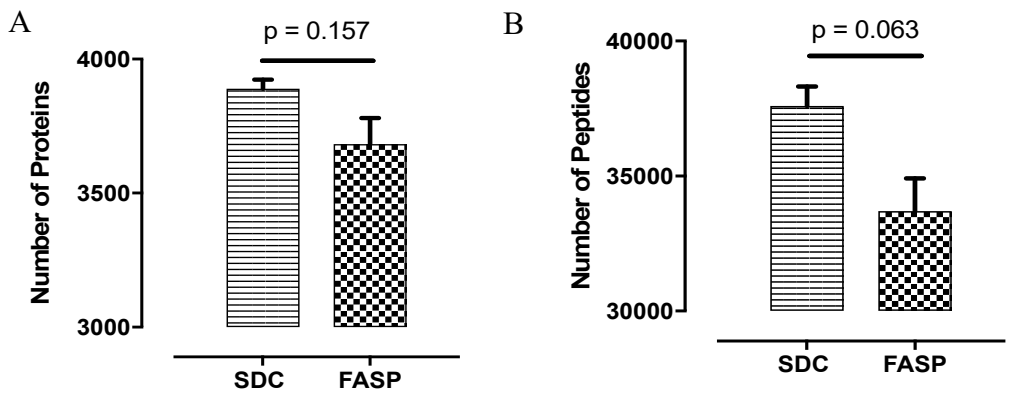


Figure 1.2. Comparison of the numbers of (A) proteins and (B) peptides identified using the SDC and SDS-FASP protein treatment protocols. All data is presented as mean ± SEM, n=3.

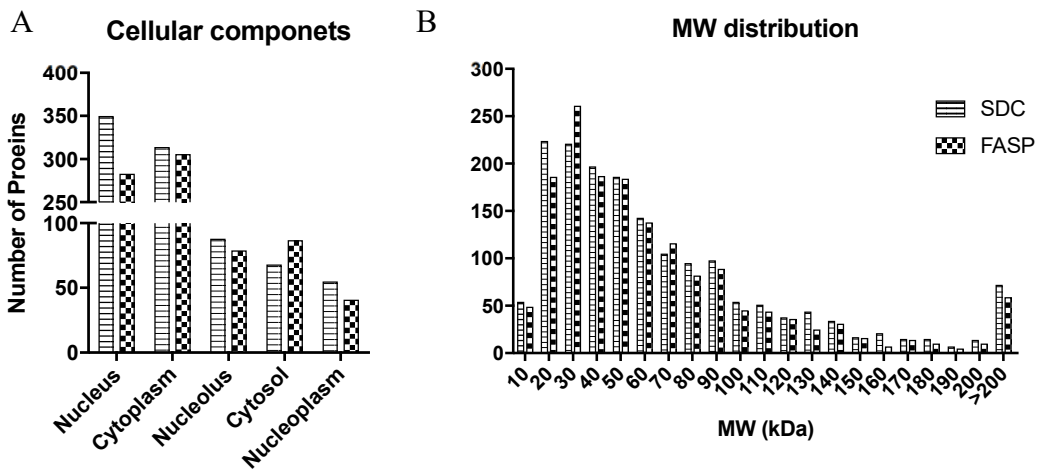


Figure 1.3. Cellular components (A) and MW distribution (B) comparison of unique proteins between SDC and FASP.

1.1.4. Discussion: Choice of SDC for Protein Extraction

As shown above, SDC solubilisation slightly but consistently outperformed the conventional SDS/FASP procedure in our hands, albeit without the difference reaching statistical significance. We have therefore chosen to use this sample preparation method for all the experiments described in this thesis. Below we discuss some of the properties of SDS and SDC that may contribute to differences in their usefulness in proteomic studies.

SDS is a strong anionic surfactant, which is a common reagent to be routinely used for helping to solubilise proteins; however a small amount of SDS left in the sample can be detrimental to protease digestion and to peptide detection in the mass spectrometer. The application of spin filter-aided SDS/detergent removal to proteomic sample preparation was initially published by Liebler's group in 2005 [372]. All the required chemical reactions, such as protein reduction, alkylation and digestion, worked well when tested with a simple protein solution (BSA), with sequence coverage over 70%. However, the SDS could not be completely removed. Although they claimed that SDS could be reduced to a concentration that did not affect digestion and mass spectrometric analysis, the number of BSA peptides identified was decreased comparing to the sample without SDS in. It is possible that this was due to the presence of SDS in the sample interfering with the ionisation of peptides.

In 2009, Mann's group optimized the filter-aided SDS/detergent removal method by replacing SDS with urea in the spin filter and named the methods the technique FASP [378]. It appeared that urea was efficient at replacing SDS after several centrifuge steps and boosted the protein identification from a few hundreds to more than 7,000. The major part of this substantial increase can be owing to the improvement of mass spectrometer, but FASP does has been shown to be superior in completely depleting SDS. Furthermore, it improved the identification of membrane proteins and can be scale up to for phosphopeptide identification or scaled down to 1-2 μg protein. The popularity of FASP can be reflected by its over 2,000 citations.

SDC is an anionic and bile-acid detergent, which is frequently used in assisting hydrophobic or membrane protein solubilisation [373,379]. SDC has several properties, such as acid-lability and efficiency in disrupting membranes and protein complexes, that make it attractive for routine use as a cell lysis buffer. Compared to SDS/FASP

procedures, the SDC method is relatively simple, as SDC can precipitate out under acidic conditions to be removed by centrifugation and extraction into organic solvent, such as ethyl acetate. More importantly, we found that the SDC method gave a similar number of proteins identified and did not introduce any sample bias during the preparation when compared to the SDS/FASP method (Figure 3.2). This has been confirmed by comparing the gene ontology analysis of the proteins and analysing the MW distribution between SDC and SDS/FASP (Figure 3.3).

Improved from previous publication that SDC was only applied in cultured cells [377], the Monash Biomedical Proteomics Facility has now successfully applied the SDC method in proteomic studies of human tissue, yeast, *E. coli* and *Plasmodium falciparum* and for phosphopeptide identification (data not shown). However, SDC alone was not effective at removing lipids from samples sourced from adipocytes or adipose tissue; for these samples SDC or SDS combined with FASP was still a better option (data not shown). Nevertheless, with this one exception, SDC extraction appears to be a general a robust method and was used successfully for the studies described in this thesis.

1.2. Fractionation Studies

1.2.1. Introduction: The Need for Peptide Fractionation

It has been reported that the human core proteome contains roughly 10,000 proteins [104]. Due to the limited scan speed of a mass spectrometer, a single-shot mass spectrometric analysis normally can characterise around 3,000 human proteins. In order to achieve a better coverage of the proteome, peptide pre-fractionation with a chromatography column or cartridge is always the first choice, as a single sample is distributed into a few fractions, increasing the chance of identification of low abundant peptides.

There are many types of chromatography resins available on the market and different conditions have been applied in many publications. To select the best method for fractionation of our samples, we have conducted an experiment to compare the performance of reverse phase C18 and strong cation exchange (SCX) resins, which are two of the most cited methods.

1.2.2. Experimental Design

HEK293 cells were lysed in SDC buffer and phase transfer was used to remove SDC, as described in materials and methods. Equal amounts of digested peptide were fractionated with SCX or C18 and analysed by LC-MS/MS. The experimental design is presented in Figure 1.4. The tested conditions can be found in Table 1.1.

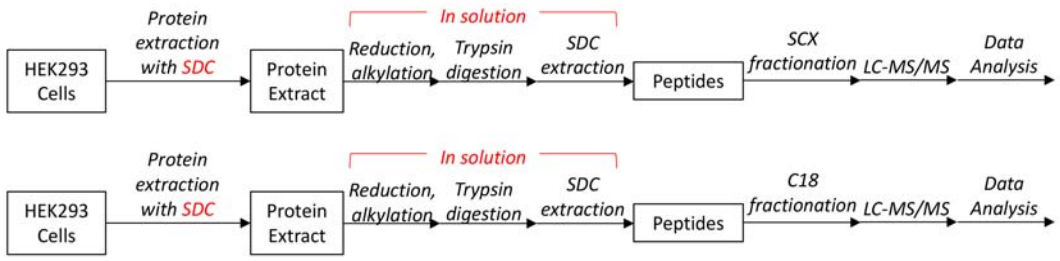


Figure 1.4 The experimental design of the fractionation comparison experiment.

Table 1.1 Fractionation conditions

Sample name	Condition
SCX-A or B	Buffer 1: 100mM Ammonium Acetate or NaCl, 20% ACN, 5% Formic acid
A: Ammonium Acetate	Buffer 2: 200mM Ammonium Acetate or NaCl, 20% ACN, 5% Formic acid
B: NaCl	Buffer 3: 300mM Ammonium Acetate or NaCl, 20% ACN, 5% Formic acid
	Buffer 4: 5% Ammonium Hydroxide, 80% ACN
reverse phase C18-Low pH	Buffer 1: 60mM Ammonium Formate, 10% ACN, 1% Formic Acid
	Buffer 2: 100mM Ammonium Formate, 35% ACN, 1% Formic Acid
	Buffer 3: 150mM Ammonium Formate, 50% ACN, 1% Formic Acid
	Buffer 4: 5% Ammonium Hydroxide, 80% ACN
reverse phase C18-High pH	Buffer 1: 60mM Ammonium Formate, 10% ACN, pH 10
	Buffer 2: 100mM Ammonium Formate, 35% ACN, pH 10
	Buffer 3: 150mM Ammonium Formate, 50% ACN, pH 10
	Buffer 4: 5% Formic Acid, 80% ACN

1.2.3. Results

As shown in Figure 1.5, all four conditions tested here allowed identification of at least 5000 proteins. Although only a single experiment was done under each set of conditions so statistical comparisons are not possible, reverse phase C18-High pH fractionation slightly outperformed the other three fractionation techniques. This is consistent with the results of a systematic comparison reported previously by Batth et al. [237]. Therefore, we chose the reverse phase C18-High pH fractionation technique as the standard method if fractionation was required in experiments described in this thesis.

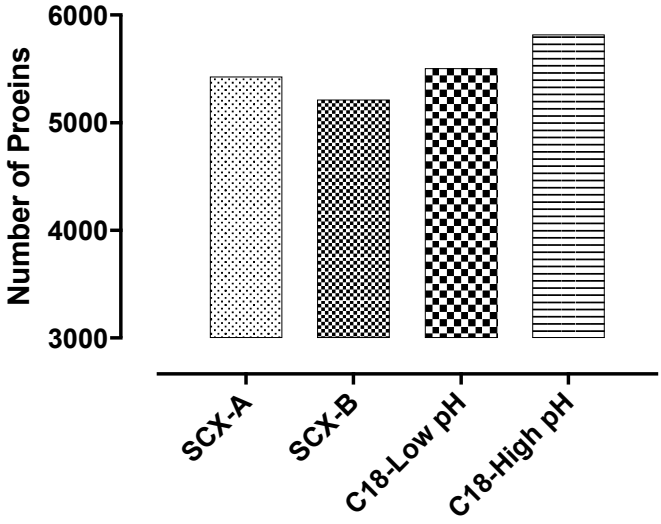


Figure 1.5. Identified proteins in different fractionation conditions.

1.2.4. Discussion: Advantages of C18 Fractionation

SCX can be an effective method for peptide fractionation in proteomic studies. Gygi's lab combined SCX with IMAC and improved phosphopeptide identification substantially from 2,000-3,000 to more than 10,000 [142]. Most tryptic peptides contain two positive charges under acidic condition, which enables them to be separated them from peptides with net charge of 1, such as phosphorylated peptides. However, SCX fractionation needs to be followed by a desalting step to prepare samples for mass spectrometric analysis, resulting in increased effort and cost. Furthermore, most phosphorylated peptides are present in the flow-through instead of binding to SCX column, phosphopeptide enrichment procedures are required for the flow-through.

In addition to being an effective method for peptide separation and recovery, reverse phase C18 fractionation some other benefits over SCX fractionation. First, the resulting peptide fractions can be directly lyophilized and subjected to mass spectrometric analysis, whereas an additional desalting step is required in case of SCX fractionation. Second, the peptides tend to be more evenly distributed over the different fractions in C18 fractionation, which can increase the peptide/protein coverage [380]. Therefore, using high pH C18 fractionation in the studies described in Chapters 3, 4 and 5, we successfully characterised in-depth proteomes, with more than 8,000 proteins identified in each case.

1.3. Assessment of the Reproducibility of Proteomic and Phosphoproteomic Analyses for Proteins Extracted Using SDC

1.3.1. Introduction: The Importance of Reliable Quantification

Successful detection and identification of protein-specific peptides confirm their presence within the sample; however, it is more important to quantitate the changes of the proteins. As extensively discussed in Chapter 1, phosphorylation is the most studied post-translational modification and crucial in signal transduction pathways. Exploiting the advantages of state-of-the-art mass spectrometry-based technologies, tens of thousands of phosphopeptides have been characterised in a single experiment in the past. However, it is becoming increasingly important to not only identify as many phosphopeptides as possible, but also reliably quantify these identified phosphopeptides. We therefore tested whether our SDC solubilization protocol is suitable for phosphoproteomic studies.

1.3.2. Experimental Design

HEK293 cells were lysed in SDC lysis buffer and the proteins were digested with trypsin. Phase transfer was subsequently used to remove SDC. One half of the digested peptides were used for phosphopeptide enrichment and analysis, while the other half was used for global proteome identification, as indicted by Figure 1.6. All samples were then analysed by LC-MS/MS.

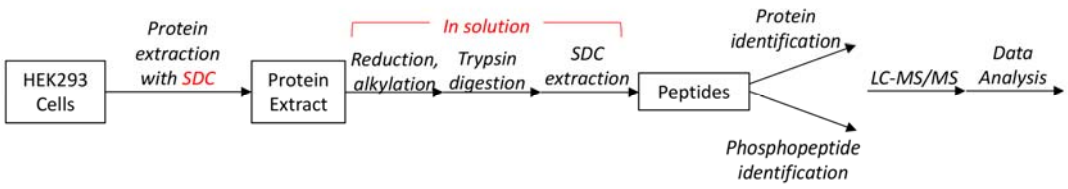


Figure 1.6. The experimental design of phosphopeptide or protein LFQ quantification.

1.3.3. Results

From the small-scale experiment, we started with 1 mg of protein and we identified a total of ~2,500 phosphopeptides across all samples (localisation probability of the phospho-site of > 0.75) as well as ~3,500 proteins as shown in Figure 1.7. The excellent reproducibility of the LFQ methods is illustrated for the repeats of the MCP-2 (one of the chemokines that we are interested in)-treated cells by the scatter plots in Figure 1.8 for both phosphopeptide and protein quantification, with Pearson correlation of 0.88 and 0.99, respectively. Similar results were obtained between all pairs of independent experiments under the same treatment conditions, with coefficient of determination higher than 0.85 (data not shown). Overall, the number of identified phosphopeptides or proteins was relatively low, but we suspect that this is due to the small amount of stating material that was used (~1 mg). Nevertheless, these experiments show that our SDC solubilization protocol is suitable for phosphoproteomic studies.

1.4. Summary of Selected Proteomic Methods

The pilot studies presented in this appendix, along with studies reported by other groups, guided our design of the proteomic methods used in the studies described in Chapters 3, 4 and 5 or this thesis. In summary, sodium deoxycholate was selected as the cell lysis detergent, high pH C18 fractionation was used to improve protein and phosphopeptide identification, and phosphopeptide and protein quantification was based on label free quantification. In addition to guiding the studies reported here, our pilot experiments have also contributed to the design of various other successful studies conducted (or ongoing) in the Monash Biomedical Proteomics Facility.

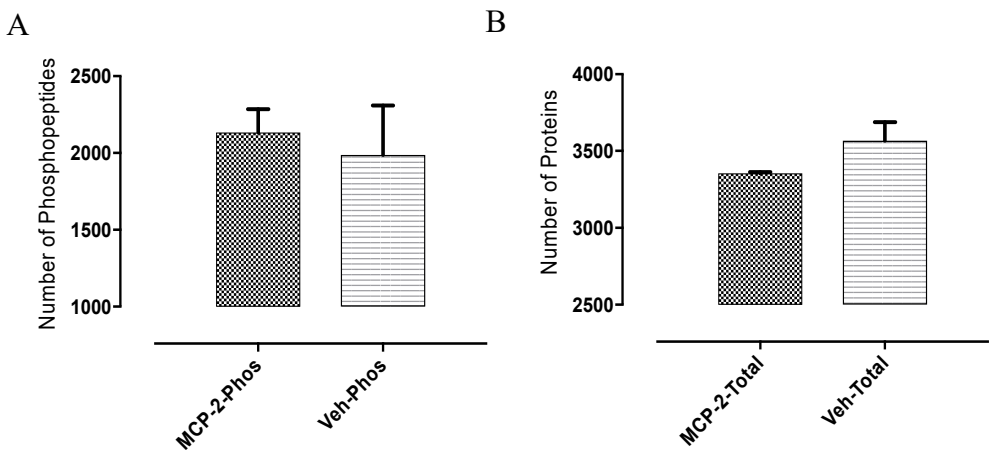


Figure 1.7 Reliably quantified phosphopeptides (A) and proteins (B) from MCP-2 or vehicle treated cells.

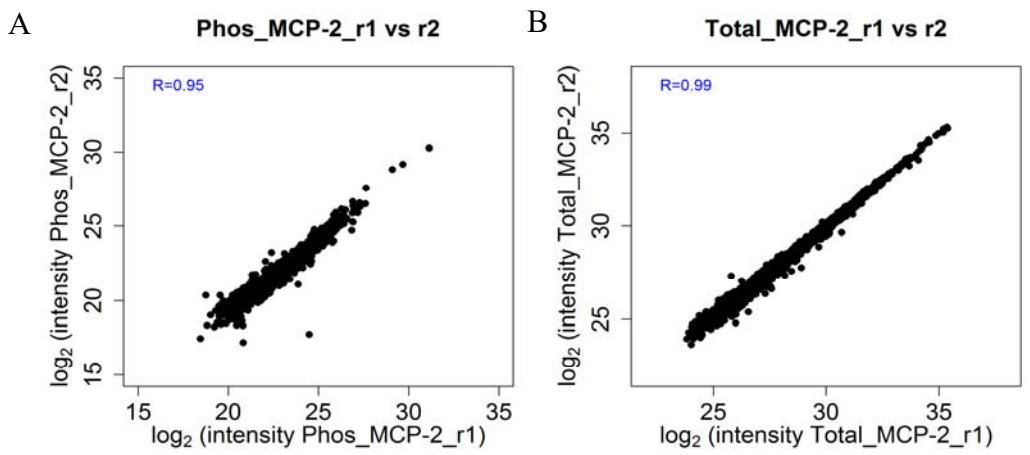


Figure 1.8. Scatter plot to show the reproducibility of LFQ. The figure presents the Pearson correlation of phosphopeptide quantification (A) and protein quantification (B) between repeat 1 and repeat 2. The Pearson correlation is indicated with blue number on the top left corner.

Appendix II Supplementary Material for Chapter 3

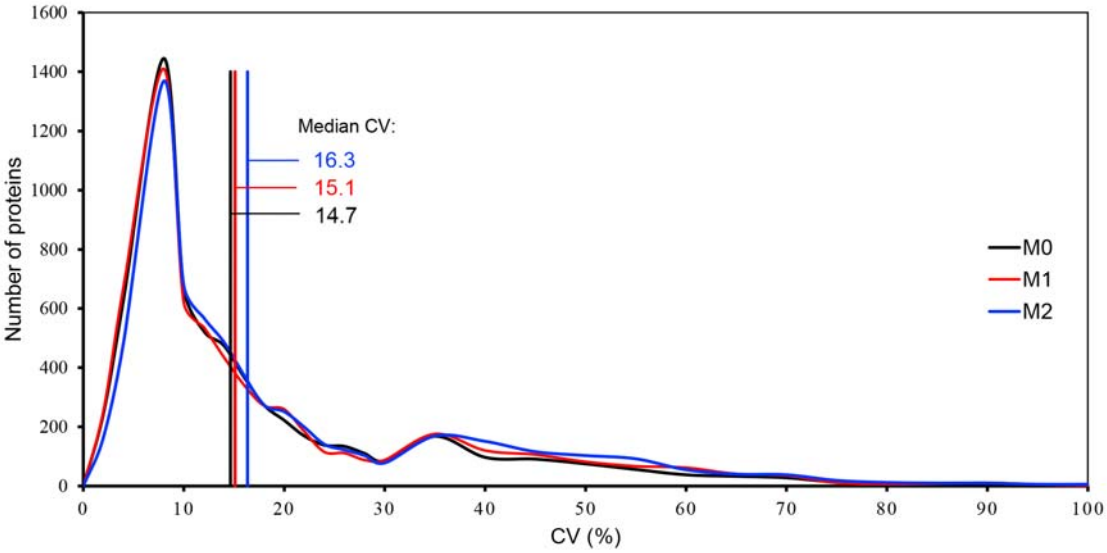


Figure S1. Coefficient of variation (CV) histograms for protein intensities under each treatment condition (black, M0; red, M1; blue, M2).

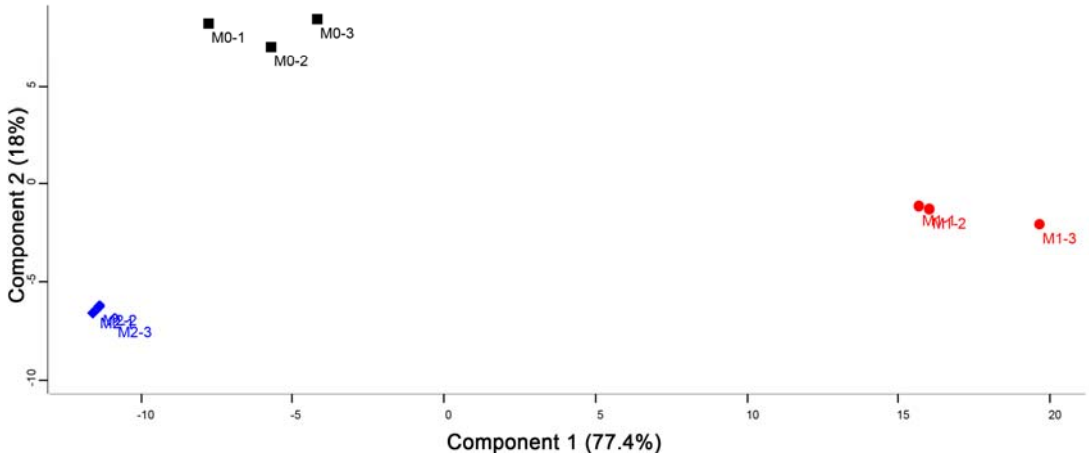


Figure S2. Principal component analysis, generated in Perseus, of significantly differentially expressed proteins. The graph shows the values of the first and second principal components for three repeats of each treatment condition (black, M0; red, M1; blue, M2).

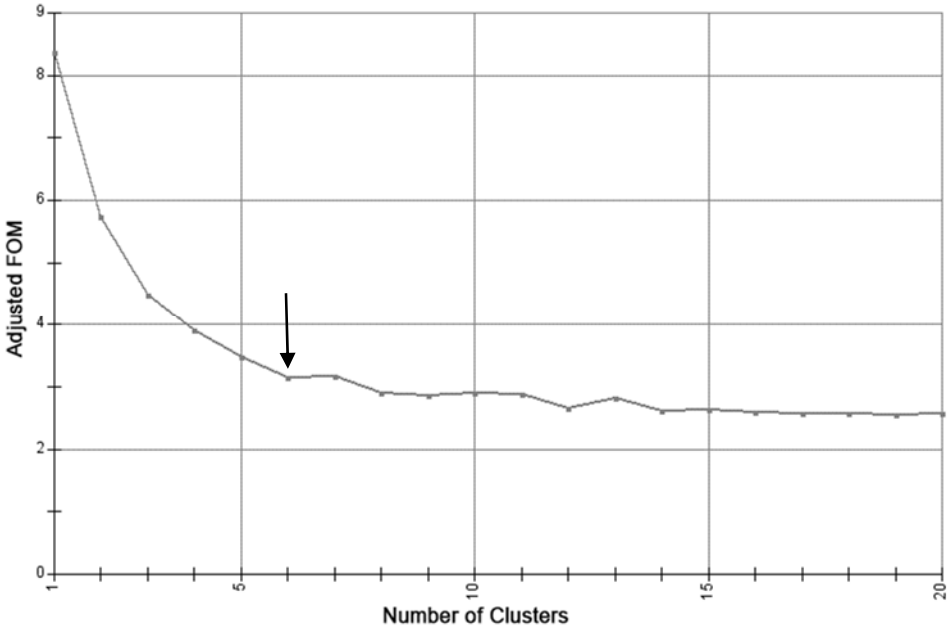


Figure S3. Figure of merit (FOM) for the clustering analysis obtained using TM4 MeV. The thick arrow indicates the FOM score for the six clusters used for classification of differentially expressed proteins in this study.

Appendix III Supplementary Material for Chapter 4

Table S4.1. Pearson correlation (R) for phosphopeptide quantification between each replicate.

	Veh_r1 [1]	Veh_r2	Veh_r3	MCP-1_r1 [2]	MCP-1_r2	MCP-1_r3	MCP-3_r1 [3]	MCP-3_r2	MCP-3_r3
Veh_r1	1	0.93	0.92	0.89	0.88	0.91	0.88	0.89	0.88
Veh_r2	0.93	1	0.95	0.91	0.9	0.92	0.9	0.91	0.91
Veh_r3	0.92	0.95	1	0.91	0.91	0.92	0.9	0.92	0.92
MCP-1_r1	0.89	0.91	0.91	1	0.94	0.93	0.91	0.92	0.92
MCP-1_r2	0.88	0.9	0.91	0.94	1	0.92	0.92	0.92	0.92
MCP-1_r3	0.91	0.92	0.92	0.93	0.92	1	0.91	0.92	0.92
MCP-3_r1	0.88	0.9	0.9	0.91	0.92	0.91	1	0.94	0.94
MCP-3_r2	0.89	0.91	0.92	0.92	0.92	0.92	0.94	1	0.95
MCP-3_r3	0.88	0.91	0.92	0.92	0.92	0.92	0.94	0.95	1

[1] Veh_r1 indicates cells were treated with vehicle control and replicate 1. [2] MCP-1_r1 indicates cells were treated with MCP-1 and replicate 1.

[3] MCP-3_r1 indicates cells were treated with MCP-3 and replicate 1.

Table S4.2. Pearson correlation (R) for protein quantification between each replicate.

	Veh_r1 [1]	Veh_r2	Veh_r3	MCP-1_r1 [2]	MCP-1_r2	MCP-1_r3	MCP-3_r1 [3]	MCP-3_r2	MCP-3_r3
Veh_r1	1	0.96	0.95	0.96	0.95	0.96	0.94	0.95	0.94
Veh_r2	0.96	1	0.96	0.96	0.96	0.97	0.96	0.96	0.96
Veh_r3	0.95	0.96	1	0.96	0.96	0.96	0.96	0.96	0.96
MCP-1_r1	0.96	0.96	0.96	1	0.96	0.96	0.95	0.96	0.95
MCP-1_r2	0.95	0.96	0.96	0.96	1	0.96	0.96	0.97	0.96
MCP-1_r3	0.96	0.97	0.96	0.96	0.96	1	0.95	0.96	0.95
MCP-3_r1	0.94	0.96	0.96	0.95	0.96	0.95	1	0.96	0.96
MCP-3_r2	0.95	0.96	0.96	0.96	0.97	0.96	0.96	1	0.96
MCP-3_r3	0.94	0.96	0.96	0.95	0.96	0.95	0.96	0.96	1

[1] Veh_r1 indicates cells were treated with vehicle control and replicate 1. [2] MCP-1_r1 indicates cells were treated with MCP-1 and replicate 1.

[3] MCP-3_r1 indicates cells were treated with MCP-3 and replicate 1.

Table S4.3. A table containing all the significantly up or down regulated phosphopeptides [1]

UniprotID	Gene name	Phosphopeptide Sequence	Phosphorylation residue	Phosphorylation Site	-log10(pvalue_MCP-1 vs Vehicle) [2]	log2(Fold change_MCP-1 vs Vehicle) [3]	-log10(pvalue_MCP-3 vs Vehicle)	log2(Fold change_MCP-3 vs Vehicle)	Involved in chemokine/receptor signalling [4]
O60716	CTND1	GSLASLDS[+80]LRK	['S']	[352]	2.26	1.45	1.41	1.67	Y (SDF1 [381])
P07947	YES	YRPENTPEPVSTSVSHYGAEPTTVS[+80]PC[+57]PS[+80]SSAK	['S', 'S']	[40, 44]	1.75	-1.29	3.01	1.35	Y (SDF1 [291])
Q13177	PAK2	S[+80]VIDPVPAPVGDSHVDGAAK	['S']	[197]	0.73	1.9	2.58	4.06	Y (RANTES [293])
P21333	FLNA	RAPS[+80]VANVGSHC[+57]DLSLK	['S']	[2152]	0.53	-0.46	2.42	1.25	Y (MCP-1 [292])
Q14247	SRC8	LPSS[+80]PVYEDAAS[+80]FK	['S', 'S']	[418, 426]	2.45	1.98	2.25	2.04	Y (MCP-1 [290])
P46940	IQGA1	ALQS[+80]PALGLR	['S']	[330]	3.03	-1.79	0.17	0.22	Y (CXCR4 [382])
P67870	CSK2B	IHPM[+16]AYQLQLQAASNFKS[+80]PVK	['S']	[209]	0.65	-0.17	2.06	-0.74	Y (CXCL8 [322])
Q14185	DOCK1	S[+80]QVMNVIGSER	['S']	[1743]	1.48	0.86	3.08	1.49	Y (CXCL12 [383])

Q14155	ARHG7	KS[+80]TAALEEDAQILK	['S']	[703]	1.28	0.49	2.73	1.1	N
Q14155	ARHG7	S[+80]TAALEEDAQILK	['S']	[703]	2.41	0.63	3.15	0.99	N
Q96DR7	ARHGQ	S[+80]PANGAVTLPAPPPPPVLRPPR	['S']	[127]	2.43	2.44	1.46	1.59	N
P12270	TPR	AADSQNSGEGNTGAAESS[+80]FSQEVSR	['S']	[2048]	1.87	-0.78	2.1	-0.9	N
P12270	TPR	TDGFAEAIHS[+80]PQVAGVPR	['S']	[2155]	3.84	1.92	3.64	1.68	N
P35658	NU214	S[+80]PGSTPTTPTSSQAPQK	['S']	[430]	2.57	-0.82	2.77	-1.67	N
P35658	NU214	TPS[+80]IQPSLLPHAAPFAK	['S']	[1023]	2.06	0.84	0.68	0.27	N
P49790	NU153	EGSVLDILKS[+80]PGFASPK	['S']	[614]	5.4	2.7	4.49	2.17	N
P49790	NU153	IPSIVSSPLNS[+80]PLDR	['S']	[338]	3.69	1.51	3.91	1.05	N
P49790	NU153	VQM[+16]TS[+80]PSSTGSPMFK	['S']	[516]	2.04	1.15	2.12	0.61	N
P49790	NU153	VQMTS[+80]PSSTGSPMFK	['S']	[516]	3.29	0.76	3.99	0.86	N
P52948	NUP98	ELAVGS[+80]L	['S']	[1816]	1.33	2.09	3.94	3.67	N
Q8WUM0	NU133	KGLPLGSAVSSPVLFS[+80]PVGR	['S']	[50]	3.03	-0.64	3.56	-0.74	N
Q9UKX7	NUP50	NSESESNKVAAETQS[+80]PSLFGSTK	['S']	[221]	2.59	0.97	3.04	0.86	N
Q9UKX7	NUP50	VAAETQS[+80]PSLFGSTK	['S']	[221]	3.14	1.26	3.07	1.09	N
P19634	SL9A1	ARIGS[+80]DPLAYEPK	['S']	[703]	2.37	1.28	2.22	0.75	N
P29317	EPHA2	VS[+80]IRLPSTSGSEGVPFR	['S']	[892]	2.99	-0.68	1.27	-1.13	N
Q04759	KPCT	ALINS[+80]M[+16]DQNMFR	['S']	[685]	3.54	0.68	0.57	0.13	N
Q04759	KPCT	ALINS[+80]MDQNMFR	['S']	[685]	3.07	0.68	3.42	0.45	N

Q04759	KPCT	NFS[+80]FM[+16]NPGM[+16]ER	['S']	[695]	1.32	-0.54	2.4	-1.35	N
Q04759	KPCT	NFS[+80]FM[+16]NPGMER	['S']	[695]	2.81	-0.52	3.17	-0.66	N
Q04759	KPCT	NFS[+80]FMNPGM[+16]ER	['S']	[695]	2.27	-0.52	2.59	-0.67	N
Q9P212	PLCE1	SGLLSTFGGS[+80]TGR	['S']	[490]	2.15	-0.72	2.64	-0.48	N
Q9Y2X7	GIT1	GVSASAVPFTPSS[+80]PLLSC[+57]SQEGSR	['S']	[571]	1.34	-0.69	2.74	-1.39	N
Q9Y2X7	GIT1	HGS[+80]GADSDYENTQSGDPLLGLEGK	['S']	[592]	2.71	1.06	2.18	0.64	N
Q9Y2X7	GIT1	HGS[+80]GADSDYENTQSGDPLLGLEGKR	['S']	[592]	2.19	0.9	3.66	0.64	N
O14639	ABLM1	GVS[+80]M[+16]PNM[+16]LEPK	['S']	[706]	0.32	-0.56	2.35	-2.14	N
O14639	ABLM1	GVS[+80]M[+16]PNMLEPK	['S']	[706]	0.28	-0.14	2.76	-1.02	N
O60716	CTND1	APS[+80]RQDVYGPQPQVR	['S']	[252]	3.65	-1.66	1.36	-0.56	N
O60716	CTND1	HYEDGYPGGSDNYGS[+80]LSR	['S']	[230]	0.5	-0.39	3.01	-0.98	N
O60716	CTND1	VGGS[+80]SVDLHR	['S']	[268]	1.21	-0.82	2.64	-1.33	N
O60716	CTND1	VGGSS[+80]VDLHR	['S']	[269]	0.65	-1.86	2.36	-1.27	N
O60716	CTND1	VRVS[+80]PQDANPLMANGT[+80]LTR	['S', 'T']	[47, 59]	2.36	-1.22	0.78	-0.55	N
O60716	CTND1	VSPQDANPLMANGT[+80]LTR	['T']	[59]	2.73	-0.78	3.01	-0.43	N
O75044	SRGP2	SAS[+80]TAGDIAC[+57]AFRPVK	['S']	[1013]	1.98	-0.6	2.94	-0.85	N
P08581	MET	S[+80]VS[+80]PTTEMVSNESVDYR	['S', 'S']	[988, 990]	3.66	2.1	0.38	0.87	N

P08581	MET	S[+80]VSPTTEMVSNESVDYR	['S']	[988]	0.53	-0.55	2.29	-0.9	N
P25054	APC	FQPHFGQQEC[+57]VS[+80]PYR	['S']	[1100]	1.45	-0.66	2.84	-1.18	N
P25054	APC	QAS[+80]SDSDSILSLK	['S']	[2129]	2.43	-0.95	2.22	-0.68	N
P25054	APC	QASS[+80]DSDSILSLK	['S']	[2130]	2.44	-0.96	2.12	-0.66	N
P25054	APC	VTPFNYNPS[+80]PR	['S']	[2789]	2.85	-0.76	2.87	-0.48	N
P35221	CTNA1	S[+80]RTSVQTEDDQLIAGQSAR	['S']	[652]	1.28	-0.96	3.27	1.52	N
P55196	AFAD	EELSSGDSLS[+80]PDPWKR	['S']	[1512]	2.2	-2.58	2.4	-3.8	N
P55196	AFAD	KS[+80]DS[+80]DMWINQSSSLDSSTSSQEHLNHSSK	['S', 'S']	[1300, 1302]	1.53	-1.53	2.4	-2.38	N
P55196	AFAD	SS[+80]PNVANQPPS[+80]PGGK	['S', 'S']	[1173, 1182]	2.29	1.05	2.45	-3.23	N
Q07157	ZO1	S[+80]REDLSAQPVQTK	['S']	[617]	1.24	-0.71	2.47	-1.15	N
Q8N8S7	ENAH	QNS[+80]QLPAQVQNGPSQEELEIQR	['S']	[125]	2.64	-0.74	1.76	-0.28	N
Q8N8S7	ENAH	QNSQLPAQVQNGPS[+80]QEELEIQR	['S']	[136]	2.77	-0.77	2.24	-0.35	N
Q92558	WASF1	PIPTC[+57]ISSATGLIENRPQS[+80]PATGR	['S']	[310]	2.1	-0.51	2.45	-0.6	N
Q9UHD8	SEPT9	S[+80]FEVEEVETPNS[+80]TPPR	['S', 'S']	[30, 41]	3.2	0.96	2.47	0.96	N
Q9UHD8	SEPT9	S[+80]FEVEEVETPNSTPPR	['S']	[30]	3.38	0.6	3.39	0.41	N
P21333	FLNA	C[+57]SGPGLS[+80]PGM[+16]VR	['S']	[1459]	0.11	-0.15	2.04	-1.35	N
P01106	MYC	KFELLPT[+80]PPLS[+80]PSR	['T', 'S']	[58, 62]	1.84	0.49	2.81	0.69	
P01106	MYC	SGLC[+57]S[+80]PSYVAVTPFSLR	['S']	[71]	2.23	0.91	1.61	0.65	

P10398	ARAF	QHEAPS[+80]NRPLNELLTPQGPS[+80]PR	['S', 'S']	[172, 186]	3.49	3.88	3.68	4.07	
P10398	ARAF	QHEAPSNRPLNELLT[+80]PQGPS[+80]PR	['T', 'S']	[181, 186]	3.27	4.51	3.48	4.67	
P10398	ARAF	QQFYHS[+80]VQDLSGGSR	['S']	[157]	1.07	-1.02	3.61	-1.63	
P27361	MK03	IADPEHDHTGFLT[+80]EY[+80]VATR	['T', 'Y']	[202, 204]	4.73	4.46	5.06	4.41	
P27361	MK03	IADPEHDHTGFLTEY[+80]VATR	['Y']	[204]	3.86	1.93	4.53	1.87	
Q07889	SOS1	KSDHGNAFFPNSPSPFTPPPPQT[+80]PS[+80]PHGTR	['T', 'S']	[1263, 1265]	0.58	1.47	2.85	2.08	
P05412	JUN	AKNS[+80]DLLTS[+80]PDVGLLK	['S', 'S']	[58, 63]	2.47	3.03	2.17	2.3	
P40763	STAT3	FIC[+57]VT[+80]PTTC[+57]SNTIDLPM[+16]S[+80]PR	['T', 'S']	[714, 727]	2.62	1.46	2.33	1.16	
P40763	STAT3	FIC[+57]VT[+80]PTTC[+57]SNTIDLPMS[+80]PR	['T', 'S']	[714, 727]	1.86	0.91	2.02	0.94	
P40763	STAT3	FIC[+57]VTPTTC[+57]SNTIDLPM[+16]S[+80]PR	['S']	[727]	3.78	1.49	2.76	0.92	
Q9Y2I7	FYV1	AEGGQGEQQPLSGSWTS[+80]PQLPSR	['S']	[76]	3.97	-1.46	4.08	-1.4	
O43524	FOXO3	[+42]AEAPAS[+80]PAPLSPLEVELDPEFEPQSRPR	['S']	[7]	1.19	-3.95	2.29	-4.53	
O43524	FOXO3	AALQTAPESADDS[+80]PSQLSK	['S']	[284]	1.35	0.99	2.37	1.34	
O43524	FOXO3	S[+80]SDELDAWTDFR	['S']	[299]	2.15	-0.62	1.63	-0.46	
O43524	FOXO3	SS[+80]DELDAWTDFRS[+80]R	['S', 'S']	[300, 311]	1.31	-2.52	3.27	-3.39	

O75385	ULK1	AGGTSSPSPVVFTVGS[+80]PPSGST[+80]PPQGPR	['S', 'T']	[758, 764]	0.97	-2.41	2.15	-1.8	
P20042	IF2B	[+42]S[+80]GDEM[+16]IFDPTM[+16]SKK	['S']	[2]	0.27	-0.1	2.17	-0.85	
P20042	IF2B	IES[+80]DVQEPTEPEDDLDIM[+16]LGNK	['S']	[105]	3.1	-0.63	2.19	-0.89	
P20042	IF2B	IESDVQEPT[+80]EPEDDLDIM[+16]LGNK	['T']	[111]	2.56	-0.37	3.66	-1.07	
P23443	KS6B1	T[+80]PVS[+80]PVKFSPGDFWGR	['T', 'S']	[444, 447]	4.54	1.87	3.71	1.55	
P23443	KS6B1	TPVS[+80]PVKFS[+80]PGDFWGR	['S', 'S']	[447, 452]	2.77	2.06	1.74	1.4	
P23588	IF4B	ERHPS[+80]WR	['S']	[406]	2.05	-1.12	1.8	1.25	
Q13541	4EBP1	RVVLGDGVQLPPGDY[+80]STTPGGTLFSTT[+80]PGGTR	['Y', 'T']	[34, 46]	1.71	-0.89	3.1	-2.58	
Q13541	4EBP1	VVLGDGVQLPPGDYSTTPGGTLFSTT[+80]PGGTR	['T']	[46]	2.67	-4.8	0.89	-2.45	
Q14152	EIF3A	AREES[+80]WGPPR	['S']	[1188]	2.28	-2.64	1.49	-1.74	
Q92934	BAD	HSS[+80]YPAGTEDDEGM[+16]GEEPSPFR	['S']	[75]	0.24	-0.14	2.11	-0.64	
Q9HBH9	MKNK2	AS[+80]LSSAPVVLVGDHA	['S']	[452]	3.42	1.31	4.01	1.53	
Q9Y6A5	TACC3	AM[+16]TLS[+80]PQEEVAAGQMASSSR	['S']	[317]	1.34	-0.28	3.06	-0.76	

[1] The table contains the information, including Uniprot accession number, gene name, which network the protein belong to, phosphorylated residue and position; [2] $-\log_{10}$ transformed (p value); [3] $\log_2$ transformed (fold change); and [4] whether the phosphosite is connected to chemokine:receptor signalling, indicated by Y (Yes) and N (No).

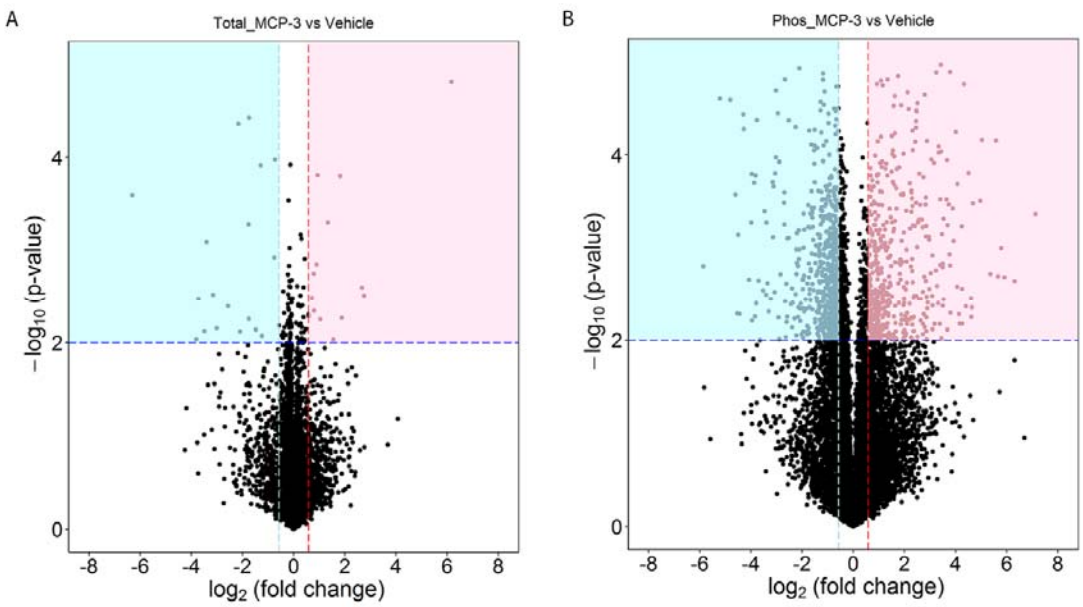


Figure S4.1. Volcano plot to compare two different conditions; A: comparison between vehicle and MCP-3 activated cells in protein quantification; B: comparison between vehicle and MCP-3 activated cells in phosphopeptide quantification. The x-axis is $\log_2$-transformed fold change and the y-axis is $-\log_{10}$ (p value); each black dot indicates one protein or phosphopeptide. The horizontal blue dash line indicates p = 0.01; the vertical grey dashed line (left) indicates a 1.5-fold decrease and the vertical red dash line (right) indicates a 1.5-fold increase. sThe blue and pink areas highlight proteins or phosphopeptides that are significantly decreased or increased, respectively.

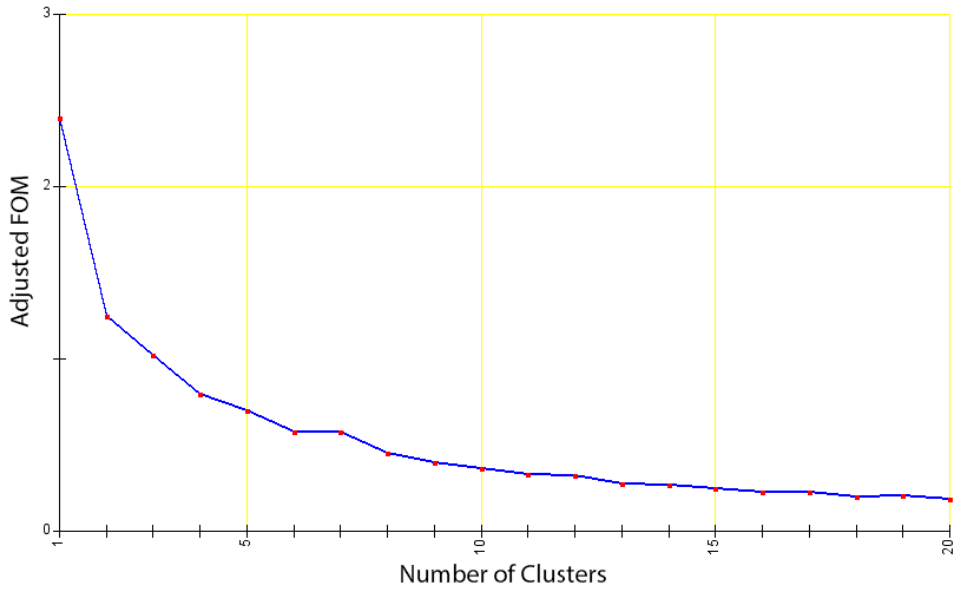


Figure S4.2. Figure of merits for clustering analysis. The significantly up or down regulated phosphopeptides were analysed with figure of merits to indicate the best numbers of clustering. When clustering the up- and down-regulated phosphopeptides to more than 6 clusters, the adjusted FOM or the variance between clusters become smaller, indicating less difference will be observed when assigning to more than 6 clusters.

Appendix IV Supplementary Material for Chapter 5

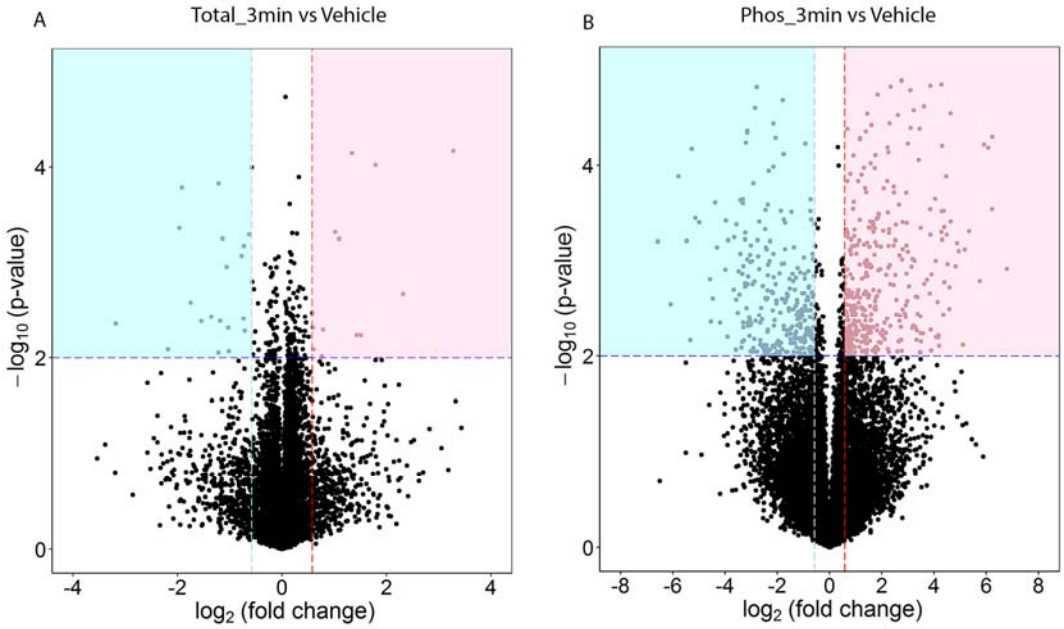


Figure S5.1. Volcano plots showing the effects of 3 minutes' treatment.

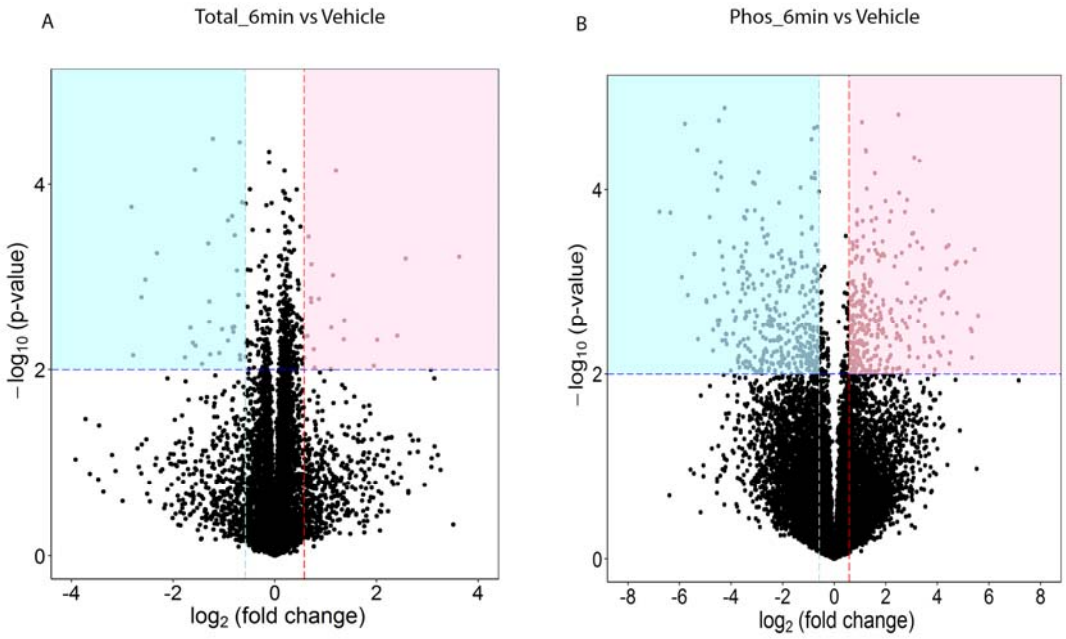


Figure S5.2. Volcano plots showing the effects of 6 minutes' treatment.

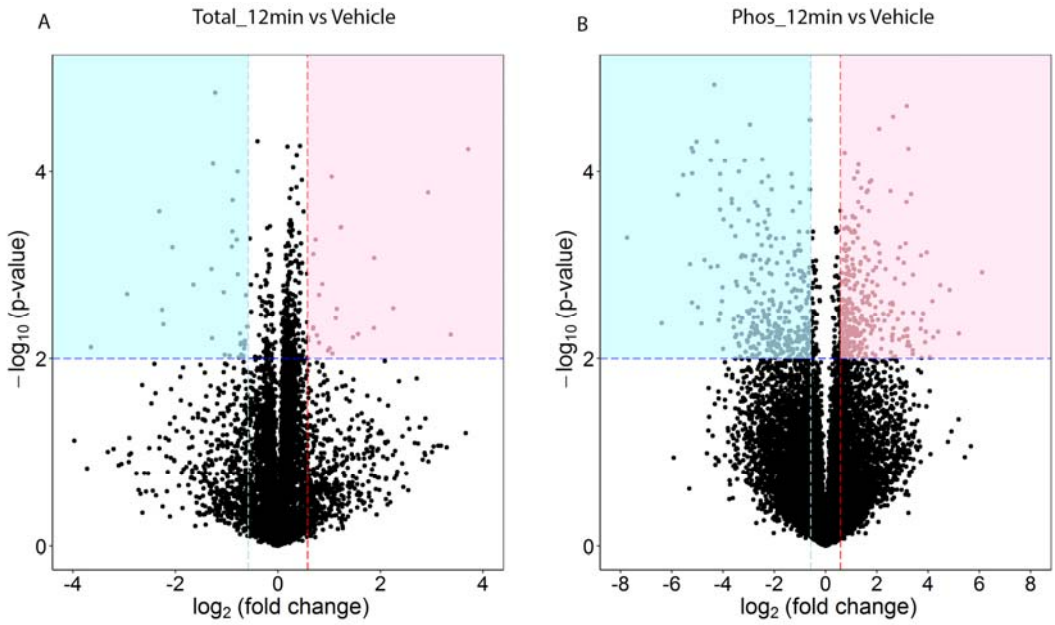


Figure S5.3. Volcano plots showing the effects of 12 minutes' treatment.

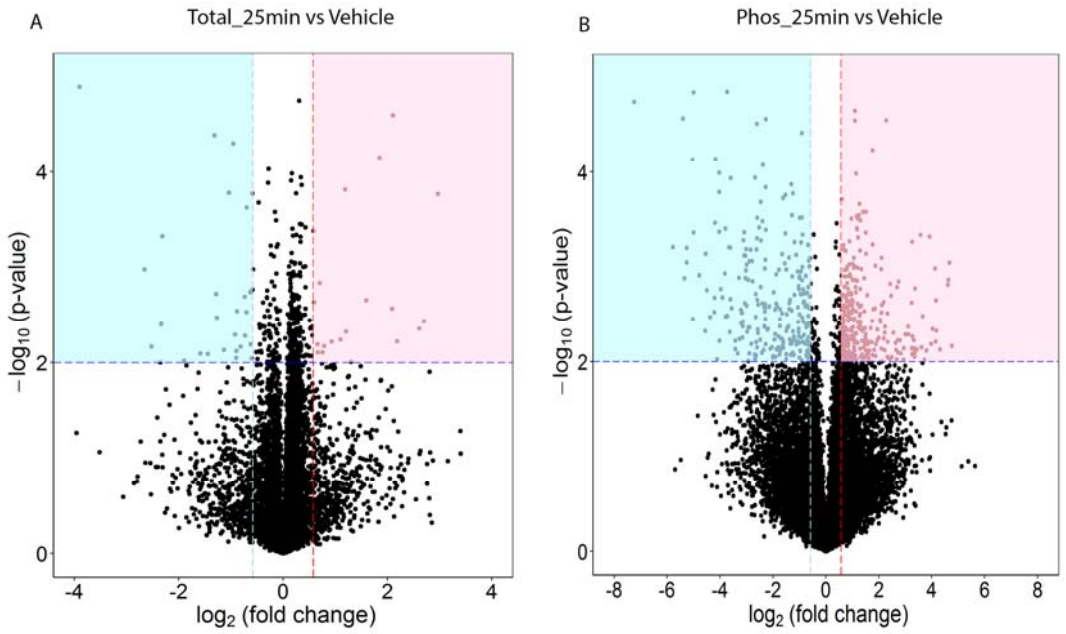


Figure S5.4. Volcano plots showing the effects of 25 minutes' treatment.

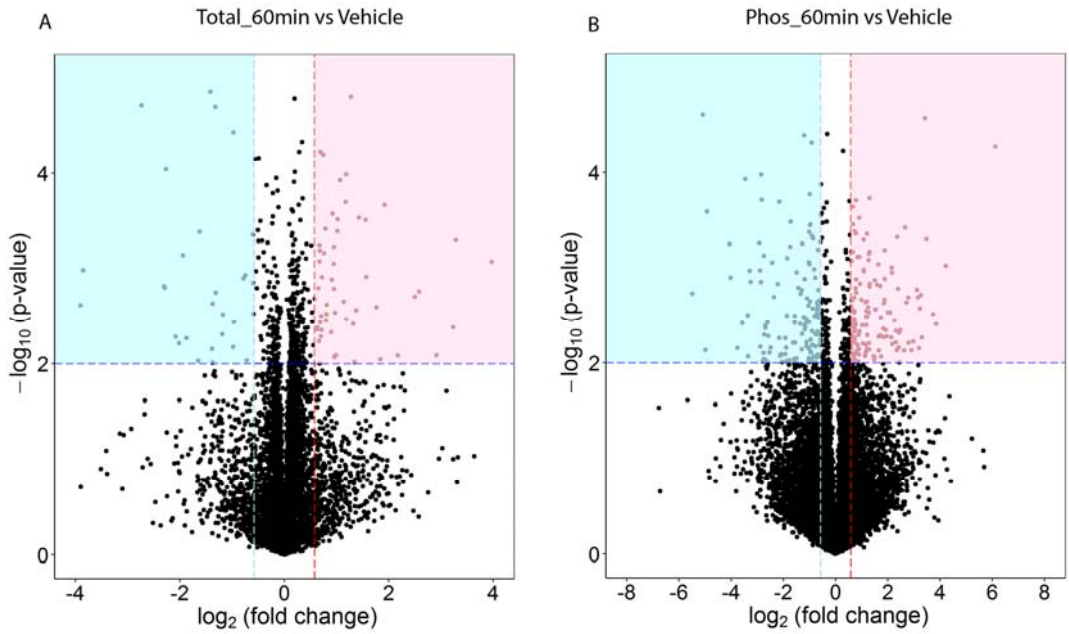


Figure S5.5. Volcano plots showing the effects of 60 minutes' treatment.

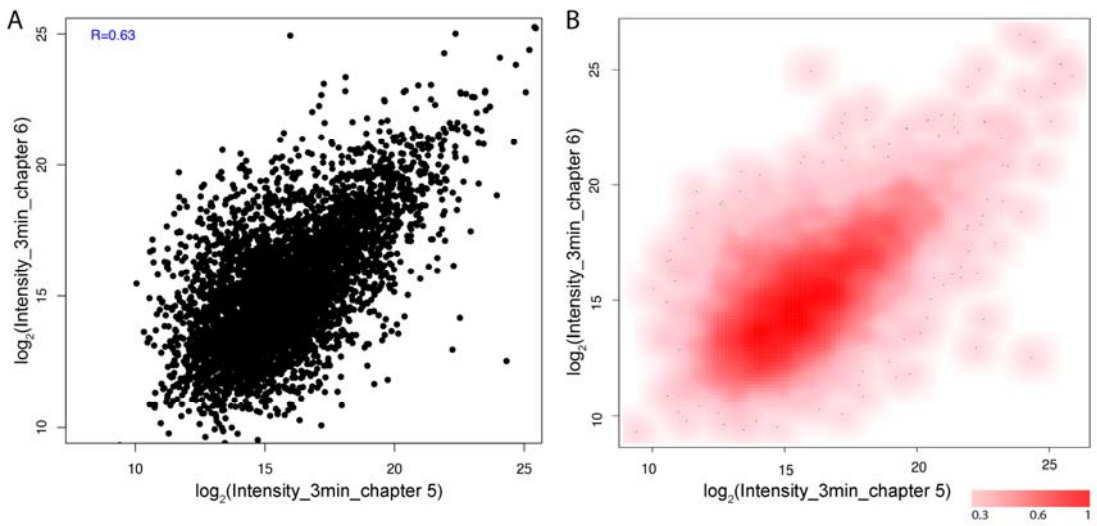


Figure S5.6 Scatter plot (A) and density plot (B) to present correlation of quantified phosphopeptides between chapter 4 and 5. The $\log_2$ transformed of overlaid phosphopeptides were plotted in R (A) and the Pearson correlation is as indicated on the top left with blue value; the density was estimated based on Pearson correlation and plotted in R (B).

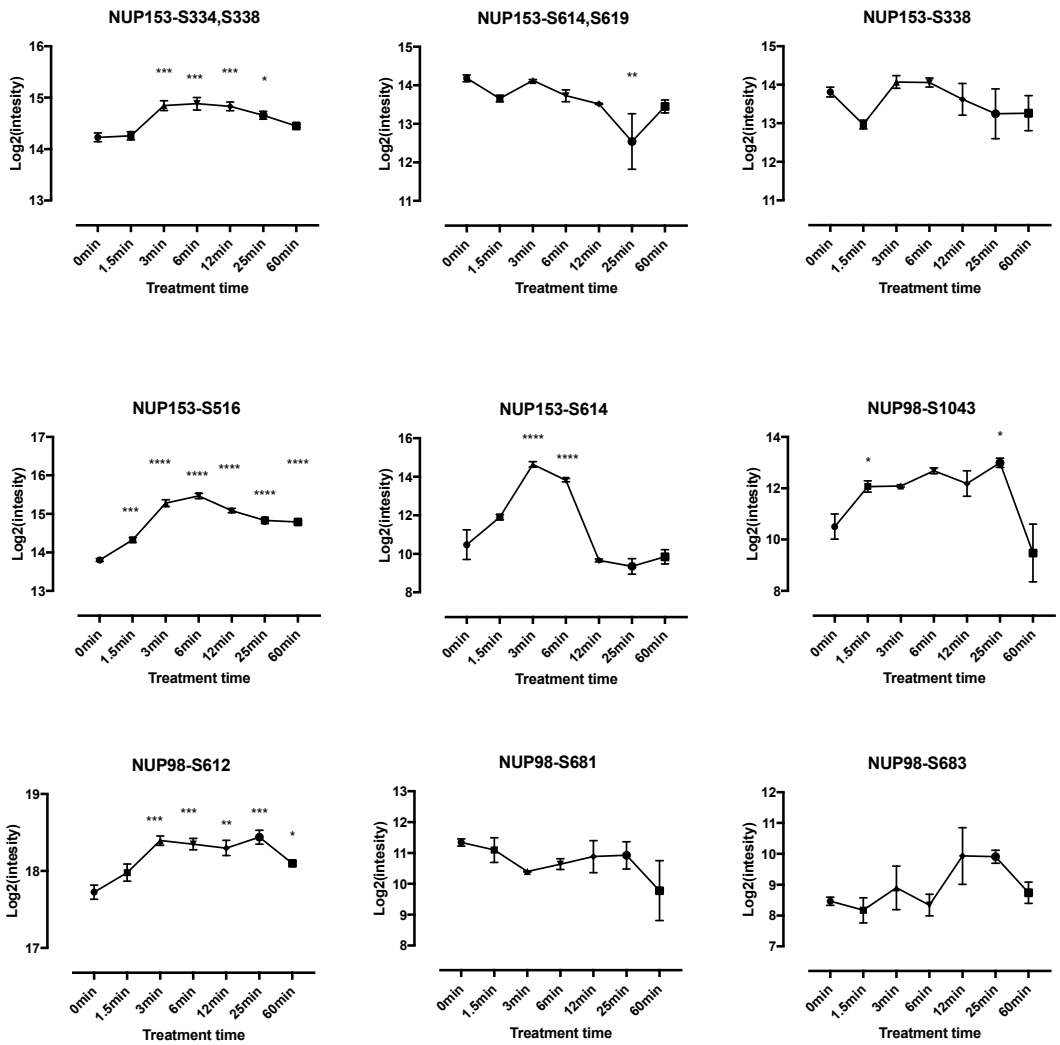


Figure S5.7. Phosphorylation dynamics of NUP153 and NUP98 at different residues. All data are presented as mean ± SEM, n=3. * P < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.00001 vs time 0min (one-way ANOVA followed by Dunnett's post hoc test)

Appendix V Python Script to Extract Phospho-residue

```
def fastaseqs(fastadb):
    acc = ""
    sequence=""
    with open(fastadb,"r") as db:
        for line in db:
            if line[0]==">":
                if not acc==="":
                    yield acc,sequence
                    sequence = ""
                acc = line[1:].strip().split("|",2)[1]
            else:
                sequence = sequence+line.strip()
    yield acc, sequence

fastadict = {a:s for (a,s) in fastaseqs("D:\Analysis_Data\chengh\S16_0129\Positions\uniprot_human_SwissProt_100516.fasta")}

targets = ["[+80]"]
start=False
with open("D:\Analysis_Data\chengh\S16_0129\Positions\Position_TC.tsv","wb") as fo:
    with open("D:\Analysis_Data\chengh\S16_0129\Positions\TC.csv","r") as fi:
        for line in fi:
            if start:
                entry = line.split(",")
                acc = entry[0].split("_")[0]
                pep = entry[1].translate(None,'][+1234567890_')
                phosphos = entry[1].split("[+80]")
                nophos = len(phosphos)-1
                for i,p in enumerate(phosphos):
                    phosphos[i]=p.translate(None,'][+1234567890_')
```

```
                    peploc = fastadict[acc].find(pep)
                    phosloc = []
                    phosprotloc = []
                    phosres = []
                    phosresloc = []
                    for i in range(len(phosphos)-1):
                        phosloc.append(len("".join(phosphos[:i+1])))
                        phosprotloc.append(peploc+phosloc[-1])
                        phosres.append(phosphos[i][-1])
                        phosresloc.append("-".join((phosres[-1],str(phosprotloc[-1]))))
                    fo.write("\t".join(str(x)                for                x                in
[acc,entry[1],phosres,phosprotloc,str(len(phosres)),phosresloc]))
                    fo.write("\n")
                else:
                    #Header

fo.write("ProteinAcc\tPeptideSeq\tPhosphoRes\tPosition\tMultiplicity\tResPos\n")

                start = True
```