Post translational modifications regulate the function of E4bp4

Tomasz Kostrzewski

Imperial College London
Department of Life Sciences
Supervised by Dr Hugh Brady

Submitted for the degree of Doctor of Philosophy

2014

Dedicated to

Ryszard Boguslaw Kostrzewski

(1952 - 2010)

Declaration of originality

The data presented in this thesis is all my own work, where information and data has been derived from other sources it is indicated within the thesis, with the following exceptions:

- Mass spectrometry experiments described in Chapters 3 and 4 were performed in collaboration with Dr Peter Di Maggio and Aaron Borg (Department of Chemical Engineering, Imperial College London).
- In Chapter 5, undergraduate student Alexis Shih contributed to the cloning of E4bp4 binding sites into the pGL3 luciferase vector (Figure 5.6).

Tomasz Kostrzewski

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Abstract

The basic leucine zipper transcription factor E4bp4 is essential for various immunological processes, most notably the development of natural killer (NK) cells. E4bp4 is required at the point of lineage commitment when NK cell progenitors develop from common lymphoid progenitors. E4bp4 promotes NK cell development by directly regulating the expression of other transcription factors, including Eomes and Id2. Despite its critical role, little is known about how the functions of the E4bp4 protein are regulated within the cellular environment. This study demonstrates that E4bp4 is post translationally modified by phosphorylation and SUMOylation and that these modifications directly regulate the protein's function. In the absence of either modification, E4bp4 is both a more potent transcriptional activator and transcriptional repressor. In NK cells, the absence of post translational modifications on E4bp4 promotes the expression of target genes Notch1 and E2A. Most strikingly, when expressed in hematopoietic progenitor cells, versions of E4bp4 lacking phosphorylation or SUMOylation sites, promote the development of up to three times more NK cells than the wild type form of E4bp4. This work has implications for the production of NK cells for use in immunotherapy and provides the first evidence that SUMOylation of an individual transcription factor protein can regulate a highly complex cellular process.

Acknowledgements

First I would like to thank my PhD supervisor, Dr Hugh Brady, for giving me the opportunity to undertake research in his lab. I am hugely grateful for all of his advice, for his unwavering enthusiasm and for the freedom he has given me in my research, which has all contributed to making me the scientist I am today.

Thank you to my two advisors Professor Erhard Hohenester and Professor Murray Selkirk, for their guidance and support. A big thank you also to Dr Pete Di Maggio for providing assistance with mass spectrometry experiments and for critical input on this part of my project, our discussions were extremely useful for the progress of the project.

I am grateful to all the members of the Brady lab, past and present, who have supported me and provided an engaging and enjoyable research environment. My thanks go to Dr Victoria Male and Dr Ilaria Nisoli for their technical assistance and support during the early part of my PhD and to Yiran Meng for his help with experiments, useful discussions and friendship. I would also like to acknowledge the various undergraduate and postgraduate students that I mentored during my PhD; working with all of you was very enjoyable. My thanks also go to Catherine Simpson and Jane Srivastava who ran the flow cytometry and FACS sorting facility.

Thank you to all my colleges and friends from across Imperial College, who have made it such an enjoyable place to work, in particular members of the Lo Celso and Guerra labs. A special mention must go to Kathryn Lagrue and Charlotte Barthen my fellow PhD students, who despite leaving for ‘The North’ have been great friends since the first day that I arrived at Imperial. Also to Dr Michael Povelones who taught me to strive for “the perfect western”.

I would like to thank my family, without whom nothing would be possible. Thank you for your constant love and encouragement. Finally, to Nicola, I am so grateful for your unwavering support and understanding. You had amazing patience during the busy, tough times and your reading of my thesis was invaluable. You are my rock; I could not have done this without you.

Funding was provided by the Biotechnology and Biological Sciences Research Council (BBSRC).

Table of Contents

	Declaration of originality	3
	Copyright Declaration	4
	Abstract	5
	Acknowledgements	6
	Table of Contents	7
	List of Figures	11
	List of Tables	13
	List of Abbreviations	14

Chapter 1 Introduction

1.1	Natural killer cells	18
1.1.1	NK cell functions	18
1.1.2	Innate lymphoid cells	20
1.1.3	NK cells and immunotherapy	23
1.2	NK cell development	24
1.2.1	Murine NK cell development	25
1.2.2	Human NK cell development	27
1.3	Regulation of NK cell development	28
1.3.1	E4bp4	30
1.3.2	Id2	31
1.3.3	Ets1	32
1.3.4	Tox	33
1.3.5	Tbet and Eomes	33
1.3.6	Runx3	34
1.3.7	PU.1	35
1.3.8	Gata3	35
1.3.9	Ikaros	35
1.3.10	MEF	36
1.3.11	IRF-2	36
1.3.12	Blimp1	36
1.3.13	Helios	37
1.3.14	The transcriptional network controlling NK cell development	37
1.4	E4bp4/NFIL3	38
1.5	Other functions of E4bp4	41
1.5.1	$CD8\alpha^+$ conventional dendritic cell development	41
1.5.2	B cell class switching	42
1.5.3	T cell cytokine production and activity	43
1.5.4	Macrophage activation and cytokine production	44
1.5.5	Circadian rhythm and T_H17 development	44
1.5.6	Cell survival	45
1.5.7	Motor neuron development and survival	46
1.5.8	Osteoblast function	46

1.6 Regulation of E4bp4 47
1.6.1 Transcriptional regulation of *E4bp4* 47
1.6.2 Post transcriptional regulation of E4bp4 48
1.7 Project aims and research questions 54

Chapter 2 Materials and Methods

2.1 Buffers and reagents 56
2.2 Cells and cell culture 60
2.2.1 Cell lines 60
2.2.2 Transfections 60
2.2.3 Mice 61
2.3 Molecular biology 62
2.3.1 DNA electrophoresis 62
2.3.2 Polymerase chain reaction 62
2.3.3 Restriction enzyme digests 63
2.3.4 Ligation and transformation of competent bacteria 63
2.3.5 Colony PCR 63
2.3.6 Plasmid preparations and sequencing 63
2.3.7 Site directed mutagenesis 64
2.3.8 Cloning of mutant E4bp4 cDNA into pMSCV-IRES-hCD2 65
2.3.9 Cloning of mutant E4bp4 cDNA into pCSGW 65
2.3.10 Cloning of FLAG-E4bp4 66
2.4 SDS-PAGE and Western blotting 67
2.5 Flow cytometry and cell sorting 67
2.6 Analysis of SUMOylation *in vivo* 69
2.7 Immunoprecipitation of FLAG-E4bp4 70
2.8 Optimisation of FLAG-E4bp4 expression and purification 71
2.9 Mass spectrometry analysis 72
2.10 *In vitro* development of NK cells from transduced lineage negative bone marrow cells 74
2.11 Cycloheximide time course analysis of protein stability 75
2.11.1 Creation of stably transduced cell lines 75
2.11.2 Cycloheximide time course 75
2.12 Luciferase assay 76
2.13 Regulation of E4bp4 target genes in NK cells 76
2.13.1 Phenotyping of MNK-1 cells 76
2.13.2 Lentiviral transduction of MNK-1 cells 77
2.13.3 Quantitative PCR 77
2.14 *In vitro* development of NK cells from transduced human umbilical cord blood stem cells 78

Chapter 3 Regulation of E4bp4 by SUMOylation

3.1 Introduction 81
3.1.1 Regulation of E4bp4 81
3.1.2 SUMOylation 81
3.1.3 SUMOylation of target proteins 82
3.1.4 Effect of SUMOylation 84
3.1.5 SUMOylation and transcription 86
3.1.6 Studying SUMOylation 87
3.1.7 SUMOylation of E4bp4 88
3.1.8 Aims and objectives 89
3.2 E4bp4 has five potential SUMOylation sites 89
3.3 E4bp4 is multiply SUMOylated 90
3.4 Mutating putative SUMOylation sites of E4bp4 abolishes modification of the protein 98
3.5 Analysis of E4bp4 SUMOylation by mass spectrometry 104
3.6 Removing SUMOylation sites of E4bp4 causes increased NK cell production 116
3.7 Discussion 122

Chapter 4 Regulation of E4bp4 by phosphorylation

4.1 Introduction 130
4.1.1 Regulation of transcription factors by phosphorylation 130
4.1.2 Phosphorylation of E4bp4 133
4.1.3 Aims and objectives 136
4.2 Prediction of E4bp4 phosphorylation 136
4.3 E4bp4 is multiply phosphorylated 138
4.4 Mutating E4bp4 phosphorylation sites 144
4.5 Removing phosphorylation sites of E4bp4 causes increased NK cell production 145
4.6 Phosphorylation by Casein Kinase 1ε is important during NK cell development 149
4.7 Discussion 150

Chapter 5 Post translational modifications influence the transcriptional activity of E4bp4

5.1 Introduction 157
5.2 The PTMs of E4bp4 do not affect its stability or proteasomal degradation 158
5.3 PTM of E4bp4 regulate its transcriptional activity 163
5.3.1 Transcriptional repression of target genes by E4bp4 163
5.3.2 Transcriptional activation of target genes by E4bp4 166
5.4 PTMs of E4bp4 influence target gene expression in NK cells 171

5.5 Discussion 177

Chapter 6 Conclusions and future work

6.1 Conclusions 185
6.2 Signals that influence the PTMs and activity of E4bp4 186
6.3 E4bp4 and human NK cell development 187
6.4 E4bp4 SUMOylation and infection 191
6.5 Interactions between E4bp4 PTMs 193
6.6 Summary 194

References 195

List of Figures

Figure 1.1	Murine innate lymphoid cell subsets	22
Figure 1.2	Stages of conventional murine Natural Killer cell development	26
Figure 1.3	Stage of human Natural Killer cell development	28
Figure 1.4	Transcription factors that regulate conventional NK cell development	30
Figure 1.5	Transcriptional network regulating NK cell development	38
Figure 1.6	Domain structure of E4bp4	39
Figure 1.7	Functions of E4bp4 in the immune system	42
Figure 1.8	Complexity of the mammalian proteome	50
Figure 2.1	Genotyping *E4bp4*$^{-/-}$ mice	61
Figure 2.2	Phenotype of *E4bp4*$^{-/-}$ mice	62
Figure 3.1	Three-dimensional structure of SUMO isoforms	82
Figure 3.2	The SUMO conjugation cycle	83
Figure 3.3	Effects of SUMO modification on protein activity	85
Figure 3.4	Putative SUMOylation site of E4bp4 are highly conserved across a range of species	92
Figure 3.5	Detection of E4bp4 SUMOylation *in vivo*	93
Figure 3.6	E4bp4 has multiple SUMO modifications by both SUMO1 and SUMO2/3	94
Figure 3.7	Cloning and expression of FLAG-E4bp4	96
Figure 3.8	FLAG-E4bp4 purification confirms E4bp4 has multiple SUMO modifications	97
Figure 3.9	Mutating lysine residues to arginine abolishes SUMO modifications	99
Figure 3.10	Mutation of putative E4bp4 SUMOylation sites abolishes SUMO2 modification	101
Figure 3.11	Mutation of putative E4bp4 SUMOylation sites abolishes SUMO3 modification	102
Figure 3.12	Mutation of putative E4bp4 SUMOylation sites abolishes SUMO1 modification	103
Figure 3.13	Analysing protein post translational modification using mass spectrometry	105
Figure 3.14	Trypsin and Glu-C double digest provides a method to identify SUMO modified peptides	107
Figure 3.15	Optimisation of FLAG-E4bp4 expression	109
Figure 3.16	Optimisation of FLAG-E4bp4 purification for mass spectrometry analysis	111
Figure 3.17	Purification of E4bp4 for mass spectrometry analysis	113
Figure 3.18	Confirmation of E4bp4 SUMO2/3 modification at lysine 219	114
Figure 3.19	Analysis of synthetic E4bp4-SUMO2/3 peptide standard	115
Figure 3.20	NK cell production from lineage negative bone marrow cells, with transduction	117
Figure 3.21	E4bp4 SUMO mutants were cloned into the MSCV-IRES-hCD2 vector	118
Figure 3.22	Removing SUMO modifications from E4bp4 increases *in vitro* NK cell production from *E4bp4*$^{-/-}$ Lin^{-} BM	120

Figure 3.23	Removing SUMO modifications from E4bp4 increases *in vitro* NK cell production from WT Lin^- BM	121
Figure 4.1	Phosphorylation affects the function of transcription factors	131
Figure 4.2	Putative phosphorylation sites of E4bp4 are highly conserved	137
Figure 4.3	E4bp4 is phosphorylated	139
Figure 4.4	Serine 301 is a site of E4bp4 phosphorylation	141
Figure 4.5	Serine 353 is a site of E4bp4 phosphorylation	142
Figure 4.6	Serine 285 or 286 is a site of E4bp4 phosphorylation	143
Figure 4.7	Mutating serine residues to alanine abolishes phosphorylation	145
Figure 4.8	Removing phosphorylation sites from E4bp4 increases *in vitro* NK cell production from *E4bp4*$^{-/-}$ Lin^- BM	147
Figure 4.9	Removing phosphorylation sites from E4bp4 increases *in vitro* NK cell production from WT Lin^- BM	148
Figure 4.10	Inhibition of Casein Kinase I promotes NK cell development	150
Figure 5.1	Development of stable cell lines expressing E4bp4 mutants	161
Figure 5.2	SUMOylation and phosphorylation do not influence the stability of E4bp4	162
Figure 5.3	Dual luciferase reporter assay system	164
Figure 5.4	SUMOylation affects transcriptional repression by E4bp4	166
Figure 5.5	Phosphorylation affects transcriptional repression by E4bp4	167
Figure 5.6	Cloning of Zhang *et al.* E4bp4 binding sites into the pGL3 luciferase vector	169
Figure 5.7	E4bp4 can activate and repress transcription of target genes, depending on variations in the sequence of DNA binding sites	170
Figure 5.8	SUMOylation and phosphorylation affects transcriptional activation by E4bp4	171
Figure 5.9	MNK-1 cells have an NK cell phenotype	172
Figure 5.10	Cloning E4bp4 mutant cDNA into pCSGW lentiviral expression vector	173
Figure 5.11	MNK-1 cells were highly transduced and E4bp4 was stably expressed in these cells	175
Figure 5.12	SUMOylation and phosphorylation affect the activation of E4bp4 target genes in NK cells	176
Figure 6.1	Development of NK cells from transduced umbilical cord blood $CD34^+$ progenitor cells	189
Figure 6.2	Variations in the number of NK cells in human peripheral blood	191
Figure 6.3	Listeriolysin O affects the SUMOylation of E4bp4	192

List of Tables

Table 1.1	E4bp4 interacting partners from yeast-two-hybrid screen	52
Table 2.1	Primers used for sequencing	64
Table 2.2	Primers used for site-directed mutagenesis	65
Table 2.3	Primers used for cloning FLAG-E4bp4	66
Table 2.4	Antibodies used for Western blotting	68
Table 2.5	Fluorochrome conjugated antibodies	69
Table 2.6	TaqMan gene expression probes used for QPCR	78
Table 3.1	Identification of putative E4bp4 SUMOylation sites using bioinformatic prediction tools	90
Table 3.2	Mutant forms of E4bp4 created lacking putative SUMOylation sites	98
Table 3.3	Predicted masses of E4bp4 SUMO modified peptides	112
Table 4.1	Phosphoproteomic screens that identify E4bp4 as a target of phosphorylation	135
Table 4.2	Prediction of E4bp4 phosphorylation sites and modifying kinases	138
Table 4.3	Predicted masses of potential E4bp4 phosphopeptides	140
Table 4.4	E4bp4 phosphorylation-mutant cDNAs generated	144
Table 5.1	E4bp4 DNA binding sequences	168
Table 5.2	*In vivo* E4bp4 binding sites	179
Table 5.3	Predicted E4bp4 binding sites in the promoter regions of Notch1 and E2A (Tcf3)	181

List of Abbreviations

ALS	Amyotrophic lateral sclerosis
AML	Acute myeloid leukaemia
BM	Bone marrow
BMDM	Bone marrow derived macrophages
bZIP	Basic leucine zipper
cDC	Conventional dendritic cell
cDNA	Complementary DNA
ChIP	Chromatin immunoprecipitation
cNK	Conventional NK cell
CID	Collision-induced dissociation
CLP	Common lymphoid progenitor
CKIε	Casein kinase Iε
DMEM	Dulbecco’s modified eagle medium
DN	Dominant negative
dsDNA	Double stranded DNA
E4bp4	E4-binding protein 4
Eomes	Eomesodermin
FACS	Fluorescence activated cell sorting
FBS	Fetal bovine serum
Fw	Forward
GC	Glucocorticoid
GLε	Germ-line ε
GM-CSF	Granulocyte macrophage colony-stimulating factor
HDAC	Histone deacetylase
HEK	Human embryonic kidney
HRP	Horseradish peroxidase
HSC	Hematopoietic stem cell
HSF1	Heat shock factor 1
IMDM	Iscove's Modified Dulbecco's Medium
Id2	Inhibitor of DNA binding 2
IFN-γ	Interferon-gamma
IL	Interleukin
ILC	Innate lymphoid cell
IP	Immunoprecipitation

IRES	Internal ribosome entry site
IRF	Interferon regulatory factor
kb	Kilobase
kDa	Kilo dalton
KIR	Killer cell immunoglobulin-like receptor
Lin^-	Lineage negative
LLO	Listeriolysin O
LMPP	Lymphoid primed multipotent progenitor
LPS	Lipopolysaccharide
LTi	Lymphoid tissue inducer
Ltr	Long terminal repeats
MCMV	Murine cytomegalovirus
MeCN	Acetonitrile
MHC	Major histocompatibility complex
mNK	Mature NK cell
MS	Mass spectrometry
NEM	*N*-Ethylmaleimide
NK	Natural killer
O/N	Over night
ORF	Open reading frame
PBS	Phosphate buffered saline
PBST	Phosphate buffered saline + Tween20
PcG	Polycomb group bodies
PCR	Polymerase chain reaction
PE	Phycoerythrin
PML	Promyelocytic leukaemia protein
PTM	Post translational modification
QPCR	Quantitative Polymerase Chain Reaction
RLA	Relative luciferase activity
RT	Room temperature
Rv	Reverse
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SENP	Sentrin/SUMO-specific protease
SIM	SUMO-interacting motif
SLE	Systemic lupus erythematosus
SNP	Single nucleotide polymorphism

SUMO	Small ubiquitin-like modifier
TBS	Tris buffer saline
T/C	Tissue culture
TFA	Trifluoroactic acid
TNF	Tumour necrosis factor
trNK	Tissue resident NK cell
U	Unit of activity
UCB	Umbilical cord blood
WB	Western blot
WT	Wild-type

CHAPTER 1

Introduction

1.1 Natural Killer cells

Natural Killer (NK) cells are large granular lymphocytes that sit at the interface between innate and adaptive immunity. NK cells were first identified nearly forty years ago by their ability to kill certain target cells *in vitro* without prior sensitisation and were described as 'cytolytic effector lymphocytes' (Herberman et al., 1975) (Kiessling et al., 1975). NK cells were distinguished from B and T lymphocytes by their lack of B and T cell surface antigens and their ability to develop in mice and humans that were unable to rearrange their antigen receptor genes (Hackett et al., 1986). NK cells rely on a collection of germline encoded activating and inhibitory receptors that recognise target cells and determine whether NK cell effector functions are triggered, killing the target cell (Vivier et al., 2008).

NK cells are classically part of the innate immune system, where they represent one of the first lines of defence against infection and are important for tumour immunosurveillance, graft rejection and during pregnancy (Di Santo, 2008). NK cells produce an array of cytokines and chemokines, often in response to interactions with antigen-presenting cells, and these can influence and augment adaptive immune responses (Vivier et al., 2011). Some studies have shown that NK cells can also mount antigen-specific memory responses and long-lived 'memory' NK cells can arise following encounters with certain pathogens (Sun et al., 2011) (Sun et al., 2012). Therefore, NK cells have a unique biological niche covering both innate and adaptive halves of the immune system. NK cells have been extensively studied in mice, in particular how they function and develop and this has proven a good model for human NK biology.

1.1.1 NK cell functions

NK cells are characterised by a number of effector functions. Most notably, NK cells can spontaneously lyse susceptible target cells by the exocytosis of perforin and cytoplasmic granules containing numerous granzymes. Perforins create holes in the cell surface of target cells, whilst granzymes activate apoptosis pathways in the targeted cell (Topham and Hewitt, 2009). NK cells can also eliminate targets via Fas-L and Trail-mediated apoptosis pathways, which are stimulated following ligand binding, receptor oligomersiation and recruitment of adaptor proteins to death domain receptors (Warren and Smyth, 1999).

Kärre *et al*. performed pioneering early work into the activities of NK cells, demonstrating that NK cells are able to lyse cells that lack expression of major histocompatibility complex

class I (MHC-I) (Karre et al., 1986). This work led to the formation of the 'missing self' hypothesis, where the predominant targets for NK cell cytolytic activity are cells which down regulate MHC-I, such as virally infected cells and tumours (Karre et al., 1986). Inhibitory receptors (such as Ly49 receptors in mice and killer cell immunoglobulin-like receptors (KIRs) in humans) on the surface of NK cells recognise MHC-I ligands on healthy cells. Upon interaction with MHC ligands, inhibitory receptors recruit tyrosine phosphatases (SHP-1 and SHP-2) to their cytoplasmic domains and these are able to block activation signals, therefore preventing a target cell from being killed (Cooper et al., 2001) (Pegram et al., 2011). Different subsets of NK cells express different inhibitory receptors, including many non-MHC-binding inhibitory receptors, with most cells expressing more than one, as this allows the global pool of NK cells to carefully monitor self-MHC molecules and react when MHC expression is perturbed by disease (Cooper et al., 2001) (Pegram et al., 2011).

Inhibitory receptors do not work in isolation but alongside a large collection of activating receptors, including NKG2D, NKp46 and the CD94/NKG2C complex (Koch et al., 2013). Through a large number of activating receptor-coupled-adapters NK cells are able to achieve their diverse number of functional activities (Koch et al., 2013). Not all activating receptor ligands have been defined but some are known to be unregulated on stressed cells, e.g. MICA and MICB on tumour cells, which are recognised by NKG2D (Weiss-Steider et al., 2011). The activating receptor CD16 can trigger antibody-dependent cellular cytotoxicity by encountering target cells opsonised with IgG (Moretta et al., 2008). NK cell activity is a fine balance between activating and inhibitory signals and target cell killing is initiated when this balance is perturbed (Topham and Hewitt, 2009) (Vivier et al., 2008). NK cell triggering is not only caused by a lack of MHC-I, but can also be caused by cytokine stimulation or the presence of large numbers of stimulatory molecules (Colucci et al., 2002). NK cells can respond to missing, altered, stressed or non-self situations, and the 'missing self hypothesis' is now more commonly described as the dynamic equilibrium concept (Vivier et al., 2011).

An equally important function of NK cells is their ability to rapidly produce cytokines, most commonly interferon-gamma (IFN-γ) and tumour necrosis factor-alpha (TNF-α), but also in some instances granulocyte macrophage colony-stimulating factor (GM-CSF) and IL-10 (Perussia et al., 2005). These cytokines can reinforce T_H1 differentiation of T cells in secondary lymphoid organs and also promote the maturation of dendritic cells (Walzer et al., 2005). NK cells can also produce a number of chemokines, such as MIP-1α, MIP-1β and RANTES, which are used to promote inflammatory responses and attract leukocytes to sites of infection or injury (Perussia et al., 2005) (Vivier et al., 2008). Using these functions NK

cells can eliminate unwanted target cells, through various non-redundant pathways, and promote inflammatory immune responses.

NK cells are formally members of the innate immune response, but recent evidence suggests that, in certain situations, NK cells may mediate some long-lived memory-like responses, which is a central attribute of the adaptive immune system (Sun et al., 2009). In C57BL/6 mice infected with murine cytomegalovirus (MCMV), a population of NK cells expressing the activating receptor Ly49H were found to be expanded and these cells were long-lived, possibly surviving for up to 70 days (Sun et al., 2009). On secondary challenge with MCMV, these cells proliferated in an IL-12-dependent manner, were highly cytotoxic and produced large quantities of cytokines, which provided enhanced protection against secondary viral challenge (Sun et al., 2009) (Sun et al., 2012). A similar occurrence has been observed in human transplant patients, where a mature NKG2C^{+} memory-like NK cell population expands, is long-lived and has enhanced responsiveness to secondary CMV infection (Foley et al., 2012) (Lopez-Verges et al., 2011).

NK cells have essential activities in preventing tumour initiation, halting tumour development and controlling several types of viral infection, including MCMV and herpes virus in humans (Vivier et al., 2008). This is demonstrated by individuals who selectively lack NK cells and exhibit increased susceptibility to viral infections, especially herpes viruses and have an increased tumour incidence (Orange, 2006). For example, a large cohort study of Japanese people aged over 40 indicated a direct correlation between low NK cytotoxicity and elevated cancer incidence (Imai et al., 2000).

1.1.2 Innate lymphoid cells

NK cells are the founding members of the recently defined family of leukocytes that are collectively known as Innate Lymphoid Cells (ILCs) (Spits et al., 2013). ILCs are a functionally diverse collection of cells that have been grouped according to their functional characteristics, as well as the transcription factors (TFs) which drive their differentiation and cytokine production (Walker et al., 2013). ILCs are clearly differentiated from T and B lymphocytes by their inability to undergo somatic rearrangement of T cell receptors or immunoglobulin genes.

Three main subsets of ILCs have now been distinguished (ILC1, ILC2, ILC3) based on phenotypic and functional characteristics, mainly the expression of key cytokines and TFs (Spits et al., 2013). NK cells are the prototypical members of the ILC1 subset, which are defined by their production of type 1 cytokines such as IFN-γ and TNF-α, and their expression of the TFs Tbet and/or Eomes **(Figure 1.1)**. ILC2 cells are dependent on the TF Gata3 and express the type 2 cytokines IL-5 and IL-13 **(Figure 1.1)**, whilst ILC3 cells express RORγT and produce type 17 cytokines including IL-17 and IL-22 **(Figure 1.1)**. ILCs are commonly located at barrier surfaces and are important for maintaining tissue homeostasis and providing protection against invading pathogens (Diefenbach et al., 2014). As well as contributing to immune responses, lymphoid tissue inducer (LTi) cells, part of the ILC3 subset, are also essential for lymphorganogenesis (Walker et al., 2013). LTi cells are required during fetal life to initially develop lymphoid organs and are present in adult tissues where they promote lymphoid follicle formation (Diefenbach et al., 2014).

All lymphocytes of the adaptive immune system and all ILC subsets derive from common lymphoid progenitor (CLP) cells in the bone marrow (BM) **(Figure 1.1).** For many years it was postulated whether a lineage restricted progenitor existed, for some or all of the ILC subsets, which was downstream of the CLP. Recently, two independent studies have identified ILC progenitor populations that give rise to many of the ILC subsets. Constantinides *et al*. first discovered the committed ILC progenitor (CILP) that can give rise to all ILC subsets, except for NK cells and LTi cells, and was identified by its expression of the TF promyelocytic leukaemia zinc finger (PLZF) (Constantinides et al., 2014) **(Figure 1.1)**. Subsequently, Klose *et al*. identified the common helper-like ILC progenitor (CHILP), which can give rise to all ILC subsets, except NK cells and was distinguished by its expression of integrin $\alpha_4\beta_7$ and the TF Id2 (Klose et al., 2014) **(Figure 1.1)**. A population of CHILP cells, present in fetal liver and BM, were discovered to express PLZF (Klose et al., 2014) and due to their less restricted differentiation potential, current models suggests that the $\alpha_4\beta_7^+$ $Id2^+$ CHILP gives rise to the $PLZF^+$ CILP cell **(Figure 1.1)**.

There is also evidence for plasticity in the way ILC subsets develop. ILC3s normally display no cytotoxic activity and generally secrete IL-17 or IL-22 (Walker et al., 2013). However, it has been reported that some ILC3 cells can respond to IL-12 and IL-23, causing them to downregulate expression of RORγT and acquire the capacity to produce IFNγ (Vonarbourg et al., 2010). These cells were demonstrated to be analogous to ILC1s and contributed to the induction of colitis (Vonarbourg et al., 2010). The disappearance of RORγt expression in these cells was accompanied by a strong increase in T-bet expression and mice that are deficient for T-bet lack the ability to produce this cell type (Bernink et al., 2013) (Vonarbourg

et al., 2010). T-bet was found to positively regulate IFNγ production and negatively regulate IL-17 production by ILCs and together, these studies indicate that there is plasticity in group 1 and group 3 ILC development (Bernink et al., 2013). T-bet and RORγt are important regulators of the plasticity of ILC subsets, which is similar to the plasticity observed between T_H1 and T_H17 cells.

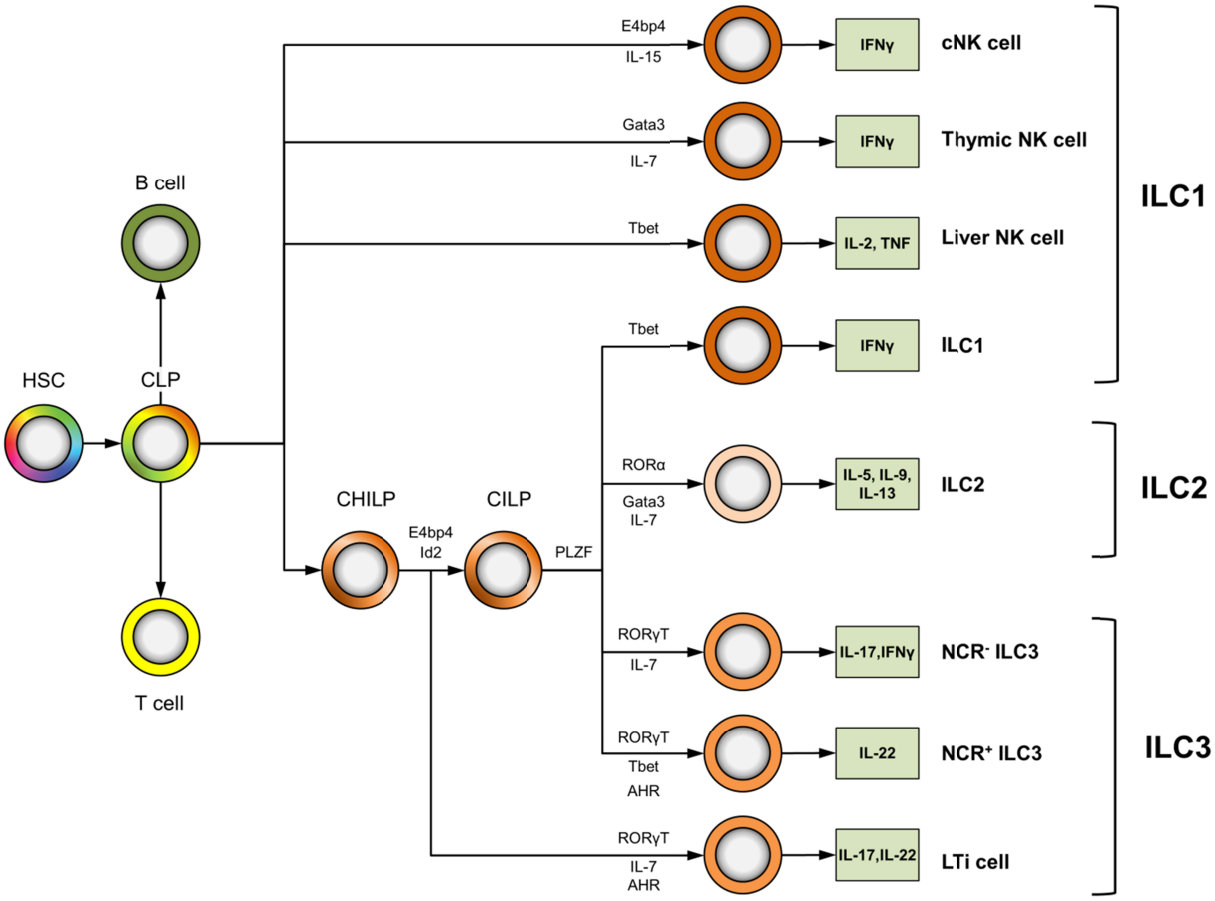


Figure 1.1 – Murine innate lymphoid cell subsets

Innate lymphoid cells (ILC) arise from hematopoietic stem cells (HSC), which develop into common lymphoid progenitors (CLPs) and these can also give rise to T and B cells. ILC1, ILC2 and certain ILC3 subsets also all arise from the common helper-like innate lymphoid cell progenitor (CHILP) and the committed innate lymphoid cell progenitor (CILP). NK cell subsets do not derive from the CHILP or CILP and lymphoid tissue inducer (LTi) cells do not develop from the CILP. ILC subsets are organised based on functional characteristics, the cytokines and transcription factors required for their development (indicated) and the cytokines they produce when fully mature (green boxes). cNK = conventional NK cell.

1.1.3 NK cells and immunotherapy

The importance of NK cells for immune responses against virally infected and cancerous cells has meant that there is an ever increasing interest in the development NK cell-based immunotherapies (Knorr et al., 2014) (Ljunggren and Malmberg, 2007). Numerous studies have shown that NK cells have the ability to kill mouse and human tumour cell lines *in vitro*, as well as freshly isolated human tumour cells (Hayakawa and Smyth, 2006b) (Wu and Lanier, 2003). Studies in mice have shown that NK cells are able to eliminate grafted mouse tumour cells *in vivo* and NK cells have been shown to be involved in the eradication of induced and spontaneously developing tumours (Hayakawa and Smyth, 2006b) (Wu and Lanier, 2003). Significantly, a number of studies have now demonstrated that NK cell immunotherapy can to be effective in controlling and/or eradicating some human haematological tumours (Curti et al., 2011) (Ruggeri et al., 2002). In particular, adoptively transferred NK cells have been effective against acute myeloid leukaemia (AML) when given to patients in combination with a hematopoietic stem cell transplant (Curti et al., 2011) (Ruggeri et al., 2002). The NK cells are able to attack and kill the tumour cells whilst not causing any graft-versus-host disease, as observed when using cytotoxic T cells in similar approaches (Geller and Miller, 2011).

Numerous advances have been made in understanding how adoptively transferred NK cells behave when transfused into a patient and why in some situations they effectively attack tumour cells and in others they do not (Geller and Miller, 2011). For example, NK cell killing can fail due to inhibitory KIRs recognising self MHC-1 on the surface of tumour cells (Miller et al., 2007). Patients and NK cell donors are now routinely screened for their KIR genotype or ligand status, so that patients receive KIR-ligand mismatched alloreactive NK cells that can more effectively target cancer cells (Miller et al., 2007) (Ruggeri et al., 2002). However, before NK cells can be used regularly in the clinic major challenges remain to increase the cytotoxic effects of adoptively transferred NK cells, making them consistently expand *in vivo* and making their activity more reliable (Scheper et al., 2014).

Producing the large numbers of highly cytotoxic NK cells required for every patient is difficult, especially as most recent clinical studies involved the use of adoptively transferred NK cells, isolated from adult donors by lymphapheresis and transfused into patients (Knorr et al., 2014). Although these cells have been educated in healthy hosts and have high anti-tumour activity, this strategy has the major drawback of requiring very large numbers of healthy volunteers to donate their NK cells. Strategies have been suggested for expanding these cells *in vitro* before transfusion, but to date no standard procedure has been adopted for

achieving this (Gong et al., 2010) (Rujkijyanont et al., 2013). Numerous alternative sources of NK cells are now being investigated including the generation of NK cells from hematopoietic stem cells (HSCs) derived from BM (Knorr et al., 2013) (Woll et al., 2005) or from umbilical cord blood (Luevano et al., 2012) (Spanholtz et al., 2010). Using these strategies potentially unlimited numbers of highly cytotoxic NK cells could be produced and with correct cryopreservation could be made as an 'off-the-shelf' product (Knorr et al., 2014) (Luevano et al., 2014). There is also the potential to genetically modify these NK cells allowing tailored therapeutic products to be produced expressing tumour-specific receptors (Themeli et al., 2013). However, if these or other strategies are to be successful, a greater understanding is required of how NK cells naturally develop and how this process is controlled and regulated.

1.2 NK cell development

Hematopoiesis is the process by which all mature blood cells develop from HSCs in the BM of adult mammals (Orkin and Zon, 2008). Mature blood cells are relatively short lived and so self-renewing HSCs are required to replenish multi-lineage progenitors and committed precursor cells that give rise to all hematopoietic lineages (Orkin and Zon, 2008). Common lymphoid progenitors (CLPs) develop from HSCs and have precursor potential to give rise to T, B and NK cells, as well as all other innate lymphoid cells (Papatriantafyllou, 2013) **(Figure 1.1)**.

The understanding of NK cell development has mainly derived from studies on human peripheral blood and mouse splenic NK cells and these are now commonly referred to as conventional NK (cNK) cells. These cNK cells develop in the BM through a number of well-defined stages to produce mature NK (mNK) cells that are released to the periphery to acquire effector functions (Yu et al., 2013a). However, numerous studies have now identified NK cell subsets in the liver, uterus, thymus and lymph nodes and these tissue-resident NK (trNK) cells have alternative developmental pathways and can perform different roles to cNK cells (Sojka et al., 2014) (Yokoyama et al., 2013).

1.2.1 Murine NK cell development

NK cell development in the mouse BM requires CLPs, which are lineage marker negative (Lin^-) and express c-Kit (CD117), Sca-1 (Ly6A/E), IL-7Rα (CD127), 2B4 (CD244), TNF receptor (CD27) and Flt3 (CD135) **(Figure 1.2)**. Like T and B cells, NK cells go through an ordered and linear maturation process **(Figure 1.2)**, but unlike T and B cells, NK cells do not require significant time to develop effector functions and, once produced, can rapidly be activated (Huntington et al., 2007). Two reports have defined the earliest NK cell committed precursor (NKP) residing within the Lin^- population. Fathman *et al.* used adoptive transfer experiments and *in vitro* assays to define two populations, pre-NKPs (Lin^- $CD135^-$ $CD27^+$ $CD244^+$ $CD127^+$ $CD122^-$) as the earliest committed NK cell stage, which differentiate to become refined NKPs (Lin^- $CD135^-$ $CD27^+$ $CD244^+$ $CD127^+$ $CD122^+$), also referred to as rNKPs (Fathman et al., 2011) **(Figure 1.2)**. In the second report, Carotta *et al.* used an $Id2^{gfp}$ reporter mouse to identify similar NK committed precursors, as the TF Id2 is essential for NK cell development (Carotta et al., 2011). They identified a pre-pro NK cell population as the earliest NK cell committed stage, defined as: Lin^- $CD122^-$ $Id2^+$ $CD127^+$ $CD135^-$ $Sca1^+$ $CD117^+$ $NKG2D^{low}$ (Carotta et al., 2011). The pre-pro NK cell population seems to largely overlap with the pre-NKP population identified by Fathman *et al.* (Fathman et al., 2011). However, it should be noted that immature NK (iNK) cells can develop in the absence of Id2 (Boos et al., 2007). Therefore the $Id2^+$ pre-pro NK cell population may not include every single committed NK cell precursor.

The acquisition of the IL-15 receptor β chain (IL-15Rβ; CD122) at the NKP stage is an important step in the development of NK cells, as IL-15 specifically promotes the differentiation, maturation and survival of NK cells (Huntington, 2014). Maturing NK cells acquire functional receptors in a systematic fashion with iNK cells acquiring NK1.1, CD94 and NKp46 **(Figure 1.2)**. This is followed by Ly49 receptors, DX5, CD43 and CD11b by fully mature NK cells **(Figure 1.2)**. Mature NK cells further differentiate from $CD11b^-$ $CD27^+$ NK cells, into double positive $CD11b^+$ $CD27^+$ and finally to the most mature $CD11b^+$ $CD27^-$ NK cells (Kim et al., 2002). Both the double positive and $CD11b^+$ $CD27^-$ NK cells have capacity to kill target cells and secrete IFN-γ, but the double positive population displays greater effector function and responsiveness to chemokines (Hayakawa and Smyth, 2006a). The three subsets also differ in tissue distribution, with the $CD11b^-$ $CD27^+$ NK cells primarily present in the bone marrow and lymph nodes, whilst the $CD11b^+$ $CD27^-$ population are more abundant in the spleen, lung, liver and blood (Kim et al., 2002).

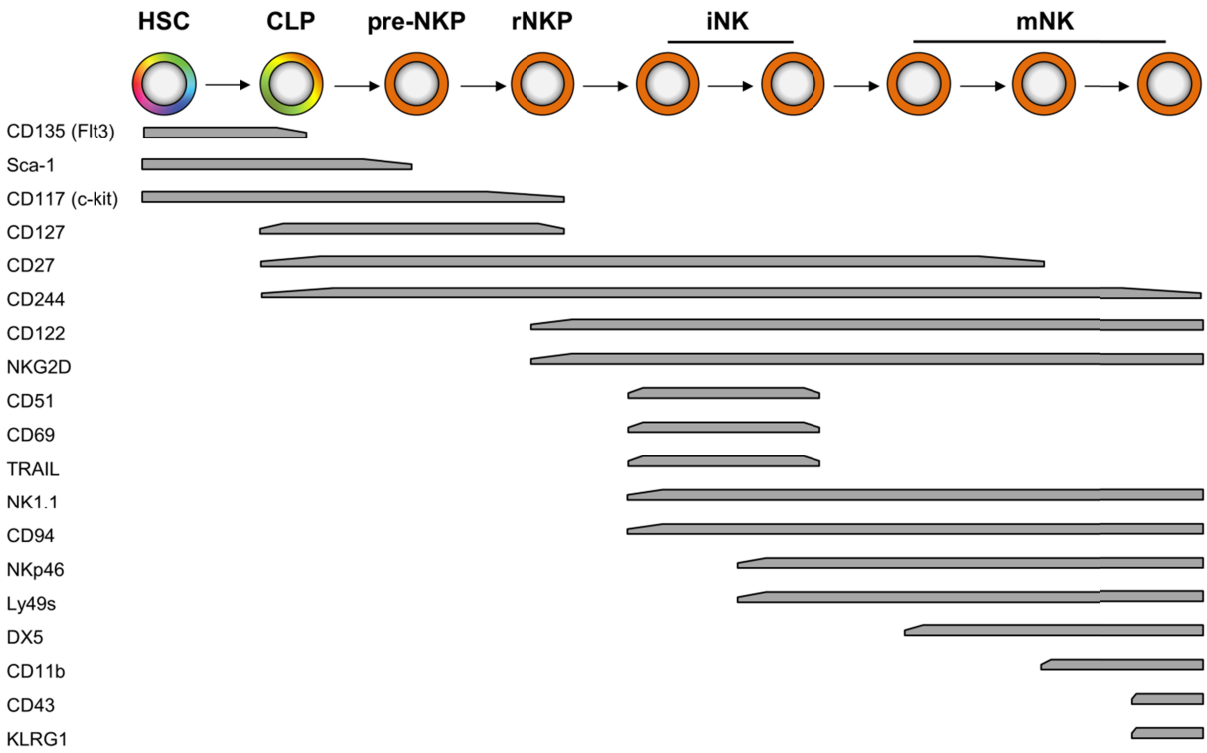


Figure 1.2 – Stages of conventional murine Natural Killer cell development

NK cells develop from hematopoietic stem cells (HSCs) and common lymphoid progenitors (CLPs) in murine bone marrow. Each stage of conventional NK cell development is determined by the expression of cell surface markers, indicated on the left. Pre-NKPs represent the earliest precursor cells committed to the NK lineage. Immature NK (iNK) cells are subdivided based on the expression of NKp46 and Ly49s. Any given Ly49 is expressed at a characteristic frequency in mature (mNK) cells, depending on genetic background. mNK cells are subdivided based on the expression of CD27, CD11b and CD43. Adapted from (Vosshenrich and Di Santo, 2013).

Alternative murine NK cells subsets are also found in other organs. These trNK cells are found in the lining of the uterus and are the most abundant lymphocyte at the site of embryo implantation, where they regulate trophoblast invasion and vascular remodelling (Manaster and Mandelboim, 2008). The thymus also contains a phenotypically unique NK cell subset, which has been described as $CD127^+$ $CD69^+$ $Ly49^-$ $CD11b^-$, in contrast to cNK cells which are generally $CD127^-$ $CD69^-$ $Ly49^+$ $CD11b^+$ (Vosshenrich et al., 2006). Interestingly these thymic NK cells don't develop from NK cell progenitors, but rather from double negative thymocyte progenitors, which normally develop into T cells (Vargas et al., 2011). These cells are capable of producing more cytokines than cNK cells, but are less efficient at killing target cells (Vosshenrich et al., 2006).

For many years the liver was known to contain large numbers of NK cells, but more recently it has been discovered that many of these NK cells are not like cNK cells (Crispe, 2009). Up to 50% of the NK cells in the liver have an immature TRAIL^{+} DX5^{-} phenotype; whereas NK cells in most other peripheral organs are TRAIL^{-} DX5^{+} and are phenotypically mature. Peng *et al*. recently identified that these immature NK cells are a liver-specific trNK cell population that exclusively express CD49a, reside in liver sinusoids and are important for antigen-specific memory-like responses (Peng et al., 2013).

1.2.2 Human NK cell development

It has been suggested that human NK cell development goes through a similar set of developmental stages to murine NK cell development, but this is currently less well defined **(Figure 1.3)**. Human CLPs are defined as CD34^{+} CD38^{+} CD45RA^{+} CD10^{+} and are generally found in the BM, but similar CLP-like populations have also been identified in extramedullary tissues, e.g. the tonsils and lymph nodes (McClory et al., 2012). In these tissues the earliest NK progenitors have been identified (stage 1 and stage 2), but these cells also appear to retain some potential for T cell and DC development, under certain conditions (Freud et al., 2006) **(Figure 1.3)**. Therefore these cells may not represent true NK progenitors that are committed specially to the NK lineage. Similarly to murine NK cell development, IL-15 is key driver of differentiation, and IL-15 responsive CD45RA^{+} CD117^{+} CD161^{+} CD94^{-} cells (stage 2 cells) are found in secondary lymphoid tissues and these give rise to CD56^{+} CD94^{+} CD122^{+} progeny; stages 3 and 4 (Freud et al., 2006) **(Figure 1.3)**.

Peripheral human NK cells are split into CD56^{bright} and CD56^{dim} populations, which differ in their functional activity and localisation (Stages 4 and 5) **(Figure 1.3)**. CD56^{dim} NK cells are fully mature, highly cytotoxic, produce IFN-γ and perforin and constitute 90% of the NK cells in peripheral blood and spleen (Cooper et al., 2001) (Robertson and Ritz, 1990). In contrast, CD56^{bright} NK cells are more immature and are considered as cytokine producers and commonly found in lymph nodes and tonsils (Cooper et al., 2001). Human NK cells have also been discovered in the uterus, liver, intestine and thymus, often with varying phenotypes, suggesting different anatomical locations can support specialised NK cell development (Yu et al., 2013a).

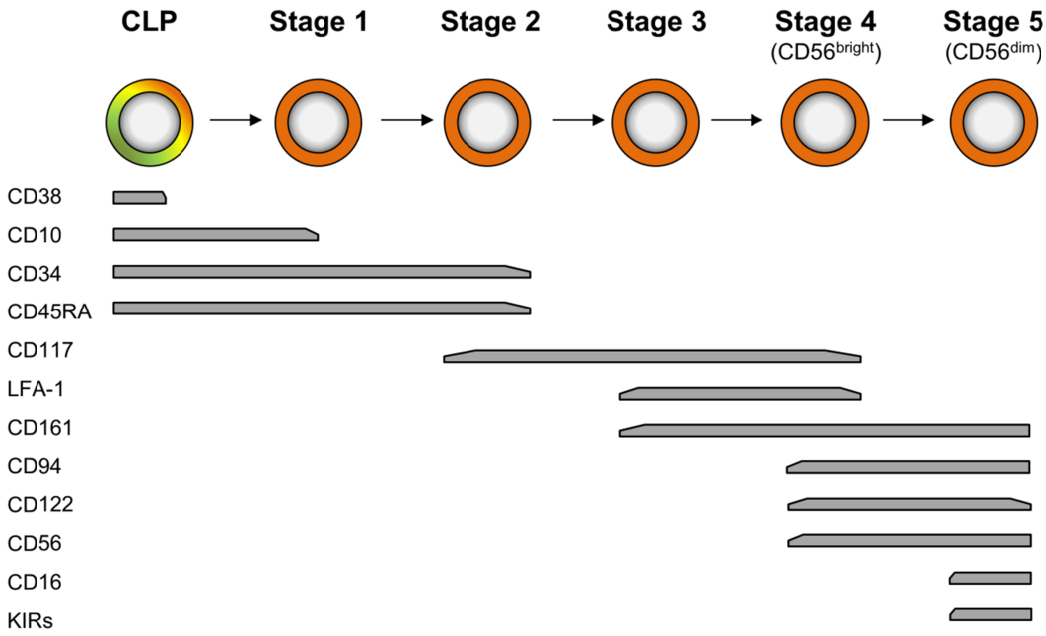


Figure 1.3 – Stages of human Natural Killer cell development

Human NK cells develop from common lymphoid progenitors to mature stage 5 $CD56^{dim}$ NK cells is a linear process. Each developmental stage is distinguished by expression of specific cell surface markers, indicated on the left. Mature stage 5 NK cells can express any given KIR, depending on genetic background, with each KIR expressed at characteristic frequency in the whole NK cell population.

1.3 Regulation of NK cell development

Lineage restriction of hematopoietic progenitor cells involves major proteomic and transcriptomic changes, so that alternative lineage-potentials are restricted and desired cell fates are produced (Vosshenrich and Di Santo, 2013). Numerous signals can influence these changes, including the microenvironment of the cell and the activation status (resting or infectious/inflammatory) (Vosshenrich and Di Santo, 2013). The stochastic expression of various TFs drives the lineage-restricted differentiation of different cell types, so that all types are produced correctly. Commitment to the NK lineage and development of NK cells in the BM is dependent upon the expression of NK specific genes, whilst T and B cell genes must be repressed (Huntington et al., 2013).

The regulation of murine T and B cell development has been well understood for a number of years, and sets of genetic regulators have been identified that consign progeny to each particular lineage (Rothenberg, 2014). In T cells, Notch1, E2A and TCF7 work together to activate T cell genes and repress non-T cell genes; whilst in B cells Pax5 and Ebf1 performs a similar role (Warren and Rothenberg, 2003). In murine NK cell development the picture has been less clear and only in recent years have a number of key transcriptional regulators been identified **(Figure 1.4)**. Epigenetic regulation is an additional layer of complexity that can influence TF access to binding sites (Cichocki et al., 2013). The loci of important regulators such as Ebf1 and Pax5 have been found to carry epigenetic marks in HSCs and, following lineage commitment, these marks can either change to become activating or repressing marks depending on which lineage the cell commits to (Heinz et al., 2010).

The development of NK cells is highly dependent on the cytokine IL-15. Developing NKPs express IL-15Rβ, which is required for IL-15 signal transduction and development (Rosmaraki et al., 2001). IL-15Rβ dimerises with the common gamma chain (IL-2Rγ, CD132) allowing IL-15 binding and signalling via Jak1/Jak3 and STAT5, driving NK cell differentiation (Huntington, 2014). The third component of the IL-15 receptor, IL-15Rα, can either be expressed intrinsically by developing NK cells or more commonly by trans-presentation from other hematopoietic or parenchymal cells (Schluns et al., 2004). The mechanism of trans-presentation allows IL-15 to be chaperoned to the surface of developing NK cells by IL-15 producing cells.

Several key TFs have now been identified as essential for the development of NK cells and other ILC subsets **(Figure 1.4)**. They control the transition from each developmental stage and promote the expression of specific NK cell markers **(Figure 1.4)**. How these factors work together to implement an NK cell-specific transcriptional programme is still to be fully elucidated, especially as many appear to cross-regulate each other.

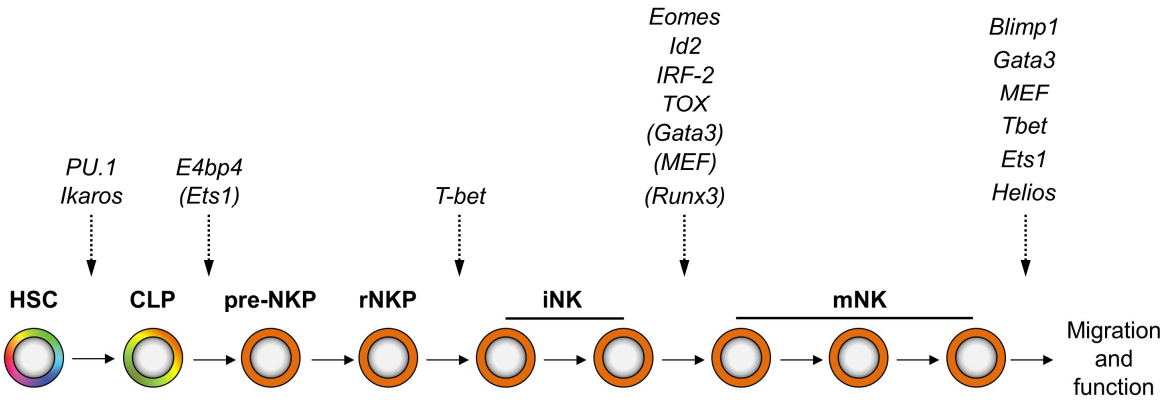


Figure 1.4 – Transcription factors that regulate conventional NK cell development

Numerous transcription factors are essential for NK cell development, but few have been discovered to regulate the earliest steps that commit common lymphoid precursors (CLPs) to the NK lineage. Each transcription factor is shown at the earliest transition they are known to mediate. Those shown in brackets display a partial phenotype at the indicated transition.

1.3.1 E4bp4

Of all the TFs identified to date, the only one that is clearly essential for the commitment of NK cells to the NK lineage is the basic leucine zipper (bZIP) transcription factor E4-binding protein 4 (E4bp4) (Male et al., 2014). After the development of the *E4bp4*$^{-/-}$ mice it was reported that E4bp4 was required for the development of iNK cells from NKPs (Gascoyne et al., 2009) (Kamizono et al., 2009). However, this work was performed using the definition of NKPs described by Rosmaraki *et al*. where only 1 in 12 of these NKP cells had NK cell potential (Rosmaraki et al., 2001). Using the refined NKP (rNKP) phenotype, where 1 in 2 cells has NK cell potential, E4bp4 was observed to be indispensable for the production of the earliest NK cell progenitors as *E4bp4*$^{-/-}$ BM contained almost no rNKP or pre-NKP cells (Male et al., 2014). E4bp4 was also shown to be expressed in CLP cells and was absolutely required for these cells to develop into NK cells *in vitro* (Male et al., 2014). When expressed in NK cells E4bp4 regulates the expression of other important NK-specific TFs, in particular Id2 and Eomes, which were shown to be direct transcriptional targets of E4bp4 and are expressed at later stages of NK cell development (Male et al., 2014) (Nandakumar et al., 2013).

E4bp4 is particularly required for the development of cNK cells in the BM, as certain populations of trNK cells are not dependent on E4bp4 and are found in the liver, thymus and spleen of *E4bp4*$^{-/-}$ mice (Crotta et al., 2014) (Seillet et al., 2014a). These trNK cells in the

thymus have an immature phenotype ($CD3^-$ $NK1.1^+$ $NKp46^+$ $CD122^+$), express low levels of Eomes and have lower levels of fitness compared to cNKs (Crotta et al., 2014). In an independent study, Seillet *et al.* identified $CD49b^-$ $TRAIL^+$ trNK cells in the liver of *E4bp4*$^{-/-}$ mice, and these cells also do not express Eomes (Seillet et al., 2014a). Development of these E4bp4 and Eomes-independent NK cells represents a distinct pathway of NK cell development that is specific to the liver and thymus.

E4bp4 is specifically required for the commitment to the NK cell lineage and the initial development of NK cells, but it is not needed for the maintenance of peripheral NK cells or the maturation of mNK cells in the bone marrow (Firth et al., 2013). Firth *et al*. used a tamoxifen inducible Cre/loxP-mediated deletion of *E4bp4* to also show that E4bp4 is also not involved in NK cell memory responses, following MCMV infection (Firth et al., 2013).

Additional analysis of the *E4bp4*$^{-/-}$ mouse has recently revealed that E4bp4 is essential for the development of all ILC subsets, as well as cNK cells (Fuchs et al., 2013) (Geiger et al., 2014) (Seillet et al., 2014b). E4bp4 is highly expressed in all ILC subsets and is required for correct formation of Peyer's patches (Seillet et al., 2014b). In the absence of *E4bp4,* the numbers of ILCs are greatly reduced in the small intestine (all subsets), colon (all subsets), lung (ILC1 and ILC2) and fat tissue (ILC2) (Fuchs et al., 2013) (Geiger et al., 2014) (Seillet et al., 2014b). This phenotype made *E4bp4*$^{-/-}$ mice highly susceptible to intestinal bacterial infection with *Citrobacter rodentium* and *Clostridium difficile* (Geiger et al., 2014) (Seillet et al., 2014b). Finally, E4bp4 was found to be required for the development of the earliest ILC progenitor (CHILP), defined as Lin^- $CD45^+$ $CD27^+$ $CD127^+$ $c\text{-}Kit^+$ $CD135^-$ $\alpha4\beta7^+$ confirming its central role in controlling the development of all innate lymphocytes (Geiger et al., 2014). However, it should be noted that mice lacking *E4bp4* still have lymph nodes, which indicates that some form ILC3 development is not completely blocked in its absence (Gascoyne et al., 2009) (Kamizono et al., 2009). Further studies are required to define exactly how E4bp4 controls ILC fate on a molecular level.

1.3.2 Id2

The transcriptional repressor Inhibitor of DNA binding 2 (Id2) is important for the development of NK cells and other ILC subsets (Verykokakis et al., 2014). In the absence of Id2, very few NK cells develop and the residual cells that do remain display numerous maturation defects (Boos et al., 2007) (Ikawa et al., 2001). In the BM and periphery of *Id2*$^{-/-}$ mice, major decreases in the number of mNK cells are observed, yet numbers of NKPs and

iNK cells in the BM are not affected (Boos et al., 2007). Id2 therefore appears to be required for the later stages of NK cell development. However, some evidence suggests that Id2 may be important earlier in the developmental process. Using an $Id2^{gfp}$ reporter line Carrota *et al*. showed the earliest NK cell precursors, pre-pro NK cells, express high levels of Id2 (Carotta et al., 2011), but whether these cells can develop in the absence of Id2 has not been confirmed. It is possible that Id2 is important for the earlier stages of NK cell development but its role may not be essential until later in the developmental pathway.

Id2 functions by counteracting the activity of E-box binding proteins (e.g. E2A, E2-2 and HEB) that are required for T and B cell development (Di Santo, 2014). As Id2 does not have a DNA binding domain, it binds directly heterodimerically to E-box proteins, preventing them from activating target genes and promoting T or B cell transcriptional programs (Ramirez and Kee, 2010). The expression of Ets1, another TF involved in NK cell development, is reduced in the absence of Id2; suggesting Id2 may function in multiple ways to promote the development of NK cells (Boos et al., 2007). Expression of Id2 in human thymic progenitor cells *in vitro* has also been found to promote NK cell production, providing further evidence of a central role for Id2 in NK cell development (Schotte et al., 2010).

Id2 is also highly expressed in all other ILC populations (Seillet et al., 2014b) and is required for the production of certain ILC subsets, including LTi cells (Boos et al., 2007) and IL-22-producing $NKp46^{+}$ cells, present in the gut (Satoh-Takayama et al., 2010). This observation has led some to suggest that Id2 may be essential for the development of all ILC subsets and expressed by a potential ILC progenitor (Spits et al., 2013). Constantinides *et al*. and Klose *et al*. recently identified novel ILC precursor populations, both of which expressed Id2 and could give rise to multiple ILC populations (Constantinides et al., 2014) (Klose et al., 2014). However, neither of these precursor populations was able to develop into cNK cells and it is not clear if Id2 is an essential requirement for ILC lineage commitment or if alternative developmental pathways exist.

1.3.3 Ets1

Ets1 is expressed in many hematopoietic lineages and Ets1-deficent mice have defects in T, B and NK cells. *Ets1*$^{-/-}$ mice have reduced pre-NKP, rNKP and mature NK cells demonstrating that Ets1 is a regulator of NK cell development (Ramirez et al., 2012). The loss of NK cells in *Ets1*$^{-/-}$ mice is, however, less severe than that observed in *E4bp4*$^{-/-}$ or *Id2*$^{-/-}$ mice. Despite this, Ets1 is expressed early in NK cell development and chromatin

immunoprecipitation (ChIP) experiments suggest that Ets1 helps to promote NK cell development by regulating the expression of *Id2*, *Tbet* and *IL2rβ* (CD122) (Ramirez et al., 2012). As mNKs are still produced in *Ets1*$^{-/-}$ mice, the activity of Ets1 is not essential and there may be some overlap of function with other TFs of the Ets family that are involved in NK cell development and maturation, such as PU.1 and MEF.

The NK cells that develop independently of Ets1 are able to produce IFNγ, but are unable to degranulate correctly in response to activation of NK1.1 (Ramirez et al., 2012). They also have higher expression of the activation marker CD69, higher expression of granzyme B and are hyper-proliferative in response to IL-15 stimulation (Ramirez et al., 2012). E4bp4 and Helios expression were upregulated in the absence Ets1, providing evidence that Ets1 may be required to limit NK cell responses following cytokine stimulation (Ramirez et al., 2012). Whether Ets1 directly or indirectly regulates E4bp4 and/or Helios expression remains to be explored.

1.3.4 Tox

Tox is a high mobility group box TF which is involved in regulating chromatin accessibility for other TFs (Aliahmad et al., 2012). It is an important modulator of CD4 T cell development and also NKT cell development (Aliahmad et al., 2012). In the absence of Tox, mNKs are not present in the spleen or BM, numbers of iNK cells are reduced by 50%, but NKP cell numbers are unchanged (Aliahmad et al., 2010). Allied to this, Tox expression increases in iNK and mNKs, compared to that seen in NKPs (Aliahmad et al., 2010). However, the analysis of pre-pro NKPs and rNKPs has not been reported in *Tox*$^{-/-}$ mice and so the mechanism by which Tox promotes NK cell development remains unclear. As well as regulating NK cell development, Tox is also required for the development of LTi cells and the formation of lymph nodes and Peyer's patches (Aliahmad et al., 2010).

1.3.5 Tbet and Eomes

The T-box family TFs Tbet and Eomesodermin (Eomes) work together to regulate sequential checkpoints of NK cell maturation (Gordon et al., 2012). Eomes deficient mice lack all DX5^{+} mNK cells, whilst Tbet deficient mice lack TRAIL^{+} DX5^{-} iNK cells, present mainly in the liver, as well as all mNKs (Gordon et al., 2012). The combined deletion of *Tbet* and *Eomes* lead to the loss of all NK cells beyond the NKP stage (Gordon et al., 2012). Tbet was therefore thought to be required for the generation of iNK cells and Eomes required for mNK

development. However, further analysis using an Eomes-GFP reporter line has now ascertained that in the BM Eomes is expressed in TRAIL$^-$ DX5$^+$ NK cells and Tbet expression is repressed in these cells (Daussy et al., 2014). When these Eomes$^+$ cells leave the BM for the periphery they can then express Tbet, which promotes the final stages of development and maturation (Daussy et al., 2014). TRAIL$^+$ DX5$^-$ Eomes$^-$ iNK cells in the liver express high levels of Tbet and develop into a distinct Eomes$^-$ NK cell lineage that resemble NKT cells in terms of cytokine production (Daussy et al., 2014). The development of this alternative lineage is blocked in the BM due to the repression of Tbet expression, allowing the development of Eomes$^+$ cNK cells. Whilst the liver permits Tbet expression, which subsequently represses Eomes expression and promotes development of TRAIL$^+$ DX5$^-$ Eomes$^-$ trNK cells.

1.3.6 Runx3

Runx family proteins have numerous roles in hematopoietic lineages, with Runx3 being dominantly expressed in NK cells and CD8$^+$ T cells (Ohno et al., 2008). Runx3 expression begins at the NKP stage and is maintained throughout NK cell development (Ohno et al., 2008). During NK cell development Runx3 is important for regulating the expression of CD122 (Ohno et al., 2008) and NKp46 (Lai and Mager, 2012). Knockout mice of all Runx family genes are embryonic lethal, but the function of Runx family members can be determined by expressing a dominant negative (DN) form of Runx1 in haematopoietic progenitor cells, which interferes with the function of all Runx proteins (Ohno et al., 2008). The expression of DN Runx1 causes lower numbers of mNKs to be produced and fewer of these cells express Ly49 receptors (Ohno et al., 2008). Expression of DN Runx1 was actually observed to promote IFN-γ production in response to cytokine stimulation, but these cells were not found to be more cytotoxic than WT cells (Ohno et al., 2008). Using this approach it is difficult to be certain that other Runx family members are not responsible for these phenotypes, but Runx3 is highly expressed in NK cells, suggesting it is the most likely candidate. Recent evidence acquired using genome-wide ChIP-seq experiments suggests that Runx3 may act alongside Ets family TFs to regulate gene expression in NK cells, promoting cytotoxicity, migration and cytokine production (Lotem et al., 2013).

1.3.7 PU.1

PU.1 is expressed in HSCs and is important for the development and function of numerous hematopoietic lineages (Carotta et al., 2010a). Previously, it was thought that PU.1 was required for NK cell development as it is strongly expressed in IL-2 expanded NK cells (Colucci, 2001). However, studies involving the conditional deletion of PU.1 in adult HSCs have confirmed that PU.1 is a key regulator of CLP development as in its absence all lymphoid cells and macrophages fail to develop (Carotta et al., 2010a). Analysis of PU.1^{GFP} reporter mice has shown that PU.1 is expressed in hematopoietic progenitor cells but not in pre-pro or mNK cells (Nutt et al., 2005) (Carotta et al., 2010b). PU.1 is proposed to regulate the expression of cytokine receptors Flt3 and IL7Rα, both essential for T, B and NK cell development (Carotta et al., 2010b)

1.3.8 Gata3

Gata3 is a zinc finger transcription factor with a well-defined role in T cell development (Hosoya et al., 2010). Gata3 is expressed in developing NK cells, but loss of Gata3 specifically affects the production of CD127^{+} thymic NK cells; it does not affect cNK cell development in the BM and spleen (Samson et al., 2003) (Vosshenrich et al., 2006). *Gata3*$^{-/-}$ NK cells have an immature phenotype, express lower levels of CD11b and CD43, and have a decreased ability to produce IFN-γ (Samson et al 2003). In the absence of Gata3, NK cell homing to the liver was also affected and all hepatic NK cell subsets were reduced (Samson et al., 2003). How Gata3 functions in trNK cells is still not entirely clear but it is suggested that it may control the expression of both Tbet and Eomes (Vosshenrich and Di Santo, 2013). As well as trNK cells, Gata3 is required for the differentiation and function of all other IL-7Rα-expressing ILC subsets (Samson et al., 2003) (Serafini et al., 2014) (Vosshenrich et al., 2006) (Yagi et al., 2014). Gata3 has a number of specific roles in different ILC subsets, for example in ILC2s it was observed to control the expression of the retinoic acid receptor RORα (Hoyler et al., 2012) (Mjosberg et al., 2012). As Gata3 is essential for all ILC subsets except for cNK cells, this provides further evidence that these “killer” ILCs are in fact an entirely separate lineage.

1.3.9 Ikaros

Ikaros is part of the small family of Krüppel-type zinc finger TFs and is expressed in many hematopoietic progenitor cells, including HSCs (Huntington et al., 2013). Studies on *Ikaros*$^{-/-}$ mice have demonstrated that Ikaros has an essential role in lymphoid development as these

mice that lack all T, B and NK cells (Georgopoulos et al., 1997) (Yoshida et al., 2006). Ikaros is expressed in developing NK cells and Ikaros-deficient progenitors cannot develop into NK cells *in vitro*, suggesting Ikaros is a critical regulator of NK cell development (Georgopoulos et al., 1997) (Wang et al., 1996). However, as *Ikaros*$^{-/-}$ mice also lack T and B cells, it is possible that Ikaros is needed for the development of CLPs and is not specifically required for the development of NK cells, although this is yet to be determined.

1.3.10 MEF

Myeloid ELF1-like factor (MEF) is a member of the Ets TF family and loss of MEF affects the development of NK and NKT cells (Lacorazza et al., 2002). *Mef* $^{-/-}$ mice have 60% less mNK cells in their spleen compare to WT littermates and the NK cells they do have are significantly less cytotoxic (Lacorazza et al., 2002). Lack of cytotoxicity in *Mef* $^{-/-}$ NK cells is due to an inability to produce perforin and IFN-γ, as MEF directly regulates the expression of the perforin gene (Lacorazza et al., 2002).

1.3.11 IRF-2

Mice lacking the TF interferon regulatory factor (IRF)-2 have a deficiency in the development of both T helper type 1 cells and NK cells (Lohoff et al., 2000). *Irf2*$^{-/-}$ mice have a selective loss of mNK cells because they have relatively normal levels of iNK cells in their spleen and liver (Lohoff et al., 2000) (Taki et al., 2005). In the BM of *Irf2*$^{-/-}$ mice, NK cells were able to develop to a relatively mature CD11b^{-} DX5^{+} phenotype, but these cells do not contribute to the peripheral NK cell pool as they appear to go through premature apoptosis (Taki et al., 2005). IRF-2 is therefore critical at a late checkpoint in NK cell maturation.

1.3.12 Blimp1

Blimp1 is a transcriptional repressor required for the differentiation of plasma cells and CD8^{+} effector T cells. Blimp1 is expressed by NK cells throughout their development and IL-15 induces its expression in CD122^{+} NKPs (Kallies et al., 2011). Blimp1 is needed for the maturation and homeostasis of NK cells in the periphery, as in its absence mNK are still produced, but at a lower frequency to WT (Kallies et al., 2011). The maturation defect observed in Blimp1-deficent NK cells is similar to the phenotype seen in mice lacking T-bet and Kallies *et al*. demonstrated that the expression of Blimp1 is at least partly regulated by T-bet (Kallies et al., 2011).

1.3.13 Helios

The zinc finger TF Helios is not required for the development of NK cells, but has an important role in NK cell activation. Noé mice have a mutation in Ncr1 (NKp46), preventing its surface expression. NK cells in these mice produce high levels of IFNγ, degranulate more effectively than WT NK cells and are hyper-responsive during MCMV infection (Narni-Mancinelli et al., 2012). Helios is expressed at higher levels in Noé mNK cells than in WT cells and short-hairpin RNA knockdown of Helios in Noé mNK cells reversed their hyper responsive phenotype (Narni-Mancinelli et al., 2012). This suggested that Helios down-regulation is involved in the control of NK cell reactivity via NKp46 (Narni-Mancinelli et al., 2012). Helios is also overexpressed in Ets1 deficient mNK cells, suggesting Ets1 may regulate NK cell activation, in part by acting through Helios (Ramirez et al., 2012).

1.3.14 The transcriptional network controlling NK cell development

The precise points at which each TF influences NK cell development has become clearer, yet how they interact with each other and work together to regulate the production of NK cells remains to be fully established. Combining the data from the studies discussed above, a picture can be built up of the network of TFs controlling NK cell development **(Figure 1.5)**. Most of the interactions in this network have been observed through changes in expression of a downstream target, in the absence of the TF of interest or by the use of ChIP to confirm direct binding to promoter regions **(Figure 1.5)**. There is cross-regulation between the central TFs in this network, namely E4bp4, Ets1 and Id2, which provides precise control of the initiation and progression of NK cell development **(Figure 1.5)**. Understanding how other TFs such as Gata3, Runx3 and TOX, fit into this regulatory network requires further investigation.

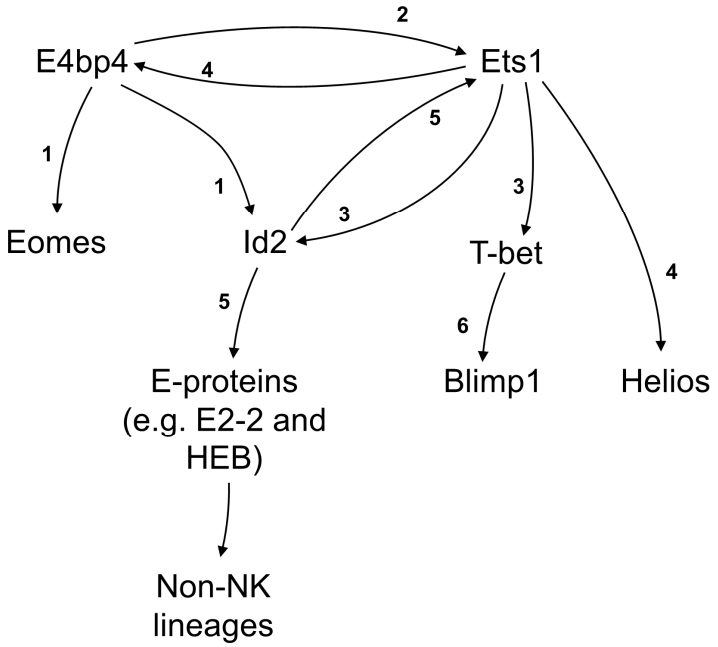


Figure 1.5 – Transcriptional network regulating NK cell development

Summary of the known interactions between transcription factors during NK cell development. 1) Id2 and Eomes are direct transcriptional targets of E4bp4, are under expressed in its absence and can rescue NK cell development in its absence (Male et al., 2014). 2) Ets1 can rescue NK cell development in the absence of E4bp4 (Male et al., 2014). 3) T-bet and Id2 are direct transcriptional targets of Ets1 and are under expressed in its absence (Ramirez et al., 2012). 4) Helios and E4bp4 are over expressed in the absence of Ets1 (Ramirez et al., 2012). 5) Id2 binds to E-proteins to prevent commitment to the T or B cell lineages and Ets1 is also under expressed in its absence (Boos et al., 2007). 6) Blimp1 expression is reduced in the absence of T-bet. Adapted from (Male and Brady, 2014).

1.4 E4bp4/NFIL3

Whilst many TFs have been identified as important for the development of NK cells, a complete picture of the transcriptional regulation of this process remains unresolved. How the activities of individual TFs are controlled is not well understood, in particular, little is known about the molecular mechanisms that control E4bp4 activity during NK cell development.

E4bp4, also referred to as nuclear factor regulated by IL-3 (NFIL3), was first identified as a transcription repressor which bound to an activating transcription factor DNA consensus sequence site in the adenovirus E4 promoter (Cowell et al., 1992). It was independently discovered to be a transcriptional activator of the IL-3 promoter in human T cells (Zhang et al., 1995). Sequence homology studies identified E4bp4 as a bZIP transcription factor, with a

DNA binding domain at its amino-terminus, adjacent to a heptad repeat of leucine residues that can mediate dimerisation **(Figure 1.6)**. As well as being able to homodimerise, it was found that E4bp4 can heterodimerise with another bZIP transcription factor CREB (Acharya et al., 2006). The C-terminus of E4bp4 contains a 65 amino acid domain that a mutational study revealed was essential for its ability to repress transcriptional activity (Cowell and Hurst, 1994) **(Figure 1.6)**. This domain was later shown to interact with the global transcriptional repressor protein Dr1, which suggested a possible mechanism by which E4bp4 could promote transcriptional repression (Cowell and Hurst, 1996). However, it is not known if these interactions with CREB and Dr1 are crucial for its activities *in vivo*.

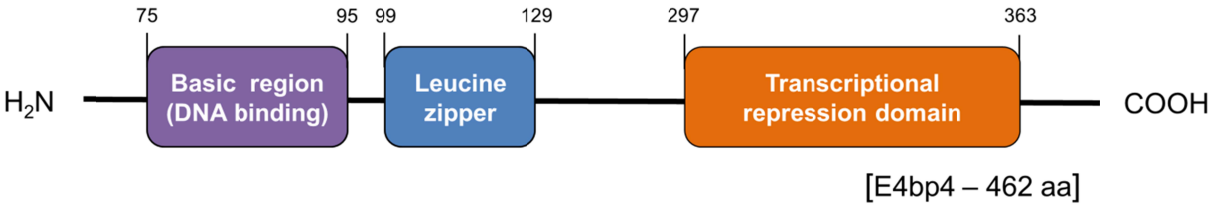


Figure 1.6 – Domain structure of E4bp4

Schematic representation of the conserved domain structure of E4bp4 (not to scale), which has been identified through sequence homology and mutational studies. The basic region has been identified to be essential for E4bp4 to bind to its DNA target sequence, whilst the leucine zipper domain will allow E4bp4 to interact with other transcription factors or potentially homo-dimerise. The transcriptional repression domain is required for E4bp4's transcriptional activity. Amino acid position is indicated.

As well as early studies using *in vitro* data and transcriptional reporter assays, a number of studies have now revealed that E4bp4 can both activate the expression of target genes, as well as repress the transcription of others. Using a combination of ChIP and target gene expression analysis, E4bp4 has been shown to repress the expression of numerous target genes *in vivo*, including IL-12 in macrophages and IL-13 in Type 2 T helper (T_H2) cells (Kashiwada et al., 2011a) (Kobayashi et al., 2011). Conversely, a number of studies have demonstrated that E4bp4 can promote the expression of target genes (Kashiwada et al., 2010) (Male et al., 2014) (Motomura et al., 2011) (Tong et al., 2013) (Zhang et al., 1995). Zhang *et al*. originally identified E4bp4 and found that it promoted the expression of IL-3 in T cells (Zhang et al., 1995), whilst in B cells E4bp4 activates the transcription of Ig germ-line ε,

required for IgE class switching (Kashiwada et al., 2010). Most significantly in NK cells, E4bp4 binds to the Eomes and Id2 loci and promotes their expression (Male et al., 2014).

It remains unclear how in different contexts E4bp4 can have opposing effects on target gene expression. Other TFs involved with regulating immune cell development are known to be multi-functional with the ability to both activate and repress target genes. Ikaros is a transcriptional activator that promotes T cell development and recruits chromatin remodelling complexes to lineage-specific genes, such as CD8 (Koipally and Georgopoulos, 2002). Ikaros is also able to promote transcriptional repression by associating with histone deacetylase-containing complexes (Koipally et al., 1999) or by interacting with the transcriptional co-repressor CtBP (Koipally and Georgopoulos, 2000). Gata3 is essential for a number of stages in T cell development and is required to activate a number of T-cell specific genes, including T cell receptor genes in the intermediate stages of T cell developmental (Hosoya et al., 2010). However, it is also important for repression transcription of certain target genes, such as Foxp3, in naïve T cells, preventing them from becoming regulatory T cells (Hosoya et al., 2010).

Recently, a number of reports have demonstrated that E4bp4 can influence the epigenetic modifications of chromatin at promoter regions of target genes. Many of these studies involve E4bp4 promoting a transcriptionally repressive state by inducing deacetylation of histone H3 and/or methylation of histone H3K9. In T_H2 cells, E4bp4 directly binds the *IL-13* locus and in its absence increased histone H3 acetylation and increase methylation of histone H3K4 was observed (Kashiwada et al., 2011a). In a separate study, E4bp4 was shown to bind to the *CD40L* locus in $CD4^+$ T cells where it promoted an increase in H3K9 methylation, a reduction in H3 acetylation and a reduction in H3K4 methylation, all hallmarks of transcriptional repression (Zhao et al., 2013). MacGillavry *et al*. also found that many E4bp4 bound promoters in neuronal cells were hypoacetylated, suggesting E4bp4 serves to keep these genes in a suppressed, transcriptionally silent state (MacGillavry et al., 2011).

In a report looking at cell survival in breast cancer cells, E4bp4 was found to prevent transcription of TRAIL and some FOXO genes, preventing apoptosis of these cells (Keniry et al., 2013). This study was the first to suggest a mechanism by which E4bp4 can regulate epigenetic modifications, as it was shown to promote deacetylation by recruiting histone deacetylase-2 and reducing levels of H3 acetylation (Keniry et al., 2013). Another possible mechanism of action for E4bp4 was proposed by Tong *et al*. who showed E4bp4 was able to interact with histone methyltransferase G9a and this interaction caused increased H3K9 methylation at the loci of the E4bp4 target gene *Fgf21* in liver cells (Tong et al., 2013).

However, the interactions observed in both of these studies were not performed on endogenous protein in the cell types of interest and so further investigations should confirm that these interactions occur *in vivo* and effect the transcription of E4bp4 target genes.

As well as repressing transcription of target genes by modulating chromatin structure E4bp4 has also been implicated with promoting transcription by epigenetic modulation. In chronically stimulated T_H1 cells E4bp4 can bind to the *IL-13* locus and promote its expression and this phenotype was observed alongside an increase in H3K4 methylation in these cells, a common mark of gene activation (Motomura et al., 2011). Additionally, the same study found that in T_H2 cells, in the absence of E4bp4, the *IL-10* locus had less H3 acetylation at K9 and K14, suggesting E4bp4 is able to promote an active state of this locus in this cell type (Motomura et al., 2011). Nandakumar *et al*. recently showed that E4bp4 can interact with a histone H2A deubiquitinase, MYSM1, and this can promote a more active transcription state in the promoter of Id2 in NK cells (Nandakumar et al., 2013). They proposed that this interaction helps to keep the Id2 locus active during NK cell development, especially as in the absence of MYSM1 NK cell development and maturation was severely affected (Nandakumar et al., 2013). E4bp4's interactions with chromatin remodelling enzymes may therefore be important for its activities in numerous different contexts.

1.5 Other functions of E4bp4

E4bp4 is essential for the development of cNK cells, commitment of developing lymphocytes to the NK lineage and the development of other ILC subsets including LTi cells. However, E4bp4 has been implicated in a variety of other roles in numerous cells types **(Figure 1.7)**. Particularly, the characterisation of *E4bp4*$^{-/-}$ mice has revealed roles for E4bp4 in many immune lineages **(Figure 1.7)**.

1.5.1 CD8α$^+$ conventional dendritic cell development

E4bp4$^{-/-}$ mice are deficient in CD8α$^+$ conventional dendritic cells (cDCs), but other DC subsets and progenitors are not affected by the absence of E4bp4, suggesting it has a role late in the developmental pathway specific for CD8α$^+$ DCs (Kashiwada et al., 2011b). A vital role for CD8α$^+$ DCs is their ability to cross present antigen from endocytosed particles on MHC-I molecules and therefore *E4bp4*$^{-/-}$ mice are unable to mediate cross presentation and induce CD8$^+$ T cell responses (Kashiwada et al., 2011b). More recently, Seillet *et al*. used

BM reconstitution experiments to show that CD8α$^+$ cDCs could still be developed in the absence of E4bp4, suggesting it is not an absolute requirement for the development of this cell type, but rather an important regulator (Seillet et al., 2013).

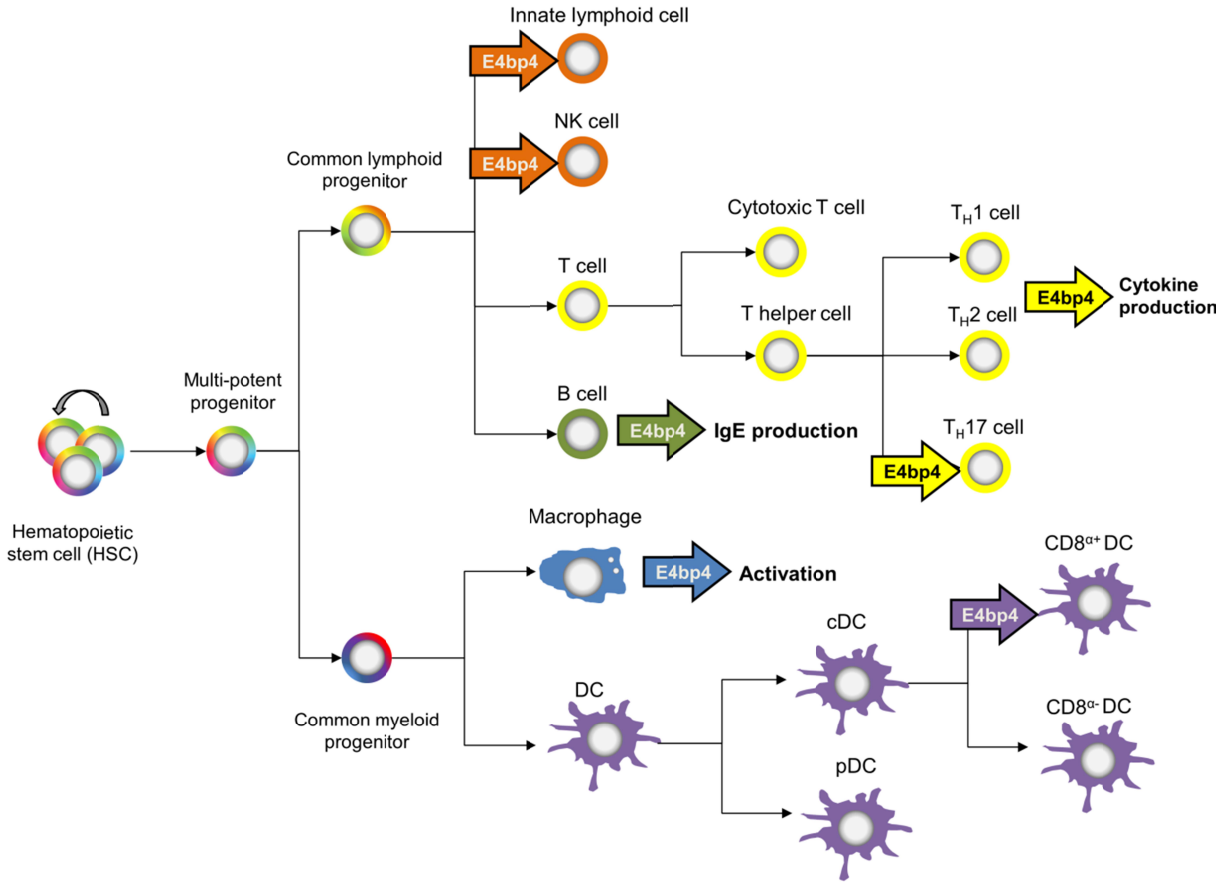


Figure 1.7 – Functions of E4bp4 in the immune system

E4bp4 is important for numerous immunological processes. E4bp4 is essential for the development of NK cells, all ILC subsets, CD8α$^+$ cDCs and T$_H$17 T cells. E4bp4 is required for B cells to produce IgE and in the activation of macrophages.E4bp4 also regulates the cytokines that are produced by T$_H$1 and T$_H$2 cells. cDC - conventional dendritic cell, pDC - plasmacytoid dendritic cell. Adapted from (Male et al., 2012).

1.5.2 B cell class switching

E4bp4 regulates the IL-4 driven process of immunoglobulin heavy chain class switching required for IgE production (Kashiwada et al., 2010). *E4bp4*$^{-/-}$ mice have impaired IgE class switching and germ-line ε (GLε) transcripts are not induced in *E4bp4* deficient B cells (Kashiwada et al., 2010). E4bp4 expression in B cells is induced by IL-4 via STAT6

activation (Schroder et al., 2002) and, by directly binding to the Ig GLε promoter, E4bp4 activates its transcription (Kashiwada et al., 2010). Despite this clear phenotype it remains unclear whether E4bp4 is a regulator of the biological processes involving IgE, e.g. immune responses to pathogens at mucosal surfaces and allergic responses.

1.5.3 T cell cytokine production and activity

T helper cells are categorised into different subsets based on their ability to produce effector cytokines and regulate immune responses. Two conflicting reports have described a role for E4bp4 in $CD4^+$ T helper cells, where it directly affects cytokine production. Motomura *et al*. reported that T_H2 cells require E4bp4 to produce IL-10, but not for the production of IL-4, IL-5 and IL-13 (Motomura et al., 2011). In contrast, Kashiwada *et al*. found that E4bp4 promoted the expression of IL-4 in T_H2 cells developing *in vitro* and actively repressed the production of IL-5 and IL-13 (Kashiwada et al., 2011a). A similar phenotype was observed *in vivo* when looking at T_H2 responses in mice immunised with ovalbumin and alum (Kashiwada et al., 2011a). Kashiwada *et al*. did observed that in the earliest stages of T_H2 differentiation E4bp4 actually suppressed IL-4 production and this may explain why Motomura *et al*. observed no effect on IL-4 production (Kashiwada et al., 2011a). However, it is not clear why the two studies disagree about the effect of E4bp4 on IL-5 and IL-13 expression, although E4bp4 does appear to directly bind to the *IL-13* promoter in T_H2 cells and actively repress transcription (Kashiwada et al., 2011a).

T_H1 cells are also able to produce the T_H2 cytokines IL-10 and IL-13 during chronic infection and the production of both of these cytokines also appears to be driven by E4bp4 (Motomura et al., 2011). In T_H1 cells E4bp4 was shown to again bind directly to the IL-13 promoter, but rather than repress transcription, as seen in T_H2 cells, it appears to promote transcription (Motomura et al., 2011). *E4bp4*$^{-/-}$ T_H1 cells produce less IL-13 and also less IL-10, when chronically stimulated, and T_H1 cells from mice transgenically overexpressing E4bp4 are able to produce more IL-10 and IL-13 than WT cells (Motomura et al., 2011). E4bp4 was not observed to bind directly to the IL-10 promoter but appeared to affect transcription by modulating chromatin structure (Motomura et al., 2011). These effects on cytokine production may have important clinical implications because *E4bp4*$^{-/-}$ mice spontaneously develop intestinal inflammation and colitis, which at least in part is due to the deregulation of T cell cytokine production (Motomura et al., 2011).

It has also been discovered that E4bp4 inhibits the activity of self-reactive T cells during systemic lupus erythematosus (SLE) (Zhao et al., 2013). E4bp4 was shown to down regulate the expression of CD40L in SLE $CD4^+$ T cells which is required to promote autoantibody production from B cells (Zhao et al., 2013). E4bp4 was therefore suggested to be important for inhibiting $CD4^+$ T cell activation and subsequent autoimmune responses. However, it is not clear if E4bp4's activity is crucial *in vivo* as SLE $CD4^+$ T cells isolated from patients had increased E4bp4 expression and are known to also have hyperexpression of CD40L (Desai-Mehta et al., 1996).

1.5.4 Macrophage activation and cytokine production

Macrophages are central to inflammatory innate immune defences at mucosal surfaces, but they are also involved in the regulation of responses to enteric microbiota. E4bp4 has a role in controlling macrophage activity, since stimulation with lipopolysaccharide (LPS) or commensal bacteria causes bone marrow derived macrophages (BMDMs) to upregulate *E4bp4* expression (Kobayashi et al., 2011) (Smith et al., 2011). The anti-inflammatory cytokine IL-10 also promotes the transcription of *E4bp4* in BMDMs, suggesting E4bp4 is a component of the IL-10 anti-inflammatory signalling pathway (Lang et al., 2002) (Smith et al., 2011). E4bp4 produced in BMDMs controls the expression of the pro-inflammatory cytokine IL-12 by binding directly to a DNA enhancer upstream of the *IL-12b* transcriptional start site and repressing its activity (Smith et al., 2011). In the absence of E4bp4, mice challenged with *Toxoplasma gondii* fared far worse than WT littermates and produced greater levels of systemic IL-12 (Smith et al., 2011). E4bp4 is therefore important for maintaining intestinal homeostasis by acting as a microbiota-induced *IL-12b* transcriptional inhibitor.

A recent study by Kobayashi *et al.* has confirmed the spontaneous colitis that develops in *E4bp4*$^{-/-}$ mice is specifically due to the dysregulation of *IL-12b* in intestinal macrophages and is not dependent on the presence of other cytokines produced by lymphocytes (Kobayashi et al., 2014). Allied to this, $CD14^+$ macrophages from patients with Crohn's disease or ulcerative colitis, which produce large amounts of IL-12, express lower levels of E4bp4 compared with those from non-inflammatory controls (Kobayashi et al., 2011).

1.5.5 Circadian rhythm and T_H17 development

Internal circadian clocks are found in a diverse range of organisms, controlling physiological and behavioural responses and are driven by autoregulatory transcriptional and translational

feedback loops that generally oscillate over a 24 hour period (Ohno et al., 2007b). E4bp4 mRNA and protein expression fluctuate in a regular pattern over a 24 hour period and in an opposing manner to essential circadian rhythm genes such as *DBP* (Mitsui et al., 2001) and *period2* (*Per2*) (Doi et al., 2001). In the chick pineal gland E4bp4 expression is induced by light and its expression was to shown to suppress the expression of *Per2* in a recombinant system (Doi et al., 2001). This repression was later confirmed *in vivo* and it was confirmed that E4bp4 regulation of *Per2* is required for correct expression of *Per2* in the cell-autonomous clock (Ohno et al., 2007b). E4bp4 was also found to interact with Per2 protein and another mammalian circadian clock protein Cryptochrome2 (Cry2), suggesting E4bp4 is part of a negative regulator complex of mammalian circadian clock proteins (Ohno et al., 2007a).

E4bp4's role in regulating the circadian clock has been linked to a number of biological processes, most prominently the development of interleukin-17-producing $CD4^+$ T helper (T_H17) cells. T_H17 cells are pro-inflammatory modulators that protect mucosal surfaces from bacterial and fungal infections (Yu et al., 2013b). *E4bp4*$^{-/-}$ mice have significantly higher numbers of T_H17 cells than WT mice in both the small intestine and colon (Yu et al., 2013b). T_H17 development requires the orphan nuclear receptor, RORγt, and E4bp4 represses transcription of RORγt by binding directly to its promoter (Yu et al., 2013b). Yu *et al*. were also able to show that E4bp4 expression in T_H17 cells is regulated by the circadian clock (Yu et al., 2013b). Primary circadian clock activators BMAL1 and CLOCK promote the activity of the TF REV-ERBα, which represses E4bp4 expression. Therefore T_H17 development is directly linked to the circadian clock and occurs in a diurnal fashion (Yu et al., 2013b).

1.5.6 Cell survival

E4bp4 was first shown to inhibit apoptosis in IL-3-dependent pro-B cell lines, as enforced expression of E4bp4 provided protection from cell death upon cytokine withdrawal (Ikushima et al., 1997). Oncogenic proteins Ras and Gata-1 were both suggested to promote E4bp4 expression in these pro-B cell lines, in response to IL-3 (Kuribara et al., 1999). However, the anti-apoptotic effects of E4bp4 in these cells were not consistently observed and the oncoprotein E2A-HLF, which causes childhood acute lymphoblastic leukaemia, was found to be a more potent pro-survival factor for pro-B cells (Yeung et al., 2004).

Recently, Keniry *et al*. demonstrated that E4bp4 expression promotes cell survival in breast cancer cells and promotes resistance to oxidative stress (Keniry et al., 2013). E4bp4

prevented FOXO TFs from accessing the promoter of TRAIL (the death receptor ligand capable of inducing extrinsic apoptosis) by binding to nearby chromatin, recruiting histone deactelyase-2, reducing histone acetylation and ultimately preventing transcriptional activation (Keniry et al., 2013). E4bp4 was also found to be overexpressed in a number of different breast cancer cell types and its expression was associated with lower patient survival (Keniry et al., 2013).

1.5.7 Motor neuron development and survival

E4bp4 is expressed by embryonic chicken and rat motor neurons and high levels of expression have been associated with long term survival of neurons (Junghans et al., 2004). Overexpression of E4bp4 was found to promote neuronal growth *in vitro* and *in vivo*, reducing the number of dying neurons in chicken embryos by 45% (Junghans et al., 2004). Recently, it has been shown that this effect on neuronal survival may have clinical implications because expression of E4bp4 in neurons appears to protect against the neurodegenerative disease Amyotrophic Lateral Sclerosis (ALS) (Tamai et al., 2014). E4bp4 expression is increased following neurotoxic damage and over expression of E4bp4 in ALS mice was found to reduce motor axon degeneration and delay the onset of disease (Tamai et al., 2014).

In contrast, two studies have shown that E4bp4 has a different role in neurons more closely associated with the central nervous system. MacGillavry *et al*. found that E4bp4 represses regeneration and cellular outgrowth of dorsal root ganglion cells following neuronal injury (MacGillavry et al., 2009). In dorsal root ganglion cells, E4bp4 acts antagonistically, alongside the TFs CREB and C/EBP, to regulate expression of regeneration-associated genes (MacGillavry et al., 2009) (MacGillavry et al., 2011). These studies clearly demonstrate how E4bp4 has cell type specific roles and can perform different roles in cells at different developmental stages.

1.5.8 Osteoblast function

Myeloma bone disease is a common symptom in multiple myeloma patients, caused by defects in osteoblasts, which prevents new bone formation and repair (Silvestris et al., 2008). E4bp4 is involved with this impairment of osteoblast function as it transcriptionally represses the activity of important osteoblast TFs Runx2 and Osterix (Ozkurt and Tetradis, 2003) (Silvestris et al., 2008). E4bp4 expression is induced in osteoblasts by parathyroid

hormone-related protein, which is commonly released by myeloma cells (Ozkurt and Tetradis, 2003) (Silvestris et al., 2008). This observation was validated by looking at E4bp4 expression in osteoblast isolated directly from myeloma patients, where high levels of E4bp4 were correlated with low levels of Osterix expression and high levels of serum parathyroid hormone-related protein (Silvestris et al., 2008).

1.6 Regulation of E4bp4

E4bp4 is involved in a broad range of biological processes and is able to produce differing effects on target gene expression, in different cell types, and in different contexts. In many instances it is clearly essential to repress gene transcription, but in others it is required to promote transcription. Arguably the most striking phenotype of the *E4bp4*$^{-/-}$ mouse is its absolute lack of NK cells, NK cells precursors and its dramatic loss of all ILC subsets (Gascoyne et al., 2009) (Geiger et al., 2014) (Male et al., 2014) (Seillet et al., 2014b). Over the past five years since the *E4bp4*$^{-/-}$ mouse was first reported, the importance of E4bp4 as a central regulator of numerous aspects of the immune system has continued to grow. Despite this only a small number of studies have begun to analyse how the activities of E4bp4 are regulated.

1.6.1 Transcriptional regulation of *E4bp4*

Due to its importance the activity of E4bp4 in a particular cell type is likely to be tightly regulated, but the precise molecular mechanisms by which this occurs remain unclear. In particular during NK cell development it is not known how E4bp4 activity is regulated and how it is able to promote an NK cell specific programme of gene expression. Various cytokines have been found to regulate the expression of *E4bp4*, including IL-3 in pro-B cells (Ikushima et al., 1997) (Yeung et al., 2004); IL-4, via STAT6 signalling, in T and B cells (Chen et al., 2003) (Schroder et al., 2002); IL-10, via STAT3 signalling in activated macrophages (Lang et al., 2002), and IL-15 in T cells (Ramsborg and Papoutsakis, 2007) and NK cells (Gascoyne et al., 2009). E4bp4 is expressed in CLPs and is required for the development of committed pre-NKPs, but NK cells don't express the IL-15 receptor CD122 until the rNKP development stage (Fathman et al., 2011). Therefore IL-15 is important for driving the expression of E4bp4 at a later stage of NK cell development, but it is likely to not be the only regulator. Other mechanisms must be utilised to control its activity early in NK cell development and in CLPs.

Glucocorticoids (GCs) have also been found to regulate the expression of *E4bp4* (Priceman et al., 2006) (Wallace et al., 1997). In human and murine T cells, GCs can upregulate *E4bp4* expression leading to up regulation of Bim (member of Bcl-2 family, required for apoptosis) and eventually leading to GC-induced apoptosis (Beach et al., 2011) (Carey et al., 2013) (Priceman et al., 2006) (Zhao et al., 2013). A similar phenotype has also now been observed in osteoblasts, where E4bp4 induction, results in apoptosis (Chen et al., 2014). Whether GCs can regulate E4bp4 expression in NK cells and ILCs is unknown, but as these cells are involved in numerous inflammatory diseases, often treated by GCs, it would be interesting to determine if GCs influence E4bp4 activity in this cell type.

1.6.2 Post transcriptional regulation of E4bp4

Regulating TFs often occurs on the protein level, rather than at the level of transcription, so that rapid changes can occur to gene expression, allowing cells to adapt quickly to environmental signals. The activity of TFs is therefore essential for controlling cellular development and activation (Brivanlou and Darnell, 2002) (Deribe et al., 2010). The activities of TFs are controlled by four major mechanisms:

A. Regulation by protein-ligand binding; e.g. the transcriptional repressor DREAM is directly modulated by Ca^{2+} levels. Ca^{2+} binds directly to the DREAM protein, reducing its ability to repress transcription of target genes (Rivas et al., 2004). The DREAM protein has four Ca^{2+} binding motifs that, in the absence of Ca^{2+}, interact as a tetramer with regulatory elements to inhibit transcription of target genes (Rivas et al., 2004). DREAM has a number of roles, but it is prominently involved with sensing pain in specific areas of the spinal cord (Cheng et al., 2002).

B. Regulation by protein-protein interactions; e.g. the TF NFκB associates with IκB in the cytoplasm of unstimulated cells and cannot promote the activation of target genes (Viatour et al., 2005). Following cellular stimuli (e.g. by cytokines or bacterial lipopolysaccharides), NFκB can be rapidly activated by cellular signalling that induces phosphorylation and ultimately degradation of IκB, allowing NFκB to freely localise to the nucleus and promote gene expression (Viatour et al., 2005). Activation of NFκB is important for inflammatory responses as it induces cell proliferation and anti-apoptotic gene expression, but aberrant NF-kB activity is common during inflammatory disease (Viatour et al., 2005).

C. Regulation by proteasomal degradation; e.g. The gene silencing TF REST is expressed in pluripotent stem cells and actively represses the transcription of neuronal-specific genes (Ballas et al., 2005). As stem cells develop into neuronal precursors REST is degraded more rapidly, reducing the amount of protein in the cell, allowing neuronal genes to be initially expressed (Ballas et al., 2005). Once neurons fully develop synthesis of REST is switched off completely and neuronal genes are fully active (Ballas et al., 2005).

D. Regulation by post translational modifications (PTMs); hundreds of TFs have been identified to be modified by numerous different PTMs that regulate their activity in a positive or negative fashion. For example, acetylation of the TF p53, by the acetyl-transferase p300, disrupts interactions with its inhibitor Mdm2, increases its DNA binding capacity, leading to the activation of stress-related genes (Tang et al., 2008).

Many of these mechanisms for regulating TF activity overlap; PTMs can promote or prevent interactions between proteins, whilst other PTMs can lead to the degradation of a modified protein (Filtz et al., 2014). PTMs serve to exponentially increase the functionality of the proteome in vertebrates; by creating modified forms of pre-existing proteins, multiple functions can therefore be provided by a single gene product **(Figure 1.8)**. PTMs help to make up for the relatively small number of genes encoded for by the genomes of complex mammals and permit the tissue and cell specific activities of individual TFs **(Figure 1.8)**. Large numbers of different PTMs have now been identified, but some of the most common include: phosphorylation, ubiquitination, acetylation, glycosylation, methylation and SUMOylation. Some modifications are simple small chemical modifications whilst others are much larger more complex molecules, but all can have significant effects on the activity of a target protein. PTMs can alter every aspect of a TFs activity, including its localisation, stability, interactions with other proteins, DNA binding capacity, secondary and/or tertiary structure and ultimately transcriptional activity (Benayoun and Veitia, 2009).

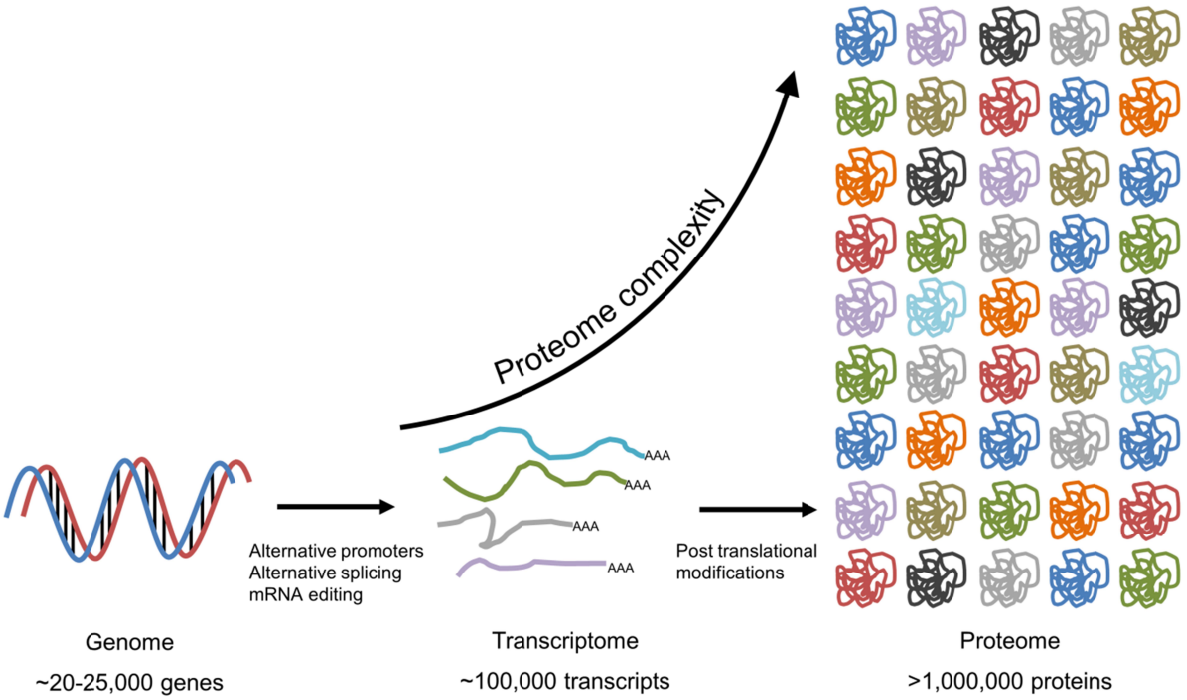


Figure 1.8 – Complexity of the mammalian proteome

Human and murine genomes contain 20-25,000 genes, but more than 1 million proteins are estimated to be part of the proteome. The transcriptome is relatively larger than the genome due to changes at the transcriptional and mRNA level. However, numerous different post translational modifications of proteins exponentially increase the size of the proteome relative to both genome and transcriptome. This ultimately allows complex organisms to form with vast numbers of differentiated cells and tissues. Adapted from: http://www.piercenet.com/method/overview-post-translational-modification

Despite its importance, especially for NK cells and other innate lymphocytes, very little is known about how the activity of the E4bp4 protein is regulated. E4bp4 was originally proposed to be able to repress gene transcription by interacting with the TATA box binding protein Dr1 (Cowell and Hurst, 1996). This interaction was suggested to result in an interruption of the basal transcriptional machinery at the promoter region of E4bp4 target genes (Cowell and Hurst, 1996). However, the interaction between E4bp4 and Dr1 was only observed *in vitro* and this interaction has yet to be shown to have any phenotypic consequences. As mentioned above (section 1.4) a number of epigenetic modification enzymes have been proposed to interact with E4bp4, resulting in changes to chromatin structure at E4bp4 target gene *loci*. In hepatocytes E4bp4 was shown to interact with the histone methyltransferase G9a, whilst in breast cancer cell lines it was observed to interact with histone deacetylase-2 (Keniry et al., 2013) (Tong et al., 2013). Both of these

interactions caused E4bp4 to repress target gene expression, by inducing a repressive chromatin structure. Whether these interactions are important in other cell types, and in other contexts, remains to be explored.

The only mechanism of E4bp4 regulation that is associated with NK cells is its interaction with the histone H2A deubiqutinase MYSM1 (Nandakumar et al., 2013). Nandakumar *et al*. were able to show that MYSM1 associates with E4bp4 at the *Id2* locus in NK cells and both are required for Id2 transcription (Nandakumar et al., 2013). E4bp4 co-immunoprecipitated with MYSM1, but only when the proteins were over expressed together in 293T cells, rather than directly from NK cells (Nandakumar et al., 2013). It would be interesting to ascertain if this interaction is required for the transcriptional regulation of other NK cell specific genes and if this is a mechanism employed by E4bp4 to regulate target gene expression in other cell types.

As E4bp4 is central to numerous biological processes it is likely to be regulated by complex mechanisms and if these can be understood in more detail it may allow for the activity of E4bp4 to be modulated. For example, increased expression of E4bp4 in NK cell development has been shown to promote the development of greater numbers of NK cells (Gascoyne et al., 2009). This could potentially be important when it comes to developing a system for producing large quantities of cytotoxic NK cells for immunotherapeutic purposes. Numerous strategies have been developed for creating clinical-grade NK cells for adoptive transfer into patients, but modulating the activity of E4bp4 in these systems could help to further increase quantity and potentially cytotoxicity of the NK cells produced.

To begin to ascertain in more detail how the activities of E4bp4 may be regulated, the Brady laboratory performed a yeast-two-hybrid screen to identify proteins that interact with E4bp4 (Brady laboratory, unpublished). Full length E4bp4 protein was used as the bait and screened for interactions with the human leucocyte/activated mononuclear cell library (Hybrigenics). Twenty one proteins were identified in the screen, but only 11 of these received multiple positive identifications **(Table 1.1)**. Many of the identified proteins have no known function or were common contaminants in yeast-two-hybrid analysis such as cytoskeletal proteins. No interactions were observed between E4bp4 and previously suggested binding partners including Dr1 (Cowell and Hurst, 1996), CREB (Acharya et al., 2006) or any histone modification enzymes.

Strikingly the protein with the highest number of positive identifications was PIAS1 **(Table 1.1)**. PIAS1 is a SUMO E3 ligase, which is involved in the addition of post translational

SUMO modifications onto target proteins. Additionally, positive hits were also achieved for PIAS3 and RanBP2, both of which are also SUMO E3 ligases **(Table 1.1)**. This data provided the first evidence that E4bp4 may be post translationally SUMOylated. As very few others proteins were identified to interact with E4bp4 in the yeast-two-hybrid screen, other than SUMO E3 ligases, it was postulated that these interaction could be important for its functions *in vivo*.

Table 1.1 – E4bp4 interacting partners from yeast-two-hybrid screen.

Gene Name	Positive Identifications	Protein function
C18orf25	1	Unknown
CHD1	1	DNA helicase binding protein
DSP	2	Cytoskeletal protein
FLJ13057	5	Unknown
FLJ544447	1	Unknown
FLNA	1	Cytoskeletal protein
HIPK1	1	Serine/threonine protein kinase
HIPK3	5	Serine/threonine protein kinase
PIAS1	20	SUMO E3 ligase
PIAS3	2	SUMO E3 ligase
RANBP2	1	SUMO E3 ligase
RNF111	1	SUMO-targeted ubiquitin ligase
SETX	2	DNA helicase
SNRP70	2	Splicesomal ribonucleoprotein
SORL1	1	Neuronal apolipoprotein E receptor
TLK2	2	Tousled-like serine/threonine kinase
U5-200KD	1	RNA helicase
ZBTB16	1	Zinc finger transcription factor
ZMYM5	5	Zinc Finger MYM-Type Protein
ZNF198	5	Zinc Finger Transcriptional cofactor
ZNF237	5	Zinc Finger Transcriptional cofactor

The only previous evidence of a PTM regulating E4bp4 activity is for phosphorylation and the data about this PTM is limited and contradictory. Chen *et al*. first identified E4bp4 as a phosphoprotein almost 20 years ago and suggested phosphorylation was important for E4bp4's ability to bind DNA (Chen et al., 1995). In contrast, a study by Doi *et al*. studying E4bp4 in the chick pineal gland, suggested E4bp4 is phosphorylated by Casein Kinase 1ε (CK1ε) and modification by this kinase reduces the activity of E4bp4 (Doi et al., 2004). However, both of these studies used recombinant and *in vitro* systems to study E4bp4 and their observations may not be indicative of what occurs to E4bp4 *in vivo*. Additionally, no specific phosphorylation sites for E4bp4 have been conclusively identified. As PTMs can have profound effects on the activity of TFs, understanding the modifications of E4bp4 could be vital for fully establishing how it regulates various processes in the immune system. Whether phosphorylation or potentially SUMOylation are required for the functions of E4bp4, particularly in NK cell development is an important area of research that is currently unexplored.

1.7 Project aim and research questions

E4bp4 is a regulator of various immunological processes including class switching in B cells to produce IgE, cytokine production by $CD4^+$ T helper cells and in the development of $CD8\alpha^+$ cDCs (Male et al., 2012). Its most prominent role appears to be in the development of NK cells and other ILC subsets, where it is essential for developing lymphocytes to commit to become innate lymphocytes rather than T or B cells (Geiger et al., 2014) (Male et al., 2014) (Seillet et al., 2014b). However, very little is known about the molecular mechanisms that control the activities of E4bp4, particularly in NK cells.

The main aim of this thesis was to ascertain a greater understanding of the regulation of the bZIP transcription factor E4bp4, particularly by PTMs and determine if this affected NK cell development and function. The specific research questions of this thesis were:

1) Determine if E4bp4 is SUMOylated *in vivo* using HeLa cell lines stably expressing 6His-SUMO1, 6His-SUMO2 and 6His-SUMO3.
2) Ascertain the exact sites of E4bp4 SUMOylation using high accuracy mass spectrometry and mutate these sites to create versions of E4bp4 that lack all SUMO modifications.
3) Using an *in vitro* NK cell development assay to establish whether SUMOylation of E4bp4 is required for regulating its activity.
4) Confirm that E4bp4 is phosphorylated and use mass spectrometry to accurately determine which amino acids are phosphorylated.
5) Mutate E4bp4 phosphorylation sites and establish whether the phosphorylation of E4bp4 is important for its function in NK cell development.
6) Determine how the post translational modifications of E4bp4 influence its function as a TF, using transcriptional reporter assays and quantitative PCR.

CHAPTER 2

Materials and Methods

2.1 Buffers and reagents

Buffer	Components
TAE buffer	40 mM Tris 20 mM Acetic Acid 1 mM EDTA pH 8.0
LB medium	10 g/L Peptone 10 g/L Yeast extract 5 g/L Sodium chloride
SOC medium	2% (w/v) Peptone 0.5% (w/v) Yeast Extract 10 mM NaCl 2.5 mM KCl 10 mM $MgCl_2$ 10 mM $MgSO_4$ 20 mM Glucose
NZY+ medium	10% (w/v) NZ amine 5% (w/v) Yeast extract 5% (w/v) NaCl 12.5 mM $MgCl_2$ 12.5 mM $MgSO_4$ 20 mM Glucose pH 7.5 – adjusted with NaOH
LB agar medium	10 g/L Peptone 5 g/L Yeast extract 10 g/L Sodium chloride 15 g/L Agar
SDS-PAGE running buffer	25 mM Tris 192 mM Glycine 0.1% (w/v) SDS
Transfer buffer	48 mM Tris 39 mM Glycine 0.04% (w/v) SDS 20% (v/v) Methanol

Buffer	Components
Laemmli sample buffer	125 mM Tris-HCl pH 6.8 2% (w/v) SDS 25% (v/v) Glycerol 0.005% (w/v) Bromophenol Blue 5% (v/v) β-mercaptoethanol
PBST	PBS 0.1% (v/v) Tween20
Blocking buffer	PBS 0.1% (v/v) Tween20 3% (w/v) Skim Milk Powder
CE buffer	10 mM Hepes pH 8.0 1.5 mM $MgCl_2$ 10 mM KCl 1% (v/v) Trition X-100 0.05% (v/v) β-mercaptoethanol Complete protease inhibitors
NE buffer	10 mM Hepes pH 8.0 1.5 mM $MgCl_2$ 0.2 mM EDTA 25% (v/v) glycerol 450 mM NaCl 0.05% (v/v) β-mercaptoethanol Complete protease inhibitors
Low salt TBS	50 mM Tris-HCl pH 7.4 100 mM NaCl 0.01% (v/v) Trition X-100
High salt TBS	50 mM Tris-HCl pH 7.4 300 mM NaCl 0.01% (v/v) Trition X-100
FLAG elution buffer	150 ng/ul FLAG peptide 50 mM Tris-HCl pH 7.4 300 mM NaCl 0.01% (v/v) Trition X-100

Buffer	Components
0.1 M Sodium phosphate buffer pH 6.3	45 mM Na_2HPO_4 155 mM NaH_2PO_4 50% (v/v) H_2O
0.1 M Sodium phosphate buffer pH 8.0	190 mM Na_2HPO_4 10 mM NaH_2PO_4 50% (v/v) H_2O
SUMO cell lysis buffer	6 M Guanidinium-HCl 10 mM Tris 0.1 M Sodium phosphate buffer pH8
pH 8.0 wash buffer	8 M Urea 10 mM Tris 0.1 M Sodium phosphate buffer pH 8.0 0.1% (v/v) Trition X-100 5 mM β-mercaptoethanol
pH 6.3 wash buffer	8 M Urea 10 mM Tris 0.1 M Sodium phosphate buffer pH 6.3 0.1% (v/v) Trition X-100 5 mM β-mercaptoethanol
SUMO elution buffer	200 mM Imidazole 5% (w/v) SDS 150 mM Tris-HCl pH 6.7 30% (v/v) glycerol 720 mM β-mercaptoethanol 0.005% (w/v) bromophenol blue
Two-step lysis buffer	20 mM Tris-HCl pH 7.4 10 mM $MgCl_2$ 10 mM KCl 2 mM EDTA 10% (v/v) glycerol 1% (v/v) Trition X-100 Protease and phosphatase inhibitors 25 mM N-Ethylmaleimide

Buffer	Components
MACS buffer	PBS 2 mM EDTA pH 8.0 0.5% (w/v) BSA
FACS buffer	PBS 1% (v/v) FBS
Complete cytokine medium	DMEM 10% (v/v) FBS 100 U/ml Penicillin 100 µg/ml Streptomycin 10 ng/ml Flt3 ligand 10 ng/ml IL-7 100 ng/ml Stem cell factor 50 µM β-mercaptoethanol
IL-15 medium	Minimal essential medium 20% (v/v) FBS 100 U/ml Penicillin 100 µg/ml Streptomycin 50 µM β-mercaptoethanol 30 ng/ml IL-15
NK cell differentiation medium	DMEM 20% (v/v) Human serum 100 U/ml Penicillin 100 µg/ml Streptomycin 50 µM β-mercaptoethanol 10 ng/ml IL-15 5 ng/ml IL- 3 (first week of culture only) 20 ng/ml IL-7 20 ng/ml Stem cell factor 10 ng/ml Flt3 ligand

2.2 Cells and cell culture

2.2.1 Cell lines

Immortalised cell lines and primary cells used in this study were maintained at 37˚C and 5% CO_2. HEK-293T, HeLa, NIH-3T3, Cos-7, MNK-1 and LinXE cell lines were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Sigma Aldrich). OP9-GFP stromal cells were cultured in Iscove's Modified Dulbecco's Media (IMDM) (Sigma Aldrich) and BaF3 cells were cultured in RPMI-1640 (Sigma Aldrich). The media for all cell lines was supplemented with 10% fetal bovine serum (FBS) (Gibco), 2 mM L-glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin (Sigma Aldrich).

MNK-1 cells (gift from David S.J. Allan and James R. Carlyle, University of Toronto, Canada) are a murine NK cell line and were cultured in the presence of 25 mM HEPES, 50 µM β-mercaptoethanol, 1 mM sodium pyruvate and 10 ng/mL IL-2 (Peprotech). HeLa cells stably expressing 6His-SUMO-1, 6His-SUMO-2 and 6His-SUMO-3 (gift from Ronald T Hay, University of Dundee, UK) were cultured in the presence of 5 µM puromycin, to maintain selection of 6His-SUMO proteins. LinXE cells were cultured under 15 µM hygromycin selection to maintain the expression of retroviral packaging genes.

E108.1D2 stromal cells were cultured in Alpha-MEM with Glutamax (Gibco) supplemented with 50% Myelocult M5300 (Stem Cell Technologies), 7.5% FBS, 50µM β-mercaptoethanol, 1 µM hydrocortisone, 100 U/ml penicillin and 100 µg/ml streptomycin (Sigma Aldrich). For $CD34^+$ stem cell differentiation experiments E108.1D2 cells were plated into 96-well EmbryoMax Gelatin (Millipore) coated plates, irradiated at 3000 rads/30 Gy and maintained at 32°C, before $CD34^+$ were plated onto of them.

2.2.2 Transfections

Cell lines were transfected with a number of plasmid constructs and in each case a 100 mm dish of cells was grown to 60 - 85% confluencey on the day of transfection. HeLa and HEK-293T cells were transfected with 8 µg of plasmid DNA and 40µl of the transfection reagent Lipofectamine (Invitrogen). LinXE cells were transfected with 2 µg of plasmid DNA using 20µl of the transfection reagent Effectene (Qiagen). 3T3 cells were cultured in 24 well plates prior to transfection (for Luciferase reporter assay) and each well of cells was transfected with 200 - 500 ng of plasmid DNA and 5 µl Effectene. All transfectants were incubated for 48 hours before transgene expression was analysed and/or cell lysates were prepared.

2.2.3 Mice

All animals used were on a C57BL/6 background, between six and twelve weeks old and were matched for age and gender. Where possible, littermates were used. *E4bp4*$^{-/-}$ mice were generated as previously described (Gascoyne et al., 2009). Mice were genotyped with the forward primer 5'~CTCTGAGCTTGGCTGATGTG~3' (primer A) and reverse primer 5'~GCTTCAAGTCTCCACCAAGC~3' (primer B) for detection of the wild type allele or 5'~CCATGCTCCTGTCTTGATGA~3' (primer C) for detection of the knockout allele **(Figure 2.1)**. The phenotype of all mice used was confirmed by flow cytometry, with knockouts defined by the absence of NK1.1^{+} CD3^{-} splenocytes **(Figure 2.2)**. Whole mouse spleens were crushed through a 40 µM strainer into 1 ml PBS. 10% of the cell suspension was ACK lysed for 3 minutes, washed in PBS and resuspended in 200 µl FACS buffer for staining with antibodies against NK1.1 and CD3. All animal husbandry and experimental procedures were carried out according to UK Home Office regulations and institute guidelines.

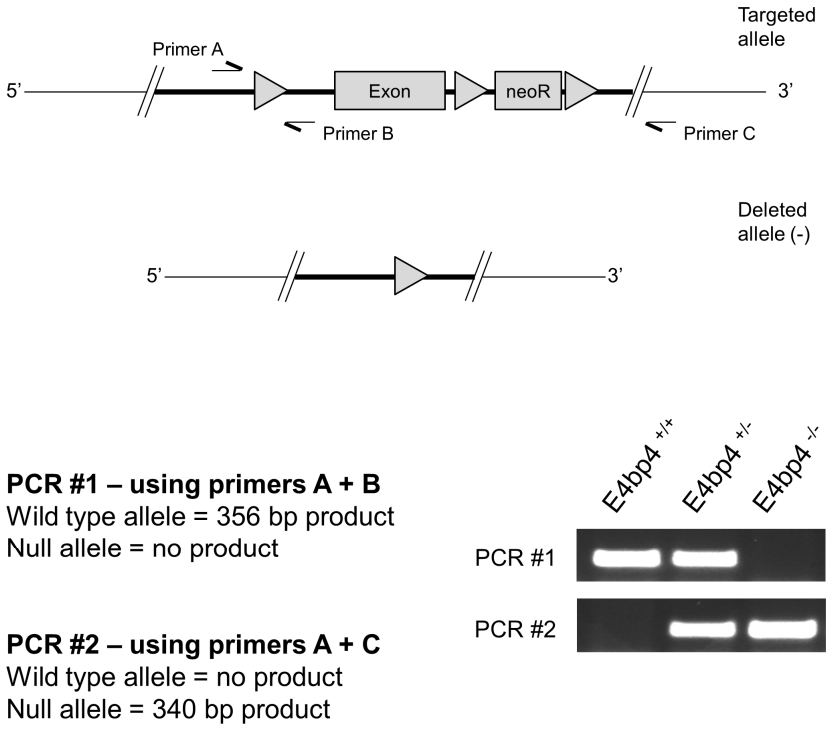


Figure 2.1 – Genotyping *E4bp4*$^{-/-}$ mice.

Genomic DNA was isolated *E4bp4*$^{-/-}$ mice tail or ear clips. Two PCR reactions were performed on each sample to determine the presence of the WT and KO allele.

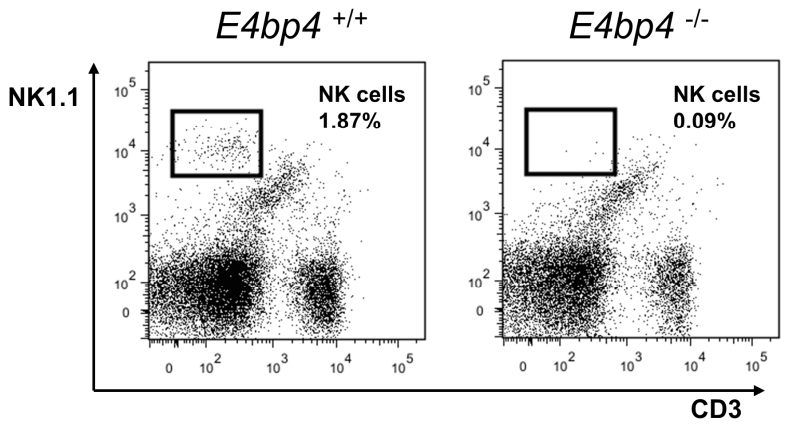


Figure 2.2 – Phenotype of *E4bp4*$^{-/-}$ mice.

Splenocytes from WT and *E4bp4*$^{-/-}$ mice were ACK lysed, before staining with antibodies against NK1.1 and CD3 to identify the presence of NK cells.

2.3 Molecular biology and cloning

2.3.1 DNA electrophoresis

DNA electrophoresis was performed using 1% agarose (Sigma), dissolved in TAE buffer and stained with 500 ng/ml Ethidium Bromide (Sigma). DNA samples were mixed with 5X DNA loading buffer (Bioline) and loaded into gel wells before electrophoresis was performed at 100 volts for approximately 60 minutes. Gels were imaged using an EC3 Imaging System (Ultra Violet Products Ltd).

2.3.2 Polymerase chain reaction

E4bp4 complementary DNA (cDNA) was amplified by Polymerase chain reaction (PCR) using specific sets of primers and Q5 High-Fidelity DNA Polymerase (New England Biolabs). Individual PCR reactions contained: 200 µM dNTPs, 0.5 µM forward primer, 0.5 µM reverse primer, approximately 100 ng template DNA and 1 U DNA polymerase. PCR reactions were subjects to thermal cycling: 98°C for 30 seconds (1 cycle); 98°C for 15 seconds, 55°C for 15 seconds, 72°C for 30 seconds (30 cycles); 72°C for 2 minutes (1 cycle); hold at 4°C. Following thermal cycling PCR products were purified using QIAprep PCR Purification Kit (Qiagen) and eluted in 30 µl H_2O.

2.3.3 Restriction enzyme digests

Restriction enzymes (Promega) were used to cleave double stranded DNA (dsDNA) and reactions were set up using manufacturer's instructions. Digests were incubated for 3 hours at 37°C before products were analysed on a 1% agarose gel. Bands of interest were excised from the gel using a scalpel and DNA was purified using the QIAquick Gel Extraction kit (Qiagen) and eluted in 30 µl H_2O.

2.3.4 Ligation and transformation of competent bacteria

Digested PCR products were ligated into the desired vector using T4 DNA ligase (Promega) using the manufacturer's instructions and an insert:vector ratio of 3:1. Ligations were performed at room temperature (RT) for 3 hours or at 16°C overnight (O/N). Ligations products were transformed into One Shot® TOP10 Chemically Competent *E. coli* cells (Life Technologies) to maintain the ligation product. 1 µl of ligation reaction was added to 25 µl of *E. coli* cells and incubated on ice for 30 minutes. Cells were heat shocked at 42°C for 40 seconds, incubated on ice for 2 minutes and then 250 µl SOC media was added to each reaction. Cells were allowed to recover at 37°C, shaking at 250 rpm, for 1 hour, before being plated on to LB agar plates containing the correct antibiotic to select for successful transformants. Bacterial cultures were selected with either 100 µg/ml Ampicillin (Sigma) or 50 µg/ml Kanamycin (Sigma).

2.3.5 Colony PCR

To determine successful cloning of PCR products, transformed colonies were screened by colony PCR. Individual transformed colonies were used to inoculate 10 µl of LB medium and 1 µl of this bacterial suspension was analysed in each PCR reaction. PCR reactions were performed using DreamTaq Green DNA polymerase (Thermo Scientific) and reactions were set up using manufacturer's instructions. PCR conditions were: 98°C - 5 mins; 1 cycle of 98°C – 1 min, 55°C – 1 min, 72°C – 2 mins; followed by 30 cycles of 98°C – 30 sec, 55°C – 35 sec, 72°C – 90 sec; and finally 72°C – 7 mins. Following thermal cycling, reactions were directly loaded onto 1% agarose gels for analysis.

2.3.6 Plasmid preparations and sequencing

For small scale plasmid preparations, 3 ml LB media (plus antibiotic) was inoculated with a single colony of transformed *E. coli* and incubated at 37°C, shaking, for approximately 16

hours. Plasmid DNA was purified using the QIAprep Spin Miniprep Kit (Qiagen) and DNA was eluted in 30 µl H_2O. Larger scale plasmid preparations used 100 ml LB media (plus antibiotic) and the plasmid DNA was purified using the QIAfilter Plasmid Midi Kit (Qiagen). Using this kit, DNA was eluted in 200 µl H_2O. The concentrations of DNA samples were determined using a Nanodrop ND-1000 and DNA samples were sent to Eurofins Genomics (Ebersberg, Germany) to be sequenced using a number of specific primers **(Table 2.1)**. Each sequencing reaction contained 1.5 µg template DNA and 2 pmol/µl primer.

Table 2.1 – Primers used for sequencing.

Name	Sequence (5' – 3')
T7	TAATACGACTCACTATAGGG
T3	ATTAACCCTCACTAAAGGGA
LMP5	CCCTTGAACCTCCTCGTTCGACC
LMP3	CACATTGCCAAAAGACGG
CSGW-seq-Fw	CAATCAGCCTGCTTCTCGCTTC
CSGW-seq-Rv	AAGAATACCAGTCAATCTTTCAC
GEM-seq-Fw	TAATACGACTCACTATAGGG
GEM-seq-Rv	CTTTCTAAATGCCCAGCTCTTACTCAG

2.3.7 Site directed mutagenesis

Single base pair mutations were made in the E4bp4 cDNA (K10R, K116R, K219R, K337R, K394R, S182A, S285A S286A, S301A, S352A and S353A), in pCMV-script expression vector, using the QuikChange® XL site-directed mutagenesis kit (Agilent). Specifically designed primers **(Table 2.2)** were used to incorporate specific single base mutations during temperature cycling. The *Dpn* I restriction enzyme was used to remove methylated (unmutated) DNA, before double stranded DNA (dsDNA) was transformed into XL10-Gold ultracompetent cells (Agilent). 2 µl of mutated DNA was added to 45 µl of competent cells and incubated on ice for 30 minutes. Cells were heat shocked at 42°C for 30 seconds, incubated on ice for 2 minutes and then 500 µl NZY+ media was added to each reaction. Cells were allowed to recover at 37°C, shaking at 250 rpm, for 1 hour, before being plated on to LB agar plates. Each mutated construct was sequenced to ensure that the mutation had been correctly incorporated. Mutations for double, triple, quadruple and quintuple mutants were made using the same method.

Table 2.2 – Primers used for site-directed mutagenesis.

Mutation	Forward primer (5' – 3')	Reverse primer (5' – 3')
K10R	AAAATGCAGACCATCAGAAAGGAGCCCGCACCC	GGGTGCGGGCTCCTTTCTGATGGTCTGCATTTT
K116R	GGGAGAAGAAAATGCCACTTTAAGAGCTGAGCTGCT	AGCAGCTCAGCTCTTAAAGTGGCATTTTCTTCTCCC
K219R	AACAAGTTCCCTGTGATCAGGCAGGAGCCCG	CGGGCTCCTGCCTGATCACAGGGAACTTGTT
K337R	GCCAAGGCCATGCAGGTCAGAGTGGAGGCT	AGCCTCCACTCTGACCTGCATGGCCTTGGC
K394R	TCAAGATTGGTCCCTCAGATCGGAACACTGGCATC	GATGCCAGTGTTCCGATCTGAGGGACCAATCTTGA
S182A	CAGTCATCAAGCACGCTCCCCAGAGCTCG	CGAGCTCTGGGGAGCGTGCTTGATGACTG
S210A	GGGAGGCTGCCGGGCCCCTGAGAACAAG	CTTGTTCTCAGGGGCCCGGCAGCCTCCC
S285A	GGTGTAGTGGGCAAGGCTTCTGATGGGGAAG	CTTCCCCATCAGAAGCCTTGCCCACTACACC
S286A	GTAGTGGGCAAGTCTGCTGATGGGGAAGACG	CGTCTTCCCCATCAGCAGACTTGCCCACTAC
S301A	TAAGGGCCCCATCCATGCTCCAGTGGAGC	GCTCCACTGGAGCATGGATGGGGCCCTTA
S353A	ATGCAGAAACTCTCTGCACCCGCCGATGCG	CGCATCGGCGGGTGCAGAGAGTTTCTGCAT
	(Mutated codons highlighted)	

2.3.8 Cloning of mutant E4bp4 cDNA into pMSCV-IRES-hCD2

WT E4bp4 cDNA and each mutant E4bp4 cDNA were cloned into the pMSCV-IRES-hCD2 retroviral expression vector (Gascoyne et al., 2009) (Male et al., 2014). Each cDNA was excised from the pCMV-script vector using NotI and XhoI restriction enzymes (Promega) and the pMSCV-IRES-hCD2 vector was digested in the same way. Each of the E4bp4 fragments was separately ligated into pMSCV-IRES-hCD2 and ligation products were transformed into competent *E. coli* cells.

2.3.9 Cloning of mutant E4bp4 cDNA into pCSGW

The WT E4bp4 cDNA and a number of mutant E4bp4 cDNAs were cloned into the pCSGW lentiviral expression vector. The E4bp4 cDNAs were amplified using specific forward and reverse primers, designed to incorporate BamHI and NotI digest sites at the 5' and 3' ends of the PCR product respectively:

pCSGW-E4bp4 forward primer: 5'-CTAGGATCCACCATGCAGCTGAGAAAAATG-3'
pCSGW-E4bp4 reverse primer: 5'-GTAGCGGCCGCGAGTCACAAGTCCTCTTC-3'

The PCR products and pCSGW vector were digested with BamHI and NotI restriction enzymes, purified, ligated together and transformed into competent *E. coli*. Successfully transformed colonies were tested for the presence of E4bp4 using colony PCR. Positive colonies were expanded and plasmid DNA was purified and sequenced.

2.3.10 Cloning of FLAG-E4bp4

E4bp4 cDNA (in the pCMV-script vector) was amplified by PCR using primers designed to incorporate: 5' NotI restriction digest site, 5' ATG start codon, 5' FLAG tag, 3' stop codon, and 3' XhoI restriction digest site **(Table 2.3)**. The PCR amplification was performed in two stages: 1) E4bp4 cDNA was amplified from pCMV vector template using primers FLAG-E4bp4-FW #1 and FLAG-E4bp4-RV; 2) The product from stage 1 was amplified further using primers FLAG-E4bp4-FW #2 and Flag-E4bp4-RV. The resulting product was purified, digested with NotI and XhoI restriction enzymes, ligated into the pCMV-script vector and the resulting ligation product was transformed into competent *E. coli*. Successfully transformed colonies were tested for the presence of FLAG-E4bp4 using colony PCR. Positive colonies were expanded and plasmid DNA was purified and sequenced.

Table 2.3 – Primers used for cloning FLAG-E4bp4.

Name	Primer sequence (5' – 3')
FLAG-E4bp4-FW #1	CAAGGATGACGACGATAAGCAGCTGAGAAAAATGCAGACCAT
FLAG-E4bp4-FW #2	ATTGCGGCCGCCACCATGGATTACAAGGATGACGACGATAAG
FLAG-E4bp4-RV	ATACTCGAGTCACCTGGAGTCCGAAGCCGAGATCGGT

Features:
E4bp4 cDNA FLAG tag Start/Stop codon NotI digest site XhoI digest site

2.4 SDS-PAGE and Western blotting

Cell lysates and protein samples were mixed 1:1 with Laemmli sample buffer and reduced by boiling in 5% β-mercaptoethanol for 5 minutes. Samples were separated on 8% polyacrylamide gels, submersed in SDS-PAGE running buffer, subjected to a constant 100 volts for approximately 90 minutes. To visualise the whole protein content of samples gels were stained with Imperial Protein Stain (Thermo Scientific) using manufacturer's instructions.

For Western blotting (WB), gels were not stained and samples were transferred from SDS-PAGE gels onto PVDF membranes using a wet transfer, run at constant 100 volts for 1 hour at 4˚C. Transferred membranes were blocked with blocking buffer for 1 hour at RT or at 4˚C O/N. Primary antibodies **(Table 2.4)** were used to probe blots for 1 hour at RT or O/N at 4˚C. All primary antibodies were used at the appropriate concentration and diluted in blocking buffer **(Table 2.4)**. All blots were washed 3 x 10 minutes in PBST buffer prior to incubation with the appropriate HRP-conjugated secondary antibody **(Table 2.4)**. Membranes were washed again for 3 x 10 minutes before chemiluminescence was detected using Western Lightning® Plus-ECL detection reagents (Perkin Elmer). Images of exposed blots were digitally acquired using the ChemiDoc™ XRS+ system (Bio-Rad).

2.5 Flow cytometry and cell sorting

Cells to be analysed by flow cytometry were washed in PBS buffer, centrifuged at 500 *g* for 5 minutes at 4°C and suspended in 200 µl FACS buffer. Cells were stained with the appropriate fluorochrome conjugated antibodies, all of which were used at a concentration of 1/100 **(Table 2.5)**. Cells were stained for 15 - 30 minutes, at 4°C, in the dark. Cells were washed with 2 ml FACS buffer, centrifuged again and suspended in 300 µl FACS buffer for analysis. Flow cytometry was performed using an LSRFortessa™ cell analyser (Becton Dickinson Bioscience) and off-line data analysis was performed using FlowJo software.

Cells for sorting were prepared in the same way and resuspended in FACS buffer at a concentration < 5 x 10^6 cells ml^{-1}. Cells were sorted using a FACSAria II (Becton Dickinson Bioscience) and positive cells were sorted in bulk into 15 ml falcon tubes containing appropriate growth media, supplemented with 20% FBS.

Table 2.4 – Antibodies used for Western blotting.

PRIMARY ANTIBODIES

Target	Origin	Clone	Source	Concentration
E4bp4	Goat	C18	Santa Cruz Biotech	1/1000
SUMO2/3	Rabbit	3742	AbCam	1/1000
FLAG	Mouse	M2	Sigma Aldrich	1/500
c-Myc	Mouse	9E10	Roche	1/1000
6X His	Mouse	4D11	AbCam	1/5000
Alpha-Tubulin	Mouse	DM1A	eBioscience	1/2000
RanGAP1	Rabbit	ab4784	AbCam	1/1000
Phospho-serine	Mouse	PSR-45	Sigma	1/500
Histone H3	Rabbit	ab1791	AbCam	1/10000
HA	Rabbit	ab9110	AbCam	1/1000
SUMO1	Rabbit	PW9460	Enzo Life Sciences	1/1000

HRP-CONJUGATED SECONDARY ANTIBODIES

Antibody	Origin	Clone	Source	Concentration
Anti-mouse	Rabbit	ab98790	AbCam	1/10000
Anti-rabbit	Goat	ab97080	AbCam	1/10000
Anti-goat	Donkey	ab7125	AbCam	1/10000

Table 2.5 – Fluorochrome conjugated antibodies.

Antigen	Clone	Conjugate	Origin	Isotype	Source
CD3	17A2	e450	Rat	IgG2b	eBioscience
CD19	eBio1D3	e450	Rat	IgG2a	eBioscience
hCD2	RPA-2.10	APC	Mouse	IgG1	eBioscience
NK1.1	PK136	PE	Mouse	IgG2a	eBioscience
CD11b	M1/70	PE	Rat	IgG2b	eBioscience
GR-1	RB6-8c5	PE	Rat	IgG2b	eBioscience
CD8	29A1.4	PerCP e710	Rat	IgG2a	eBioscience
B220	RA3-6B2	PE	Rat	IgG2a	eBioscience
Ter119	Ter119	PE	Rat	IgG2b	eBioscience
CD2	RM2-5	PE	Rat	IgG2b	eBioscience
CD11b	M1/70	PE-Cy7	Rat	IgG2b	eBioscience
CD4	GK1.5	FITC	Rat	IgG2b	eBioscience
CD127	A7R34	FITC	Rat	IgG2a	eBioscience
hCD56	CMSSB	APC	Mouse	IgG1	eBioscience
hCD45	H130	APC-eFluor® 780	Mouse	IgG1	eBioscience

2.6 Analysis of SUMOylation *in vivo*

Protocol adapted from (Tatham et al., 2009). HeLa cell lines 6His-SUMO-1, 6His-SUMO-2, 6His-SUMO-3 and the parental HeLa cells were transfected with pCMV-E4BP4 or pCMV (empty vector). After transfection each plate of cells was resuspended, washed with PBS and 20% of the cell suspension was removed as an input sample. The remaining cells were lysed in 5 ml SUMO cell lysis buffer containing 6 M Guanidinium-HCl, whilst the input sample cells were lysed in 100 µl 2x Laemmli sample buffer and analysed by WB for the expression of E4bp4 and 6His-SUMO.

Ni^{2+} affinity purification was performed on the larger cell lysates to enrich for all of the SUMO modified proteins. 50 µl of Ni^{2+} NTA agarose beads (Qiagen) were incubated with each cell lysate O/N at 4˚C. The following day, the unbound material was discarded before the agarose beads were washed extensively with SUMO cell lysis buffer, pH 8.0 wash buffer

and three times with pH 6.3 wash buffer. His-tagged proteins were eluted from agarose beads with 50 µl elution buffer, containing 200 mM imidazole. The samples were boiled for 5 minutes to ensure all material had been eluted and samples were analysed by WB for the presence of E4bp4, 6His-SUMO and RanGAP1 (positive control for the purification of SUMO-modified proteins).

To analyse the effect of Listerolysin O (LLO) on the SUMOylation of E4bp4, transfected 6His-SUMO2/3 cells were transferred into serum-free media 2 hours before being treated with 3 nM purified LLO (gift from David Ribet, Institute Pasteur, France) for 20 minutes at RT (Ribet et al., 2010). The purified LLO was added directly to the culture medium. Following LLO treatment cells were resuspended in PBS and lysed as described above.

2.7 Immunoprecipitation of FLAG-E4bp4

HeLa cell lines 6His-SUMO1, 6His-SUMO2 and 6His-SUMO3 were transfected with pCMV-FLAG-E4bp4 or pCMV (empty vector). After transfection, cells were lysed using a two-step lysis protocol (Klenova et al., 2002). Cells were lysed in two-step lysis buffer (containing protease inhibitors, phosphatase inhibitors and 25 mM *N*-Ethylmaleimide to inhibit SUMO proteases), incubated on ice for 15 minutes, sonicated for 30 seconds, NaCl was added to each sample, at a final concentration of 420 mM and then each lysate was incubated for 60 minutes rotating at 4°C. Following this incubation the lysates were sonicated again, centrifuged for 30 minutes at 10,000 *g* and the supernatant collected.

To efficiently immunoprecipitate (IP) FLAG-E4bp4, cleared lysates were diluted 1:1 with low salt TBS and incubated with Anti-FLAG® M2 Affinity Gel (Sigma Aldrich) O/N at 4˚C. Samples were centrifuged at 5000 *g* for 30 seconds and the supernatant was removed (flow through sample). Each sample was washed three times with low salt TBS and three times with high salt TBS to remove any non-specially bound proteins. Laemmli sample buffer was boiled with each sample for 5 minutes to elute purified material.

2.8 Optimisation of FLAG-E4bp4 expression and purification

To produce large quantities of E4bp4 for biochemical analysis, different cell lines were transfected to determine their ability to produce recombinant E4bp4. One 100 mm dish of 3T3, Cos-7, LinXE and 293T cells was transfected with FLAG-E4bp4 using the Lipofectamine transfection reagent. 48 hours after transfection, 2 x 10^6 cells from each plate were lysed in 100 µl Laemmli sample buffer and 25 µl of each was analysed by WB for the expression of E4bp4.

To increase expression of E4bp4 in 293T cells, each plate of cells was also transfected with a plasmid expressing VP35. Each 100 mm dish of cells was transfected with a total of 8 µg DNA, but with varying amounts of VP35. One plate of cells was transfected with: 0, 1, 2, 3 and 4 µg of VP35 plasmid DNA and 8, 7, 6, 5 and 4 µg of pCMV-FLAG-E4bp4 DNA. Forty-eight hours after transfection, 2 x 10^6 cells from each plate were lysed in 100 µl Laemmli sample buffer and 25 µl of each was analysed by WB for the expression of E4bp4. WBs were analysed using ImageLab software (Biorad) to quantify intensity of bands on blots.

FLAG-E4bp4 was transfected into 293T cells, with 5 µg pCMV-FLAG-E4bp4 and 3 µg VP35 per 100 mm tissue culture dish, and after 48 hours cells were lysed using the two-step lysis protocol (section 2.7). Cleared lysates were diluted 1:1 with low salt TBS and incubated with Anti-FLAG® M2 Affinity Gel O/N at 4˚C. Samples were centrifuged at 5000 *g* for 30 seconds and the supernatant was removed (flow through sample). Each sample was washed three times with low salt TBS and three times with high salt TBS to remove any non-specially bound proteins. Initially, a number of elution techniques were tested including: Laemmli sample buffer; 0.1 M glycine pH 2.5; 0.1 M glycine pH 10; 1 M Arg-HCl pH 2.5; 0.5 M Arg-HCl pH 2.5; 150 ng/µl FLAG peptide and 300 ng/µl FLAG peptide.

After initial optimisation all elutions were carried out using 150 ng/µl FLAG peptide in high salt TBS. Four elution fractions were taken from each resin and when elution buffer was removed, fresh buffer wash added to replace it. Each elution fraction was incubated with the resin for 4 - 16 hours, at 4°C, on a rotating wheel. For each sample all elution fractions were combined and concentrated using Amicon Ultra 0.5 ml centrifuge filters with a 30 kDa molecular weight cut off (Millipore). Purified samples were stored at -80°C, but 10% of each was always removed and mixed 1:1 with Laemmli sample buffer for analysis by WB and coomassie stained SDS-PAGE gel.

2.9 Mass spectrometry analysis

FLAG-E4bp4 was transfected into 10 x 100 mm tissue culture dishes of HEK-293T cells and 48 hours after transfection cells were lysed (using the two-step lysis protocol - section 2.7) and FLAG-E4BP4 was purified separately from each lysate. The eluted samples were combined was concentrated using Amicon Ultra 0.5 ml centrifuge filters (Millipore). 10% of the final purified product was removed for analysis by WB and coomassie stained SDS-PAGE gel, whilst the remaining was analysed by mass spectrometry (MS).

Purified FLAG-E4bp4 was concentrated using vacuum centrifugation and resuspended in 100 mM ammonium bicarbonate (Sigma Aldrich) pH 8. Reduction of disulfide bonds was carried out using a final concentration of 5 mM dithiothreitol (Fisher Scientific) for 1 hour at 51°C and iodoacetamide (Sigma-Aldrich) was then added at a final concentration of 14 mM and allowed to react in the dark for 45 mins to label the reduced cysteines (Villen and Gygi, 2008). Proteomics-grade trypsin (Promega) was added at a 1:20 w/w ratio to digest the E4bp4 protein for 6 hours at 37 °C. For SUMOylated peptide analysis the sample was sequentially digested with endoproteinase GluC (Roche), which was added at a 1:10 w/w ratio and incubated for 6 hours at RT. After quenching the digestions by freezing to -80°C the peptides were dried by vacuum centrifugation. The SUMOylated peptides were desalted with ZipTips (EMD Millipore) (Rappsilber et al., 2003) and reconstituted into the correct loading solvent for nano LC–MS whilst phosphopeptides were processed further.

For phosphopeptide analysis, phosphorylated peptides were enriched using titanium dioxide (GL sciences) as previously described (Li et al., 2009). Titanium dioxide beads were first dissolved at a ratio of 1:50 (w/v) into a loading buffer comprised of saturated glutamic acid in H_2O, 65% acetonitrile (MeCN), 2% trifluoroactic acid (TFA) (v/v/v) and incubated at RT on a shaker for 10 mins. Tryptic peptides were resuspended in loading buffer and then added to the titanium dioxide slurry at a ratio of 1:4 (w/w) using the original protein weight, and incubated at RT on a shaker for 30 mins. The slurry was spun down to pellet the titanium dioxide bound phosphopeptides and the supernate discarded. To wash the phosphopeptides, the pellet was resuspended in loading buffer and incubated at RT on a shaker for 15 mins, spun down and the supernate discarded. This step was repeated using 34.5% H_2O, 65% MeCN and 0.5 TFA (v/v/v), and followed by a repetition with 34.9% H_2O, 65% MeCN and 0.1 TFA (v/v/v). To elute the phosphopeptides, the pellet was resuspended in 50% 300 mM ammonium hydroxide (NH_4OH) in H_2O and 50% MeCN (v/v) and incubated at RT on a shaker for 15 mins, spun down and the supernate collected. This step was repeated using 50% 500 mM NH_4OH and 50% MeCN (v/v) pooling the supernate each time

for a further 2 times. The eluted phosphopeptides were then desalted with ZipTips (Rappsilber et al., 2003) and reconstituted into the correct loading solvent for nano LC–MS.

Peptides were chromatographically resolved on an Ultimate 3000 RS-LC-nano System (Dionex), with an Acclaim PepMap100, C18 stationary phase, 2 µm particle size, 100 Å pore size, 75 µm internal diameter x 15 cm length column (Thermo Fisher). The LC gradient conditions comprised of a flow rate of 0.3 µL/min and a gradient starting at 1% B (5% H_2O, 95% MeCN, and 0.1% formic acid) and 99% A (0.1% formic acid and 100% H_2O) and increased to 50% B over 80 min followed by an increase to 95% B over 25 mins. Real-time tandem mass spectra were acquired on an LTQ Velos Pro linear ion trap (Thermo Scientific) with a 120 min acquisition time. Phosphopeptides were analysed using a 375–2000 m/z scan range and CID fragmentation collision energy of 35%. Tandem mass spectra were collected using top 6 data-dependent acquisition, with a dynamic exclusion list (repeat count of 2, repeat duration of 30 seconds, exclusion list size 100, and exclusion duration of 60 seconds). Initial phosphopeptide identification from the nanoLC–MS/MS data was performed using a Sequest search in Proteome discoverer 1.3 (Thermo Fisher) against the Mouse Uniprot database. Reverse decoy false-discovery-rate values of 0.01/0.05 were used (strict/relaxed), allowing for 4 missed cleavages and a 2 Da precursor mass tolerance. Dynamic side-chain modifications included in the search were carbamidomethyl (+57.021; from iodoacetamide treatment), phosphorylation (+79.966) on serine, threonine, and tyrosine residues, and oxidation (+15.995) on methionine residues. Putative phosphopeptides were then validated using an algorithm for phosphorylation site identification (Wu et al., 2013) and manually assessed for accuracy.

The SUMOylated peptides were analysed using a mixture of data-dependent acquisitions and targeted MS/MS scans to search for peptides containing putative sites of SUMOylation based on a SUMO tag of GGTQQFV. Specifically, targeted MS/MS scans were always collected for the following m/z values, which correspond to E4bp4 peptides predicted to have SUMOylated lysines: K10 m/z^{2+} = 531.99; K116 m/z^{2+} = 796.49; K219 m/z^{2+} = 853.54; K337 m/z^{2+} = 825.52). De novo analysis of a tandem MS corresponding to K219 SUMO (m/z^{2+} = 853.54) revealed fragment ion peaks to support this modification. To validate the observed fragmentation, a synthetic version of this peptide containing the SUMOylated tag on K219 was purchased: FPVIK(GGTQQQFV)QE (Anaspec, Eurogenetec). The retention time and MS/MS fragment ions observed for the peptide standard were found to match the experimental tandem MS, thus confirming the identification of K219 SUMOylation on E4bp4.

2.10 *In vitro* development of NK cells from transduced lineage negative bone marrow cells

To isolate lineage negative (Lin^-) bone marrow (BM) cells, mouse leg bones were crushed in PBS + 2% FBS to release bone marrow cells. These cells were resuspended in MACS buffer and stained with a phycoerythrin (PE)-conjugated lineage cocktail of antibodies (α-B220, α-CD2, α-Ter119, α-NK1.1, α-Mac-1 and α-GR1 **(Table 2.5)**) for 5 minutes at 4°C. Stained cells were incubated with anti-PE Microbeads (Miltenyi) for 15 minutes at 4°C to allow negative selection of lineage negative cells. The cells and cell-bead conjugates were washed with MACS buffer before being applied to an LS column (Miltenyi) attached to a MidiMACS™ magnet (Militenyi). The unstained Lin^- BM cells flowed directly through the column and were retained for culture in complete cytokine medium at a concentration of 5 x 10^5 cells ml^{-1}. Lin^- BM cells were cultured for 48 hours before retroviral transduction.

LinXE cells were transfected with pMSCV-E4BP4-hCD2 (or E4bp4 mutant versions, or empty vector) and after 48 hours the supernatant was collected from each plate of cells. Virus particles were concentrated by centrifugation at 18,000 *g* for 45 minutes at 4°C; resuspended in complete cytokine media and mixed with cultured Lin^- BM cells, in the presence of 8 µg/ml polybrene, at a concentration of 10^5 cells ml^{-1}. The cell-virus suspension was aliquoted into 96-well plates (100 µl per well) and centrifuged for 45 minutes at 700 *g* and RT. Following centrifugation, 100 µl of complete cytokine media was added to each well and the cells were cultured for 72 hours.

Transduced Lin^- BM cells were resuspended in IL-15 medium at a concentration of 3 x 10^4 cells ml^{-1} and distributed into wells of a 24-well plate containing cultured OP9-GFP stromal cells. The transduced Lin^- BM cells were cultured on the OP9-GFP stroma for 7 days, with a medium change on day 4. After 7 days of culture, the cells were washed with PBS and stained with antibodies against NK1.1, CD19, CD11b and hCD2 **(Table 2.5)** for analysis by flow cytometry.

Percentages of NK cells in the transduced populations were compared. For *E4bp4*$^{-/-}$ experiments, each sample was normalised to WT E4bp4 and combined to give an overall mean ± S.E.M. (n = 3). For experiments using WT Lin^- BM, each sample was normalised to the empty vector, before a mean was calculated. Results were verified to be significant using the Student's paired T-test (Graphpad Prism), where values of $p < 0.05$ were considered as statistically significant.

To analyse the effects of kinase inhibitors on NK cell development WT Lin^- BM cells were isolated (by negative selection with anti-PE microbeads) and cultured in complete cytokine medium for 5 days before being transferred into 24-well plate containing cultured OP9-GFP stromal cells. Cells were cultured for 7 days in IL-15 medium, in the presence or absence of 30 ng/ml IL-15, 10 μM 4-(4-(2,3-Dihydrobenzodioxin-6-yl)-5-pyridin-2-yl-1H-imidazol-2-yl)benzamide (D4476) (Sigma Aldrich) and 20 μM 4,5,6,7-Tetrabromobenzotriazole (TBB) (Sigma Aldrich). The media was refreshed on day 4 and after 7 days of culture, the cells were washed with PBS and stained with antibodies against NK1.1 and CD19 and analysed by flow cytometry.

2.11 Cycloheximide time course analysis of protein stability

2.11.1 Creation of stably transduced cell lines

LinXE cells were transfected with pMSCV-E4bp4-hCD2 (or E4bp4 mutant versions) and after 48 hours the supernatant was collected from each plate of cells. Virus particles were concentrated by centrifugation at 18,000 *g* for 45 minutes at 4°C; resuspended in 4 ml 3T3 culture medium and then added to 3T3 cells growing in 100 mm dishes. Each plate of 3T3 cells was incubated with concentrated virus for 4 hours in the presence of 8 μg/ml polybrene. Fresh 3T3 culture media was added to each plate of cells before they were cultured for 48 hours. Transduced cells were analysed for surface hCD2 expression using flow cytometry to determine the transduction efficiency. Transduced 3T3 cells were cultured for a further 14 days (with a passage every 2-3 days), before the cells were resuspended in FACS buffer, stained (α-hCD2-APC) and the hCD2 positive population was bulk sorted for each cell line. Sorted cells were cultured for a maximum of 10 passages and hCD2 expression was regularly checked by flow cytometry.

2.11.2 Cycloheximide time course

Stably transduced 3T3 cell lines were incubated with 50 μg/ml cycloheximide (Sigma Aldrich) for 0, 2, 4, 6, 8, 12 and 16 hours. At the end of the time course each sample (~5 x 10^6 cells) was suspended in 400 μl CE buffer and lysed by passing repeatedly through a fine gauge needle and incubated for 30 minutes on ice. Samples were centrifuged for 15 minutes at 3500 *g* (4°C) and the supernatant removed (cytoplasmic extract). The remaining nuclei were suspended in 200 μl NE buffer, sonicated for 15 seconds and incubated at 4°C, rotating, for 1 hours. Samples were centrifuged for 20 minutes at 20000 *g* (4°C) to remove any insoluble material and supernatant maintained (nuclear extract). Nuclear extracts were

analysed by WB for the presence of E4bp4 and Histone H3 was used a loading control to compare samples. WBs were analysed using ImageLab software (Biorad) to quantify intensity of bands on blots.

2.12 Luciferase assay

E4bp4 transcriptional activity was analysing using numerous pGL3 vectors (containing the *Firefly luciferase* reporter gene) and the Dual-Luciferase® reporter assay system. The vectors used were: pGL3-E4bp4-CS (previously created by Yeung et al., 2004), which contained the E4bp4 consensus sequence identified by Cowell *et al*. (1992); pGL3-E4bp4-CS-ZH, which contained the E4bp4 consensus sequence identified by Zhang *et al*. (1995); pGL3-5'-Eomes and pGL3-5'-Id2, which both contained the 2 kb regions upstream of these genes from the mouse genome (created by Dr Victoria Male, Brady Laboratory).

3T3 cells were cultured in 24-well plates and each well was transfected with 200 ng pMSCV-E4BP4-hCD2 (or mutant versions, or empty vector), 200 ng pGL3-E4BP4 (or empty vector) and 50 ng pRL-CMV. After 48 hours, cells were lysed with passive lysis buffer and the Dual-Luciferase® reporter assay system (Promega) was used to determine the activities of Firefly and Renilla Luciferases. Cell lysates were combined with LAR II solution and Stop & Glo® reagent in 96-well plates and the luminescence produced by each sample was recorded by a FLUOstar Omega plate reader (BMG Labtech). The plate reader was set to analyse each well for 8 seconds, with a 2 second pre-measurement delay between samples. Using Renilla luciferase as an internal control Relative Luciferase activities (RLAs) were calculated for each sample. RLAs from independent experiments were combined by expressing individual samples relative to empty vector to give an overall mean ± S.E.M. Results were verified to be significant using the Student's unpaired T-test (Graphpad Prism), where values of $p < 0.05$ were considered as statistically significant.

2.13 Regulation of E4bp4 target genes in NK cells

2.13.1 Phenotyping of MNK-1 cells

MNK-1 cells were cultured in the presence or absence of 10 µg/ml IL-2 for 36 hours and at 4 hour time points the cell viability was analysed. Cells were stained with trypan blue and the percentage of live/dead cells was determined using a haemocytometer. MNK-1 cells were also analysed for their expression of numerous cell surface markers. Cells were washed with PBS, suspended in 200 µl FACS buffer and stained separately with antibodies against NK1.1, CD8, CD3, CD127, CD19, CD11b, CD122 and CD4 **(Table 2.5)**.

2.13.2 Lentiviral transduction of MNK-1 cells

HEK-293T cells were transfected with pCSGW-E4bp4 (or E4bp4 mutants), pSPAX2 and pMD2.G (each 100 mm dish: 6 µg pCSGW, 6 µg pSPAX2, 2.5 µg pMD2.G) and after 48 hours the supernatant was collected from each plate of cells. Virus particles were concentrated by centrifugation at 18,000 *g* for 45 minutes at 4°C, suspended in MNK-1 growth media and mixed with MNK-1 cells. For each condition two plates of 293T cells were transfected and virus particles were used to transduce 1.5 x 10^6 cells in 2 ml culture media, in the presence of 8 µg/ml polybrene. The virus/cell mix was aliquoted into 96-well plates and centrifuged at 700 *g* for 45 minutes, at RT. After centrifugation the virus/cells were transferred to 100 mm tissue culture dishes to be cultured for 48 hours.

2.13.3 Quantitative PCR

After transduction, MNK-1 cells were lysed and RNA was isolated for transcriptional analysis using the RNeasy mini kit (Qiagen). RNA from each sample was eluted in 50 µl RNase-free H_2O and was converted to cDNA using 1st Strand transcriptor cDNA synthesis kit (Roche). Each cDNA synthesis reaction contained 2 µg total RNA and 2.5 µM anchored-oligo$(dT)_{18}$ primer, which were initially denatured at 65ºC for 10 minutes. Subsequently, 5X reaction buffer, 20 U RNase inhibitor, 1 mM deoxynueclotide mix and 10 U Transcriptor Reverse Transcriptase were added to each reaction. Reactions were then incubated at 55ºC for 30 minutes, 85ºC for 5 minutes, before being quenched on ice for 15 minutes.

Quantitative PCR (QPCR) reactions were performed using TaqMan primer/probes **(Table 2.5)** to determine gene expression levels of potential E4bp4 target genes. Each cDNA sample was analysed in technical duplicate, with each individual reaction prepared containing 5 µl 2X TaqMan Fast Advanced Master Mix (Life Technologies), 0.5 µl TaqMan Gene Expression Assay **(Table 2.5)**, 2.5 µl RNase-free H_2O and 2 µl cDNA sample.

Samples were analysing using an Applied Biosystems 7500 Fast Real-Time PCR system with the following thermal cycling profile: 50ºC for two minutes, 95ºC for 20 seconds and then 45 cycles of 95ºC for 3 seconds and 60ºC for 30 seconds. Ct values from samples were compared with a standard curve made from a known concentration of plasmid DNA (Eomes, Id2, T-bet, Gata3, Gfi1) or murine splenocyte genomic DNA (Notch, E2A, EBF, HPRT1).

Three independent experiments were performed analysing the expression of all genes and each sample was normalised to HPRT1, before a mean ± S.E.M. was calculated for each gene. Results were verified to be significant using the Student's paired T-test (Graphpad Prism), where values of $p < 0.05$ were considered as statistically significant.

Table 2.6 – TaqMan gene expression probes used for QPCR.

Gene target	Entrez Gene ID	Probe number	Amplicon length (bp)
Eomes	13813	Mm01351985_m1	58
Id2	15902	Mm00711781_m1	152
T-bet	57765	Mm00450960_m1	69
Gata3	14462	Mm00484683_m1	57
Gfi1	14581	Mm00515855_m1	70
Notch1	18128	Mm00435249_m1	76
TCF3 (E2A)	69714	Mm00469848_m1	81
EBF2	13592	Mm00438625_m1	66
HPRT	15452	Mm00446968_m1	65

2.14 *In vitro* development of NK cells from transduced human umbilical cord blood stem cells

$CD34^+$ umbilical cord blood (UCB) stem cells were kindly provided by Aurore Saudemont (Anthony Nolan Research Institute, and University College London). These cells had been previously isolated from whole cord blood by ficoll gradient centrifugation, to isolate all mononuclear cord blood cells and by MACS column magnetic separation to enrich for $CD34^+$ cells. Isolated cells were cryopreserved in liquid nitrogen for long term storage and transport. $CD34^+$ cells were cultured using protocol previously described by Luevano *et al*. (Luevano et

al., 2014). Cells were thawed and resuspended in NK cell differentiation media at a concentration of 2.5 x 10^3 cell ml^{-1}. $CD34^+$ cells were cultured on irradiated E108.1D2 feeder cells for 48 hours before lentiviral transduction.

HEK-293T cells were transfected with pCSGW-IRES-td-tomato (created by Dr Ilaria Nisoli, Brady laboratory), pSPAX2 and pMD2.G (each 100 mm dish: 6 µg pCSGW, 6 µg pSPAX2, 2.5 µg pMD2.G) and after 48 hours the supernatant was collected from each plate of cells. Virus particles were concentrated by centrifugation at 18,000 *g* for 45 minutes at 4°C, suspended in NK cell differentiation media and mixed with $CD34^+$ cells. Each plate of virus is used to transduce 1.25 x 10^5 cells in 5 ml of culture media in the presence of 8 µg/ml polybrene. The virus/cell mix was aliquoted into 96-well plates and centrifuged at 700 *g* for 45 minutes, at RT. After centrifugation the virus/cells were transferred back onto E108.1D2 feeder cells and cultured for 3 days before all transduced td-tomato positive cells were FACS sorted in bulk. Cells were cultured for a further 9 days (total 14 days in culture) with a half media change every 3 days. Following the culture period, cells were washed with PBS and stained with antibodies against CD56 and CD45 **(Table 2.5)** for analysis by flow cytometry to determine NK cell development.

CHAPTER 3

Regulation of E4bp4 by SUMOylation

3.1 Introduction

3.1.1 Regulation of E4bp4

The transcription factor (TF) E4bp4 is an essential regulator of many biological processes, including the circadian rhythm, the development of CD8α cDC's and cytokine production by T_H2 cells (Male et al., 2012). Most prominently, it is absolutely required for the production of NK cells and commitment to the NK lineage during lymphocyte development (Male et al., 2014). Despite its importance, very little is known about how the actions of E4bp4 are regulated, particular inside NK cells. PTMs are a key mechanism for regulating TF activity, as they can rapidly and reversible alter how a TF recognises DNA or interacts with a partner protein and they allow external signals to directly influence gene expression (Benayoun and Veitia, 2009). However, little is known about the PTMs of E4bp4 and whether these regulate its function. Using a yeast-two-hybrid screen, it was discovered that E4bp4 can interact with numerous SUMO E3 ligases, in particular PIAS1 **(Table 1.1)**. From this data it was postulated that E4bp4 may be post translationally SUMOylated, but it was not known whether this modification was required for its function.

3.1.2 SUMOylation

Proteins of the small ubiquitin-like modifier (SUMO) family are reversible post-translational protein modifiers that regulate diverse cellular processes including DNA repair, chromosome segregation and transcription (Gareau and Lima, 2010). SUMO proteins are ~11 kDa in size and have a very similar three-dimensional structure to ubiquitin **(Figure 3.1)**. In mammals there are four SUMO isoforms, designated SUMO1 to SUMO4. The mature forms of SUMO2 and SUMO3 only differ from each other by three N-terminal amino acids and are therefore commonly referred to as SUMO2/3 (Wilkinson and Henley, 2010). SUMO1 shares approximately 50% sequence identify with SUMO2/3 **(Figure 3.1)** and serves distinct functions as it often conjugates to different target proteins *in vivo* (Vertegaal et al., 2006) The SUMO4 isoform has 87% similarity to SUMO2, but its conjugation to target proteins has only been reported in conditions of extreme stress (Wei et al., 2008). SUMO4 has a critical proline residue that prevents efficient maturation and conjugation to target proteins and is therefore not thought to be functional in mammalian cells (Owerbach et al., 2005).

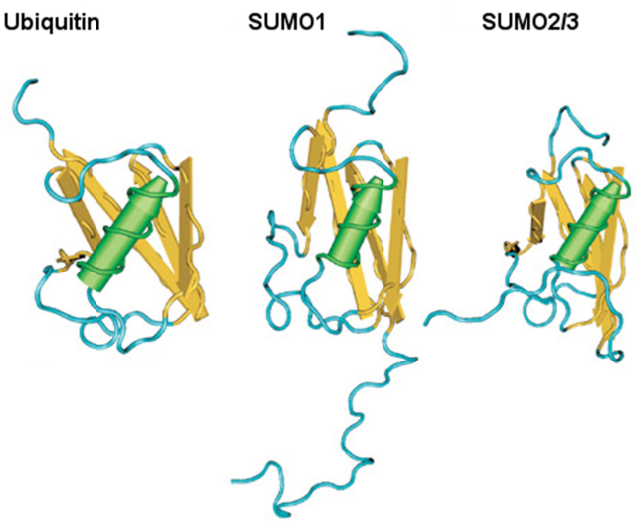


Figure 3.1 – Three-dimensional structure of SUMO isoforms

Representative NMR ribbon diagrams comparing the three-dimensional structures of ubiquitin, SUMO1 and SUMO2/3. The structure of SUMO2 and SUMO3 are identical as they only differ by three residues and are therefore referred to as SUMO2/3. There is low sequence similarity between ubiquitin and SUMO isoforms but they have very comparable three-dimensional structures. Adapted from (Martin et al., 2007).

3.1.3 SUMOylation of target proteins

SUMO proteins are covalently attached to lysine residues of target proteins, by a process analogous to, but distinct from, ubiquitination (Anckar and Sistonen, 2007). During this process, an isopeptide bond is formed between the ε-amino group of the lysine residue and the C-terminal Gly of the SUMO protein (Geiss-Friedlander and Melchior, 2007). The attachment of SUMO requires an enzymatic cascade that begins with the cleavage of immature inactive SUMO precursors to expose a di-glycine motif **(Figure 3.2)**. Mature SUMO proteins are activated in an ATP-dependent manner by the E1 activating enzyme (a heterodimer of SAE1 and UBA2) and a thioester bond forms between the C-terminal carboxyl group of SUMO and the catalytic Cys residue of UBA2 **(Figure 3.2)**. SUMO is then transferred to the E2 conjugating enzyme, Ubc9, via a thioester linkage. The E2 enzyme can interact with SUMO protein substrates and transfer the SUMO to an accept lysine residue **(Figure 3.2)**. This process is usually facilitated by a SUMO E3 ligase, which promotes the transfer of SUMO from Ubc9 onto the target substrate **(Figure 3.2)**. Importantly, SUMO conjugation is a highly reversible process, due to the action of SENP (sentrin/SUMO-specific protease) enzymes, which cleave the isopeptide bond, releasing the SUMO protein to be conjugated to another target **(Figure 3.2)**.

One characteristic feature of SUMO conjugation is that target lysine residues are commonly found within the consensus sequence Ψ-K-*x*-E (Ψ denoting a hydrophobic residue and *x*

denoting any amino acid). These residues directly interact with the Ubc9 enzyme, allowing it to select accurately which Lys residues to modify (Anckar and Sistonen, 2007). However, some alternative SUMO modification sites have also been identified, for example Lys164 in PCNA (Hoege et al., 2002) and Lys14 in the ubiquitin E2 enzyme E2-25K (Pichler et al., 2005). SUMO1 conjugates to E2-25K at Lys14 due to the α-helical secondary structure in the surrounding protein and prevents ubiquitin thioester formation (Pichler et al., 2005). Whether these non-consensus sites are rare exceptions or a common occurrence remains to be confirmed, as is an understanding of how the Ubc9 enzyme can recognise and interact with these sites.

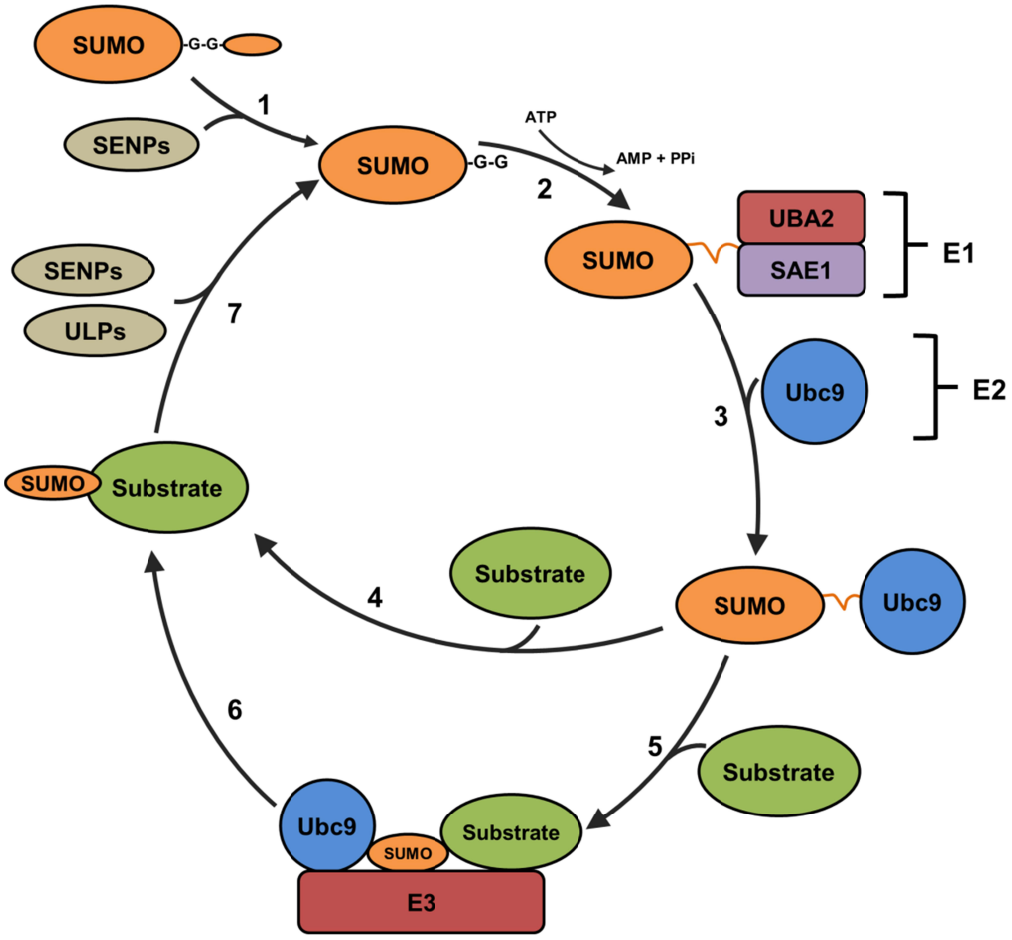


Figure 3.2 – The SUMO conjugation cycle

Schematic representation of the cyclic process that results in the conjugation and deconjugation of SUMO onto substrate proteins. SUMO proteins are processed in the cell by sentrin-specific proteases (SENPs) to reveal a C-terminal –GG motif (step 1). Mature SUMO proteins are sequestered by the heterodimeric E1 enzyme (step 2), which transfers them to the SUMO E2 enzyme (step 3). SUMO2 E2 can directly interact with the substrate protein to transfer the SUMO conjugate to a target lysine residue (step 4), but this process is usually facilitated by an E3 SUMO ligase (step 5). Either way, a SUMO-conjugated form of the protein is freely released into the cell (step 6). SUMO conjugates are subsequently removed from proteins by the action of proteases (step 7) including ubiquitin-like-protein-specific proteases (ULPs) and SENPs. Adapted from (Gareau and Lima, 2010).

Some substrates are modified by a singular SUMO isoform (e.g. the nuclear dot protein Sp100 is modified by SUMO2/3 only), whilst many are modified by both SUMO1 and SUMO2/3 (Vertegaal et al., 2006). However, it is unclear how the SUMOylation machinery can distinguish between SUMO1 and SUMO2/3 for conjugation onto target proteins and how modifications with the different isoforms will impact on the function of a particular substrate protein (Wilkinson and Henley, 2010). SUMO2 and SUMO3 are also able to form chains on target proteins, due to an internal Ψ-K-*x*-E acceptor site (Knipscheer et al., 2007). This characteristic is very similar to that observed for ubiquitin, but the extent of polySUMO modification in the proteome has yet to be established (Golebiowski et al., 2009). The significance of these SUMO chains and whether they have specific biological functions also is not understood (Bekes et al., 2011).

3.1.4 Effect of SUMOylation

The first reported targets of SUMO modification were the GTPase-activating protein 1 (RanGAP1) (Mahajan et al., 1997) and the promyelocytic leukaemia protein (PML) (Muller et al., 1998), but since then SUMOylation has been found to influence proteins involved in almost every aspect of cellular activity. SUMOylation has been found to be important for regulating TF activity, nuclear transport, chromosome segregation, DNA repair and mitochondrial dynamics (Geiss-Friedlander and Melchior, 2007).

There is no way to predict how SUMOylation will affect the function of a particular protein, as any aspects of a target protein could be affected. However, on a molecular level, generally there are three consequences of SUMOylation **(Figure 3.3)**. SUMOylation can block the binding site of protein, preventing interactions with a partner protein, for example, E2-25k cannot interact with its partner ubiquitin E1 enzyme when it is modified (Pichler et al., 2005). Conversely, SUMO modifications may help to promote new interactions with other proteins, either by direct non-covalent interaction or via a novel interaction domain created at the interface between the SUMO and substrate **(Figure 3.3)**. SUMOylation of RanGAP1 promotes its interaction with RanBP2 and its relocation to the nuclear pore (Mahajan et al., 1997). Thirdly, SUMOylation can induce a conformational change in the SUMOylated target protein, directly affecting its function **(Figure 3.3)**. A good example of this is the DNA modifying enzyme TDG (thymine-DNA glycosylase), which recognises and removes mismatched base pairs (Baba et al., 2005) (Hardeland et al., 2002). Modification by SUMO1 induces a conformational change in TDG, decreasing its affinity for its target DNA allowing

the TDG to release into the nucleoplasm and be recycled to go and find more target DNA sequences (Baba et al., 2005) (Hardeland et al., 2002).

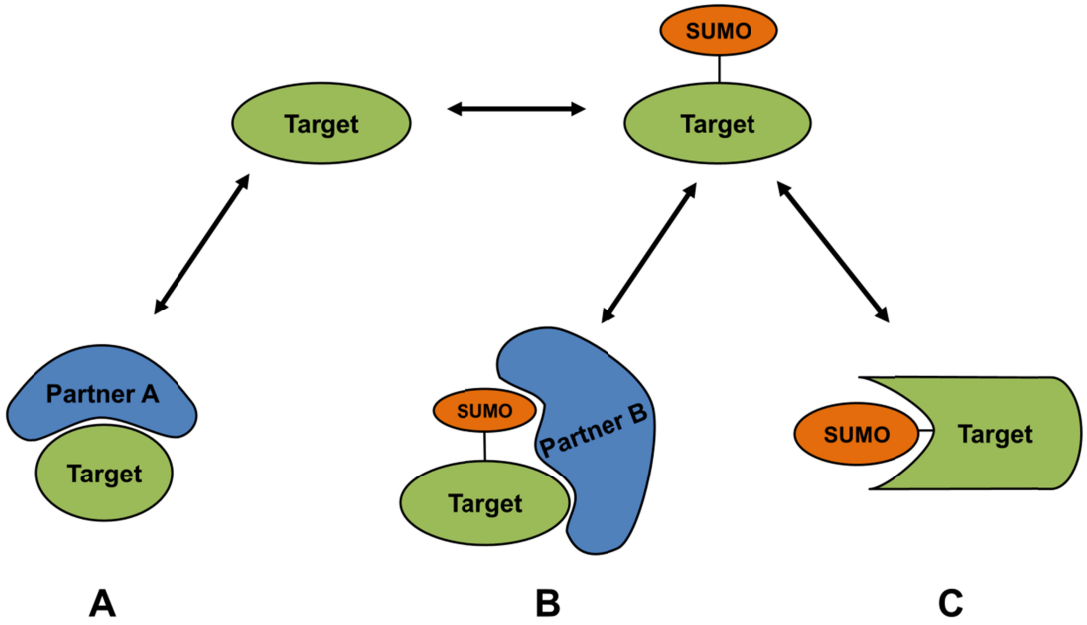


Figure 3.3 – Effects of SUMO modification on protein activity

Schematic representation of the effects of SUMOylation on the activity of a target protein. SUMO modification of a protein can prevent interactions with other partner molecules (A) or it can promote interactions with different molecules that the unmodified protein cannot interact with (B). SUMOylation can also alter the structure of a protein completely (C), having numerous consequences including altered localisation or activity. Partner molecules shown here can represent proteins or in the case of transcription factors, they also represent DNA sequences. Adapted from (Gareau and Lima, 2010).

SUMO modification can often allow the target protein to form novel interactions, but unless a conformational change occurs, these interactions are only possible if the SUMO itself contributes (Geiss-Friedlander and Melchior, 2007). SIMs (SUMO-interacting motifs) have been identified on numerous proteins that non-covalently interact with SUMO, including PML, p73α (tumour suppressor protein, related to p53) and the transcriptional repressor DAXX (death domain associated protein) (Kerscher, 2007). SIMs generally contain a small core of hydrophobic amino acids, flanked by acidic residues or phosphorylated serine residues (Song et al., 2004). The hydrophobic residues form a β-strand conformation that can interact with a hydrophobic pocket on the SUMO surface (Song et al., 2004). Studies on DAXX have shown that the SIM is required for it to be SUMOylated and is essential for its ability to repress transcription (Lin et al., 2006).

3.1.5 SUMOylation and transcription

TFs and transcriptional co-regulators are one of the most abundant classes of SUMOylated proteins (Cubenas-Potts and Matunis, 2013). SUMOylation of TFs appears to often be associated with enhanced transcriptional repression and it has also been suggested that there is a general causal relationship between transcription and SUMOylation (Cubenas-Potts and Matunis, 2013). Inducing hypersumoylation in cells has been observed to increase gene repression (Chupreta et al., 2005) and conversely inducing hyposumoylation, by depleting Ubc9 or SUMO, enhances ectopic gene expression (Ouyang et al., 2009) (Spektor et al., 2011). One mechanism by which these transcriptional changes may occur was identified by Ouyang *et al*. who showed that recruitment of the LSD1-CoREST1-HDAC complex to chromatin required SUMO2/3 (Ouyang et al., 2009). SUMO2/3 interacts directly with a SIM in CoREST1, which promotes the formation of the transcriptionally repressive complex (Ouyang et al., 2009). Secondly, SUMOylation has been shown to enhance DNA methyltransferase activity, which causes transcriptional repression in target genes by blocking DNA accessibility (Lee and Muller, 2009). Finally, SUMO has been found to localise to polycomb group (PcG) bodies which are small subnuclear hubs of transcriptional repression (Cubenas-Potts and Matunis, 2013). The assembly of PcG bodies is dependent on SUMO and without it they cannot correctly repress inappropriate gene expression, for example during embryonic development (Kang et al., 2010).

SUMOylation appears to regulate the activity of a wide range of TFs, of which many are integral to development, e.g. IRF-2 (Han et al., 2008), Ets1 (Ji et al., 2007) and Bcl11b (Zhang et al., 2012); apoptosis, e.g. FLASH (Alm-Kristiansen et al., 2009) and ATF3 (Wang et al., 2013), and cancer, e.g. MITF (Bertolotto et al., 2011) and Myc (Kessler et al., 2012). For many of these TFs, the underlying mechanism by which they are regulated is unclear. However, in the case of Bcl11b, SUMOylation causes the recruitment of the transcriptional co-activator p300 and the subsequent induction of target genes (Zhang et al., 2012). This turns Bcl11b from a repressor to an activator and is critical in developing thymocytes for promoting a T-cell specific gene program (Zhang et al., 2012). An alternative example of the importance of SUMOylation is provided by the melanoma oncogene, MITF, which becomes more active when it cannot be SUMOylated (Bertolotto et al., 2011). A germline missense mutation in the MITF SUMOylation site produces a more transcriptionally active protein that is more oncogenic and promotes cell growth, migration and invasion (Bertolotto et al., 2011). This mutation occurs significantly higher in patients with melanoma and renal cell carcinoma, meaning an individual carrying this mutation has a fivefold increased risk of developing either cancer (Bertolotto et al., 2011).

3.1.6 Studying SUMOylation

One key observation regarding SUMOylation is that only a small proportion of an available substrate needs to be SUMOylated to produce a large effect. Indeed, only small percentages (< 5%) of most SUMOylation targets appear to be modified at steady state, mainly due to the rapid action of SENPs and SUMO conjugating enzymes (Wilkinson and Henley, 2010). The highly dynamic nature of the SUMOylation cycle is at least in part responsible for this, as rapid cycles of modification and demodification allow the whole pool of a target protein to be affected by SUMOylation in a short time frame (Geiss-Friedlander and Melchior, 2007). The underlying mechanisms for this effect has not been determined for many SUMOylated proteins, but two mechanisms have been proposed for how low level SUMOylation can have a dramatic effect on TF activity.

Firstly, SUMOylation of a TF could induce chromatin-remodelling activity to induce a repressive or active chromatin state that will remain, even after SUMOylation of the TF is removed. Secondly, the presence of the SUMO modification will promote interactions with other transcriptional co-factors which can come together to form a repressive or active transcriptional complex that, once assembled, will remain stable even if the SUMO modifications are lost (Hay, 2005). One well studied example of this is DAXX, which interacts with histone deacetylase (HDAC) enzymes that remove acetyl groups from histone proteins, allowing them to bind more tightly to DNA, making it more inaccessible (Kuo et al., 2005). SUMO-mediated recruitment of HDACs therefore induces transcriptional repression of DAXX target genes (Kuo et al., 2005).

Understanding how SUMOylation can control the activity of a particular protein presents a number of challenges. As mentioned above, the activity of SENPs, which rapidly remove SUMO from target proteins, is one of the main challenges when studying SUMOylation, especially as these enzymes work inside the cell and also after cell lysis (Da Silva-Ferrada et al., 2012). It is therefore important to either lyse cells under denaturing conditions or use broad cysteine protease inhibitors, e.g. *N*-ethylmaleimide (Horowitz et al., 2013), in cell lysis buffers. As only a small percentage of a particular protein will be modified at steady state and many proteins, particularly TFs, constitute a tiny proportion of all the protein within a cell, detecting endogenously modified protein is often not possible, even with highly sensitive mass spectrometry (MS).

Many studies involve the over expression of SUMO isoforms or protein of interest and many have used *in vitro* assays to identify the modification of protein. SUMOylation levels inside

the cell can also be manipulated to help aid detection of a modified target protein, as global SUMOylation levels increase with cell stress responses. SUMOylation levels have been shown to increase following heat shock, osmotic and oxidative stress and inhibition of the proteasome (Golebiowski et al., 2009) (Saitoh and Hinchey, 2000) (Schimmel et al., 2008). These experimental approaches are beneficial, but must be used in combination to ensure that false positive modifications are not identified. MS of endogenously modified protein and unambiguous mass accuracy is the only definitive way to confirm the modification of a protein *in vivo* and determine the exact sites of SUMO modification (Da Silva-Ferrada et al., 2012).

3.1.7 SUMOylation of E4bp4

Following the identification of a possible interaction between E4bp4 and PIAS1 **(Table 1.1)**, other evidence was saught from the literature to suggest if E4bp4 was SUMOylated. Numerous proteomic screens have been performed looking to identify novel SUMOylation targets and analyse global changes in the SUMOylation status of mammalian cells. Interestingly, four of these studies have positively identified E4bp4 as a potential target of SUMO2/3 modification. Golebiowski *et al*. and Schimmel *et al*. both analysed all the SUMOylated proteins in HeLa cells expressing 6His-SUMO2 or TAP-SUMO2 respectively and looked for changes induced by cell stress, either heat shock or proteasomal inhibition (Golebiowski et al., 2009) (Schimmel et al., 2008). Using MS, both studies identified multiple E4bp4 peptides alongside thousands of other peptides that were enriched following SUMO2 purification. Both studies also suggested that E4bp4 SUMOylation was not affected by cell stress; neither heat shock or proteasomal inhibition.

Schou *et al*. also performed a similar set of experiments, but looked at dynamic SUMOylation changes during mitosis (Schou et al., 2014). They found that E4bp4 was enriched following purification of 6His-SUMO2 and an increase in E4bp4 SUMOylation was observed at the end of mitosis, consistent with many others TFs (Schou et al., 2014). However, none of these studies were able to identify the presence of specific SUMOylated E4bp4 peptides in their MS analysis and therefore do not provide definitive evidence of SUMOylation at specific lysine residues. Additionally, in these studies, SUMO is over expressed in cells before the analysis and this has the potential for false positive identification of SUMOylation sites. Further targeted approaches are required to confirm the SUMOylation of E4bp4 and, more importantly, determine if it is essential for the function of E4bp4.

3.1.8 Aims and objectives

There is now evidence to suggest that E4bp4 may be SUMOylated, but this has not been determined *in vivo*. It is not known whether the SUMOylation of E4bp4 is required for its activity as a TF and, in particular, in the regulation of NK cell development. Therefore this work aimed to confirm if E4bp4 is SUMOylated and determine if this modification is essential for its function. In order to achieve this, the following objectives were set:

1. Analyse the amino acid sequence of E4bp4 and look for potential SUMOylation sites.
2. Confirm that E4bp4 is SUMOylated *in vivo* using a number of approaches.
3. Determine the sites of E4bp4 modification using high accuracy MS and create mutant versions of E4bp4 that lacks the ability to be SUMO modified.
4. Determine if SUMOylation is required for the activity of E4bp4 during the development of NK cells, using an *in vitro* NK cell development assay.

3.2 E4bp4 has five potential SUMOylation sites

Initial bioinformatic analysis was performed to determine if the primary amino acid sequence of E4bp4 contained any predicted sites for SUMOylation. As SUMOylation occurs principally at the consensus sequence Ψ-K-x-E, it is possible to search for potential SUMOylation sites using a number of web based bioinformatic tools. SUMOplot (Xu et al., 2008), SUMOsp 2.0 (Xue et al., 2006) and seeSUMO (Teng et al., 2012) identified five potential SUMOylation sites in the E4bp4 primary amino acid sequence at lysine residues 10, 116, 219, 337 and 394 **(Table 3.1)**. All three searches provided a confidence score for the likelihood of a positive identification, based on the sequence arrangement around the centrally modified lysine. All three tools identified K10, K219 and K337 as the most likely sites for SUMOylation **(Table 3.1)**.

Recently, another online tool for predicting mammalian SUMOylation sites, SUMOhydro, was developed. This tool included amino acid hydrophobicity as a parameter in the prediction process and had improved prediction performance (Chen et al., 2012). In some instances, hydrophobicity has been found to be critical for SUMOylation site recognition (Matic et al., 2010). Using SUMOhydro, five potential SUMOylation sites were identified, but K180 was predicted to be SUMOylated rather than K116 which had previously been predicted by the alternative prediction tools **(Table 3.1)**. However, the prediction score for

K180 was significantly lower than the other four sites, suggesting it is unlikely to be modified **(Table 3.1)**. In line with the predictions made by the alternative prediction tools, SUMOhydro identified K10 and K219 as the two sites most likely to be SUMOylated.

Table 3.1 – Identification of putative E4bp4 SUMOylation sites using bioinformatic prediction tools.

Position	Peptide	SUMOplot score (0 – 1)	SUMO sp2.0 score (0 – 6)	seeSUMO score (0 – 1)	SUMOhydro score (0 – 2)
K10	MQTIKKEPAPL	0.94	5.422	0.902	1.205
K116	NATLKAELLSL	0.91	1.578	0.851	x
K180	CISVIKHSPQSS	x	x	x	0.053
K219	FPVIKQEPVEL	0.94	5.611	0.938	1.681
K337	AMQVKVEALDS	0.93	3.393	0.936	0.882
K394	DWSLKSEHWHH	0.91	1.009	0.876	0.751

SUMOylation consensus sequence highlighted

Most of the potential SUMOylation sites of E4bp4 were also found to be highly conserved across a range of vertebrate species, from humans to Zebrafish and Xenopus **(Figure 3.4)**. K394 was not as well conserved as the other sites and was only present in three out of the six species analysed **(Figure 3.4)**. This initial bioinformatic data suggested that K10, K219 and K337 were likely to be the sites of E4bp4 SUMOylation, with K116 and K394 also possible, but less likely sites of modification.

3.3 E4bp4 is multiply SUMOylated

To ascertain which isoform(s) of SUMO can interact with E4bp4, HeLa cell lines which express 6xHis tagged forms of each SUMO isoform were used. Using a protocol adapted from Tatham *et al*. (Tatham et al., 2009), E4bp4 was transiently expressed in HeLa cell lines that stably produce 6His-SUMO1, 6His-SUMO2 and 6His-SUMO3 **(Figure 3.5)**. Cells were lysed under denaturing conditions, 20% of the sample was removed as an input sample and all the remaining SUMO-modified proteins were purified using Ni^{2+} affinity chromatography **(Figure 3.5)**. Subsequently, input and Ni^{2+} affinity purified samples were analysed by Western blot (WB) for the presence of E4bp4.

E4bp4 (~60 kDa) was detected in all the 6His-SUMO samples following purification, as well as additional higher molecular weight forms of the protein at ~80 kDa, ~100 kDa and ~120 kDa **(Figure 3.6)**. These higher molecular weight forms were not observed in the input samples and were more prominent in the 6His-SUMO2 and 6His-SUMO3 samples than in the 6His-SUMO1 sample **(Figure 3.6)**. As each SUMO conjugate adds 10-15 kDa onto the apparent molecular weight of a protein, it is likely that these higher molecular weight forms of E4bp4 correspond with various SUMOylated and multiply SUMOylated forms of the protein. SUMO2 and SUMO3 share 95% sequence homology, and are generally functionally homologous therefore, as expected, the banding pattern observed for these two samples was almost identical **(Figure 3.6)**. As E4bp4 has five potential SUMOylation sites **(Figure 3.4)**, it is plausible that it is modified by SUMO at multiple sites. SUMO2 and SUMO3 are also known to form chains as they have an internal SUMO acceptor site, therefore E4bp4 could have multiple SUMO modifications attached to a single lysine residue (Tatham et al., 2001). RanGAP1 was used as the control in these experiments as it is known to be highly SUMOylated by both SUMO1 and SUMO2/3 and the presence of its SUMOylated form (~80 kDa) in the Ni^{2+} purified samples, confirms the successful enrichment of SUMOylated protein **(Figure 3.6)**.

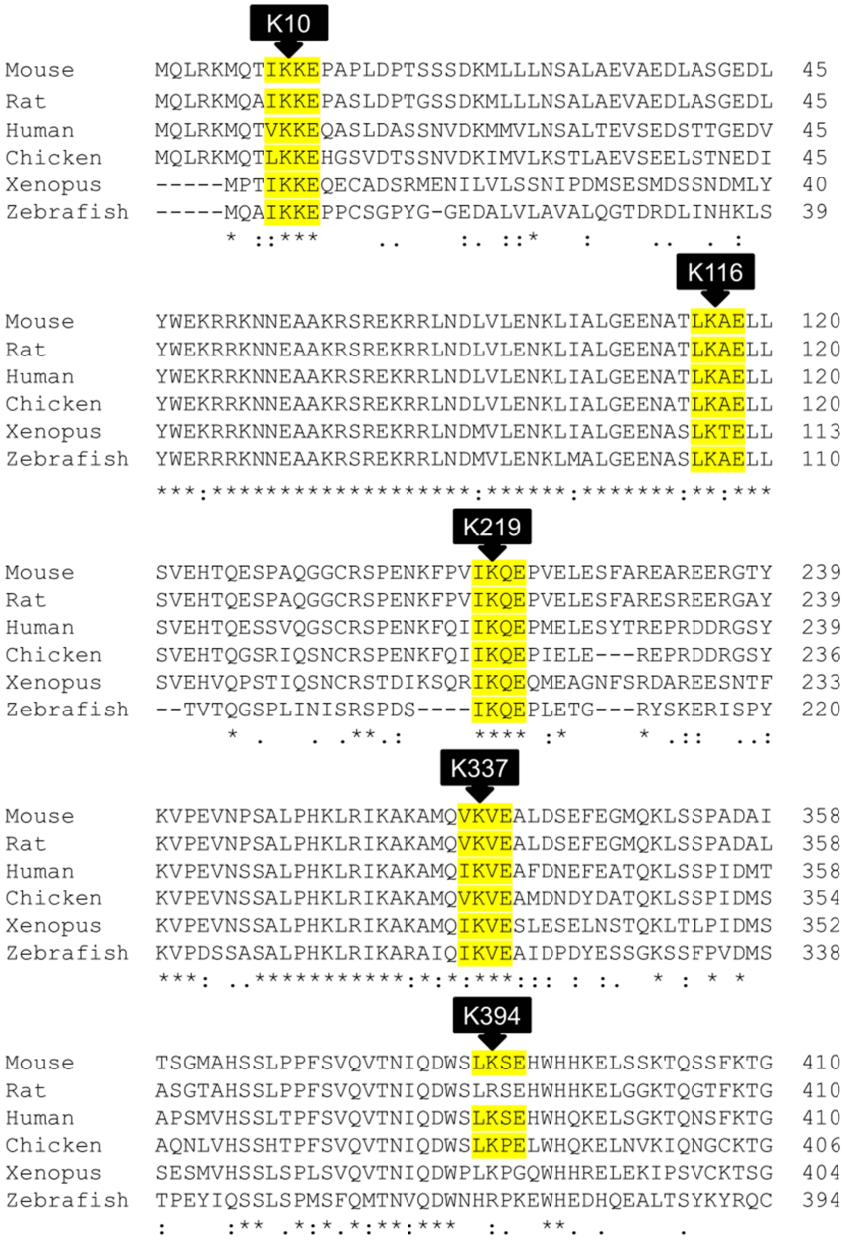


Figure 3.4 – Putative SUMOylation site of E4bp4 are highly conserved across a range of species

The E4bp4 amino acid sequence from various species was aligned using the web-based tool Clustal Omega (Goujon et al., 2010). Sites of potential SUMO modification are highlighted. Asterisks highlight perfectly conserved residues and dots highlight partially conserved residues.

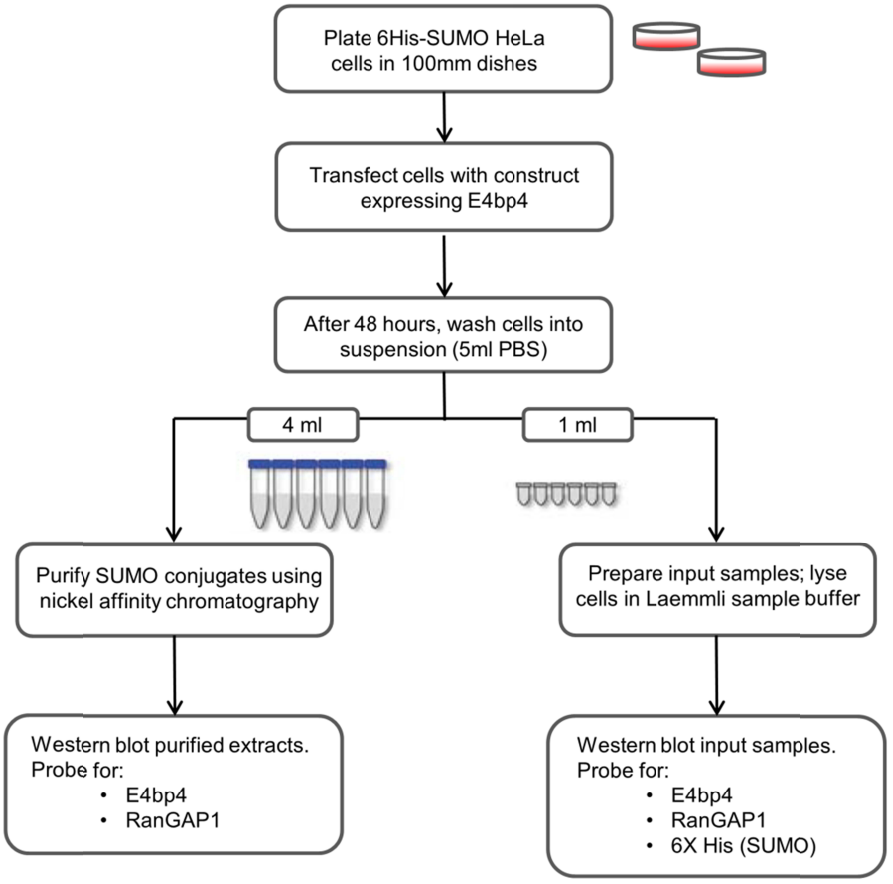


Figure 3.5 – Detection of E4bp4 SUMOylation *in vivo*

Schematic representation of the protocol used to determine if E4bp4 is SUMOylated *in vivo*. HeLa cell lines stably expressing 6His-SUMO1, 6His-SUMO2 and 6His-SUMO3 were grown and transfected with a pCMV construct expressing E4BP4. Input samples were used to confirm expression of E4bp4 in all cells. Ni^{2+} affinity chromatography was used to purify all SUMO conjugated proteins, which were subsequently analysed by WB. The presence of higher molecular weight species of E4bp4 on WBs was due to the addition of SUMO modifications onto the protein. All WBs were additionally probed for RanGAP1 to confirm loading and show that SUMO purification was successful. Adapted from (Tatham et al., 2009).

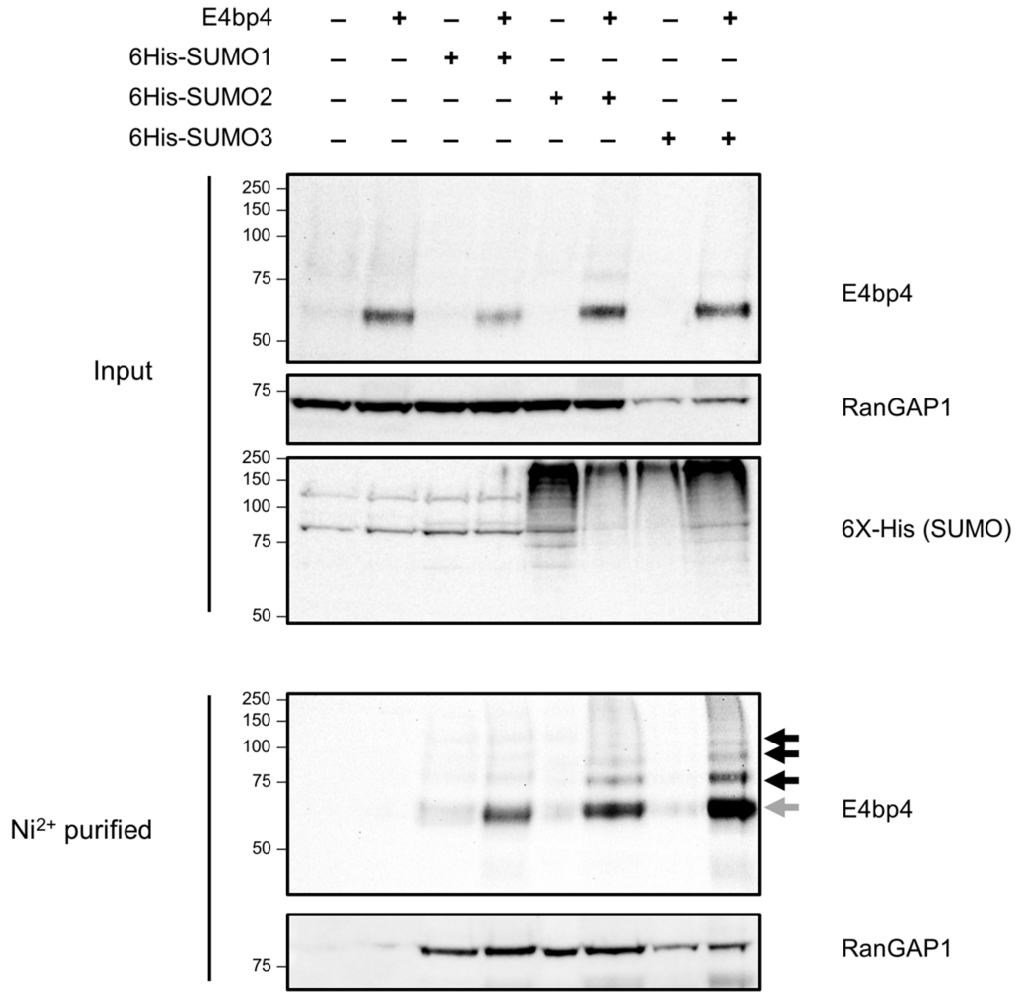


Figure 3.6 – E4bp4 has multiple SUMO modifications by both SUMO1 and SUMO2/3

E4bp4 was expressed in HeLa cells stably expressing 6His-SUMO1, 6His-SUMO2 or 6His-SUMO3. As a control the empty vector (pCMV) was also transfected into each cell line. Input samples were taken from each sample, prior to Ni^{2+} affinity purification performed under denaturing conditions. Equivalent volumes of purified samples were separated by reducing SDS-PAGE and blotted. Blotting for RanGAP1 provided a positive control for the purification process as it is known to be modified by both SUMO1 and SUMO2/3. Blotting for 6X-His demonstrated the expression of 6His-SUMO proteins in each cell line. Grey arrows highlight unmodified E4bp4 and black arrows highlight SUMO modified forms of E4bp4 with higher molecular weights. Molecular weight markers show on the left of each WB in kDa. Representative of three independent experiments.

To confirm these findings, it was important to perform the inverse experiment by which E4bp4 is specifically immunoprecipitated (IP) from cell lysates containing 6His-SUMO and the presence of SUMO modified E4bp4 is then detected by WB. To allow E4bp4 protein to be efficiently purified from cell lysates, an expression construct was created that expressed E4bp4 with a FLAG® tag at its amino terminus. The FLAG peptide is a recombinant, hydrophilic 8-amino acid peptide that is specially recognised by several anti-FLAG monoclonal antibodies, allowing for simple, high specificity purification of the tagged protein (Einhauer and Jungbauer, 2001). Additionally, the small size of the FLAG tag means it is unlikely to obscure any epitopes or alter protein function, transport or secretion (Einhauer and Jungbauer, 2001). To create this expression construct a two-stage PCR amplification procedure was used to produce an E4bp4 cDNA with the attached FLAG tag **(Figure 3.7)**. Specific primers first amplified the E4bp4 cDNA whilst incorporating a 5' FLAG tag, a 3' stop codon and a 3' XhoI restriction digest site (Stage 1). The second pair of primers incorporated an ATG start codon and NotI restriction digest site at the 5' end of the cDNA (Stage 2). Using the incorporated restriction digest sites, the cDNA was ligated into the pCMV-script expression vector for expression in the HeLa cells **(Figure 3.7)**. The construct was initially transfected into parental HeLa cells and FLAG-E4bp4 was detected in whole cell lysate when analysed by WB **(Figure 3.7)**.

In order to allow E4bp4 to be purified from cell lysates using the FLAG tag, lysates have to be prepared in non-denaturing conditions to allow the use of anti-FLAG monoclonal antibodies for the purification. Therefore, it was essential to use a SUMO protease inhibitor in all cell lysis buffers, as it is known that SENPs act efficiently to remove SUMO modifications from substrates (Gareau and Lima, 2010). The broad cysteine protease inhibitor, NEM, has been used effectively to inhibit SENP activity (Sarge and Park-Sarge, 2009). NEM was used in cell lysis buffers and to observe its importance HEK-293T cells were transfected with FLAG-E4bp4 and lysed in the presence or absence of NEM **(Figure 3.8A)**. High molecular weight (SUMOylated) forms of E4bp4 could only be observed in cell lysates when 25 mM NEM was present in the cell lysis buffer **(Figure 3.8A)**.

FLAG-E4bp4 was expressed in HeLa 6His-SUMO1, 6His-SUMO2 and 6His-SUMO3 cell lines and cell lysates were prepared under non-denaturing conditions (with NEM) to allow purification of FLAG-E4bp4 using anti-FLAG® M2 affinity resin (anti-FLAG M2 monoclonal antibody permanently conjugated to agarose). FLAG purified samples were analysed by WB and, when probed with an anti-E4bp4 antibody, additional higher molecular weight forms of E4bp4 were again observed in the 6His-SUMO2 and 6His-SUMO3 samples **(Figure 3.8B)**. The same bands were clearly observed when the same WB was probed with an anti-His

antibody and an anti-SUMO2/3 antibody **(Figure 3.8B)**. Higher molecular forms of E4bp4 were observed in the 6His-SUMO1 sample but these were very low in abundance **(Figure 3.8B)**. These results confirmed that E4bp4 can acquire multiple SUMO modifications and the protein can be modified primarily by SUMO2/3 and to a lesser extent by SUMO1.

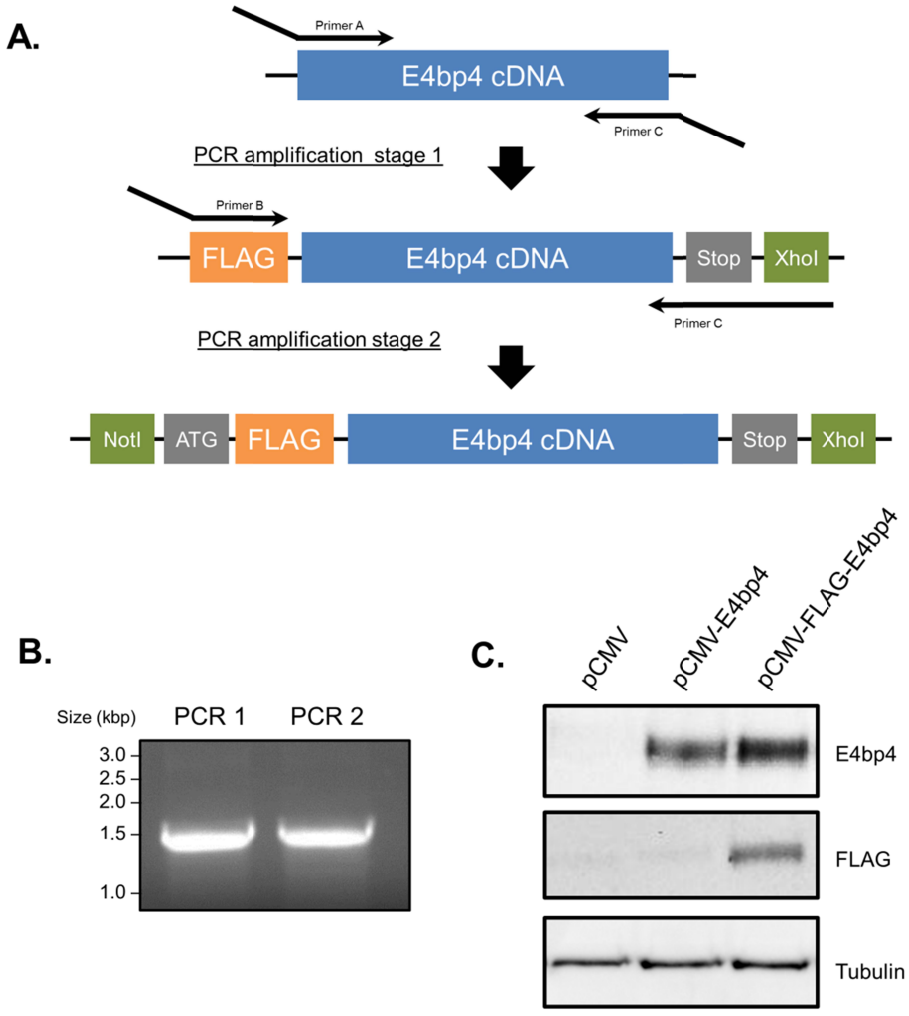


Figure 3.7 – Cloning and expression of FLAG-E4bp4

A) Schematic representation of two-step PCR amplification of E4bp4 cDNA to add a 5' FLAG tag. PCR amplification stage 1 used primers to add the FLAG tag, a 3' XhoI digest site and a 3' stop codon. PCR amplification stage 2 used primers to add a 5' NotI digest site and an ATG start codon. Primers are donated by black arrows. B) Following each PCR reaction, products were run on a 1% agarose gel and visualised using ethidium bromide. E4bp4 cDNA = 1460 bp. C) FLAG-E4bp4 was expressed in HEK-293T cells and whole cell lysates were analysed by WB. The blot was sequentially probed for anti-FLAG, anti-E4bp4 and anti-tubulin.

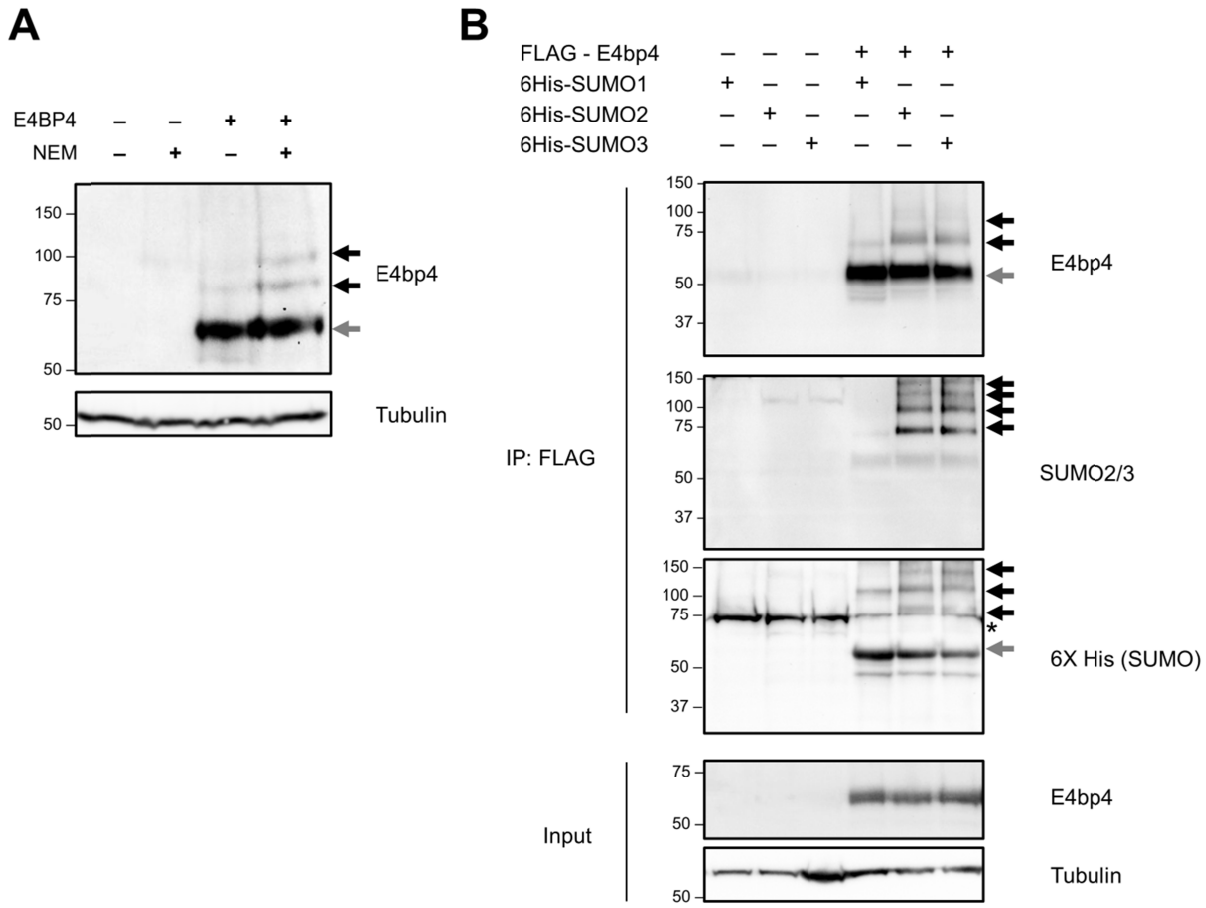


Figure 3.8 – FLAG-E4bp4 purification confirms E4bp4 has multiple SUMO modifications

A) FLAG-E4bp4 (or empty vector) was expressed in HEK-293T cells and whole cell lysates were prepared in the presence or absence of the SUMO protease inhibitor NEM. Lysates were separated by SDS-PAGE and blotted. Grey arrows highlight unmodified E4bp4 and black arrows highlight SUMO modified forms of E4bp4 with higher molecular weights. B) FLAG-E4bp4 (or empty vector) was expressed in 6His-SUMO1, 6His-SUMO2 and 6His-SUMO3 HeLa cells and FLAG-E4bp4 was immunoprecipitated using anti-FLAG M2 affinity gel (Sigma). IP samples and input samples were separated by SDS-PAGE and blotted. Blotting for anti-His identified the SUMOylated protein in each IP sample. Blotting for anti-SUMO2/3 further identified SUMOylated E4bp4. Grey arrows highlight unmodified E4bp4 and black arrows highlight SUMO modified forms of E4bp4 with higher molecular weights. Molecular weight markers show on the left of each WB in kDa.

3.4 Mutating putative SUMOylation sites of E4bp4 abolishes modification of the protein

To determine if a particular lysine side chain is a site of SUMO modification, the amino acid of interest can be mutated to an arginine residue **(Figure 3.9)**. This mutation will abolish any SUMO modification at that site, whilst maintaining the structural integrity of the protein (Tatham et al., 2009) **(Figure 3.9)**. E4bp4 had five putative SUMOylation sites and the central lysine at each site was individually mutated in the E4bp4 cDNA to an arginine using site directed mutagenesis **(Table 3.2)**. To ensure the mutations had been created correctly, all mutant cDNAs were sequenced. Numerous double, triple and multi-site E4bp4 SUMO mutants were created using this approach **(Table 3.2)**. To ensure the mutations had been made correctly in each cDNA and no unwanted extra mutations had occurred, each cDNA was sequenced. Analysing the SUMO modifications of these mutants was used to establish which residues of the E4bp4 protein are SUMOylated.

Table 3.2 – Mutant forms of E4bp4 created lacking putative SUMOylation sites

Name	Lysine – Arginine mutations	Name	Lysine – Arginine mutations
E4bp4-K10R	K10	E4bp4-K219-337R	K219, K337
E4bp4-K116R	K116	E4bp4-K337-394R	K337, K394
E4bp4-K219R	K219	E4bp4-3X-1-SUMO	K10, K116, K219
E4bp4-K337R	K337	E4bp4-3X-2-SUMO	K10, K219, K337
E4bp4-K394R	K394	E4bp4-3X-3-SUMO	K116, K219, K337
E4bp4-K10-219R	K10, K219	E4bp4-4X-SUMO	K10, K116, K219, K337
E4bp4-K116-219R	K116, K219	E4bp4-5X-SUMO	K10, K116, K219, K337, K394

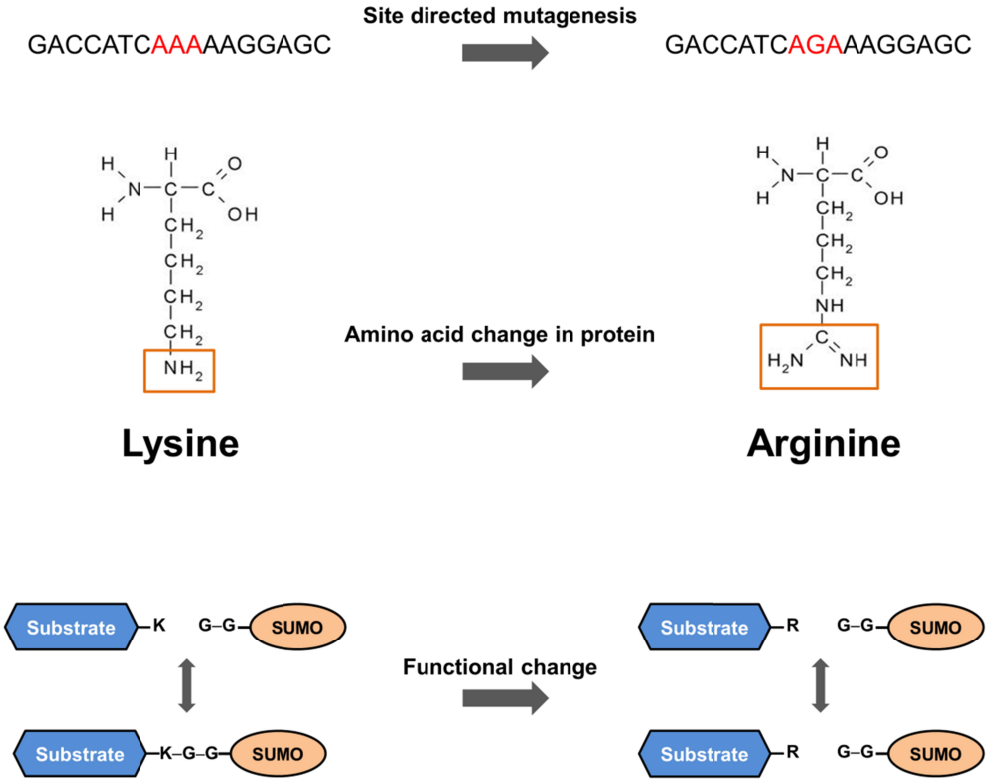


Figure 3.9 – Mutating lysine residues to arginine abolishes SUMO modifications

Site directed mutagenesis was used to mutate individual lysine residues with arginine residues. Mutation does not affect the three-dimensional structure of the protein but will prevent any SUMO modification at that site. Lysine side chains form an isopeptide bond with the –GG motif of SUMO, but this cannot occur with the side chain of arginine.

To assess how each of these mutations and combinations of mutations affected the SUMOylation of E4bp4, each mutant was expressed in the 6His-SUMO1, 6His-SUMO2 and 6His-SUMO3 HeLa cell lines. Cell lysates were made from each sample and all the SUMOylated proteins were purified using Ni^{2+} affinity chromatography, before input and purified samples were analysed by WB **(Figure 3.5)**.

Expression of wild-type (WT) E4bp4 in the 6His-SUMO2 and 6His-SUMO3 cell lines produced three higher molecular weight SUMOylated forms of E4bp4, that were not present in the input samples **(Figure 3.10) (Figure 3.11)**. E4bp4-K219R was the only individual mutant which had any effect on the SUMOylation pattern of E4bp4 **(Figure 3.10) (Figure 3.11)**. All of the other individual mutants had the sample banding pattern as WT E4bp4

suggesting K219 is an important site of SUMOylation. However, the K219R mutation did not abolish all of the higher molecular weight bands, suggesting multiple sites of E4bp4 are modified and a potential hierarchy exists between the sites **(Figure 3.10) (Figure 3.11)**. As K219R was observed to be a significant SUMOylation site, nearly all of the multi-site mutants that were created contained K219R **(Table 3.2)**.

All of the multi-site mutations affected the SUMOylation of E4bp4 **(Figure 3.10) (Figure 3.11)**. The 5X SUMO mutant lacked all of the putative E4bp4 SUMOylation sites and this resulted in ablation of all the higher molecular weight forms of E4bp4 **(Figure 3.10) (Figure 3.11)**. This same effect was seen with the 4X SUMO mutant, as well as the K10-219-337R SUMO mutant **(Figure 3.10) (Figure 3.11)**. This confirmed the bioinformatic prediction data which suggested that these three sites are the SUMOylation sites of E4bp4.

Each E4bp4 mutant was also expressed in the 6His-SUMO1 cell line **(Figure 3.12)**. Expression of WT E4bp4 in these cells resulted in the production of two higher molecular weight forms of E4bp4. However, these were very low in abundance and not as prominent as those produced in the 6His-SUMO2 and 6His-SUMO3 cell lines **(Figure 3.12)**, suggesting that E4bp4 is more commonly modified by SUMO2/3 than SUMO1. As previously observed for SUMO2/3, the E4bp4-K219R mutation resulted in a loss of the higher molecular weight forms of E4bp4, but additionally, the K337R mutation also affected E4bp4 SUMOylation **(Figure 3.12)**. All of the multi-site E4bp4 mutants also lacked any higher molecular weight SUMOylated forms **(Figure 3.12)**. This confirmed that K219 and K337 are sites of E4bp4 SUMO1 modification but, as this modification did not appear abundant, E4bp4 is more commonly modified by SUMO2/3. These mutant forms of E4bp4 that lacked any detectable SUMOylation could now be tested to determine if SUMOylation directly affects the function of the E4bp4 protein.

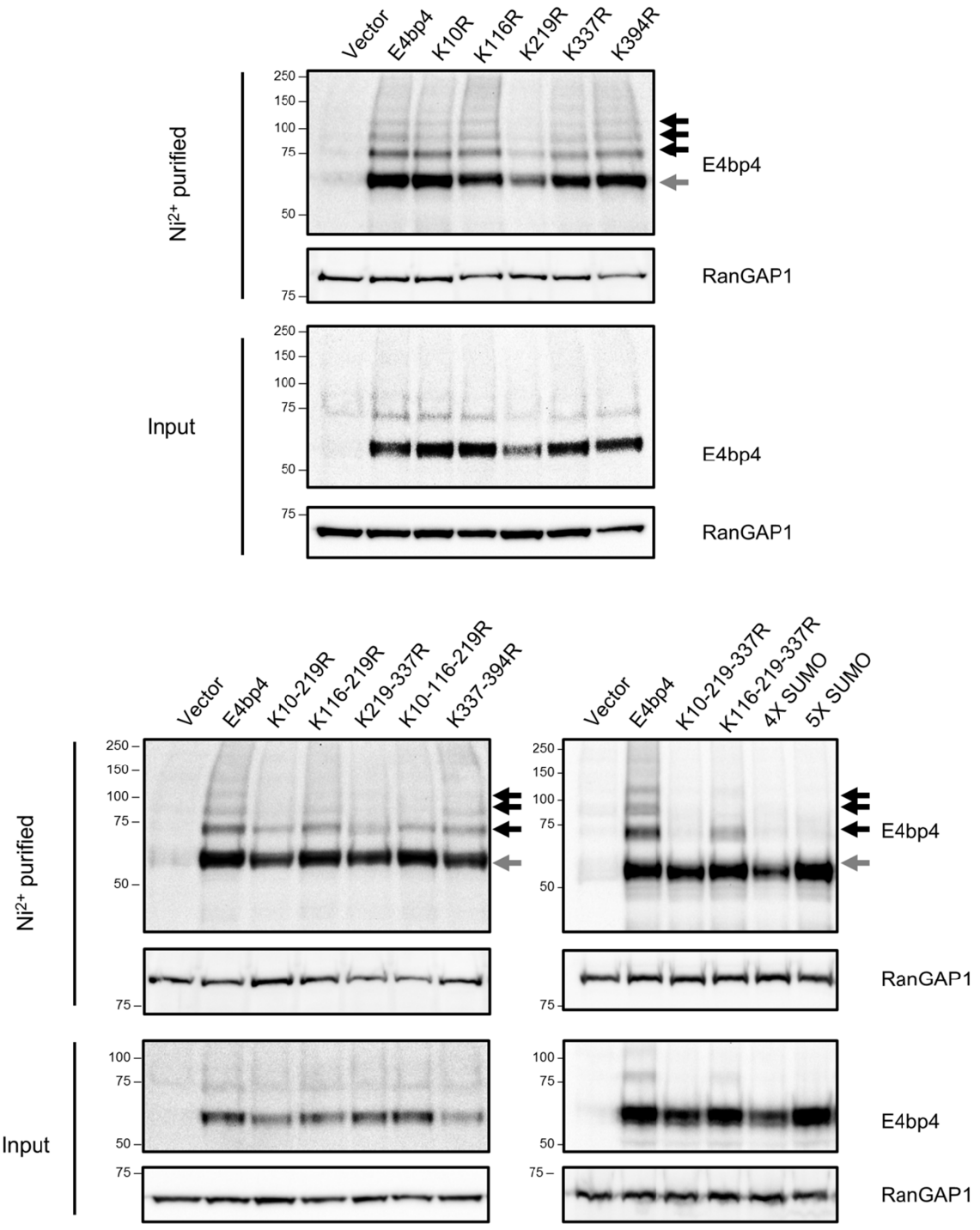


Figure 3.10 – Mutation of putative E4bp4 SUMOylation sites abolishes SUMO2 modification

Five single and various multiple mutants, including E4bp4-5X SUMO were expressed in the 6His-SUMO2 HeLa cell line. Forty eight hours after transfection cells were lysed and Ni^{2+} affinity purification was used to enrich the SUMO modified protein. Prior to affinity purification, input samples were prepared for each mutant and were compared by WB for the expression of E4bp4. Equivalent volumes of Ni^{2+} purified samples were separated by SDS-PAGE and compared by WB for the presence of higher molecular weight modified forms of E4bp4. Blots were additionally probed for RanGAP1 to confirm loading and successful purification of SUMO modified protein. Grey arrows highlight unmodified E4bp4 and black arrows highlight higher molecular weight, SUMO-modified forms of E4bp4. Molecular weight markers show on the left of each WB in kDa. Representative of two independent experiments.

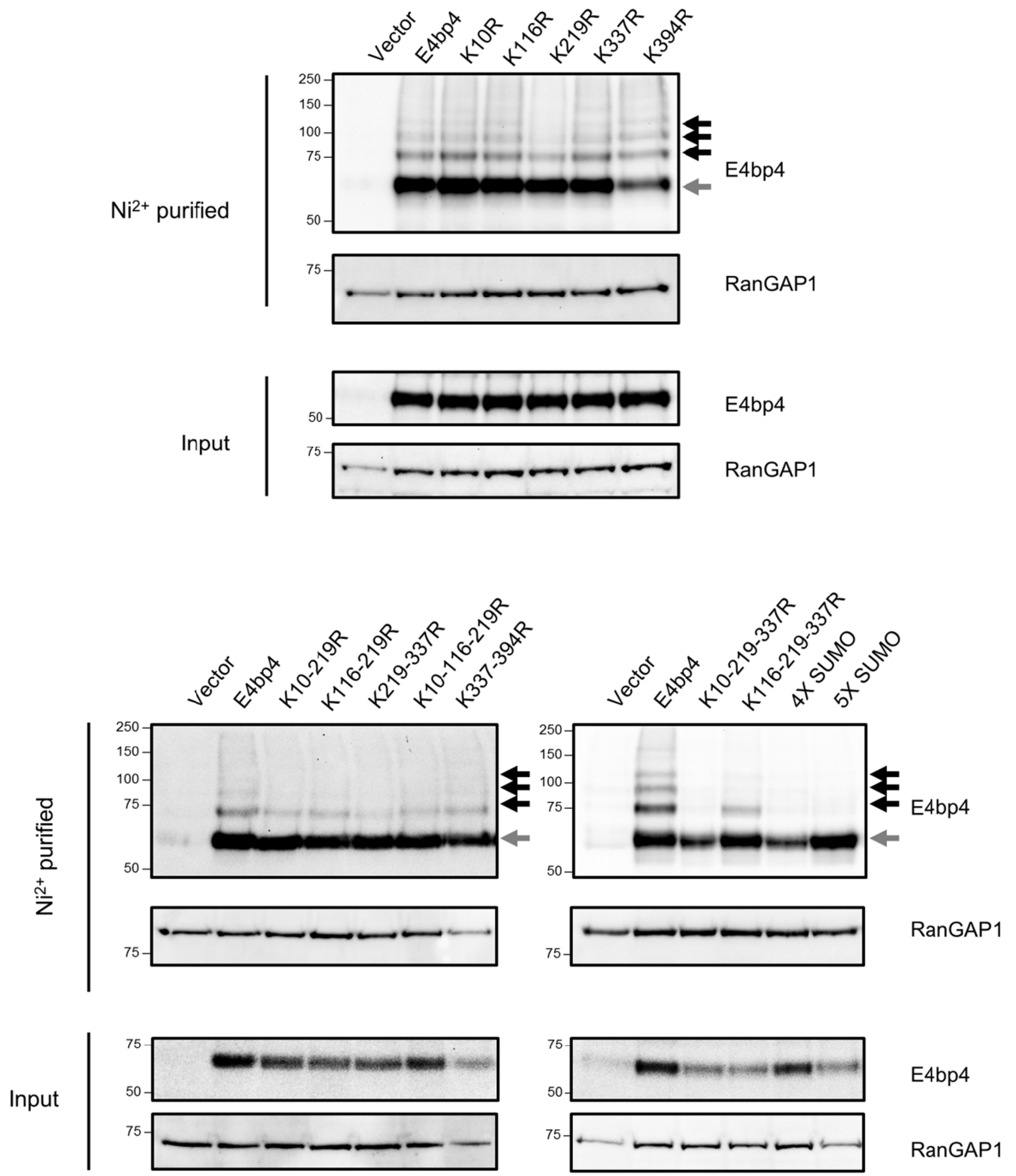


Figure 3.11 – Mutation of putative E4bp4 SUMOylation sites abolishes SUMO3 modification

Five single and various multiple mutants, including E4bp4-5X SUMO were expressed in the 6His-SUMO3 HeLa cell line. Forty eight hours after transfection cells were lysed and Ni^{2+} affinity purification was used to enrich the SUMO modified protein. Prior to affinity purification, input samples were prepared for each mutant and were compared by WB for the expression of E4bp4. Equivalent volumes of Ni^{2+} purified samples were separated by SDS-PAGE and compared by WB for the presence of higher molecular weight modified forms of E4bp4. Blots were additionally probed for RanGAP1 to confirm loading and successful purification of SUMO modified protein. Grey arrows highlight unmodified E4bp4 and black arrows highlight higher molecular weight, SUMO-modified forms of E4bp4. Molecular weight markers show on the left of each WB in kDa. Representative of two independent experiments.

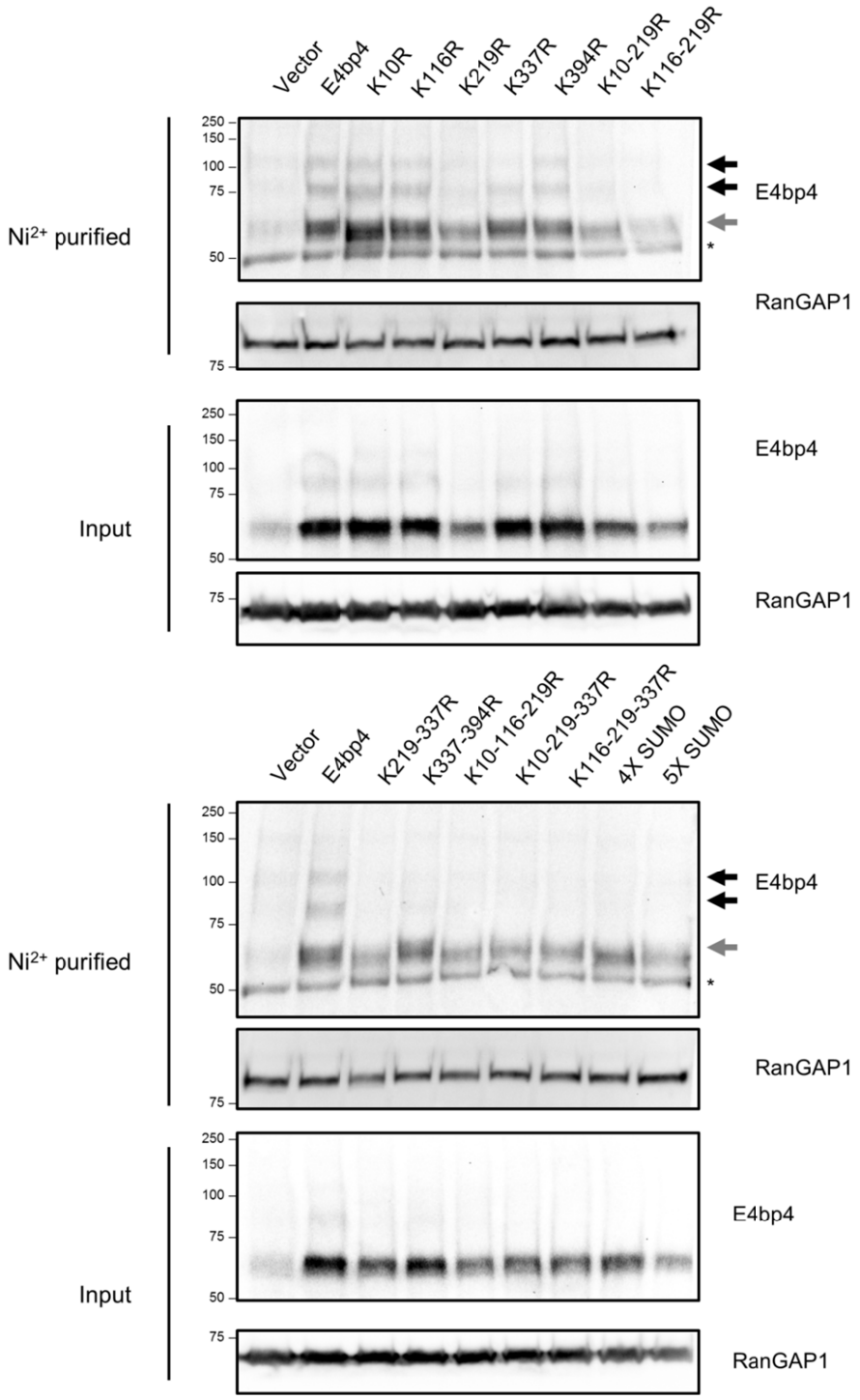


Figure 3.12 – Mutation of putative E4bp4 SUMOylation sites abolishes SUMO1 modification

Five single and various multiple mutants including E4bp4-5X SUMO were expressed in the 6His-SUMO1 HeLa cell line. Forty eight hours after transfection cells were lysed and Ni^{2+} affinity purification was used to enrich the SUMO modified protein. Prior to affinity purification, input samples were prepared for each mutant and were compared by WB for the expression of E4bp4. Equivalent volumes of Ni^{2+} purified samples were separated by SDS-PAGE and compared using WB for the presence of higher molecular weight modified forms of E4bp4. Blots were additionally probed for RanGAP1 to confirm loading and successful purification of SUMO modified protein. Grey arrows highlight unmodified E4bp4 and black arrows highlight higher molecular weight, SUMO-modified forms of E4bp4. * highlight non-specific band. Molecular weight markers show on the left of each WB in kDa. Representative of two independent experiments.

3.5 Analysis of E4bp4 SUMOylation by mass spectrometry

To support the data acquired using the 6His-SUMO HeLa cells, it was important to accurately determine which residues of E4bp4 are sites of SUMO modification. MS has emerged as a powerful technology for accurately determining the sites of various post translational modifications, including phosphorylation, acetylation and ubiquitination on a target proteins (Witze et al., 2007). Global proteomic MS studies have now revealed how PTMs change during the cell cycle, development and disease states (Jeram et al., 2009).

Standard MS analysis for PTMs involves trypsin digestion of the protein(s) of interest and the resulting peptides are separated by chromatography before analysis on the MS. After an initial MS scan, peptides of high abundance, or those of a molecular weight of interest are fragmented by collision-induced dissociation (CID) to produce a collection of resulting ions that are detected by MS/MS, allowing the peptide to be sequenced **(Figure 3.13)**. During CID many peptide fragments are generated, but the major product ions are derived from fragmentation events occurring at consecutive peptide bonds, referred to as the *b*- and *y*-ion series (Jeram et al., 2009). *b*-ions derive from fragmentation at the N-terminus of the peptide and *y*-ions from fragmentation at the C-terminus **(Figure 3.13)**. Simple chemical PTMs do not undergo significant fragmentation during CID, they simply alter the mass of one or more fragment ions by a predictable mass e.g. acetylation causes a mass increase of 42.01 Da **(Figure 3.13B)**. Identification of PTM modification sites where the PTM is a peptide is more complex, as the PTM itself will fragment during CID, resulting in a greater number of product ions in the MS/MS spectra. Trypsin digest of a ubiquitin modified peptide results in a –GG remnant side chain on the modified lysine residue and the trypsin does not cleave at the modified lysine as the ubiquitin modification blocks the cleavage site (Jeram et al., 2009) **(Figure 3.13C)**. This –GG modified peptide with a missed cleavage site can therefore be detected in MS/MS spectra and the resulting *b* and *y* ions, with the –GG isopeptide adding an additional 114.04 Da **(Figure 3.13C).**

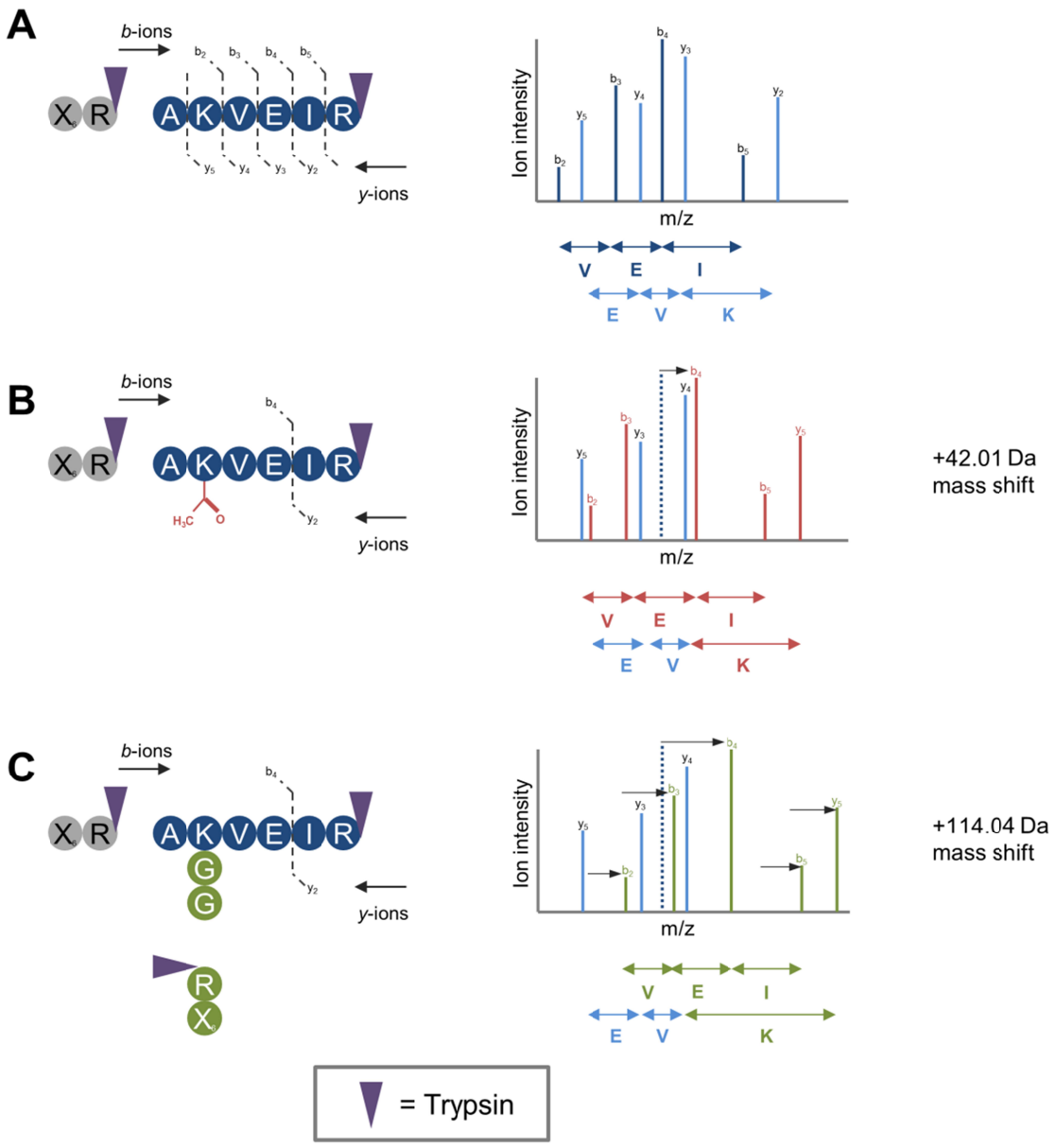


Figure 3.13 – Analysing protein post translational modification using mass spectrometry

Simplified theoretical CID spectra of unmodified and modified peptides. A) CID spectra of an unmodified peptide following trypsin digestion, with b and y ions indicated by light and dark blue peaks. Trypsin digest sites are indicated with purple arrows. B) CID spectra of an acetylated peptide, with all the b ions and y5 ion undergo a +42.01 mass shift (modified ions in red). Dotted blue line represents the unmodified b4 ion from panel A. C) CID spectra of an ubiquitin modified peptide, with GG- modification. All the b ions and y5 ion undergo a +114.04 mass shift (modified ions in green). Dotted blue line represents the unmodified b4 ion from panel A. Adapted from (Jeram et al., 2009).

When studying SUMO modifications by MS, three main problems arise. Firstly, SUMO modified forms of target proteins are generally very low abundance within a cell (Becker et al 2013). Secondly, SUMO modifications are rapidly turned over and are highly susceptible to degradation. It is therefore essential to use general protease inhibitors in cell lysis buffers (e.g. NEM) or some groups have used proteasome inhibitors (e.g. MG132) to promote levels of modified protein (Schimmel et al., 2008). However, even if enough material is attainable, trypsin cleavage of SUMO results in long peptide 'tails' remaining conjugated to the target protein peptides **(Figure 3.14A)**. For example, human SUMO2/3 isopeptides are 32 amino acids in length after trypsin digest, which adds 3567.53 Da to the molecular weight of a modified peptide **(Figure 3.14A)**. This large size and inevitable high charge state makes any modified peptide very difficult to detect in standard MS (Da Silva-Ferrada et al., 2012). If the peptide is detectable the 'tail' peptide can produce multiple fragment ions during CID making the resulting MS/MS spectra almost impossible to interpret (Da Silva-Ferrada et al., 2012). Due to these reasons limited numbers of proteins have been analysed in detail for their SUMOylation sites; however some strategies have been developed to try to conquer this problem.

Mutated forms of SUMO that contain an Arg residue next to the –GG at the C' terminus of the SUMO2 protein have been created, meaning that trypsin digest of SUMO2 produces modified peptides with a simple –GG modification (Galisson et al., 2011) (Tammsalu et al., 2014). This strategy allows for simple identification of SUMO modified peptides via MS/MS spectra and Tammsalu *et al.* have combined this approach with multi protease digests to identify over 1000 SUMOylated peptides in more than 500 human proteins (Tammsalu et al., 2014). Other strategies have involved the development of specialised database searching programs, including ChopNSpice (Hsiao et al., 2009) and SUMmOn (Jeram et al., 2010), which use automated pattern recognition and database searching to identify potential SUMO acceptor sites from complex MS/MS spectra. Some studies have also employed novel digest strategies to reduce the size of the SUMO isopeptide, including the use of elastase (Chicooree et al., 2013) or aspartic acid cleavage (Osula et al., 2012) alongside trypsin, whilst others have used stressed conditions to increase SUMOylation levels in cells (Blomster et al., 2009) (Golebiowski et al., 2009) (Schimmel et al., 2008).

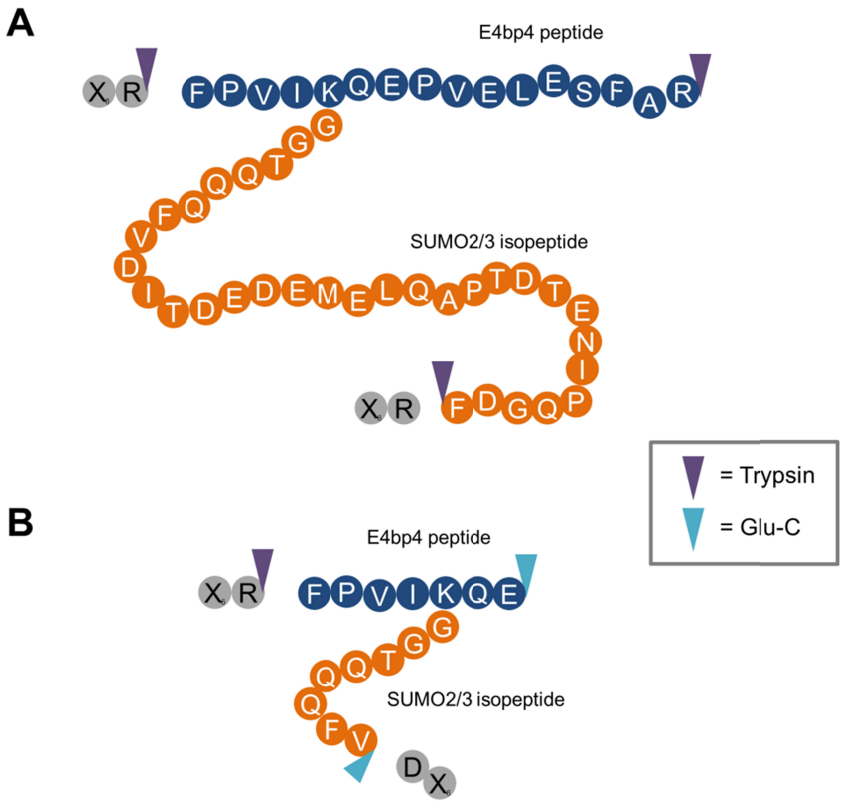


Figure 3.14 – Trypsin and Glu-C double digest provides a method to identify SUMO modified peptides

A) Schematic representation of a typical E4bp4 peptide with a SUMO2/3 modification after trypsin digestion. B) The same peptide, following trypsin and Glu-C double digest, is significantly reduced in size, making it much more amenable for accurate MS/MS identification. Both the E4bp4 backbone peptide and SUMO2/3 isopeptide have been reduced in size. Trypsin digest sites are shown with purple arrows and Glu-C digests sites with light blue arrows.

In order to identify the SUMO modifications of E4bp4, a new strategy for MS analysis was devised that did not involve the over expression of SUMO or the use of stress conditions to promote SUMOylation. This was developed to allow the accurate identification of endogenous E4bp4 SUMOylation sites *in vivo*. In order to do this it was important to have a peptide digest strategy that would reduce the size of the SUMO modification isopeptide 'tail' to make identification of SUMOylated peptides by MS possible. Sequential digestion with the endoproteinase Glu-C and trypsin was predicted to produce a SUMO isopeptide 'tail' of 8 amino acids, because Glu-C digests at the C-terminus of either aspartic or glutamic acid

residues at pH 8 **(Figure 3.14B)**. Additionally, digesting with Glu-C and trypsin reduced the size of the E4bp4 backbone peptides and in particular, digested after the glutamic acid residue which is part of the SUMO consensus modification site **(Figure 3.14B)**. This strategy produced 'T' shaped, branched, SUMOylated peptides which were amenable for MS/MS identification.

For the Glu-C/trypsin double digest strategy to be successful, large quantities of modified E4bp4 protein were required for MS analysis. As E4bp4 is a TF, it is not abundant endogenously in any cell type, even in NK cells where it has such critical functions. Therefore, to produce material for MS analysis, E4bp4 needed to be over expressed and then purified from a mammalian cell line to enable the protein to acquire all of its correct PTMs. Four different mammalian cell lines (Cos-7, 3T3, HEK-293T and LinXE) were transiently transfected with FLAG-E4bp4 and equivalent numbers of transfected cells were analysed by WB for their expression of E4bp4 **(Figure 3.15A)**. HEK-293T cells were observed to produce the highest levels of recombinant protein and so these were chosen to produce material for MS experiments.

Gantke *et al*. recently demonstrated that the expression of the Ebola virus protein, VP35, in 293T cells, blocked plasmid activation of double-stranded-RNA-dependent protein kinase, therefore increasing translation and recombinant protein production (Gantke et al., 2013). To increase levels of FLAG-E4bp4 expression, VP35 protein was co-transfected into 293T cells at varying ratios **(Figure 3.15B)**. At a ratio of 5:3 (FLAG-E4bp4:VP35 plasmid DNA), three times the amount of FLAG-E4bp4 was expressed by the 293T cells, compared to FLAG-E4bp4 on its own **(Figure 3.15B)**. Using these conditions, large quantities of FLAG-E4bp4 were produced by 293T cells. However, for MS analysis this material needed to be highly purified.

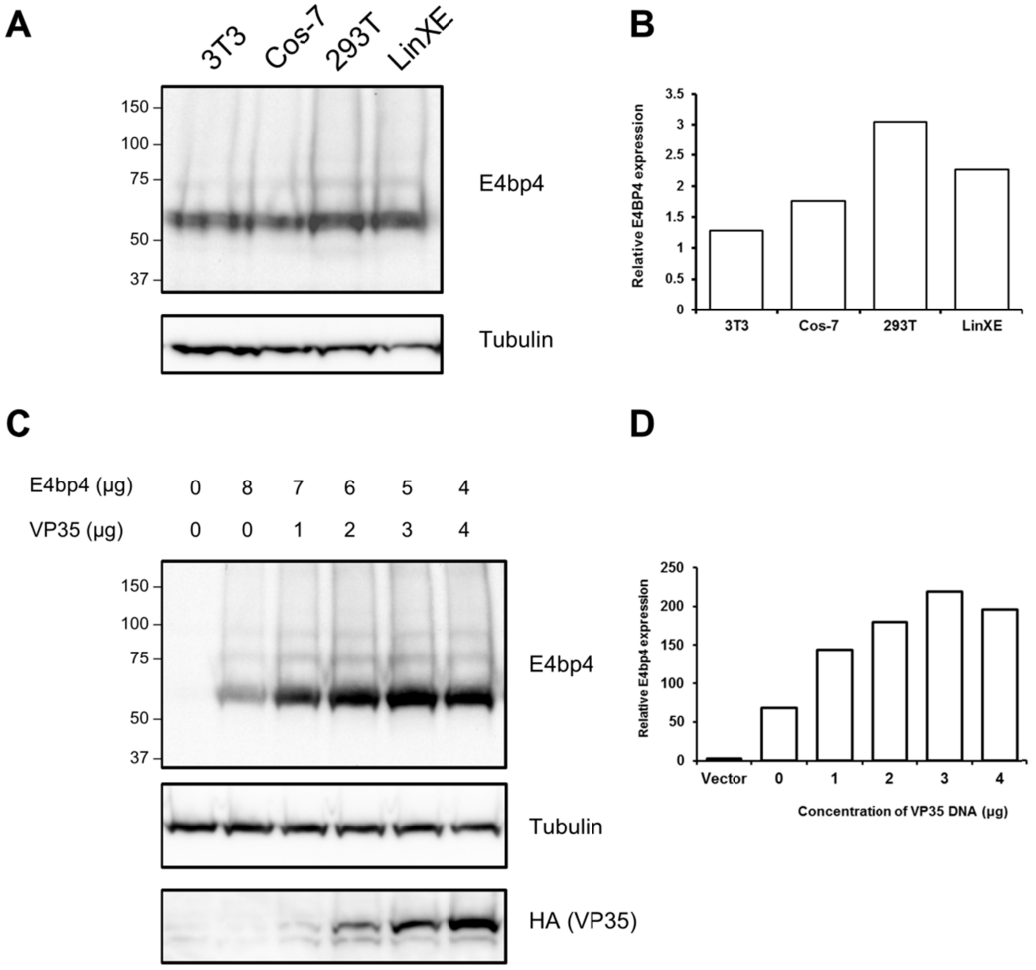


Figure 3.15 – Optimisation of FLAG-E4bp4 expression

A) 3T3, Cos-7, 293T and LinXE cells were all transfected with pCMV-FLAG-E4bp4 and equivalent numbers of transfected cells were lysed in Laemmli sample buffer, separated by SDS-PAGE and blotted. Blotting for tubulin provided a loading control. B) Samples were compared using densitometry analysis using ImageLab software. C) FLAG-E4bp4 was transfected into 293T with increasing amounts of pcDNA-HA-VP35. Each transfection contained 8 µg of DNA in total and each sample had a varying amount of pCMV-FLAG-E4bp4 and pcDNA-HA-VP35 DNA. Transfected cells were lysed in Lamelli sample buffer, separated by SDS-PAGE and blotted. Probing for HA confirmed expression of VP35 and probing for tubulin provided a loading control. D) Samples were compared using densitometry analysis using ImageLab software. Molecular weight markers show on the left of each WB in kDa.

As the SUMO modified forms of E4bp4 are low in abundance, prior to MS analysis the trypsin/Glu-C digest, was performed on a purified E4bp4 sample in solution rather than from a section of SDS-PAGE gel. In-gel protease digestion can be advantageous as the proteins are separated on a gel and cutting out the band of interest, means that any unwanted contaminating proteins can be removed from the sample. However, in-gel digestion can

result in substantial sample loss and as E4bp4 has multiple modified forms, of different molecular weights it was important to ensure that they were all in the sample that was to be analysed by MS. Therefore, to produce a highly purified sample of FLAG-E4bp4 for MS analysis the FLAG purification process was optimised and in particular the method for eluting the purified protein.

Cell lysates containing FLAG-E4bp4 were prepared under non-denaturing conditions using a protocol adapted from Klenova *et al*. (Klenova et al., 2002) to allow IP of FLAG-E4bp4 using anti-FLAG® M2 affinity resin. A variety of methods were initially tested to elute the FLAG-E4bp4 from the anti-FLAG gel, with 0.1 M glycine pH 2.5 and SDS sample buffer proving most effective **(Figure 3.16A)**. However, SDS sample buffer could not be used for preparing a MS sample as it would not be appropriate for an in-solution protease digest prior to MS analysis as it reduces all proteins.

As an alternative strategy, the 3X FLAG peptide, a synthetic peptide of 23 amino acids, containing three repeats of the Tyr-Lys-Xaa-Xaa-Asp motif recognised by the anti-FLAG monoclonal M2 antibody, was used to competitively elute E4bp4 under native conditions. This elution method was compared to 0.1 M glycine pH 2.5, using two concentrations of peptide and was performed at a number of salt concentrations as higher concentration of salt can influence the competition effects of the 3X FLAG peptide **(Figure 3.16B)**. Equivalent samples were compared and FLAG-E4bp4 was purified under all conditions, but the most FLAG-E4bp4 protein was purified using 3X FLAG peptide in the presence of 300 mM NaCl **(Figure 3.16B)**. Increasing the concentration of 3X FLAG peptide from 150 ng/µl to 300 ng/µl did not seem to affect the elution **(Figure 3.16B)**.

Using these elution conditions, a larger sample of FLAG-E4bp4 was purified and analysed by both WB and also on a coomassie gel **(Figure 3.17)**. The sample contained large amounts of E4bp4, including numerous higher molecular weight SUMO modified forms **(Figure 3.17)**. When the WB was probed for SUMO2/3, the same E4bp4-SUMO bands were observed in the purified sample. However, when probing with an anti-SUMO1 antibody these same bands were not detected **(Figure 3.17)**. This again confirmed that E4bp4 is preferentially modified by SUMO2/3. Analysing a fraction of the purified sample on the SDS-PAGE gel, stained with Imperial protein stain, showed the E4bp4 protein had been highly purified, with only a small number of contaminating proteins in the sample **(Figure 3.17)**.

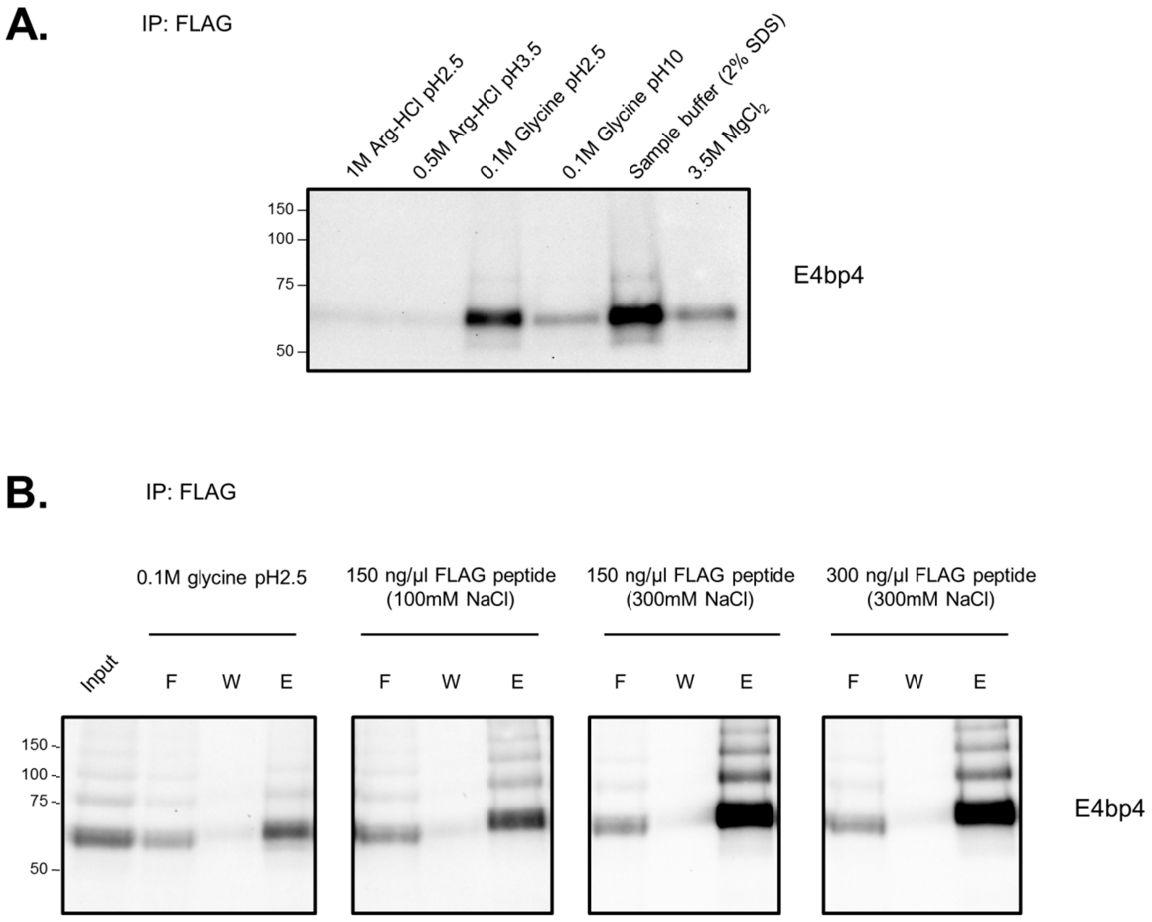


Figure 3.16 – Optimisation of FLAG-E4bp4 purification for mass spectrometry analysis

For mass spectrometry analysis a purified sample was required that was eluted under non-denaturing conditions, to allow for simple downstream processing. A) FLAG-E4bp4 was immunoprecipitated using anti-FLAG M2 affinity resin and the purified protein was eluted from the gel using a variety of methods. An equal quantity of each elution was compared by SDS-PAGE and blotted for the presence of E4bp4. B) The 3X FLAG peptide was analysed as a method for elution of FLAG-E4bp4. Four equivalent FLAG-E4bp4 immunoprecipitation samples were prepared, with the flow through and initial wash of each sample was retained for analysis. Elution fractions were collected from each sample using different concentrations of 3X peptide, as well as varying concentrations of NaCl. After elution fractions were collected, Laemmli sample buffer was added to each resin to elute off any material still bound to the resin. Equivalent samples of each elution were separated by SDS-PAGE and blotted for the presence of E4bp4. The same input was used for all the samples and is shown once. I = Input, F = Flow through, W = Wash, E = Elution, S = Sample buffer elution. Molecular weight markers show on the left of each WB in kDa.

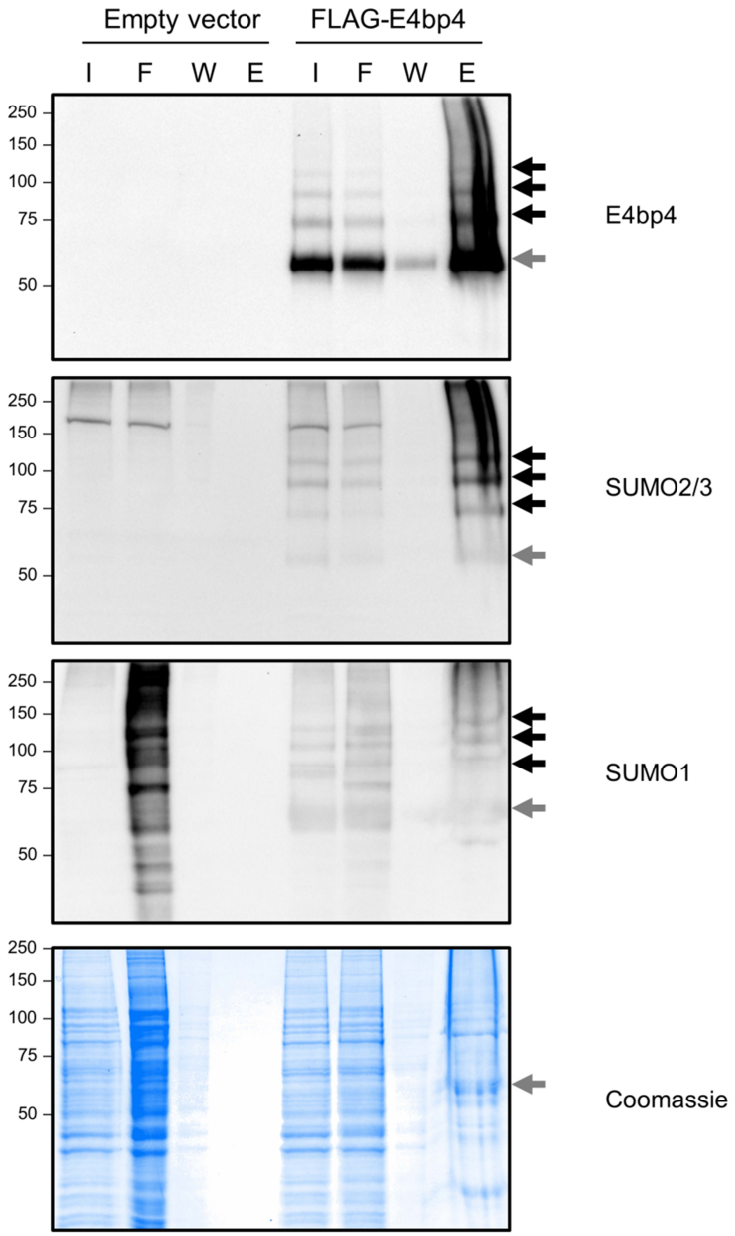
The purified E4bp4 sample was trypsin digested for 8 hours and sequentially digested with Glu-C at pH 8 and at RT. At this pH the enzyme digested after both glutamic acid and aspartic acid residues, whereas at pH 4 Glu-C only digests after glutamic acid residues. This digest strategy aimed to produce SUMO isopeptides with a sequence of GGTQQQFV,

attached to the central modified lysine residue. The digested peptides were purified and analysed by LC-MS/MS using an LTQ Velos Pro linear ion trap (Thermo Scientific).

The predicted masses for a range of E4bp4 SUMO-conjugated peptides were calculated based on predicted trypsin and Glu-C digest sites **(Table 3.3)**. The MS data was scanned for the presence of any of the predicted E4bp4-SUMO peptides **(Table 3.3)** and a peptide with a mass of 853.51 Da was detected, corresponding to a modified form of FPVIKQE **(Figure 3.18)**. Tandem mass spectra for this peptide confirmed the presence of a SUMO2/3 modification at K219 **(Figure 3.18)**. The spectra did contain a number of unidentified peaks, some of which could be as a result of fragmentation of the GGTQQQFV isopeptide 'tail', but some may also have resulted from fragmentation of a contaminated peptide **(Figure 3.18)**. To confirm the identification of the K219 SUMO modified peptide, a synthetic peptide was designed with the same predicted amino acid sequence as the modified K219 peptide following trypsin and Glu-C sequential digest **(Figure 3.19A)**. The MS/MS spectra of the recombinant peptide was analysed and showed nearly all the same ion peaks as the K219 modified peptide identified in the FLAG-E4bp4 sample **(Figure 3.19B)**. This confirmed K219 as an endogenous site of E4bp4 SUMOylation. Despite this, it was not possible to confirm SUMO modification at the other putative modification sites, K10 and K337. Further work is required to fully ascertain the modification status of these sites.

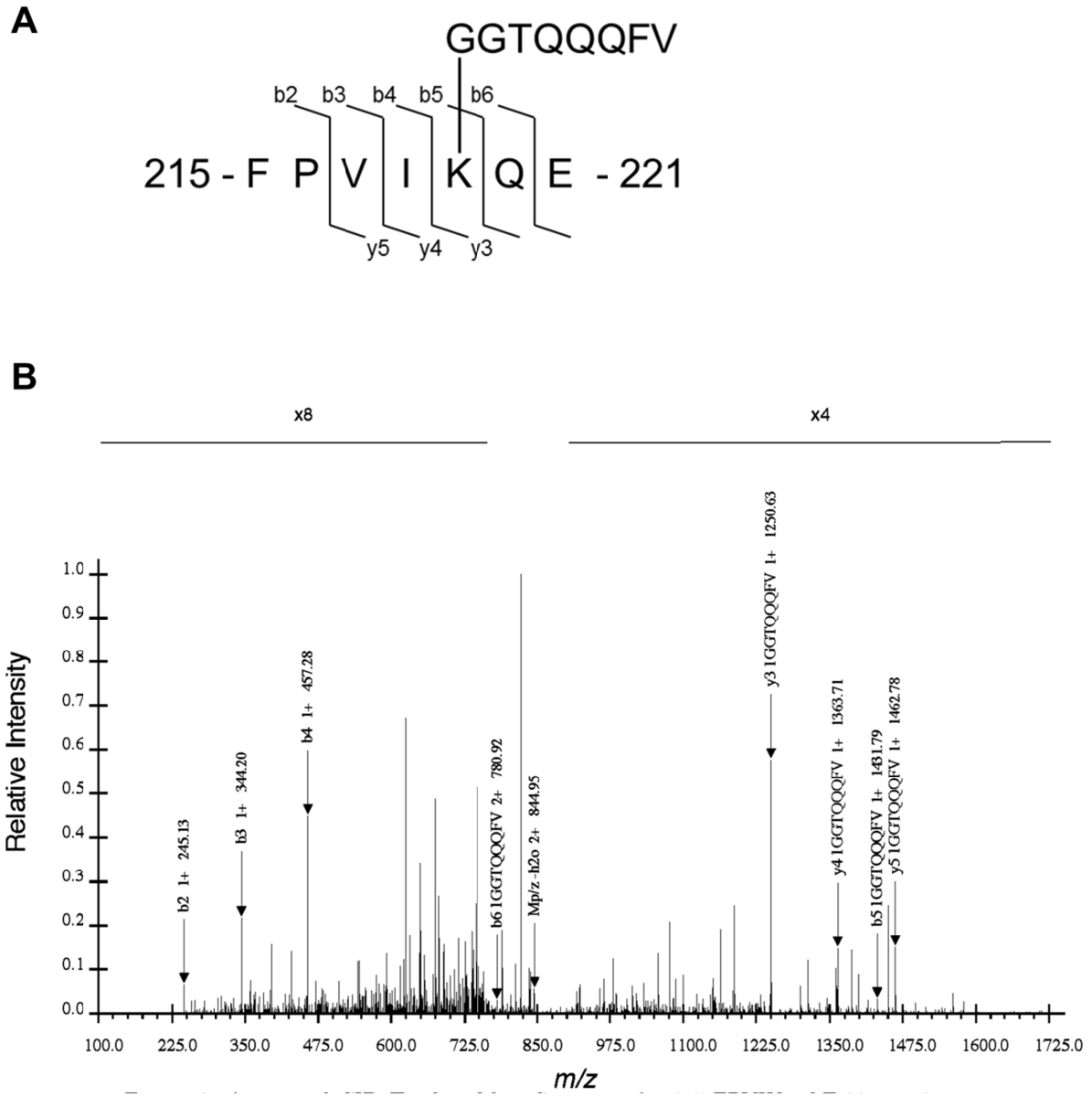
Table 3.3 – Predicted masses of E4bp4 SUMO modified peptides

Modified site	E4bp4 peptide	Peptide (m/z) (2+ charge)	Peptide plus GGTQQQFV (m/z) (2+ charge)	Peptide plus GGTQQQFV (m/z) (3+ charge)
K10	MQTIKK	374.77	797.52	531.99
K116	NATLKAE	373.75	796.49	531.33
K219	FPVIKQE	430.80	853.54	569.37
K337	AMQVKVE	402.77	825.52	550.68
K394	WSLKSE	375.19	797.89	532.26



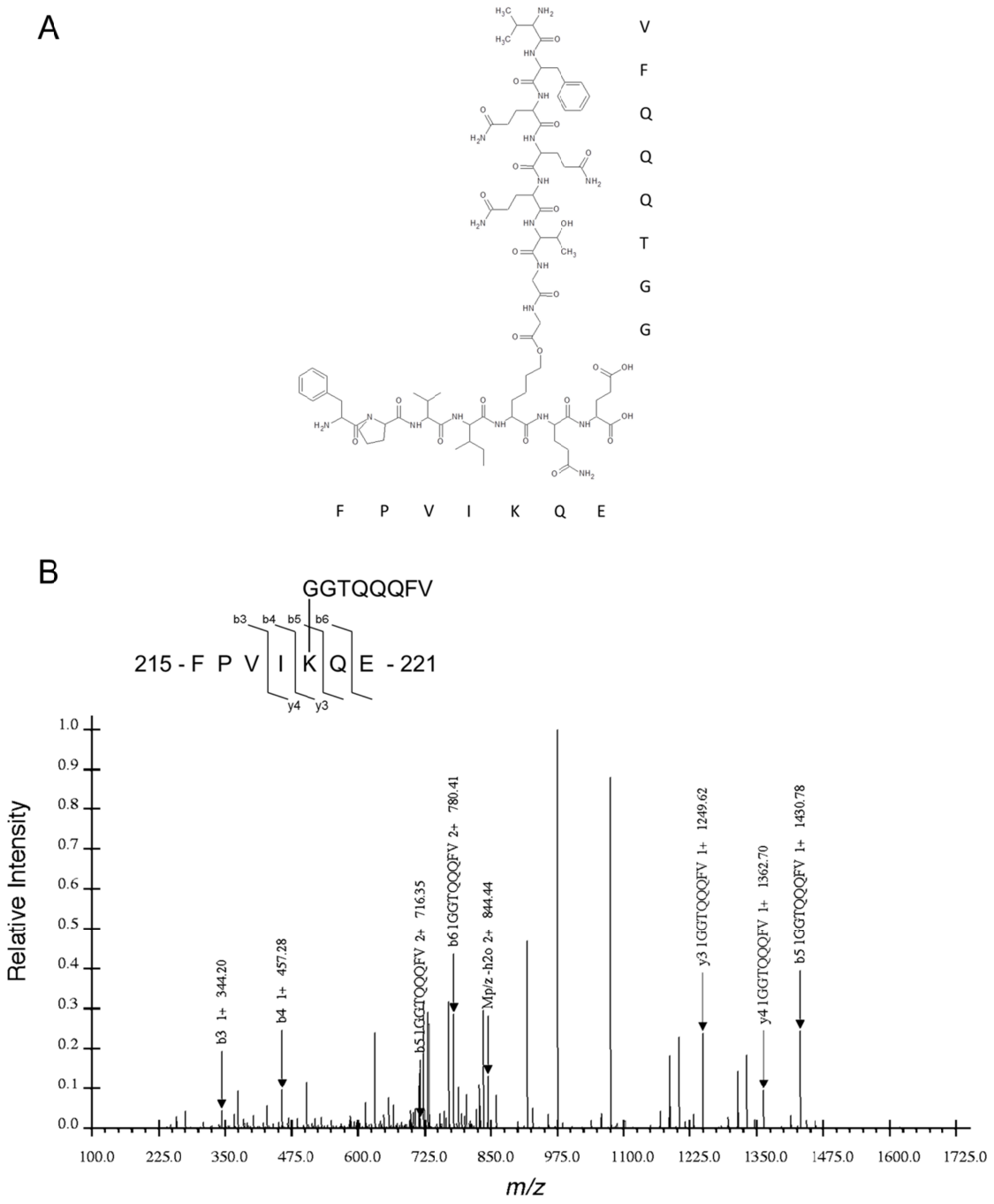
3.17 – Purification of E4bp4 for mass spectrometry analysis

FLAG-E4bp4 was expressed in 293T cells and immunoprecipitated using anti-FLAG M2 affinity resin and the purified protein was competitively eluted using 3X FLAG peptide. An equivalent sample was also prepared with cells transfected with an empty vector. Elution fractions were combined and concentrated and 10% of the purified sample was analysed on a WB and on an SDS-PAGE gel stained with Imperial protein stain. WB was probed for the presence of E4bp4, SUMO1 and SUMO2/3. I = Input, F = Flow through, W = Wash, E = Elution. Grey arrows highlight unmodified E4bp4 and black arrows highlight higher molecular weight, SUMO-modified forms of E4bp4. Molecular weight markers show on the left of each WB in kDa.



3.18 – Confirmation of E4bp4 SUMO2/3 modification at lysine 219

FLAG-E4bp4 was expressed in HEK-293T cells and purified by immunoprecipitation with anti-FLAG M2 affinity resin and the purified protein competitively eluted using 3X FLAG peptide. Purified E4bp4 was sequentially digested with trypsin and Glu-C to reduce the size of modified E4bp4 peptide and also the isopeptide SUMO side chain. This digest strategy produced a SUMO2/3 isopeptide fragment of GGTQQQFV, attached to the modified lysine side chain. Tandem mass spectra was acquired with an electrospray ionization LTQ/Orbitrap mass spectrometer. A) Schematic representation of the identified peptide, with detected *b* and *y* ions labelled and amino acid positions within the E4bp4 protein. The GGTQQQFV isopeptide sidechain is present on b5 and b6 ions, as well as all y ions. B) Annotated CID tandem mass spectra of the +2 ion at *m/z* 853.95, confirming SUMO2/3 modification of E4bp4 at K219. Magnification of certain regions of the spectra are highlighted.



3.19 – Analysis of synthetic E4bp4-SUMO2/3 peptide standard

A) Design of synthetic peptide based on observations of E4bp4-SUMO peptide *in vivo*, with E4bp4 backbone of FPVIKQE modified on the lysine side chain with a peptide tag of GGTQQQFV. B) Annotated CID tandem mass spectrum of the +2 ion at *m/z* 853.44, confirming SUMO2/3 modification of E4bp4 at K219. Tandem mass spectra acquired with an electrospray ionization LTQ/Orbitrap mass spectrometer from purified FLAG-E4bp4. FLAG-E4bp4 was expressed in HEK-293T cells and purified by immunoprecipitation with anti-FLAG M2 affinity resin and the purified protein competitively eluted using 3X FLAG peptide. Purified E4bp4 was sequentially digested with trypsin and Glu-C to reduce the size of modified E4bp4 peptide and also the isopeptide SUMO side chain. This digest strategy produced a SUMO2/3 isopeptide fragment of GGTQQQFV, which is present on b5 and b6 ions, as well as all y ions.

3.6 Removing SUMOylation sites of E4bp4 causes increased NK cell production

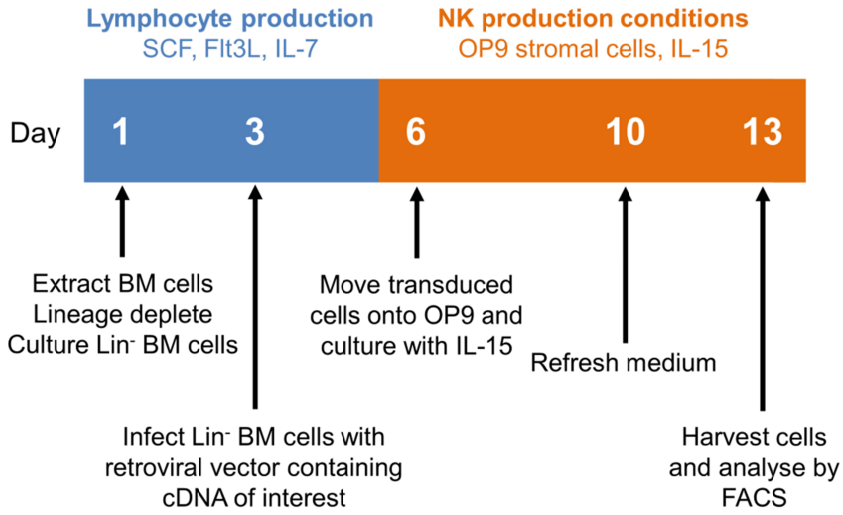
SUMOylation can affect the function of TFs in a number of different ways; such as their localisation within the cell, their interactions with other proteins and ultimately their transcriptional activity (Gill, 2005). SUMOylation of TFs is important for a number of cellular processes, for example, SUMOylation of Bcl11b is required for the later stages of T-cell development (Zhang et al., 2012), whilst SUMOylation of TFs Myc and MITF is linked to metastatic cancer development (Bertolotto et al., 2011) (Kessler et al., 2012).

E4bp4 is an essential TF for numerous cellular processes, but arguably, the most striking phenotype of *E4bp4*$^{-/-}$ mice is their complete lack of cNK cells (Gascoyne et al., 2009). E4bp4 is specifically required at the commitment point for NK cell development to allow CLPs to develop into NK cell committed progenitors (Male et al., 2014). Interestingly, E4bp4 is also a limiting factor in NK cell development because over expression of E4bp4 in developing WT progenitors increases the number of NK cells that develop *in vitro* (Gascoyne et al., 2009) (Male et al., 2014).

As E4bp4 has multiple SUMO modifications, it was postulated whether these were required for its ability to promote NK cell development. In order to answer this question, an NK cell development assay, which had previously been developed in the Brady laboratory, was utilised (Gascoyne et al., 2009) (Male et al., 2014). Lineage negative (Lin$^-$) BM cells were isolated from *E4bp4*$^{-/-}$ mice and cultured in lymphocyte inducing conditions before transduction with a retrovirus, expressing a gene of interest **(Figure 3.20)**. The transduced cells were cultured on stromal OP9 cells, in the presence of IL-15, to promote NK cell development **(Figure 3.20)**. *E4bp4*$^{-/-}$ Lin$^-$ BM cells cannot develop into NK cells due to their lack of E4bp4, but by transducing the cells with a retrovirus expressing E4bp4, NK cell development can be rescued (Gascoyne et al., 2009). Using this assay the activity of E4bp4 with and without SUMO modifications was compared to determine if the PTMs are essential for its function.

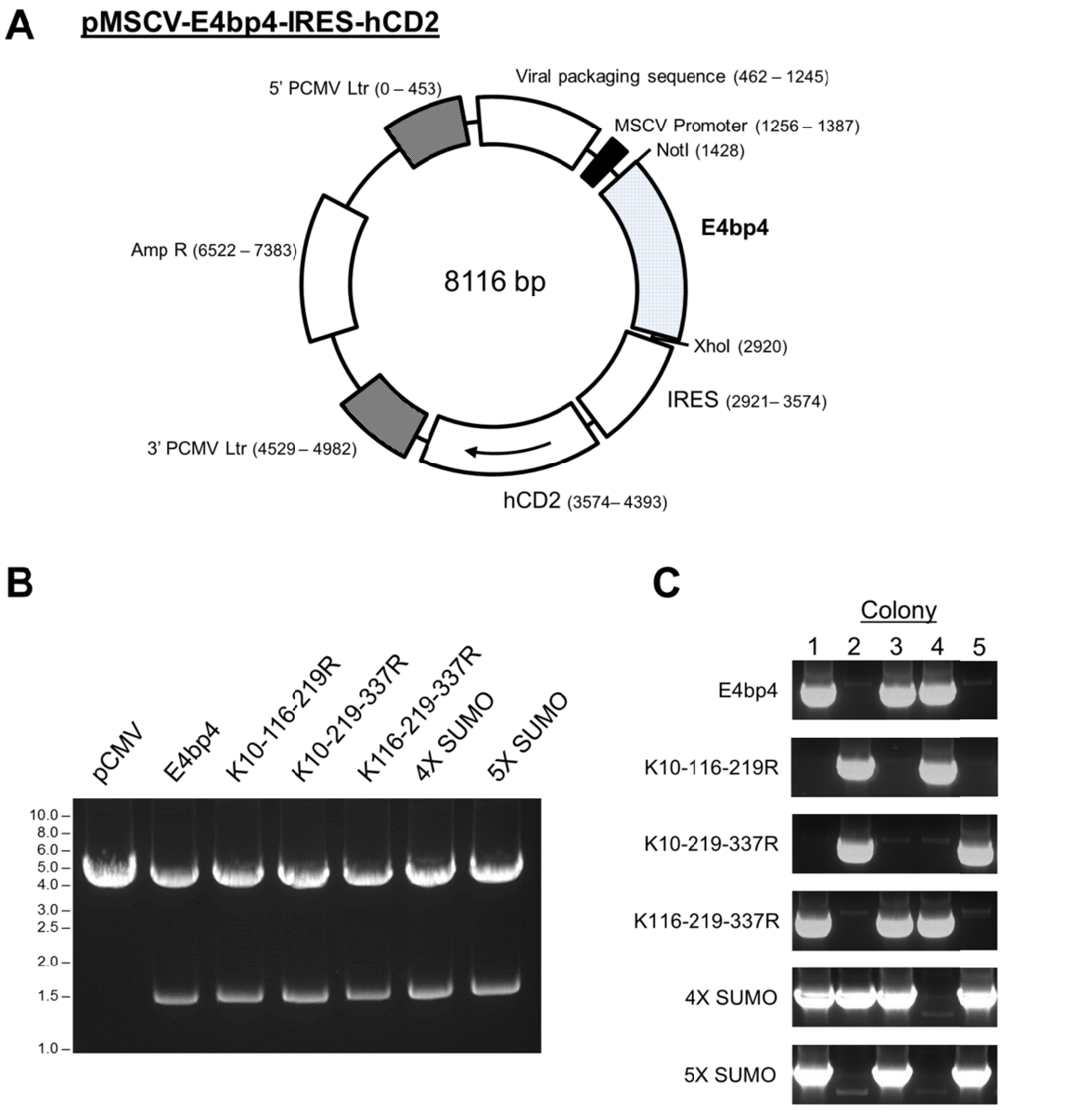
Initially, the E4bp4 mutant cDNAs which lacked SUMOylation sites were cloned into the pMSCV-IRES-hCD2 retroviral expression vector **(Figure 3.21A)**. Five different mutant cDNAs were cloned into MSCV: K10-116-219R, K10-219-337R, K116-219-337R, 4X SUMO, 5X SUMO. These mutants were cloned as they were found to have very little or no SUMO modifications. Each cDNA was cut out of the pCMV-script expression vector using NotI and XhoI restriction enzymes **(Figure 3.21B)**. They were then purified before being ligated into

the pMSCV-hCD2 vector. Ligation products were transformed into chemically competent *E. coli* and successfully transformed colonies were identified by colony PCR **(Figure 3.21C)**. Each pMSCV-IRES-hCD2 construct was sequenced prior to use in the NK cell development assay to ensure that the correct mutations were maintained at the SUMO modification sites of E4bp4.



3.20 – NK cell production from lineage negative bone marrow cells, with transduction

Schematic representation of the assay system used to produce NK cells *in vitro* from Lin^- BM cells extracted from mouse BM. During the first of two phases (blue), lymphocytes production was stimulated from Lin^- BM cells. These were subsequently infected with a retroviral MSCV vector expressing the cDNA of interest and additionally hCD2 to allow identification of transduced cells. During phase two (orange), transduced cells were transferred to OP9 stromal cells and cultured with IL-15 to promote NK cell development. After 13 days cells were harvested and analysed by flow cytometry for the presence of transduced cells ($hCD2^+$) that have developed into NK cells ($NK1.1^+$ $CD19^-$). SCF – stem cell factor; Flt3L – Fms-related tyrosine kinase 3 ligand.



3.21 – E4bp4 SUMO mutants were cloned into pMSCV-IRES-hCD2 vector

A) Schematic representation of pMSCV-IRES-hCD2, with E4bp4 cDNA cloned between the promoter and the internal ribosome entry site (IRES). Long terminal repeats (Ltr) allow integration of the DNA into the genome of transduced cells and ampicillin resistance (Amp R) gene allows for amplification of the vector in *E. coli*. Position of each element within the vector is shown in bp. B) E4bp4 mutant cDNAs were digested out of the pCMV vector using NotI and XhoI restriction enzymes. Digest products were analysed on a 1% agarose gel (molecular weight markers shown in kb) and insert fragments (~1.48 kbp) were cut out of the gel and purified to then be ligated into the pMSCV-IRES-hCD2 vector. C) E4bp4 SUMO mutant cDNA were ligated into the pMSCV-IRES-hCD2 vector and ligation products were transformed into competent *E. coli*, successful transformants were analysed by colony PCR for the presence of E4bp4; colony PCR reactions were run on a 1% agarose gel. LMP5 and LMP3 primers were used during colony PCR and amplified across NotI/XhoI cloning site.

The K10-116-219R, 4X SUMO and 5X SUMO mutants were all expressed in *E4bp4*$^{-/-}$ Lin^{-} BM cells and, following culture in IL-15, the cells were analysed by flow cytometry to determine the number of transduced cells that had developed into NK cells. Using the IRES hCD2 expressed by the retrovirus it was possible to analyse only the transduced cells **(Figure 3.22)**. The OP9 stromal cells were excluded from the analysis as they are GFP positive **(Figure 3.22)**. Strikingly, there were consistently more than double the number of NK cells produced in the samples transduced with the SUMO mutants, compared to that of WT E4bp4 **(Figure 3.22)**. No NK cells developed from the *E4bp4*$^{-/-}$ Lin^{-} BM cells that had been transduced with an empty retrovirus, as previously reported **(Figure 3.22)**. The phenotype of the NK cells that developed in the SUMO mutant samples was similar to the WT E4bp4 sample, with a small percentage (4-7%) of the cells developing into mature NKs and expressing CD11b **(Figure 3.22C)**. The lack of SUMO modifications clearly appeared to promote the activity of E4bp4 and induce greater numbers of NK cells to develop.

A similar experiment was performed using WT Lin^{-} BM cells that were transduced and cultured in the same conditions **(Figure 3.20)**. Transduction of these cells with a retrovirus expressing E4bp4 had previously been shown to increase NK cell production *in vitro* (Gascoyne et al., 2009). Transduction of WT Lin^{-} BM cells with the K10-116-219R SUMO and 5X SUMO mutants again resulted in a significant increase in the production of NK cells when compared to WT E4bp4 **(Figure 3.23)**. This further confirms that removing SUMO modifications from E4bp4 promotes its activity, which is helping to drive the production of NK cells. Additionally, this assay demonstrates how E4bp4 is a limiting factor for NK cell development, which has potentially important implications for the development or expansion of human NK cells for clinical use.

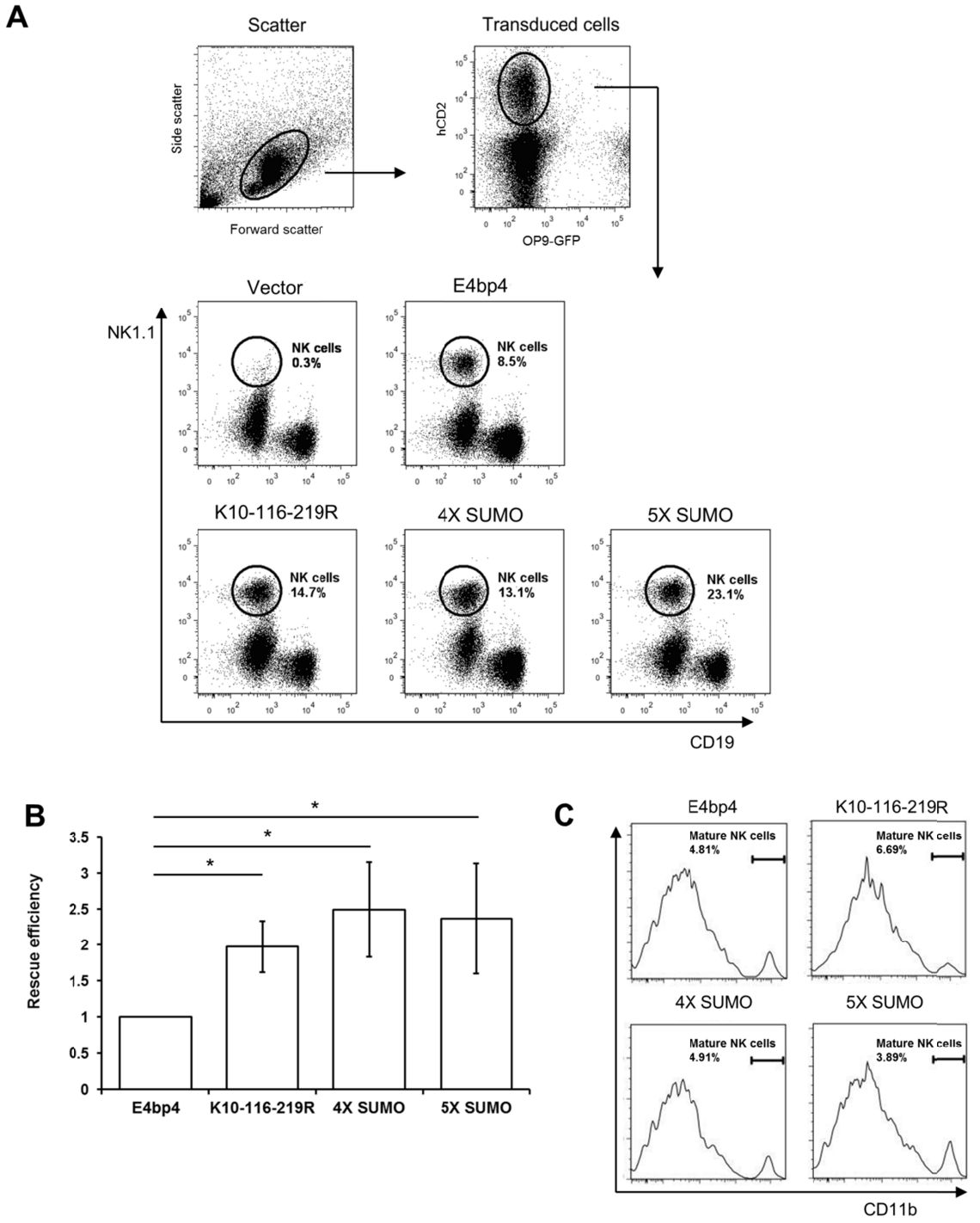


Figure 3.22 – Removing SUMO modification sites from E4bp4 increases *in vitro* NK cell production from *E4bp4*$^{-/-}$ Lin^{-} BM

Lin^{-} BM cells were isolated from *E4bp4*$^{-/-}$ mouse BM and cultured for two days in the presence of IL-7, Flt3 ligand and SCF before retroviral transduction with a MSCV-IRES-hCD2 construct (expressing E4bp4, E4bp4 3X SUMO, E4bp4 4X SUMO or E4bp4 5X SUMO). Transduced cells were cultured for three days, before being transferred onto OP9 stromal cells and cultured in the presence of IL-15 for 7 days. A) Flow cytometry was used to analyse cultured cells to identify live lymphocytes that were positive for hCD2 (marker of transduced cells) and subsequently to examine their expression of NK1.1 and CD19. B) Efficiency of the rescue affected by the different E4bp4 SUMO mutants, normalised to that of WT E4bp4. Data are representative of four independent experiments. Error bars show SEM and statistical differences were calculated using paired Student's t-tests, * = $P < 0.05$. C) NK cell populations were additionally analysed for their percentage of Cd11b^{+} mature NK cells, representative of two independent experiments.

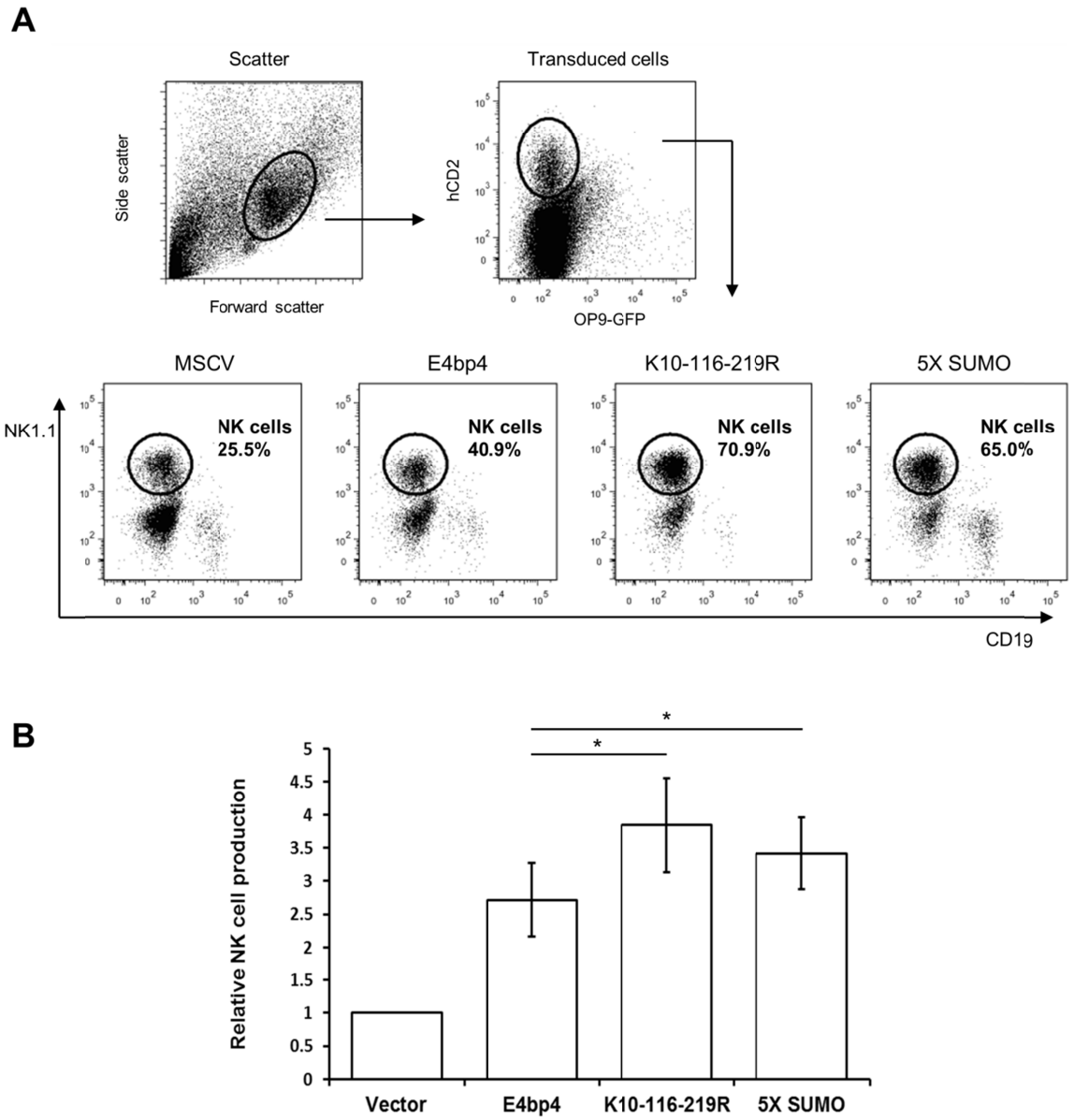


Figure 3.23 – Removing SUMO modification sites from E4bp4 increases *in vitro* NK cell production from WT Lin$^-$ BM

Lin$^-$ BM cells were isolated from WT mouse BM and cultured for two days in the presence of IL-7, Flt3 ligand and SCF before retroviral transduction with a MSCV-IRES-hCD2 construct (expressing E4bp4, E4bp4 3X SUMO, or E4bp4 5X SUMO). Transduced cells were cultured for three days, before being transferred onto OP9 stromal cells and cultured in the presence of IL-15 for 7 days. A) Flow cytometry was used to analyse cultured cells to identify live lymphocytes that were positive for hCD2 (marker of transduced cells) and subsequently to examine their expression of NK1.1 and CD19. B) Relative level of NK cell production by the different E4bp4 SUMO mutants and WT E4bp4 normalised to empty vector. Data are representative of four independent experiments. Error bars show SEM and statistical differences were calculated using paired Student's t-tests, * = $P < 0.05$.

3.7 Discussion

SUMOylation is a highly dynamic PTM that is critical for many cellular functions including cell cycle progression, DNA repair and transcription (Hay, 2005). Many TFs have been discovered as targets for SUMO modification and this is often associated with enhanced transcriptional repressive activity (Filtz et al., 2014) (Gill, 2005). However, the underlying mechanism by which SUMOylation affects transcriptional activity of most TF's is yet to be fully determined. E4bp4 is an essential TF, involved in numerous cellular processes, many involving the immune system, and this work has shown for the first time that it is a target of SUMOylation.

Using HeLa cells expressing 6His-SUMO, E4bp4 was discovered to be SUMOylated *in vivo* by SUMO1 and to a much greater extent by SUMO2/3. Many proteins have been found to be SUMOylated more prominently by one SUMO isoform over the other, but the reasons for this are still unclear (Vertegaal et al., 2006). The same modified forms of E4bp4 were observed whether E4bp4 was the target of the pull down or if it was for 6His-SUMO. Up to four different modified forms of E4bp4 were present on the WBs, suggesting E4bp4 either has multiple SUMO modification sites and/or multiple modifications are occurring at the same lysine residue where they have formed poly-SUMO chains (Knipscheer et al., 2007). As E4bp4 has multiple functions in different cell types, it is possible that it has a variety of SUMOylation states, which vary depending upon the context in which it is functioning.

Initial bioinformatic analysis revealed that E4bp4 had five highly conserved potential SUMOylation sites and also suggested that three of these sites (K10, K219 and K337) were the most likely to be modified. However, it should be noted that the prediction of SUMOylation sites does not take into account the three-dimensional structure of the protein. Predicted SUMOylation sites could be identified that are not exposed at the surface of the protein and therefore *in vivo* could not be modified (Da Silva-Ferrada et al., 2012). To identify which of the predicted sites was modified lysine-arginine mutations were made to abolish the potential for SUMOylation at those sites. Individual mutants were created as well as forms of E4bp4 that lacked three, four or all five putative SUMO modification sites. Using a similar approach, the SUMO modification sites of the TFs XBP1 (Chen and Qi, 2010), FOG-2 (Perdomo et al., 2012) and NRL (Roger et al., 2010) were identified. For example, Perdomo *et al.* created ten different mutants of FOG-2 to identify the four SUMO modified lysine residues in the protein (Perdomo et al., 2012).

When these mutants were all expressed in the 6His-SUMO2 cell line, K219R was immediately identified as a key site for modification, as it was the only single site mutant that had any perturbation of the SUMOylated forms of E4bp4. However, this mutation alone was not sufficient to remove all the higher molecular weight forms of E4bp4. When combined with K10R and K337R, all SUMO modification was abolished, as was the case for the 5X SUMO mutant and 4X SUMO mutant. This clearly demonstrated that E4bp4 is SUMO modified at K10, K219 and K337, with K219 being the most prominent site of modification. Interestingly, these are the same three sites that were predicted as the most likely sites for E4bp4 SUMOylation in the initial bioinformatic analysis. These observations suggest that there may be a hierarchy or some synergy between the modification sites as three must be mutated in order to remove all SUMO modifications. However, mutations of K10 and K337 individually do not seem to affect SUMOylation at all. In many instances SUMOylation results in the addition of single SUMO moieties onto an acceptor Lys residue, yet the formation of polySUMO chains of SUMO2/3 has been observed (Tatham et al., 2001). It may be that a polySUMO chain occurs at K219, however, it was not possible to determine if this was the case using this assay.

Currently there is no structural information available about E4bp4, indeed no full three dimensional structure has been determined for any bZIP TF. The structure of the DNA-binding, basic region and leucine zipper region, of the TF CREB was previously solved (Schumacher et al., 2000), and this is likely to be similar to the N-terminus of E4bp4. However, the structure of the majority of the E4bp4 protein is unknown and it is therefore difficult to know the distances between the E4bp4 SUMOylation sites in three dimensional space. The three SUMOylation sites (K10, K219 and K337) are not found in the DNA binding or dimerisation domains of E4bp4 and both K219 and K337 are found within the hypothetical transcription repression domain. This domain was identified through mutational studies as essential for the transcription activity of E4bp4 (Cowell and Hurst, 1994). This suggests SUMOylation at these sites may influence the transcriptional activity of E4bp4, potentially through its interactions with other transcriptional cofactors, rather than influencing its ability to bind DNA.

Identical observations were made with all mutants when they were expressed in both the 6His-SUMO2 and 6His-SUMO3 cell lines. This was expected as they share 95% sequence homology (Hay, 2005). However, when expressed in the 6His-SUMO1 cells, significantly less modified E4bp4 was detected, demonstrating that E4bp4 is predominantly modified by SUMO2/3. Two modified forms of E4bp4 were observed but these were abolished by the K219R mutation and also affected by the K337R mutation. Nearly all of the multi-site

mutants lacked any detectable SUMO1 modification. This was vital for function analysis of these mutants, as it was clear that the K10-219-337R mutant, 4X SUMO mutant and 5X SUMO mutant were not modified by any SUMO isoform. These mutants could be compared to WT E4bp4 to determine if SUMO modifications are important for E4bp4's function.

To unambiguously establish which E4bp4 residues were modified by SUMO, high accuracy MS was used to specifically identify SUMOylated E4bp4 peptides. Although MS has been used extensively to identify specific modification sites of small PTMs, such as phosphorylation, identifying endogenous SUMOylation sites is challenging (Da Silva-Ferrada et al., 2012) (Jeram et al., 2009) (Witze et al., 2007). SUMOylation results in the addition of a large peptide to the ε-amino group of modified lysine residues, creating a branched polypeptide. Identification of large branched peptides by tandem MS is complex and in some instances unfeasible as peptides of interest are too large to detect. The low stoichiometry of the SUMO-modified forms of a specific protein, and the ease with which the modifications can be removed by proteases, adds to the difficulties of direct MS analysis (Jeram et al., 2009) (Witze et al., 2007). Subsequently, few studies have been able to identify SUMO-modified peptides of endogenously modified protein by MS. However, one technique has recently been described, involving the use of monoclonal antibodies against the different SUMO isoforms (Becker et al., 2013). Many MS studies analysing SUMOylation targets have involved the over expression of SUMO in order to aid the detection of modified peptides (Da Silva-Ferrada et al., 2012). Additionally, a number of studies have exploited the link between SUMOylation and the proteasome; increasing SUMOylation levels inside cells by blocking the proteasome with inhibitors, such as MG132 (Schimmel et al., 2008).

In this study, both of those strategies were avoided so that only endogenous E4bp4 SUMOylation sites were identified. The expression of E4bp4 in 293T cells was optimised, as well as the purification process to allow the elution of E4bp4 in non-denaturing conditions. This subsequently allowed for in-solution digestion of the protein so that no loss of SUMOylated material was incurred. Analysis of purified E4bp4 by WB confirmed the presence of SUMO2/3 conjugated forms of E4bp4, with limited amounts of SUMO1 modification detected. To aid the identification of modified E4bp4 peptides, a novel trypsin/Glu-C double digest strategy was employed to reduce the size of both backbone E4bp4 peptides as well as the size of the SUMO isopeptide tag. The digested peptides were analysed by MS and the data was searched for the presence of SUMO2/3 modified E4bp4 peptides. The analysis of tandem mass spectra revealed a SUMO2/3 modification at K219 of E4bp4. As the spectra for the peptide contained a number of unidentified ions, and the trypsin/Glu-C digest strategy was novel, a recombinant peptide was acquired with the same

branched structure as the K219 SUMO modified peptide. The MS spectra of this peptide was analysed and matched that of the endogenous K219 peptide, confirming its identification.

E4bp4 was observed to have more than one modified residue in 6His-HeLa cells and mutational studies demonstrated that both K10 and K337 are sites of modification. However it was not possible to identify these modified peptides by MS. The main difficulty with this analysis was the low abundance of the SUMO-modified E4bp4 peptides and the large amount of unmodified peptides in the sample. Further optimisation is required to prepare a purified E4bp4 sample with a higher proportion of SUMO-modified peptides. This could involve either enrichment of all SUMO-modified proteins prior to FLAG-E4bp4 purification or the use of size exclusion chromatography to separate the larger modified forms of E4bp4 from the unmodified ones.

Two recent SUMO proteomic studies used mutated versions of SUMO1 and SUMO2 that produce a simple -GG motif following trypsin digest, which significantly aided identification of SUMOylated peptides (Impens et al., 2014) (Tammsalu et al., 2014). Interestingly, both of these studies identified the same modified E4bp4 SUMO modified peptide as identified here, further confirming K219 as a site an endogenous site of SUMO modification (Impens *et al* 2014) (Tammsalu et al., 2014). The study by Impens *et al.* also was able to show that K219 is modified by SUMO1, but more preferentially by SUMO2/3, again confirming the results observed here (Impens et al., 2014). K219 is therefore a critical SUMOylation site for E4bp4 and is possibly more abundant than any other E4bp4-SUMO modified peptide, as these were not identified in the proteomic studies.

The E4bp4 mutants that lacked SUMO modifications, in particular the 5X SUMO mutant which lacked all potential SUMOylation sites, provided a tool to determine if SUMOylation was required for the function of E4bp4. Gascoyne *et al*. previously demonstrated that *E4bp4*$^{-/-}$ Lin^{-} BM cells could not develop into NK cell *in vitro*, unless they were retrovirally transduced with a vector expressing E4bp4 (Gascoyne et al., 2009). Using this *in vitro* NK cell development assay, WT E4bp4 was compared to the K10-116-219R mutant, the 4X SUMO mutant and the 5X SUMO E4bp4 mutant for their ability to rescue NK cell development. All three E4bp4 SUMO mutants consistently increased the production of NK cells, compared with WT E4bp4. On average, the 4X SUMO and 5X SUMO mutants produced 2.5 times more NK cells. The NK cells produced by the SUMO mutants also had a similar phenotype to WT E4bp4 with regards to CD11b expression. This provided the first clear evidence that SUMOylation was able to regulate the activity of E4bp4 and that manipulation of PTMs could have such a dramatic effect on a complex biological process.

This notion was further support by evidence using WT Lin$^-$ BM cells in the *in vitro* NK cell assay and transducing them with WT E4bp4 or 5X SUMO mutants. Expression of E4bp4 in these cells has previously been shown to increase the production of NK cells, compared to an empty vector/untransduced cells, suggesting that E4bp4 is a limiting factor during NK cell development (Gascoyne et al., 2009) (Male et al., 2014). Both the K10-116-219R mutant and 5X SUMO mutant were able to consistently promote greater levels of NK cell production from the WT Lin$^-$ BM cells than WT E4bp4. In fact both mutants caused more than 65% of the transduced cells to become NK cells. SUMOylation therefore negatively regulates the activity of E4bp4 and by removing these modifications its activity can be significantly increased. This is the first study to demonstrate that SUMOylation is important for the development of NK cells.

In future studies it would be of significant interest to determine if the observations made in this study could be repeated *in vivo*. To test this hypothesis, *E4bp4*$^{-/-}$ Lin$^-$ BM cells could be transduced with E4bp4 or E4bp4 5X SUMO and then transferred into lethally irradiated mice. The ability of E4bp4 and the 5X SUMO mutant to promote NK cell development would then be compared by analysing NK cells numbers in their BM, spleen and blood of each mouse 3 - 4 weeks post transfer. It would be expected that the cells transduced with the 5X SUMO mutant would produce greater numbers of NK cells in all host compartments, compared to WT E4bp4.

TFs are generally expressed at relatively low levels within cells, making their direct quantitative analysis challenging (Vaquerizas et al., 2009). Even in NK cells where E4bp4 has important regulatory functions, the protein is only expressed at a low level and is barely detectable by WB (Brady laboratory, unpublished). Therefore one limitation of this study was that all the experiments on the SUMOylation of E4bp4 involved recombinant over expression of the protein. Ideally these observations would be supported by work performed on the endogenous E4bp4 protein, particularly within NK cells, however this was not feasible. The problem of detecting endogenous SUMOylation of E4bp4 is also enhanced by the fact that only a small fraction of any SUMO substrate is generally modified at steady state. Future work involving high affinity monoclonal antibodies raised against E4bp4 may make isolation of the endogenous E4bp4 protein possible and its PTMs could then be analysed.

Many TFs have now been identified to be SUMOylated and, whilst many of these require SUMO for their activity, a number have been identified that like E4bp4 are more active in the absence of SUMO modifications. These include MITF (Bertolotto et al., 2011), XBP1 (Chen and Qi, 2010), MAML1 (Lindberg et al., 2010), FOG-2 (Perdomo et al., 2012) and MafA

(Shao and Cobb, 2009). The transcriptional activities of MAML1, MafA and XBP1 are all increased when their SUMOylation sites are mutated from Lys to Arg, demonstrating that SUMOylation can repress the activity of a wide variety of TFs. However, these studies are not able to demonstrate how the SUMOylation of these particular TFs can affect a whole biological process, as has been shown here for E4bp4. The only study to date that is able to show such a dramatic effect on the phenotype of developing cells was by Bertolotto *et al*. who examined SUMOylation of MITF (Bertolotto et al., 2011). Their study demonstrated that in the absence of SUMOylation, MITF is a more potent transcriptional activator in melanoma and renal carcinoma cells and unmodified MITF makes these cancer cells more tumourgenic, with increased migration, invasion and clonogenicity (Bertolotto et al., 2011).

Before this study, SUMOylation had not been directly linked to the development of NK cells, but a number of TFs known to be involved in regulating NK cell development have been identified as targets of SUMO modification. Ikaros has two SUMOylation sites and mutation of these affected its repression of target gene transcription and influenced its interactions with transcriptional co-repressors (Gomez-del Arco et al., 2005). Ets1 is essential for the early stages of NK cell development and also has two SUMOylated lysine residues (Ji et al., 2007). Mutation of these sites increased its transcriptional activity, suggesting that SUMO also negatively regulates its activity (Ji et al., 2007). IRF-2, which is required to produce mNK cells, has three SUMOylation sites and mutating these sites had varying effects on its transcriptional activity (Han et al., 2008). SUMOylation of IRF-2 increases it ability to inhibit IRF-1 transcriptional activity and decreases it ability to activate transcription of type I interferon genes (Han et al., 2008). However, none of these studies were able to show if the SUMO modifications of Ikaros, Ets1 and IRF-2 were critical for their activities *in vivo* and, in particular, if these modifications were required for NK cell development.

SUMOylation is however important for the development of other immune cell lineages. One recent study by Van Nguyen *et al*. demonstrated that SENP1 (SUMO protease) is essential for the development of T and B cells, as *SENP1*$^{-/-}$ HSCs isolated from fetal livers were unable to differentiate into T or B cell *in vitro*, yet the myeloid lineage development was unaffected (Van Nguyen et al., 2012). Interestingly, Van Nguyen *et al*. did not mention NK cells or innate lymphoid cells in their report, so whether the development of these cells is affected by an absence of SENP1 is still to be determined. SUMOylation was also linked to the development of T cells through its interactions with the TFs Bcl11b and JunB (Garaude et al., 2008) (Zhang et al., 2012). Bcl11b is required at three checkpoints in the development of functional T cells in the thymus and SUMOylation functions to inhibit transcriptional repression by Bcl11b (Zhang et al., 2012). SUMOylation also results in the recruitment of the

transcriptional co-activator p300 to the promoters of Bcl11b repressed genes (Zhang et al., 2012). This activity is essential to allow T cell specific genes, supressed by Bcl11b, to be expressed in immature thymocytes to help drive T cell maturation (Zhang et al., 2012). SUMOylation of JunB is required for activation of mature T cells and the regulation of cytokine genes, particularly IL-2 (Garaude et al., 2008). In the absence of SUMO modifications, JunB is not able to associate with open chromatin structures and promote the expression of IL-2 and IL-4 genes (Garaude et al., 2008).

SUMOylation of E4bp4 is an important regulatory process and provides evidence for the complex mechanisms that exist to control E4bp4 activity during cellular processes. E4bp4 has very recently been discovered to be critical for the development of other ILC subsets and T_H17 T cells, as well as NK cells, so it would now be interesting to determine if SUMOylation of E4bp4 has a role in any of these developmental processes. If SUMO modifications affect the transcriptional activity of E4bp4 is a vital question to answer and could provide an insight into the mechanism by which E4bp4 can produce such a strong phenotype. Additionally, it is not clear what other PTMs regulate E4bp4 and whether these are also significant in the NK cell development process.

CHAPTER 4

Regulation of E4bp4 by phosphorylation

4.1 Introduction

4.1.1 Regulation of transcription factors by phosphorylation

Phosphorylation is the most abundant and therefore most well studied of all PTMs, with approximately one third of all the mammalian proteome thought to be a target for phosphorylation (Mumby and Brekken, 2005). Phosphorylation involves the formation of an ester bond between a phosphate group (PO_4^{3-}) and the side chain of a polar amino acid (serine, threonine or tyrosine). Phosphorylation is rapidly reversible, with hundreds of known kinases able to modify target proteins and phosphatases able to subsequently remove the modifications (Holmberg et al., 2002). The actions of these enzymes allow biological processes to be tightly controlled and quickly manipulated in response to changes in the cellular environment (Whitmarsh and Davis, 2000). Phosphoproteomic studies have revealed how a wide variety of cellular proteins are phosphorylated, with TFs a commonly modified class of protein (Wu and Liao, 2007).

TFs can have diverse effects on the expression of different target genes, depending on cell type, the presence/absence of certain stimuli, and interactions with tissue specific cofactors and accessory proteins (Filtz et al., 2014). Phosphorylation can increase the functionality of TFs and allow them to respond to external stimuli, via signalling pathways, to induce changes in gene expression (Holmberg et al., 2002). The phosphorylation of a TF can modulate its activity in a number of ways: 1) DNA binding activity; 2) localisation; 3) protein interactions; 4) stability; 5) transcriptional activity **(Figure 4.1)**.

The DNA binding activity of the bZIP TF c-Jun is directly and indirectly regulated by phosphorylation (Papavassiliou et al., 1995). c-Jun has five phosphorylation sites, three of which are close to its N-terminal DNA binding domain and phosphorylation at these sites directly interferes with DNA binding (Papavassiliou et al., 1995). However, phosphorylation of the other two sites at the C-terminus of the protein, by JNK MAP kinase, causes a conformational change that prevents phosphorylation at the N-terminus of the protein and therefore promotes DNA binding (Papavassiliou et al., 1995).

Phosphorylation of the oncogene c-Myc regulates its stability as phosphorylation at both T58 and S62 causes it to be rapidly degraded by the proteasome. However, phosphorylation of S62 alone causes c-Myc to promote cell cycle progression and ultimately tumorigenesis (Hann, 2006). Phosphorylation at this site is caused by extracellular proliferative signals that increase the transcriptional activity of c-Myc and also promote its stability (Hann, 2006).

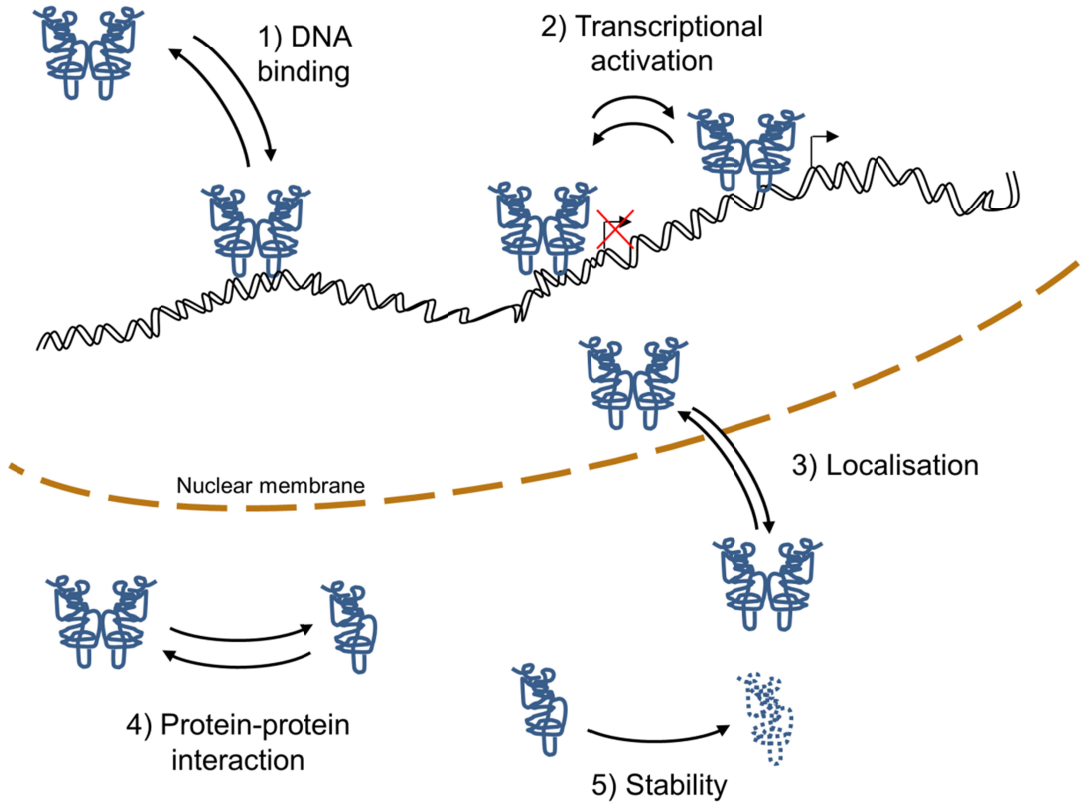


Figure 4.1 – Phosphorylation affects the function of transcription factors

Phosphorylation and dephosphorylation provide at least five different general methods to influence the functions of transcription factors (TF). The phosphorylation of a TF can be modified in response to changes in cellular environment, providing a direct link between external stimuli and gene expression. Phosphorylation can influence the DNA binding capability of a TF (1) or directly affects its transcriptional activity (2). Phosphorylation can also influence the localisation of TF to the nucleus (3), influence its interactions with other proteins (4) or it can affect the stability of a protein and its degradation by the proteasome (5). The TF is shown in blue and chromatin in black. Adapted from Holmberg et al., 2002.

Phosphorylation of the TF RUNX1 in megakaroyctes negatively regulates its transcriptional activity and is important for its interactions with other proteins within the nucleus (Huang et al., 2012). RUNX1 is phosphorylated, by Src family kinases, at multiple tyrosine residues located within its negative regulatory DNA-binding domains (Huang et al., 2012). De-phosphorylation of RUNX1, by the tyrosine phosphatase SHP-2, increases RUNX1 association with the SWI/SNF chromatin remodelling complex and reduces its interaction with Gata1 (Huang et al., 2012). Both of these changes promote transcriptional activation, helping to drive the development of megakaryocytes (Huang et al., 2012).

Multi-site phosphorylation of Nuclear Factor of Activated T cells (NFAT) demonstrates how phosphorylation can influence the localisation of TF. The dephosphorylation of NFAT by calcernurin, a Ca^{2+} sensitive serine/threonine phosphatase, causes conformation changes in the protein, exposing a nuclear localisation sub-domain, allowing the TF to accumulate in the nucleus and activate gene expression (Okamura et al., 2000). When Ca^{2+} levels fall and cells are no longer stimulated, calcernurin will no longer be active, causing NFAT to be phosphorylated by multiple kinases and therefore once again sequester in the cytoplasm (Okamura et al., 2000).

Phosphorylation can provide more than a static on/off regulation of TF activity; multiple phosphorylation events can dynamically and precisely tune the transcriptional activity of the protein (Pirkkala et al., 2001). The addition or removal of several phosphate groups can have a dramatic effect on the activity of a TF as each phosphate group adds a negative charge to the modified amino acid, potentially inducing conformation changes and attracting or preventing interactions with other molecules (Holmberg et al., 2002). Phosphorylation of Heat Shock Factor 1 (HSF1) regulates the scale of its transactivation in response to numerous protein-damaging stresses, including heat shock (Pirkkala et al., 2001). HSF1 has four phosphorylation sites, which are modified by different kinases in response to stress, causing it to localise to the nucleus. The more phosphorylated HSF1 becomes, the higher its transcriptional activity (Holmberg et al., 2001). The regulation of HSF1 is complex as three of its phosphorylation sites (S303, S307 and S363) are also modified under normal conditions, as well as in response to stress (Holmberg et al., 2001). However, phosphorylation of S230 is specific to stress responses and is absolutely essential for the activation of HSF1 (Holmberg et al., 2001).

Multiple phosphorylation events of the same TF can also allow integration of two different signalling pathways to cause the same changes in gene expression (Filtz et al., 2014). The process by which substrate specificity is achieved by signal transduction pathways is very important, particularly when a single protein has multiple phosphorylation sites. This generally involves the recognition of particular consensus amino acid sequences or three-dimensional protein structures by kinases (Whitmarsh and Davis, 2000). Additionally, protein complexes can help to promote precise modification, by bringing together a specific TF and kinase combination, allowing them to interact.

Ets1 is an essential TF for the development of multiple immune lineages and its activity is regulated by numerous phosphorylation sites and kinases, allowing it to integrate signals

from a number of signal transduction pathways. Mitogen activated protein kinase (MAPK) mediated phosphorylation of Ets1 positively regulates its transcriptional activity and promotes its interaction with the transcriptional co-activators p300 and CBP (Tootle and Rebay, 2005). This phosphorylation does not affect the stability of Ets1, its ability to bind DNA or its subcellular localisation (Tootle and Rebay, 2005). In contrast, phosphorylation of Ets1 by calmodulin-dependent protein kinase II at multiple sites near its DNA binding domain, interfere with its ability to bind DNA and cause a conformational change, reducing the protein's stability (Pufall et al., 2005). The activity of Ets1 can therefore be fine-tuned by distinct phosphorylation events, allowing it to respond to different external stimuli.

4.1.2 Phosphorylation of E4bp4

E4bp4 was first proposed to be phosphorylated by Chen *et al*. who observed that phosphatases affected the apparent molecular weight of E4bp4 on SDS-PAGE (Chen et al., 1995). They suggested that the phosphorylation of E4bp4 promoted its ability to recognise DNA, as phosphatase treatment interfered with the binding of E4bp4 to target DNA sequences *in vitro* (Chen et al., 1995). Doi *et al*. studied E4bp4 in the chick pineal gland and its role in regulating the circadian clock (Doi et al., 2004). They identified E4bp4 as a target of phosphorylation, suggesting it was modified by Casein Kinase 1ε (CK1ε) and that serine 182 was a site of phosphorylation (Doi et al., 2004). Expression of CK1ε with E4bp4 promoted proteasomal degradation of E4bp4 and also affected its transcriptional activity (Doi et al., 2004). This finding seems to contradict with the previous observation of Chen *et al*. (1995) who suggested phosphorylation was required for E4bp4 to be functional.

However, neither of these studies provided direct biochemical evidence that E4bp4 is phosphorylated and the only evidence of modification was a small shift in molecular weight on SDS-PAGE. Since then, many phosphoproteomic studies have been performed and these have aimed to identify all of the phosphorylated proteins in certain cells and tissues. These studies often analyse global changes in phosphorylation during various cellular events, such as mitosis or cellular responses to stress. The large data sets that these studies produce are openly available and searching these provides further evidence that E4bp4 is phosphorylated.

To date, E4bp4 has been identified as a target of phosphorylation by ten phosphoproteomic studies that analysed the phosphoproteome of various cell types and tissues **(Table 4.1)**. These studies have identified nine serine residues and one threonine residue as possible

sites of phosphorylation in the E4bp4 protein: S130, T132, S210, S258, S285, S286, S301, S352 and S353 **(Table 4.1)**. Both S285 and S286 were present on the same phosphopeptide and most of the studies did not generate sufficient data to determine exactly which serine residue is modified. The same situation occurred with S352 and S353, but a number of studies were able to ascertain that S353 is a more likely site of phosphorylation. In all the studies that positively identified E4bp4 phosphopeptides, none identified phosphorylation at S182, the phosphorylation site previously identified by Doi *et al*. (Doi et al., 2004). Across all ten studies, S286, S301 and S353 were identified numerous times, providing the strongest evidence that these three residues are sites of E4bp4 phosphorylation. However, these studies do not provide full conclusive evidence of phosphorylation at a particular residue. They often identify thousands of potential modification sites and are looking at global changes in cellular phosphorylation levels; therefore there is the potential for false positive identification. It should also be noted that, of the 462 amino acids in the murine E4bp4 protein, 64 residues are serines and 18 threonines, suggesting there is great potential for E4bp4 to have multiple phosphorylation sites.

Despite its growing importance, particularly in the development of innate lymphocyte subsets, very little is known about the mechanisms that regulate E4bp4 function. In particular, little is known about how PTMs affect the function of E4bp4. Despite it being nearly 20 years since the first evidence of E4bp4 phosphorylation, nothing is known about how this PTM may affect its function. No studies have looked specifically at the phosphorylation of E4bp4 *in vivo* and how this affects its functions, for example during NK cell development.

Table 4.1 – Phosphoproteomic screens that identify E4bp4 as a target of phosphorylation

Study	Species	Cell type	Peptides identified	Phosphorylated residues
(Olsen et al., 2010)	*H. sapiens*	HeLa cell line	GPIHS*PVELK	S301
(Villen et al., 2007)	*M. musculus*	Primary liver cells	TSEADEGVVGKSS*DGEDEQQVPK	S285, S286
(Dephoure et al., 2008)	*H. sapiens*	HeLa cell line	TSEADEGVVGKSS*DGEDEQQVPK	S285, S286
			GPIHS*PVELK	S301
(Chen et al., 2009)	*H. sapiens*	HeLa cell line	GSYTASIYQNYMGNSFSGYSHS*PPLLQVNR	S258
			GPIHS*PVELK	S301
			LSS*PIDMTSK	S353
(Choudhary et al., 2009)	*M. musculus*	32D myeloid progenitor cell line	FGLIS*STAYAQEIQK	S130
			FGLISST*AYAQEIQK	T132
			S*PENKFPVIK	S210
			TSEADEGVVGKSS*DGEDEQQVPK	S286
			GPIHS*PVELQR	S301
			LSS*PADAIAK	S353
(Mayya et al., 2009)	*H. sapiens*	Jurkat T cell leukaemia cell line	ETDDGVVGKSS*DGEDEQQVPK	S285, S286
			NEFEATQKLSS*PIDMTSKRHF	S353
(Olsen et al., 2010)	*H. sapiens*	HeLa cell line	GPIHS*PVELK	S301
			LSS*PIDMTSK	S352, S353
(Weintz et al., 2010)	*M. musculus*	Bone marrow derived macrophages	S*PENKFPVIK	S210
			TSEADEGVVGKSS*DGEDEQQVPK	S285, S286
			GPIHS*PVELQR	S301
(Kettenbach et al., 2011)	*H. sapiens*	HeLa cells	TSEADEGVVGKSS*DGEDEQQVPK	S285, S286
			GPIHS*PVELK	S301
(Yu et al., 2011)	*M. musculus*	Murine embryonic fibroblasts	TSEADEGVVGKSS*DGEDEQQVPK	S285, S286

* highlights phosphorylated residues

4.1.3 Aim and objectives

This study aimed to provide unambiguous evidence for the phosphorylation of E4bp4, determine exactly which sites are modified and establish if these regulate the function of the protein. In order to achieve this, the following objectives were set:

1. Determine if E4bp4 is phosphorylated and use mass spectrometry to identify exactly which amino acid residues are modified.
2. Create mutated forms of E4bp4 lacking phosphorylation sites.
3. Determine if phosphorylation is important for the function of E4bp4, particularly during NK cell development by analysing activity of E4bp4 mutants.
4. Determine if the inhibiting kinase activity can influence E4bp4 and NK cell development.

4.2 Prediction of E4bp4 phosphorylation

Bioinformatic analysis was performed to analyse the potential phosphorylation sites previously identified for E4bp4 in phosphoproteomic screens **(Table 4.1)**. All of the putative phosphorylation sites were found to be highly conserved across a range of vertebrate species from humans through to zebrafish, including S182 the phosphorylation site suggested by Doi *et al*. (Doi et al 2004) **(Figure 4.2)**. Three prediction servers were used to analyse the putative phosphorylation sites of E4bp4 to determine which are the most likely to be modified and which kinase may recognise each site **(Table 4.2)**. The murine E4bp4 primary amino acid sequence was analysed by GPS 2.1 (Xue et al., 2011), NetPhosK (Blom et al., 2004) and Scansite 3.0 (Obenauer, 2003), which all make predictions based on the amino acid sequence surrounding the phosphorylated serine, comparing the input sequence to known sites of phosphorylation **(Table 4.2)**. The results from the three prediction tools varied but GPS 2.1 predicted S210 and S353 as the most likely sites of phosphorylation; NetPhosK predicted S286 and S353 and Scansite predicted S182 and S210 **(Table 4.2)**. All three tools predicted S130, S285 and S352 as the least likely to be phosphorylated **(Table 4.2)**. They also predicted that E4bp4 would be modified by Mitogen-activated protein kinases, cyclin-dependent kinases and casein kinases at the various potential phosphorylation sites **(Table 4.2)**. This analysis provided further evidence that phosphorylation of E4bp4 is likely to occur.

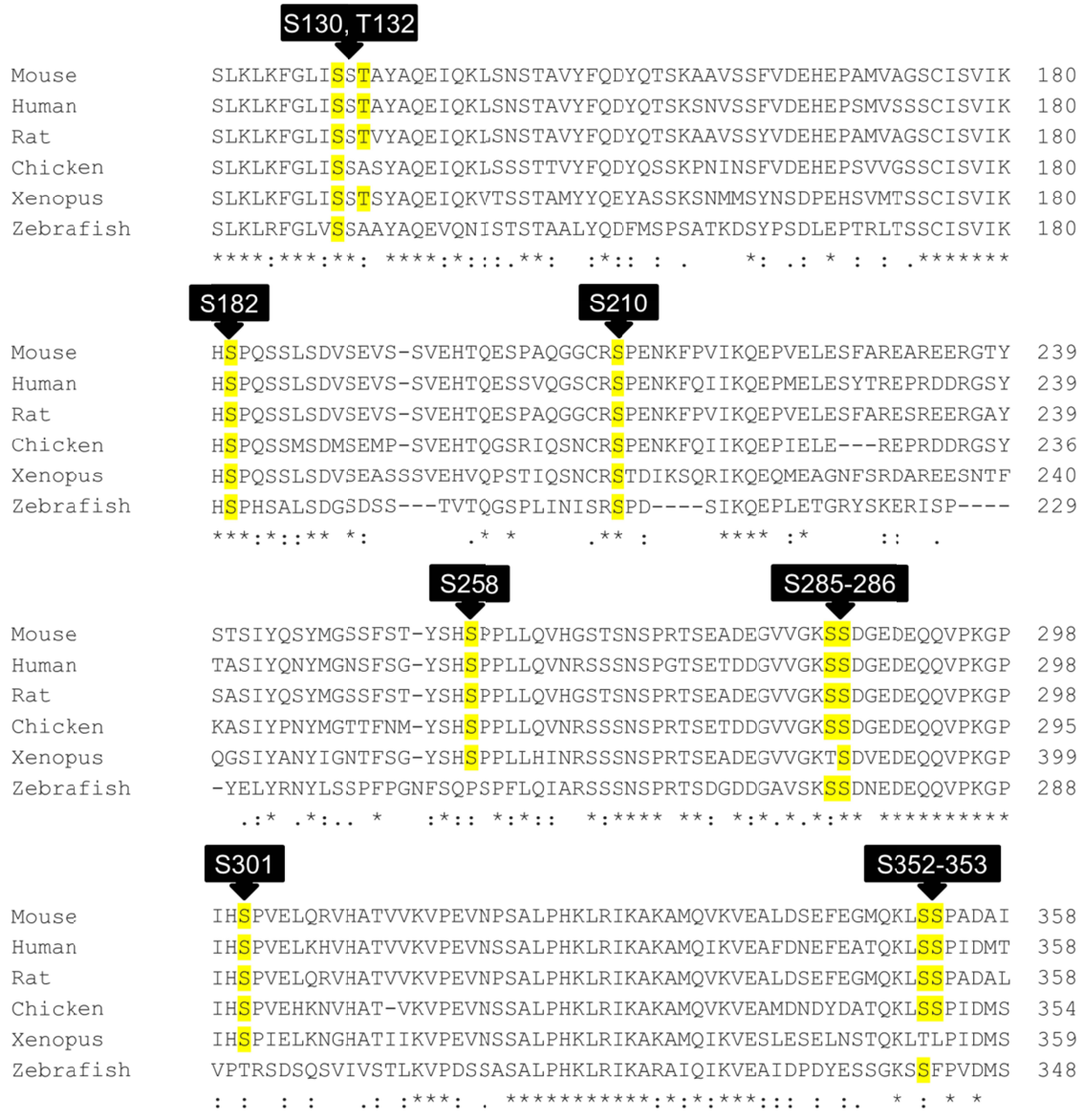


Figure 4.2 – Putative phosphorylation sites of E4bp4 are highly conserved

The E4bp4 amino acid sequence from various species was aligned using the web-based tool Clustal Omega (Goujon et al., 2010). Potential phosphorylation sites are highlighted. Asterisks highlight perfectly conserved residues and dots highlight partially conserved residues.

Table 4.2 – Prediction of E4bp4 phosphorylation sites and modifying kinases

Position	Peptide	GPS Score (0-20)	NetPhosK Score (0-1)	Scansite Score (0-1)	Predicted Kinase
S130	KFGLISSTAYAQ	4.62	0.023	0.243	GRK
T132	GLISSTAYAQEI	11.0	0.011	0.285	MAPK
S182	SVIKHSPQSSLS	6.47	0.967	0.552	CK 1/2
S210	QGGCRSPENKFP	12.0	0.866	0.707	CDK
S258	STYSHSPPLLQV	10.6	0.071	0.425	CDK
S285	GVVGKSSDGEDE	4.02	0.937	0.442	CK 1/2
S286	VVGKSSDGEDEQ	6.38	0.990	0.427	CK 1/2
S301	KGPIHSPVELQR	9.33	0.958	0.459	CDK
S352	GMQKLSSPADAI	5.45	0.098	0.275	Auora B
S353	MQKLSSPADAIA	16.41	0.997	0.422	MAPK

Phosphorylated residue highlighted

GRK = G-protein coupled receptor kinase; MAPK = Mitogen-activated protein kinase; CDK = Cyclin-dependent kinase; CK = Casein kinase

4.3 E4bp4 is multiply phosphorylated

High accuracy MS was used to definitively determine the phosphorylation sites of E4bp4. Previously, a similar approach was used to identify SUMOylation sites of E4bp4 (section 3.4). The MS analysis aimed to unambiguously determine exactly which of the putative E4bp4 phosphorylation sites existed and also search for any unknown modification sites. For this analysis a purified sample of E4bp4 was required, therefore FLAG-E4bp4 was expressed in 293T cells and purified by immunoprecipitation. FLAG-E4bp4 was competitively eluted from the anti-FLAG M2 affinity resin using the 3X FLAG peptide and elution fractions were combined and concentrated. Analysis of the purified sample by WB revealed that E4bp4 was clearly phosphorylated, because a strong signal was observed when probing for anti-phosphoserine **(Figure 4.3A)**. The sample was also analysed on an SDS-PAGE gel stained with Imperial Protein Stain (Thermo Scientific) and this confirmed that the purification of FLAG-E4bp4 was successful because few contaminating proteins were detected in the sample **(Figure 4.3B)**.

As phosphorylation is rapidly reversible and modified isoforms are typically low in abundance, the detection of phosphoproteins is not straightforward. Therefore, following trypsin digestion of the purified FLAG-E4bp4 sample, all phosphopeptides were enriched using titanium dioxide beads that specifically bind phosphorylated peptides. This purification step removed a large amount of unmodified E4bp4 peptides that would provide background noise during the MS analysis.

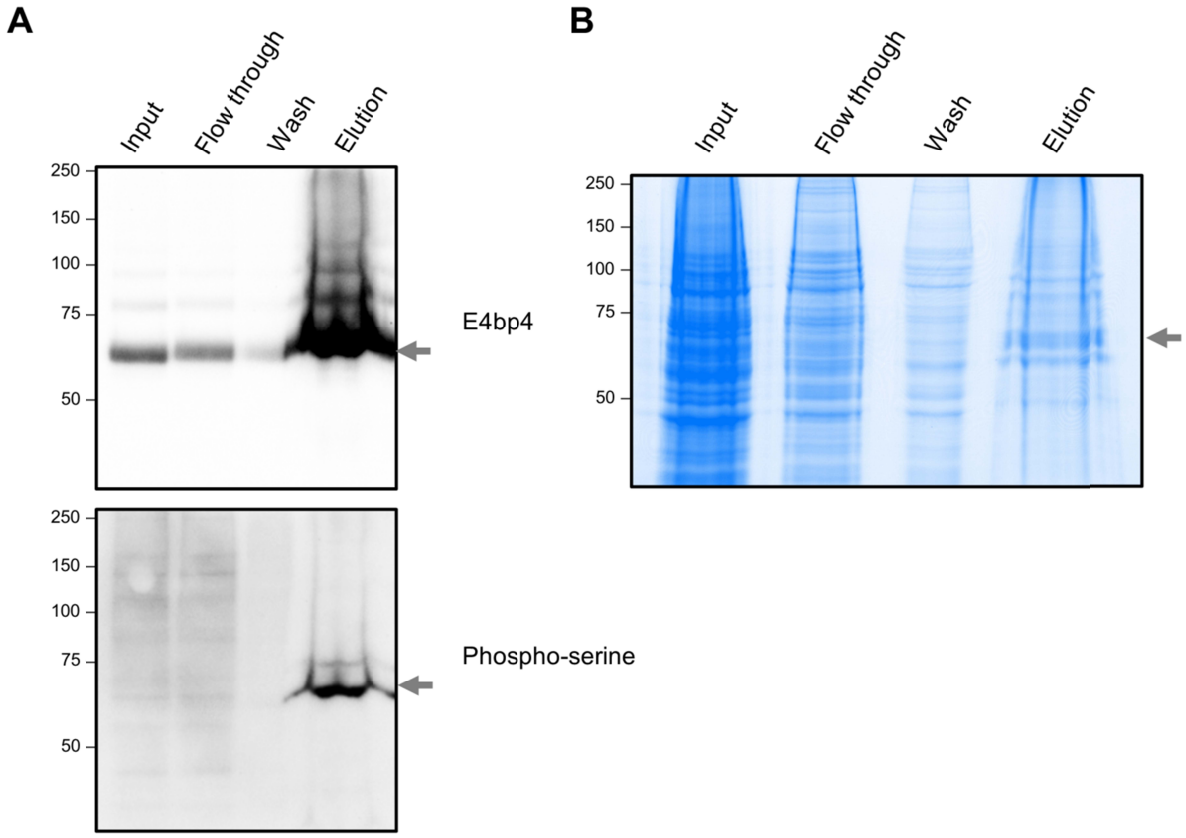


Figure 4.3 – E4bp4 is phosphorylated

FLAG-E4bp4 was expressed in 293T cells and immunoprecipitated using anti-FLAG M2 affinity resin and the purified protein was competitively eluted using 3X FLAG peptide. Elution fractions were combined and concentrated and 10% of the purified sample was analysed on a WB (A) and on an SDS-PAGE gel stained with Imperial protein stain (B). WB was probed for the presence of E4bp4 and phosphoserine. I = Input, F = Flow through, W = Wash, E = Elution. Grey arrows highlight E4bp4. Molecular weight markers are shown on the left of all blots in kDa.

The phosphopeptide enriched sample was analysed by a Thermo Scientific™ Velos Pro ion Trap mass spectrometer and the resulting data was searched against the mouse Uniprot database using Sequest, looking for phosphorylated peptides, indicated by a mass increase of 79.966 Da **(Table 4.3)**. All reported identifications were validated by manual inspection of the tandem mass spectra data and spectra were searched for the presence of predicted ions.

Table 4.3 – Predicted masses of potential E4bp4 phosphopeptides

Modified residue	Peptide	Peptide mass (m/z, 1+ charge)	Peptide + phosphate (m/z, 1+ charge)	Peptide + phosphate (m/z, 2+ charge)
S130	FGLISSTAYAQEIQK	1655.853	1735.853	908.430
S182	HSPQSSLSDVSEVSSVEHTQES PAQGGCR	3025.344	3105.344	1593.176
S210	SPENK	574.272	654.272	367.640
S258	GTYSTSIYQSYMGSSFSTYSHS PPLLQVHGSTSNSPR	4011.834	4091.834	2086.421
S286	SSDGEDEQQVPK	1318.575	1398.576	699.788
S301	GPIHSPVELQR	1232.674	1312.675	656.837
S353	LSSPADAIAK	972.536	1052.536	526.768

Analysis of the MS/MS data revealed clear evidence of three E4bp4 phosphorylation sites. Peptides were identified with phosphorylated serine residues at serine 301 **(Figure 4.4)**, S353 **(Figure 4.5)** and at either serine 285 or serine 286 **(Figure 4.6)**. The tandem mass spectra for the S301 and S353 peptides unambiguously identified phosphorylation at these sites, with most *b* and *y* ions observed for both peptides **(Figure 4.4) (Figure 4.5)**. The final peptide identified confirmed E4bp4 phosphorylation at either S285 or S286 **(Figure 4.6)**. Analysis of the CID tandem mass spectra for this peptide could not unambiguously determine whether the phosphorylated residue was S285 or S286, as these two residues are adjacent to each other at the N-terminus of the peptide and some of the ions were low in abundance **(Figure 4.6)**. However, phosphorylation can alter the ability of trypsin to identify digestion sites, causing mis-cleavages when phosphorylated serine residues are location adjacent to Arg or Lys residues (Steen et al 2005) (Previs et al 2008). As S285 is at the very

N-terminus of the identified peptide, next to the trypsin cleavage site, it is unlikely this site is modified as a mis-cleavage was not observed. This MS analysis clearly demonstrated that E4bp4 is multiply phosphorylated, but whether these modifications are required for its function is yet to be explored. All of the phosphorylation sites should be tested to determine if they regulate the activity of E4bp4.

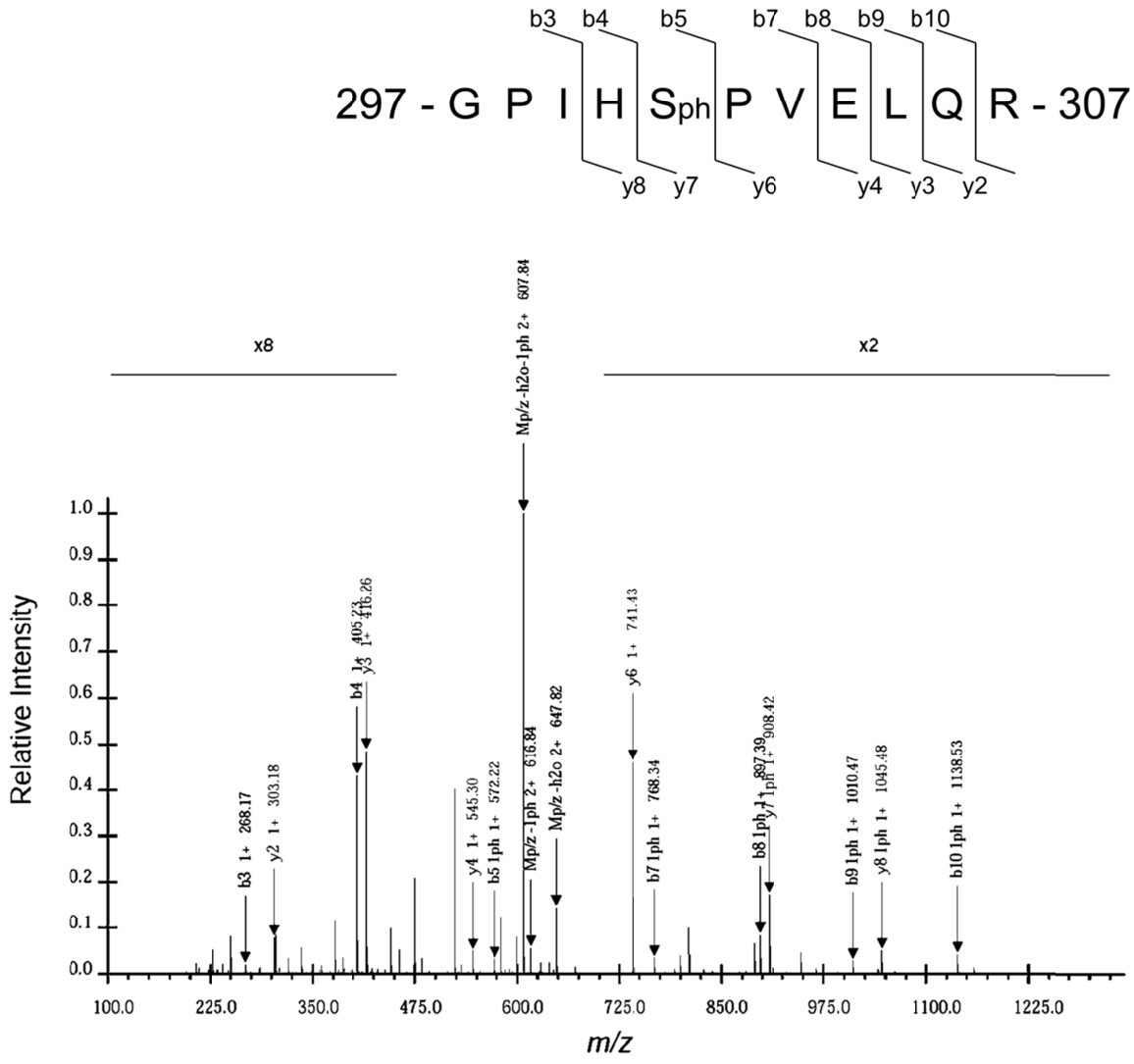


Figure 4.4 – Serine 301 is a site of E4bp4 phosphorylation

Annotated CID tandem mass spectrum of the +2 ion at *m/z* 656.83, confirming phosphorylation of E4bp4 at S301. Tandem mass spectra acquired with an electrospray ionization LTQ/Orbitrap mass spectrometer from purified FLAG-E4bp4. FLAG-E4bp4 was expressed in HEK-293T cells and purified by immunoprecipitation with anti-FLAG M2 affinity resin and the purified protein competitively eluted using 3X FLAG peptide. Purified E4bp4 was digested with trypsin and subjected to phosphopeptide enrichment using TiO_2 beads to reduce background from unphosphorylated E4bp4 peptides. Magnification of certain regions of the spectra is highlighted. Above spectra is schematic representation of fragmented peptide with identified ions labelled.

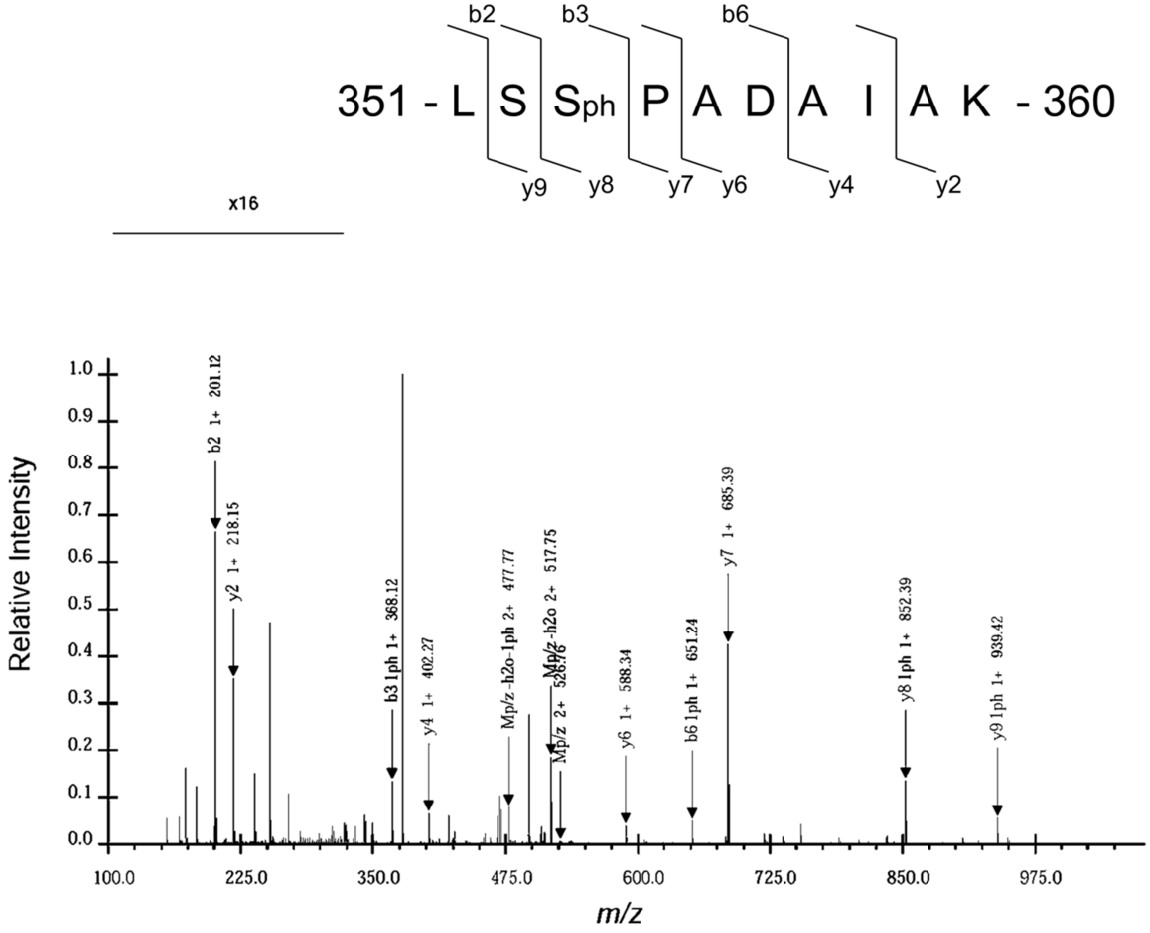


Figure 4.5 – Serine 353 is a site of E4bp4 phosphorylation

Annotated CID tandem mass spectrum of the +2 ion at *m/z* 526.76, confirming phosphorylation of E4bp4 at S353. Tandem mass spectra acquired with an electrospray ionization LTQ/Orbitrap mass spectrometer from purified FLAG-E4bp4. FLAG-E4bp4 was expressed in HEK-293T cells and purified by immunoprecipitation with anti-FLAG M2 affinity resin and the purified protein competitively eluted using 3X FLAG peptide. Purified E4bp4 was digested with trypsin and subjected to phosphopeptide enrichment using TiO_2 beads to reduce background from unphosphorylated E4bp4 peptides. Magnification of certain regions of the spectra is highlighted. Above spectra is schematic representation of fragmented peptide with identified ions labelled.

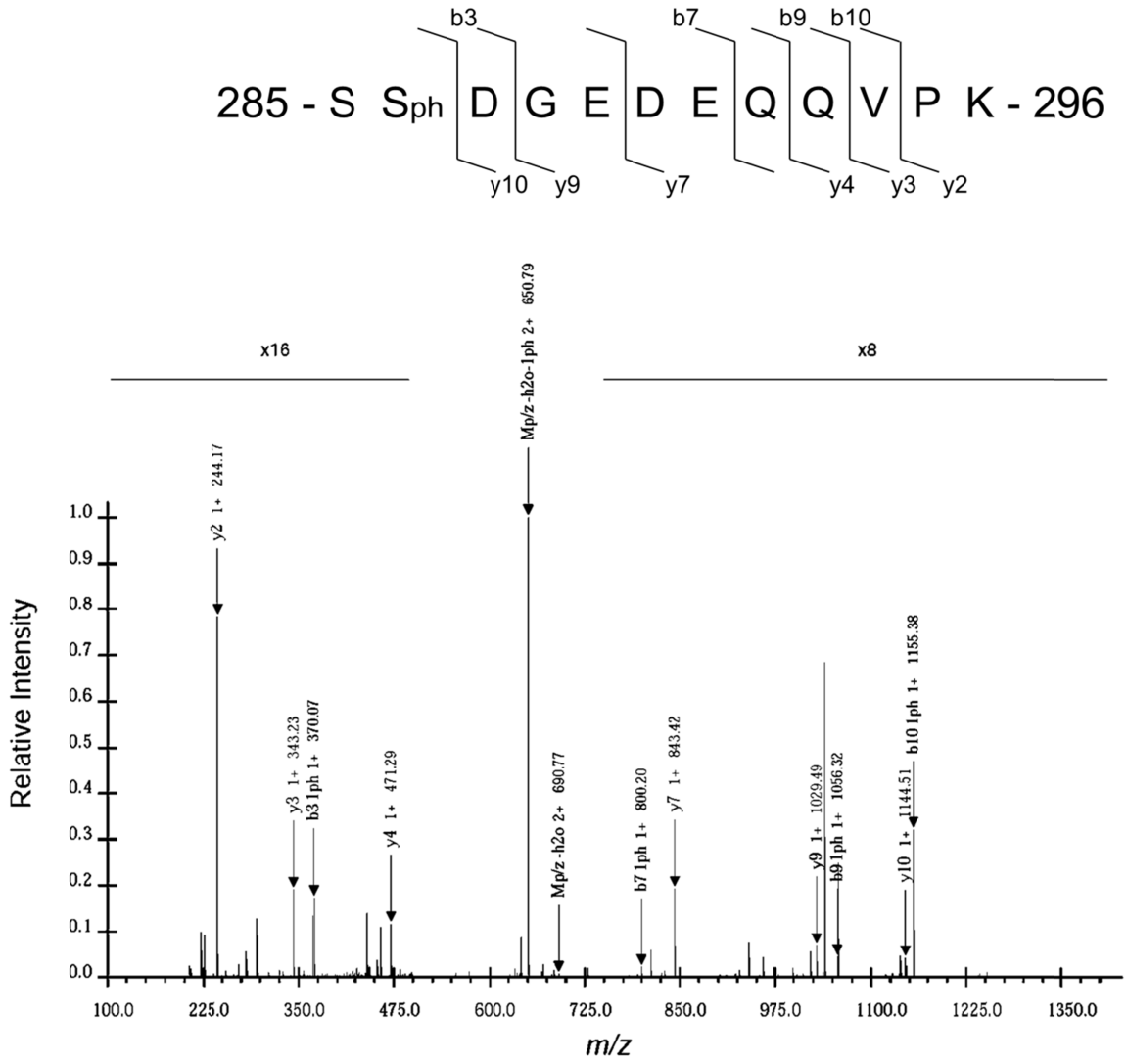


Figure 4.6 – Serine 285 or 286 is a site of E4bp4 phosphorylation

Annotated CID tandem mass spectrum of the +2 ion at *m/z* 699.77, confirming phosphorylation of E4bp4 at S285 or S286. Tandem mass spectra acquired with an electrospray ionization LTQ/Orbitrap mass spectrometer from purified FLAG-E4bp4. FLAG-E4bp4 was expressed in HEK-293T cells and purified by immunoprecipitation with anti-FLAG M2 affinity resin and the purified protein competitively eluted using 3X FLAG peptide. Purified E4bp4 was digested with trypsin and subjected to phosphopeptide enrichment using TiO_2 beads to reduce background from unphosphorylated E4bp4 peptides. Magnification of certain regions of the spectra is highlighted. Above spectra is a schematic representation of fragmented peptide with identified ions labelled.

4.4 Mutating E4bp4 phosphorylation sites

To determine if a particular serine or threonine residue is a site of phosphorylation, it can be mutated to an alanine residue, which will abolish any phosphorylation at that site, whilst maintaining the three-dimensional structure of the protein (Holmberg et al., 2002) **(Figure 4.7)**. Analysing the functionality of serine-alanine mutants allows the importance of particular phosphorylation sites to be determined. This method has been successfully used to analyse the phosphorylation sites of numerous TFs including p53 (Maclaine and Hupp, 2011), HSF1 (Holmberg et al., 2001) and NFκB (Viatour et al., 2005).

All of the confirmed E4bp4 phosphorylation sites were individually mutated in the E4bp4 cDNA using site direct mutagenesis **(Table 4.4)**. Both S285 and S286 sites were separately mutated as it was not possible to unambiguously determine exactly which site was phosphorylated. Mutations were made at S182 and S210 as they were predicted to be phosphorylated, but modification of neither site was confirmed in this study. Doi *et al*. previously suggested that S182 was a site of phosphorylation that influenced the function of E4bp4, so it was mutated to determine if this was correct (Doi et al., 2004). Numerous double mutants lacking two phosphorylation sites were also generated, primarily involving S286, S301 and S353 as these sites were consistently identified in the phosphoproteomic screens and were all confirmed here to be modified in this study **(Table 4.4)**. All of the mutations were made in the E4bp4 cDNA, which had been previously cloned into the MSCV-IRES-hCD2 expression vector **(Figure 3.21)**. This meant that all the mutants could be expressed in numerous cell types, but specially in developing Lin^- BM cells, so that their effect on NK cell development could be determined.

Table 4.4 – E4bp4 phosphorylation-mutant cDNAs generated

Name	Serine – Alanine mutations	Name	Serine – Alanine mutations
E4bp4-S182A	S182	E4bp4-S353A	S353
E4bp4-S210A	S210	E4bp4- S286-301A	S286, S301
E4bp4-S285A	S285	E4bp4-S286-353A	S286, S353
E4bp4-S286A	S286	E4bp4-S301-353A	S301, S353
E4bp4-S301A	S301		

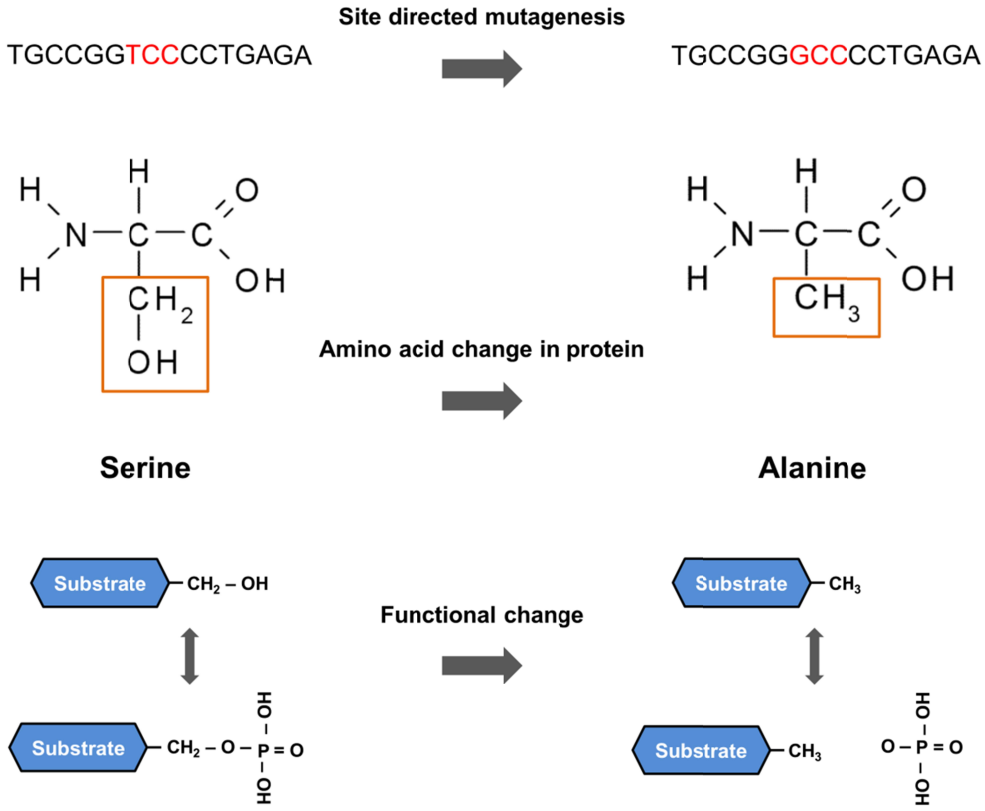


Figure 4.7 – Mutating serine residues to alanine abolishes phosphorylation

Site directed mutagenesis was used to mutate individual serine residues to alanine residues. This mutation does not affect the three-dimensional structure of the protein but will prevent any phosphorylation at that site. Serine side chains form an ester bond with the PO_4^{3-} group but this cannot occur with the side chain of alanine.

4.5 Removing phosphorylation sites of E4bp4 causes increased NK cell production

As E4bp4 has numerous phosphorylated serine residues it was predicted that these may influence its activity. In particular, as it is a central regulator of NK cell development it was proposed that phosphorylation, like SUMOylation, may influence the activity of E4bp4 during this process. To determine whether phosphorylation was important for the function of E4bp4, each serine-alanine phosphorylation mutant was expressed in *E4bp4*$^{-/-}$ Lin^{-} BM cells, cultured *in vitro,* to determine if it had the ability to rescue the development of NK cells **(Figure 3.20)**. Cells were transduced with a retrovirus expressing the different mutants, grown in lymphocyte inducing conditions, and in the presence of IL-15, before flow cytometry analysis was used to determine the number of NK cells that had developed **(Figure 3.20)**.

In the absence of E4bp4, the Lin$^-$ BM cells were unable to develop into NK cells, but transduction with WT E4bp4 rescued the development of NK cells **(Figure 4.8)**. All of the single site serine-alanine mutants were analysed and most did not affect NK cell development, except for S286A, which, on average, produced almost double the number of NK cells compared to WT E4bp4 **(Figure 4.8)**. The S353A mutant also had a smaller, but insignificant effect on NK cell production **(Figure 4.8)**. The S286-301A and S286-353A mutants even greater increases in the production of NK cells, compared to WT E4bp4, in particular, the S286-353A mutant increased NK cell production three fold **(Figure 4.8)**. This demonstrated that phosphorylation of S286 is potentially a key regulatory mechanism for controlling the activity of E4bp4 and that to a lesser extent phosphorylation of S353 is also required. Phosphorylation of S353 may also be dependent or influenced by the phosphorylation status of S286, as it had a larger affect on NK cell development in combination with the S286A mutation.

Serine 301 was clearly demonstrated to be phosphorylated by MS analysis in this study, but mutating this site had no effect on NK cell development, even in combination with S286A the level of NK cell production was the same as the single S286A mutant **(Figure 4.8)**. Mutating the S182 and S210 phosphorylation sites did not affect the function of E4bp4 in any way, this suggests that these sites are not likely to be modified and if they are, it is not important for NK cell development. The S285A mutation also did not affect NK cell development, further confirming that this site is not phosphorylated.

To expand on these findings, a number of the phosphorylation mutants were also expressed in WT Lin$^-$ BM cells to determine if they could promote greater levels of NK cell production than WT E4bp4. Transduction of WT Lin$^-$ BM with WT E4bp4 has been previously shown to increase the production of NK cell *in vitro*, demonstrating that E4bp4 is a limiting factor in NK cell development (Gascoyne et al., 2009) (Male et al., 2014) and this observation was confirmed here. More significantly, the expression of E4bp4 phosphorylation mutants increased NK cell production to a greater extent than WT E4bp4 **(Figure 4.9)**. The expression of S286A and S286-353A consistently produced greater numbers of NK cells than WT E4bp4 or the S353A mutant **(Figure 4.9)**. However, in this assay (unlike with *E4bp4*$^{-/-}$ Lin$^-$ BM) the S286-301A mutant did not affect NK cell production **(Figure 4.9)**. The differences observed between WT E4bp4 and phosphorylation mutants were greater with *E4bp4*$^{-/-}$ Lin$^-$ BM cells, than when using WT cells, most likely due to the presence of endogenous E4bp4 in these cells.

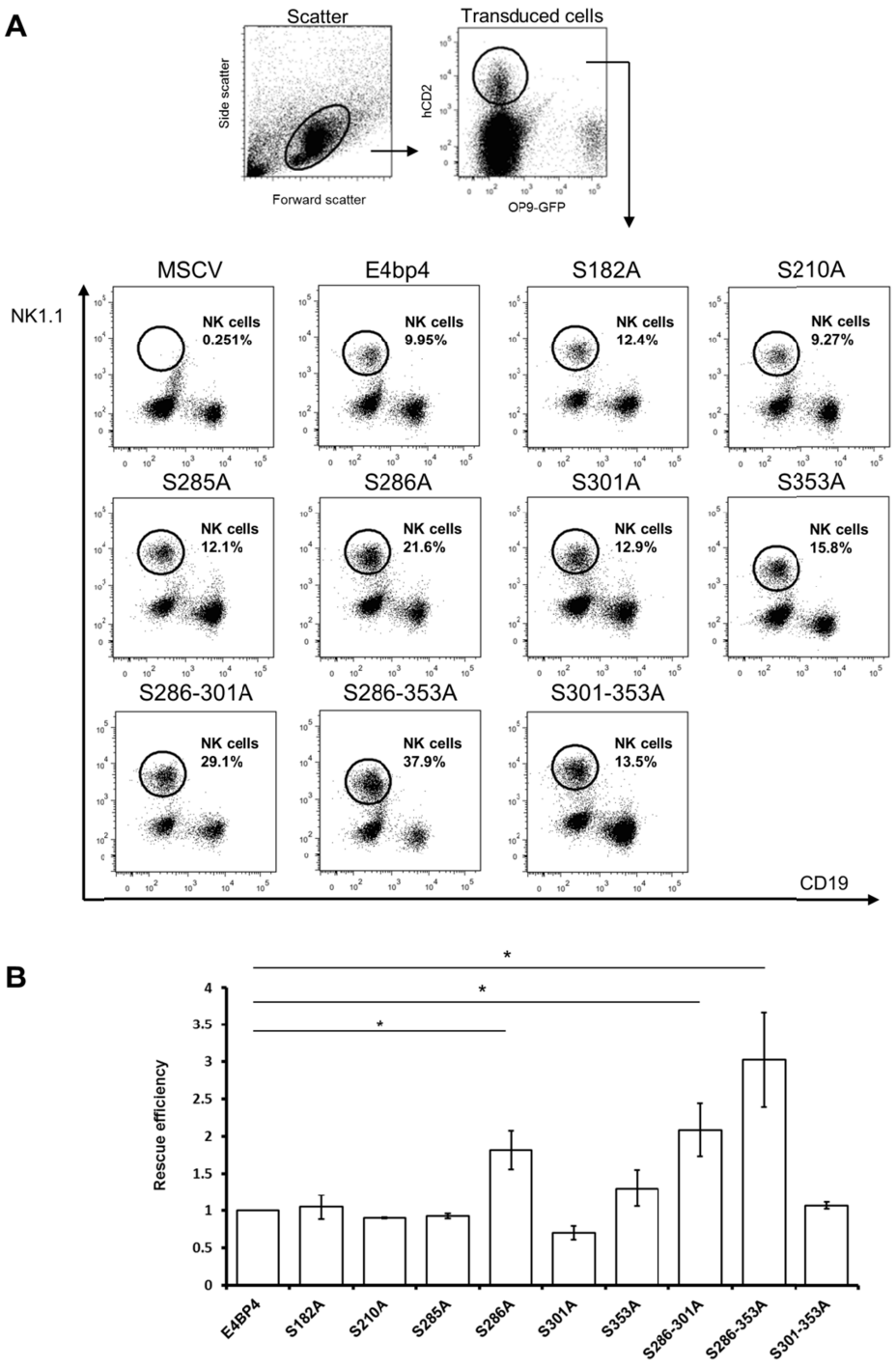


Figure 4.8 – Removing phosphorylation sites from E4bp4 increases *in vitro* NK cell production from *E4bp4*$^{-/-}$ Lin^{-} BM

Lin^{-} BM cells were isolated from *E4bp4*$^{-/-}$ mouse BM and cultured for two days in the presence of IL-7, Flt3 ligand and SCF before retroviral transduction with a MSCV-IRES-hCD2 construct (expressing WT E4bp4 or E4bp4 phosphorylation site mutants). Transduced cells were cultured for three days, before being transferred onto OP9 stromal cells and cultured in the presence of IL-15 for 7 days. A) Flow cytometry was used to analyse cultured cells to identify live lymphocytes that were positive for hCD2 (marker of transduced cells) and subsequently to examine their expression of NK1.1 and CD19. B) Efficiency of the rescue affected by the different E4bp4 phosphorylation mutants, normalised to that of WT E4bp4. Data are representative of three independent experiments for each mutant. Results were verified to be significant using paired Student's t-tests, error bars show SEM. *, $P < 0.05$.

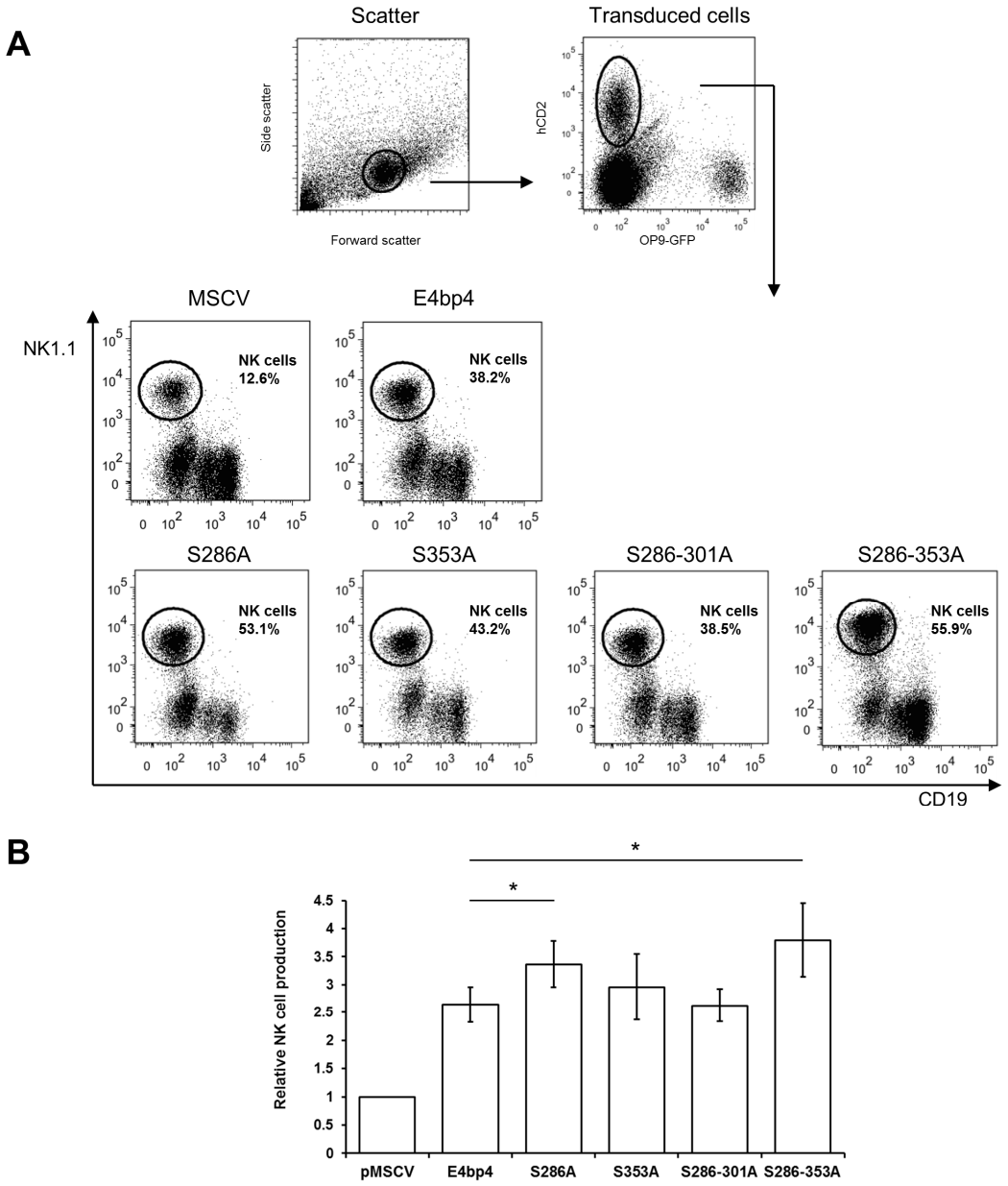


Figure 4.9 – Removing phosphorylation sites from E4bp4 increases *in vitro* NK cell production from WT Lin$^-$ BM

Lin$^-$ BM cells were isolated from WT BM and cultured for two days in the presence of IL-7, Flt3 ligand and SCF before retroviral transduction with a MSCV-IRES-hCD2 construct (WT E4bp4 or E4bp4 phosphorylation site mutants). Transduced cells were cultured for three days, before being transferred onto OP9 stromal cells and cultured in the presence of IL-15 for 7 days. A) Flow cytometry was used to analyse cultured cells to identify live lymphocytes that were positive for hCD2 (marker of transduced cells) and subsequently to examine their expression of NK1.1 and CD19. B) Relative level of NK cell production by the different E4bp4 phosphorylation mutants and WT E4bp4 normalised to the empty vector. Data are representative of three independent experiments for each mutant. Results were verified to be significant using paired Student's t-tests, error bars show SEM. *, $P < 0.05$.

4.6 Phosphorylation by Casein Kinase 1ε is important during NK cell development

As the phosphorylation of E4bp4 has a regulatory role during NK cell development, it was postulated that inhibition of specific kinases may influence this process. Doi *et al*. previously suggested that E4bp4 was phosphorylated by casein kinase Iε (CKIε) (Doi et al., 2004), so inhibition of this enzyme may affect E4bp4 and subsequently NK cell development. To test this hypothesis, NK cells were developed *in vitro* in the presence or absence of an inhibitor of casein kinase I activity, D4476 (Rena et al., 2004).

Lin^- BM cells were isolated from WT BM, grown for 5 days in lymphocyte inducing conditions **(Figure 3.20)**, before being transferred onto OP9 stromal cells and grown in the presence or absence of IL-15 and D4476. In the presence of the kinase inhibitor, the number of NK cells that developed dramatically increased, regardless of whether the cells were grown with IL-15 **(Figure 4.10)**. Incidentally, the kinase inhibitor seemed to have a greater effect on the cells when they were grown without IL-15 **(Figure 4.10)**. When this assay was repeated with a casein kinase II inhibitor, 4,5,6,7-tetrabromobenzotriazole (TBB), there was no effect on NK cell development **(Figure 4.10)**. In fact TBB actually caused a small reduction in the number of NK cells that developed, but this may simply be due to toxicity issues. Phosphorylation of E4bp4 and potentially other TFs is clearly a mechanism by which NK cell development can be regulated and chemical kinase inhibitors represent a simple method by which the development of NK cells can be influenced.

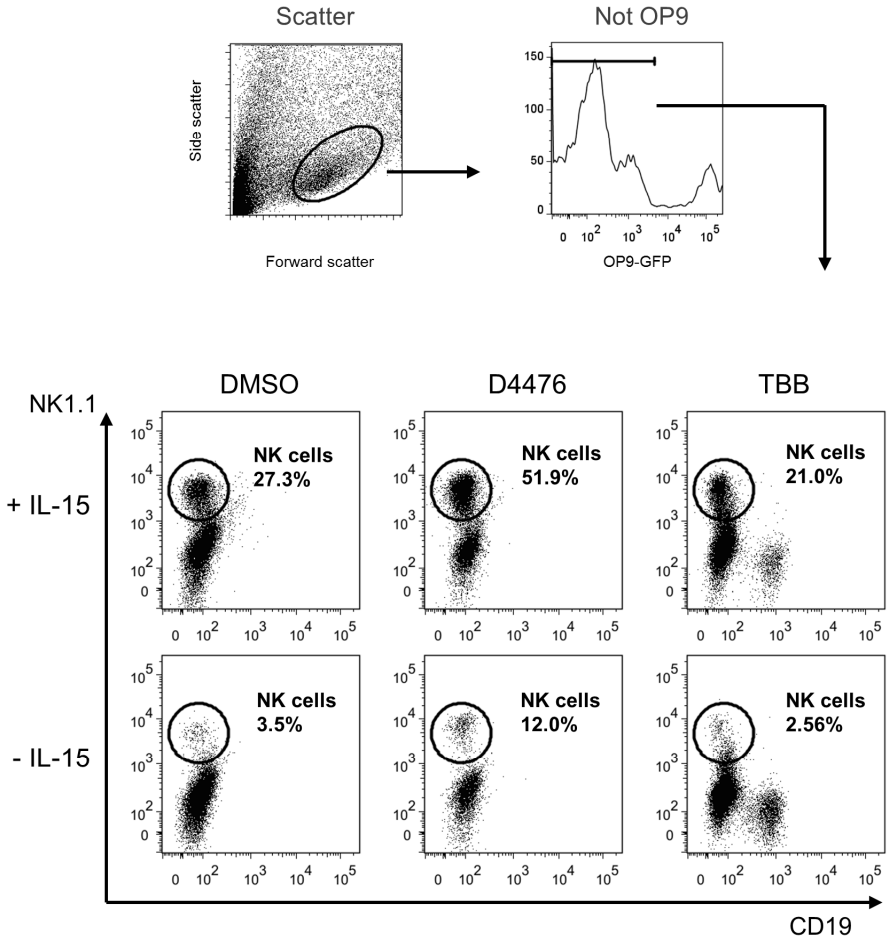


Figure 4.10 – Inhibition of Casein Kinase I promotes NK cell development.

WT Lin$^-$ BM cells were grown in lymphocyte producing conditions (IL-7, Flt3L, SCF) for five days, before being transferred onto OP9 stromal cells. Cells were grown in presence or absence of IL-15 and 10 μM D4476 (an inhibitor of casein kinase I) or 20 μM TBB (an inhibitor of casein kinase II). After seven days of culture cells were analysed by flow cytometry to identify live lymphocytes (scatter), which were GFP$^-$ (stromal OP9 cells are GFP$^+$) and subsequently their expression of NK1.1 and CD19. Results shown are representative of two independent experiments.

4.7 Discussion

Phosphorylation is a critical regulator of TF function that can rapidly and reversibly alter the function of a TF and lead to changes in gene expression (Filtz et al., 2014). TFs can have multiple phosphorylated residues allowing them to receive signals from different signalling pathways, enabling them to respond to different stimuli, or alternatively only respond if two separate signals are received (Holmberg et al., 2002). Multiple phosphorylation sites can

also allow the activity of a TF to be tuned with incremental phosphorylation, leading to equivalent changes in gene expression (Filtz et al., 2014). The TF E4bp4, which is essential for the development of NK cells, has previously been suggested as a target for phosphorylation, but whether this is important for its function is not known.

In this study, E4bp4 has been conclusively established to have multiple phosphorylation sites and more significantly removing the modified serine residues influenced the activity of E4bp4 during NK cell development. Initial searches of the data sets produced by numerous phosphoproteomic studies revealed a number of highly conserved potential E4bp4 phosphorylation sites. When a purified sample of E4bp4 was analysed on a WB, a clear signal was observed when probing for phosphoserine, providing clear evidence that E4bp4 is phosphorylated *in vivo*. However, to precisely determine the exact locations of E4bp4 phosphorylation, the purified E4bp4 was analysed using high accuracy MS. This was not a straightforward process as phosphorylated forms of most proteins constitute < 5% of the total protein within a cell. Therefore, to aid the identification of E4bp4 phosphorylation sites, phosphopeptides were enriched using TiO_2 beads. The MS analysis provided unambiguous evidence that E4bp4 has at least three phosphorylation sites at serine 301, 353 and either 285 or 286. This confirmed the data of previous phosphoproteomic screens (Choudhary et al., 2009) (Dephoure et al., 2008) (Kettenbach et al., 2011) (Olsen et al., 2010) (Weintz et al., 2010) (Yu et al., 2011) which identified the same phosphorylated E4bp4 peptides. Doi *et al*. previously proposed that S182 was a site of E4bp4 phosphorylation (Doi et al., 2004), but in this study no clear evidence was obtained to support this claim and this site does not appear to be important for the function of E4bp4.

As the S285-286 peptide was of low abundance and the S285 residue was at the N-terminus of the peptide, it was not possible to categorically determine which serine residue in the peptide was modified. Close proximity of a phosphorylated serine residues to a trypsin cleavage site can cause miscleavage by trypsin, due to the presence of the negative charge of the phosphate group changing the protein chemistry around the trypsin digest site (Steen et al., 2005) (Previs et al., 2008). This suggested that S285 is unlikely to be the site of modification as it is at the very N-terminus of the peptide and therefore next to a trypsin digest site. More significantly, the phosphoproteomic study by Choudhary *et al.* identified the same E4bp4 phosphopeptide and their observations suggested that S286 is indeed the site of phosphorylation on this peptide (Choudhary et al., 2009).

The MS analysis revealed no other phosphorylated E4bp4 peptides in the sample. However one of the limitations of this approach means that it is not possible to rule out

phosphorylation at other sites. It may be that other sites do exist but are extremely low in abundance and therefore could not be detected in this analysis. Alternatively, phosphorylation sites can be missed if they exist on tryptic peptides that are too small or too large to be accurately identified by the MS. To be certain that all E4bp4 phosphorylation sites have been identified, alternative protein digest strategies could be employed using enzymes such as Glu-C or Lys-C. A more complex but potentially useful strategy would be to use top-down MS, which involves the analysis of whole proteins, allowing all the PTMs on a single protein to be observed at one time (Lanucara and Eyers, 2013) (Zhang and Ge, 2011). Using this approach the whole protein is fragmented inside the mass spectrometer, meaning complete sequence coverage can be achieved, as the protein is not subject to enzymatic digestion prior to MS analysis (Zhang and Ge, 2011). All individual modification sites can potentially be mapped on the protein, with some PTMs present on multiple peptides which can help to confirm their existence (Zhang and Ge, 2011). Additionally this approach allows the absence of a modification at a specific site to also be confirmed. However, top-down MS does require highly specialised hardware and software to analyse the complex detailed spectra that are produced and a number of technical challenges still remain with dealing the large data sets acquired (Lanucara and Eyers, 2013) (Zhang and Ge, 2011).

All of the E4bp4 phosphorylation sites that were identified in this study are within or very close to the C-terminal transcriptional repression domain of E4bp4. This means they potentially influence the transcriptional activity of E4bp4, but are less likely to influence its DNA binding ability. However, minimal information about the possible three-dimensional structure of E4bp4 is available; only the structure of the DNA binding domain of CREB (bZIP TF) has been solved (Schumacher et al., 2000). E4bp4 is likely to have a similar structure at its N-terminus where it binds to DNA, but the structure of the majority of the protein is unknown. It is therefore difficult to know how the E4bp4 phosphorylation sites come together in three-dimensional space. It maybe that they all cluster together at a particular region of the protein's surface, but currently there is no way to predict this.

As E4bp4 was found to be multiply phosphorylated, it was essential to determine whether phosphorylation was important for its function. In order to do this, numerous serine-alanine mutants were created by site-directed mutagenesis and the functionality of these mutants was compared to WT E4bp4 to determine how the loss of a phosphorylation site affected its function. These mutants were compared to WT E4bp4 for their ability to rescue the development of NK cells from *E4bp4*$^{-/-}$ Lin^{-} BM. The S286A, S286-301A and S286-353A mutants all consistently rescued NK cell development to a greater extent that WT E4bp4. In

fact, the S286-353A mutant on average promoted three times the number of NK cells to develop compared to WT E4bp4. Strikingly, mutating a single serine residue, S286, almost doubled the number of NK cells that were produced, demonstrating that phosphorylation at this residue is critical for regulating E4bp4's activity.

Another intriguing observation was that mutation of S353 alone only made a small increase in NK cell production, but in combination with S286A increased NK cell production compared to WT E4bp4 and the S286A mutant alone. This suggested there may be a hierarchy of E4bp4 phosphorylation sites. The S301 phosphorylation sites did not seem to be essential for NK cell development and may be involved in regulating its function in different cell types or in different contexts. S182 and S210 which were proposed to be phosphorylation sites of E4bp4, but were not identified by MS in this study, also did not influence its function as mutation of these residues did not affect the production of NK cells from *E4bp4*$^{-/-}$ Lin^{-} BM.

These observations were supported by data from a similar experiment using WT Lin^{-} BM. In these cells the E4bp4 phosphorylation mutants also promoted a greater level of NK cell production compared to WT E4bp4. However the differences observed between the mutants and WT E4bp4 were not as substantial in this experiment as the transduced cells already expressed endogenous E4bp4. Nevertheless more than 50% of all the cells transduced with the S286-353A mutant developed into NK cells.

These two experiments clearly demonstrated that phosphorylation is a key regulatory mechanism for controlling the activity of E4bp4. This finding is supported by the data from the original study on E4bp4 phosphorylation by Doi *et al*. who also suggested that phosphorylation inhibited the activity of E4bp4 (Doi et al., 2004). Doi *et al*. observed that when E4bp4 was phosphorylated by CKIε it became less transcriptionally active and more prone to proteasomal degradation (Doi et al., 2004). However, the first study to identify E4bp4 as a phosphoprotein by Chen *et al*. suggested that phosphorylation was required for E4bp4's ability to bind DNA (Chen et al., 1995). This suggestion contradicts what is observed here and what was observed by Doi *et al.* (Doi et al., 2004), as E4bp4 is more active without phosphorylation you would expect it to be able to still bind target DNA sequences more effectively. Significantly, Chen *et al*. made their observations from experiments performed *in vitro* with E4bp4 produced from non-mammalian cells (Chen et al., 1995), which may why explain their findings do not agree with what was observed here.

Like E4bp4, the activities of numerous TFs are negatively regulated by phosphorylation and multiple phosphorylation events of the same protein can often sequentially reduce activity.

For example in T cells, NFAT has 13 phosphorylation sites that are dephosphorylated by calcineurin when cells are stimulated (Okamura et al., 2000). Progressive dephosphorylation allows NFAT to localise to the nucleus and increases its transcriptional activity (Okamura et al., 2000). Similarly, phosphorylation of the TF c-Jun, at three sites near its N' terminal DNA-binding domain, act to negatively control its ability to bind DNA and must be removed for the transcription of target genes to be activated (Papavassiliou et al., 1995). This type of regulatory mechanism is essential to control the activity of TFs and prevents aberrant transcription that can often lead to human disease (Cohen, 2014).

Doi *et al*. suggested that CKIε was specifically involved with modifying E4bp4 and was a key regulator of its function (Doi et al., 2004). It was postulated whether disrupting the activity of this kinase would have an effect on NK cell production. Interestingly, the CKIε inhibitor, D4476, promoted the development of NK cells, in the presence and absence of IL-15. Further confirming that phosphorylation has an inhibitory role during NK cell development. However, it was not possible to determine if the inhibitor directly affected the phosphorylation of E4bp4 during NK cell development and it is possible that the inhibitor is having a number of off target effects. Development of a specific antibody against the phosphorylated form of E4bp4 would allow this to be ascertained. Despite the simplicity of this approach, it demonstrates that basic pharmacological intervention can have a significant impact on the NK cell development. This has a number of important applications, but most prominently in the development of NK cells for immunotherapy. It would also be interesting to determine whether this observation could be repeated *in vivo*. If the administration of this type of inhibitor was able to boost an individual's NK cell development, even for a short period of time, it could have significant therapeutic potential.

Many TFs are multiply phosphorylated like E4bp4 and this can sometimes allow signals to be integrated from a number of stimuli as each one can result in a different modification on the TF (Benayoun and Veitia, 2009) (Deribe et al., 2010). There is redundancy of signalling mechanisms involving phosphorylation within the cell, therefore it is likely that many kinases and phosphatases are involved in E4bp4 phosphorylation and de-phosphorylation respectively. Whether E4bp4 is phosphorylated by CKIε *in vivo* is still to be ascertained, but as inhibiting this enzyme can directly affect NK cell development, it is possible that it is important for E4bp4 function. Doi *et al*. suggested CKIε modified S182 (Doi et al., 2004), but the bioinformatic prediction tools used here predicted that S286 may also be phosphorylated by a casein kinase. As this study has shown that this is a critical phosphorylation site for E4bp4 activity, it will be vital to ascertain which kinase modifies this residue. The bioinformatic prediction tools also predicted that E4bp4 may be phosphorylated by a number

of other kinases, for example S301 was predicted to be modified by a type of cyclin dependent kinase, based on the amino acid sequence around the phosphorylated serine, whilst S353 was predicted to be modified by a mitogen activated kinase. Potentially these two different types of kinase may phosphorylate E4bp4, at different residues, following the receipt of different external signals.

Prior to the initiation of this study, a yeast-two-hybrid screen was performed to identify interacting partners of E4bp4. Surprisingly few proteins were identified in the experiment, but an interaction with the SUMO E3 ligase PIAS1 was observed, leading to the investigation into the SUMOylation of E4bp4. However, another protein identified in that screen was homeodomain-interacting kinase 3 (Hipk3). Whether this enzyme phosphorylates E4bp4 has yet to be explored, but it is known to regulate transcription and phosphorylation *in vivo*, and is implicated in drug resistance and inhibition of apoptosis in a number of tumours (Curtin and Cotter, 2004).

Phosphorylation appears to strongly regulate the function of E4bp4, which is analogous to what is observed for other major transcription factors such as p53, NFκB and c-Myc (Holmberg et al., 2002). It is now essential to ascertain how phosphorylation affects the activity of E4bp4 on a molecular level; for example it may affect the stability of the protein, as suggested by Doi *et al*. (Doi et al., 2004) or alternatively it may influence its ability to bind DNA or other components of the general transcriptional complex. As phosphorylation has a similar effect to SUMOylation on the function of E4bp4, it will be interesting to determine if both of these PTMs utilise the same molecular mechanism to induce changes during NK cell development.

CHAPTER 5

Post translational modifications influence the transcriptional activity of E4bp4

5.1 Introduction

The alteration of TF activity by PTMs is a key feature of signal transduction pathways that allow cells to rapidly respond to environmental cues by varying levels of gene expression (Brivanlou and Darnell, 2002) (Deribe et al., 2010). TF activities are tightly controlled by PTMs, as misregulation of transcriptional responses is a fundamental contributor to various human diseases, particularly cancer (Cohen, 2014). A vast array of different PTMs have now been identified which influence the activity of TFs and these commonly alter DNA binding capacity, localisation, stability, protein interactions or transcriptional activity (Benayoun and Veitia, 2009) (Filtz et al., 2014). An individual TF can be modified by multiple PTMs, which work together to achieve a common outcome; one PTM may even promote the modification of TF by another PTM (Benayoun and Veitia, 2009) (Deribe et al., 2010). Alternatively, different PTMs may negatively cooperate and actively prevent modification of a TF by another PTM. TFs can often control gene expression of unrelated genes in different cell types and so it is important they are not globally activated. Multiple PTMs of the same TF can help to produce tissue or cell specific effects, but studying the significance and frequency of these interdependent sequential modifications is still at an early stage (Filtz et al., 2014).

E4bp4 is a key TF for a range of biological processes, many involving cells of the immune system (Male et al., 2012). In particular, E4bp4 has been identified as an essential regulator of NK cell and ILC development (Geiger et al., 2014) (Male et al., 2014) (Seillet et al., 2014b). Despite its prominence, it is still not clear how E4bp4 functions on a molecular level to regulate the expression of different target genes, in different cell types. E4bp4 was first identified as a transcriptional repressor and has been shown to repress the expression of numerous target genes *in vivo*, including *IL-13* in T_H2 cells (Kashiwada et al., 2011a). Equally, E4bp4 has been found to transactivate the expression of various target genes, including *Id2* and *Eomes* in NK cells (Male et al., 2014) and the immunoglobulin germ-line ε gene in B cells, required for IgE class switching (Kashiwada et al., 2010). Some evidence now suggests that E4bp4 is able to interact with a number of chromatin modifying enzymes that may help to influence the transcription of target genes (Keniry et al., 2013) (Motomura et al., 2011) (Tong et al., 2013). E4bp4 was observed to interact with the histone H2A ubiquitinase MYSM1, and this interaction was identified to be crucial for the expression of *Id2* and ultimately the development of NK cells (Nandakumar et al., 2013).

It is unclear how E4bp4 is able to influence transcriptional activation and repression of different target genes, and how the activities of this TF are regulated. In this study E4bp4 was found to be post translationally modified by both SUMOylation and phosphorylation. Both of these PTMs appear to function in a similar manner, repressing the activity of E4bp4. When phosphorylation or SUMOylation sites of E4bp4 are mutated, the activity of the protein increased. This was observed by the greater number of NK cells produced from Lin^- BM cells, grown *in vitro* and transduced with E4bp4 mutants, compared to WT E4bp4. However the underlying mechanism for this phenotype was not known, or whether the mechanism was the same for both phosphorylation and SUMOylation.

This study aimed to establish the molecular mechanism underlying the post translational regulation of E4bp4 during NK cell development. In order to achieve this, the following objectives were set:

1. Ascertain whether PTMs affect the stability of E4bp4.
2. Determine whether the PTMs of E4bp4 influence its transcriptional activity using a luciferase reporter assay.
3. Use quantitative PCR to determine whether the PTMs of E4bp4 are important for its transcriptional activity in NK cells.

5.2 The PTMs of E4bp4 do not affect its stability or proteasomal degradation

SUMOylation and phosphorylation both influence the stability of TFs. Multi-site phosphorylation of the TF ATF4 affects its ubiquitination and ultimately its stability, which is essential for regulation of the cell cycle (Frank et al., 2010). Phosphorylation of ATF4 acts progressively to control its degradation, allowing for dose-dependent regulation of target genes during neurogenesis (Frank et al., 2010). Despite being similar in structure to ubiquitin, SUMOylation does not directly target proteins for proteasomal degradation, but it can promote ubiquitin-mediated degradation, which has been observed for the TF PML (Lallemand-Breitenbach et al., 2008). Arsenic trioxide, used to treat acute myeloid leukaemia, can induce SUMOylation of PML in leukemic cells and these SUMO modifications help to recruit the ubiquitin-ligase RNF4 (Lallemand-Breitenbach et al., 2008).

This leads to polyubiquitination of PML, its subsequent proteasomal degradation and therefore a reduction in the oncogenic activity of PML in leukemic cells (Lallemand-Breitenbach et al., 2008). Conversely, SUMOylation can stabilise target proteins by blocking ubiquitination at a particular lysine residue, which has been shown for IκBα (Desterro et al., 1998). SUMO1 modification of IκBα at lysine 21 prevents ubiquitination and protects the protein from degradation, preventing signal-induced activation of NFκB-dependent transcription (Desterro et al., 1998).

As both phosphorylation and SUMOylation negatively regulate the activity of E4bp4, it was hypothesised that both of these modifications could promote the proteasomal degradation of the protein. Alternatively, the Lys to Arg mutations that were made in the E4bp4 SUMOylation mutants could prevent E4bp4 from being ubiquitinated if one of the mutated lysine residues was also a site of ubiquitination (Tatham et al., 2009). Doi *et al.* previously suggested the phosphorylation of E4bp4 induces its proteasomal degradation (Doi et al., 2004). However, this was only inferred by the fact that, when E4bp4 was coexpressed with Casein kinase I, a small reduction in the production of E4bp4 was observed. This was subsequently abolished when the proteasome was chemically inhibited by MG132 (Doi et al., 2004). This experiment was also performed with transiently overexpressed protein which may not accurately represent how the protein functions *in vivo*. To determine whether SUMOylation and phosphorylation regulate the degradation of E4bp4, a cycloheximide time course assay was used to compare the half-life of E4bp4 in the presence and absence of its PTMs.

Cycloheximide is an inhibitor of the mammalian translational machinery and once active inside cells new peptides cannot be synthesised (Schneider-Poetsch et al., 2010). To determine a protein's stability, samples of cycloheximide treated cells can be analysed every few hours to establish the point at which the protein is no longer detectable (Schneider-Poetsch et al., 2010). Cycloheximide assays have previously been used to analyse the stability of numerous TFs including the oncogene Myc (Yeh et al., 2011), the circadian clock protein Per2 (Asher et al., 2008) and the tumour suppressor p53 (Shen et al., 2014).

To perform the cycloheximide time course assay a source of cells that expressed detectable levels of E4bp4 and the E4bp4 PTM mutants were required. It was important to have cells that expressed the E4bp4 mutants at close to endogenous levels so that if either PTM affected stability, it would be relevant to the situation *in vivo*. Therefore, murine 3T3 cells were stably transduced with retroviruses produced from the pMSCV-IRES-hCD2 construct

(the same as was used in the NK cell development assay) to create stable cell lines that expressed WT E4bp4, E4bp4 S286-353A or E4bp4 5X SUMO **(Figure 5.1)**. The S286-353A and 5X SUMO mutants were analysed as these had produced the strongest phenotypes in the NK cell development assays.

Following transduction, all hCD2 positive cells in each sample were FACS sorted, maintained in culture and the level of hCD2 expression was checked using flow cytometry every 3 - 4 days **(Figure 5.1)**. To confirm successful transduction, the cell lines were analysed by WB for their expression of E4bp4. Nuclear extracts were created from each cell line and E4bp4 was present in each of the nuclear extracts **(Figure 5.1B)**. This confirmed that the SUMOylation and phosphorylation site mutations did not interfere with the ability of E4bp4 to translocate to the nucleus **(Figure 5.1B)**. However, this was not unexpected as the mutants both have greater functionality than WT E4bp4.

Each transduced cell line was subjected to 50 µg/ml cycloheximide for 0 - 16 hours, with samples lysed at regular time intervals. Nuclear extracts were prepared and compared by WB for E4bp4 expression **(Figure 5.2)**. The 5X SUMO mutant, the S286-353A mutant and WT E4bp4 all had very similar half-lives; the protein level reduced by almost half after 8 hours and after 12 hours of cycloheximide treatment, E4bp4 was undetectable in all samples **(Figure 5.2)**. Neither SUMOylation nor phosphorylation affected the stability of E4bp4 and so it was concluded that these PTMs must use a different mechanism to influence the activity of E4bp4.

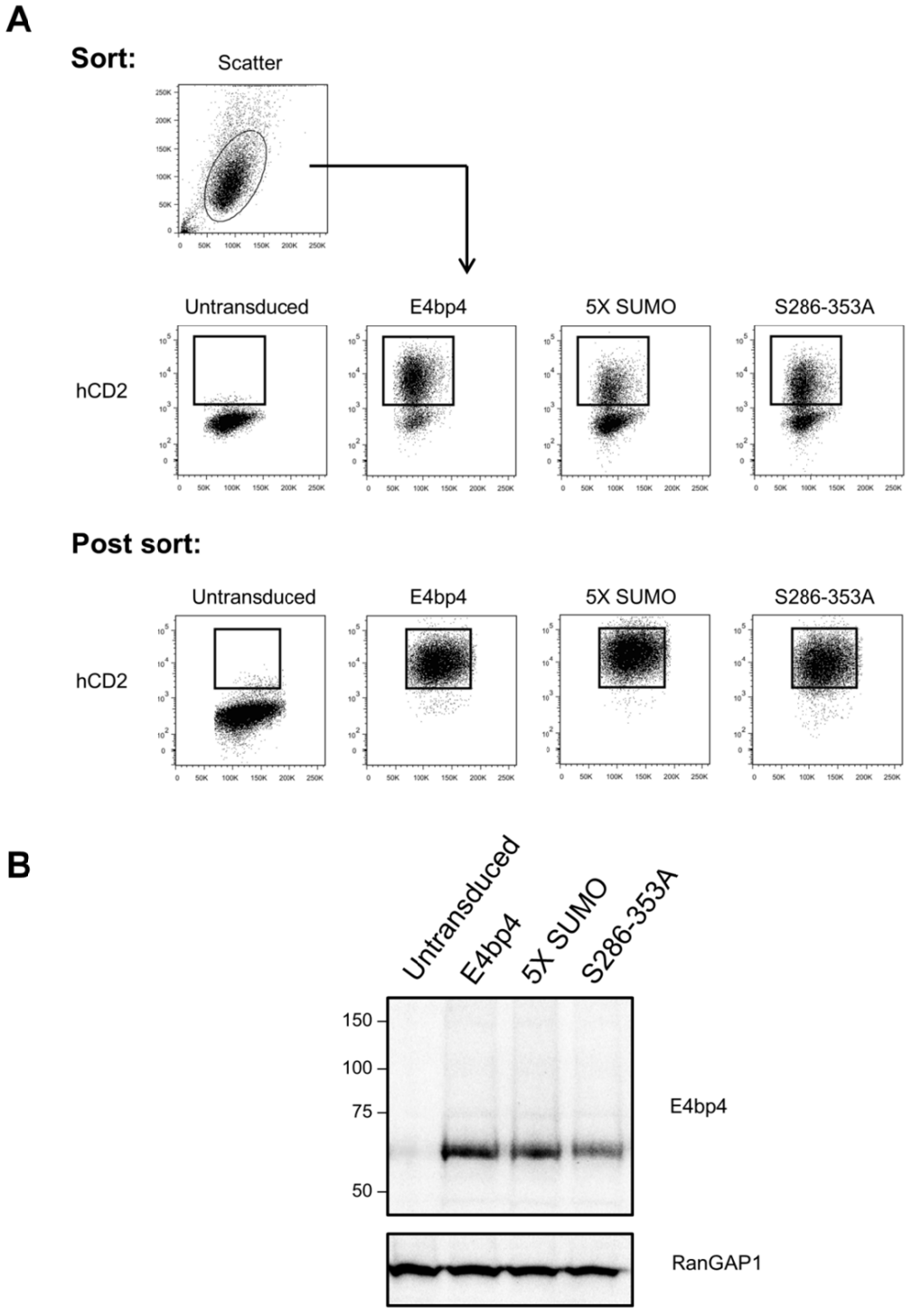


Figure 5.1 – Development of stable cell lines expressing E4bp4 mutants

Murine 3T3 were transduced with MSCV-IRES-hCD2 expressing WT E4bp4, 5X SUMO or S286-353A E4bp4. A) Seven days after transduction cells were bulk sorted based on hCD2 expression using a FACS Aria II. Sorted cells were cultured for a further five days before hCD2 expression was analysed by flow cytometry and (B) E4bp4 expression was analysed by WB, comparing nuclear extracts of each cell line. Probing with RanGAP1 was used as a loading control. Molecular weight markers are showing in kDa.

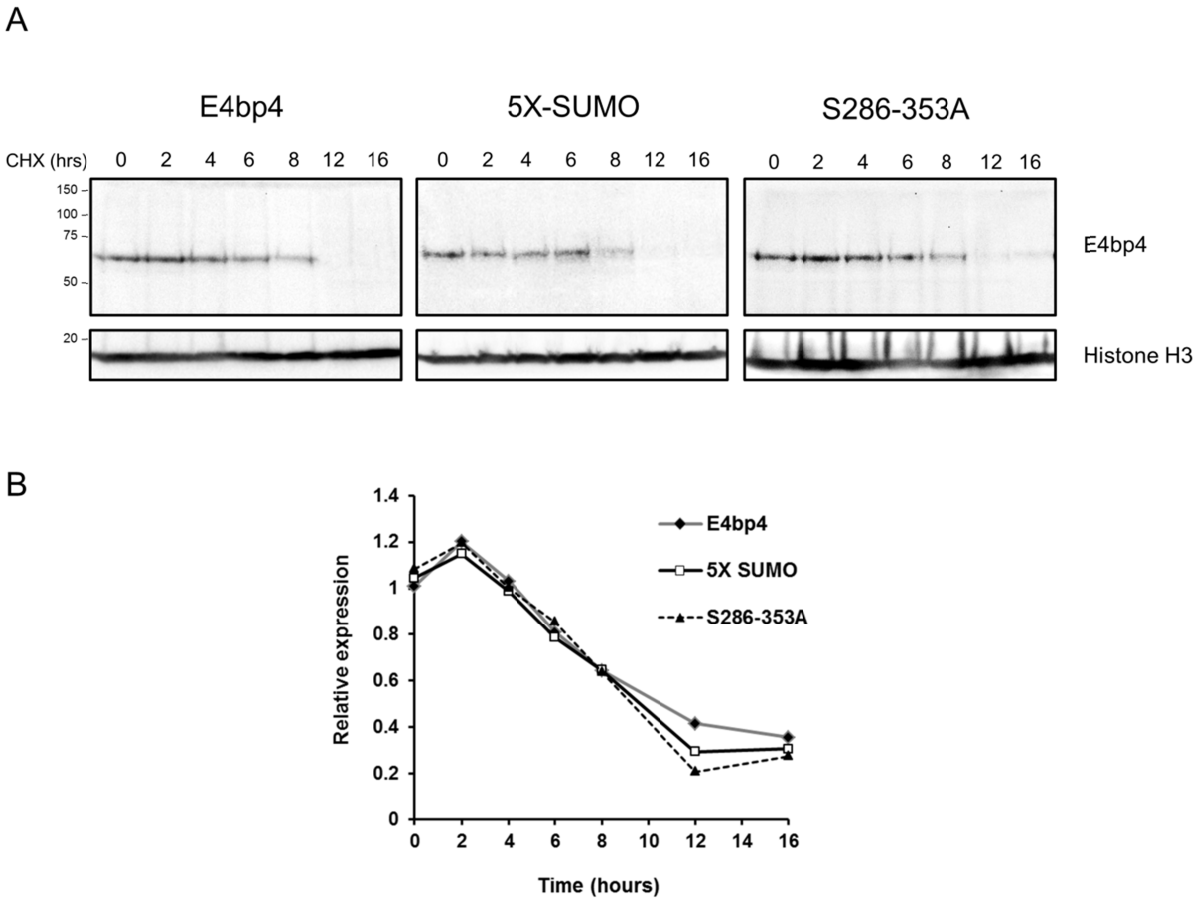


Figure 5.2 – SUMOylation and phosphorylation do not influence the stability of E4bp4

Stable 3T3 cell lines expressing WT E4bp4, 5X SUMO E4bp4 and S286-353A E4bp4 were cultured in the presence of 50 µg cycloheximide for up to 16 hours. At regular intervals cells from each cell line were lysed and nuclear extracts were prepared. A) Nuclear extracts were compared by WB for the presence of E4bp4 and probing for histone H3 was used as a loading control. Molecular weight markers are showing in kDa. B) Relative expression levels of E4bp4 were determined using Image Lab software (Bio-rad) to calculate densitometry of bands.

5.3 PTMs of E4bp4 regulate its transcriptional activity

5.3.1 Transcriptional repression of target genes by E4bp4

There are many of reports of SUMOylation influencing the transcriptional activities of TFs and there appears to be a strong association between SUMOylation and enhanced transcriptional repression (Gill, 2005). There are suggestions that this may be due to SUMO promoting interactions between TFs and chromatin remodelling proteins that induce transcriptionally repressive chromatin structures (Lee and Muller, 2009) (Ouyang et al., 2009). There are also reports of SUMOylation stimulating transcriptional activity, or at least dampening transcriptional repression, mediated by a number of TFs (Yan et al., 2010) (Zhang et al., 2012). Phosphorylation has also commonly been found to regulate TF function by influencing interactions with transcriptional cofactors, promoting transactivation through phosphorylation dependent activation domains and, in some cases, repression domains (Whitmarsh and Davis, 2000).

To establish if SUMOylation or phosphorylation regulates the transcriptional activity of E4bp4, a transcription reporter gene assay was established **(Figure 5.3)**. Using the pMSCV-IRES-hCD2 construct, 3T3 cells were transfected and were able to express E4bp4. The expression of hCD2 allowed the levels of transfection to be analysed using flow cytometry, to ensure E4bp4 was expressed in all cells **(Figure 5.3)**. The 3T3 cells were also transfected with a pGL3-Luc construct that contained a promoter controlling the expression of the Firefly Luciferase gene **(Figure 5.3)**. Upstream of the promoter was either three copies of the E4bp4 DNA-binding consensus sequence, identified by Cowell *et al*. (Cowell et al., 1992) or the 2 kb region upstream of *Eomes* or *Id2*, in the murine genome, both of which contain E4bp4 binding sites (Male et al., 2014). As an additional control, the 3T3 cells were also transfected with a second luciferase construct containing the Renilla Luciferase gene, which acted as an internal control for luciferase activity **(Figure 5.3)**. 3T3 cells were used in this assay system as they are of murine origin and they allowed consistent transfections of the three DNA constructs in each cell. Ideally the murine NK cell line MNK-1 (Allan et al., 2014) would have been used for this assay, but previous work from the Brady laboratory had shown transgene expression in these cell could only be achieved using lentiviral transduction.

Initially the activity of E4bp4 in the luciferase assay was compared using the different pGL3-Luc constructs **(Figure 5.4A)**. The upstream genomic regions of *Eomes* and *Id2* did not cause E4bp4 to alter the transcription of the luciferase gene, but as previously reported

(Yeung et al., 2004), E4bp4 significantly reduced the transcriptional activity of luciferase when the E4bp4 DNA-binding consensus sequence (E4bp4-CS) was present upstream of the promoter **(Figure 5.4A)**. Therefore this construct was used to compare the transcriptional activity of the different E4bp4 SUMOylation and phosphorylation site mutants.

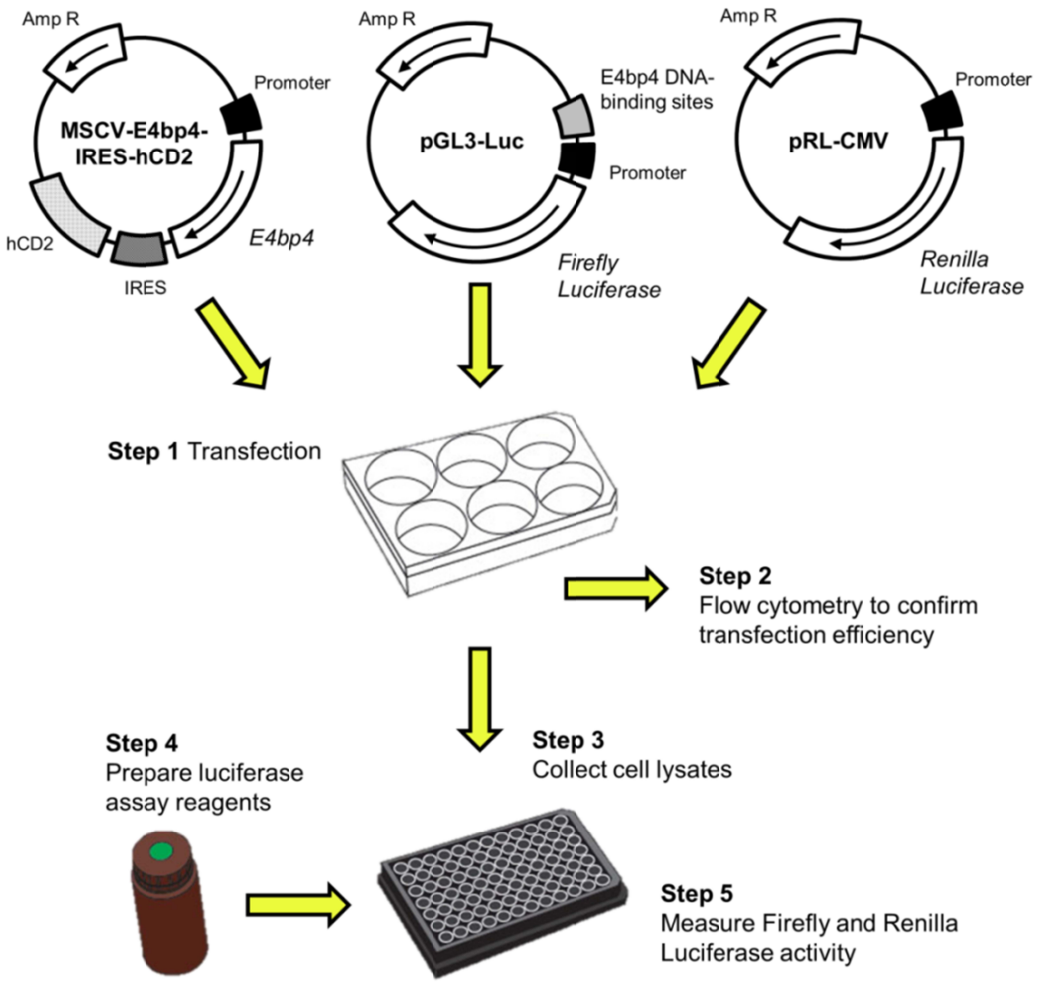


Figure 5.3 – Dual luciferase reporter assay system

Schematic representation of the assay system used to measure transcriptional activity of E4bp4. Murine 3T3 cells were transfected with three constructs; MSCV-IRES-hCD2 expressing E4bp4 (or mutants), pGL3 expressing the Firefly luciferase gene with three copies of the E4bp4 DNA-binding site upstream (or empty vector); and pRL-CMV expressing the Renilla luciferase gene. Forty-eight hours after transfection, some cells were removed and analysed by flow cytometry for the expression of hCD2. Remaining cells were lysed and cell lysates were transferred to a 96-well plate where they were mixed with luciferase assay reagents, allowing the luminescent signals to be quantified. The assay system allowed sequential measurements of both Firefly and Renilla luciferase activities from a single sample.

The WT form of E4bp4, the K10-116-219R E4bp4 mutant and 5X SUMO mutant were transfected into 3T3 cells with pGL3-E4bp4-CS to compare their transcriptional activity. Forty-eight hours after transfection, the cells were analysed by flow cytometry, which demonstrated all cells were consistently transfected with the MSCV-IRES-hCD2 constructs, due to their expression of hCD2 **(Figure 5.4B)**. Both the K10-116-219R and 5X SUMO mutants caused greater transcriptional repression of the luciferase gene than the WT E4bp4 **(Figure 5.4D)**. The 5X SUMO mutant, which lacked all SUMO modifications gave the strongest transcriptional repression and this was significantly lower than the empty vector and WT E4bp4 **(Figure 5.4D)**. This transcriptional repression was only observed when the E4bp4 DNA-binding sites were present upstream of the luciferase gene promoter **(Figure 5.4)**.

Five of the E4bp4 phosphorylation mutants were also analysed using this assay, in particular those that were previously found to affect NK cell development. The S286A, S286-353A and S286-301A mutants caused greater transcriptional repression than WT E4bp4, in a similar manner to the 5X SUMO mutant **(Figure 5.5)**. However, the S210A, S301A and S353A mutants did not affect E4bp4's ability to repress transcription of the luciferase gene **(Figure 5.5)**. Removing the phosphorylation and SUMOylation modifications from E4bp4 appeared to increase its transcriptional activity, making it a more potent transcriptional repressor. Additionally, the same E4bp4 mutants that affected NK cell development also affected the transcriptional activity of E4bp4, suggesting that this is a potential mechanism by which E4bp4 PTMs may regulate NK cell development.

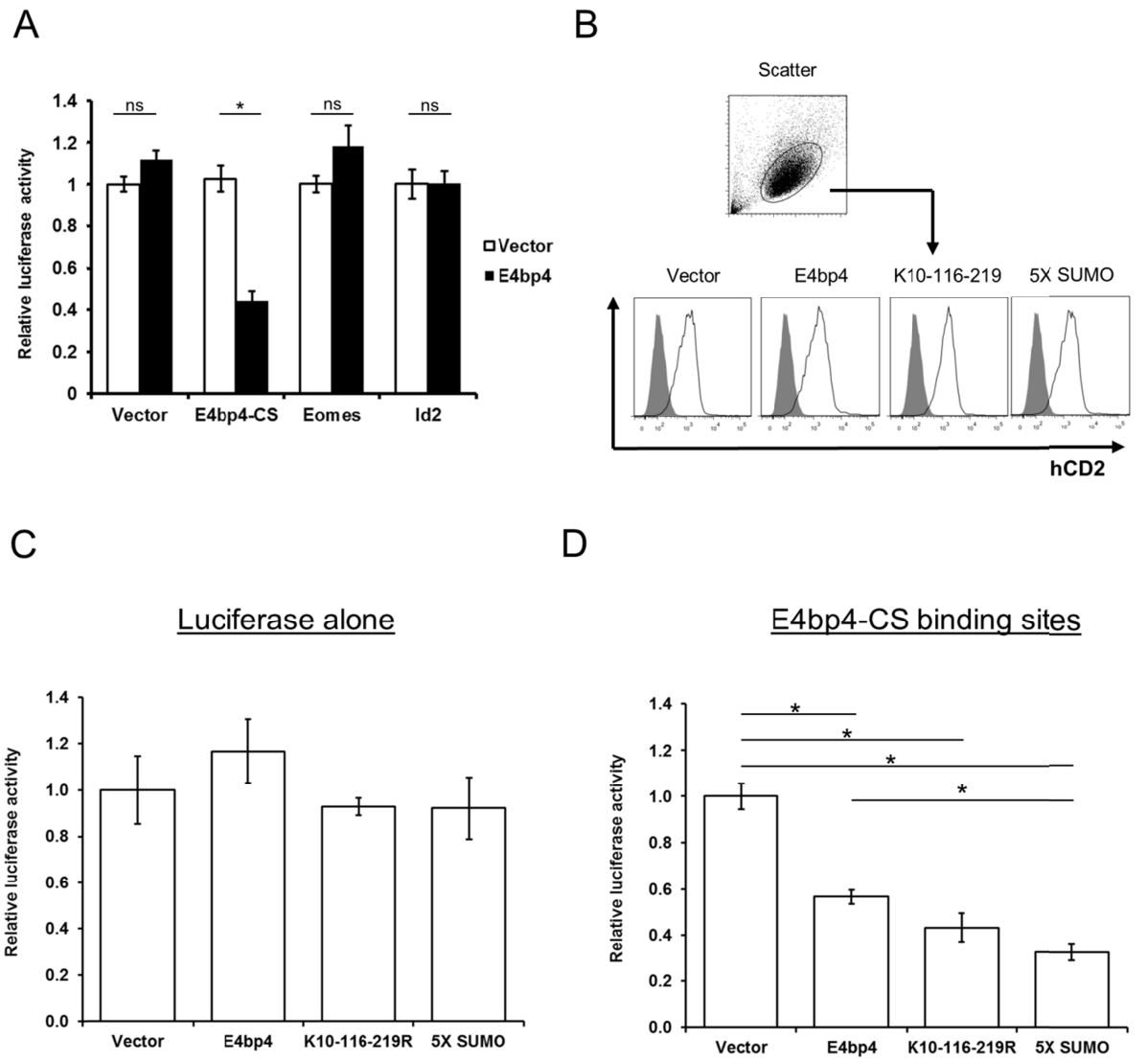


Figure 5.4 – Removing SUMOylation sites of E4bp4 affects its transcriptional activity

E4bp4's transcriptional activity was analysed using a luciferase reporter gene. Murine 3T3 cells were transfected with pMSCV-E4bp4-IRES-hCD2 (or empty vector) and the pGL3 reporter gene construct, containing either three copies of the Cowell *et al*. E4bp4 DNA-binding consensus sequence (E4bp4-CS), the upstream region of *Eomes* or the upstream region of *Id2*, located upstream of the promoter controlling the Firefly luciferase gene (or an empty vector – with luciferase alone). Cells were also transfected with the normalisation vector pRL-CMV (Renilla luciferase under control of a CMV promoter). After 48 hours luciferase activities of each sample were measured and all Firefly activities were normalised against Renilla activity. A) The transcriptional activity of E4bp4 was analysed with the different pGL3-Luc constructs. B - D) The E4bp4 K10-116-219R and E4bp4 5X SUMO mutants were compared to WT E4bp4 for their ability to influence the transcriptional activity of a luciferase reporter gene from the pGL3-E4bp4-CS construct. B) 48 hours after transfection, some of the cells from each sample were analysed by flow cytometry for the expression of hCD2, to confirm E4bp4 expression. Untransduced cells were used as a negative control (grey). C - D) Luciferase activities were measured and all Firefly activities were normalised against Renilla activity. Results shown are combined from four independent experiments, each result was determined relative to MSCV-IRES-hCD2 (vector) and are expressed as a mean ± S.E.M. Statistical differences between samples were determined using unpaired student T-tests (*, $p < 0.05$).

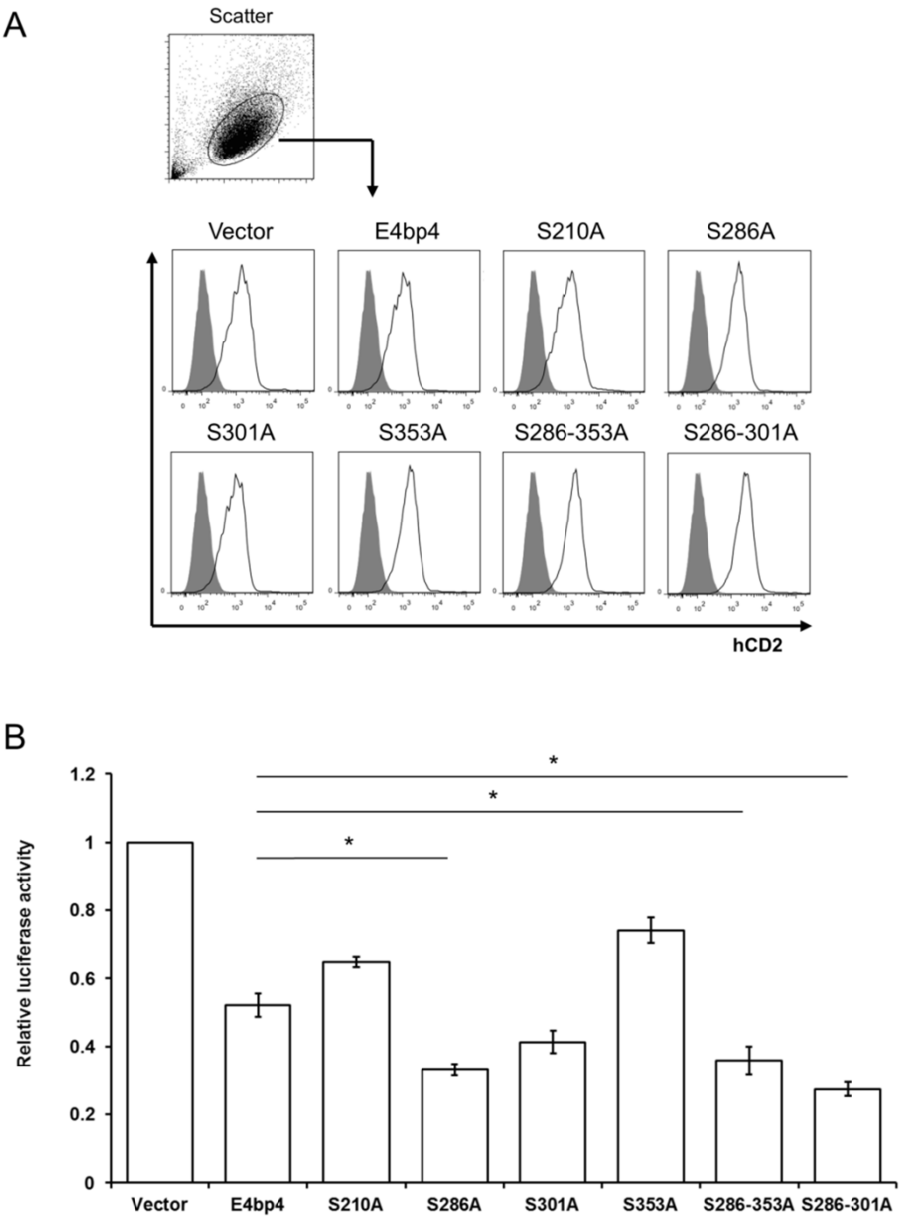


Figure 5.5 – Removing phosphorylation sites of E4bp4 affects its transcriptional activity

The E4bp4 S210A, S286A, S301A, S353A, S286-353A and S286-301A mutants were compared to WT E4bp4 for their ability to influence the transcriptional activity of a luciferase reporter gene. Mouse 3T3 cells were transfected with pMSCV-E4bp4-IRES-hCD2 (or E4bp4 phosphorylation mutants), the pGL3 reporter gene construct, containing the three copies of the Cowell *et al*. E4bp4 DNA-binding consensus sequence (Cowell et al., 1992) upstream of the promoter controlling the Firefly luciferase gene and the normalisation vector pRL-CMV (Renilla luciferase under control of a CMV promoter). A) 48 hours after transfection, some of the cells from each sample were analysed by flow cytometry for the expression of hCD2, to confirm E4bp4 expression. Untransduced cells were used as a negative control (grey). B) Luciferase activities were measured and all Firefly activities were normalised against Renilla activity. Results shown are combined from four independent experiments, each result was determined relative to MSCV-IRES-hCD2 (vector) and are expressed as a mean ± S.E.M. Statistical differences between samples were determined using unpaired student T-tests (*, $p < 0.05$).

5.3.2 Transcriptional activation of target genes by E4bp4

As well as acting as a transcriptional repressor, E4bp4 has also been reported to activate the transcription of its target genes (Kashiwada et al., 2011a) (Male et al., 2014) (Zhang et al., 1995). As E4bp4 did not induce transcriptional activation of the luciferase gene in the presence of promoter regions from *Eomes* and *Id2* **(Figure 5.4A)** an alternative approach was developed. Zhang *et al*. identified E4bp4 in human T cells as a regulator of IL-3 expression and was the first to demonstrate that it could promote transcription of target genes (Zhang et al., 1995). In this study, Zhang *et al*. identified a minimal E4bp4 DNA-binding consensus sequence that was required for transactivation of target genes (Zhang et al., 1995). Interestingly, this sequence differed slightly from the E4bp4 DNA-binding consensus sequence identified by Cowell *et al*. who discovered E4bp4 as a transcriptional repressor **(Table 5.1)**. As SUMOylation and phosphorylation both affected E4bp4's ability to repress transcription using the Cowell *et al*. binding sites (in pGL3-E4bp4-CS), it was postulated whether these would have a similar effect with transcriptional activation.

Table 5.1 – E4bp4 DNA binding sequences

Name	Minimal DNA binding sequence	Activity	Reference
E4bp4-CS	TTATGTAAC	Repression	Cowell *et al*. (1992)
E4bp4-CS-ZH	CAGACGTAATTA	Activation	Zhang *et al*. (1995)

To clone the Zhang *et al*. E4bp4 DNA-binding consensus sequence into the pGL3 Firefly Luciferase vector (pGL3-E4bp4-CS-ZH), an oligonucleotide was synthesised of the desired sequence which contained three copies of the minimal consensus DNA-binding sequence **(Table 5.1) (Figure 5.6)**. The oligonucleotide was amplified by PCR, restriction digested and ligated into the pGL3 vector, upstream of the Luciferase promoter **(Figure 5.6)**. The ligation product was transformed into competent *E. coli* and colony PCR was used to identify transformed bacterial colonies containing the correct DNA insert **(Figure 5.6)**. Additionally, to confirm that the cloning was successful vector DNA was fully sequenced. The pGL3-E4bp4-CS-ZH construct was identical to the previously used pGL3-E4bp4-CS construct except for the three copies of the twelve base-pair minimal E4bp4 DNA-binding site **(Figure 5.6)**.

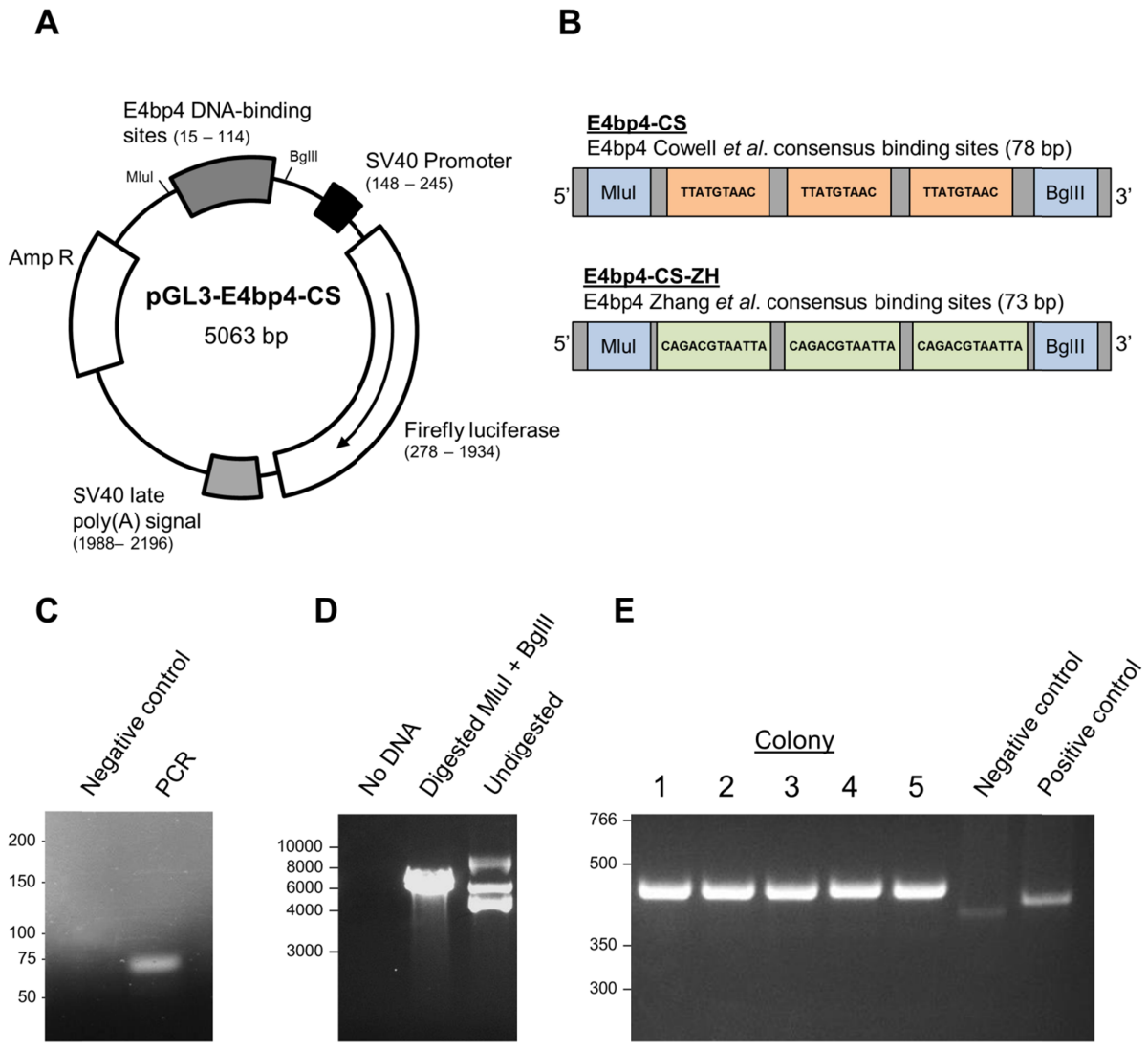


Figure 5.6 – Cloning of Zhang *et al.* E4bp4 binding sites into the pGL3 luciferase vector

A) Schematic representation of the pGL3 Firefly luciferase vector containing the E4bp4 consensus DNA-binding sites upstream of the SV40 promoter controlling expression of the Firefly Luciferase gene. B) Schematic representation of the different oligonucleotides cloned into the pGL3 vector, containing E4bp4 consensus DNA-binding sites identified by Cowell *et al.* (Cowell et al., 1992) and Zhang *et al.* (Zhang et al., 1995). The Cowell *et al.* sequence was previously cloned into the vector (Yeung et al., 2004). C) The Zhang *et al*. oligonucleotide was amplified by PCR, with the PCR product analysed on an agarose gel, before digestion with MluI and BglII restriction enzymes. D) The pGL3 vector was digested with the same restriction enzymes, with the products analysed on an agarose gel to confirm successful digestion. E) Digest PCR product and vector were ligated together and transformed in competent *E. coli* and colony PCR was used to identify successfully transformed colonies. Colonies containing the correct construct produced a PCR product of 444 bp, produced by primers that amplified across the MluI/BglII cloning site. Molecular weight markers are shown in bp.

Using the pGL3-E4bp4-CS-ZH construct in the Luciferase reporter gene assay caused WT E4bp4 to promote the activity of the Firefly Luciferase gene when compared to the empty vector **(Figure 5.7)**. This contrasted with the phenotype observed with the E4bp4-CS construct, which caused E4bp4 to repress transcription of the luciferase gene **(Figure 5.7)**. This demonstrated that E4bp4 was able to recognise more than one DNA-binding site and these different sites can influence its transcriptional activity in different ways. When the E4bp4 5X SUMO and S286-353A mutants were expressed with the pGL3-E4bp4-CS-ZH luciferase construct, they caused greater transcriptional activation than WT E4bp4 and empty vector **(Figure 5.8)**. The activation by the S286-353A mutant was both higher than with the 5X SUMO mutant and was also significantly higher than WT E4bp4 **(Figure 5.8)**. The SUMOylation and phosphorylation of E4bp4 therefore negatively regulated its transcriptional activity, regardless of whether target gene activity was being activated or repressed. Removal of either PTM allowed E4bp4 to become a more potent transcriptional repressor and to a lesser extent, a more potent transcriptional activator.

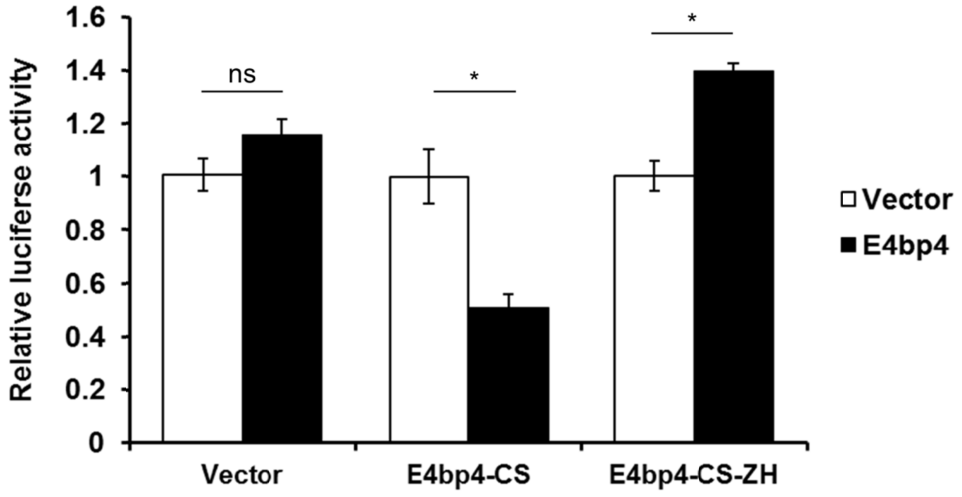


Figure 5.7 – E4bp4 can activate and repress transcription of target genes, depending on variations in the sequence of DNA binding sites

The different E4bp4 consensus DNA-binding sites were compared for their ability to influence transcription in the presence of E4bp4. 3T3 cells were transfected with MSCV-IRES-hCD2 (with and without E4bp4), along with pGL3-E4bp4-CS (containing the Cowell *et al.* binding sites), pGL3-E4bp4-CS-ZH (containing the Zhang *et al.* binding sites) or no binding sites (vector). Cells were also transfected with pRL-CMV to act as an internal control. 48 hours after transfection Luciferase activities were measured and all Firefly activities were normalised against Renilla activity. Results shown are combined from three independent experiments, each result was determined relative to MSCV-IRES-hCD2 (vector) and are expressed as a mean ± S.E.M. Statistical differences between samples were determined using unpaired student T-tests (*, $p < 0.05$).

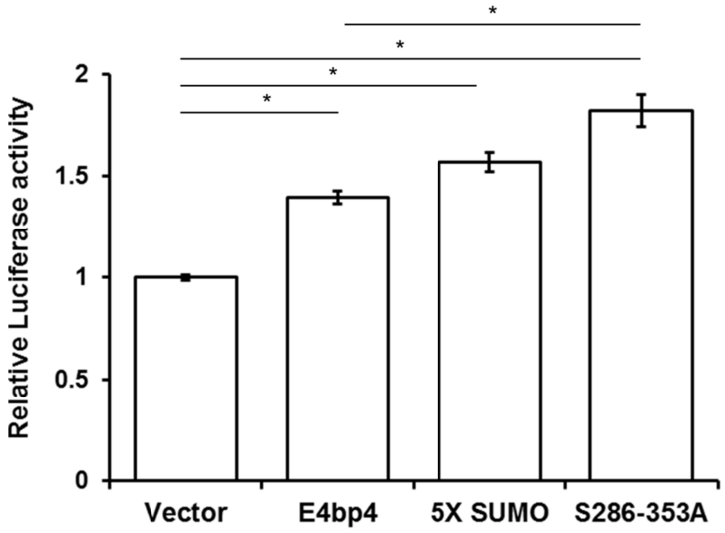


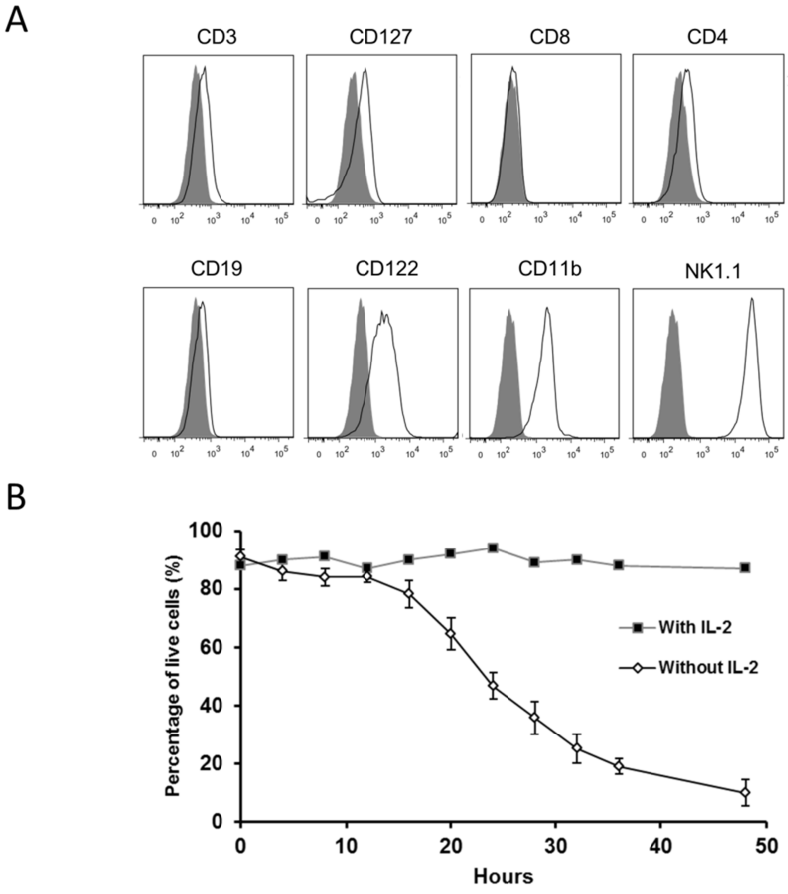
Figure 5.8 – Mutation of the SUMOylation and phosphorylation sites of E4bp4 influences its ability to cause transcriptional activation

WT E4bp4, the 5X SUMO mutant and S286-353A mutant were transfected into 3T3 cells along with the pGL3-E4bp4-CS-ZH construct, containing three E4bp4 DNA-binding sites identified by Zhang *et al.* (Zhang et al., 1995). Cells were also transfected with pRL-CMV to act as an internal control. 48 hours after transfection Luciferase activities were measured and all Firefly activities were normalised against Renilla activity. Results shown are combined from three independent experiments, each result was determined relative to MSCV-IRES-hCD2 (vector) and are expressed as a mean ± S.E.M. Statistical differences between samples were determined using unpaired student T-tests (*, $p < 0.01$).

5.4 PTMs of E4bp4 influence target gene expression in NK cells

As SUMOylation and phosphorylation influenced the transcriptional activity of E4bp4 during the luciferase reporter assay, it was critical to explore whether a similar effect would be observed with E4bp4 target genes *in vivo*. As E4bp4 PTMs influence NK cell development it was postulated that target gene expression in NK cells would be modulated. Previously, Male *et al*. demonstrated that when E4bp4 is transduced into the murine NK cell line MNK-1 (Allan et al., 2014) it is able to transactivate the expression of target genes *Id2* and *Eomes* (Male et al., 2014). To further analyse the effect of SUMOylation and phosphorylation on the transcriptional activity of E4bp4, the 5X SUMO and S286-353A mutants were expressed in MNK-1 cells and the transcription of target genes was compared to WT E4bp4 or an empty vector. The MNK-1 cells were confirmed to have an NK cell phenotype as they expressed a

number of NK cell surfaces markers, including high levels of the NK cell marker NK1.1 **(Figure 5.9A)**. The MNK-1 cells were also highly dependent on IL-2, as 24 hours after it was removed from the culture medium cell viability had reduced by 50% **(Figure 5.9B)**. The MNK-1 cells were originally derived from NK1.1^{+} thymocytes and despite their NK cell phenotype expressed very low levels of endogenous E4bp4 (Allan et al., 2014). This meant they were ideal for analysing the different E4bp4 mutants as high levels of endogenous E4bp4 would interfere with the transcriptional effects of the mutant forms that were transduced into the MNK-1 cells.

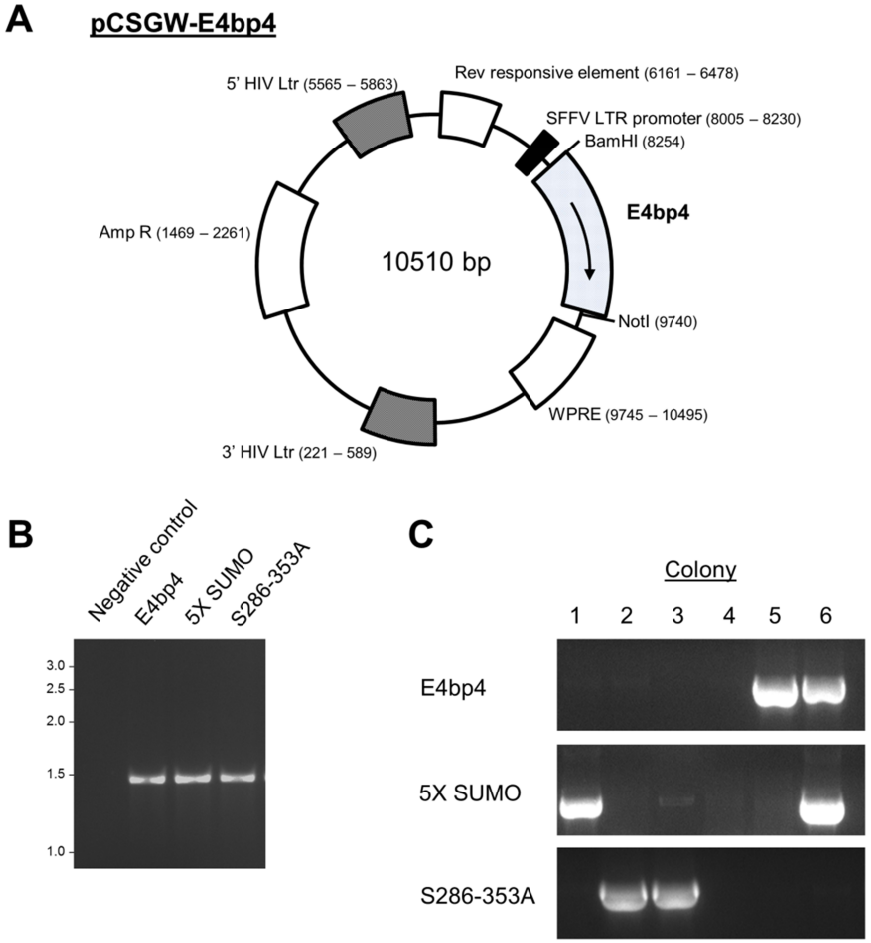


5.9 – MNK-1 cells have an NK cell phenotype.

A) MNK-1 cells were analysed by flow cytometry for the expression of numerous common lymphocyte cell surface markers. Unstained cells were used as a negative control (grey). B) MNK-1 cells were cultured in the presence or absence of IL-2 for 48 hours and at regular time points the cell viability was determined using trypan blue staining and a hemocytometer. Results shown are combined from three independent experiments and expressed as a mean ± S.E.M.

In order to express E4bp4, or the E4bp4 PTM mutants, in the MNK-1 cells, the cDNA of each was cloned into a lentiviral expression vector **(Figure 5.10)**. This vector allows high levels of transduction of a wide range of cell types and readily integrates into host genomes

to allow long-term expression of transgenes at near endogenous levels (Tiscornia et al., 2006). Previous work in the Brady laboratory had also established that retroviruses could not be used to successfully transduce the MNK-1 cells. The 5X SUMO, S286-353A and WT E4bp4 cDNAs were PCR amplified from the pMSCV-IRES-hCD2 vector, restriction digested and ligated into the pCSGW lentiviral vector **(Figure 5.10)**. Ligation products were transformed into competent *E. coli* and colonies successfully transformed with the correct vector were identified by colony PCR **(Figure 5.10)**. Additionally, vector DNA was fully sequenced to confirm successful cloning.

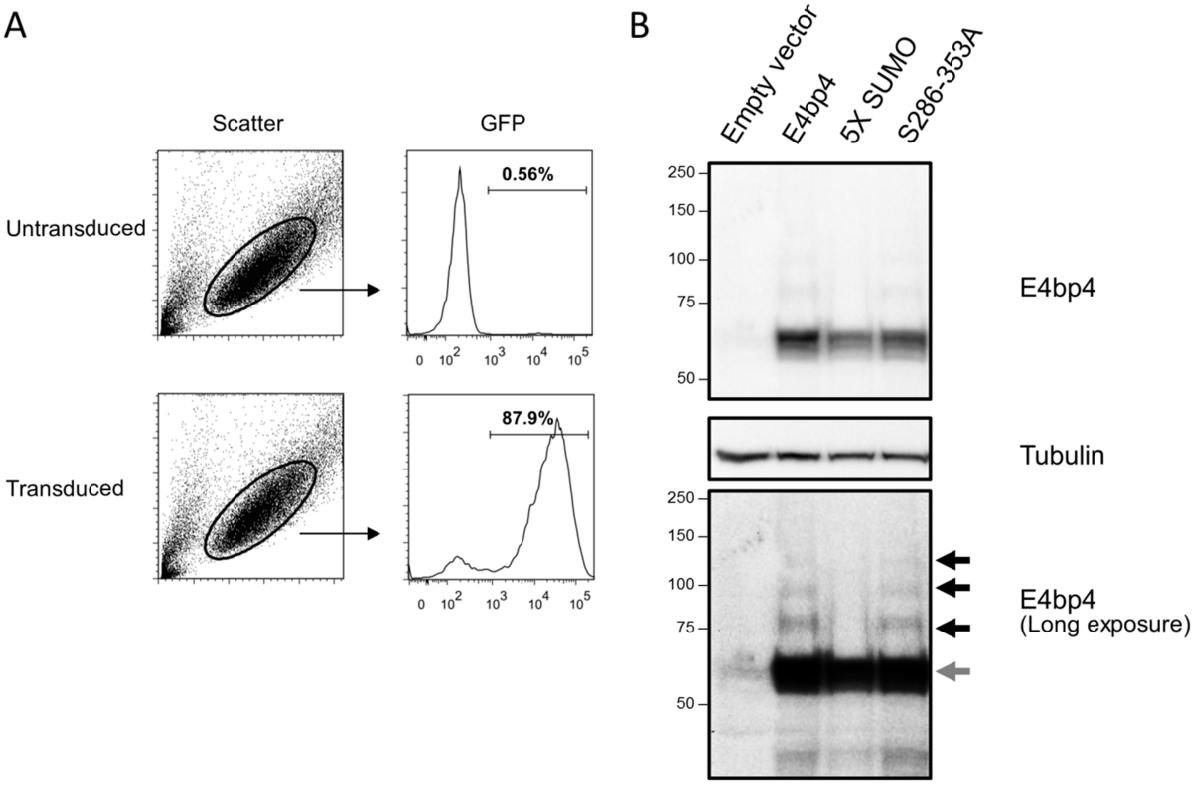


5.10 – Cloning E4bp4 mutant cDNA into pCSGW lentiviral expression vector

WT E4bp4, the 5X SUMO mutant and the S286-353A mutant cDNA were all cloned into the pCSGW lentiviral expression vector to allow expression in MNK-1 cells and other cells types. A) Schematic representation of the pCSGW vector which contains: a spleen focus-forming virus LTR promoter (SFFV LTR); 5' and 3' long terminal repeats (Ltr) from the HIV virus to allow integration in host genomes; the Rev responsive element which allows export of viral mRNA from the nucleus; and the Woodchuck posttranscriptional response element (WPRE) which enhances transgene expression from viral vectors. B) E4bp4 WT and mutant cDNAs were amplified by PCR using specifically designed primers and the PCR products were analysed on an agarose gel. PCR products were restriction digested and ligated into the pCSGW vector and transformed into competent *E. coli* cells. C) Five transformed colonies for each construct were analysed for the presence of E4bp4 using colony PCR (cloning primers utlised) and the products of the PCR were analysed on an agarose gel. Positive colonies were grown O/N and plasmid DNA purified and sequenced. Molecular weight markers are shown in kb.

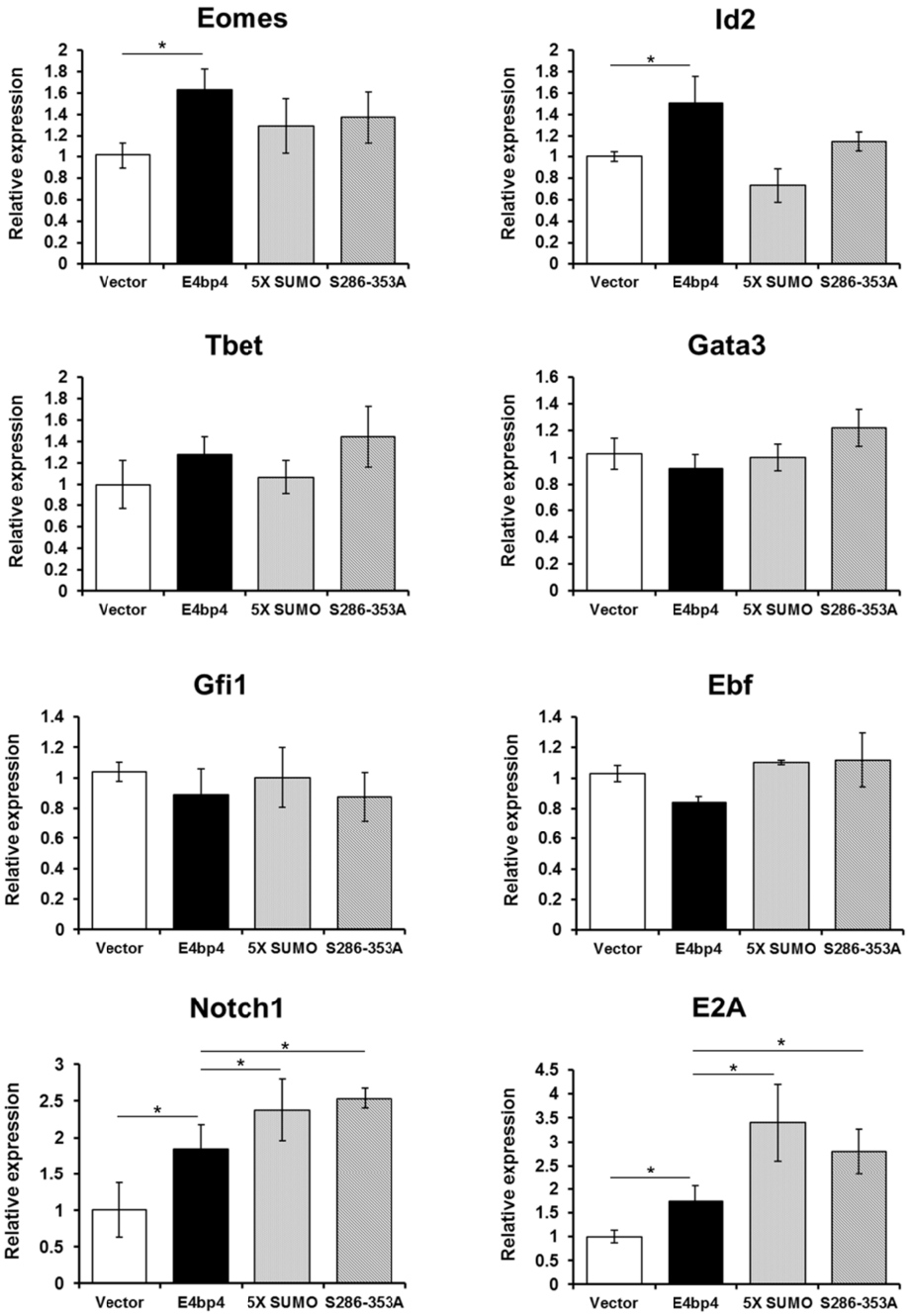
Using the lentiviral expression system it was possible to consistently achieve high levels of transduction of the MNK-1 cells. When analysing transduction by a GFP control construct, over 85% of the cells were GFP^+ **(Figure 5.11A)**. When MNK-1 cells were transduced with pCSGW-E4bp4 or mutant versions of E4bp4, high levels of E4bp4 were detectable in whole cell lysates of transduced cells when analysed by WB **(Figure 5.11B)**. Analysis of the same WB with a long exposure also confirmed that WT E4bp4 expressed in these NK cells is SUMOylated and, as expected, the 5X SUMO mutant does not have any SUMO modifications, as determined by the presence of higher molecular weight forms of E4bp4 **(Figure 5.11B)**. This observation supported the findings in chapter 3, that SUMOylation of E4bp4 is important in NK cell development as it confirms that E4bp4 is SUMOylated in NK cells. The S286-353A mutant had the same higher molecular weight forms as WT E4bp4, suggesting that phosphorylation does not affect SUMOylation **(Figure 5.11B)**.

Using quantitative-PCR (qPCR), the expression of potential E4bp4 target genes were analysed in the transduced MNK-1 cells **(Figure 5.12)**. WT E4bp4 effectively transactivated the expression of Id2 and Eomes, as previously reported (Male et al., 2014) and also Notch1 and E2A **(Figure 5.12)**. In comparison to WT E4bp4, the 5X SUMO and S286-353A mutants did not affect the transactivation of Eomes and Id2 **(Figure 5.12)**. However, both mutants caused greater transcriptional activation of both Notch1 and E2A, when compared to WT E4bp4 **(Figure 5.12)**. This provided evidence that the SUMOylation and phosphorylation of E4bp4 negatively regulate the transcriptional activity of E4bp4 in NK cells and, in the absence of these modifications, E4bp4 is a more potent transcriptional activator. Why the absence of PTMs on E4bp4 affected the transactivation of some genes and not others remains unclear and should be further investigated.



5.11 – MNK-1 cells were highly transduced and E4bp4 was stably expressed in these cells.

A) MNK-1 cells were transduced with a lentiviral expression vector, pCSGW, containing the GFP gene. Forty-eight hours after transduction cells were analysed by flow cytometry for GFP expression. B) MNK-1 cells were transduced with pCSGW constructs containing WT E4bp4, E4bp4 5X SUMO and E4bp4 S286-353A cDNA. MNK-1 cells were also transduced with an empty pCSGW construct. Forty-eight hours after transduction, equivalent numbers of cells were lysed and whole cell lysates were analysed by WB for E4bp4 expression. Probing for tubulin acted as a loading control. Molecular weight markers are shown in kDa. Grey arrow highlights unmodified E4bp4; black arrows highlight SUMO-modified forms of E4bp4.



5.12 – SUMOylation and phosphorylation affect the activation of E4bp4 target genes in NK cells.

MNK-1 cells were transduced with the lentiviral expression vector containing either the WT E4bp4, E4bp4 5X SUMO, E4bp4 S286-353A cDNA or an empty vector. Transduced cells were lysed and all cellular RNA was purified for each sample. The RNA was converted to cDNA using reverse transcriptase and the expression of various potential E4bp4 target genes was analysed using quantitative PCR with Taqman primer/probe sets. Data is representative of three independent experiments. Error bars show SEM. Statistical differences were determined between E4bp4 and 5X SUMO or S285-353A mutants using unpaired student T tests. *, $P < 0.05$.

5.5 Discussion

The bZIP TF E4bp4 has multiple PTMs; it is phosphorylated on at least three serine residues and at least three lysine residues carry SUMO modifications. Both phosphorylation and SUMOylation regulate the activity of E4bp4 and suppress its functionality. When phosphorylation or SUMOylation sites are mutated, the activity of E4bp4 during NK cell development is dramatically increased, resulting in greater numbers of NK cells being produced *in vitro*. The inhibition of CK1ε, which is thought to modify E4bp4, was also able to promote an increase in NK cell production. However, the molecular mechanism underlying the regulation by PTMs is yet to be explored. As a TF E4bp4 is able to active and repress the transcription of target genes, in different cell types and in varying contexts (Male et al., 2012). Here, the PTMs of E4bp4 were found to directly influence its ability to repress and activate target genes expression, but did not influence the stability of the protein.

As the PTMs of E4bp4 were able to influence the activity of E4bp4 during NK cell development, it was important to establish if this was a direct effect of increased protein levels being present within the cell due to the PTMs influencing the stability of E4bp4. Both phosphorylation and SUMOylation have been reported to influence ubiquitin-mediated proteasomal degradation of TFs (Filtz et al., 2014). The stability of WT E4bp4, 5X SUMO E4bp4 and S286-353A E4bp4 were compared using a cycloheximide time course assay. Although this assay is only semi-quantitative, it clearly demonstrated that the PTMs of E4bp4 do not affect its stability and both mutants and WT E4bp4 are degraded at the same rate. This finding contradicted the previous work of Doi *et al*. who had suggested that the phosphorylation of E4bp4 influenced its proteasomal degradation (Doi et al., 2004). However, their approach simply involved the overexpression of CK1ε and E4bp4 together and is unlikely to be indicative of what occurs endogenously. As the PTMs of E4bp4 did not affect the stability of the protein, they must influence its activity through another mechanism.

E4bp4 is a TF which regulates the expression of numerous target genes. Using a reporter gene assay, with E4bp4 binding sites cloned upstream of the promoter for Firefly Luciferase, E4bp4 was able to repress transcriptional activity. Mutants lacking both SUMOylation and phosphorylation sites caused greater transcriptional repression than WT E4bp4, demonstrating that both modifications regulate E4bp4 activity in the same way. Different phosphorylation mutants were analysed in this assay and all containing the S286A mutation caused greater repression than WT E4bp4. The effects of the various mutants in this assay mirrored what was previously observed in the NK cell development assays, suggesting this

is an important mechanism by which PTMs regulate E4bp4 and ultimately influence NK cell development. However, in the luciferase assay, the combination of S286A and S353A did not seem to increase the activity of E4bp4 further than the S286A mutant alone, as was the case in the NK cell development assay. This regulation of E4bp4 is very similar to what has been observed for other TFs such as FOG-2, a critical regulator of heart morphogenesis and coronary vessel formation (Perdomo et al 2012). The SUMOylation of FOG-2 negatively regulated its ability to repress transcription of target genes, as mutating FOG-2 SUMOylation sites resulted in stronger transcriptional repression activity (Perdomo et al., 2012).

It was noted that more than one minimal consensus-DNA binding sequence had been discovered for E4bp4. The binding site identified by Cowell *et al*. varied slightly from the one later identified by Zhang *et al*. but both contained the consensus $T\frac{T}{G}A\frac{T}{C}GTAA$ at the centre of the minimal consensus-DNA binding site. These different binding sequences caused E4bp4 to induce opposing effects on transcription when used in reporter gene assays (Cowell et al., 1992) (Zhang et al., 1995). Using the Zhang *et al.* sequence, it was possible to analyse the effect of PTMs on transactivation by E4bp4 (Zhang et al., 1995). Removing PTMs from E4bp4 increased its transactivation, in similar manner to what was observed with the transcriptional repression. Both the 5X SUMO and S286-353A mutants caused greater transactivation than WT E4bp4. However, the transactivation observed for WT E4bp4 in the assay was lower than previously reported by Chen *et al*., who showed that reporter gene activity was 4.8x higher in the presence of E4bp4 (Chen et al., 1995). Further optimisation of assay conditions may be required to achieve such high levels of transcriptional activation.

These findings demonstrated that E4bp4 can recognise different DNA-binding sites and can cause differing effects on the expression of target genes. However, the regulation of E4bp4 by phosphorylation and SUMOylation is the same, regardless of whether it is activating or repressing transcription. A number of other bZIP TFs can also recognise two distinct DNA-binding sites, including XBP1 and CREB3 (Jolma et al., 2013). Comparison of experimentally verified E4bp4 binding sites, upstream of *in vivo* target genes, also demonstrates that E4bp4 can bind to varying DNA sequences **(Table 5.2)**. The exact mechanism that determines if E4bp4 will activate or repress a particular target gene is not clear, but it is likely to be combination of the exact DNA binding site it identifies, the wider nucleotide sequence surrounding that binding site and the presence or absence of other interacting proteins and cell specific factors. The tumour suppressor TF, p53, has been extensively studied and has also been found to bind different DNA binding sites as well as

having different transcriptional activity on its target genes (Riley et al., 2008). In fact, p53 uses both direct and indirect methods to repress gene transcription (Riley et al., 2008).

Table 5.2 – *In vivo* E4bp4 binding sites

Gene	Cell type	E4bp4 binding site	Transcriptional activity	Reference
Id2	NK cell	GTGTGTATGTAAATGT	Activation	(Male et al., 2014)
Eomes	NK cell	AAAATAATGTAAATTT	Activation	(Male et al., 2014)
IL3	T cell	CCACAGACGTAATTAA	Activation	(Zhang et al., 1995)
Per2	Pineal gland	TCACTTATGTAAAATG	Repression	(Doi et al., 2001)
IL-12b	Macrophages	ATGATGATGTAAGAAA	Repression	(Smith et al., 2011)
IL-13	T cells	GGCCTTATGTAACGGG	Repression	(Kashiwada et al., 2011a)

Residues matching consensus E4bp4 binding sequence highlighted

As the PTMs of E4bp4 influenced its transcriptional activity in the reporter gene assay, it was postulated whether the same would occur for endogenous target genes *in vivo*. The 5X SUMO mutant, S286-353A mutant and WT E4bp4 were expressed in the newly developed MNK-1 murine NK cell line (Allan et al., 2014) using lentiviral transduction and then the expression of potential E4bp4 target genes was analysed by qPCR. The MNK-1 cells did not express detectable levels of endogenous E4bp4, had a clear NK cell phenotype and could be transduced to high levels, so was ideal for this analysis. Expression of E4bp4 in the MNK-1 cells caused the transactivation of a number of genes, but interestingly the transactivation of Notch1 and E2A was significantly higher with the 5X SUMO and S286-353A mutants compared to WT E4bp4. These observations supported the findings from the reporter gene assay and demonstrated that SUMOylation and phosphorylation are key regulators of E4bp4 activity, influencing its transcriptional activity *in vivo* and specifically in NK cells.

Notch1 and E2A have not previously been suggested as transcriptional targets for E4bp4 and are more commonly associated with T cell development, but there is evidence to suggest that they play a role in NK cell development. Carotta *et al*. demonstrated that Notch signalling can directly induce NK cell development in murine *Pax5*$^{-/-}$ pro-B cells (Carotta et

al., 2006). Specifically, Notch signalling was required at the earliest stages of NK cell development, as prolonged signalling decreased the number of NK cells produced and promoted the development of T cells (Carotta et al., 2006). A separate study also showed that the Notch ligand, Jagged 2, was able to induce NK cell development *in vitro* from murine HSCs (DeHart et al., 2005).

Similar observations have also been made for the development of human NK cells, as the activation of Notch at an early point in NK cell development can promote the appearance of NK cells (Bachanova et al., 2009) (Haraguchi et al., 2009) (Schotte et al., 2010). The interruption of Notch signalling, at an early stage of development, also resulted in the loss of human NK cell development from HSCs *in vitro* (Haraguchi et al., 2009). Human umbilical cord blood $CD34^+$ progenitors have also been developed into $CD7^+$ lymphoid precursors and functional NK cells in the absence of both stromal cells and IL-15, when sustained Notch1 signalling was present (Bachanova et al., 2009). However, for NK cells to fully mature, produce a full range of cytokines and acquire all NK cell receptors required for effector functions, they do require stromal cells as well as IL-15 and cannot simply rely on Notch1 signalling (Bachanova et al., 2009).

These findings demonstrate that Notch is required at the early stages of NK cell development, which is strikingly similar to what is observed for E4bp4 which is indispensable for the early stages of NK cell development but it is not required for NK cell maturation and functionality (Firth et al., 2013). Boos *et al*. also demonstrated that Id2, which is required to suppress E-proteins in developing NK cells, does not target E2A in developing NK cells, suggesting that this too could be involved in this process (Boos et al., 2007). E2A expression is essential for the production of LTi cells (Boos et al., 2007), a close relative of NK cells and E4bp4 is also required for their development (Geiger et al., 2014). Finally, promoter regions of both *Notch1* and E2A (*Tcf3*) genes in the mouse genome, were discovered to contain predicted E4bp4 binding sites, in particular five sites were found in different promoter regions of the *Notch1* gene **(Table 5.3)**.

E4bp4 was found to transactivate a number of genes in the MNK-1 cells, but Eomes and Id2 transactivation was not influenced by the lack of SUMOylation or phosphorylation. Therefore, the regulation of E4bp4 transcriptional activity by its PTMs may be more complex, with only the transcription of a subset of E4bp4 target genes being regulated by its PTMs. The MNK-1 experiment did not reveal any transcriptional repression by E4bp4 *in vivo*, however only a small number of potential target genes were tested. As both phosphorylation and SUMOylation had a significant effect on transcriptional repression in the Luciferase reporter

assay, it would be interesting to determine if this also occurred *in vivo*. Analysing the effects of the E4bp4 PTM mutants during the development of *ex vivo* NK cells would also be of interest. However, to analyse changes in developing cells would require large numbers of FACS-sorted CLP cells to be transduced in order to see clear differences in gene activation between E4bp4 mutants, making this experiment technically and financially impractical. An alternative approach could involve FACS-sorting NK cells that develop from the transduced *ex vivo* Lin^- BM cells and analysing global changes in gene expression by microarray. Gene expression patterns could be compared between cells transduced with the WT form of E4bp4 or those transduced with a mutant form. However, E4bp4 functions at the earliest stages of NK cell development, so analysing gene expression in fully mature NK cells may not reveal important changes caused by the presence/absence of PTMs on E4bp4 that lead to increased levels of NK cell production.

Table 5.3 – Predicted E4bp4 binding sites in the promoter regions of *Notch1* and E2A (*Tcf3*)

Gene	Promoter	Position in genome	Predicted E4bp4 binding site
Notch1	GXP_164006	Chromosome 2: 26477525 - 26477545	CTACGGTCATGTAACTCATC
Notch1	GXP_164006	26477509 - 26477529	TCAGGCTGAGGTAAAGCTGG
Notch1	GXP_164007	26458924 – 26458944	TCACACTTCTGTAACCACCA
Notch1	GXP_164010	26468217 – 26468237	GTGCAAGTAAGTAAGGCTGA
Notch1	GXP_3078977	26466245 – 26466265	TAAAAAATATGTAATGTGTAT
Notch1	GXP_3078978	26466821 – 26466841	ATTCTTTGAAGTAAGGAAAAG
Tcf3	GXP_3062655	Chromosome 10: 80416768 - 80416788	CAGCCCTGAGGTAACACATGG

Residues matching consensus E4bp4 binding sequence highlighted

As well as E4bp4, other bZIP TFs are known to be able to transactivate and repress different target genes. The bZIP transcription factor, MafA, which is responsible for β-islet cell maturation in the pancreas, has also been identified as a transcriptional activator and repressor (Hang and Stein, 2011) (Pouponnot et al., 2006). MafA activity is tightly controlled by phosphorylation and changes to specific phosphorylation sites are associated with

pathologies including Type 1 Diabetes Mellitus (Hang and Stein, 2011). MafA is also SUMOylated and, like E4bp4, SUMOylation negatively regulates MafA transcriptional activity and ultimately β-islet cell function (Shao and Cobb, 2009). The TF Ets1, which, like E4bp4, is essential for NK cell development, is also negatively regulated by both SUMOylation and phosphorylation. Ets1 has two SUMOylation sites at K15 and K227 and modification of these sites reduced transactivation of target genes by Ets1, but did not affect its stability (Ji et al., 2007). Additionally, Ets1 can be phosphorylated by calmodulin-dependent protein kinase II near its DNA binding domain and this influences it ability to bind DNA and activate target gene expression (Pufall et al., 2005).

Both the SUMOylation and phosphorylation mutants displayed the same phenotypes throughout this study. Mutants lacking both PTMs were able to promote a greater level of transcription repression and transcriptional activation. Therefore, it is possible that these two PTMs work together to regulate the transcriptional activity of E4bp4. A number of studies have reported that phosphorylation can influence SUMOylation, through a phosphorylation-dependent consensus motif ΨKxExxS/T, so the phosphorylated serine or threonine is in close proximity to the SUMOylated lysine (Hietakangas et al., 2006) (Shalizi et al., 2006) (Su et al., 2012). The TF, MEF2A, which is required for myogenesis and neuronal morphogenesis, is regulated by phosphorylation dependent SUMOylation at one of these consensus sites (Shalizi et al., 2006). Dephosphorylation of Ser408, which is catalysed by neuronal activity and calcium-dependent phosphatases, promotes the desumoylation of Lys403, resulting in the activation of MEF2A and the subsequent expression of target genes (Shalizi et al., 2006). However, a similar process is unlikely to occur with E4bp4 because none of the phosphorylation and SUMOylation sites identified in this study are in close proximity and E4bp4 does not contain any ΨKxExxS/T consensus motifs for phosphorylation-dependent SUMOylation. The E4bp4 S286-353A mutant also appeared to have the same SUMO modifications as WT E4bp4, suggesting that phosphorylation does not influence the SUMOylation of E4bp4. However, analysing whether the phosphorylation and SUMOylation of E4bp4 influence each other should be part of future studies into the regulation of E4bp4.

Phosphorylation and SUMOylation are both potent regulators of E4bp4's activity during NK cell development and both PTMs directly influence its transcriptional activity, without affecting its stability. The activity of E4bp4 during NK cell development is therefore likely to be controlled by PTMs, allowing it to activate and repress different target genes to different extents. In the absence of phosphorylation or SUMOylation E4bp4 can be more active as a TF, causing greater levels of transcriptional activation or repression, both of which will help

to drive an NK cell specific genetic programme. In particular it was observed that in the absence of PTMs E4bp4 promoted transcriptional activation of Notch1 and E2A in NK cells, which may help to influence NK cell development. The full relationship between E4bp4 PTMs remains to be explored as it could have alternative modification states *in vivo* that may indeed allow it to act as a transcriptional repressor and activator in different contexts.

CHAPTER 6

Conclusions and future work

6.1 Conclusions

NK cells are innate lymphocytes that are capable of producing inflammatory cytokines and can spontaneously kill infected, stressed or malignant cells. NK cells develop in the BM from HSCs and lymphoid progenitor cells by a tightly regulated process, requiring the expression of numerous TFs as well as the presence of cytokines and other signal molecules (Vosshenrich and Di Santo, 2013). Of the TFs identified as essential for NK cell development so far, E4bp4 is the only one required for commitment to the NK lineage and must be expressed in CLPs in order for the earliest pre-NKPs to develop (Male et al., 2014). During this process, E4bp4 regulates the expression of other important TFs that control NK cell development, such as Eomes and Id2 (Male et al., 2014). This is a highly specialised function as E4bp4 activity is expendable for the maintenance, homeostasis and anti-viral responses of mature NK cells (Firth et al., 2013). E4bp4 is also essential for other immunological processes including the cytokine production by T_H1 and T_H2 T helper cells (Kashiwada et al., 2011a) (Kobayashi et al., 2014) (Motomura et al., 2011), class switching to IgE in B cells (Kashiwada et al., 2010) and the development of T_H17 T helper cells (Yu et al., 2013b).

The research presented here demonstrates that the E4bp4 protein is multiply SUMOylated and multiply phosphorylated *in vivo* and both function in a similar manner to negatively regulate its transcriptional activity. This influence on transcriptional activity affected the function of E4bp4 during NK cell development. The expression of E4bp4 lacking either SUMOylation or phosphorylation sites in *E4bp4*$^{-/-}$ Lin$^-$ BM cells caused a dramatic increase in NK cell production, compared with cells expressing WT E4bp4. An almost identical phenotype was observed when the mutant versions of E4bp4 were expressed in WT Lin$^-$ BM cells. This negative regulation of E4bp4 by phosphorylation and SUMOylation is similar to what has been reported for a number of other TFs, for example transcriptional repression of target genes by FOG-2 is negatively regulated by SUMOylation (Perdomo et al., 2012). The activity of Ets1, which is important for NK cell development, is also regulated by both phosphorylation and SUMOylation (Ji et al., 2007) (Tootle and Rebay, 2005). However, until this study very few have demonstrated that the SUMOylation of a single TF can have such a significant impact on a complex, dynamic biological process. Additionally, mutation of a single phosphorylation site, S286, was found to have a profound effect on E4bp4 activity and ultimately NK cell development.

6.2 Signals that influence the PTMs and activity of E4bp4

The signals required to induce E4bp4 expression and activity in NK cells, as well as the signals that influence its PTMs, are largely unknown. IL-15, the critical cytokine for NK cell production, was originally proposed to function 'upstream' of E4bp4 in NK cell development as IL-15 signalling promoted its expression (Gascoyne et al., 2009). Ectopic expression of E4bp4 also promoted NK cell development even in the absence of IL-15 signalling (Gascoyne et al., 2009). However, the relationship between E4bp4 and IL-15 now appears more complex; in the absence of E4bp4, IL-15 responsive NKPs expressing the IL-2Rβ receptor (CD122) do not develop (Male et al., 2014). This suggests that E4bp4 may also act upstream of IL-15 signalling and may function in some form of self-reinforcing positive feed-back loop that drives cells to commit to the NK lineage (Male et al., 2014). IL-15 signalling involves the phosphorylation of JAK3, which in turn phosphorylate STATs, such as STAT5, causing its dimerisation and translocation to the nucleus (Orr et al., 2010). It is possible that this signalling influences the PTMs of E4bp4 as well as its transcription.

The evolutionarily conserved serine-threonine kinase, mTOR ('mammalian target of rapamycin'), is a central component in many metabolic and inflammatory pathways and recent work by Marcais *et al*. has revealed that its activation is important for proliferation of NK cells in the BM (Marcais et al., 2014). mTOR activation was found to be an important component of IL-15 signalling and therefore it is possible that it too could influence the PTMs and activity of E4bp4. Identifying whether or not these signalling pathways influence the modifications of E4bp4 will be challenging, as modified forms of endogenous E4bp4 are very low in abundance. However, analysis could be made possible through the development of specific antibodies that recognise the phosphorylated/SUMOylated isoforms of E4bp4.

E4bp4 is a circadian protein; its expression fluctuates in an overt diurnal manner and is opposite to that of numerous PAR proteins, which form part of the mammalian circadian clock (Mitsui et al., 2001) (Ueda et al., 2005). The differentiation of T_H17 cells is controlled by E4bp4 and the circadian clock, with the expression of RORγt being repressed in a diurnal manner by E4bp4 (Yu et al., 2013b). In the absence of the primary circadian activator CLOCK, naïve $CD4^+$ T cells had higher expression of *E4bp4* and a reduced capacity for T_H17 cell differentiation (Yu et al., 2013b). Whether E4bp4 expression, its activity or its PTMs are controlled by the circadian clock in developing NK cells has yet to be determined, however this could be an important mechanism by which NK cell production and immune

responses are coordinated. To establish if this regulatory mechanism exists, NK cell populations and their E4bp4 expression should be analysed in *CLOCK*$^{-/-}$ mice and in other mice that lack key circadian rhythm genes.

6.3 E4bp4 and human NK cell development

The regulation of E4bp4 by its PTMs has significant implications for the production of human NK cells for use in various forms of immunotherapy. NK cells spontaneous cytotoxicity against a wide variety of tumour cells means there is intense interest in using NK cells in various forms of anti-cancer treatment (Knorr et al., 2014). NK cells exist at fairly low numbers in the blood so, after isolation from volunteers, these cells must be expanded *in vitro* to produce enough cells to give to each patient (Rujkijyanont et al., 2013). However, this process can lead to NK cell exhaustion meaning they become less effective at killing target cells and do not survive more than a few days in a recipient (Knorr et al., 2014). Alternative strategies have now been developed with the aim of generating large numbers of cytotoxic NK cells from embryonic stem cells (Woll et al., 2005), induced-pluripotent stem cells (Kaufman, 2009) (Knorr et al., 2013) (Ni et al., 2011) and umbilical cord blood (UCB) (Luevano et al., 2014) (Pinho et al., 2012) (Spanholtz et al., 2010). The production of NK cells from these different sources involves the use of cytokines and 'feeder layer' stromal cells to commit stem cells to the NK cell lineage (Luevano et al., 2012). Influencing E4bp4 activity, potentially via its PTMs, in developing human NK cells could be an important step towards increasing NK cell production.

Both previous research (Gascoyne et al., 2009) (Male et al., 2014) and this study have shown that the overexpression of E4bp4 in developing murine NK cells leads to an increase in NK cell production, demonstrating that E4bp4 is a limiting factor for this process. E4bp4 is expressed in human NK cells developed from UCB and its expression significantly increases as the NK cells mature (Luevano et al., 2014). It is highly likely that E4bp4 is important for this process, but this has yet to be studied in detail. However, a recent report by Ning *et al*. has begun to suggest the potential impact that E4bp4 can have on developing human NK cells (Ning et al., 2014). NK cells were developed from adipose tissue-derived stem cells and one strategy used to do this involved the lentiviral transduction of the stem cells with a virus expressing E4bp4 (Ning et al., 2014). These E4bp4 expressing NK cells were more cytotoxic than NK cells produced with just cytokines and 'feeder layer' stromal cells and they significantly delayed tumour progression when injected into nude mice with induced prostate

or breast cancer (Ning et al., 2014). The manipulation of E4bp4 expression levels is likely to have an important role in producing future NK cell immunotherapeutic products. Influencing the PTMs of E4bp4, for example through the use of a specific kinase inhibitors, could be the simplest and most effective way to help boost the development of large numbers of highly cytotoxic NK cells.

To determine if the activity of E4bp4 and its PTMs are important for the production of human NK cells, the Brady laboratory has established a human NK cell development assay, based on the work of the Saudemont laboratory (Anthony Nolan Institute, London) (Luevano et al., 2014), to allow genes of interest to be ectopically expressed in developing human NK cells. Using an *in vitro* culture system NK cells can be produced from UBC $CD34^+$ progenitor cells and a lentiviral expression vector can be used to stably and efficiently transduce them with a cDNA of interest **(Figure 6.1)**. The pCSGW lentiviral vector has been adapted to drive expression of a gene of interest linked to an IRES sequence upstream of a cDNA for the tandem Tomato Red fluorescent protein. Following transduction of $CD34^+$ cells and subsequent culture, more than 70% of cells were stably transduced and more than 20% were $CD56^+$ mature human NK cells **(Figure 6.1)**. Future studies should involve the transduction of these cells with genes of interest, in particular E4bp4 and E4bp4 mutants lacking PTM sites, to determine what affect these have on human NK cell development.

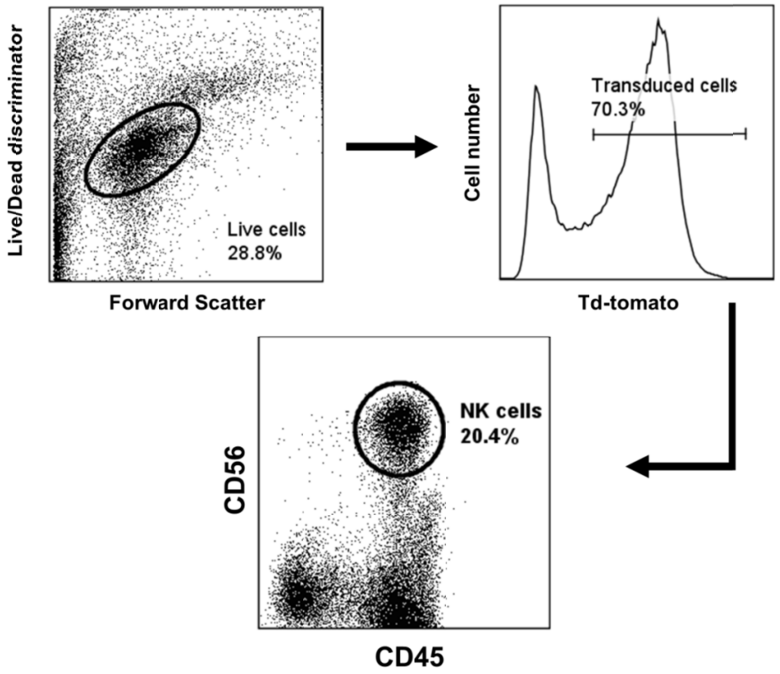


Figure 6.1 – Development of NK cells from transduced umbilical cord blood CD34$^+$ progenitor cells

Purified CD34$^+$ cells were transduced with a lentiviral vector expressing the reporter fluorescent protein tandem Tomato Red (Td-tomato). Transduced cell were FACS sorted in bulk and grown on irradiated EL08-1D2 stromal cells with recombinant cytokines (IL-7, SCF, IL-3, Flt3 ligand and IL-15) for 14 days, before flow cytometry analysis to determine the presence of transduced cells that had successfully developed into NK cells. Performed in collaboration with Dr Ilaria Nisoli.

As mutating a very small number of E4bp4 PTM sites had a large impact on its function, it would be of interest to determine if single nucleotide polymorphisms (SNPs) exist at these sites in the human genome. These mutations may increase the function of E4bp4 for a particular individual and could impact on their NK cell populations and wider immune system. To identify E4bp4 SNPs and analyse their impact, a collaboration has been established with Karl-Johan Malmberg's laboratory at the University of Oslo and the Karolinska Institute in Stockholm, who have access to large numbers of samples from volunteer blood donors. Blood samples can be analysed to determine the constitution of each individual's immune system, including the quantity and phenotype of their peripheral blood NK cells. Future studies will involve the isolation of genomic DNA from each blood sample and sequencing the E4bp4 gene and surrounding region. If mutations are discovered, then the impact on the individual's immune system will be analysed.

This analysis may also reveal other dysfunctional mutations that are carried in their *E4bp4* alleles. *E4bp4*$^{-/-}$ mice have almost no detectable cNK cells, but strikingly the *E4bp4*$^{+/-}$ mice have almost exactly half the number of NK cells of a WT mouse (Gascoyne et al., 2009). As the heterozygote mice display haploinsufficiency for NK cell production, it would be interesting if SNP analysis could identify *E4bp4* heterozygote humans, who have a dysfunctional mutation in one of their *E4bp4* alleles. Analysis of the NK cell populations and wider immune system of these individuals may demonstrate they have a similar phenotype to the heterozygote mice. An initial cohort of 38 blood samples has been acquired and analysis of the peripheral blood NK cell numbers showed the large variability in the number of NK cells in each person **(Figure 6.2)**. The percentage of NK cells of all the lymphocytes in the blood varied greaterly from 3 – 20% **(Figure 6.2)**. Age and acute or chronic viral infection are known to influence the numbers of peripheral blood NK cells populations *in vivo* (Zhang et al., 2007). However, *E4bp4* expression and the presence of mutations in *E4bp4* alleles may also contribute to this high level of variablity in human NK cell populations, which could have implications for human health.

The 1000 Genomes Project (Genomes Project et al., 2010) and the NHLBI Exome Sequencing Project (Tennessen et al 2012) have provided unprecedented resources for studying the genetic variation of human genes and the data from these projects contains information on the variability of the E4bp4 gene. To date, 43 missense variants, two frame shift variants and three stop gain variants have been identified in the E4bp4 ORF. The data showed that 13 missense variants may damage the structure and/or function of the E4bp4 protein, meaning an individual with any of these mutations could effectively be *E4bp4* heterozygous. Most interestingly a SNP was identified that caused a mutation is serine 286, changing it to a phenylalanine. The results presented in Chapter 4 demonstrate that this residue is phosphorylated and is important for NK cell development. If individuals carrying this polymorphism could be identified, then analysis of their NK cell populations and wider immune system would be of great interest. However, the NHLBI Exome Sequencing database states that this polymorphism was only found in one individual of the 6500 analysed (Exome Variant Server, NHLBI GO Exome Sequencing Project, Seattle, WA (URL: http://evs.gs.washington.edu/EVS/) [accessed 26/08/2104]).

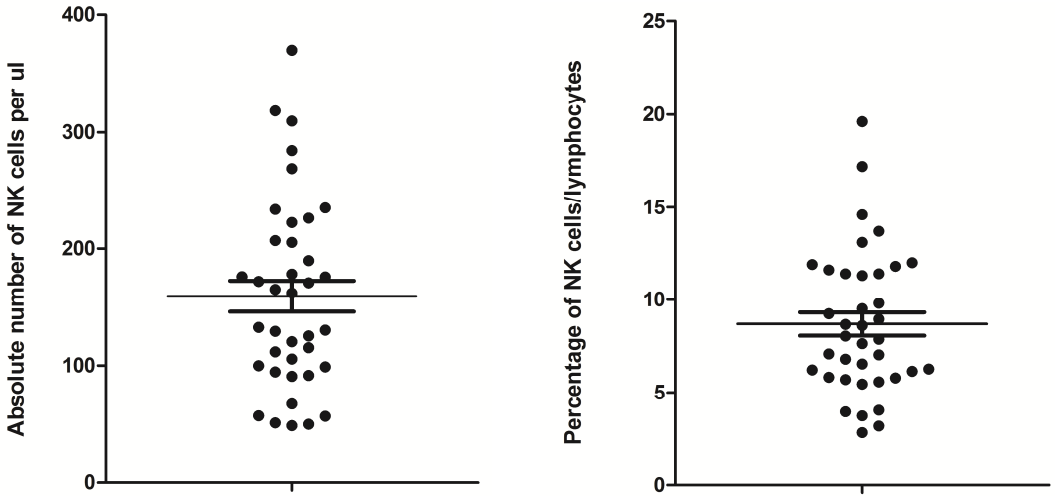


Figure 6.2 – Variations in the number of NK cells in human peripheral blood

Blood donated by 38 healthy volunteers was analysed by flow cytometry for the presence of human NK cells ($CD3^-$ $CD19^-$ $CD16^+$ $CD56^{bright/dim}$). The absolute number of NK cells was quantified and compared to the overall lymphocyte population. Data produced by Lisa Liu and Karl-Johan Malmberg, University of Oslo and Karolinska Institute, Stockholm.

6.4 E4bp4 SUMOylation and infection

Some intracellular pathogens influence SUMOylation of host cells during infection (Fritah et al., 2014) (Ribet et al., 2010). *Listeria monocytogenes* produces the virulence factor, Listeriolysin O (LLO), which attacks the Ubc9 enzyme central to the SUMOylation pathway in all cells (Ribet et al., 2010). By targeting this enzyme *L. monocytogenes* can use the LLO toxin to simultaneously modify the activity of many proteins within a host cell, which may help to dampen host immune responses (Ribet et al., 2010). NK cells and other innate lymphoid cells are important for controlling *L. monocytogenes* infections (Naper et al., 2011), however these cells can become sites of infection and may also serve as reservoirs for *L. monocytogenes* growth (McElroy et al., 2009). As E4bp4 is a prominent molecule for the function of NK cells, it is possible that it is targeted by *L. monocytogenes* and LLO as part of the mechanism of infection.

In an attempt to establish whether LLO affected E4bp4, 6His-SUMO2 HeLa cells were transfected and exposed to purified LLO **(Figure 6.3)**. LLO clearly affected the SUMOylation of E4bp4 in the 6His-SUMO2 cells **(Figure 6.3)**. The global SUMO2 level within these cells was also affected, as previously reported (Ribet et al., 2010) **(Figure 6.3)**. Recent data from

a SUMO proteomic screen also identified that modification of E4bp4 by both SUMO1 and SUMO2 was significantly reduced, following treatment with LLO (Impens et al., 2014). This study also demonstrated that not all host factors are sensitive to LLO-induced loss of SUMOylation, but many nuclear factors, such as E4bp4, are targeted by LLO (Impens et al., 2014). LLO-induced modification of the SUMOylation state of TFs will ultimately lead to changes in the transcriptome of the cell in response to the toxin, but exactly how this promotes infection is yet to be determined. The importance of changes to E4bp4 SUMOylation during an infection process is unknown and should also be further studied. Targeting the SUMOylation of E4bp4 may influence the growth or survival of infected NK cells, providing a more robust reservoir for the *L. monocytogenes* to develop. Alternatively, it could affect the NK cells ability to recognise and respond to the invading bacteria.

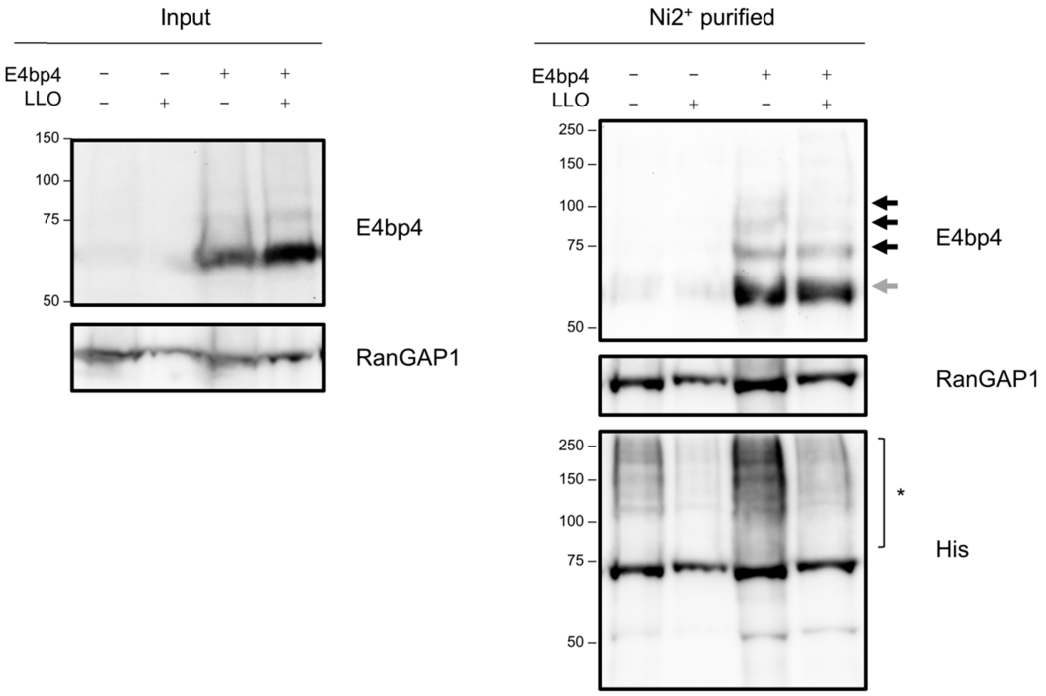


Figure 6.3 – Listeriolysin O affects the SUMOylation of E4bp4

HeLa cells stably expressing 6His-SUMO2 were transfected with pCMV-E4bp4 vector or corresponding empty vector. Cells were incubated in the presence or absence of 3 nM purified LLO for 20 minutes at RT. Subsequently, all cells were lysed and SUMOylated protein was purified by Ni^{2+} affinity chromatograph from all cell lysates. Equivalent amounts of purified material from each sample were analysed by WB, which was probed for the presence of E4bp4, His (SUMO2/3) and RanGAP1 as a loading control. Grey arrow highlights unmodified E4bp4, black around highlights modified forms of E4bp4 and * highlights all SUMOylated protein within the sample. Molecular weight markers are indicated on the left of each blot in kDa.

6.5 Interactions between E4bp4 PTMs

Both SUMOylation and phosphorylation function to repress the transcriptional activity of E4bp4 as the protein is more active once these modifications are removed. E4bp4 is crucial for many biological processes; therefore, it is important that its function is tightly regulated by multiple mechanisms because aberrant activity could have damaging consequences. As the phenotypes observed for the phosphorylation and SUMOylation site mutants were so similar this suggests that the two PTMs may work together to regulate the activity of E4bp4.

Some studies suggest that some PTMs occur progressively; SUMOylation can affect acetylation (Wu and Chiang, 2009), SUMOylation can lead to ubiquitination (Guo et al., 2012) and phosphorylation may increase (or decrease) levels of SUMOylation (Mohideen et al., 2009) (Hietakangas et al., 2006). Other reports have demonstrated how PTMs function reciprocally to promote specific functions of a TF. This was observed by Zhang *et al*. who found that phosphorylation and SUMOylation act to reduce the activities of Bcl11b and prevent it from binding to target DNA sites (Zhang et al., 2012). The interplay between E4bp4's PTMs is yet to be explored in detail, but the S286-353A mutant expressed in the MNK-1 cells was observed to have the same SUMO modifications as WT E4bp4, suggesting that phosphorylation does not influence SUMOylation. MS experiments involving various E4bp4 PTM mutants will reveal if the absence of a particular modification impacts on the presence of another. Additionally, *in vitro* sequential modification assays can also be used to help reveal if some modifications are dependent on others.

Further MS analysis will also establish if E4bp4 has other PTMs such as acetylation, methylation or glycosylation, which may also be important for regulating its activity. Knowledge of the three-dimensional structure of E4bp4 would supplement this study and is an essential future step towards fully understanding the mechanism underlying its post translational regulation.

6.6 Summary

This study has demonstrated that SUMOylation and phosphorylation both control the transcriptional activity of E4bp4 and its function during NK cell development. In the absence of these modifications the transcriptional activation and repression of target genes by E4bp4 was significantly increased and this heightened functional activity caused greater levels of NK cell production, compared to WT E4bp4. Determining what signals influence these modifications *in vivo*, should now be ascertained as this may provide a simple mechanism by which the activity of E4bp4 can be modulated. The findings of this study have immediate application in the production of human NK cells for use in various forms of immunotherapy. Modulation of E4bp4 activity in developing human NK cells may prove to be a key tool in producing large numbers of highly cytotoxic cells to treat patients with various malignancies and viral infections.

References

Acharya, A., Rishi, V., Moll, J., and Vinson, C. (2006). Experimental identification of homodimerizing B-ZIP families in Homo sapiens. J Struct Biol *155*, 130-139.

Aliahmad, P., de la Torre, B., and Kaye, J. (2010). Shared dependence on the DNA-binding factor TOX for the development of lymphoid tissue-inducer cell and NK cell lineages. Nat Immunol *11*, 945-952.

Aliahmad, P., Seksenyan, A., and Kaye, J. (2012). The many roles of TOX in the immune system. Curr Opin Immunol *24*, 173-177.

Allan, D.S., Kirkham, C.L., Aguilar, O.A., Qu, L.C., Chen, P., Fine, J.H., Serra, P., Awong, G., Gommerman, J.L., Zuniga-Pflucker, J.C.*, et al.* (2014). An in vitro model of innate lymphoid cell function and differentiation. Mucosal Immunol.

Alm-Kristiansen, A.H., Norman, I.L., Matre, V., and Gabrielsen, O.S. (2009). SUMO modification regulates the transcriptional activity of FLASH. Biochem Biophys Res Commun *387*, 494-499.

Anckar, J., and Sistonen, L. (2007). SUMO: getting it on. Biochemical Society Transactions *35*, 1409-1413.

Asher, G., Gatfield, D., Stratmann, M., Reinke, H., Dibner, C., Kreppel, F., Mostoslavsky, R., Alt, F.W., and Schibler, U. (2008). SIRT1 regulates circadian clock gene expression through PER2 deacetylation. Cell *134*, 317-328.

Baba, D., Maita, N., Jee, J.G., Uchimura, Y., Saitoh, H., Sugasawa, K., Hanaoka, F., Tochio, H., Hiroaki, H., and Shirakawa, M. (2005). Crystal structure of thymine DNA glycosylase conjugated to SUMO-1. Nature *435*, 979-982.

Bachanova, V., McCullar, V., Lenvik, T., Wangen, R., Peterson, K.A., Ankarlo, D.E., Panoskaltsis-Mortari, A., Wagner, J.E., and Miller, J.S. (2009). Activated notch supports development of cytokine producing NK cells which are hyporesponsive and fail to acquire NK cell effector functions. Biol Blood Marrow Transplant *15*, 183-194.

Ballas, N., Grunseich, C., Lu, D.D., Speh, J.C., and Mandel, G. (2005). REST and its corepressors mediate plasticity of neuronal gene chromatin throughout neurogenesis. Cell *121*, 645-657.

Beach, J.A., Nary, L.J., Hirakawa, Y., Holland, E., Hovanessian, R., and Medh, R.D. (2011). E4BP4 facilitates glucocorticoid-evoked apoptosis of human leukemic CEM cells via upregulation of Bim. J Mol Signal *6*, 13.

Becker, J., Barysch, S.V., Karaca, S., Dittner, C., Hsiao, H.H., Berriel Diaz, M., Herzig, S., Urlaub, H., and Melchior, F. (2013). Detecting endogenous SUMO targets in mammalian cells and tissues. Nat Struct Mol Biol *20*, 525-531.

Bekes, M., Prudden, J., Srikumar, T., Raught, B., Boddy, M.N., and Salvesen, G.S. (2011). The dynamics and mechanism of SUMO chain deconjugation by SUMO-specific proteases. J Biol Chem *286*, 10238-10247.

Benayoun, B.A., and Veitia, R.A. (2009). A post-translational modification code for transcription factors: sorting through a sea of signals. Trends Cell Biol *19*, 189-197.

Bernink, J.H., Peters, C.P., Munneke, M., te Velde, A.A., Meijer, S.L., Weijer, K., Hreggvidsdottir, H.S., Heinsbroek, S.E., Legrand, N., Buskens, C.J.*, et al.* (2013). Human type 1 innate lymphoid cells accumulate in inflamed mucosal tissues. Nat Immunol *14*, 221-229.

Bertolotto, C., Lesueur, F., Giuliano, S., Strub, T., de Lichy, M., Bille, K., Dessen, P., d'Hayer, B., Mohamdi, H., Remenieras, A.*, et al.* (2011). A SUMOylation-defective MITF germline mutation predisposes to melanoma and renal carcinoma. Nature *480*, 94-98.

Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S., and Brunak, S. (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics *4*, 1633-1649.

Blomster, H.A., Hietakangas, V., Wu, J., Kouvonen, P., Hautaniemi, S., and Sistonen, L. (2009). Novel Proteomics Strategy Brings Insight into the Prevalence of SUMO-2 Target Sites. Molecular & Cellular Proteomics *8*, 1382-1390.

Boos, M.D., Yokota, Y., Eberl, G., and Kee, B.L. (2007). Mature natural killer cell and lymphoid tissue-inducing cell development requires Id2-mediated suppression of E protein activity. J Exp Med *204*, 1119-1130.

Brivanlou, A.H., and Darnell, J.E., Jr. (2002). Signal transduction and the control of gene expression. Science *295*, 813-818.

Carey, K.T., Tan, K.H., Ng, J., Liddicoat, D.R., Godfrey, D.I., and Cole, T.J. (2013). Nfil3 is a glucocorticoid-regulated gene required for glucocorticoid-induced apoptosis in male murine T cells. Endocrinology *154*, 1540-1552.

Carotta, S., Brady, J., Wu, L., and Nutt, S.L. (2006). Transient Notch signaling induces NK cell potential in Pax5-deficient pro-B cells. Eur J Immunol *36*, 3294-3304.

Carotta, S., Dakic, A., D'Amico, A., Pang, S.H., Greig, K.T., Nutt, S.L., and Wu, L. (2010a). The transcription factor PU.1 controls dendritic cell development and Flt3 cytokine receptor expression in a dose-dependent manner. Immunity *32*, 628-641.

Carotta, S., Pang, S.H.M., Nutt, S.L., and Belz, G.T. (2011). Identification of the earliest NK-cell precursor in the mouse BM. Blood *117*, 5449-5452.

Carotta, S., Wu, L., and Nutt, S.L. (2010b). Surprising new roles for PU.1 in the adaptive immune response. Immunol Rev *238*, 63-75.

Chen, F., Zhang, L., OuYang, Y., Guan, H., Liu, Q., and Ni, B. (2014). Glucocorticoid induced osteoblast apoptosis by increasing E4BP4 expression via up-regulation of Bim. Calcif Tissue Int *94*, 640-647.

Chen, H., and Qi, L. (2010). SUMO modification regulates the transcriptional activity of XBP1. Biochem J *429*, 95-102.

Chen, R.Q., Yang, Q.K., Lu, B.W., Yi, W., Cantin, G., Chen, Y.L., Fearns, C., Yates, J.R., 3rd, and Lee, J.D. (2009). CDC25B mediates rapamycin-induced oncogenic responses in cancer cells. Cancer Res *69*, 2663-2668.

Chen, W.-J., Lewis, K.S., Chandra, G., Cogswell, J.P., Stinnett, S.W., Kadwell, S.H., and Gray, J.G. (1995). Characterization of human E4BP4, a phosphorylated bZIP factor. Biochim Biophys Acta *1264*, 388-396.

Chen, Y.Z., Chen, Z., Gong, Y.A., and Ying, G. (2012). SUMOhydro: a novel method for the prediction of sumoylation sites based on hydrophobic properties. PLoS ONE *7*, e39195.

Chen, Z., Lund, R., Aittokallio, T., Kosonen, M., Nevalainen, O., and Lahesmaa, R. (2003). Identification of novel IL-4/Stat6-regulated genes in T lymphocytes. J Immunol *171*, 3627-3635.

Cheng, H.Y.M., Pitcher, G.M., Laviolette, S.R., Whishaw, I.Q., Tong, K.I., Kockeritz, L.K., Wada, T., Joza, N.A., Crackower, M., Goncalves, J.*, et al.* (2002). DREAM is a critical transcriptional repressor for pain modulation. Cell *108*, 31-43.

Chicooree, N., Griffiths, J.R., Connolly, Y., Tan, C.T., Malliri, A., Eyers, C.E., and Smith, D.L. (2013). A novel approach to the analysis of SUMOylation with the independent use of trypsin and elastase digestion followed by database searching utilising consecutive residue addition to lysine. Rapid Commun Mass Spectrom *27*, 127-134.

Choudhary, C., Olsen, J.V., Brandts, C., Cox, J., Reddy, P.N., Bohmer, F.D., Gerke, V., Schmidt-Arras, D.E., Berdel, W.E., Muller-Tidow, C.*, et al.* (2009). Mislocalized activation of oncogenic RTKs switches downstream signaling outcomes. Mol Cell *36*, 326-339.

Chupreta, S., Holmstrom, S., Subramanian, L., and Iniguez-Lluhi, J.A. (2005). A small conserved surface in SUMO is the critical structural determinant of its transcriptional inhibitory properties. Mol Cell Biol *25*, 4272-4282.

Cichocki, F., Miller, J.S., Anderson, S.K., and Bryceson, Y.T. (2013). Epigenetic regulation of NK cell differentiation and effector functions. Front Immunol *4*, 55.

Cohen, P. (2014). Immune diseases caused by mutations in kinases and components of the ubiquitin system. Nat Immunol *15*, 521-529.

Colucci, F. (2001). Differential requirement for the transcription factor PU.1 in the generation of natural killer cells versus B and T cells. Blood *97*, 2625-2632.

Colucci, F., Di Santo, J.P., and Leibson, P.J. (2002). Natural killer cell activation in mice and men: different triggers for similar weapons? Nat Immunol *3*, 807-813.

Constantinides, M.G., McDonald, B.D., Verhoef, P.A., and Bendelac, A. (2014). A committed precursor to innate lymphoid cells. Nature *508*, 397-401.

Cooper, M.A., Fehniger, T.A., and Caligiuri, M.A. (2001). The biology of human natural killer-cell subsets. Trends Immunol *22*, 633-640.

Cowell, I.G., and Hurst, H.C. (1994). Transcriptional repression by the human bZIP factor E4BP4: definition of a minimal repression domain. Nucleic Acids Res *22*, 59-65.

Cowell, I.G., and Hurst, H.C. (1996). Protein-protein interaction between the transcriptional repressor E4BP4 and the TBP-binding protein Dr1. Nucleic Acids Res *24*, 3607-3613.

Cowell, I.G., Skinner, A., and Hurst, H.C. (1992). Transcriptional repression by a novel member of the bZIP family of transcription factors. Mol Cell Biol *12*, 3070-3077.

Crispe, I.N. (2009). The liver as a lymphoid organ. Annu Rev Immunol *27*, 147-163.

Crotta, S., Gkioka, A., Male, V., Duarte, J.H., Davidson, S., Nisoli, I., Brady, H.J., and Wack, A. (2014). The Transcription Factor E4BP4 Is Not Required for Extramedullary Pathways of NK Cell Development. J Immunol *192*, 2677-2688.

Cubenas-Potts, C., and Matunis, M.J. (2013). SUMO: a multifaceted modifier of chromatin structure and function. Dev Cell *24*, 1-12.

Curti, A., Ruggeri, L., D'Addio, A., Bontadini, A., Dan, E., Motta, M.R., Trabanelli, S., Giudice, V., Urbani, E., Martinelli, G.*, et al.* (2011). Successful transfer of alloreactive haploidentical KIR ligand-mismatched natural killer cells after infusion in elderly high risk acute myeloid leukemia patients. Blood *118*, 3273-3279.

Curtin, J.F., and Cotter, T.G. (2004). JNK regulates HIPK3 expression and promotes resistance to Fas-mediated apoptosis in DU 145 prostate carcinoma cells. J Biol Chem *279*, 17090-17100.

Da Silva-Ferrada, E., Lopitz-Otsoa, F., Lang, V., Rodriguez, M.S., and Matthiesen, R. (2012). Strategies to Identify Recognition Signals and Targets of SUMOylation. Biochem Res Int *2012*, 875148.

Daussy, C., Faure, F., Mayol, K., Viel, S., Gasteiger, G., Charrier, E., Bienvenu, J., Henry, T., Debien, E., Hasan, U.A.*, et al.* (2014). T-bet and Eomes instruct the development of two distinct natural killer cell lineages in the liver and in the bone marrow. J Exp Med *211*, 563-577.

DeHart, S.L., Heikens, M.J., and Tsai, S. (2005). Jagged2 promotes the development of natural killer cells and the establishment of functional natural killer cell lines. Blood *105*, 3521-3527.

Dephoure, N., Zhou, C., Villen, J., Beausoleil, S.A., Bakalarski, C.E., Elledge, S.J., and Gygi, S.P. (2008). A quantitative atlas of mitotic phosphorylation. Proc Natl Acad Sci U S A *105*, 10762-10767.

Deribe, Y.L., Pawson, T., and Dikic, I. (2010). Post-translational modifications in signal integration. Nat Struct Mol Biol *17*, 666-672.

Desai-Mehta, A., Lu, L., Ramsey-Goldman, R., and Datta, S.K. (1996). Hyperexpression of CD40 ligand by B and T cells in human lupus and its role in pathogenic autoantibody production. J Clin Invest *97*, 2063-2073.

Desterro, J.M., Rodriguez, M.S., and Hay, R.T. (1998). SUMO-1 modification of IkappaBalpha inhibits NF-kappaB activation. Mol Cell *2*, 233-239.

Di Santo, J.P. (2008). Natural killer cells: diversity in search of a niche. Nat Immunol *9*, 473-475.

Di Santo, J.P. (2014). Staying innate: transcription factor maintenance of innate lymphoid cell identity. Immunol Rev *261*, 169-176.

Diefenbach, A., Colonna, M., and Koyasu, S. (2014). Development, Differentiation, and Diversity of Innate Lymphoid Cells. Immunity *41*, 354-365.

Doi, M., Nakajima, Y., Okano, T., and Fukada, Y. (2001). Light-induced phase-delay of the chicken pineal circadian clock is associated with the induction of cE4bp4, a potential transcriptional repressor of cPer2 gene. Proc Natl Acad Sci U S A *98*, 8089-8094.

Doi, M., Okano, T., Yujnovsky, I., Sassone-Corsi, P., and Fukada, Y. (2004). Negative control of circadian clock regulator E4BP4 by casein kinase Iepsilon-mediated phosphorylation. Curr Biol *14*, 975-980.

Einhauer, A., and Jungbauer, A. (2001). The FLAG peptide, a versatile fusion tag for the purification of recombinant proteins. J Biochem Biophys Methods *49*, 455-465.

Fathman, J.W., Bhattacharya, D., Inlay, M.A., Seita, J., Karsunky, H., and Weissman, I.L. (2011). Identification of the earliest natural killer cell-committed progenitor in murine bone marrow. Blood *118*, 5439-5447.

Filtz, T.M., Vogel, W.K., and Leid, M. (2014). Regulation of transcription factor activity by interconnected post-translational modifications. Trends Pharmacol Sci *35*, 76-85.

Firth, M.A., Madera, S., Beaulieu, A.M., Gasteiger, G., Castillo, E.F., Schluns, K.S., Kubo, M., Rothman, P.B., Vivier, E., and Sun, J.C. (2013). Nfil3-independent lineage maintenance and antiviral response of natural killer cells. J Exp Med *210*, 2981-2990.

Foley, B., Cooley, S., Verneris, M.R., Curtsinger, J., Luo, X., Waller, E.K., Anasetti, C., Weisdorf, D., and Miller, J.S. (2012). Human cytomegalovirus (CMV)-induced memory-like NKG2C(+) NK cells are transplantable and expand in vivo in response to recipient CMV antigen. J Immunol *189*, 5082-5088.

Frank, C.L., Ge, X., Xie, Z., Zhou, Y., and Tsai, L.H. (2010). Control of activating transcription factor 4 (ATF4) persistence by multisite phosphorylation impacts cell cycle progression and neurogenesis. J Biol Chem *285*, 33324-33337.

Freud, A.G., Yokohama, A., Becknell, B., Lee, M.T., Mao, H.C., Ferketich, A.K., and Caligiuri, M.A. (2006). Evidence for discrete stages of human natural killer cell differentiation in vivo. J Exp Med *203*, 1033-1043.

Fritah, S., Lhocine, N., Golebiowski, F., Mounier, J., Andrieux, A., Jouvion, G., Hay, R.T., Sansonetti, P., and Dejean, A. (2014). Sumoylation controls host anti-bacterial response to the gut invasive pathogen Shigella flexneri. EMBO Rep *15*, 965-972.

Fuchs, A., Vermi, W., Lee, J.S., Lonardi, S., Gilfillan, S., Newberry, R.D., Cella, M., and Colonna, M. (2013). Intraepithelial type 1 innate lymphoid cells are a unique subset of IL-12- and IL-15-responsive IFN-gamma-producing cells. Immunity *38*, 769-781.

Galisson, F., Mahrouche, L., Courcelles, M., Bonneil, E., Meloche, S., Chelbi-Alix, M.K., and Thibault, P. (2011). A novel proteomics approach to identify SUMOylated proteins and their modification sites in human cells. Mol Cell Proteomics *10*, M110 004796.

Gantke, T., Boussouf, S., Janzen, J., Morrice, N.A., Howell, S., Muhlberger, E., and Ley, S.C. (2013). Ebola virus VP35 induces high-level production of recombinant TPL-2-ABIN-2-NF-kappaB1 p105 complex in co-transfected HEK-293 cells. Biochem J *452*, 359-365.

Garaude, J., Farras, R., Bossis, G., Charni, S., Piechaczyk, M., Hipskind, R.A., and Villalba, M. (2008). SUMOylation Regulates the Transcriptional Activity of JunB in T Lymphocytes. J Immunol *180*, 5983-5990.

Gareau, J.R., and Lima, C.D. (2010). The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. Nat Rev Mol Cell Biol *11*, 861-871.

Gascoyne, D.M., Long, E., Veiga-Fernandes, H., de Boer, J., Williams, O., Seddon, B., Coles, M., Kioussis, D., and Brady, H.J. (2009). The basic leucine zipper transcription factor E4BP4 is essential for natural killer cell development. Nat Immunol *10*, 1118-1124.

Geiger, T.L., Abt, M.C., Gasteiger, G., Firth, M.A., O'Connor, M.H., Geary, C.D., O'Sullivan, T.E., van den Brink, M.R., Pamer, E.G., Hanash, A.M.*, et al.* (2014). Nfil3 is crucial for development of innate lymphoid cells and host protection against intestinal pathogens. J Exp Med *211*, 1723-1731.

Geiss-Friedlander, R., and Melchior, F. (2007). Concepts in sumoylation: a decade on. Nat Rev Mol Cell Biol *8*, 947-956.

Geller, M.A., and Miller, J.S. (2011). Use of allogeneic NK cells for cancer immunotherapy. Immunotherapy *3*, 1445-1459.

Genomes Project, C., Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E., and McVean, G.A. (2010). A map of human genome variation from population-scale sequencing. Nature *467*, 1061-1073.

Georgopoulos, K., Winandy, S., and Avitahl, N. (1997). The role of the Ikaros gene in lymphocyte development and homeostasis. Annu Rev Immunol *15*, 155-176.

Gill, G. (2005). Something about SUMO inhibits transcription. Curr Opin Genet Dev *15*, 536-541.

Golebiowski, F., Matic, I., Tatham, M.H., Cole, C., Yin, Y., Nakamura, A., Cox, J., Barton, G.J., Mann, M., and Hay, R.T. (2009). System-wide changes to SUMO modifications in response to heat shock. Sci Signal *2*, ra24.

Gomez-del Arco, P., Koipally, J., and Georgopoulos, K. (2005). Ikaros SUMOylation: switching out of repression. Mol Cell Biol *25*, 2688-2697.

Gong, W., Xiao, W., Hu, M., Weng, X., Qian, L., Pan, X., and Ji, M. (2010). Ex vivo expansion of natural killer cells with high cytotoxicity by K562 cells modified to co-express major histocompatibility complex class I chain-related protein A, 4-1BB ligand, and interleukin-15. Tissue Antigens *76*, 467-475.

Gordon, S.M., Chaix, J., Rupp, L.J., Wu, J., Madera, S., Sun, J.C., Lindsten, T., and Reiner, S.L. (2012). The transcription factors T-bet and Eomes control key checkpoints of natural killer cell maturation. Immunity *36*, 55-67.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res *38*, W695-699.

Guo, Z., Kanjanapangka, J., Liu, N., Liu, S., Liu, C., Wu, Z., Wang, Y., Loh, T., Kowolik, C., Jamsen, J.*, et al.* (2012). Sequential Posttranslational Modifications Program FEN1 Degradation during Cell-Cycle Progression. Mol Cell *47*, 444-456.

Hackett, J., Jr., Bennett, M., Koo, G.C., and Kumar, V. (1986). Origin and differentiation of natural killer cells. III. Relationship between the precursors and effectors of natural killer and natural cytotoxic activity. Immunol Res *5*, 16-24.

Han, K.J., Jiang, L., and Shu, H.B. (2008). Regulation of IRF2 transcriptional activity by its sumoylation. Biochem Biophys Res Commun *372*, 772-778.

Hang, Y., and Stein, R. (2011). MafA and MafB activity in pancreatic beta cells. Trends Endocrinol Metab *22*, 364-373.

Hann, S.R. (2006). Role of post-translational modifications in regulating c-Myc proteolysis, transcriptional activity and biological function. Semin Cancer Biol *16*, 288-302.

Haraguchi, K., Suzuki, T., Koyama, N., Kumano, K., Nakahara, F., Matsumoto, A., Yokoyama, Y., Sakata-Yanagimoto, M., Masuda, S., Takahashi, T.*, et al.* (2009). Notch activation induces the generation of functional NK cells from human cord blood CD34-positive cells devoid of IL-15. J Immunol *182*, 6168-6178.

Hardeland, U., Steinacher, R., Jiricny, J., and Schar, P. (2002). Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. EMBO J *21*, 1456-1464.

Hay, R.T. (2005). SUMO: a history of modification. Mol Cell *18*, 1-12.

Hayakawa, Y., and Smyth, M.J. (2006a). CD27 dissects mature NK cells into two subsets with distinct responsiveness and migratory capacity. J Immunol *176*, 1517-1524.

Hayakawa, Y., and Smyth, M.J. (2006b). Innate immune recognition and suppression of tumors. Adv Cancer Res *95*, 293-322.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Herberman, R.B., Nunn, M.E., and Lavrin, D.H. (1975). Natural cytotoxic reactivity of mouse lymphoid cells against syngeneic acid allogeneic tumors. I. Distribution of reactivity and specificity. Int J Cancer *16*, 216-229.

Hietakangas, V., Anckar, J., Blomster, H.A., Fujimoto, M., Palvimo, J.J., Nakai, A., and Sistonen, L. (2006). PDSM, a motif for phosphorylation-dependent SUMO modification. Proc Natl Acad Sci U S A *103*, 45-50.

Hoege, C., Pfander, B., Moldovan, G.L., Pyrowolakis, G., and Jentsch, S. (2002). RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. Nature *419*, 135-141.

Holmberg, C.I., Hietakangas, V., Mikhailov, A., Rantanen, J.O., Kallio, M., Meinander, A., Hellman, J., Morrice, N., MacKintosh, C., Morimoto, R.I.*, et al.* (2001). Phosphorylation of serine 230 promotes inducible transcriptional activity of heat shock factor 1. EMBO J *20*, 3800-3810.

Holmberg, C.I., Tran, S.E., Eriksson, J.E., and Sistonen, L. (2002). Multisite phosphorylation provides sophisticated regulation of transcription factors. Trends Biochem Sci *27*, 619-627.

Horowitz, A., Strauss-Albee, D.M., Leipold, M., Kubo, J., Nemat-Gorgani, N., Dogan, O.C., Dekker, C.L., Mackey, S., Maecker, H., Swan, G.E.*, et al.* (2013). Genetic and environmental determinants of human NK cell diversity revealed by mass cytometry. Sci Transl Med *5*, 208ra145.

Hosoya, T., Maillard, I., and Engel, J.D. (2010). From the cradle to the grave: activities of GATA-3 throughout T-cell development and differentiation. Immunol Rev *238*, 110-125.

Hoyler, T., Klose, C.S., Souabni, A., Turqueti-Neves, A., Pfeifer, D., Rawlins, E.L., Voehringer, D., Busslinger, M., and Diefenbach, A. (2012). The transcription factor GATA-3 controls cell fate and maintenance of type 2 innate lymphoid cells. Immunity *37*, 634-648.

Hsiao, H.H., Meulmeester, E., Frank, B.T., Melchior, F., and Urlaub, H. (2009). "ChopNSpice," a mass spectrometric approach that allows identification of endogenous small ubiquitin-like modifier-conjugated peptides. Mol Cell Proteomics *8*, 2664-2675.

Huang, H., Woo, A.J., Waldon, Z., Schindler, Y., Moran, T.B., Zhu, H.H., Feng, G.S., Steen, H., and Cantor, A.B. (2012). A Src family kinase-Shp2 axis controls RUNX1 activity in megakaryocyte and T-lymphocyte differentiation. Genes Dev *26*, 1587-1601.

Huntington, N.D. (2014). The unconventional expression of IL-15 and its role in NK cell homeostasis. Immunol Cell Biol *92*, 210-213.

Huntington, N.D., Nutt, S.L., and Carotta, S. (2013). Regulation of murine natural killer cell commitment. Front Immunol *4*, 14.

Huntington, N.D., Vosshenrich, C.A., and Di Santo, J.P. (2007). Developmental pathways that generate natural-killer-cell diversity in mice and humans. Nat Rev Immunol *7*, 703-714.

Ikawa, T., Fujimoto, S., Kawamoto, H., Katsura, Y., and Yokota, Y. (2001). Commitment to natural killer cells requires the helix-loop-helix inhibitor Id2. Proc Natl Acad Sci U S A *98*, 5164-5169.

Ikushima, S., Inukai, T., Inaba, T., Nimer, S.D., Cleveland, J.L., and Look, A.T. (1997). Pivotal role for the NFIL3/E4BP4 transcription factor in interleukin 3-mediated survival of pro-B lymphocytes. Proc Natl Acad Sci U S A *94*, 2609-2614.

Imai, K., Matsuyama, S., Miyake, S., Suga, K., and Nakachi, K. (2000). Natural cytotoxic activity of peripheral-blood lymphocytes and cancer incidence: an 11-year follow-up study of a general population. The Lancet *356*, 1795-1799.

Impens, F., Radoshevich, L., Cossart, P., and Ribet, D. (2014). Mapping of SUMO sites and analysis of SUMOylation changes induced by external stimuli. Proc Natl Acad Sci U S A *111*, 12432-12437.

Jeram, S.M., Srikumar, T., Pedrioli, P.G., and Raught, B. (2009). Using mass spectrometry to identify ubiquitin and ubiquitin-like protein conjugation sites. Proteomics *9*, 922-934.

Jeram, S.M., Srikumar, T., Zhang, X.D., Anne Eisenhauer, H., Rogers, R., Pedrioli, P.G., Matunis, M., and Raught, B. (2010). An improved SUMmOn-based methodology for the identification of ubiquitin and ubiquitin-like protein conjugation sites identifies novel ubiquitin-like protein chain linkages. Proteomics *10*, 254-265.

Ji, Z., Degerny, C., Vintonenko, N., Deheuninck, J., Foveau, B., Leroy, C., Coll, J., Tulasne, D., Baert, J.L., and Fafeur, V. (2007). Regulation of the Ets-1 transcription factor by sumoylation and ubiquitinylation. Oncogene *26*, 395-406.

Jolma, A., Yan, J., Whitington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G.*, et al.* (2013). DNA-binding specificities of human transcription factors. Cell *152*, 327-339.

Junghans, D., Chauvet, S., Buhler, E., Dudley, K., Sykes, T., and Henderson, C.E. (2004). The CES-2-related transcription factor E4BP4 is an intrinsic regulator of motoneuron growth and survival. Development *131*, 4425-4434.

Kallies, A., Carotta, S., Huntington, N.D., Bernard, N.J., Tarlinton, D.M., Smyth, M.J., and Nutt, S.L. (2011). A role for Blimp1 in the transcriptional network controlling natural killer cell maturation. Blood *117*, 1869-1879.

Kamizono, S., Duncan, G.S., Seidel, M.G., Morimoto, A., Hamada, K., Grosveld, G., Akashi, K., Lind, E.F., Haight, J.P., Ohashi, P.S.*, et al.* (2009). Nfil3/E4bp4 is required for the development and maturation of NK cells in vivo. J Exp Med *206*, 2977-2986.

Kang, X., Qi, Y., Zuo, Y., Wang, Q., Zou, Y., Schwartz, R.J., Cheng, J., and Yeh, E.T. (2010). SUMO-specific protease 2 is essential for suppression of polycomb group protein-mediated gene silencing during embryonic development. Mol Cell *38*, 191-201.

Karre, K., Ljunggren, H.G., Piontek, G., and Kiessling, R. (1986). Selective rejection of H-2-deficient lymphoma variants suggests alternative immune defence strategy. Nature *319*, 675-678.

Kashiwada, M., Cassel, S.L., Colgan, J.D., and Rothman, P.B. (2011a). NFIL3/E4BP4 controls type 2 T helper cell cytokine expression. EMBO J *30*, 2071-2082.

Kashiwada, M., Levy, D.M., McKeag, L., Murray, K., Schroder, A.J., Canfield, S.M., Traver, G., and Rothman, P.B. (2010). IL-4-induced transcription factor NFIL3/E4BP4 controls IgE class switching. Proc Natl Acad Sci U S A *107*, 821-826.

Kashiwada, M., Pham, N.L., Pewe, L.L., Harty, J.T., and Rothman, P.B. (2011b). NFIL3/E4BP4 is a key transcription factor for CD8alpha dendritic cell development. Blood *117*, 6193-6197.

Kaufman, D.S. (2009). Toward clinical therapies using hematopoietic cells derived from human pluripotent stem cells. Blood *114*, 3513-3523.

Keniry, M., Pires, M.M., Mense, S., Lefebvre, C., Gan, B., Justiano, K., Lau, Y.K., Hopkins, B., Hodakoski, C., Koujak, S.*, et al.* (2013). Survival factor NFIL3 restricts FOXO-induced gene expression in cancer. Genes Dev *27*, 916-927.

Kerscher, O. (2007). SUMO junction-what's your function? New insights through SUMO-interacting motifs. EMBO Rep *8*, 550-555.

Kessler, J.D., Kahle, K.T., Sun, T., Meerbrey, K.L., Schlabach, M.R., Schmitt, E.M., Skinner, S.O., Xu, Q., Li, M.Z., Hartman, Z.C.*, et al.* (2012). A SUMOylation-dependent transcriptional subprogram is required for Myc-driven tumorigenesis. Science *335*, 348-353.

Kettenbach, A.N., Schweppe, D.K., Faherty, B.K., Pechenick, D., Pletnev, A.A., and Gerber, S.A. (2011). Quantitative phosphoproteomics identifies substrates and functional modules of Aurora and Polo-like kinase activities in mitotic cells. Sci Signal *4*, rs5.

Kiessling, R., Klein, E., and Wigzell, H. (1975). "Natural" killer cells in the mouse. I. Cytotoxic cells with specificity for mouse Moloney leukemia cells. Specificity and distribution according to genotype. Eur J Immunol *5*, 112-117.

Kim, S., Iizuka, K., Kang, H.S., Dokun, A., French, A.R., Greco, S., and Yokoyama, W.M. (2002). In vivo developmental stages in murine natural killer cell maturation. Nat Immunol *3*, 523-528.

Klenova, E., Chernukhin, I., Inoue, T., Shamsuddin, S., and Norton, J. (2002). Immunoprecipitation techniques for the analysis of transcription factor complexes. Methods *26*, 254-259.

Klose, C.S., Flach, M., Mohle, L., Rogell, L., Hoyler, T., Ebert, K., Fabiunke, C., Pfeifer, D., Sexl, V., Fonseca-Pereira, D.*, et al.* (2014). Differentiation of Type 1 ILCs from a Common Progenitor to All Helper-like Innate Lymphoid Cell Lineages. Cell *157*, 340-356.

Knipscheer, P., van Dijk, W.J., Olsen, J.V., Mann, M., and Sixma, T.K. (2007). Noncovalent interaction between Ubc9 and SUMO promotes SUMO chain formation. EMBO J *26*, 2797-2807.

Knorr, D.A., Bachanova, V., Verneris, M.R., and Miller, J.S. (2014). Clinical utility of natural killer cells in cancer therapy and transplantation. Semin Immunol *26*, 161-172.

Knorr, D.A., Ni, Z., Hermanson, D., Hexum, M.K., Bendzick, L., Cooper, L.J., Lee, D.A., and Kaufman, D.S. (2013). Clinical-scale derivation of natural killer cells from human pluripotent stem cells for cancer therapy. Stem Cells Transl Med *2*, 274-283.

Kobayashi, T., Matsuoka, K., Sheikh, S.Z., Elloumi, H.Z., Kamada, N., Hisamatsu, T., Hansen, J.J., Doty, K.R., Pope, S.D., Smale, S.T.*, et al.* (2011). NFIL3 is a regulator of IL-12 p40 in macrophages and mucosal immunity. J Immunol *186*, 4649-4655.

Kobayashi, T., Steinbach, E.C., Russo, S.M., Matsuoka, K., Nochi, T., Maharshak, N., Borst, L.B., Hostager, B., Garcia-Martinez, J.V., Rothman, P.B.*, et al.* (2014). NFIL3-deficient mice develop microbiota-dependent, IL-12/23-driven spontaneous colitis. J Immunol *192*, 1918-1927.

Koch, J., Steinle, A., Watzl, C., and Mandelboim, O. (2013). Activating natural cytotoxicity receptors of natural killer cells in cancer and infection. Trends Immunol *34*, 182-191.

Koipally, J., and Georgopoulos, K. (2000). Ikaros interactions with CtBP reveal a repression mechanism that is independent of histone deacetylase activity. J Biol Chem *275*, 19594-19602.

Koipally, J., and Georgopoulos, K. (2002). Ikaros-CtIP interactions do not require C-terminal binding protein and participate in a deacetylase-independent mode of repression. J Biol Chem *277*, 23143-23149.

Koipally, J., Renold, A., Kim, J., and Georgopoulos, K. (1999). Repression by Ikaros and Aiolos is mediated through histone deacetylase complexes. EMBO J *18*, 3090-3100.

Kuo, H.Y., Chang, C.C., Jeng, J.C., Hu, H.M., Lin, D.Y., Maul, G.G., Kwok, R.P., and Shih, H.M. (2005). SUMO modification negatively modulates the transcriptional activity of CREB-binding protein via the recruitment of Daxx. Proc Natl Acad Sci U S A *102*, 16973-16978.

Kuribara, R., Kinoshita, T., Miyajima, A., Shinjyo, T., Yoshihara, T., Inukai, T., Ozawa, K., Look, A.T., and Inaba, T. (1999). Two distinct interleukin-3-mediated signal pathways, Ras-NFIL3 (E4BP4) and Bcl-x(L), regulate the survival of murine pro-B lymphocytes. Mol Cell Biol *19*, 2754-2762.

Lacorazza, H.D., Miyazaki, Y., Di Cristofano, A., Deblasio, A., Hedvat, C., Zhang, J., Cordon-Cardo, C., Mao, S., Pandolfi, P.P., and Nimer, S.D. (2002). The ETS protein MEF plays a critical role in perforin gene expression and the development of natural killer and NK-T cells. Immunity *17*, 437-449.

Lai, C.B., and Mager, D.L. (2012). Role of runt-related transcription factor 3 (RUNX3) in transcription regulation of natural cytotoxicity receptor 1 (NCR1/NKp46), an activating natural killer (NK) cell receptor. J Biol Chem *287*, 7324-7334.

Lallemand-Breitenbach, V., Jeanne, M., Benhenda, S., Nasr, R., Lei, M., Peres, L., Zhou, J., Zhu, J., Raught, B., and de The, H. (2008). Arsenic degrades PML or PML-RARalpha through a SUMO-triggered RNF4/ubiquitin-mediated pathway. Nat Cell Biol *10*, 547-555.

Lang, R., Patel, D., Morris, J.J., Rutschman, R.L., and Murray, P.J. (2002). Shaping gene expression in activated and resting primary macrophages by IL-10. J Immunol *169*, 2253-2263.

Lanucara, F., and Eyers, C.E. (2013). Top-down mass spectrometry for the analysis of combinatorial post-translational modifications. Mass Spectrom Rev *32*, 27-42.

Lee, B., and Muller, M.T. (2009). SUMOylation enhances DNA methyltransferase 1 activity. Biochem J *421*, 449-461.

Li, Q.R., Ning, Z.B., Tang, J.S., Nie, S., and Zeng, R. (2009). Effect of peptide-to-TiO2 beads ratio on phosphopeptide enrichment selectivity. J Proteome Res *8*, 5375-5381.

Lin, D.Y., Huang, Y.S., Jeng, J.C., Kuo, H.Y., Chang, C.C., Chao, T.T., Ho, C.C., Chen, Y.C., Lin, T.P., Fang, H.I*., et al.* (2006). Role of SUMO-interacting motif in Daxx SUMO modification, subnuclear localization, and repression of sumoylated transcription factors. Mol Cell *24*, 341-354.

Lindberg, M.J., Popko-Scibor, A.E., Hansson, M.L., and Wallberg, A.E. (2010). SUMO modification regulates the transcriptional activity of MAML1. FASEB J *24*, 2396-2404.

Ljunggren, H.G., and Malmberg, K.J. (2007). Prospects for the use of NK cells in immunotherapy of human cancer. Nat Rev Immunol *7*, 329-339.

Lohoff, M., Duncan, G.S., Ferrick, D., Mittrucker, H.W., Bischof, S., Prechtl, S., Rollinghoff, M., Schmitt, E., Pahl, A., and Mak, T.W. (2000). Deficiency in the transcription factor interferon regulatory factor (IRF)-2 leads to severely compromised development of natural killer and T helper type 1 cells. J Exp Med *192*, 325-336.

Lotem, J., Levanon, D., Negreanu, V., Leshkowitz, D., Friedlander, G., and Groner, Y. (2013). Runx3-mediated transcriptional program in cytotoxic lymphocytes. PLoS ONE *8*, e80467.

Luevano, M., Domogala, A., Blundell, M., Jackson, N., Pedroza-Pacheco, I., Derniame, S., Escobedo-Cousin, M., Querol, S., Thrasher, A., Madrigal, A*., et al.* (2014). Frozen cord blood hematopoietic stem cells differentiate into higher numbers of functional natural killer cells in vitro than mobilized hematopoietic stem cells or freshly isolated cord blood hematopoietic stem cells. PLoS ONE *9*, e87086.

Luevano, M., Madrigal, A., and Saudemont, A. (2012). Generation of natural killer cells from hematopoietic stem cells in vitro for immunotherapy. Cell Mol Immunol *9*, 310-320.

MacGillavry, H.D., Cornelis, J., van der Kallen, L.R., Sassen, M.M., Verhaagen, J., Smit, A.B., and van Kesteren, R.E. (2011). Genome-wide gene expression and promoter binding analysis identifies NFIL3 as a repressor of C/EBP target genes in neuronal outgrowth. Mol Cell Neurosci *46*, 460-468.

MacGillavry, H.D., Stam, F.J., Sassen, M.M., Kegel, L., Hendriks, W.T., Verhaagen, J., Smit, A.B., and van Kesteren, R.E. (2009). NFIL3 and cAMP response element-binding protein form a transcriptional feedforward loop that controls neuronal regeneration-associated gene expression. J Neurosci *29*, 15542-15550.

Maclaine, N.J., and Hupp, T.R. (2011). How phosphorylation controls p53. Cell Cycle *10*, 916-921.

Mahajan, R., Delphin, C., Guan, T., Gerace, L., and Melchior, F. (1997). A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. Cell *88*, 97-107.

Male, V., and Brady, H.J. (2014). Transcriptional control of NK cell differentiation and function. Curr Top Microbiol Immunol *381*, 173-187.

Male, V., Nisoli, I., Gascoyne, D.M., and Brady, H.J. (2012). E4BP4: an unexpected player in the immune response. Trends Immunol *33*, 98-102.

Male, V., Nisoli, I., Kostrzewski, T., Allan, D.S., Carlyle, J.R., Lord, G.M., Wack, A., and Brady, H.J. (2014). The transcription factor E4bp4/Nfil3 controls commitment to the NK lineage and directly regulates Eomes and Id2 expression. J Exp Med *211*, 635-642.

Manaster, I., and Mandelboim, O. (2008). The unique properties of human NK cells in the uterine mucosa. Placenta *29 Suppl A*, S60-66.

Marcais, A., Cherfils-Vicini, J., Viant, C., Degouve, S., Viel, S., Fenis, A., Rabilloud, J., Mayol, K., Tavares, A., Bienvenu, J.*, et al.* (2014). The metabolic checkpoint kinase mTOR is essential for IL-15 signaling during the development and activation of NK cells. Nat Immunol *15*, 749-757.

Martin, S., Wilkinson, K.A., Nishimune, A., and Henley, J.M. (2007). Emerging extranuclear roles of protein SUMOylation in neuronal function and dysfunction. Nat Rev Neurosci *8*, 948-959.

Matic, I., Schimmel, J., Hendriks, I.A., van Santen, M.A., van de Rijke, F., van Dam, H., Gnad, F., Mann, M., and Vertegaal, A.C. (2010). Site-specific identification of SUMO-2 targets in cells reveals an inverted SUMOylation motif and a hydrophobic cluster SUMOylation motif. Mol Cell *39*, 641-652.

Mayya, V., Lundgren, D.H., Hwang, S.I., Rezaul, K., Wu, L., Eng, J.K., Rodionov, V., and Han, D.K. (2009). Quantitative phosphoproteomic analysis of T cell receptor signaling reveals system-wide modulation of protein-protein interactions. Sci Signal *2*, ra46.

McClory, S., Hughes, T., Freud, A.G., Briercheck, E.L., Martin, C., Trimboli, A.J., Yu, J., Zhang, X., Leone, G., Nuovo, G.*, et al.* (2012). Evidence for a stepwise program of extrathymic T cell development within the human tonsil. J Clin Invest *122*, 1403-1415.

McElroy, D.S., Ashley, T.J., and D'Orazio, S.E. (2009). Lymphocytes serve as a reservoir for Listeria monocytogenes growth during infection of mice. Microb Pathog *46*, 214-221.

Miller, J.S., Cooley, S., Parham, P., Farag, S.S., Verneris, M.R., McQueen, K.L., Guethlein, L.A., Trachtenberg, E.A., Haagenson, M., Horowitz, M.M.*, et al.* (2007). Missing KIR ligands are associated with less relapse and increased graft-versus-host disease (GVHD) following unrelated donor allogeneic HCT. Blood *109*, 5058-5061.

Mitsui, S., Yamaguchi, S., Matsuo, T., Ishida, Y., and Okamura, H. (2001). Antagonistic role of E4BP4 and PAR proteins in the circadian oscillatory mechanism. Genes Dev *15*, 995-1006.

Mjosberg, J., Bernink, J., Golebski, K., Karrich, J.J., Peters, C.P., Blom, B., te Velde, A.A., Fokkens, W.J., van Drunen, C.M., and Spits, H. (2012). The transcription factor GATA3 is essential for the function of human type 2 innate lymphoid cells. Immunity *37*, 649-659.

Mohideen, F., Capili, A.D., Bilimoria, P.M., Yamada, T., Bonni, A., and Lima, C.D. (2009). A molecular basis for phosphorylation-dependent SUMO conjugation by the E2 UBC9. Nat Struct Mol Biol *16*, 945-952.

Moretta, A., Marcenaro, E., Parolini, S., Ferlazzo, G., and Moretta, L. (2008). NK cells at the interface between innate and adaptive immunity. Cell Death Differ *15*, 226-233.

Motomura, Y., Kitamura, H., Hijikata, A., Matsunaga, Y., Matsumoto, K., Inoue, H., Atarashi, K., Hori, S., Watarai, H., Zhu, J.*, et al.* (2011). The transcription factor E4BP4 regulates the production of IL-10 and IL-13 in CD4+ T cells. Nat Immunol *12*, 450-459.

Muller, S., Matunis, M.J., and Dejean, A. (1998). Conjugation with the ubiquitin-related modifier SUMO-1 regulates the partitioning of PML within the nucleus. EMBO J *17*, 61-70.

Mumby, M., and Brekken, D. (2005). Phosphoproteomics: new insights into cellular signaling. Genome Biol *6*, 230.

Nandakumar, V., Chou, Y., Zang, L., Huang, X.F., and Chen, S.Y. (2013). Epigenetic control of natural killer cell maturation by histone H2A deubiquitinase, MYSM1. Proc Natl Acad Sci U S A *110*, E3927-3936.

Naper, C., Shegarfi, H., Inngjerdingen, M., and Rolstad, B. (2011). The role of natural killer cells in the defense against Listeria monocytogenes lessons from a rat model. J Innate Immun *3*, 289-297.

Narni-Mancinelli, E., Jaeger, B.N., Bernat, C., Fenis, A., Kung, S., De Gassart, A., Mahmood, S., Gut, M., Heath, S.C., Estelle, J.*, et al.* (2012). Tuning of natural killer cell reactivity by NKp46 and Helios calibrates T cell responses. Science *335*, 344-348.

Ni, Z., Knorr, D.A., Clouser, C.L., Hexum, M.K., Southern, P., Mansky, L.M., Park, I.H., and Kaufman, D.S. (2011). Human pluripotent stem cells produce natural killer cells that mediate anti-HIV-1 activity by utilizing diverse cellular mechanisms. J Virol *85*, 43-50.

Ning, H., Lei, H.E., Xu, Y.D., Guan, R.L., Venstrom, J.M., Lin, G., Lue, T.F., Xin, Z., and Lin, C.S. (2014). Conversion of adipose-derived stem cells into natural killer-like cells with anti-tumor activities in nude mice. PLoS ONE *9*, e106246.

Nutt, S.L., Metcalf, D., D'Amico, A., Polli, M., and Wu, L. (2005). Dynamic regulation of PU.1 expression in multipotent hematopoietic progenitors. J Exp Med *201*, 221-231.

Obenauer, J.C. (2003). Scansite 2.0: proteome-wide prediction of cell signaling interactions using short sequence motifs. Nucleic Acids Res *31*, 3635-3641.

Ohno, S., Sato, T., Kohu, K., Takeda, K., Okumura, K., Satake, M., and Habu, S. (2008). Runx proteins are involved in regulation of CD122, Ly49 family and IFN-gamma expression during NK cell differentiation. Int Immunol *20*, 71-79.

Ohno, T., Onishi, Y., and Ishida, N. (2007a). The negative transcription factor E4BP4 is associated with circadian clock protein PERIOD2. Biochem Biophys Res Commun *354*, 1010-1015.

Ohno, T., Onishi, Y., and Ishida, N. (2007b). A novel E4BP4 element drives circadian expression of mPeriod2. Nucleic Acids Res *35*, 648-655.

Okamura, H., Aramburu, J., Garcia-Rodriguez, C., Viola, J.P., Raghavan, A., Tahiliani, M., Zhang, X., Qin, J., Hogan, P.G., and Rao, A. (2000). Concerted dephosphorylation of the transcription factor NFAT1 induces a conformational switch that regulates transcriptional activity. Mol Cell *6*, 539-550.

Olsen, J.V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M.L., Jensen, L.J., Gnad, F., Cox, J., Jensen, T.S., Nigg, E.A.*, et al.* (2010). Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. Sci Signal *3*, ra3.

Orange, J.S. (2006). Human natural killer cell deficiencies. Curr Opin Allergy Clin Immunol *6*, 399-409.

Orkin, S.H., and Zon, L.I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. Cell *132*, 631-644.

Orr, S.J., Roessler, S., Quigley, L., Chan, T., Ford, J.W., O'Connor, G.M., and McVicar, D.W. (2010). Implications for gene therapy-limiting expression of IL-2R gamma c delineate differences in signaling thresholds required for lymphocyte development and maintenance. J Immunol *185*, 1393-1403.

Osula, O., Swatkoski, S., and Cotter, R.J. (2012). Identification of protein SUMOylation sites by mass spectrometry using combined microwave-assisted aspartic acid cleavage and tryptic digestion. J Mass Spectrom *47*, 644-654.

Ouyang, J., Shi, Y., Valin, A., Xuan, Y., and Gill, G. (2009). Direct binding of CoREST1 to SUMO-2/3 contributes to gene-specific repression by the LSD1/CoREST1/HDAC complex. Mol Cell *34*, 145-154.

Owerbach, D., McKay, E.M., Yeh, E.T., Gabbay, K.H., and Bohren, K.M. (2005). A proline-90 residue unique to SUMO-4 prevents maturation and sumoylation. Biochem Biophys Res Commun *337*, 517-520.

Ozkurt, I.C., and Tetradis, S. (2003). Parathyroid hormone-induced E4BP4/NFIL3 down-regulates transcription in osteoblasts. J Biol Chem *278*, 26803-26809.

Papatriantafyllou, M. (2013). Haematopoiesis: Early decisions. Nat Rev Immunol *13*, 304-305.

Papavassiliou, A.G., Treier, M., and Bohmann, D. (1995). Intramolecular signal transduction in c-Jun. EMBO J *14*, 2014-2019.

Pegram, H.J., Andrews, D.M., Smyth, M.J., Darcy, P.K., and Kershaw, M.H. (2011). Activating and inhibitory receptors of natural killer cells. Immunol Cell Biol *89*, 216-224.

Peng, H., Jiang, X., Chen, Y., Sojka, D.K., Wei, H., Gao, X., Sun, R., Yokoyama, W.M., and Tian, Z. (2013). Liver-resident NK cells confer adaptive immunity in skin-contact inflammation. J Clin Invest *123*, 1444-1456.

Perdomo, J., Jiang, X.M., Carter, D.R., Khachigian, L.M., and Chong, B.H. (2012). SUMOylation Regulates the Transcriptional Repression Activity of FOG-2 and Its Association with GATA-4. PLoS ONE *7*, e50637.

Perussia, B., Chen, Y., and Loza, M.J. (2005). Peripheral NK cell phenotypes: multiple changing of faces of an adapting, developing cell. Mol Immunol *42*, 385-395.

Pichler, A., Knipscheer, P., Oberhofer, E., van Dijk, W.J., Korner, R., Olsen, J.V., Jentsch, S., Melchior, F., and Sixma, T.K. (2005). SUMO modification of the ubiquitin-conjugating enzyme E2-25K. Nat Struct Mol Biol *12*, 264-269.

Pinho, M.J., Marques, C.J., Carvalho, F., Punzel, M., Sousa, M., and Barros, A. (2012). Genetic regulation on ex vivo differentiated natural killer cells from human umbilical cord blood CD34+ cells. J Recept Signal Transduct Res *32*, 238-249.

Pirkkala, L., Nykanen, P., and Sistonen, L. (2001). Roles of the heat shock transcription factors in regulation of the heat shock response and beyond. FASEB J *15*, 1118-1131.

Pouponnot, C., Sii-Felice, K., Hmitou, I., Rocques, N., Lecoin, L., Druillennec, S., Felder-Schmittbuhl, M.P., and Eychene, A. (2006). Cell context reveals a dual role for Maf in oncogenesis. Oncogene *25*, 1299-1310.

Previs, M.J., VanBuren, P., Begin, K.J., Vigoreaux, J.O., LeWinter, M.M., and Matthews, D.E. (2008). Quantification of protein phosphorylation by liquid chromatography-mass spectrometry. Analytical Chemistry *80*, 5864-5872.

Priceman, S.J., Kirzner, J.D., Nary, L.J., Morris, D., Shankar, D.B., Sakamoto, K.M., and Medh, R.D. (2006). Calcium-dependent upregulation of E4BP4 expression correlates with glucocorticoid-evoked apoptosis of human leukemic CEM cells. Biochem Biophys Res Commun *344*, 491-499.

Pufall, M.A., Lee, G.M., Nelson, M.L., Kang, H.S., Velyvis, A., Kay, L.E., McIntosh, L.P., and Graves, B.J. (2005). Variable control of Ets-1 DNA binding by multiple phosphates in an unstructured region. Science *309*, 142-145.

Ramirez, K., Chandler, K.J., Spaulding, C., Zandi, S., Sigvardsson, M., Graves, B.J., and Kee, B.L. (2012). Gene deregulation and chronic activation in natural killer cells deficient in the transcription factor ETS1. Immunity *36*, 921-932.

Ramirez, K., and Kee, B.L. (2010). Multiple hats for natural killers. Curr Opin Immunol *22*, 193-198.

Ramsborg, C.G., and Papoutsakis, E.T. (2007). Global transcriptional analysis delineates the differential inflammatory response interleukin-15 elicits from cultured human T cells. Exp Hematol *35*, 454-464.

Rappsilber, J., Ishihama, Y., and Mann, M. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Analytical Chemistry *75*, 663-670.

Rena, G., Bain, J., Elliott, M., and Cohen, P. (2004). D4476, a cell-permeant inhibitor of CK1, suppresses the site-specific phosphorylation and nuclear exclusion of FOXO1a. EMBO Rep *5*, 60-65.

Ribet, D., Hamon, M., Gouin, E., Nahori, M.A., Impens, F., Neyret-Kahn, H., Gevaert, K., Vandekerckhove, J., Dejean, A., and Cossart, P. (2010). Listeria monocytogenes impairs SUMOylation for efficient infection. Nature *464*, 1192-1195.

Riley, T., Sontag, E., Chen, P., and Levine, A. (2008). Transcriptional control of human p53-regulated genes. Nat Rev Mol Cell Biol *9*, 402-412.

Rivas, M., Mellstrom, B., Naranjo, J.R., and Santisteban, P. (2004). Transcriptional repressor DREAM interacts with thyroid transcription factor-1 and regulates thyroglobulin gene expression. J Biol Chem *279*, 33114-33122.

Robertson, M.J., and Ritz, J. (1990). Biology and clinical relevance of human natural killer cells. Blood *76*, 2421-2438.

Roger, J.E., Nellissery, J., Kim, D.S., and Swaroop, A. (2010). Sumoylation of bZIP transcription factor NRL modulates target gene expression during photoreceptor differentiation. J Biol Chem *285*, 25637-25644.

Rosmaraki, E.E., Douagi, I., Roth, C., Colucci, F., Cumano, A., and Di Santo, J.P. (2001). Identification of committed NK cell progenitors in adult murine bone marrow. Eur J Immunol *31*, 1900-1909.

Rothenberg, E.V. (2014). Transcriptional control of early T and B cell developmental choices. Annu Rev Immunol *32*, 283-321.

Ruggeri, L., Capanni, M., Urbani, E., Perruccio, K., Shlomchik, W.D., Tosti, A., Posati, S., Rogaia, D., Frassoni, F., Aversa, F.*, et al.* (2002). Effectiveness of donor natural killer cell alloreactivity in mismatched hematopoietic transplants. Science *295*, 2097-2100.

Rujkijyanont, P., Chan, W.K., Eldridge, P.W., Lockey, T., Holladay, M., Rooney, B., Davidoff, A.M., Leung, W., and Vong, Q. (2013). Ex vivo activation of CD56(+) immune cells that eradicate neuroblastoma. Cancer Res *73*, 2608-2618.

Saitoh, H., and Hinchey, J. (2000). Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. J Biol Chem *275*, 6252-6258.

Samson, S.I., Richard, O., Tavian, M., Ranson, T., Vosshenrich, C.A.J., Colucci, F., Buer, J., Grosveld, F., Godin, I., and Di Santo, J.P. (2003). GATA-3 promotes maturation, IFN-gamma production, and liver-specific homing of NK cells. Immunity *19*, 701-711.

Sarge, K.D., and Park-Sarge, O.K. (2009). Detection of proteins sumoylated in vivo and in vitro. Methods Mol Biol *590*, 265-277.

Satoh-Takayama, N., Lesjean-Pottier, S., Vieira, P., Sawa, S., Eberl, G., Vosshenrich, C.A., and Di Santo, J.P. (2010). IL-7 and IL-15 independently program the differentiation of intestinal CD3-NKp46+ cell subsets from Id2-dependent precursors. J Exp Med *207*, 273-280.

Scheper, W., Grunder, C., Straetemans, T., Sebestyen, Z., and Kuball, J. (2014). Hunting for clinical translation with innate-like immune cells and their receptors. Leukemia *28*, 1181-1190.

Schimmel, J., Larsen, K.M., Matic, I., van Hagen, M., Cox, J., Mann, M., Andersen, J.S., and Vertegaal, A.C.O. (2008). The Ubiquitin-Proteasome System Is a Key Component of the SUMO-2/3 Cycle. Molecular & Cellular Proteomics *7*, 2107-2122.

Schluns, K.S., Nowak, E.C., Cabrera-Hernandez, A., Puddington, L., Lefrancois, L., and Aguila, H.L. (2004). Distinct cell types control lymphoid subset development by means of IL-15 and IL-15 receptor alpha expression. Proc Natl Acad Sci U S A *101*, 5616-5621.

Schneider-Poetsch, T., Ju, J., Eyler, D.E., Dang, Y., Bhat, S., Merrick, W.C., Green, R., Shen, B., and Liu, J.O. (2010). Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. Nat Chem Biol *6*, 209-217.

Schotte, R., Dontje, W., Nagasawa, M., Yasuda, Y., Bakker, A.Q., Spits, H., and Blom, B. (2010). Synergy between IL-15 and Id2 promotes the expansion of human NK progenitor cells, which can be counteracted by the E protein HEB required to drive T cell development. J Immunol *184*, 6670-6679.

Schou, J., Kelstrup, C.D., Hayward, D.G., Olsen, J.V., and Nilsson, J. (2014). Comprehensive Identification of SUMO2/3 Targets and Their Dynamics during Mitosis. PLoS ONE *9*, e100692.

Schroder, A.J., Pavlidis, P., Arimura, A., Capece, D., and Rothman, P.B. (2002). Cutting edge: STAT6 serves as a positive and negative regulator of gene expression in IL-4-stimulated B lymphocytes. J Immunol *168*, 996-1000.

Schumacher, M.A., Goodman, R.H., and Brennan, R.G. (2000). The structure of a CREB bZIP.somatostatin CRE complex reveals the basis for selective dimerization and divalent cation-enhanced DNA binding. J Biol Chem *275*, 35242-35247.

Seillet, C., Huntington, N.D., Gangatirkar, P., Axelsson, E., Minnich, M., Brady, H.J., Busslinger, M., Smyth, M.J., Belz, G.T., and Carotta, S. (2014a). Differential Requirement for Nfil3 during NK Cell Development. J Immunol *192*, 2667-2676.

Seillet, C., Jackson, J.T., Markey, K.A., Brady, H.J., Hill, G.R., Macdonald, K.P., Nutt, S.L., and Belz, G.T. (2013). CD8alpha+ DCs can be induced in the absence of transcription factors Id2, Nfil3, and Batf3. Blood *121*, 1574-1583.

Seillet, C., Rankin, L.C., Groom, J.R., Mielke, L.A., Tellier, J., Chopin, M., Huntington, N.D., Belz, G.T., and Carotta, S. (2014b). Nfil3 is required for the development of all innate lymphoid cell subsets. J Exp Med *211*, 1733-1740.

Serafini, N., Klein Wolterink, R.G., Satoh-Takayama, N., Xu, W., Vosshenrich, C.A., Hendriks, R.W., and Di Santo, J.P. (2014). Gata3 drives development of RORgammat+ group 3 innate lymphoid cells. J Exp Med *211*, 199-208.

Shalizi, A., Gaudilliere, B., Yuan, Z., Stegmuller, J., Shirogane, T., Ge, Q., Tan, Y., Schulman, B., Harper, J.W., and Bonni, A. (2006). A calcium-regulated MEF2 sumoylation switch controls postsynaptic differentiation. Science *311*, 1012-1017.

Shao, C., and Cobb, M.H. (2009). Sumoylation regulates the transcriptional activity of MafA in pancreatic beta cells. J Biol Chem *284*, 3117-3124.

Shen, J., Sheng, X., Chang, Z., Wu, Q., Wang, S., Xuan, Z., Li, D., Wu, Y., Shang, Y., Kong, X.*, et al.* (2014). Iron metabolism regulates p53 signaling through direct heme-p53 interaction and modulation of p53 localization, stability, and function. Cell Rep *7*, 180-193.

Silvestris, F., Cafforio, P., De Matteo, M., Calvani, N., Frassanito, M.A., and Dammacco, F. (2008). Negative regulation of the osteoblast function in multiple myeloma through the repressor gene E4BP4 activated by malignant plasma cells. Clinical Cancer Research *14*, 6081-6091.

Smith, A.M., Qualls, J.E., O'Brien, K., Balouzian, L., Johnson, P.F., Schultz-Cherry, S., Smale, S.T., and Murray, P.J. (2011). A distal enhancer in Il12b is the target of transcriptional repression by the STAT3 pathway and requires the basic leucine zipper (B-ZIP) protein NFIL3. J Biol Chem *286*, 23582-23590.

Sojka, D.K., Plougastel-Douglas, B., Yang, L., Pak-Wittel, M.A., Artyomov, M.N., Ivanova, Y., Zhong, C., Chase, J.M., Rothman, P.B., Yu, J.*, et al.* (2014). Tissue-resident natural killer (NK) cells are cell lineages distinct from thymic and conventional splenic NK cells. Elife (Cambridge) *3*, e01659.

Song, J., Durrin, L.K., Wilkinson, T.A., Krontiris, T.G., and Chen, Y. (2004). Identification of a SUMO-binding motif that recognizes SUMO-modified proteins. Proc Natl Acad Sci U S A *101*, 14373-14378.

Spanholtz, J., Tordoir, M., Eissens, D., Preijers, F., van der Meer, A., Joosten, I., Schaap, N., de Witte, T.M., and Dolstra, H. (2010). High log-scale expansion of functional human natural killer cells from umbilical cord blood CD34-positive cells for adoptive cancer immunotherapy. PLoS ONE *5*, e9221.

Spektor, T.M., Congdon, L.M., Veerappan, C.S., and Rice, J.C. (2011). The UBC9 E2 SUMO conjugating enzyme binds the PR-Set7 histone methyltransferase to facilitate target gene repression. PLoS ONE *6*, e22785.

Spits, H., Artis, D., Colonna, M., Diefenbach, A., Di Santo, J.P., Eberl, G., Koyasu, S., Locksley, R.M., McKenzie, A.N., Mebius, R.E.*, et al.* (2013). Innate lymphoid cells--a proposal for uniform nomenclature. Nat Rev Immunol *13*, 145-149.

Steen, H., Jebanathirajah, J.A., Springer, M., and Kirschner, M.W. (2005). Stable isotope-free relative and absolute quantitation of protein phosphorylation stoichiometry by MS. Proc Natl Acad Sci U S A *102*, 3948-3953.

Su, Y.F., Shyu, Y.C., Shen, C.K., and Hwang, J. (2012). Phosphorylation-Dependent SUMOylation of the Transcription Factor NF-E2. PLoS ONE *7*, e44608.

Sun, J.C., Beilke, J.N., and Lanier, L.L. (2009). Adaptive immune features of natural killer cells. Nature *457*, 557-561.

Sun, J.C., Lopez-Verges, S., Kim, C.C., DeRisi, J.L., and Lanier, L.L. (2011). NK cells and immune "memory". J Immunol *186*, 1891-1897.

Sun, J.C., Madera, S., Bezman, N.A., Beilke, J.N., Kaplan, M.H., and Lanier, L.L. (2012). Proinflammatory cytokine signaling required for the generation of natural killer cell memory. J Exp Med *209*, 947-954.

Taki, S., Nakajima, S., Ichikawa, E., Saito, T., and Hida, S. (2005). IFN regulatory factor-2 deficiency revealed a novel checkpoint critical for the generation of peripheral NK cells. J Immunol *174*, 6005-6012.

Tamai, S., Imaizumi, K., Kurabayashi, N., Nguyen, M.D., Abe, T., Inoue, M., Fukada, Y., and Sanada, K. (2014). Neuroprotective role of the basic leucine zipper transcription factor NFIL3 in models of amyotrophic lateral sclerosis. J Biol Chem *289*, 1629-1638.

Tammsalu, T., Matic, I., Jaffray, E.G., Ibrahim, A.F., Tatham, M.H., and Hay, R.T. (2014). Proteome-Wide Identification of SUMO2 Modification Sites. Sci Signal *7*, rs2.

Tang, Y., Zhao, W., Chen, Y., Zhao, Y., and Gu, W. (2008). Acetylation is indispensable for p53 activation. Cell *133*, 612-626.

Tatham, M.H., Jaffray, E., Vaughan, O.A., Desterro, J.M., Botting, C.H., Naismith, J.H., and Hay, R.T. (2001). Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. J Biol Chem *276*, 35368-35374.

Tatham, M.H., Rodriguez, M.S., Xirodimas, D.P., and Hay, R.T. (2009). Detection of protein SUMOylation in vivo. Nat Protoc *4*, 1363-1371.

Teng, S., Luo, H., and Wang, L. (2012). Predicting protein sumoylation sites from sequence features. Amino Acids *43*, 447-455.

Themeli, M., Kloss, C.C., Ciriello, G., Fedorov, V.D., Perna, F., Gonen, M., and Sadelain, M. (2013). Generation of tumor-targeted human T lymphocytes from induced pluripotent stem cells for cancer therapy. Nat Biotechnol *31*, 928-933.

Tiscornia, G., Singer, O., and Verma, I.M. (2006). Production and purification of lentiviral vectors. Nat Protoc *1*, 241-245.

Tong, X., Zhang, D., Buelow, K., Guha, A., Arthurs, B., Brady, H.J., and Yin, L. (2013). Recruitment of histone methyltransferase G9a mediates transcriptional repression of Fgf21 gene by E4BP4 protein. J Biol Chem *288*, 5417-5425.

Tootle, T.L., and Rebay, I. (2005). Post-translational modifications influence transcription factor activity: a view from the ETS superfamily. Bioessays *27*, 285-298.

Topham, N.J., and Hewitt, E.W. (2009). Natural killer cell cytotoxicity: how do they pull the trigger? Immunology *128*, 7-15.

Ueda, H.R., Hayashi, S., Chen, W., Sano, M., Machida, M., Shigeyoshi, Y., Iino, M., and Hashimoto, S. (2005). System-level identification of transcriptional circuits underlying mammalian circadian clocks. Nat Genet *37*, 187-192.

Van Nguyen, T., Angkasekwinai, P., Dou, H., Lin, F.M., Lu, L.S., Cheng, J., Chin, Y.E., Dong, C., and Yeh, E.T. (2012). SUMO-Specific Protease 1 Is Critical for Early Lymphoid Development through Regulation of STAT5 Activation. Mol Cell *45*, 210-221.

Vaquerizas, J.M., Kummerfeld, S.K., Teichmann, S.A., and Luscombe, N.M. (2009). A census of human transcription factors: function, expression and evolution. Nat Rev Genet *10*, 252-263.

Vargas, C.L., Poursine-Laurent, J., Yang, L., and Yokoyama, W.M. (2011). Development of thymic NK cells from double negative 1 thymocyte precursors. Blood *118*, 3570-3578.

Vertegaal, A.C., Andersen, J.S., Ogg, S.C., Hay, R.T., Mann, M., and Lamond, A.I. (2006). Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics. Mol Cell Proteomics *5*, 2298-2310.

Verykokakis, M., Zook, E.C., and Kee, B.L. (2014). ID'ing innate and innate-like lymphoid cells. Immunol Rev *261*, 177-197.

Viatour, P., Merville, M.P., Bours, V., and Chariot, A. (2005). Phosphorylation of NF-kappaB and IkappaB proteins: implications in cancer and inflammation. Trends Biochem Sci *30*, 43-52.

Villen, J., Beausoleil, S.A., Gerber, S.A., and Gygi, S.P. (2007). Large-scale phosphorylation analysis of mouse liver. Proc Natl Acad Sci U S A *104*, 1488-1493.

Villen, J., and Gygi, S.P. (2008). The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. Nat Protoc *3*, 1630-1638.

Vivier, E., Raulet, D.H., Moretta, A., Caligiuri, M.A., Zitvogel, L., Lanier, L.L., Yokoyama, W.M., and Ugolini, S. (2011). Innate or adaptive immunity? The example of natural killer cells. Science *331*, 44-49.

Vivier, E., Tomasello, E., Baratin, M., Walzer, T., and Ugolini, S. (2008). Functions of natural killer cells. Nat Immunol *9*, 503-510.

Vonarbourg, C., Mortha, A., Bui, V.L., Hernandez, P.P., Kiss, E.A., Hoyler, T., Flach, M., Bengsch, B., Thimme, R., Holscher, C.*, et al.* (2010). Regulated expression of nuclear receptor RORgammat confers distinct functional fates to NK cell receptor-expressing RORgammat(+) innate lymphocytes. Immunity *33*, 736-751.

Vosshenrich, C.A., and Di Santo, J.P. (2013). Developmental programming of natural killer and innate lymphoid cells. Curr Opin Immunol *25*, 130-138.

Vosshenrich, C.A., Garcia-Ojeda, M.E., Samson-Villeger, S.I., Pasqualetto, V., Enault, L., Richard-Le Goff, O., Corcuff, E., Guy-Grand, D., Rocha, B., Cumano, A.*, et al.* (2006). A thymic pathway of mouse natural killer cell development characterized by expression of GATA-3 and CD127. Nat Immunol *7*, 1217-1224.

Walker, J.A., Barlow, J.L., and McKenzie, A.N. (2013). Innate lymphoid cells--how did we miss them? Nat Rev Immunol *13*, 75-87.

Wallace, A.D., Wheeler, T.T., and Young, D.A. (1997). Inducibility of E4BP4 suggests a novel mechanism of negative gene regulation by glucocorticoids. Biochem Biophys Res Commun *232*, 403-406.

Walzer, T., Dalod, M., Robbins, S.H., Zitvogel, L., and Vivier, E. (2005). Natural-killer cells and dendritic cells: "l'union fait la force". Blood *106*, 2252-2258.

Wang, C.M., Brennan, V.C., Gutierrez, N.M., Wang, X., Wang, L., and Yang, W.H. (2013). SUMOylation of ATF3 alters its transcriptional activity on regulation of TP53 gene. J Cell Biochem *114*, 589-598.

Wang, J.H., Nichogiannopoulou, A., Wu, L., Sun, L., Sharpe, A.H., Bigby, M., and Georgopoulos, K. (1996). Selective defects in the development of the fetal and adult lymphoid system in mice with an Ikaros null mutation. Immunity *5*, 537-549.

Warren, H.S., and Smyth, M.J. (1999). NK cells and apoptosis. Immunol Cell Biol *77*, 64-75.

Warren, L.A., and Rothenberg, E.V. (2003). Regulatory coding of lymphoid lineage choice by hematopoietic transcription factors. Curr Opin Immunol *15*, 166-175.

Wei, W., Yang, P., Pang, J., Zhang, S., Wang, Y., Wang, M.H., Dong, Z., She, J.X., and Wang, C.Y. (2008). A stress-dependent SUMO4 sumoylation of its substrate proteins. Biochem Biophys Res Commun *375*, 454-459.

Weintz, G., Olsen, J.V., Fruhauf, K., Niedzielska, M., Amit, I., Jantsch, J., Mages, J., Frech, C., Dolken, L., Mann, M.*, et al.* (2010). The phosphoproteome of toll-like receptor-activated macrophages. Mol Syst Biol *6*, 371.

Weiss-Steider, B., Soto-Cruz, I., Martinez-Campos, C.A., and Mendoza-Rincon, J.F. (2011). Expression of MICA, MICB and NKG2D in human leukemic myelomonocytic and cervical cancer cells. J Exp Clin Cancer Res *30*, 37.

Whitmarsh, A.J., and Davis, R.J. (2000). Regulation of transcription factor function by phosphorylation. Cell Mol Life Sci *57*, 1172-1183.

Wilkinson, K.A., and Henley, J.M. (2010). Mechanisms, regulation and consequences of protein SUMOylation. Biochem J *428*, 133-145.

Witze, E.S., Old, W.M., Resing, K.A., and Ahn, N.G. (2007). Mapping protein post-translational modifications with mass spectrometry. Nat Methods *4*, 798-806.

Woll, P.S., Martin, C.H., Miller, J.S., and Kaufman, D.S. (2005). Human embryonic stem cell-derived NK cells acquire functional receptors and cytolytic activity. J Immunol *175*, 5095-5103.

Wu, H., and Liao, P. (2007). Analysis of Protein Phosphorylation Using Mass Spectrometry. Chang Gung Med J *31*, 11.

Wu, J., and Lanier, L.L. (2003). Natural killer cells and cancer. Adv Cancer Res *90*, 127-156.

Wu, S.Y., and Chiang, C.M. (2009). Crosstalk between sumoylation and acetylation regulates p53-dependent chromatin transcription and DNA binding. EMBO J *28*, 1246-1259.

Wu, Y., DiMaggio, P.A., Jr., Perlman, D.H., Zakian, V.A., and Garcia, B.A. (2013). Novel phosphorylation sites in the S. cerevisiae Cdc13 protein reveal new targets for telomere length regulation. J Proteome Res *12*, 316-327.

Xu, J., He, Y., Qiang, B., Yuan, J., Peng, X., and Pan, X.M. (2008). A novel method for high accuracy sumoylation site prediction from protein sequences. BMC Bioinformatics *9*, 8.

Xue, Y., Liu, Z., Cao, J., Ma, Q., Gao, X., Wang, Q., Jin, C., Zhou, Y., Wen, L., and Ren, J. (2011). GPS 2.1: enhanced prediction of kinase-specific phosphorylation sites with an algorithm of motif length selection. Protein Eng Des Sel *24*, 255-260.

Xue, Y., Zhou, F., Fu, C., Xu, Y., and Yao, X. (2006). SUMOsp: a web server for sumoylation site prediction. Nucleic Acids Res *34*, W254-257.

Yagi, R., Zhong, C., Northrup, D.L., Yu, F., Bouladoux, N., Spencer, S., Hu, G., Barron, L., Sharma, S., Nakayama, T.*, et al.* (2014). The Transcription Factor GATA3 Is Critical for the Development of All IL-7Ralpha-Expressing Innate Lymphoid Cells. Immunity *40*, 378-388.

Yan, Q., Gong, L., Deng, M., Zhang, L., Sun, S., Liu, J., Ma, H., Yuan, D., Chen, P.C., Hu, X.*, et al.* (2010). Sumoylation activates the transcriptional activity of Pax-6, an important transcription factor for eye and brain development. Proc Natl Acad Sci U S A *107*, 21034-21039.

Yeh, P.Y., Lu, Y.S., Ou, D.L., and Cheng, A.L. (2011). IkappaB kinases increase Myc protein stability and enhance progression of breast cancer cells. Mol Cancer *10*, 53.

Yeung, J., O'Sullivan, E., Hubank, M., and Brady, H.J. (2004). E4BP4 expression is regulated by the t(17;19)-associated oncoprotein E2A-HLF in pro-B cells. Br J Haematol *125*, 560-567.

Yokoyama, W.M., Sojka, D.K., Peng, H., and Tian, Z. (2013). Tissue-resident natural killer cells. Cold Spring Harb Symp Quant Biol *78*, 149-156.

Yoshida, T., Ng, S.Y., Zuniga-Pflucker, J.C., and Georgopoulos, K. (2006). Early hematopoietic lineage restrictions directed by Ikaros. Nat Immunol *7*, 382-391.

Yu, J., Freud, A.G., and Caligiuri, M.A. (2013a). Location and cellular stages of natural killer cell development. Trends Immunol *34*, 573-582.

Yu, X., Rollins, D., Ruhn, K.A., Stubblefield, J.J., Green, C.B., Kashiwada, M., Rothman, P.B., Takahashi, J.S., and Hooper, L.V. (2013b). TH17 cell differentiation is regulated by the circadian clock. Science *342*, 727-730.

Yu, Y., Yoon, S.O., Poulogiannis, G., Yang, Q., Ma, X.M., Villen, J., Kubica, N., Hoffman, G.R., Cantley, L.C., Gygi, S.P*., et al.* (2011). Phosphoproteomic analysis identifies Grb10 as an mTORC1 substrate that negatively regulates insulin signaling. Science *332*, 1322-1326.

Zhang, H., and Ge, Y. (2011). Comprehensive analysis of protein modifications by top-down mass spectrometry. Circ Cardiovasc Genet *4*, 711.

Zhang, L.J., Vogel, W.K., Liu, X., Topark-Ngarm, A., Arbogast, B.L., Maier, C.S., Filtz, T.M., and Leid, M. (2012). Coordinated regulation of transcription factor Bcl11b activity in thymocytes by the mitogen-activated protein kinase (MAPK) pathways and protein sumoylation. J Biol Chem *287*, 26971-26988.

Zhang, W., Zhang, J., Kornuc, M., Kwan, K., Frank, R., and Nimer, S.D. (1995). Molecular-cloning and characterization of NF-IL3A, a transcriptional activator of the human interleukin-3 promoter. Mol Cell Biol *15*, 6055-6063.

Zhang, Y., Wallace, D.L., de Lara, C.M., Ghattas, H., Asquith, B., Worth, A., Griffin, G.E., Taylor, G.P., Tough, D.F., Beverley, P.C*., et al.* (2007). In vivo kinetics of human natural killer cells: the effects of ageing and acute and chronic viral infection. Immunology *121*, 258-265.

Zhao, M., Liu, Q., Liang, G., Wang, L., Luo, S., Tang, Q., Zhao, H., Su, Y., Yung, S., Chan, T.M*., et al.* (2013). E4BP4 overexpression: a protective mechanism in CD4+ T cells from SLE patients. J Autoimmun *41*, 152-160.