Investigating the role of NF-κB p50 serine 80 phosphorylation in the regulation of inflammation

Emma Louise Smith
BSc (Hons)

Submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

College of Medical, Veterinary, and Life Sciences
Institute of Infection, Immunity and Inflammation
University of Glasgow

September 2018

Abstract

NF-κB is a key transcription factor involved in the regulation of inflammation. The transcriptional activity of NF-κB is regulated by a number of post-translational modifications, including phosphorylation. Phosphorylation of NF-κB subunits may regulate transcription in a gene selective manner. The NF-κB p50 subunit is an important dual regulator of inflammatory responses, which can either promote or repress gene expression depending on the formation of heterodimer or homodimer complexes, respectively. Although p50 is a critical regulator of the immune system, phosphorylation of this subunit is largely understudied. In this thesis, the role of phosphorylation of NF-κB p50 on serine 80 (S80) in regulating transcriptional responses is investigated, using two human *NFKB1*^{S80A} knock-in cell lines generated by CRISPR/Cas9 genome editing techniques. Transcriptomic analyses reveal a critical role for S80 in selectively regulating the expression of a subset of NF-κB target genes in response to TNFα and LPS, including pro-inflammatory chemokines and cytokines. DNA binding analyses demonstrate that S80 regulates the binding of p50 to NF-κB binding sites in a sequence-specific manner. Specifically, phosphorylation of S80 reduces the affinity of p50 for κB sites that have an adenine at the -1 position. These data indicate that p50 S80 phosphorylation predominantly regulates transcription through the p50:p65 heterodimer, where S80 phosphorylation acts in *trans* to limit the NF-κB mediated transcription of pro-inflammatory genes. This advances our understanding of the regulation of transcriptional programmes by the NF-κB p50 subunit.

Table of Contents

Abstract 2
List of figures 6
List of tables 8
Acknowledgements 9
Author's declaration 10
Abbreviations 11
Amino acid codes 19
Nucleotide codes 19
1 Introduction 21
1.1 General introduction 21
1.2 Components of the NF-κB pathway 21
1.2.1 NF-κB 21
1.2.2 IκB 25
1.2.3 IKK complex 29
1.3 Signalling to NF-κB 30
1.3.1 NF-κB activation overview 30
1.3.2 NF-κB activation by specific receptors 32
1.4 NF-κB p50 subunit 38
1.4.1 Importance of *NFKB1* 38
1.4.2 p105 processing 40
1.4.3 Regulation of NF-κB p50 activity 42
1.5 Thesis aims 48
2 General materials and methods 50
2.1 Materials 50
2.1.1 Antibodies 50
2.1.2 Reagents 51
2.1.3 Buffers 51
2.1.4 Plasmids 53
2.2 Methods 53
2.2.1 Cell culture 53
2.2.2 Protein analysis 55
2.2.3 Functional assays 57
2.2.4 CRISPR/Cas9 genome editing 60
2.2.5 Transcriptomic analysis 63
2.2.6 Bioinformatics 64
2.2.7 Molecular biology techniques 66

3 Generation of knock-in *NFKB1*^{S80A} cell lines using CRISPR/Cas9 genome editing techniques 70
3.1 Abstract 70
3.2 Introduction 71
3.3 Results 75
3.3.1 *NFKB1*^{S80A} knock-in genome editing strategy 75
3.3.2 CRISPR/Cas9 editing in HEK293T cells 77
3.3.3 CRISPR/Cas9 editing in THP-1 cells 84
3.4 Discussion 88
4 The regulation of sequence specific NF-κB DNA binding and transcription by IKKβ phosphorylation of NF-κB p50 at serine 80 90
4.1 Abstract 90
4.2 Introduction 91
4.3 Materials and methods 94
4.3.1 Cell culture and transfection 94
4.3.2 Western blot analysis and immunoprecipitation 94
4.3.3 Kinase assays 95
4.3.4 Site directed mutagenesis and GST protein purification 96
4.3.5 CRISPR/Cas9 genome editing 96
4.3.6 Transcriptomic and bioinformatic analysis 97
4.3.7 Motif analysis 98
4.3.8 DNA affinity binding assay (DAPA) 98
4.3.9 Luciferase assays 99
4.3.10 Limited Proteolysis 99
4.4 Results 100
4.4.1 IKKβ phosphorylates NF-κB p50 100
4.4.2 Serine 80 is the IKKβ phosphorylation site of NF-κB p50 102
4.4.3 S80 phosphorylation is not required for NF-κB activation 104
4.4.4 S80 phosphorylation selectively regulates TNFα-induced gene expression 106
4.4.5 Specific DNA-binding motifs associated with differential regulation of NF-κB target genes by S80 108
4.4.6 S80 phosphorylation reduces p50 affinity for -1A containing κB sites 109
4.4.7 S80 phosphorylation inhibits transcription from -1A containing κB sites 111
4.5 Discussion 114
4.6 Supplementary data 117
5 *NFKB1*^{S80A} THP-1 profiling analyses 119
5.1 Abstract 119

5.2 Introduction 120
5.3 Results 123
5.3.1 NF-κB activation unaffected despite reduced p105/p50 protein levels in *NFKB1*^{S80A} THP-1 cells 123
5.3.2 ERK1/2 activation is increased in *NFKB1*^{S80A} THP-1 cells 124
5.3.3 Transcriptomic analysis of *NFKB1*^{S80A} THP-1 cells 127
5.3.4 Transcription factor and binding site enrichment analyses 131
5.3.5 Gene set enrichment analysis (GSEA) 138
5.3.6 κB site-based prediction of gene expression 143
5.4 Discussion 145
6 General discussion 150
7 Appendix 158
7.1 CRISPR/Cas9 targeting of *NFKB1* 158
7.2 Clustering analysis of differentially expressed THP-1 genes 159
7.3 THP-1 clustering analysis 160
7.4 GSEA enrichment plots 164
7.5 Upstream regulator analysis of untreated THP-1 cells 165
References 166

List of figures

Figure 1.1: NF-κB family of transcription factors..... 23

Figure 1.2: IκB protein family members 25

Figure 1.3: Canonical and non-canonical NF-κB signalling 31

Figure 1.4: MYD88 dependent and MYD88 independent TLR4 signalling 35

Figure 1.5: TNFα-induced TNFR1 signalling 37

Figure 1.6: Phosphorylation of NFKB1 45

Figure 3.1: Overview of CRISPR/Cas9 genome editing..... 73

Figure 3.2: CRISPR/Cas9 editing strategy..... 76

Figure 3.3: Cas9/gRNA CRISPR plasmid cloning and validation 78

Figure 3.4: ssODN validation and HDR optimisation..... 80

Figure 3.5: Screening and identification of *NFKB1*^{S80A} positive HEK293T cells ... 82

Figure 3.6: Off-target site analysis of knock-in *NFKB1*^{S80A} HEK293T cells 83

Figure 3.7: Screening and identification of *NFKB1*^{S80A} positive THP-1 cells..... 86

Figure 3.8: Off-target site analysis of knock-in *NFKB1*^{S80A} THP-1 cells 87

Figure 4.1: IKKβ phosphorylates NF-κB p50. 101

Figure 4.2: IKKβ phosphorylates NF-κB p50 at S80. 103

Figure 4.3: S80 phosphorylation is not required for NF-κB activation. 105

Figure 4.4: S80 phosphorylation selectively regulates TNFα-induced gene expression. 107

Figure 4.5: Specific κB sites are enriched in genes selectively regulated by S80 phosphorylation.....109

Figure 4.6: S80 phosphorylation regulates DNA binding affinity to -1A containing κB sites.....111

Figure 4.7: S80 phosphorylation regulates transcription in a κB sequence specific manner.....113

Figure 5.1: S80 phosphorylation is not required for NF-κB activation in THP-1 cells 124

Figure 5.2: p105/p50 protein levels and ERK1/2 activation altered in differentiated *NFKB1*^{S80A} THP-1 cells.....126

Figure 5.3: p50 S80 phosphorylation selectively regulates gene expression in THP-1s.....128

Figure 5.4: Upstream regulator enrichment analysis of LPS-induced genes.....133

Figure 5.5: Transcription factor binding site analysis of LPS-induced upregulated genes.....136

Figure 5.6: Transcription factor binding site analysis of LPS-induced downregulated genes.....137

Figure 5.7: Enriched hallmark gene sets in LPS-treated WT vs *NFKB1*^{S80A} THP-1 cells.....139

Figure 5.8: Enriched KEGG pathways in LPS-treated WT vs *NFKB1*^{S80A} THP-1 cells.....141

Figure 5.9: -A(-2)A(-1) κB site enrichment in GSEA refined gene sets.....142

Figure 5.10: Prediction of p50 S80 regulated gene expression based on κB site sequence.....144

Figure 7.1: CRISPR/Cas9 targeting of *NFKB1*.....158

Figure 7.2: Clustering analysis of differentially expressed genes in THP-1 cells159

Figure 7.3: GSEA enrichment plots: LPS-treated WT vs *NFKB1*^{S80A}.....164

Figure 7.4: Upstream regulator enrichment analysis of untreated THP-1 cells.165

Supplementary Figure 1: Transcription factor binding site shape analysis.....117

Supplementary Figure 2: S80 affects proteolysis of p50.....117

List of tables

Table 2.1: Primary antibodies 50

Table 2.2: Secondary antibodies 50

Table 2.3: Agarose gel electrophoresis buffer 51

Table 2.4: Western blot electrophoresis buffers 51

Table 2.5: Agarose gel recipe 51

Table 2.6: SDS-PAGE gel recipes 51

Table 2.7: SDS-PAGE sample buffers 52

Table 2.8: RIPA buffer 52

Table 2.9: DAPA buffer 52

Table 2.10: Plasmids sources 53

Table 2.11: Plating conditions for endogenous assays 54

Table 2.12: Transfection conditions for functional assays 54

Table 2.13: Transfection conditions for CRISPR/Cas9 genome editing 55

Table 2.14: DAPA oligonucleotides 58

Table 2.15: κB site repeat oligonucleotides 59

Table 2.16: Oligonucleotides used for CRISPR/Cas9 genome editing 61

Table 2.17: QuantiTect primers used for qPCR 64

Table 7.1: Groups of upregulated genes utilised for transcription factor binding site analysis using HOMER 162

Table 7.2: Groups of downregulated genes utilised for transcription factor binding site analysis using HOMER 163

Acknowledgements

Firstly, I thank my supervisor Dr Ruaidhrí Carmody. I am extremely grateful for your continuous guidance and support throughout my PhD. Thank you for dedicating time to my work, and for your encouragement.

I thank my fellow Carmody lab hens, past and present: Trish, Jenny, Izaskun, Felix, Dom and Dave. Thank you for all your help, and most importantly for your friendship inside and outside the lab. To all my GBRC companions, thanks for the laughs, and particularly for some memorable socials.

I thank Dave for being the CRISPR king and Dom for teaching me how to be a hacker, guiding me through the matrix. Your support has helped to shape this thesis. A big thanks to Jenny, for waving at me through the window of my mind palace, yielding snacks and showing me GIFs to motivate me through 12 hour thesis writing shifts. To Trish, my “how to science” role model, thank you for all of your advice and guidance through PhD life. Thank you for listening to my rants during our tissue culture debriefs. But most important of all, I can now re-assemble cardboard boxes.

Finally, I could not have completed my PhD without the continuous love and support of my Mother and Father (team base camp). Thank you for always believing in me, even when I did not believe in myself. I have climbed the snail!

Author's declaration

I declare that, except where reference is made to the contribution of others, that this thesis is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Signature:

Printed name:

Abbreviations

Ab	Antibody
AIM2	Absent in melanoma 2
ANK	Ankyrin
AP-1	Activator protein 1
ATP	Adenosine triphosphate
βTrCP	Beta-transducin repeat containing
BAFF	B-cell activating factor
BCL-3	B-cell lymphoma-3
BCR	B-cell receptor
BMDM	Bone marrow derived macrophage
BSA	Bovine serum albumin
Bp	Base pair
CCL	C-C chemokine ligand
CCR	C-C chemokine receptor
CD40	Cluster of differentiation 40
cGAS/STING	Cyclic GMP-AMP synthase stimulator of IFN genes
ChIP	Chromatin immunoprecipitation
ChIP-seq	ChIP-sequencing

Chk1	Checkpoint protein kinase 1
cIAP	Cellular inhibitor of apoptosis protein
COMMD1	Copper metabolism mouse U2af1-rs1 region 1-domain-containing protein 1
COX2	Cyclo-oxygenase
CRD	Cysteine-rich domain
CRISPR	Clusters of regularly interspaced short palindromic repeats
crRNAs	CRISPR RNAs
CSF1	Colony stimulating factor 1
C-terminal	Carboxy-terminal
CXCL	Chemokine C-X-C motif ligand
DAPA	DNA affinity purification assay
DD	Death domain
DMEM	Dulbecco's modified eagle medium
DNA-PK	DNA-dependent protein kinase
DSB	Double-strand break
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ELK1	ETS domain-containing protein 1
ERK	Extracellular signal-regulated kinases

ES	Enrichment score
FBS	Fetal bovine serum
FIMO	Find individual motif occurrences
Fmoc	(N-(9-flurenyl)methoxycarbonyl)
FPKM	Fragments per kilobase million
GCN5	General control of amino acid synthesis protein 5
GFI1B	Growth factor independent 1B transcriptional repressor
G-MSCF	Granulocyte-macrophage colony-stimulating factor
gRNA	Guide RNA
GRR	Glycine rich region
GSEA	Gene set enrichment analysis
GSH	Glutathione
GSK3β	Glycogen synthase kinase 3 beta
GST	Glutathione S-transferases
HCC	Hepatocellular carcinoma
HDAC-1	Histone deacetylase 1
HDR	Homology directed repair
HEK293T	Human embryonic kidney 293T
HRP	Horseradish peroxidase

ICAM1	Intercellular adhesion molecule 1
IFN	Interferon
IgG	Immunoglobulin G
IκB	Inhibitor of κB
IKK	IκB kinase
IL	Interleukin
IL-1R	Interleukin-1 receptor
Indel	Insertion or deletion mutation
IP	Immunoprecipitation
IPTG	Isopropyl ß-D-1-thiogalactopyranoside
IRAK	Interleukin-1 receptor-associated kinase
IRE	Interferon responsive elements
IRF	Interferon regulatory transcription factor
ISRE	Interferon-sensitive response elements
JNK	c-Jun N-terminal kinase
kDa	Kilodalton
KEGG	Kyoto encyclopedia of genes and genomes
KPC	Kip1 ubiquitylation-promoting complex
LB	Lysogeny broth

LIF	Leukemia inhibitory factor
LPS	Lipopolysaccharide
LTβR	Lymphotoxin β receptor
LZ	Leucine zipper
MAP3K	MAP kinase kinase kinase
MAPK	Mitogen activated protein kinase
M-CSF	Macrophage colony-stimulating factor
MEK	Mitogen/extracellular signal-regulated Kinase
MHC	Major histocompatibility complex
MSigDB	Molecular signatures database
MYD88	Myeloid differentiation primary-response protein 88
NEMO	NF-κB essential modulator
NES	Nuclear export sequence
NF-κB	Nuclear factor-kappa B
NHEJ	Non-homologous end joining
NIK	NF-κB-inducing kinase
NLS	Nuclear localisation sequence
N-terminal	Amino-terminal
PAGE	Polyacrylamide gel electrophoresis

PAM	Protospacer-associated motif
PAMP	Pathogen-associated molecular patterns
PBM	Protein-binding microarray
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate buffer saline
PBS-T	Phosphate buffer saline-tween
PCA	Principal component analysis
PCR	Polymerase chain reaction
PEST	Proline, glutamic acid, serine, threonine
PKA	Protein kinase A
PMA	Phorbol 12-myristate 13-acetate
PMSF	Phenylmethylsulfonyl fluoride
PTM	Post-translational modification
PWM	Position weight matrix
qPCR	Quantitative PCR
RHD	Rel homology domain
RIP	Receptor-interacting protein
RIPA	Radioimmunoprecipitation assay buffer
RNA-seq	RNA sequencing

RNP	Ribonucleoprotein
RPMI	Roswell park memorial institute
SARM	Sterile-α- and armadillo motif-containing protein
SDS	Sodium dodecyl sulphate
SEM	Standard error of the mean
sgRNA	Synthetic guide RNA
SOCS1	Suppressor of cytokine signalling 1
SP1	Specificity protein 1
ssODN	Single-stranded oligonucleotide
TAB	TAK1-binding proteins
TAD	Transactivation domain
TAK1	Transforming growth factor-ß-activated kinase-1
TALEN	Transcription activator-like effector nuclease
TBE	Tris-Borate-EDTA
TBK	TANK-binding kinase
TBP	TATA-binding protein
TCR	T-cell receptor
TEMED	Tetramethylethylenediamine
TFBS	Transcription factor binding site

Th	T helper
TIR	Toll/interleukin-1 receptor
TIRAP	TIR -domain-containing adapter protein
TLR	Toll-like receptor
TNF	Tumour necrosis factor
TNFR	Tumour necrosis factor receptor
TPL-2	Tumour progression locus 2
tracrRNA	Transactivating CRISPR RNA
TRADD	TNF receptor-associated death domain
TRAF	Tumour necrosis factor receptor–associated factor
TRAM	TRIF-related adaptor molecule
TRIF	TIR-domain-containing adaptor protein inducing IFN beta
TRRUST	Transcriptional regulatory relationships unravelled by sentence-based text mining
Ub	Ubiquitin
UBC	Ubiquitin-conjugating enzyme
UPGMA	Unweighted pair group method with arithmetic mean
UV	Ultra violet
VCAM1	Vascular cell adhesion protein 1
vD3	1α, 25-dihydroxyvitamin D3

w/v	Weight per volume
w/w	Weight per weight
WB	Western blot
WT	Wild-type
ZFN	Zinc finger nuclease
ZFP161	Zinc finger protein 161

Amino acid codes

A	Alanine	P	Proline
D	Aspartate	S	Serine
E	Glutamic Acid	T	Threonine

Nucleotide codes

A	Adenine	N	Any base
C	Cytosine	R	A or G
G	Guanine	W	A or T
T	Thymine	Y	C or T

Chapter 1

1 Introduction

1.1 General introduction

Nuclear factor-κB (NF-κB) proteins are a family of structurally related transcription factors that regulate the expression of over 500 genes involved in many important biological processes including the regulation of cell cycle, proliferation, differentiation and cell death. However, the most fundamental role that NF-κB plays is to regulate inflammation. The importance of NF-κB in these physiological processes is evident from its dysregulation in a number of disease states such as chronic inflammation, immunodeficiency and cancer (Courtois and Gilmore, 2006, Hayden and Ghosh, 2012). Due to its ability to regulate such a large number of genes, the activity of NF-κB is tightly regulated at multiple levels, including mechanisms of gene specific transcriptional regulation by NF-κB that ensure specificity of gene expression.

1.2 Components of the NF-κB pathway

1.2.1 NF-κB

The NF-κB transcription factor family is composed of five subunits: RelA (p65), RelB, c-Rel, p50 and p52. The p50 and p52 subunits are generated from the limited proteasomal processing of the precursor proteins p105 and p100, respectively, while p65, RelB and c-Rel are translated directly. The p50 and p52 subunits lack the transactivation domain (TAD) found in the C-terminal regions of the RelA, c-Rel and RelB subunits, while all NF-κB subunits contain a highly conserved Rel homology domain (RHD) (Figure 1.1A). The RHD facilitates DNA binding, dimerisation, nuclear localisation, and interaction with inhibitor of κB (IκB) proteins (Oeckinghaus and Ghosh, 2009).

The NF-κB subunits can hetero- and homo-dimerise to form dimer complexes (Figure 1.1B). While theoretically 15 combinations are possible, only 12 dimer complexes have experimentally identified *in vivo* (Huxford and Ghosh, 2009). NF-κB can promote or repress transcription depending on the subunit composition of dimer complexes. Although the p50 and p52 subunits lack a TAD, they can positively regulate transcription by forming a heterodimer with TAD

containing subunits p65, RelB and c-Rel. Alternatively, p50 and p52 homodimers function as transcriptional repressors by competing with TAD containing NF-κB dimers for the same DNA binding sites in target gene promoters (Carmody et al., 2007). While the abundance of particular NF-κB dimers is cell type-specific, the p50:p65 heterodimer represents the most abundant dimer complex, and is expressed in most cell types (Oeckinghaus and Ghosh, 2009). Various studies using knockout cell lines and knockout mouse models have elucidated both unique and overlapping biological functions of NF-κB dimers (Hoffmann et al., 2003, Gerondakis et al., 2006). For example, expression of NF-κB target gene leukemia inhibitory factor (*Lif)* was abolished in *Nfkb1*$^{-/-}$ and *Rela*$^{-/-}$ single knockout cells, and was not expressed by c-Rel containing dimers, suggesting that only the p50:p65 heterodimer is functional on this promoter (Hoffmann et al., 2003). In contrast, expression of chemokine C-X-C motif ligand 10 (*Cxcl10*), colony stimulating factor 1 (*Csf1*) and C-C chemokine ligand 5 (*Ccl5*) requires a dimer containing p65, with either p50 or p52. Furthermore, double NF-κB subunit knockout mice have more severe phenotypes compared to single subunit knockouts (Gerondakis et al., 2006).

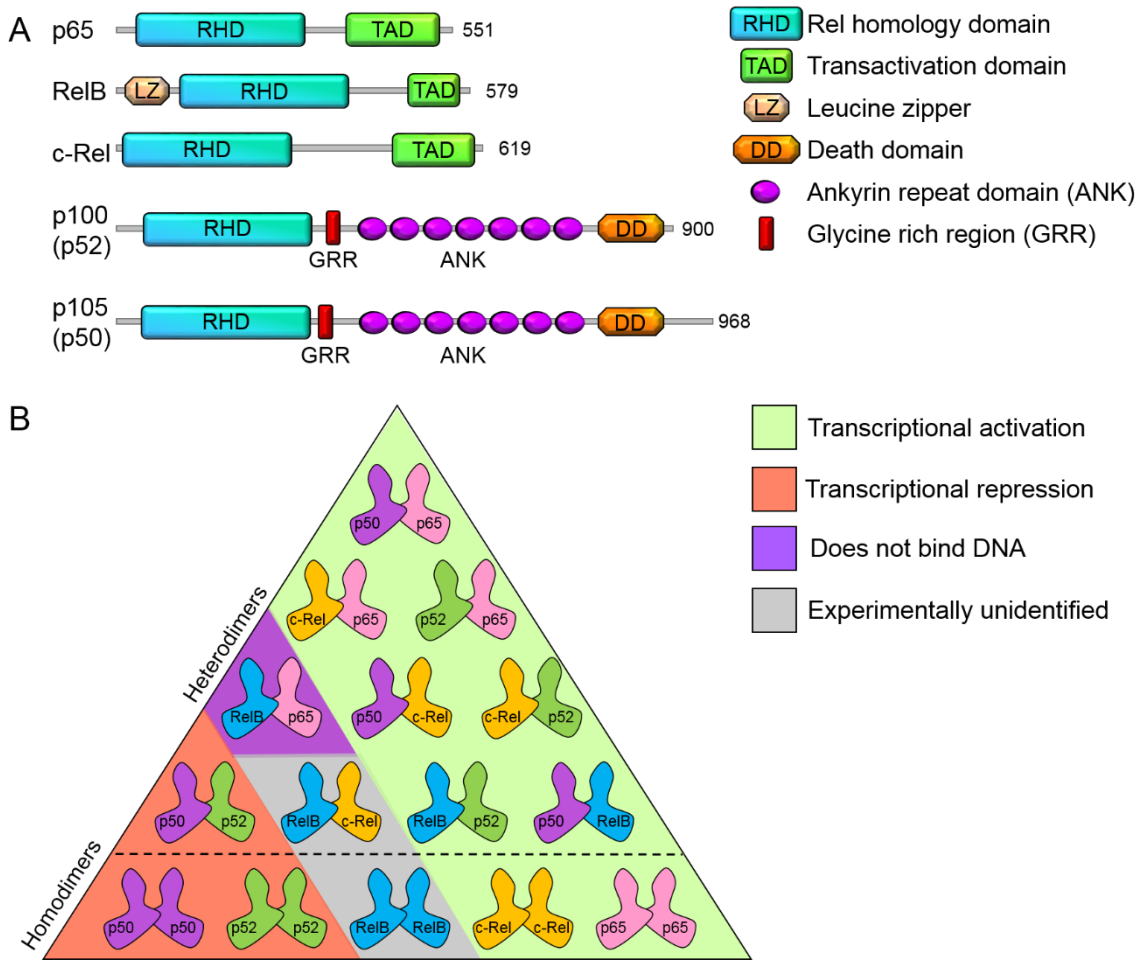


Figure 1.1: NF-κB family of transcription factors
(A) Schematic representation of the five members of the NF-κB transcription factor family: p65, RelB, c-Rel, p100/p52 and p105/p50. The principal structural domains are indicated and labelled. Amino acid residue numbering corresponds to human protein sequence. (B) 15 possible homo- and heterodimers formed by different combinations of NF-κB subunits.

NF-κB dimers bind to specific DNA sequences called κB sites in the enhancers and promoters of target genes. Most κB sites are 10 base pairs (bp) in length, with the consensus sequence 5'-$G^{-5}G^{-4}G^{-3}R^{-2}N^{-1}W^{0}Y^{+1}Y^{+2}C^{+3}C^{+4}$-3' (R represents a purine, N represents any nucleic acid, W represents an A or T and Y represents a pyrimidine). A number of structural studies have revealed that p50 and p52 bind to the 5bp 5′-GGGRN-3′ half-site of the consensus sequence, while p65, RelB and c-Rel bind to the 5′-YYCC-3′ half-site (Chen et al., 1998, Cramer et al., 1997, Ghosh et al., 1995, Huang et al., 2001). These studies suggested that NF-κB dimers follow a rule of cognate half-site recognition, separated by a 1bp spacer: p50 and p52 homodimers bind to 11bp κB site consisting of two 5bp half-sites, whereas p65 and c-Rel homodimers bind to 9bp sites consisting of two 4bp half-sites. Heterodimers on the other hand, bind to a 10bp κB site, consisting of a

5bp and 4bp half-site. DNA binding analyses utilising electrophoretic mobility shift assays and large-scale protein-binding microarrays (PBMs) have confirmed the 3 distinct classes of NF-κB DNA binding-specificity (Kunsch et al., 1992, Siggers et al., 2011). Despite this classification, NF-κB dimers have overlapping DNA binding specificities. For example, p65 and c-Rel bind to the same class of κB site, however p65 and c-Rel can elicit distinct biological functions (Hoffmann et al., 2006). By measuring DNA dissociation rates, Siggers et al. (2011) showed that c-Rel homodimers have much higher affinity to specific κB sites in comparison to p65 homodimers, contributing to the specific regulation of c-Rel-dependent functions. In addition, the NF-κB p50:p65 heterodimer has been shown to bind to a consensus κB site with an affinity of ~10nM, whereas p50 and p65 homodimers bind to the same κB site with 5- and 15-fold lower affinity, respectively (Phelps et al., 2000). These studies suggest that κB site sequences are a pre-requisite for NF-κB-DNA binding specificity. However, many of the mechanisms which regulate NF-κB affinity to particular κB sites, contributing to specificity of target gene expression, remain unknown.

1.2.2 IκB

The primary mechanism by which NF-κB activity is regulated is through interaction with IκB proteins. In unstimulated cells, NF-κB dimers are sequestered in the cytoplasm by association with an IκB family member. The IκB family of proteins can be classified into three groups: typical, precursor and atypical (

Figure 1.2). All IκB proteins are characterised by the presence of ankyrin repeat domains, which interact with the RHD of NF-κB proteins. Much like the NF-κB protein family, the IκB proteins have different affinities for particular subsets of NF-κB dimers, are regulated in different manners, and are expressed in a tissue-specific manner (Hinz et al., 2012).

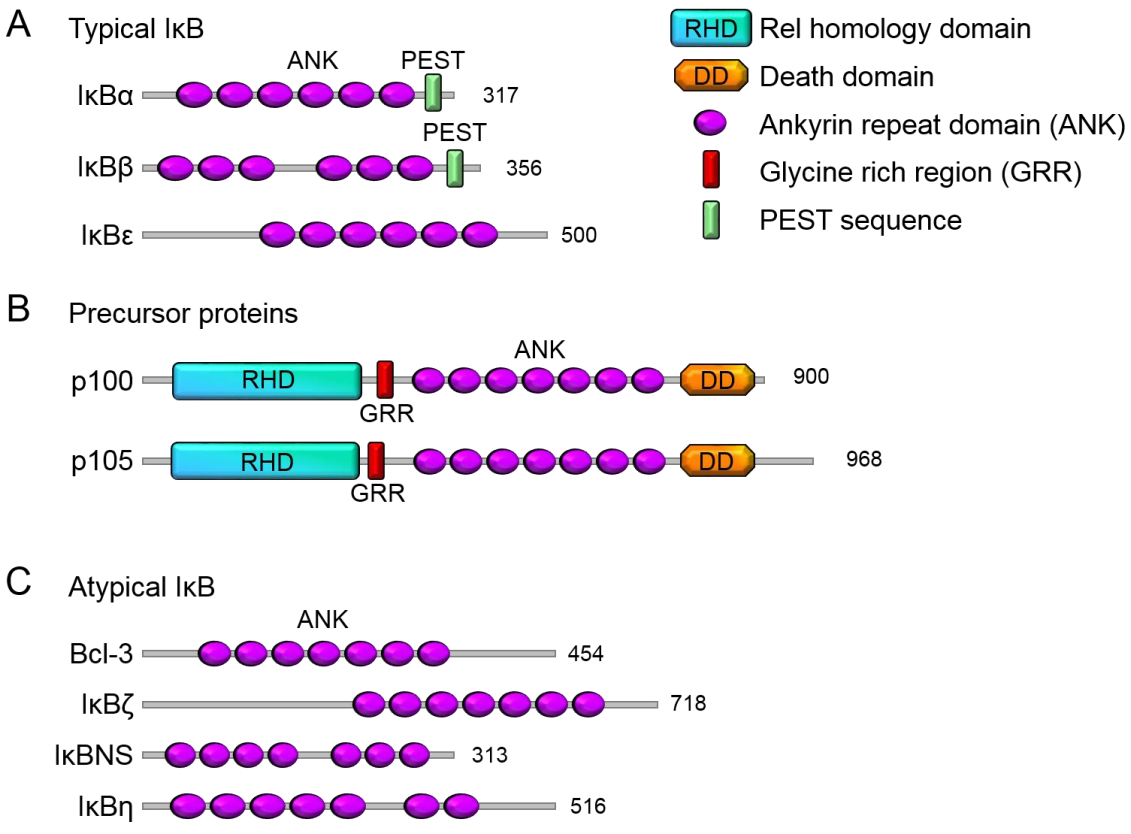


Figure 1.2: IκB protein family members
Schematic representation of the typical (A), precursor protein (B), and atypical (C) IκB proteins. The principal structural domains are indicated and labelled. Amino acid residue numbering corresponds to human protein sequence.

The typical or classical IκB proteins IκBα, IκBß and IκBε are present in the cytoplasm of unstimulated cells. All 3 classical IκB proteins share 6 ankyrin repeats, and a conserved signal response region containing phosphorylation sites required for stimulus-induced degradation. IκBα and IκBß share an acidic region, rich in proline (P), glutamic acid (E), serine (S) and threonine (T), also referred to as PEST sequences (Figure 1.2A). PEST sequences are implicated in regulating IκB protein half-life (Rogers et al., 1986, Rodriguez et al., 1995), and are thought to provide a structural mechanism for the divergent functions of IκBα and IκBß (Malek et al., 2003).

The typical IκB proteins preferentially interact with p65 and c-Rel containing NF-κB dimers, where IκBα has a higher affinity for p50:p65 heterodimers than for p65 homodimers (Malek et al., 2003). In fact, IκBα bound to p50:p65 was the first NF-κB:IκB complex to be discovered, and is the best characterised IκB protein. The IκBα:p50:p65 complex defines the general mechanism by which NF-κB dimers are inhibited (Müller and Harrison, 1995, Ghosh et al., 1995). Crystal structures of IκBα:p50:p65 revealed that IκBα masks the nuclear localisation sequence (NLS) of p65, but not p50 (Huxford et al., 1998, Jacobs and Harrison, 1998). The combination of the nuclear export sequence (NES) of IκBα and exposed p50 NLS results in shuttling of the NF-κB:IκB complex between the nucleus and cytoplasm (Johnson et al., 1999, Huang et al., 2000, Malek et al., 2001). However due to the strength of the IκBα NES, rate of nuclear export exceeds that of nuclear import, resulting in a steady-state cytoplasmic localisation (Ghosh and Karin, 2002, Huang et al., 2000). Rapid degradation of IκBα during NF-κB activation exposes the previously masked NLS of p65, facilitating nuclear translocation of the NF-κB dimer, where it can bind to specific κB sites. While IκB degradation is the primary mechanism for activation of NF-κB activity, the IκB proteins are also important for the termination of NF-κB transcriptional responses. This mechanism depends on the NF-κB dependent re-synthesis of IκB proteins. Shortly after its degradation, IκBα is rapidly re-synthesised and functions to restore the cell to its resting state by binding to NF-κB dimers in the nucleus, via a negative feedback loop (Arenzana-Seisdedos et al., 1995).

The roles of IκBβ and IκBε are less well understood. Although IκBβ has similar NF-κB dimer binding preferences to IκBα, in contrast, IκBβ masks the NLS of both NF-κB subunits. Therefore IκBβ:NF-κB complexes do not shuttle between the nucleus and the cytoplasm, and have a strict cytoplasmic localisation in unstimulated cells (Malek et al., 2001). IκBβ degradation and re-synthesis occur with delayed kinetics compared to IκBa. IκBε is lesser studied, however it is known that IκBε associates with p65 and c-Rel containing dimers. High levels of expression are seen in tissues such as spleen and lung, indicating that IκBε function appears to be limited to cells of the hematopoetic lineage (Hinz et al., 2012). In addition, like IκBβ, IκBε degradation and re-synthesis occurs at a much slower rate in comparison to IκBa. Differences in dimer preference, degradation and re-synthesis rates suggest that the typical IκB proteins have distinct roles in NF-κB inhibition.

The p100 and p105 proteins serve as both NF-κB subunit precursors, and also as IκB proteins. Both p100 and p105 undergo limited proteolysis, generating the p52 and p50 subunits, respectively. Processing can either occur constitutively, or be induced by phosphorylation by an IκB kinase (Christian et al., 2016). p100 and p105 also contain 7 ankyrin repeat domains in their C-terminus (Figure 1.2B). Therefore, p100 and p105 can bind to NF-κB in a both an NF-κB-like and IκB-like manner, via their RHD and ankyrin repeat domains, respectively. Not only does inducible processing of p100 and p105 generate NF-κB subunits, it also functions to liberate NF-κB dimers bound in this inhibitory complex, much like the typical IκB proteins. p105 is capable of binding to and inhibiting p50, p65 and c-Rel containing dimers, maintaining them in the cytoplasm (Rice et al., 1992, Naumann et al., 1993). While p100 is also capable of binding to and inhibiting many NF-κB dimers (Ishimaru et al., 2006), it is most well-known for its interaction with RelB. Inducible processing of p100 generates p52:RelB heterodimers, an important step in the non-canonical NF-κB activation pathway (Senftleben et al., 2001) (described in section 1.3.1). More recently, it was discovered that p100 and p105 also form large molecular weight complexes, defined by both NF-κB:RHD and NF-κB:ankyrin interactions which function to inhibit NF-κB activity (Savinova et al., 2009).

The atypical IκB members B-cell lymphoma 3 (BCL-3), IκBζ, IκBNS and IκBη contain 7 ankyrin repeats facilitating their ability to bind to NF-κB dimers (Figure 1.2C). The atypical IκB proteins are predominantly nuclear, and are generally not expressed in unstimulated cells with the exception of IκBη, which is constitutively expressed (Hinz et al., 2012). Upon induction, atypical IκB proteins function to modulate the activity of NF-κB dimers in the nucleus. BCL-3 was the first atypical IκB protein to be discovered, however its precise function as an NF-κB regulator remained unclear for many years. BCL-3 specifically associates with p50 and p52 homodimers, and is thought to be capable of both enhancing and repressing their transcriptional activity (Wulczyn et al., 1992, Bours et al., 1993, Fujita et al., 1993, Nolan et al., 1993). However, BCL-3 has been well characterised as an inhibitor of NF-κB, by stabilising repressive NF-κB p50 homodimers to DNA, preventing TAD containing dimers from binding (Carmody et al., 2007, Wessells et al., 2004) (described in section 1.4.3.1.1).

IκBζ shares structural and functional homology with BCL-3. IκBζ preferentially associates with and modulates the activity of p50 homodimers. IκBζ binds p50 homodimers at κB sites, most notably on the interleukin 6 (*Il6*) promoter. Expression of *Il6* is not induced in IκBζ-deficient cells, indicating that IκBζ is a co-activator of p50 homodimers (Yamamoto et al., 2004). In addition, IκBNS has been shown to preferentially associate with p50 (Fiorini et al., 2002). It is therefore assumed that IκBNS functions in a similar manner to BCL-3, stabilising p50 homodimers on DNA. However, negative regulation of NF-κB activity by IκBNS was shown to be selective since production of IL-6, but not tumour necrosis factor α (TNFα), was increased following stimulation of IκBNS knockout cells (Kuwata et al., 2006). More recently, IκBη was identified as an atypical IκB with high homology to IκBß (Yamauchi et al., 2010). IκBη associates with the p50 subunit of NF-κB, and was shown to be important for regulating a subset of pro-inflammatory genes. Knockdown of IκBη by siRNA suppressed the transcription of *Cxcl2*, *Il6* and *Il1b* in stimulated RAW264.7 cells, suggesting that IκBη might function as a co-activator of NF-κB gene expression (Yamauchi et al., 2010).

1.2.3 IKK complex

A key regulatory step in the NF-κB activation pathway is the phosphorylation of IκB proteins by the IκB kinase (IKK) complex. The IKK complex is a large protein structure (700-900 kDa) consisting of two kinase subunits IKKα and IKKß, and a regulatory subunit IKKγ (NEMO). IKKα and IKKß share 50% sequence identity, and contain a kinase domain, leucine zipper, and helix-loop-helix domain. These domains are responsible for phosphorylation of substrates, homo- or hetero-dimerisation of the kinases, and kinase activity, respectively (Israel, 2010). The IKK complex is the core component of NF-κB signalling, with most NF-κB activation signals converging on this enzymatic complex. The significance of the IKK complex has been demonstrated using mouse knockout models, whereby *Ikka*$^{-/-}$, *Ikkb*$^{-/-}$ and *Nemo*$^{-/-}$ mice all have lethal phenotypes (Gerondakis et al., 2006). Activation of the IKK complex is dependent on phosphorylation of serine residues (IKKα S176 and S180; IKKß S177 and S181) located in an activation loop within the kinase domain (Mercurio et al., 1997, Delhase et al., 1999). Following activation, the IKK complex is capable of phosphorylating a number of substrates, many of which are involved in the NF-κB pathway [as reviewed by (Chariot, 2009)]. However, the complex is best characterised for the phosphorylation of IκB proteins, targeting them for degradation in order to liberate NF-κB dimers (described in section 1.3). However, IKKα and IKKß have been shown to directly phosphorylate NF-κB subunits (Christian et al., 2016). IKKα and IKKß phosphorylate multiple serine residues on p100 and p105, required for processing to p52 and p50, respectively. IKKα and IKKß have also been implicated in phosphorylation of the p65 subunit, regulating NF-κB dependent gene expression (Christian et al., 2016). For example, serine 536 (S536) of the p65 subunit can be phosphorylated by both IKKα and IKKß, which have distinct functional outcomes. Phosphorylation of S536 by IKKß enhances p65 transactivation (Sakurai et al., 1999). In contrast, phosphorylation of the same residue by IKKα leads to ubiquitination of p65, targeting it for proteasomal degradation by the proteasome, therefore limiting p65 driven transcription (Lawrence et al., 2005). These studies indicate that the IKK complex has distinct roles in the regulation of the NF-κB pathway.

1.3 Signalling to NF-κB

1.3.1 NF-κB activation overview

NF-κB is activated by a wide variety of extracellular stimuli, involving several signalling pathways which converge on the IKK complex. Given the similarity between IKKα and IKKß, it was presumed that they would have overlapping functions. However, divergent roles have been outlined for each of the kinases, leading to the identification of two different NF-κB activation pathways: the canonical or classical pathway, and the non-canonical pathway (Figure 1.3). Each pathway is distinct with respect to the inducing stimuli, particular IKK subunit involved, and NF-κB/IκB substrates targeted.

The canonical or classical pathway is considered to be representative of the general mechanisms by which NF-κB is regulated. In this pathway, IKKß is activated by phosphorylation as a result of the engagement of various receptors such as Toll-like receptors (TLRs), tumour necrosis factor receptors (TNFRs), B-cell receptors (BCR) and T-cell receptors (TCR) (Hayden and Ghosh, 2012). Activation of IKKß leads to the specific phosphorylation of IκBα on S32 and S36, targeting IκBα for ubiquitination. In turn, ubiquitination leads to proteasomal degradation of IκBα, which frees sequestered NF-κB dimers such as the p50:p65 heterodimer, allowing translocation to the nucleus (Figure 1.3A). The non-canonical pathway is dependent on NF-κB-inducing kinase (NIK) induced activation of IKKα. NIK is activated by signals from receptors such as CD40, lymphotoxin ß receptor (LTßR) and B-cell activating factor (BAFF) receptor (Hayden and Ghosh, 2012). Activated NIK phosphorylates IKKα, and through a series of NIK/IKKα dependent phosphorylations, p100 is ubiquitinated and targeted for limited proteolysis. This generates the p52 subunit, which typically forms heterodimers with RelB. p52:RelB heterodimers are then able to translocate to the nucleus (Figure 1.3B).

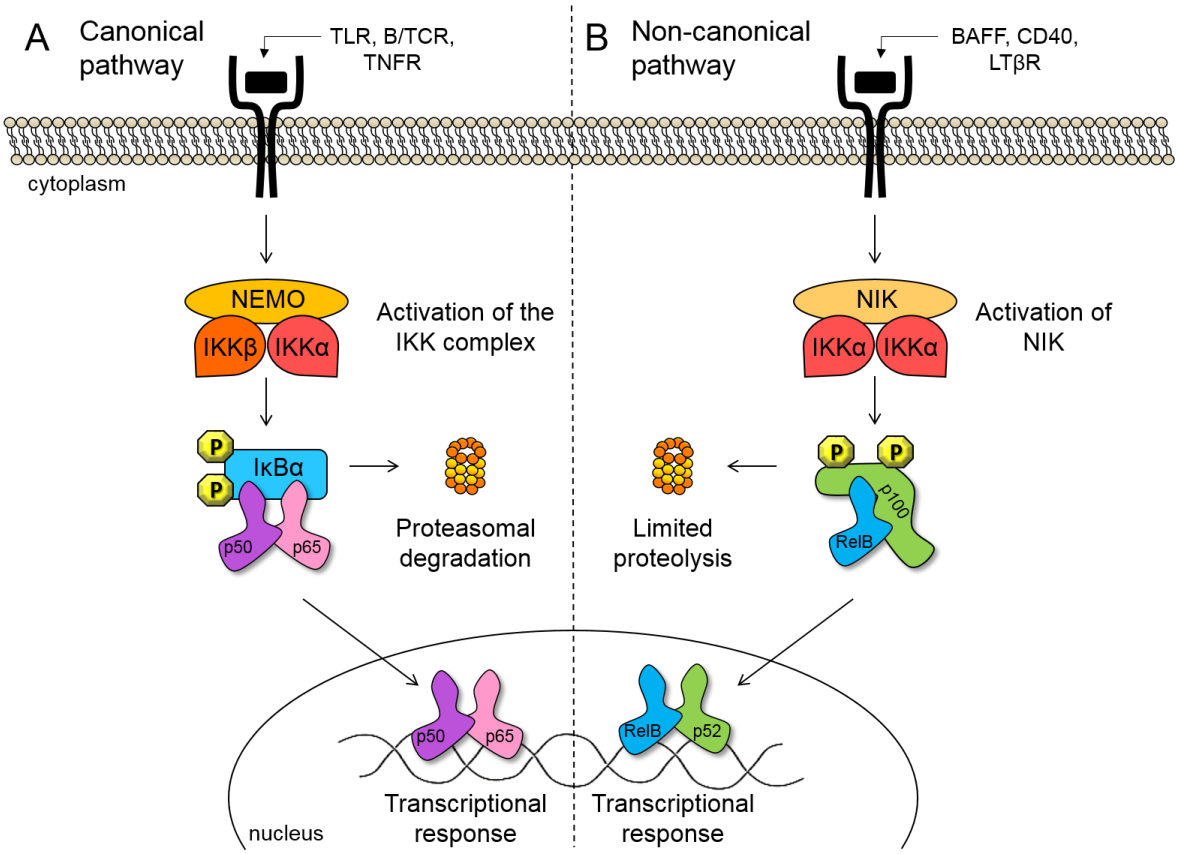


Figure 1.3: Canonical and non-canonical NF-κB signalling
(A) The canonical pathway is induced by a diverse range of stimuli, which activates a number of receptors such as Toll-like receptors (TLRs), B-cell receptors (BCR), T-cell receptors and tumour necrosis factor receptors (TNFRs). These receptors utilise a number of signalling adaptors to activate the IKK complex. Phosphorylation (yellow octagons) of specific serine residues of classical IκBs such as IκBα by IKKβ leads to IκB ubiquitination and subsequent proteosomal degradation. This results in the release of NF-κB dimers, which are free to translocate to the nucleus and induce transcription of target genes. (B) The non-canonical pathway depends on the activation of NIK by a number of receptors including B-cell activating factor (BAFF), CD40 receptor and lymphotoxin β receptor (LTβR). Activation of NIK leads to subsequent activation of IKKα. IKKα phosphorylation of the p100 NF-κB subunit (yellow octagons) leads to proteosomal processing of p100 to p52. This results in the activation of p52:RelB heterodimers, which are free to translocate to the nucleus to induce transcription of target genes.

1.3.2 NF-κB activation by specific receptors

Although a number of NF-κB activation signals converge on the IKK complex, there are fundamental differences which precede IKK activation dependent on the stimulus received, leading to specific NF-κB responses. Many receptor-induced activation pathways of IKK have been elucidated for TLR and TNFR families, which are outlined below.

1.3.2.1 Toll-like receptors (TLRs)

Toll-like receptors are a class of transmembrane proteins which belong to a family of pattern recognition receptors, which play a key role in the innate immune system. TLRs function to recognise pathogen-associated molecular patterns (PAMPs), defined as conserved molecular motifs associated with microbial pathogens. Following PAMP recognition, TLRs induce signal cascades of downstream effectors to activate innate immune responses. To date, the mammalian TLR family consists of 13 members, each detecting distinct PAMPS derived from a range of microbial pathogens such as bacteria, viruses and fungi. Of the 13 members, only TLR1-10 have been identified in humans (Kawasaki and Kawai, 2014). While each TLR varies with regards to PAMP recognition, all TLRs share a conserved Toll/IL-1R (TIR) domain, which mediates downstream signalling by recruiting TIR domain containing adapter molecules. These include the myeloid differentiation primary-response protein (MYD88), TIR-containing adaptor protein/MYD88-adaptor-like (TIRAP), TIR-domain-containing adaptor protein inducing IFNß (TRIF), TRIF-related adaptor molecule (TRAM) and sterile alpha- and armadillo-motif-containing protein (SARM) (O'Neill and Bowie, 2007). TLRs use different combinations of adaptor proteins in order to mediate downstream signals. TLR signalling can be divided into two pathways based on the requirement of the MYD88 adapter (Medzhitov et al., 1998).

1.3.2.2 TLR4

TLR4 was one of the first TLRs identified, and was found to be involved in NF-κB activation and inflammatory gene expression by inducing transcription of *IL1* and *IL8* (Medzhitov and Janeway, 1997). Subsequent studies using *Tlr4* null mice revealed an essential role for TLR4 in lipopolysaccharide (LPS) induced signalling (Poltorak et al., 1998). LPS is the outer membrane component of gram negative

bacteria, and is an extremely potent PAMP, inducing the release of critical pro-inflammatory factors to activate the immune response. While TLR4 primarily responds to LPS, TLR4 can recognise other PAMPS derived from fungi, viruses, parasites and host, highlighting the importance of TLR4 for the innate immune response (Kawasaki and Kawai, 2014). In addition, TLR4 is the only TLR which can signal through both MYD88-dependent and MYD88-independent pathways. The MYD88-dependent pathway is responsible for pro-inflammatory cytokine production through activation of NF-κB and mitogen activated protein kinase (MAPK) gene expression, while the MYD88-independent pathway activates interferon regulatory transcription factor (IRF) families, mediating the induction of interferon regulatory factors (IFNs) (Lu et al., 2008).

1.3.2.2.1 MYD88 dependent pathway

Upon ligand binding, TLR4 receptors heterodimerise and recruit adaptor molecules via interactions with their TIR domains. MYD88 recruitment is facilitated by the bridging adaptor molecule TIRAP. In addition to the TIR domain, MYD88 also contains a death domain (DD) through which it can recruit other DD containing proteins. MYD88 recruits and activates IL-1 receptor associated kinase 4 (IRAK4), a DD containing kinase. IRAK4 recruits and phosphorylates IRAK1 and IRAK2. IRAK1 undergoes auto phosphorylation which in turn recruits tumour necrosis factor receptor-associated factor 6 (TRAF6), an E3 ubiquitin ligase. Together with ubiquitin conjugating enzymes UBC13 and UEV1A, TRAF6 promotes ubiquitination of itself, which recruits a kinase complex consisting of transforming growth factor-ß-activated kinase 1 (TAK1) and TAK1 binding proteins (TABs) 2 and 3 (Figure 1.4A). TAK1 is capable of activating both NF-κB and MAPK signalling pathways, leading to pro-inflammatory gene expression (Lu et al., 2008). TAK1 binds to the regulatory subunit of the IKK complex NEMO via ubiquitin chains, which allows it to phosphorylate and activate IKKß. TAK1 phosphorylation of IKKß, leads to IκBα degradation, and subsequent NF-κB activation gene expression. TAK1 phosphorylation of MAPKs leads to the activation of extracellular signal-regulated kinases (ERK), c-Jun N-terminal kinase (JNK) and p38 pathways, and subsequent gene expression, regulated by a number of transcription factors, including AP-1 (Kawasaki and Kawai, 2014) (Figure 1.4A). However, studies have demonstrated that TAK1 activation of MAPK pathways is cell type-specific. TAK1 can activate ERK1/2, p38

and JNK in fibroblasts and B-cells (Sato et al., 2005), but does not induce ERK1/2 activation in monocytes (Lee et al., 2000).

TLR4 is also capable of utilising the TIR adaptor protein SARM. In unstimulated cells, SARM and TRIF interact with each other weakly. However, following TLR4 activation by LPS, SARM expression is upregulated and the interaction between SARM and TRIF is stabilised (Carty et al., 2006). This complex prevents TRIF from interacting with other TLR adapter proteins, thus reducing TRIF mediated TLR signalling. It has been shown that SARM inhibits the TRIF-mediated, but not MYD88-mediated, activation of NF-κB (Carty et al., 2006). It is hypothesised that this negative regulation may function to prevent excessive activation in response to LPS by limiting pro-inflammatory gene expression.

1.3.2.2.2 MYD88 independent pathway

TLR4 is able to signal without MYD88 by utilising the adaptor molecule TRIF, through association with bridging adaptor TRAM. Through the TIR domain, TRIF recruits TRAF6. In addition, TRIF contains a receptor-interacting protein (RIP) interaction motif, which mediates its interaction with RIP1. The serine/threonine kinase RIP1 is required for the activation of TAK1, which leads to subsequent activation of the IKK complex resulting in NF-κB and MAPK activation (Lu et al., 2008). TRIF signalling also induces interferon response genes through the IRF3 transcription factor. TRIF recruits TRAF3, which is responsible for activating TANK binding kinase 1 (TBK1) and IKKε. TBK1 and IKKε catalyse the dimerisation and nuclear translocation of IRF3 (Oganesyan et al., 2006). IRF3, together with NF-κB, activate transcription of type 1 interferons (Honda et al., 2006) (Figure 1.4B).

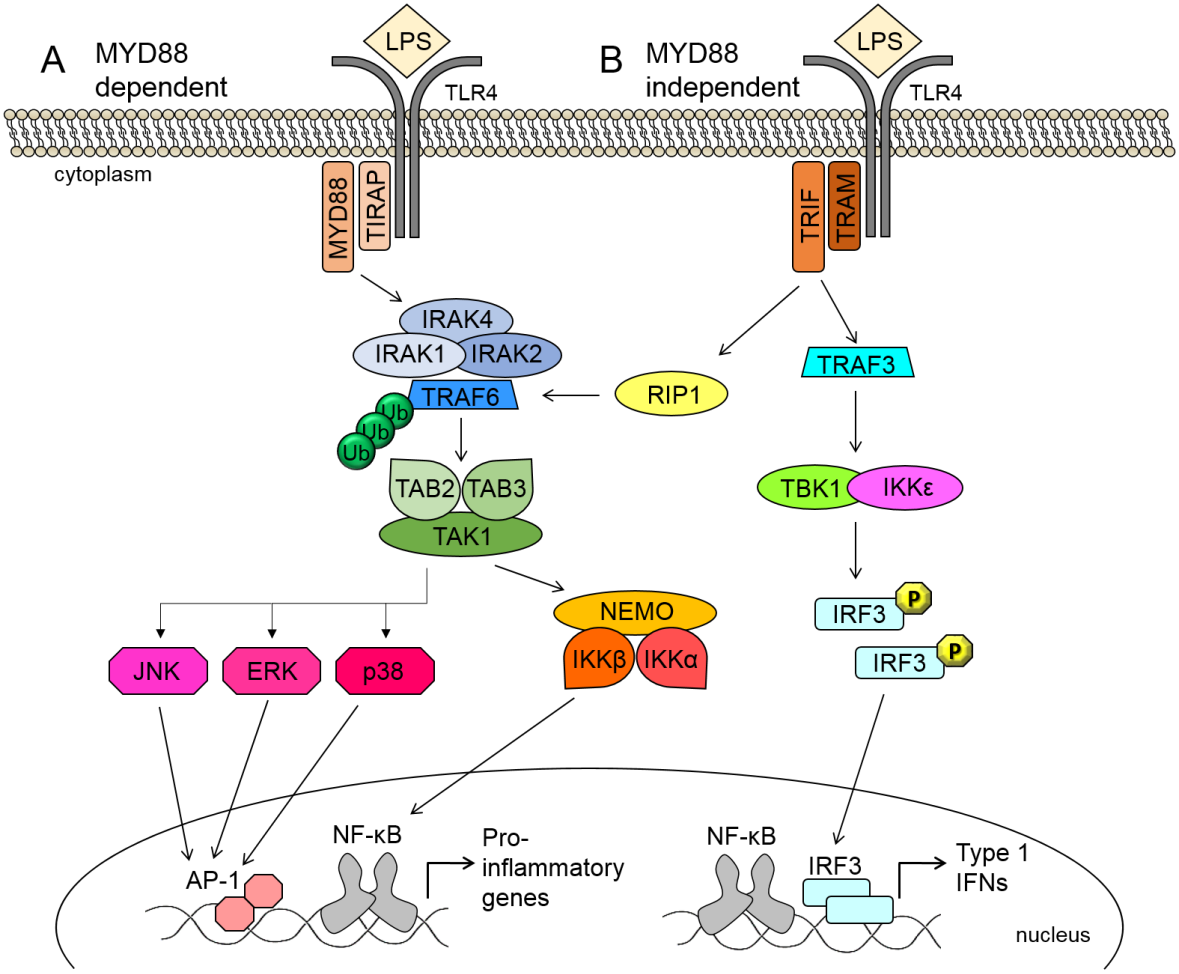


Figure 1.4: MYD88 dependent and MYD88 independent TLR4 signalling
Engagement of TLR4 by ligands such as LPS induces receptor dimerisation, activating signalling through both MYD88 dependent (A) and MYD88 independent (B) pathways. (A) TIR domains in the cytoplasmic portion of TLR4 receptors recruit bridging adaptor TIRAP, which in turn recruits MYD88. MYD88 recruits members of the IRAK family (IRAK1, IRAK2 and IRAK4) and TRAF6. Ubiquitination (Ub) of this complex recruits and activates the TAK1 complex (TAK1, TAB2 and TAB3). TAK1 binds to NEMO and phosphorylates IKKβ, leading to subsequent activation of NF-κB signalling through the canonical pathway. TAK1 also activates MAPK signalling pathways JNK, ERK and p38, inducing transcription factors such as AP-1. (B) TLR4 TIR domains also facilitate binding of bridging adaptor molecule TRAM, which recruits TRIF. TRIF interacts with RIP1, which is required for the activation of TAK1 signalling. In addition, TRIF recruits TRAF3, which activates TBK1 and IKKε kinases that phosphorylate IRF3. Phosphorylation induced IRF3 dimerisation and nuclear translocation, activating the expression of type 1 interferons.

1.3.2.3 Tumour necrosis factor receptors (TNFRs)

TNFRs are a superfamily of transmembrane receptors that play key roles in innate and adaptive immunity. The superfamily contains 29 members, characterised by the presence of cysteine-rich domain (CRD) in the extracellular region, responsible for binding 19 different tumour necrosis factor (TNF) ligands, including the potent inflammatory cytokine TNFα (Li et al., 2013). TNFRs can be divided into those containing a DD and those without, known as death receptors and activating receptors, respectively. The signalling cascades triggered by

TNFRs are diverse, and are determined by which proteins are recruited to their intracellular domains. The main pathways induced by TNFRs include the activation of NF-κB and MAPK, and apoptosis (Li et al., 2013).

TNFα binds to two receptors TNFR1 and TNFR2. TNFR1 is expressed ubiquitously in almost all cell types, whereas TNFR2 is limited to specific cell types, and expressed at lower levels in the resting state (Naude et al., 2011). A number of *in vivo* studies have demonstrated the significant role that TNFR1 plays in the immune response. TNFR1 has been shown to be primarily responsible for initiating inflammatory responses induced by TNFα (Loetscher et al., 1993, van der Poll et al., 1996). Furthermore, a study using *Tnfr1* and *Tnfr2* deficient mice revealed that *Tnfr1* deficient mice die from infection, whereas *Tnfr2* deficient mice are unaffected (Deckert-Schlüter et al., 1998). In addition, while NF-κB activation can be induced by both TNFR1 and TNFR2, is has been shown that TNFα-induced NF-κB activation occurs predominantly through TNFR1 (McFarlane et al., 2001).

TNFR1 exists in trimeric complexes as a pre-requisite to ligand binding and to prevent spontaneous receptor autoactivation (Chan et al., 2000). Upon ligand binding, TNFR1 oligomers undergo conformational change, which recruits TNF receptor-associated death domain (TRADD) mediated through DD interactions. TRADD subsequently recruits RIP1 kinase and TRAF2, where TRAF2 forms strong interaction with cellular inhibitor of apoptosis protein (cIAP) 1 and cIAP2. The formation of this signalling complex results in ubiquitination of RIP1. The exact mechanism utilised by this TNFR1-induced signalling complex to activate NF-κB remains unclear, and is the subject of ongoing research [as reviewed by (Wajant and Scheurich, 2011)]. However, a number of studies have highlighted the importance of RIP1 ubiquitination. Polyubiquitination of RIP1 was shown to be critical for IKK activation, and that polyubiquitin chains serve to bind IKK regulatory subunit NEMO (Ea et al., 2006, Wu et al., 2006). This led to the proposed model that RIP1 polyubiquitin chains provide a scaffold to recruit the TAK/TAB complex, allowing TAK1 to phosphorylate IKKß through its close proximity to the IKK complex. Therefore TNFα-induced phosphorylation of IKKß results in the degradation of IκBα, leading to the activation of NF-κB (Figure 1.5).

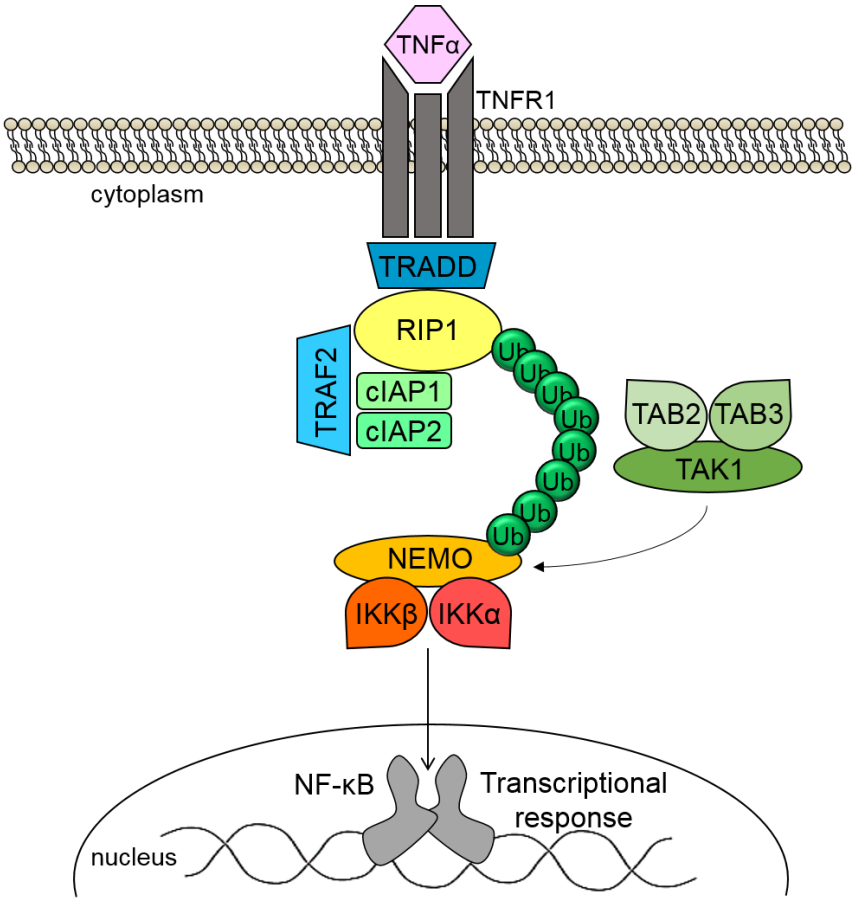


Figure 1.5: TNFα-induced TNFR1 signalling
Upon binding of TNFα, TNFR1 oligomers recruit TRADD via interaction with death domains. TRADD recruits RIP1 kinase and TRAF2, forming a strong interaction with cIAP1 and cIAP2. As a result of the formation of this complex, RIP1 is ubiquitinated (Ub). It is proposed that RIP1 binds to NEMO via polyubiquitin chains, bringing the IKK complex in close proximity with the TAK1 complex, facilitating IKKβ phosphorylation. Phosphorylation of IKKβ leads to subsequent activation of NF-κB signalling through the canonical pathway.

1.4 NF-κB p50 subunit

1.4.1 Importance of *NFKB1*

The *NFKB1* gene encodes the precursor protein p105, which undergoes limited proteasomal processing to generate the NF-κB p50 subunit. The NF-κB p50 subunit is an important dual regulator of inflammation, which can either activate or repress inflammatory gene expression depending on the formation of heterodimer or homodimer complexes, respectively. Although p50 forms the ubiquitously expressed p50:p65 heterodimer complex, p50 is often considered as an "accessory" subunit, and is consequently understudied. Research on the dynamic regulation of NF-κB function has focussed mainly on the p65 subunit. The p65 subunit has received most attention due to the lethal phenotype of *Rela*$^{-/-}$ mice (Gerondakis et al., 2006). Although *Nfkb1*$^{-/-}$ mice are viable, they display a number of complex inflammatory phenotypes, indicating the importance of p105/p50 in both innate and adaptive immune responses (Gerondakis et al., 2006).

Studies using *Nfkb1*$^{-/-}$ mice have shown that p105/p50 is essential for lymphocyte function. Mice lacking p105/p50 are defective for B-cell proliferation in response to LPS (Sha et al., 1995). In addition, B-cells from *Nfkb1*$^{-/-}$ mice are also defective for differentiation and isotype switching (Snapper et al., 1996). Therefore, *Nfkb1*$^{-/-}$ mice display impaired humoral immune responses (Sha et al., 1995). Follicular B-cell populations are turned over rapidly *in vivo*, leading to accelerated rates of B-cell apoptosis (Grumont et al., 1998). In addition, marginal zone and CD5^{+} peritoneal B-cells are reduced in *Nfkb1*$^{-/-}$ mice (Cariappa et al., 2000, Pohl et al., 2002). Lack of p105/p50 also impairs T helper (Th) cell function. Both Th1 and Th2 cells from *Nfkb1*$^{-/-}$ mice are impaired in their ability to produce effector cytokines such as IFN-γ and IL-4, respectively (Artis et al., 2005, Artis et al., 2003). It has also been shown that p105/p50 is required for granulocyte-macrophage colony-stimulating factor (GM-CSF) production by CD4^{+} T-cells (Campbell et al., 2011). Furthermore, *Nfkb1*$^{-/-}$ dendritic cells fail to promote Th2 differentiation (Artis et al., 2005). Mice lacking p105, but not p50 have shown the importance of p105 as a suppressor of inflammation. Deletion of the p105 C-terminal domain of ankyrin repeats (*Nfkb1*$^{\Delta CT/\Delta CT}$) leads to the development of enlarged spleens and lymph nodes, and inflammation of the

lungs and livers of *Nfkb1*$^{\Delta CT/\Delta CT}$ mice (Ishikawa et al., 1997). *Nfkb1*$^{-/-}$ mice studies have also revealed that p105/p50 is essential for monocyte function. Macrophages taken from *Nfkb1*$^{-/-}$ mice show impaired expression of immune effectors IL-6, IL-10 and cyclo-oxygenase 2 (COX2) when stimulated with TLR ligands, indicating the importance of p105/p50 in innate immune response (Sha et al., 1995, Waterfield et al., 2003, Banerjee et al., 2006).

Nfkb1$^{-/-}$ mice have been utilised for models of a number of diseases. An experimental model mimicking chronic liver disease revealed that *Nfkb1*$^{-/-}$ mice suffer severe chronic inflammation and fibrosis compared to WT mice (Oakley et al., 2005). The role of p50 in the resolution of LPS-induced renal inflammation was demonstrated, whereby *Nfkb1*$^{-/-}$ mice expressed elevated levels of chemokines, had prolonged renal inflammatory cell infiltration, and reduced survival rates (Panzer et al., 2009). Similar results were observed in a model of hippocampal inflammation, where *Nfkb1*$^{-/-}$ mice showed enhanced neutrophil infiltration and increased levels of pro-inflammatory gene expression in response to LPS (Rolova et al., 2014). Furthermore, during pulmonary infection with *Escherichia coli*, *Nfkb1*$^{-/-}$ mice expressed elevated levels of chemokines, increased inflammation of the lung, and had reduced survival rates, demonstrating the role of p105/p50 in a pneumonia model (Mizgerd et al., 2003). In contrast, *Nfkb1*$^{-/-}$ mice were shown to be resistant to chronic and acute models of arthritis, and mounted a weaker inflammatory response than WT mice, showing the importance of p105/p50 in the pathogenesis of inflammatory arthritis (Campbell et al., 2011). In addition to acting as suppressor of inflammation, p105/p50 has also been reported to have roles as a tumour suppressor in a model of hepatocellular carcinoma (HCC) using *Nfkb1*$^{-/-}$ mice. Loss of p105/p50 promotes the development of age-associated chronic liver disease, characterised by steatosis, neutrophilia, fibrosis, hepatocyte telomere damage and HCC (Wilson et al., 2015). In particular, development of HCC was attributed to lack of p50 by utilising an *Nfkb1*$^{S340A/S340A}$ mouse, carrying a mutation which disrupts the formation of p50 homodimers. Wilson et al. (2015) showed that p50 homodimers were important for repressing chemokine expression in neutrophils.

Functional polymorphisms in *NFKB1* are significant risk factors for the development of a number of human diseases such as inflammatory bowel disease (Karban et al., 2004), acute respiratory distress syndrome (Bajwa et al., 2011), systemic lupus erythematosus (Cen et al., 2013), chronic obstructive pulmonary disease (Huang et al., 2013), common variable immune deficiency (Tuijnenburg et al.), as well as a number of cancers (Sun and Zhang, 2007).

Despite the obvious roles that *NFKB1* plays in inflammatory responses and human disease, there is a lack of knowledge of the fundamental biochemistry of p105/p50. Therefore, understanding p105/p50 function is essential for providing insight into how NF-κB activity regulates the immune response.

1.4.2 p105 processing

p105 undergoes limited proteasomal processing to generate the p50 subunit, however the precise molecular mechanisms involved remain unclear. p105 processing to p50 occurs in both a constitutive and inducible manner, which require prior ubiquitination. Limited proteolysis is an example of a rare case in which the proteasome does not completely destroy its target, but only partially degrades it. The proteasome requires a glycine-rich region in the N-terminal half of p105 (amino acids 376–404) thought to be a poor substrate for the proteasome, which acts as a stop signal (Lin and Ghosh, 1996, Orian et al., 2000). Constitutive processing of p105 may occur co-translationally (Lin et al., 1998) and post-translationally (Fan and Maniatis, 1991). Signal induced processing requires site-specific phosphorylation of the C-terminal region (Fujimoto et al., 1995, Heissmeyer et al., 1999, Heissmeyer et al., 2001, Salmeron et al., 2001, Lang et al., 2003). In addition, complete degradation of p105 may be induced without the generation of p50 (Orian et al., 2000, Heissmeyer et al., 2001).

Initial studies identified phosphorylation of S893 and S907 residues by an unknown kinase as important for the proteasomal processing of p105. However, it is not clear whether p50 is generated as a result of p105 phosphorylation at these sites (Fujimoto et al., 1995). Stimulus induced phosphorylation of S923 by IKKβ has been reported to cause complete degradation of p105. However, there are conflicting reports as to the contribution of each of these phosphorylations

to p105 degradation, which may be cell type-specific (Heissmeyer et al., 1999, Heissmeyer et al., 2001, Salmeron et al., 2001, Lang et al., 2003). Phosphorylation of p105 at S927 and S932 by IKKß is well studied (Orian et al., 2000, Salmeron et al., 2001, Heissmeyer et al., 2001, Lang et al., 2003). However, these reports are conflicting as to whether IKKß-induced phosphorylation of S927 and S932 results in complete degradation of p105 or limited processing to generate p50. A study by Cohen et al. (2004) suggested that IKKß has two distinct functions: stimulation of degradation and stimulation of processing, mediated by two different E3 ligases (Cohen et al., 2004). Indeed recruitment of ß-transducin repeat-containing (ß-TrCP) E3 ligase (Orian et al., 2000, Lang et al., 2003) results in complete degradation of p105 (Heissmeyer et al., 2001), whereas recruitment of Kip1 ubiquitylation-promoting complex 1 (KPC1) E3 ligase results in p105 processing (Kravtsova-Ivantsiv et al., 2015). In addition, KPC1 was shown to also regulate constitutive p105 processing (Kravtsova-Ivantsiv et al., 2015). Furthermore, p105 is phosphorylated at S903 and S907 by glycogen synthase kinase 3ß (GSK3ß) (Demarchi et al., 2003). In resting cells, phosphorylation of these residues by GSK3ß stabilises p105, and cells lacking GSK3ß undergo a higher rate of constitutive p105 processing to p50. S903 and S907 are also required for signal induced p105 processing in response to TNFα (Demarchi et al., 2003).

It is important to note that p105 processing does not exclusively function to generate the p50 subunit or to liberate bound NF-κB dimers. p105 is also an important negative regulator of MAPK activation downstream of a number of receptors including the TLRs and TNFRs. A small but distinct pool of p105 is bound to TPL-2 (MAP3K8), a serine/threonine kinase belonging to the MAP3K family. p105 serves to stabilise TPL-2 in an inactive form and blocks the activation of the MAP kinase cascade (Belich et al., 1999, Beinke et al., 2003). Activation of TPL-2 is dependent on p105 processing induced by IKKß phosphorylation of p105 S927 and S932. Following phosphorylation, subsequent proteasomal degradation of p105 liberates TPL-2 kinase (Beinke et al., 2003, Waterfield et al., 2003). Free TPL-2 phosphorylates and activates MEK1/2 and triggers a cascade of MAP kinase activity including ERK1/2 (Beinke et al., 2003, Waterfield et al., 2003).

1.4.3 Regulation of NF-κB p50 activity

Since NF-κB is involved in such a diverse range of biological processes, strict regulation of NF-κB signalling is essential. In particular, NF-κB regulates a large number of pro-inflammatory genes such as chemokines (e.g. *TNFa*, *IL6*, *IL8*) cytokines (e.g. *CXCL1*, *CXCL2*, *CCL2*) immunoreceptors (e.g. C-C chemokine receptor 5 (*CCR5*), *CCR7*, *CD40*) and adhesion molecules (e.g. vascular cell adhesion protein 1 (*VCAM1*), intercellular adhesion molecule 1 (*ICAM1*), *CD44*). Aberrant or dysregulated NF-κB signalling can result in inappropriate inflammatory responses, which can be detrimental to the host. Therefore transcriptional activity must be regulated tightly. Given that the NF-κB p50 subunit is involved in both activating and repressing inflammatory gene expression, its activity is regulated through interaction with co-factors and post-translational modifications (PTMs), in particular phosphorylation.

1.4.3.1 Regulation of transcriptional activity by interacting proteins

1.4.3.1.1 BCL-3

Although rapid induction of the innate immune system is essential for mounting the appropriate response to pathogens, the magnitude and duration of pro-inflammatory genes expression must be limited in some manner. Studies using *Nfkb1*$^{-/-}$ mice revealed that both p105 and p50 have a clear role in the negative regulation of inflammation. BCL-3 has been shown to be an essential regulator of p50 homodimers, and is required to limit the expression of pro-inflammatory genes in a number of immune cells such as macrophages, dendritic cells, and B-cells (Carmody et al., 2007). BCL-3 functions to stabilise p50 homodimers on DNA by preventing their ubiquitination and degradation by the proteasome, therefore occupying the promoters of target genes. Although p50 homodimers negatively regulate expression by competing with TAD containing dimers for κB sites, p50 homodimers are ineffective in the absence of BCL-3 interaction (Collins et al., 2014). The BCL-3 mediated stability of p50 homodimers on gene promoters is also crucial for establishing TLR tolerance in macrophages (Carmody et al., 2007). TLR tolerance is defined as a state of altered responsiveness to TLR stimulation, which reduces expression of deleterious pro-inflammatory genes, while maintaining expression of pro-resolution effectors (Yan et al., 2012). Promoter tolerance is generated by stabilising the p50:p50:DNA inhibitory

complex and requires the presence of both BCL-3 and p50 (Carmody et al., 2007).

1.4.3.1.2 HDAC-1

In addition to BCL-3, histone deacetylase 1 (HDAC-1) has been shown to be an important regulator of p50 homodimers. HDAC-1 is an abundant mammalian histone deacetylase, responsible for chromatin remodelling. Studies utilising chromatin immunoprecipitation (ChIP) assays revealed that HDAC-1 is recruited to the promoters of pro-inflammatory genes in a p50 dependent manner (Zhong et al., 2002, Elsharkawy et al., 2010). These findings provide a mechanism to explain the immunosuppressive properties of the p50 subunit: p50:p50:HDAC-1 complexes repress pro-inflammatory gene expression by condensing chromatin, rendering gene promoters inaccessible for transcription. Recently, Cartwright et al. (2018) have reported that HDAC-1 directly interacts with p50 homodimers bound to the promoters of a subset of NF-κB target genes to repress their expression. Mutation of the interaction site altered histone acetylation, and resulted in elevated chemokine expression and enhanced neutrophil chemotaxis. HDAC-1 interacts with p50 at its NLS, extending the list of NLS-interacting factors that regulate p50 activity, alongside BCL-3 and IκBα. (Cartwright et al., 2018).

1.4.3.2 Regulation of transcriptional activity by phosphorylation

The activity of NF-κB subunits is dynamically regulated by a number of PTMs, in particular phosphorylation. Phosphorylation controls interaction of NF-κB with other co-factors, and the stability, degradation and transcriptional activity of NF-κB dimers. Importantly, research is beginning to reveal that certain phosphorylation events contribute to the selective regulation of NF-κB transcriptional activity in a gene specific manner. Phosphorylation of NF-κB subunits function to either enhance or downregulate target gene expression, or to modulate transcriptional responses, rather than acting as a simple on/off switch for gene transcription. Site-specific phosphorylations regulating transcriptional activity can be a consequence of signalling from upstream components of the NF-κB pathway, such as IKKα and IKKβ. NF-κB can also be

phosphorylated by kinases from other signalling pathways, providing crosstalk with multiple signalling networks (Christian et al., 2016).

To date, research of NF-κB subunit phosphorylation has focussed mainly on the p65 subunit. A number of studies have identified and functionally described 22 phosphorylation sites (www.phosphosite.org). Two of the most well characterised p65 phosphorylation sites are S276 and S536. S276 and S536 are phosphorylated by a number of different kinases, which have distinct outcomes for p65 regulation (Christian et al., 2016). These include transcriptional activation or inhibition, enhanced DNA binding, regulation of interaction with other modulators or further modification of the subunit by additional PTMs. In particular, studies have shown evidence of selective gene expression regulated by phosphorylation. Anrather et al. (2005) showed that phosphorylation of p65 at multiple residues including S276 regulates p65 transcriptional activity in a κB site-specific manner, using reporter constructs (Anrather et al., 2005). Furthermore, Hochrainer et al. (2013) demonstrated that S276 phosphorylation regulates the expression of a distinct subset of NF-κB target genes in the context of endogenous gene promoters (Hochrainer et al., 2013). In addition, Moreno et al. (2010) revealed that phosphorylation of S536 selectively regulates the induction and repression of NF-κB target genes, demonstrated by using a S536A phospho-mutant. This led to the proposal of the NF-κB barcode hypothesis which suggests that post-translational modifications of NF-κB subunits, either alone or in combination, generate distinct functional states that direct transcription in a gene specific manner (Moreno et al., 2010).

To date, the majority of the p105/p50 phosphorylation sites identified are located in the C-terminal region of p105 and regulate the processing of p105 by the proteasome (described in section 1.4.2) (Christian et al., 2016) (Figure 1.6). Phosphorylation of the p50 subunit is much less well understood, and to date only 4 p50 phosphorylation sites have been studied in detail: S337, S242, S20 and S328 (positions highlighted in Figure 1.6).

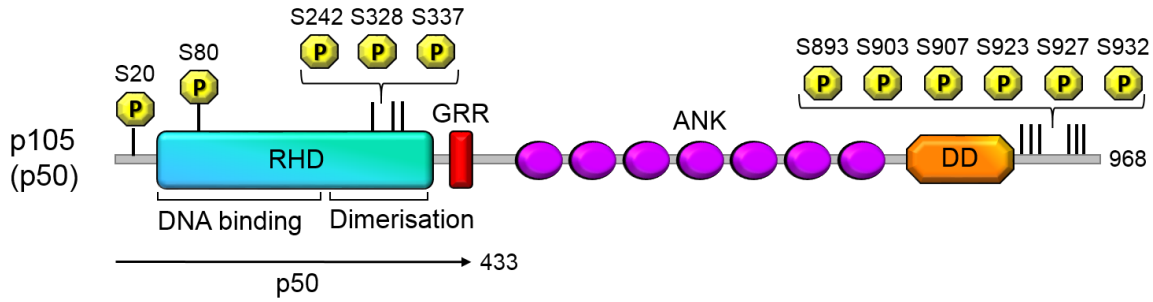


Figure 1.6: Phosphorylation of NFKB1
Schematic representation of human NFKB1 protein, highlighting the positions of experimentally identified phosphorylation sites known to regulate activity (yellow octagons). Principal structural motifs are indicated, including a Rel homology domain (RHD) consisting of a DNA binding and dimerisation region; a glycine rich region (GRR); an ankyrin repeat domain (ANK) and a death domain (DD). Amino acid residue numbering corresponds to human protein sequence.

The phosphorylation of p50 at S337 is critical for DNA binding. S337 lies in a protein kinase A (PKA) consensus site and PKA has been demonstrated to phosphorylate p50 at S337 *in vitro* and *in vivo* (Hou et al., 2003). It was also reported that p50 S337 is constitutively phosphorylated by PKA (Guan et al., 2005). Substitution of serine residues S65 and S342 with alanine (A) also abolish p50 DNA binding but there is, as of yet, no demonstration that these sites are phosphorylated (Hou et al., 2003). Recently, S242 was identified by mass spectrometry as a phosphorylation site for checkpoint protein kinase 1 (Chk1) (Vonderach et al., 2018). Phosphorylation of S242 by Chk1 is key for p50 homodimer binding. S242 lies in the DNA binding interface of p50, close to a critical lysine residue required for interacting with DNA. A phosphomimetic mutant, where p50 S242 was substituted to aspartate (D), p50 exhibited a >10-fold reduction in DNA binding affinity compared to wild type p50 (Vonderach et al., 2018). In addition, this study reported that phosphorylation of S337 by both Chk1 and PKA is critical for the regulation of p50 homodimer formation. Specifically, phosphorylation of S337 by both kinases disrupted p50 homodimerisation (Vonderach et al., 2018). These findings are in contrast with Hou et al. (2003), who report that mutation of S337 does not affect dimerisation. Phosphorylation of human p50 at S20 by DNA-dependent protein kinase (DNA-PK) is critical for expression of VCAM1 in response to TNFα (Ju et al., 2010). However, this serine residue is not conserved in mouse p50 and so it appears that the regulation of p50 function by phosphorylation at this site is specific to human.

The phosphorylation of S328 by Chk1 is triggered by DNA damage and regulates the interaction of p50 with specific NF-κB binding elements based on the κB site nucleotide sequence (Crawley et al., 2013). S328 phosphorylation reduces the affinity of p50 for κB sites with a cytosine (C) at the -1 position, without affecting binding to κB sites containing a -1 adenine (A). The differential interaction between phospho-p50 and specific κB sequences regulates the transcriptional response following DNA damage by inhibiting the expression of anti-apoptotic genes. As a consequence, p50 S328 phosphorylation promotes DNA damage induced cytotoxicity (Crawley et al., 2015). In genes with multiple κB sites, the presence of a single κB site with a C at the -1 position is sufficient to inhibit NF-κB-dependent activity. The studies of p50 S328 phosphorylation provide the strongest evidence to date that p50 phosphorylation can selectively regulate p50 transcriptional activity by modulating DNA binding in a sequence specific manner. The inhibition of gene specific transcription by p50 S328 phosphorylation presumably occurs through altered DNA binding of p50 containing heterodimers (e.g. p50:p65). Given that p50 has been shown to be an important negative regulator of inflammatory gene expression (Carmody et al., 2007, Collins et al., 2014) S328 phosphorylation of p50 homodimers may also have an important and distinct impact on transcriptional outcomes.

An additional serine residue of p50, S340, has been recently shown to be important for the formation of p50 homodimers (Wilson et al., 2015). The mutation of S340 to alanine leads to a loss of p50 homodimers. S340A knock-in mice recapitulate the *Nfkb1*$^{-/-}$ phenotype in the hepatocellular carcinoma model (described previously in section 1.4.1). However, to date, there is no direct evidence that S340 is phosphorylated. Although, given the recent findings by Vonderach et al. (2018), who report that phosphorylation of p50 S337 by Chk1 and PKA destabilises p50 dimerisation, it is plausible that p50 S340 phosphorylation may regulate dimer formation.

It is clear that phosphorylation of NF-κB subunits plays an important role in regulating transcriptional activity. In particular, site-specific phosphorylation of NF-κB subunits regulates gene expression in a selective manner. However, there is still a considerable gap in our mechanistic understanding of how phosphorylation events coordinate NF-κB transcriptional specificity. There is

evidence to suggest that the sequence of NF-κB binding sites in DNA contributes to the selective effects of certain phosphorylation events. However, it is equally clear that the phosphorylation of NF-κB subunits may also regulate the interaction with other factors which in turn may determine transcriptional outcome. Understanding gene-specific regulatory mechanisms will be key to elucidating distinct transcriptional profiles that are regulated by NF-κB phosphorylation.

1.5 Thesis aims

The NF-κB subunit p50 is an essential dual regulator of inflammation, playing important roles in human health and disease. Despite its importance, the regulation of p50 transcriptional activity is understudied. Site-specific phosphorylation has been shown to be a key regulator of p65 mediated inflammatory gene expression, however there is a lack of research investigating the importance of p50 phosphorylation. A novel phosphorylation site at S80 on the p50 subunit has been identified, and is phosphorylated by IKKβ. It is hypothesised that S80 phosphorylation regulates inflammation in a gene specific manner. The aims of this thesis were to:

1. Generate *NFKB1*^{S80A} human cell lines utilising CRISPR/Cas9 genome editing techniques to create tools to investigate the role of NF-κB p50 S80 phosphorylation *in vitro*.

2. Investigate the transcriptomic profile regulated by p50 S80 phosphorylation utilising RNA sequencing analysis.

3. Investigate the mechanism of S80 phosphorylation in the regulation of inflammatory gene expression.

Chapter 2

2 General materials and methods

2.1 Materials

2.1.1 Antibodies

Antibody	Clone	Supplier (Catalogue No.)
α-Tubulin	B-5-1-2	Sigma Aldrich (T6074)
β-Actin	ac-15	Sigma Aldrich (A5441)
HDAC1	H-51	Santa Cruz (sc-7872)
HDAC1	Polyclonal	Sigma Aldrich (AV 38530)
IκBα	44D4	Cell Signaling (4821)
Phospho-IκBα (Ser32/36)	5A5	Cell Signaling (9246)
NF-κB1 p105/p50	D7H5M	Cell Signaling (12540)
NF-κB1 p50	C-19	Santa Cruz (sc-1190)
p38 MAPK	Polyclonal	Cell Signaling (9212)
Phospho-p38 MAPK	Thr180/Tyr182	Cell Signaling (9211)
p44/42 MAPK(ERK1/2)	p44/42	Cell Signaling (9102)
Phospho-p44/42 MAPK(ERK1/2)	Thr202/Tyr204	Cell Signaling (9101)
NF-κB p65	C-20	Santa Cruz (sc-372)
RelA	Polyclonal	Bethyl Laboratories (A301-824A)
Xpress	XP	Invitrogen (R910-25)

Table 2.1: Primary antibodies

Antibody	Clone	Supplier (Catalogue No.)
Anti-mouse IgG HRP	Whole Ab	Amersham (NA931)
Anti-rabbit IgG HRP	Whole Ab	Amersham (NA934)
Rabbit-anti-goat IgG-HRP	Whole Ab	Santa Cruz (sc-2354)
Mouse-anti-goat IgG-HRP	Whole Ab	Santa Cruz (sc-2768)

Table 2.2: Secondary antibodies

2.1.2 Reagents

All general salts and chemicals were purchased from Sigma Aldrich UK unless otherwise stated, and were stored and prepared according to the manufacturer's instructions.

2.1.3 Buffers

2.1.3.1 Electrophoresis buffers

0.5X Tris-borate-EDTA (TBE) buffer
5.4g Tris (Fisher Scientific) 2.75g boric acid 2ml 0.5M EDTA

Table 2.3: Agarose gel electrophoresis buffer

1X Running buffer	1X Transfer buffer
25mM Tris 250mM glycine (Fisher Scientific) 0.1% SDS	48mM Tris 39mM glycine 0.0375% SDS 20% methanol (Fisher Scientific)

Table 2.4: Western blot electrophoresis buffers

2.1.3.2 Electrophoresis gels

Agarose gel
1.5% agarose (Bio-Rad) 0.5X TBE buffer 1X GelRed (Biotum)

Table 2.5: Agarose gel recipe

	Gel	
Reagent	5% Stacking (total 1ml)	10% Resolving (total 5ml)
H_2O	0.68ml	1.9ml
30% acrylamide	0.17ml	1.7ml
1.0M Tris-HCl (pH6.8)	0.13ml	-
1.5M Tris-HCl (pH8.8)	-	1.3ml
10% SDS	0.01ml	0.05ml
10% APS	0.01ml	0.05ml
TEMED	0.001ml	0.002ml

Table 2.6: SDS-PAGE gel recipes

2.1.3.3 Sample buffers

2X SDS Sample buffer	5X SDS Sample buffer
100mM Tris-HCl (pH6.8)	250mM Tris-HCl (pH6.8)
4% (w/v) SDS	10% (w/v) SDS
0.2% (w/v) bromophenol blue	0.02% (w/v) bromophenol blue
20% (w/v) glycerol	30% (w/v) glycerol
200mM β-mercaptoethanol	5% β-mercaptoethanol

Table 2.7: SDS-PAGE sample buffers

2.1.3.4 Assay buffers

Radioimmunoprecipitation assay (RIPA) buffer
50mM Tris-HCl pH7.4
150mM NaCl
0.25% Na Deoxycholate
1mM EDTA pH8.0
1% Nonidet P-40
1mM NaF*
1mM Na_3VO_4*
1mM PMSF*
1µg/ml leupeptin*
2µg/ml aprotinin*
2µg/ml pepstatin*

Table 2.8: RIPA buffer
*Added fresh on day of use

DNA affinity purification assay (DAPA) buffer
10mM Tris-HCl pH7.5
50mM NaCl
1mM DTT*
5% glycerol
1mM EDTA
1mM NaF*
1mM Na_3VO_4*
1mM PMSF*
1µg/ml leupeptin*
2µg/ml aprotinin*
2µg/ml pepstatin*

Table 2.9: DAPA buffer
*Added fresh on day of use

2.1.4 Plasmids

Plasmid	Source
pPGKpuro	A gift from Rudolf Jaenisch (Addgene plasmid # 11349)
pSpCas9(BB)-2A-GFP (PX458)	A gift from Feng Zhang (Addgene plasmid # 48138)
gRNA/Cas9	This study
pTAL-Luc	Purchased from Clontech (Catalogue no. 631909)
pTAL-Luc-NFKB-AA	This study
pTAL-Luc-NFKB-AC	This study
pTAL-Luc-NFKB-GA	This study
pTAL-Luc-NFKB-GC	This study
pRL-TK	Purchased from Promega (Catalogue No. E2241)
pEF4a $p50^{WT}$ Xpress	Carmody lab stocks
pEF4a $p50^{S80A}$ Xpress	Carmody lab stocks

Table 2.10: Plasmids sources

2.2 Methods

2.2.1 Cell culture

2.2.1.1 Maintenance

Human embryonic kidney 293T (HEK293T) cells were cultured in high glucose Dulbecco's Modified Eagle Medium (DMEM), supplemented with 10% heat-inactivated fetal bovine serum (FBS) (Invitrogen), 2mM L-glutamine, and 100 units/ml of streptomycin and penicillin. Cells were sub-cultured by enzymatic detachment with 0.05% Trypsin-EDTA solution (Invitrogen). THP-1 cells were cultured in Roswell Park Memorial Institute (RPMI) 1640 media, supplemented with 10% heat-inactivated FBS, 2mM L-glutamine, and 100 units/ml of streptomycin and penicillin. All cells were maintained at 37°C in a humidified environment with 5% CO_2.

2.2.1.2 Plating

Tissue culture plating conditions used for endogenous assays are described in Table 2.11.

Cell type	Assay	Plate/dish	Cell number
HEK293T	Endogenous protein	6-well plate	$1x10^6$
	Nuclear translocation	6cm dish	$2x10^6$
	DAPA	10cm dish	$7.5x10^6$
	Endogenous IP	10cm dish	$7.5x10^6$
	RNA-seq/qPCR	6-well plate	$1x10^6$
THP-1	Endogenous protein	24-well plate	$0.45x10^6$
	Nuclear translocation	6-well plate	$1.5x10^6$
	RNA seq/qPCR	24-well plate	$0.45x10^6$

Table 2.11: Plating conditions for endogenous assays

2.2.1.3 Transfection

2.2.1.3.1 Functional assays

Transfection conditions for functional assays are described in Table 2.12. WT HEK293T cells were transfected with Turbofect (ThermoFisher Scientific), according to the manufacturer's instructions.

Assay	Cell number	Plate/dish	Plasmid concentration
Dual luciferase assay	$1x10^5$	24-well plate	100ng Firefly: 10ng Renilla
Limited proteolysis	$2x10^6$	6-well plate	1µg of each plasmid

Table 2.12: Transfection conditions for functional assays

2.2.1.3.2 CRISPR/Cas9 genome editing

Transfection conditions used for CRISPR/Cas9 genome editing are described in Table 2.13.

Cell type	Buffer	Cell number	Electroporation parameters	DNA concentration
HEK293T	R	$5x10^7$ cells/ml ($5x10^6$ cells/transfection)	2 pulses 1100v 20ms	2µg pPGKpuro: 4µg pX458 sgRNA: 4µg ssODN
THP-1	T	$4x10^7$ cells/ml ($4x10^5$ cells/transfection)	2 pulses 1400v 20ms	20pmol Cas9: 20pmol sgRNA: 10pmol ssODN

Table 2.13: Transfection conditions for CRISPR/Cas9 genome editing

2.2.1.4 THP-1 differentiation

THP-1 cells were resuspended at a density of $1x10^6$/ml before plating as described in Table 2.11. THP-1 cells were treated with 25ng/ml PMA for 72 hours. Following differentiation, PMA-containing media was removed, cells were washed once in THP-1 culture media, and cells were replaced with fresh THP-1 culture media. Cells were either treated immediately, or were rested in culture for 4 days before treatment.

2.2.2 Protein analysis

2.2.2.1 Protein extraction

2.2.2.1.1 Non-denaturing whole cell extracts

Cell culture media was aspirated from tissue culture plates or dishes, and cells were washed gently with cold PBS. HEK293T cells and differentiated THP-1 cells were detached with cold PBS or cold PBS/5mM EDTA, respectively. Cells were pelleted by centrifugation at 11,000g for 45 seconds, and supernatants were removed. Cell pellets were resuspended in radioimmune precipitation assay (RIPA) buffer. Cells were incubated on ice for 30 minutes, and were vortexed every 5 minutes to lyse cells. Lysates were cleared by centrifugation at 16,000g for 10 minutes at 4°C. Supernatants were collected and analysed immediately, or stored at either -20°C or -80°C for long term storage.

2.2.2.1.2 Nuclear/cytoplasmic extracts

Nuclear and cytoplasmic extracts were isolated from cells utilising the Nuclear Extract Kit (Active Motif) according to the manufacturer's instructions with one additional step: nuclear pellets were washed with 1X hypotonic buffer before resuspension in complete lysis buffer for nuclear extraction.

2.2.2.2 Quantification

Protein concentrations for whole cell, cytoplasmic and nuclear extracts were quantified using a Bradford protein assay (Bio-Rad), where 1µl of extract was diluted in 1ml 1X Bradford reagent. A 6-point standard curve between 0-9µg/ml of bovine serum albumin (BSA) was utilised to determine unknown protein concentrations. The absorbance of each sample was measured in triplicate at 595nM using a spectrophotometric 96-well plate reader.

2.2.2.3 Western blotting

Protein samples were resolved by sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (PAGE) using the Mini-PROTEAN Tetra Cell System (Bio-Rad). Samples were diluted in either 2X or 5X SDS sample buffer (Table 2.7) and were resolved on 10% acrylamide gels (Table 2.6) at 100-120V for 90 minutes in 1X Tris-glycine running buffer (Table 2.4). Resolved proteins were subsequently transferred to Amersham Protran 0.45µM NC nitrocellulose Western blotting membranes (GE healthcare) in 1X Tris-glycine transfer buffer (Table 2.4) using the Mini Trans-Blot Electrophoretic Transfer System (Bio-Rad). To block non-specific binding, membranes were incubated in 5% fat free milk (Marvel)/PBS-Tween 20 (0.05%) (PBS-T) solution.

Membranes were probed with primary antibodies in 5% fat free milk/PBS-T solution or 5% BSA/PBS-T solution overnight at 4°C, and with secondary antibodies in 5% fat free milk/PBS-T solution for 1 hour at room temperature. Washed 3 times in PBS-T for 5 minutes. After each antibody incubation, membranes were washed in PBS-T 3 times for 5 minutes. To detect bound proteins, Western Bright ECL or Western Bright Sirius (Advansta) chemilumiescent substrates were used according to the manufacturer's instructions. Membranes were scanned using C-DiGit Blot Scanner (LI-COR) on

either standard or high sensitivity settings. Protein size was determined using a pre-stained molecular weight ladder (ThermoFisher Scientific), marked on membranes using a chemiluminescent marker (LI-COR). Western blot images were visualised and quantified using Image Studio Lite Version 5.2 (LI-COR). If proteins of interest were of similar or identical size, membranes were stripped using Restore Plus Western Blot Stripping Buffer (ThermoFisher Scientific) according to manufacturer's instructions, before re-blocking and re-probing the membrane.

2.2.3 Functional assays

2.2.3.1 Immunoprecipitation (IP)

For endogenous immunoprecipitation assays, cells were plated as described in Table 2.11. Cells were treated with 10ng/ml TNFα prior to harvest. Whole cell lysates were prepared as described in section 2.2.2.1.1. Equivalent concentrations of whole cell lysates were resuspended in a total of 1ml RIPA buffer. Samples were pre-cleared for 30 minutes at 4°C with 20µl protein G agarose beads (Millipore). Pre-cleared lysates were immunoprecipitated with primary antibody (5µl NF-κB1 p105/p50) overnight at 4°C with 20µl of fresh agarose beads. Beads were pelleted at 11,000g for 10 seconds and supernatant was discarded. Pellets were washed three times in RIPA buffer were and resuspended in equal volumes of 2X SDS loading dye. Samples were boiled for 5 minutes at 95°C and were centrifuged at 16,000g for 2 minutes. Denatured eluates were resolved by SDS-PAGE and analysed by western blot.

2.2.3.2 DNA affinity purification assay (DAPA)

5'-biotinylated and unlabelled single-stranded oligonucleotides were purchased from Eurofins Genomics (Table 2.14). Single-stranded oligonucleotides were dissolved in sterile deionised water to a final concentration of 1µg per 1 to 2µl. Equal quantities of either 5'-biotinylated sense and unlabelled antisense, or unlabelled sense and antisense single-stranded oligonucleotides were annealed by heating to 95°C for 5 minutes and allowed to cool to room temperature (>1 hour).

Oligo-nucleotide name	Sense (5'-3')	Antisense (5'-3')
AA	AGTTGAGGGGAATTTCCCAGGC	GCCTGGGAAATTCCCCTCAACT
AC	AGTTGAGGGGACTTTCCCAGGC	GCCTGGGAAAGTCCCCTCAACT
GA	AGTTGAGGGGGATTTCCCAGGC	GCCTGGGAAATCCCCCTCAACT
GC	AGTTGAGGGGGCTTTCCCAGGC	GCCTGGGAAAGCCCCCTCAACT

Table 2.14: DAPA oligonucleotides

Cells were plated as described in Table 2.11. Cells were left untreated or treated with 10ng/ml TNFα for 30 minutes. Nuclear extracts were prepared as described in section 2.2.2.1.2. DAPA reactions were prepared by mixing 1.5µg of 5'-biotinylated double-stranded oligonucleotides with 150µg of nuclear extract and 15µl streptavidin-agarose beads in a total of 500µl DAPA buffer. Unlabelled double-stranded oligonucleotides were added in 10-fold excess to confirm binding specificity. Reactions were incubated at room temperature on a rotator for 1 hour. Beads were pelleted by centrifugation at 550g for 10 seconds, and the supernatant was discarded. Beads were washed in 500µl DAPA buffer by inverting reaction tubes 10 times. The wash step was repeated twice. To elute DNA-bound proteins, beads were resuspended in 20µl of 2X SDS sample buffer, and incubated at 95°C for 5 minutes. Samples were vortexed for 10 seconds, then centrifuged at 16,000g for 2 minutes. Eluates were resolved on SDS-PAGE gels and analysed by Western blot.

2.2.3.3 Dual luciferase assay

Four different luciferase reporter constructs that contained four κB site repeats that vary at either the -2 or the -1 position immediately upstream of a minimal promoter and firefly luciferase reporter gene were generated. Oligonucleotides purchased from Eurofins Genomics (Table 2.15) were annealed by heating to 95°C for 5 minutes and allowed to cool to room temperature (>1 hour). Annealed oligonucleotides were ligated into pTAL-Luc plasmid backbone via NheI and BglII restriction site cloning, using techniques described in section 2.2.7.

Oligonucleotide name	Forward (5'-3')	Reverse (5'-3')
NFKB_AA	CTAGCGGGAATTTCCGGG AATTTCCGGGAATTTCCG GGAATTTCCA	GCCCTTAAAGGCCCTTAA AGGCCCTTAAAGGCCCTT AAAGGTCTAG
NFKB_AC	CTAGCGGGACTTTCCGGG ACTTTCCGGGACTTTCCG GGACTTTCCA	CTAGCGGGTGTTTCCGG GTGTTTCCGGGTGTTTCC GGGTGTTTCCA
NFKB_GA	CTAGCGGGGATTTCCGG GGATTTCCGGGGATTTCC GGGGATTTCCA	CTAGCGGGCTTTTCCGGG CTTTTCCGGGCTTTTCCG GGCTTTTCCA
NFKB_GC	CTAGCGGGGCTTTCCGG GGCTTTCCGGGGCTTTCC GGGGCTTTCCA	CTAGCGGGCGTTTCCGG GCGTTTCCGGGCGTTTCC GGGCGTTTCCA

Table 2.15: κB site repeat oligonucleotides

Luciferase assays were performed using the Dual-Luciferase® Reporter Assay System (Promega). Cells were transiently transfected as described in Table 2.12 for 24 hours. Co-transfection of Renilla luciferase expression vector pRL-TK was used as an internal control for all reporter assays. Cells were cultured with or without 10ng/ml TNFα for an additional 8 hours before harvest. Cells were washed once in cold PBS and lysed in 100µl passive lysis buffer. 15µl of lysate was added to 75µl LAR II buffer, and firefly luciferase activity was determined. 75µl of Stop & Glo® buffer was then added, and Renilla luciferase activity was determined. All measurements were performed for 2 seconds on a Microbeta Luminescence Counter (PerkinElmer). To normalise the transfection efficiency across all samples, firefly luciferase activity was divided by that of Renilla.

2.2.3.4 Limited proteolysis

For analysis of overexpressed p50, cells were transiently transfected with either $p50^{WT}$-XP or $p50^{S80A}$-XP plasmids as described in Table 2.12 for 24 hours. For analysis of endogenous p50, cells were plated as described in Table 2.11. Cell were treated with 10ng/ml TNFα for 30 minutes prior to harvest. Harvested and washed cells were lysed in RIPA buffer lacking protease inhibitors to extract whole cell lysates. Limited proteolysis was performed by adding varying ratios of trypsin to 50µg of protein, and incubating the reaction mixture for 30 minutes at 37°C. Proteolysis was terminated by adding 5X SDS sample buffer to the reaction and boiling for 5 minutes at 95°C. Samples were resolved by SDS-PAGE and analysed by western blot.

2.2.4 CRISPR/Cas9 genome editing

CRISPR/Cas9 genome editing was carried out with the technical assistance of Mr David Kerrigan.

2.2.4.1 Design and synthesis of CRISPR/Cas9 materials

The guide RNA (gRNA) (5'-ACAAACTTACTTTGACCTGA-3') used to target the S80 codon was designed using the Broad Institue sgRNA Designer: CRISPRko design tool (https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design). Off-target effects of the gRNA were predicted using the Benchling gRNA Design and Analysis tool (https://benchling.com/crispr). Design of the single-stranded DNA oligonucleotide (ssODN) homology directed repair (HDR) template was assisted by using the HR Template Design tool (https://benchling.com/crispr). For generation of the gRNA/Cas9 plasmid for genome editing in HEK293T cells, complementary oligonucleotides with overhangs required for cloning were purchased from Eurofins Genomics (Table 2.16). Complementary gRNA oligos were annealed, phosphorylated and cloned into PX48, according to the Zhang Lab General Cloning Protocol (http://www.addgene.org/crispr/zhang/). Sense and antisense ssODNs (Table 2.16) were ordered from IDT as 4 nmole ultramer® DNA oligos. For genome editing in THP-1 cells, a synthetic gRNA (5'-ACAAACTTACTTTGACCTGA-3') and purified Cas9 protein were purchased from Synthego.

Oligonucleotide name	Description	Sequence (5'-3')
gRNA 1F	Guide RNA	CACCGACAAACTTACTTTGACCTGA
gRNA 1R	Guide RNA	AAACTCAGGTCAAAGTAAGTTTGTC
HEK293T_F	Primer (PCR and sequencing)	ACCTGGCTTTTTAGCCATATCT
HEK293T_R	Primer (PCR and sequencing)	TTCAGCTTAGGAGCGAAGGC
THP1_F	Primer (PCR and sequencing)	TGTATGTACCTTTGTTGCCGT
THP1_R	Primer (PCR and sequencing)	CTCCAAAGAGTCAAAATCAACCA
THP1_R2	Primer (PCR)	CTTTGACCTGAGGGTAAGACT
S80A sense	ssODN HDR template	AGAGGATTTCGTTTCCGTTATGTATGTGAAG GCCCATCCCATGGTGGACTACCTGGTGCCTC TAGTGAAAAGAACAAGAAAGCTTATCCTCAG GTCAAAGTAAGTTTGTGGTAGCTCTCCTTCT ATTTGAATTCTGGAAATTTTGATTTCCTACG ATTTCCAAGGAATT
S80A antisense	ssODN HDR template	AATTCCTTGGAAATCGTAGGAAATCAAAATT TCCAGAATTCAAATAGAAGGAGAGCTACCAC AAACTTACTTTGACCTGAGGATAAGCTTTCT TGTTCTTTTCACTAGAGGCACCAGGTAGTCC ACCATGGGATGGGCCTTCACATACATAACGG AAACGAAATCCTCT
1F1R OffTar-1 F	Primer (PCR and sequencing)	TGGCTAGTTGCACATTGCTC
1F1R OffTar-1 R	Primer (PCR and sequencing)	TTTGGCAGAATTTCTCAGACTGC
1F1R OffTar-2 F	Primer (PCR and sequencing)	TGCAGTGTAGTCCACACAAACT
1F1R OffTar-2 R	Primer (PCR and sequencing)	TCGTTGAGTGCCTTCCAAGT
1F1R OffTar-3 F	Primer (PCR and sequencing)	GTTAATCCTGCCCCTCGCTC
1F1R OffTar-3 R	Primer (PCR and sequencing)	ACCCCAATTTTCATGCAGTGG
1F1R OffTar-4 F	Primer (PCR and sequencing)	ACCAAAGGTAGCAGGTAGAACC
1F1R OffTar-4 R	Primer (PCR and sequencing)	CCCCAGGAGTCAGCAGACATA
1F1R OffTar-5 F	Primer (PCR and sequencing)	GCTTCTGAGCACAGTTCCTTC
1F1R OffTar-5 R	Primer (PCR and sequencing)	GGGCATGTTCCTCTCATGGC

Table 2.16: Oligonucleotides used for CRISPR/Cas9 genome editing

2.2.4.2 CRISPR/Cas9 editing of HEK293T cells

A T150 flask of HEK293T cells (~80% confluent) were treated with 200ng/ml nocodazole for 16 hours. Following treatment, cells were washed twice with tissue culture media, trypsinised and cell density was determined. Cells were centrifuged at 300g for 5 minutes and supernatant was discarded. Cells were washed with Mg^{2+} and Ca^{2+} free PBS, centrifuged at 300g for 5 minutes and supernatant was discarded. HEK239T cells were transfected using the NEON® Transfection System 100µl kit (ThermoFisher Scientific), according to the manufacturer's instructions. Specific transfection conditions and parameters are described in Table 2.13. Electroporated cells were plated into a 10cm dish containing 10ml tissue culture media and were left to recover overnight. Transfected cells were selected for 3 days with 3µg/ml of puromycin. After selection, puromycin media was removed and HEK293T cells were washed and replaced with fresh tissue culture media. Selected cells were grown until confluent, and used for serial dilution in 96-well plates to isolate single cell clones.

2.2.4.3 CRISPR/Cas9 editing of THP-1 cells

The density of THP-1 cells grown in T75 flasks was determined before being pelleted by centrifugation at 300g for 5 minutes. Supernatant was discarded, and cells were washed with Mg^{2+} and Ca^{2+} free PBS. Cells were centrifuged at 300g for 5 minutes, and supernatant was discarded. THP-1 cells were transfected using the NEON® Transfection System 10µl kit (ThermoFisher Scientific), according to the manufacturer's instructions. Specific transfection conditions and parameters are described in Table 2.13. Cas9 nuclease protein and the synthetic gRNA were mixed and incubated at room temperature for 10 minutes to facilitate formation of ribonucleoprotein complexes (RNPs) before addition to the transfection reaction mixture. Electroporated cells were plated into a 24-well plate containing 500µl tissue culture media and were left to recover for 72 hours. To increase efficiency of identifying single cell clones, transfected cells were treated with 20µM CellTracker™ Green CMFDA dye (ThermoFisher Scientific) 24 hours prior to single cell sorting into 96-well plates by flow cytometry using a BD FACSAria III cell sorter.

2.2.4.4 Screening CRISPR/Cas9 clones

To screen single cell clones, genomic DNA was extracted from HEK293T and THP-1 cells using the DNeasy Blood & Tissue Kit (QIAGEN) according to the manufacturer's instruction. The target region of DNA was first analysed by PCR amplification using primers shown in Table 2.16 and methods described in section 2.2.7. Clones were screened for homology directed repair (HDR) by restriction digest with HindIII restriction enzyme, using methods described in 2.2.7. Clones positive for HDR were sent for Sanger sequencing (GATC-Biotech, Germany). *NFKB1*^{S80A} positive clones were expanded and cryopreserved in FBS supplemented with 10% DMSO.

2.2.5 Transcriptomic analysis

2.2.5.1 RNA sequencing (RNA-seq)

WT and *NFKB1*^{S80A} HEK293T or THP-1 cells were left untreated or were treated for 3 hours with 10ng/ml TNFα or LPS, respectively. Total RNA was extracted from cells using the RNeasy kit (QIAGEN), according to the manufacturer's instructions. Triplicate samples of each condition were submitted to the University of Glasgow Polyomics Facility for sample QC, polyA library preparation (Truseq stranded mRNA kit). Illumina NextSeq™ 500 platform was used to sequence single-end 75bp reads to a depth of 20 million.

2.2.5.2 Quantitative PCR (qPCR)

WT and *NFKB1*^{S80A} HEK293T or THP-1 cells were left untreated or were treated with 10ng/ml TNFα or LPS, respectively. Total RNA was extracted from using the RNeasy kit (QIAGEN), according to the manufacturer's instructions. Isolated RNA was primed with random hexamer oligonucleotides and reverse transcribed using Primer Design precision nanoScript2 reverse transcription kit to generate cDNA, according to the manufacturer's instructions. qPCR was performed with PerfeCTa® SYBR®Green FastMix®, ROX™ (Quantabio) and QuantiTect primers (QIAGEN) shown in Table 2.17. All data were normalised to TBP. Gene expression changes were calculated using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

Gene	Catalogue No.
BCL3	QT00012404
CCL2	QT00212730
CSF1	QT00035224
CXCL2	QT00013104
IL6	QT00083720
IL8	QT00000322
MAP3K8	QT00051730
NFKB1	QT00063791
TBP	QT00000721
TNF	QT00029162

Table 2.17: QuantiTect primers used for qPCR

2.2.6 Bioinformatics

2.2.6.1 RNA-seq analysis

2.2.6.1.1 CuffDiff pipeline

Quality control of raw sequence reads was determined using the Fast QC tool. Reads were aligned to human reference sequence iGenome NCBI GRCh38 using HISAT (v 2.0.3.2) (Kim et al., 2015). Aligned reads were assembled into transcripts using Stringtie (v 1.2.3) (Pertea et al., 2015). Significant differential fragments per kilobase per million reads (FPKM) values were calculated using CuffDiff (v. 2.2.1.3) (adjusted p-value <0.05) (Trapnell et al., 2010). The above analyses were performed using Glasgow University Galaxy server (http://heighliner.cvr.gla.ac.uk/). Clustering analysis was performed with default options (K-means clustering with Jensen-Shannon distance) using the Bioconductor package CummeRbund. Visualisation of RNA-seq data was performed using Papillon (v 0.1.1) developed by Dr Domenico Somma (University of Glasgow) (https://github.com/domenico-somma/Papillon).

2.2.6.1.2 DESeq2 pipeline

Quality control and read alignment was performed as described above (2.2.6.1.1). Aligned reads were counted using HTseq-count (v 0.6.1) (Anders et al., 2015) using Glasgow University Galaxy server. Read counts were used as input to SearchLight (Beta), an automated pipeline for analysis and visualisation of RNA-seq data, developed by Mr John Cole and colleagues (GLAZgo Discovery Centre, University of Glasgow). Significant differential gene expression was

calculated using DESeq2 (adjusted p-value <0.05 and <0.01 for HEK293T and THP-1 data sets, respectively) (Love et al., 2014).

2.2.6.1.3 Accession Numbers

RNA-seq data is available in the NCBI Gene Expression Omnibus database with the following accession numbers: HEK293T: GSE117279; THP-1: GSE120212

2.2.6.2 Transcription factor binding site (TFBS) analysis

2.2.6.2.1 Motif analysis

Motif analysis was based on the GREAT approach (McLean et al., 2010), which incorporates both proximal and distal regulatory regions for enrichment analyses. Regulatory regions for genes were obtained from the Ensembl Regulatory Build (Zerbino et al., 2015). Genomic locations were obtained using the "fetch closest non-overlapping feature" tool. JASPAR Position Weight Matrix MA0105.1 was used to identify the best matching NFKB1 binding sites in regulatory regions using the "FIMO" tool. These analyses were performed using the University of Glasgow Galaxy server (http://heighliner.cvr.gla.ac.uk/). Sequence logos were generated using WebLogo (http://weblogo.berkeley.edu/logo.cgi) (Crooks et al., 2004).

De novo motif searches were performed using the motif search program HOMER developed by C. Benner (Heinz et al., 2010), using the findMotifs.pl command with default parameters (promoter region 300bp upstream 50bp downstream of transcription start site). Custom motifs were generated using the seq2profile.pl command.

2.2.6.2.2 TFBS shape analysis

Transcription factor binding site shape analysis was performed using the TFBSshape online tool available at http://rohslab.cmb.usc.edu/TFBSshape/ (Yang et al., 2014). Minor grove width and propeller twist define base-pair parameters whereas roll and helix twist represent base pair step parameters.

2.2.6.3 Gene set enrichment analysis (GSEA)

Gene set enrichment analysis was performed using GSEA software in conjunction with the Molecular Signatures database (MSigDB) (Subramanian et al., 2005, Liberzon et al., 2011). Differentially expressed genes calculated by CuffDiff were ranked by adjusted p value, then by log2 fold change. The ranked gene list was analysed using GSEA Preranked with enrichment score set to “classic”.

2.2.7 Molecular biology techniques

2.2.7.1 Agarose gel electrophoresis

DNA samples were resolved by agarose gel electrophoresis, using the Mupid-One electrophoresis system (Eurogentect). Samples were diluted in 6X loading dye [New England Biolabs (NEB)] and were resolved on GelRed stained 1.5% agarose gels (Table 2.5) at 100-135V for 30 minutes in 0.5X TBE buffer (Table 2.3). Gels were visualised using a UV transilluminator (Alpha Innotech). DNA size was determined using 100bp ladder (Promega). If DNA was required for subsequent cloning steps, DNA was excised and purified using QIAquick gel extraction kit (QIAGEN), according to the manufacturer’s instructions.

2.2.7.2 Polymerase chain reaction (PCR)

Primers for PCR were designed using Primer3 software. Genomic DNA was amplified using HotStarTaq Master Mix Kit (QIAGEN) according to the manufacturer’s instructions. If amplified DNA was required for subsequent restriction digest DNA, PCR products were purified using QIAquick PCR purification kit (QIAGEN) according to the manufacturer’s instructions. The concentration of purified PCR products were measured using a NanoDrop spectrophotometer (ThermoFisher Scientific).

2.2.7.3 Restriction digest

Restriction digests were prepared by mixing either 1µg plasmid DNA or 500ng purified PCR product with the required enzymes (NEB), according to manufacturer’s instructions, in a total of volume of 20µl.

2.2.7.4 Ligation

Ligation reactions were prepared by mixing insert DNA and plasmid backbone and at a molar ratio of 5:1 with T4 DNA ligase (NEB) according to manufacturer's instructions, in a total of volume of 20µl. Ligation reactions were incubated for 1 hour at room temperature.

2.2.7.5 Plasmid preparation

2.2.7.5.1 DNA transformation

XL-1 Blue competent cells (Agilent technologies) were thawed on ice. Either 10ng of plasmid DNA or with 1µl of ligation reaction was mixed with 10µl cells, and incubated for 15 minutes on ice. Following incubation, cells were heat-shocked for 45 seconds at 42°C, and were returned to ice immediately. Cells were incubated on ice for a further 15 minutes before the addition of 100µl L-broth. Cells were plated onto lysogeny broth (LB) agar plates containing 100µg/ml ampicillin and were incubated overnight at 37°C.

2.2.7.5.2 Midiprep

A single bacterial colony was inoculated in 5ml L-broth containing 100µg/ml ampicillin to generate a starter culture. Starter cultures were incubated at 37°C shaking for ~8 hours. 100µl starter culture was used to inoculate 100ml L-broth containing 100µg/ml ampicillin, which was then incubated at 37°C shaking overnight. Plasmid DNA was extracted using Pure Yield Plasmid Midiprep System (Promega) according to the manufacturer's instructions. Plasmid concentration was measured using a NanoDrop spectrophotometer (ThermoFisher Scientific). Plasmid midipreps were utilised for transfection of cells.

2.2.7.5.3 Miniprep

A single bacterial colony was inoculated in 5ml L-broth containing 100µg/ml ampicillin and was incubated at 37°C shaking overnight. Plasmid DNA was extracted using QIAprep Spin Miniprep Kit (QIAGEN) according to the manufacturer's instructions. Plasmid miniprep DNA was utilised to check for successful cloning by DNA sequencing.

2.2.7.6 DNA sequencing

DNA sequencing of purified PCR products and plasmids were carried out by GATC Biotech, using either common universal primers provided by the company or custom primers (Table 2.16).

Chapter 3

3 Generation of knock-in *NFKB1*^{S80A} cell lines using CRISPR/Cas9 genome editing techniques

3.1 Abstract

The CRISPR/Cas9 system originally identified as a microbial adaptive immune system has been adapted to be used as a programmable tool for precision genome editing. Introducing targeted mutations into the genome provides means to generate effective models used to study molecular function. Here, CRISPR/Cas9 genome editing technology was utilised to generate knock-in *NFKB1*^{S80A} cell lines to investigate the role of NF-κB p50 S80 phosphorylation in the regulation of inflammatory gene expression. Plasmid-based and ribonucleoprotein complex methodologies were employed to efficiently deliver CRISPR/Cas9 components into HEK293T and THP-1 cells, respectively. *NFKB1*^{S80A} knock-in cells were confirmed by homology directed repair assays and Sanger sequencing.

3.2 Introduction

The ability to introduce specific targeted mutations into the genome has facilitated the rapid generation of animal and cell-based models used to study human disease. Although a number of genome editing techniques exist, such as zinc-finger nucleases (ZFN) (Kim et al., 1996, Bibikova et al., 2001) and transcription activator-like effector nucleases (TALENs) (Christian et al., 2010, Li et al., 2011), the CRISPR/Cas9 system is fast becoming the most utilised technique due to its simplicity, specificity, and ease of customisation.

The CRISPR-Cas9 system is a microbial adaptive immune system that protects the host against foreign nucleic acids such as plasmids or viruses (Mojica et al., 2005, Pourcel et al., 2005). Fragments from invading nucleic acids, called "protospacers" are incorporated into the bacterial genome in between clustered regularly interspaced short palindromic repeats (CRISPR). The large arrays formed are used to "remember" the DNA of foreign nucleic acids. When CRISPR arrays are transcribed, the transcripts are processed to generate small CRISPR RNAs (crRNAs), which contain sequences complementary to the invading DNA, derived from the protospacers (Brouns et al., 2008). Each crRNA must hybridise with a transactivating CRISPR RNA (tracrRNA) which acts as a scaffold for complex formation with Cas9 nuclease. The crRNA portion of the complex guides Cas9 to the complementary target DNA sequences. The endonuclease activity of Cas9 is dependent on a short conserved protospacer-associated motif (PAM) sequence NGG, where N is any base (Bolotin et al., 2005). Once bound to its target, the crRNA-tracrRNA-Cas9 complex cleaves the DNA 3bp upstream from the PAM site, destroying the invading DNA (Barrangou et al., 2007).

In 2013, multiple groups adapted the CRISPR/Cas9 system to enable its use as a programmable tool to facilitate genome engineering in eukaryotic cells (Cong et al., 2013, Jinek et al., 2012, Mali et al., 2013). The adapted system consists of two components: a "guide" RNA (gRNA) and Cas9 endonuclease. The gRNA is a fusion between the essential components of a tracrRNA and a crRNA. A 20bp sequence at the end of the gRNA (which corresponds to the protospacer portion of the crRNA) is responsible for directing Cas9 to its target site. Therefore, Cas9 endonuclease activity can be directed to any location in the genome which has the sequence N20-NGG.

The schematic shown in Figure 3.1 depicts how programmable Cas9 activity can be utilised for genome editing. Using a gRNA, Cas9 is targeted to a user defined region of the genome, where it creates double-strand breaks (DSBs) in the DNA. DSBs can induce two distinct DNA repair pathways: non-homologous end joining (NHEJ) or homology directed repair (HDR). During CRISPR/Cas9 genome editing, NHEJ is more frequently used by cell to repair DNA (Maruyama et al., 2015). While NHEJ is a highly efficient repair mechanism, it is error-prone, leading to a high frequency of insertion or deletion (indel) mutations. Indels within the coding region of a gene could potentially cause frameshift or nonsense mutations, resulting in a functional gene knockout. Alternatively, if a repair template is provided, DSBs can be processed via HDR. By designing a repair template containing desired sequence, specific genome editing can be achieved, including the creation of targeted knock-ins.

For small edits in the genome (<50bp), the optimal repair template is a single-stranded DNA oligonucleotide (ssODN) (Ran et al., 2013). The ssODN contains homology arms between 40-90bp which flank the desired mutation to be incorporated into the genome. Asymmetric homology arms have been shown to be slightly more efficient at facilitating HDR (Richardson et al., 2016). ssODNs can either be sense or antisense in orientation, relative to the target locus. A recent study has proposed that once Cas9 has been released from DNA, both sides of the double-strand break are recognised and resected by DNA repair machinery equally (Liang et al., 2017). This means that both sense and antisense ssODNs can facilitate HDR, with no preference towards a particular ssODN orientation. However, the precise mechanisms of Cas9 interaction with target DNA is not fully understood, and requires further study.

It has been reported that the efficiency of HDR-mediated genome editing can be increased by synchronising cells in the G2/M phase of the cell cycle, where eukaryotic cells undergo homologous recombination (Heyer et al., 2010). The antimitotic agent nocodazole has been shown to effectively synchronise cells, increasing HDR efficiency and donor template integration into the genome (Yang et al., 2016).

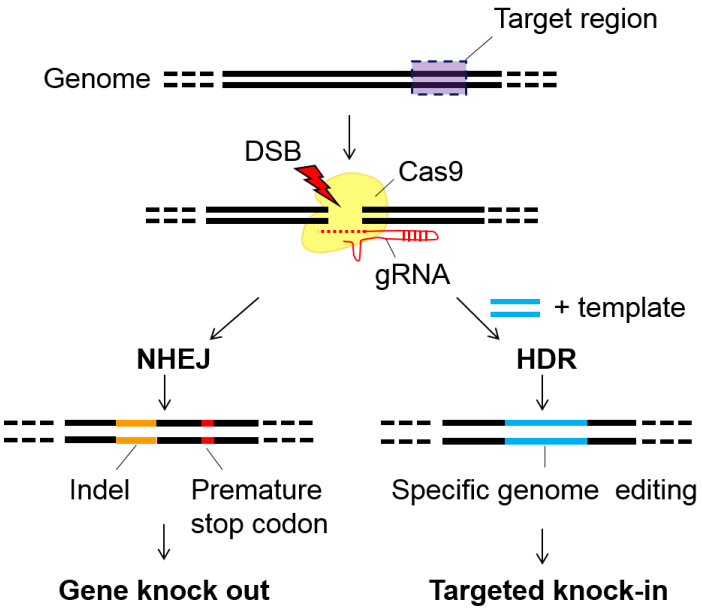


Figure 3.1: Overview of CRISPR/Cas9 genome editing
A target region is selected for genome editing. gRNAs (red) made up of a user defined 20bp protospacer (dotted red line) and a tracrRNA scaffold (solid red line) form a complex with Cas9 (yellow) which guides the endonuclease to the desired region of the genome. Cas9 induces double-strand breaks (DSBs) in the DNA, which can be repaired in one of two ways. In the error prone NHEJ pathway, DNA ends are re-joined which can result in indel mutations (orange). Indels within the coding region can cause a functional gene knockout. Alternatively, if a repair template is provided, DSBs can be processed via HDR. By designing a repair template with desired sequence (blue lines), specific genome editing can be achieved, including the creation of targeted knock-ins.

Although the CRISPR/Cas9 system presents a simple and robust method for targeted genome editing, there are a number of factors which can limit its efficiency and reliability in practise. One such factor is the delivery method of CRISPR components. Each model system has its own set of considerations to achieve efficient delivery of CRISPR components into the nucleus of cells. A common approach is the use of a mammalian expression vector, in which a promoter (typically U6), drives gRNA expression. This method is efficient for cells that are easy to transfect, such as HEK293Ts. However, not all cells efficiently take up plasmids required to express CRISPR components. For example, DNA transfection is not an effective method for expressing proteins in cells such as monocytes. Monocytic cells, such as THP-1 cells, are capable of sensing and degrading double-stranded DNA in the cytosol via the cGAS-STING and AIM2 inflammasome pathways (Cai et al., 2014). Therefore, CRISPR components are degraded before they reach the nucleus, inhibiting their ability to perform genome editing. To circumvent this, alternative methods have been developed for CRISPR/Cas9 editing of cells sensitive to nucleic acid transfection.

One example is the Cas9-gRNA ribonucleoprotein (RNP) complex method. Cas9 RNPs consist of purified Cas9 protein complexed with a gRNA which are delivered as functional complexes that evade cytosolic nucleic acid sensing pathways.

Another potential limiting factor affecting the CRISPR/Cas9 system is off-target activity. Although the targeting specificity of Cas9 is to the 20 nucleotide gRNA sequence, it has been shown that potential off-target cleavage can occur at DNA sequences similar to the gRNA (Jinek et al., 2012, Cong et al., 2013, Fu et al., 2013, Hsu et al., 2013, Mali et al., 2013). Cas9 can tolerate up to 3-5bp mismatches, however it is dependent on where the mismatches are in the gRNA sequence. Typically, if the mismatches are within the first 10-12bp adjacent to the PAM site, called the "seed sequence", off-target cleavage is less likely to occur. This is because the seed sequence determines Cas9 specificity (Jinek et al., 2012, Mali et al., 2013). It is therefore important to consider how off-target activity might affect the cell model that is to be generated. Off-target activity may cause unintended and unwanted mutations the genome. There are many tools available for predicting the most likely off-target sites for a given gRNA, which can easily be analysed by PCR amplification and DNA sequencing.

The overall aim of this chapter was to generate *NFKB1*^{S80A} human cell lines to investigate the role of NF-κB p50 S80 phosphorylation in the regulation of inflammation, by fulfilling the following objectives:

1. Design and validate CRISPR/Cas9 genome editing technology in HEK293T cells, using a plasmid-based method.

2. Utilise the Cas9-gRNA RNP method to generate *NFKB1*^{S80A} knock-in THP-1 cells.

3.3 Results

3.3.1 *NFKB1*^{S80A} knock-in genome editing strategy

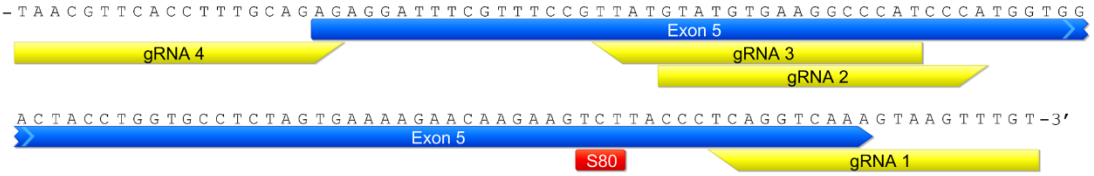
3.3.1.1 gRNA and ssODN repair template design

In order to generate *NFKB1*^{S80A} knock-in cell lines, the region of the human *NFKB1* locus containing the codon for S80 (exon 5) was selected as the target region for CRISPR/Cas9 designs. Potential gRNA targets were generated using the Broad Institute gRNA Designer tool, available online (https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design). This tool calculates an on-target efficacy score for each gRNA, indicating the likelihood of Cas9 cleavage at the target sequence. Scores are between 0 and 1, with a higher score representing higher efficacy. The position of the top four scoring gRNAs relative to *NFKB1* exon 5 are shown (Figure 3.2A). The sequence, orientation, PAM site and on-target score for each gRNA are shown (Figure 3.2B). Targeted homozygous mutations have been shown to occur more frequently if the distance of the gRNA cut site relative to the site of the desired mutation is <10bp (Paquet et al., 2016). For this reason, gRNA 1 was selected for CRISPR/Cas9 genome editing due to its proximity to the codon for S80, and highest on-target efficacy score.

To analyse any potential off-target effects of the gRNA 1, the Broad Institute Benchling online tool was utilised (https://benchling.com/crispr). This tool identifies potential off-target cutting sites throughout the genome for a selected gRNA, which are ranked based on score. Scores are between 0 and 100, with a higher score representing higher Cas9 efficacy at a given sequence. The top 5 off-target sites identified for gRNA 1 are shown in Figure 3.2C. It was found that the likelihood of Cas9 cleavage at off-target sites is low, indicated by the low scores. This is due to mismatches between the off-target site and the seed sequence region of gRNA 1, which reduces Cas9 recognition of these sites. In addition, off-target sites were found to be intergenic or intronic (Figure 3.2C). Despite low efficacy scores, if Cas9 nuclease were to cut at these sites, the mutations implemented would be unlikely to cause deleterious effects to cells. Screening for off-target mutations in both CRISPR/Cas9 cell lines is reported in section 3.3.2.4 and 3.3.3.2 for HEK293T and THP-1 *NFKB1*^{S80A} cells, respectively.

A single-stranded DNA oligonucleotide (ssODN) was designed to act as a template for HDR to introduce synonymous codon changes into the genome. The codon changes were designed to mutate serine 80 to alanine; to introduce a unique HindIII restriction site to facilitate CRISPR/Cas9 clone screening and to mutate the PAM site to prevent further Cas9 cleavage after the desired mutations introduced (Figure 3.2D). The ssODN was designed in both sense and antisense orientation in order to optimise knock-in efficiency.

A



B

gRNA	Sequence	Strand	PAM	On-target Score
1	ACAAACTTACTTTGACCTGA	Antisense	GGG	0.61
2	GTATGTGAAGGCCCATCCCA	Sense	TGG	0.47
3	ATGGGCCTTCACATACATAA	Antisense	CGG	0.23
4	TAACGTTCACCTTTGCAGAG	Sense	AGG	0.34

C

Sequence	Strand	PAM	Score	Region	Locus
ACAAATATACTTTGACCTGA	Sense	AGG	2.16	*FMN1* (intron)	Chr15:33065362-33065381
ACAATCTTACTTTGAACTGA	Antisense	AGG	1.60	Intergenic	Chr7:91918316-91918335
TCAGAATTACTTTGACCTGA	Sense	AAG	1.56	*NUAK1* (intron)	Chr12:106076161-106076180
AGAAACTCACTTTGACTTGA	Antisense	AAG	1.48	*TOX* (intron)	Chr8:58816728-58816747
AACAGCTCACTTTGACCTGA	Sense	GAG	1.43	Intergenic	Chr4:189579679-189579698

D

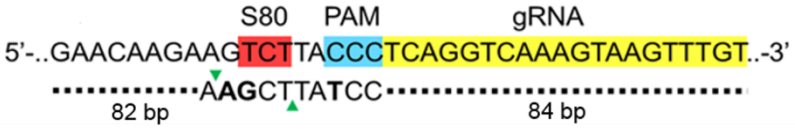


Figure 3.2: CRISPR/Cas9 editing strategy
(A) Schematic of human *NFKB1* exon 5 showing the position of four gRNAs designed (highlighted yellow) relative to the S80 codon (highlighted red). (B) The sequence, strand, PAM site and on-target efficacy score for each gRNA (score range 0-1) are shown. (C) The top five scoring off-target sites of gRNA 1 (score range 0-100) are shown. Mismatches between the sequence of gRNA 1 and each off-target site are underlined. The strand orientation, PAM site and genomic coordinates (GRCh38/hg38) are indicated for each off-target site. (D) Schematic of the ssODN oligo design, relative to the S80 codon (red), the selected gRNA (yellow) and PAM site (blue). Homology arms (dotted lines) and synonymous codon changes (bold) are shown. The HindIII restriction site is indicated (green triangles).

3.3.2 CRISPR/Cas9 editing in HEK293T cells

3.3.2.1 gRNA/Cas9 plasmid creation and validation

A plasmid-based approach was utilised for CRISPR/Cas9 genome editing of HEK293T cells, due to the high transfection efficiency of these cells. The selected gRNA was synthesised as two complementary oligonucleotides which were annealed, and cloned into the CRISPR guide expression plasmid PX458, creating a Cas9/gRNA plasmid (Figure 3.3A). In order to determine the ability of the Cas9/gRNA plasmid to create targeted DSBs, a Surveyor assay was performed (schematic shown in Figure 3.3B). This assay detects and cleaves single base mismatches or small indels in DNA caused by DSBs created by Cas9. The Cas9/gRNA plasmid was co-transfected into HEK293T cells with a puromycin selection plasmid. Untransfected cells were used as a negative control. Transfected cells were selected in puromycin for 3 days. Following puromycin selection, cells were grown in culture with fresh media. DNA was extracted from selected cells, and the region flanking the cut site was amplified by PCR using primers HEK293T_F and HEK293T_R (Table 2.16 and Figure 7.1). PCR products were purified and incubated with Surveyor nuclease, and resolved on an agarose gel. DNA from untransfected HEK293T cells was used as a negative control. The gRNA/Cas9 plasmid was found to create DSBs at the target site, indicated by cleavage of the purified PCR product from transfected cells compared to untransfected cells (Figure 3.3C). Two digestion products are expected (310bp and 126bp). However, the low molecular weight fragment is too faint to be visualised at a low exposure. To visualise the second product, gel exposure was increased and the gel image has been cropped to show the 126bp fragment (Figure 3.3C).

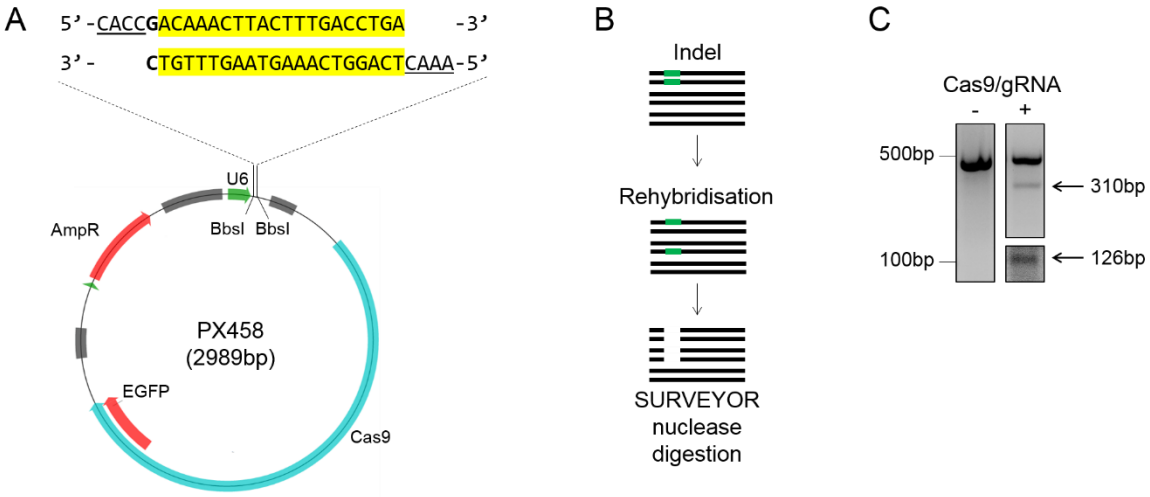


Figure 3.3: Cas9/gRNA CRISPR plasmid cloning and validation
(A) Cloning strategy for the creation of the gRNA/Cas9 plasmid used for CRISPR/Cas9 editing in HEK293T cells. Complementary gRNA oligonucleotides (highlighted yellow) were synthesised with 5'-BbsI overhangs (underlined) to facilitate cloning into BbsI digested PX458. The U6 promoter requires a G base at the transcriptional start site, therefore an additional G base (bold) was added to the gRNA. (B) Surveyor assay schematic. DNA from gRNA/Cas9 transfected cells is denatured and reannealed, forming heteroduplexes due to mismatched DNA caused by indels (green). Heteroduplexes are cleaved by Surveyor nuclease, whereas homoduplexes are not. (C) DNA extracted from untransfected (-) and gRNA/Cas9 transfected cells (+) and was amplified by PCR and purified. Purified PCR products (436bp) were digested by Surveyor nuclease, and resolved on a 1.5% agarose gel. The 310bp and 126bp fragments visualised by different exposures are indicated by arrows.

3.3.2.2 HDR efficiency optimisation

Having confirmed gRNA/Cas9 successfully cleaves the target region, the next step was to determine the ability of the ssODN to mediate HDR. HEK293T cells were co-transfected with Cas9/gRNA plasmid, puromycin selection plasmid and either the sense or antisense ssODN. Transfected cells were selected for 3 days in puromycin. Following puromycin selection, cells were grown in culture with fresh media. DNA was extracted from puromycin selected cells and amplified by PCR. Purified PCR products were digested with HindIII restriction enzyme, and resolved on an agarose gel. Since the HindIII restriction site is unique to the repair template, digested PCR products indicate successful HDR facilitated by the ssODN. It was found that the antisense ssODN was able to facilitate HDR, indicated by HindIII digestion of the purified PCR product (Figure 3.4A). However, two digestion products are expected (307bp and 129bp). Due to the low abundance of the lower molecular weight product, the 129bp fragment is too faint to be visualised.

The sense ssODN did not appear to mediate HDR indicated by the lack of HindIII digestion products (Figure 3.4A). The efficiency of HDR is proportional to the amount of digested PCR product, determined by the intensity of gel bands. Although the antisense ssODN is capable of mediating HDR, HindIII digestion products are of low intensity, and are therefore not visible (Figure 3.4A). This suggests that HDR efficiency is low. It was therefore desirable to enhance the efficiency of HDR in order to increase the probability of identifying *NFKB1*^{S80A} knock-in HEK293T cells.

The use of nocodazole for synchronising HEK293T cells to improve HDR efficiency using the antisense ssODN was investigated. HEK293T cells were treated with nocodazole 16 hours prior to transfection. Treated cells were co-transfected with Cas9/gRNA plasmid, puromycin plasmid and antisense ssODN. Transfected cells were selected with puromycin for 3 days. Following puromycin selection, cells were grown in culture in fresh media. DNA was extracted from puromycin selected cells and amplified by PCR. Purified PCR products were digested with HindIII restriction enzyme to assay for HDR. DNA from untransfected cells was used as a negative control. This revealed that HEK293T cells treated with nocodazole increased HDR efficiency, indicated by the increased levels of HindIII digestion products (Figure 3.4B) compared to unsynchronised cells (Figure 3.4A).

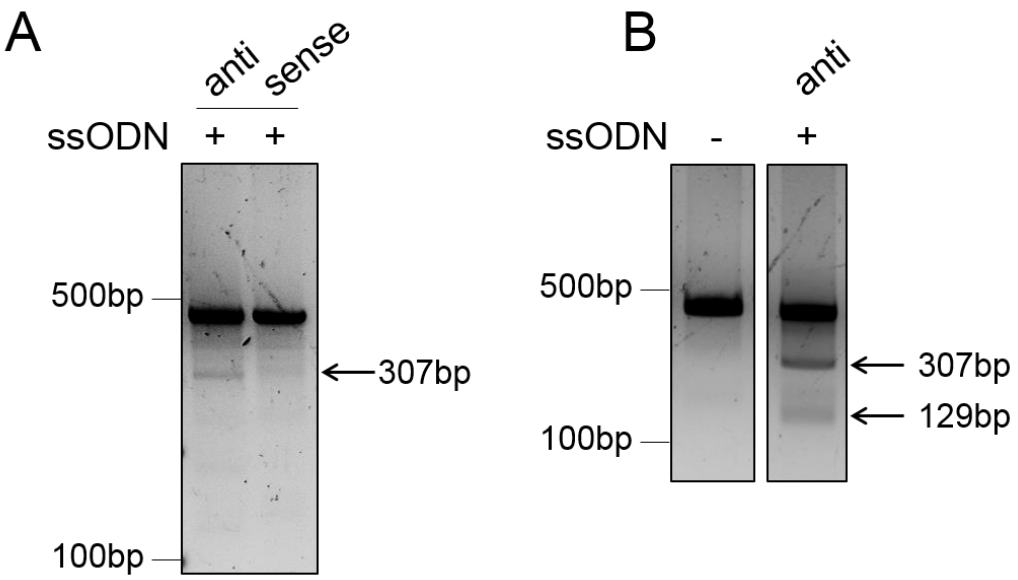


Figure 3.4: ssODN validation and HDR optimisation
(A) HEK293T cells were co-transfected with gRNA/Cas9 plasmid, puromycin plasmid and either sense or antisense ssODN as indicated. DNA was extracted and amplified by PCR. Purified PCR products (436bp) were digested with HindIII restriction enzyme, and resolved on a 1.5% agarose gel. One visible cleavage product (307bp) is indicated by an arrow. (B) HEK293T cells were treated with nocodazole 16 hours pior to transfection with gRNA/Cas9 plasmid and puromycin plasmid with or without antisense ssODN as indicated. DNA was extracted and amplified by PCR. Purified PCR products (436bp) were digested HindIII restriction enzyme, and resolved on a 1.5% agarose gel. Expected cleavage products (307bp and 129bp) are indicated by arrows.

3.3.2.3 Generation of *NFKB1*^{S80A} knock-in HEK293T cells

Having increased the efficiency of antisense ssODN-mediated HDR, transfection of G2/M synchronised HEK293T cells was repeated as described in section 3.3.2.2. Following puromycin selection for 3 days, serial dilution of transfected HEK293T cells was performed in order to isolate single clones. Single clones were expanded in culture. HEK293T clones were split for either maintenance in culture or were harvested for screening. To screen for potential *NFKB1*^{S80A} knock-in clones, DNA was extracted and the region flanking the target site was amplified by PCR. An example of PCR results obtained from clone screening is shown in Figure 3.5A. Clones with PCR products that contain indels are excluded from knock-in screening, since Cas9 indels have not been corrected by HDR. Examples of excluded clones (10 and 16) are shown in Figure 3.5A. Clones 10 and 16 are likely to be heterozygous for an insertion and deletion mutation, indicated by larger and smaller PCR products, respectively (Figure 3.5A). Clones with PCR products of expected size (463bp) are either unedited, or have been edited by HDR. These clones were screened for HDR by digesting purified PCR

products with HindIII. An example of the HDR assay results obtained from clone screening is shown in Figure 3.5B. Two clones (5 and 9) were found to be positive for HDR, indicated by HindIII digestion products of expected sizes (Figure 3.5B). PCR products from clones 5 and 9 are completely digested, indicating these clones are homozygous *NFKB1*^{S80A} knock-ins. In order to confirm the targeted knock-in mutation, clones 5 and 9 were sent for Sanger sequencing. Although the HDR assay suggested clone 5 was positive for the S80A knock-in mutation (Figure 3.5B), Sanger sequencing revealed that this clone did not contain the intended mutations (Figure 3.5C). Sanger sequencing revealed mutliple peaks in the DNA sequence chromatogram, indicating that multiple sequences are present (Figure 3.5C). Although not visible on the DNA sequence chromatogram, it is possible that clone 5 contains an allele carrying the HindIII restriction site. Peaks may have been incorrectly called due to the presence of multiple sequences, masking the sequence of allele containing the restriction site. However, it is unknown why other alleles which do not contain the HindIII restriction site were not detected in the HDR assay (Figure 3.5B). Sanger sequencing revealed that clone 9 is homozygous for the targeted knock-in mutation, confirming the generation of *NFKB1*^{S80A} knock-in HEK293T cells.

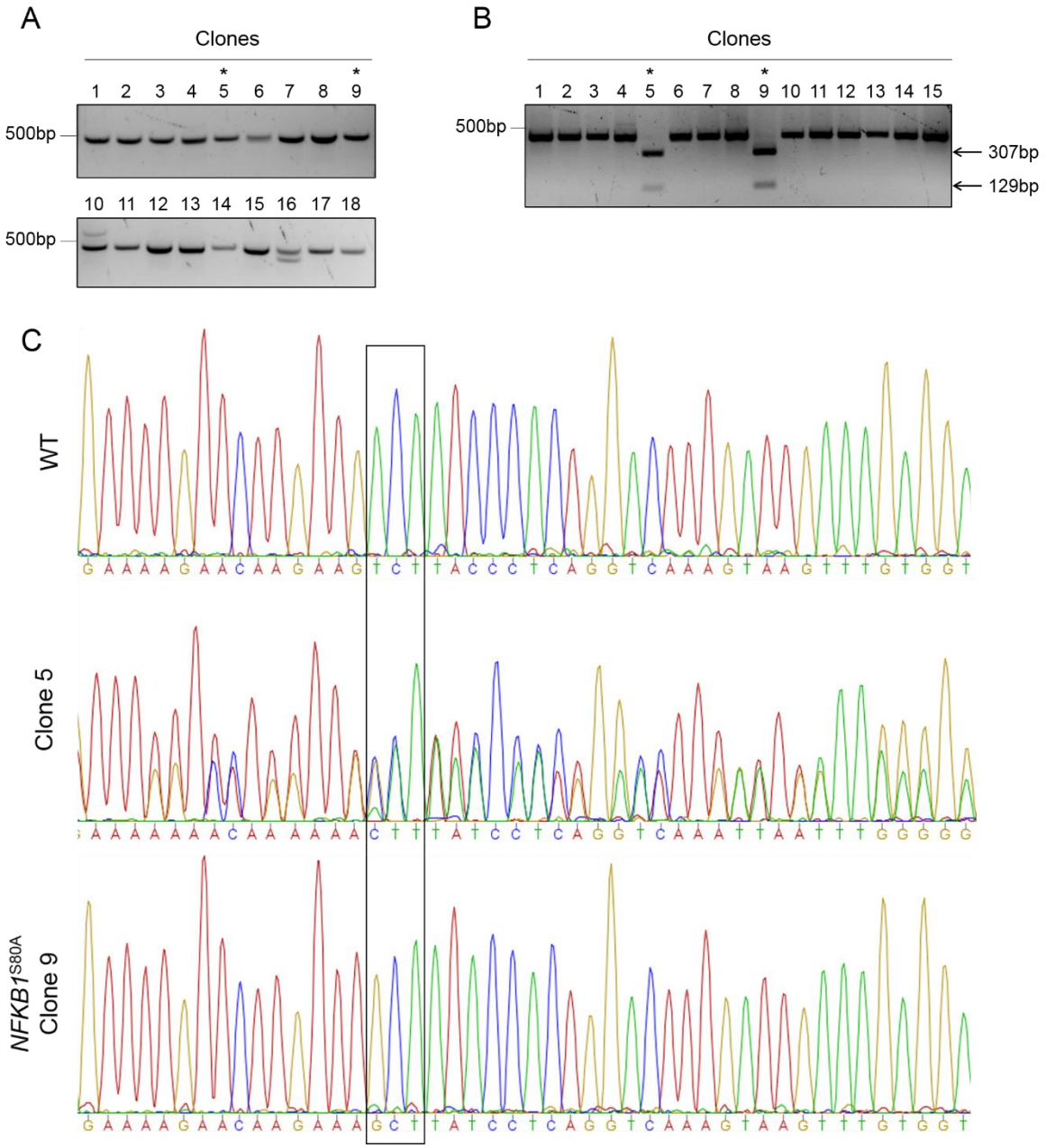


Figure 3.5: Screening and identification of *NFKB1*^{S80A} positive HEK293T cells
Clones marked with an asterisk in (A) identify which clones were shown to be homozygous for HDR in (B). (A) Example of PCR screening of clones. DNA was extracted from clones and amplified by PCR. PCR products were resolved on a 1.5% agarose gel. Clones 10 and 16 were excluded from HDR screening due to presence of indels. (B) Clones were screened for HDR by HindIII restriction digestion of purified PCR products (436bp). Digestion products were resolved on a 1.5% agarose gel. Clones 5 and 9 identified as potential *NFKB1*^{S80A} positive clones are marked with asterisk. The expected cleavage products (307bp and 129bp) are indicated by arrows. (C) DNA sequencing chromatograms of WT, clone 5 and *NFKB1*^{S80A} clone 9. The position of the codon for S80 is highlighted (black box).

3.3.2.4 Off-target analysis of *NFKB1*^{S80A} HEK293T cells

In order to detect any potential off-target mutations created during CRISPR/Cas9 genome editing, the top 5 off-target sites identified during gRNA design (Figure 3.2C) were analysed in *NFKB1*^{S80A} knock-in HEK293T cells. PCR primers were designed (Table 2.16) to amplify each off-target site to detect any Cas9 cleavage. DNA was extracted from WT and *NFKB1*^{S80A} HEK293T cells, amplified by PCR, and was analysed by Sanger sequencing. The aligned DNA sequencing reads shown in

Figure 3.6 demonstrate that no mutations were detected at the 5 potential off-target sites tested.

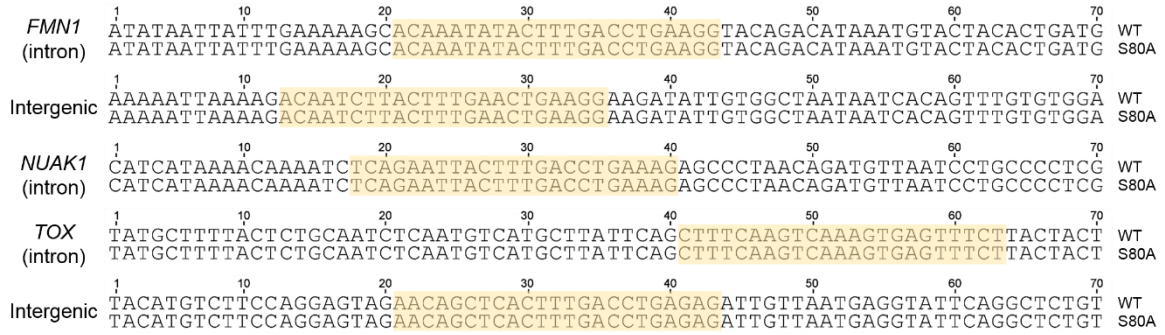


Figure 3.6: Off-target site analysis of knock-in *NFKB1*^{S80A} HEK293T cells
Each off-target site was amplified by PCR, and products were analysed by Sanger sequencing. Aligned DNA sequence reads from WT and *NFKB1*^{S80A} HEK293T cells are shown with the off-target site highlighted (yellow). (For genomic coordinates, refer to Figure 3.2C).

3.3.3 CRISPR/Cas9 editing in THP-1 cells

3.3.3.1 Generation of *NFKB1*^{S80A} knock-in THP-1 cells

In order to create *NFKB1*^{S80A} knock-in THP-1 cells, the Cas9/gRNA ribonucleoprotein (RNP) CRISPR/Cas9 delivery method was employed. A synthetic gRNA (sgRNA) with the same sequence gRNA selected in section 3.3.1 was synthesised. In order to validate the Cas9/gRNA RNP method as a successful genome editing strategy for THP-1 cells, transfected THP-1 cells were assayed for HDR. Purified Cas9 protein was incubated with the sgRNA to form a Cas9/gRNA ribonucleoprotein complex. The RNP complex was co-transfected into THP-1 cells with either sense or antisense ssODN. DNA was extracted, and the region flanking the target site was amplified by PCR using primers THP1_F and THP1_R (Table 2.16 and Figure 7.1). Purified PCR products digested with HindIII restriction enzyme and resolved on an agarose gel. DNA from untransfected THP-1 cells was used as a negative control. It was found that the sense ssODN was able to facilitate HDR, indicated by the presence of a HindIII digestion product (Figure 3.7A). Two cleavage products are expected (351bp and 154bp). However, due to the low abundance of the lower molecular weight product, the 154bp fragment is not visible on the agarose gel. The antisense ssODN did not appear to mediate HDR indicated by the lack of HindIII digestion products (Figure 3.7A). This result confirms that the Cas9/gRNA RNP method is capable of creating targeted knock-in mutations in THP-1 cells.

Transfection of THP-1 cells using the sense ssODN was repeated, and single clones were isolated by flow cytometry, using a single cell sorter. Single clones were expanded in culture. THP-1 clones were split for either maintenance in culture or were harvested for screening. To screen for potential *NFKB1*^{S80A} knock-in clones, DNA was extracted and the target region was amplified by PCR using primers HEK293T_F and THP1_R2 (Table 2.16 and Figure 7.1). The THP1_R2 primer spans the region of exon 5 where synonymous codon changes would be introduced if targeted knock-in mutation was successful (Figure 7.1). Due to the introduced DNA mismatches, the THP1_R2 primer will not anneal to the target DNA, and no amplification of DNA will occur. Therefore, clones that do not give a PCR product are included for further HDR screening. An example of PCR results obtained from clone screening is shown in Figure 3.7B. Examples of clones

included for HDR screening are marked with an asterisk. Potential *NFKB1*^{S80A} knock-in clones were screened for HDR by HindIII digest of purified PCR products, amplified by THP1_F and THP1_R primers (Table 2.16 and Figure 7.1). An example of the results obtained from HDR clone screening is shown in Figure 3.7C. Clone 3 is likely to be heterozygous for the targeted knock-in mutation, indicated by presence of both WT and HindIII digested PCR products. Clone 10 is homozygous for the *NFKB1*^{S80A} mutation, indicated by complete digestion of PCR products by HindIII (Figure 3.7C). The homozygous targeted knock-in mutation was confirmed by Sanger sequencing (Figure 3.7D). These results confirm the successful generation of *NFKB1*^{S80A} knock-in THP-1 cells.

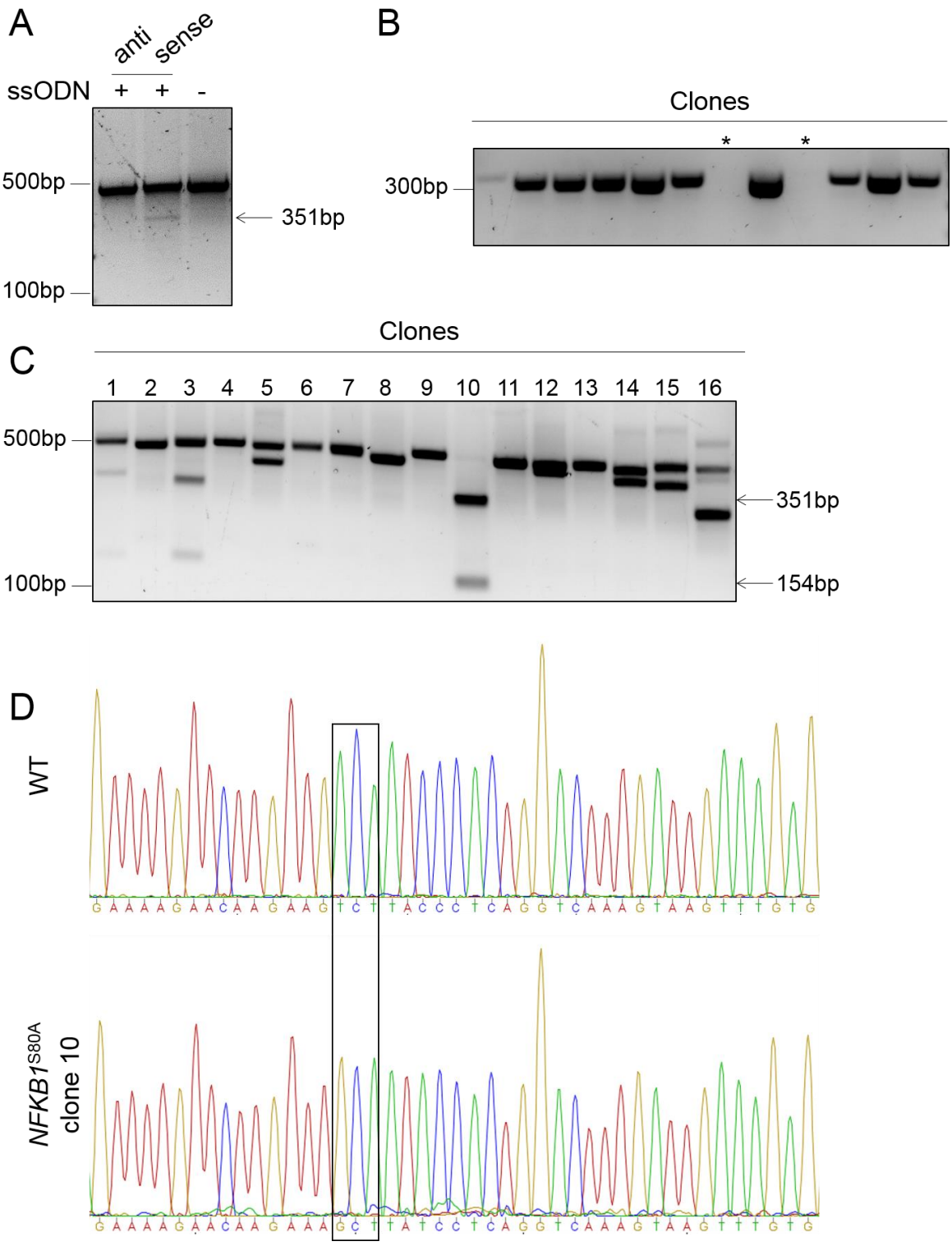


Figure 3.7: Screening and identification of *NFKB1*^{S80A} positive THP-1 cells
(A) THP-1 cells were co-transfected with Cas9/gRNA RNP complex, and either sense or antisense ssODN as indicated. DNA was extracted and amplified by PCR. Purified PCR products (505bp) were digested using HindIII restriction enzyme, and resolved on a 1.5% agarose gel. One expected cleavage product (351bp) is indicated by an arrow (B) Example of PCR screening of clones. DNA was extracted from clones and amplified by PCR (327bp). Products were resolved on a 1.5% agarose gel. Examples of clones included for HDR screening are marked with asterisk. (C) Clones were screened for HDR by HindIII restriction digestion of purified PCR products (505bp). Digestion products were resolved on a 1.5% agarose gel. The *NFKB1*^{S80A} positive clone is labelled (clone 10). The expected cleavage products (351bp and 154bp) are indicated by arrows. (D) DNA sequencing chromatograms of WT and *NFKB1*^{S80A} THP-1 cells highlighting the position of the codon for S80 (black box).

3.3.3.2 Off-target analysis of *NFKB1*^{S80A} THP-1 cells

Off-target analysis was performed as described in section 3.3.2.4 with DNA extracted from WT and *NFKB1*^{S80A} THP-1 cells. The aligned DNA sequencing reads shown in

Figure 3.8 demonstrate that no mutations were detected at the 5 potential off-target sites tested.

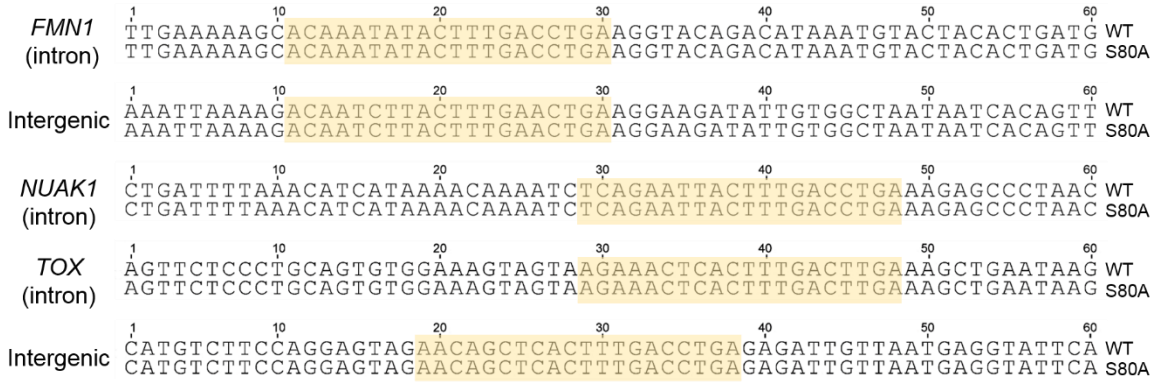


Figure 3.8: Off-target site analysis of knock-in *NFKB1*^{S80A} THP-1 cells
Each off-target site was amplified by PCR, and products were analysed by Sanger sequencing. Aligned DNA sequence reads from WT and *NFKB1*^{S80A} THP-1 cells are shown with the off-target site highlighted (yellow). (For genomic coordinates, refer to Figure 3.2C).

3.4 Discussion

In this chapter, the successful generation of knock-in HEK293T and THP-1 *NFKB1*^{S80A} cell lines using utilising two different CRISPR/Cas9 genome editing methods is reported. The gRNA designed and selected to target Cas9 to exon 5 of the *NFKB1* locus was shown to successfully introduce DSBs to the target region (Figure 3.3C). Although the ability of the gRNA/CRISPR plasmid to create DSBs was only assayed in HEK293T cells, it was assumed that it would be functional in THP-1 cells, since the target region is also human sequence. The ability of a sense and antisense ssODN repair template to fix Cas9-induced DSBs by HDR was assayed. Initially, the efficiency of HDR mediated by the antisense ssODN in HEK293T cells was low (Figure 3.4A). This was surprising, given that HEK293T cells are easily transfected and are a suitable model for plasmid-based CRISPR/Cas9 expression systems. HDR efficiency was increased by treating HEK293T cells with the G2/M phase synchroniser nocodazole, as measured by HDR assay (Figure 3.4B). DNA sequencing analysis of HDR positive clones revealed the successful incorporation of the targeted *NFKB1*^{S80A} mutation in both HEK293T cells and THP-1 cells (Figure 3.5C and Figure 3.7D).

Off-target site analysis for both *NFKB1*^{S80A} cell lines was performed. No mutations were detected at the top 5 off-target sites selected for analysis. It is important to consider that there may be off-target mutation elsewhere in the genome that are not included on predicted lists. Full genome sequencing could be performed to completely verify the absence of off-target mutation.

The generation of *NFKB1*^{S80A} human cell lines has created novel tools for investigating the importance of p50 phosphorylation for NF-κB activity. Both *NFKB1*^{S80A} HEK293T and THP-1 cells will be utilised to study the role of S80 phosphorylation in the regulation of target gene expression in an endogenous manner. These cell lines provide ways to study NF-κB driven transcription in two different contexts: *NFKB1*^{S80A} HEK293T cells provide a versatile model that can be utilised for a number of subsequent experiments due to their ease of culture, scalability and high transfection efficiency. *NFKB1*^{S80A} THP-1 cells however, provide a human macrophage model to study S80 phosphorylation in an inflammatory context. *NFKB1*^{S80A} HEK293T and THP-1 cells are utilised in Chapter 4 and 5 respectively, to investigate the role of NF-κB p50 S80 phosphorylation.

Chapter 4

4 The regulation of sequence specific NF-κB DNA binding and transcription by IKKβ phosphorylation of NF-κB p50 at serine 80

4.1 Abstract

Phosphorylation of the NF-κB transcription factor is an important regulatory mechanism for the control of transcription. Here we identify serine 80 (S80) as a novel phosphorylation site on the p50 subunit of NF-κB, and IKKß as a novel p50 kinase. Transcriptomic analysis of cells expressing a p50 S80A mutant reveal a critical role for S80 in selectively regulating the TNFα inducible expression of a subset of NF-κB target genes including pro-inflammatory cytokines and chemokines. S80 phosphorylation regulates the binding of p50 to NF-κB binding (κB) sites in a sequence specific manner. Specifically, phosphorylation of S80 reduces the binding of p50 at κB sites with an adenine at the -1 position. Our analyses demonstrate that p50 S80 phosphorylation predominantly regulates transcription through the p50:p65 heterodimer, where S80 phosphorylation acts in *trans* to limit the NF-κB mediated transcription of pro-inflammatory genes. The regulation of a functional class of pro-inflammatory genes by the interaction of S80 phosphorylated p50 with a specific κB sequence describes a novel mechanism for the control of cytokine-induced transcriptional responses.

4.2 Introduction

The transcription factor NF-κB plays an important role in a number of fundamental biological processes including cell cycle, proliferation, differentiation and cell death (Hayden and Ghosh, 2012). However, the primary role of NF-κB is as an essential regulator of the immune response through the transcriptional regulation of a large number of inflammatory genes, including chemokines, cytokines and immune effectors (Mitchell and Carmody, 2018). The NF-κB transcription factor family is comprised of five structurally related subunits: RelA (p65), RelB, c-Rel, p50 and p52. The p50 and p52 subunits are generated from the limited proteasomal processing of the precursor proteins p105 and p100 respectively, and lack the transactivation domain (TAD) found in the C terminal regions of the RelA, c-Rel and RelB subunits. All NF-κB subunits contain a highly conserved Rel homology domain (RHD) which facilitates dimerisation and DNA binding. NF-κB can promote or repress transcription depending on the subunit composition of dimer complexes. For example, although the p50 subunit lacks a TAD, it can positively regulate transcription by forming a heterodimer with a TAD containing subunit such as p65. Alternatively, p50 homodimers may function as transcriptional repressors by competing with TAD containing NF-κB dimers for the same DNA binding sites in target gene promoters (Carmody et al., 2007).

The primary mechanism regulating NF-κB activity is the cytoplasmic sequestration of NF-κB dimers by the canonical IκB proteins IκB -α, -ß and -ε, and the p105 and p100 precursor proteins. Activation of NF-κB requires the proteasomal degradation of the IκB proteins triggered by IKK complex (IKKα and IKKß) mediated phosphorylation IκB. The degradation of IκB proteins facilitates the nuclear translocation of NF-κB dimers where they bind to specific κB sites in DNA with the consensus sequence 5'-$G^{-5}G^{-4}G^{-3}R^{-2}N^{-1}W^{0}Y^{+1}Y^{+2}C^{+3}C^{+4}$-3' (R represents a purine, N represents any nucleic acid, W represents an A or T and Y represents a pyrimidine) (Sen and Baltimore, 1986, Lenardo et al., 1989). Although the nuclear localisation of NF-κB is controlled by IκB protein degradation, NF-κB transcriptional activity is regulated by a number of post-translational modifications, including acetylation (Greene and Chen, 2004), ubiquitination (Collins et al., 2016) and phosphorylation (Christian et al., 2016). The importance of phosphorylation in regulating NF-κB transcriptional activity

has been revealed mainly by studies of the p65 subunit, where phosphorylation has been demonstrated to regulate transcription in a stimulus and gene specific manner through a variety of mechanisms including the modulation of p65 interaction with IκBα and other transcription factors, and regulating p65 ubiquitination and stability (Christian et al., 2016).

Although the NF-κB p50 subunit is a critical regulator of inflammatory gene expression, its regulation by phosphorylation is much less well understood. p50 is one of the most highly expressed transcription factors in macrophages, and is central to macrophage mediated inflammatory responses (Xue et al., 2014). p50 homodimers are important repressors of inflammatory gene expression and the stability of p50 homodimers is crucial for limiting pro-inflammatory gene expression and establishing Toll-like Receptor tolerance in macrophages (Carmody et al., 2007, Yan et al., 2012). The phosphorylation of p50 at serine 337 (S337) is required for DNA binding (Hou et al., 2003), while the phosphorylation of S242 inhibits p50 homodimer DNA binding (Vonderach et al., 2018). Phosphorylation of S20 promotes DNA binding, and is required for VCAM1 expression in response to TNFα (Ju et al., 2010). p50 phosphorylation at S328 occurs in response to DNA damage and regulates the interaction of p50 with specific NF-κB binding sites to inhibit anti-apoptotic gene expression (Crawley et al., 2013, Schmitt et al., 2011). These studies indicate that transcriptional responses to specific stimuli may be shaped by the integration of signal induced NF-κB phosphorylation and binding site sequence in the regulatory elements of target genes.

In this study, we describe serine 80 as a novel phosphorylation site on the NF-κB p50 subunit. We identify IKKβ as a S80 kinase and establish p50 as a novel substrate for this kinase. Our data reveals that TNFα-induced phosphorylation of S80 selectively regulates distinct subsets of NF-κB target genes, driven by differential binding of p50 and p65 at specific DNA sequences. Our analyses demonstrate that p50 S80 phosphorylation predominantly regulates transcription through p50:p65 heterodimers and shows that p50 phosphorylation may function in *trans* to inhibit gene transcription. Our analysis demonstrates that S80 phosphorylation reduces the affinity of p50 for κB sites that have an adenine at the -1 position, limiting the expression of genes regulated by these binding sites.

Remarkably, the promoters of inflammatory cytokines and chemokines are enriched in κB sites containing a -1A and are thereby selectively regulated by phosphorylation of p50 at S80. Our data establishes IKKβ phosphorylation of p50 S80 as a novel mechanism that shapes the TNFα-induced transcriptional programme and demonstrates the control of gene expression through the interaction of NF-κB phosphorylation and DNA binding site sequence.

4.3 Materials and methods

4.3.1 Cell culture and transfection

Human embryonic kidney 293T (HEK293T) cells were cultured in DMEM (Sigma), supplemented with 10% fetal bovine serum, 2mM L-glutamine, and 100 units/ml of streptomycin and penicillin. Cells were maintained at 37ºC in a humidified environment with 5% CO_2 and sub-cultured by enzymatic detachment with 0.05% Trypsin-EDTA solution (Invitrogen). HEK293T cells were transiently transfected with Turbofect (ThermoFisher Scientific) according to manufacturer's instructions. pFLAG-CMV2 IKKß, pFLAG-CMV2 IKKß^{K44M}, and pcDNA3 IKKß^{SSEE} FLAG were kind gifts from Michael May, University of Pennsylvania. pEF4a-p50-XP and pEF4a-p50-Myc were generated by cloning murine cDNA into pEF4a empty vector. pEF4a-p50^{S80A}-XP was generated by site directed mutagenesis using the Stratagene QuikChange II SDM kit according to the manufacturer's instructions (Primers: 5'-CCTCTAGTGAGAAGAACAAGAAAGCCTACCCACAGG-3', 3'-GGAGATCACTCTTCTTGTTCTTTCGGATGGGTGTCC-5'). For generation of CRISPR/Cas9 NFKB1S80A knock-in cells, HEK293Ts were transfected using the Neon® Transfection System 100µl kit (ThermoFisher Scientific) using 2 pulses at 1100v for 20ms.

4.3.2 Western blot analysis and immunoprecipitation

Whole cell lysates were generated from cells suspended in radioimmune precipitation assay buffer (RIPA) containing 50mM Tris-HCl (pH 7.4), 1% Nonidet P-40, 0.25% deoxycholate,150mM NaCl, 1mM EDTA, 1mM PMSF, 1mM NaF, 1mM Na_3VO_4, 2µg/ml aprotinin, 1µg/ml pepstatin, and 1µg/ml leupeptin. Nuclear and cytoplasmic extracts were obtained using a nuclear extraction kit according to the manufacturer's instructions (Active Motif). Protein concentration of lysates were determined using a Bradford assay (Bio-Rad). Immunoprecipitation (IP) assays employed equal concentrations of whole cell extracts, pre-cleared for 30 min at 4°C with protein G-agarose beads (Millipore), and immunoprecipitated with antibody overnight at 4°C. Beads were washed three times in RIPA buffer, resuspended in equal volumes of 2X SDS loading dye and heated for 5 mins at 95°C to elute immunoprecipitated protein. Denatured samples were resolved by SDS-PAGE, transferred to nitrocellulose membranes and blocked for 1 hour at

room temperature using 5% fat free milk powder in PBS-Tween 20. Membranes were immunoblotted with specific antibodies: p-IκBα (cat. # 9246), IκBα (cat. # 4821) and p105/p50 (cat. # 12540) were purchased from Cell Signalling Technologies; antibodies against FLAG (cat. # F1804), HDAC-1 (cat. # AV38530), α-tubulin (cat. # T6074) and ß-actin (cat. # A5441) were purchased from Sigma; anti-p65 (cat. # A301-824A) was purchased from Bethyl Laboratories, anti-Myc (cat. # SC-40) was purchased from Santa Cruz; and anti-Xpress (cat. # R910-25) was purchased from Invitrogen. A custom anti-p50 p-S80 antibody was designed and purchased from BioGenes GmbH. Antibody binding was visualised using WesternBright HRP ECL substrate (Advansta).

4.3.3 Kinase assays

For *in vitro* kinase assays utilising an immunoprecipitated kinase source, cells were transfected with plasmid encoding constitutively active or kinase inactive IKKß and immunoprecipitations performed as described above. Protein-bound beads were used directly in the kinase assay by adding 800ng of substrate and 10µl of kinase assay buffer (400mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) pH 7.5, 400mM MgCl2, 20mM EDTA, 40mM NaF, 40mM ß-glycerophosphate, 20mM DTT, 200µM cold ATP). 10µCi γ-32P labelled ATP was added to each sample and incubated at 30°C for 15 mins. Samples were made up to 500µl with ice cold PBS, incubated with glutathione beads for 30 mins and agitated every 5 mins. Beads were washed 3 times with PBS followed by heating at 95°C for 5 mins in 40µl 2X SDS loading buffer. Samples were resolved by SDS-PAGE, and the gel was stained with GelCode Blue (Expedeon) reagent. The gel was fixed in fixation buffer (10% Glacial acetic acid/ 20% methanol/ 50% dH_2O) for 30 mins followed by a 5 min incubation with rehydration buffer (20% methanol/3% glycerol). The gel was dried onto Whatman paper using a gel-dryer, exposed to phosphor imaging screen overnight and then visualised using the Storm phosphor-imager system (Molecular devices). For recombinant *in vitro* kinase assays, 400ng recombinant IKKß kinase (Cell signalling technologies, #7458) was incubated with 400ng substrate following the protocol as described above. For the peptide array kinase assay, peptide libraries of murine p50 were generated by an automatic SPOT synthesis, as previously described (Frank, 2002) and synthesised on continuous cellulose membrane supports on Whatman 50 cellulose using Fmoc (N-(9-

flurenyl)methoxycarbonyl) chemistry with AutoSpot-Robot ASS 222 (Intavis Bioanalytical Instruments). The array was used as a substrate for immunoprecipitated and recombinant IKKß, carried out as described by (Kiely et al., 2009).

4.3.4 Site directed mutagenesis and GST protein purification

p50 was cloned into pGEX-6p1 in frame with the N terminal GST tag and transformed into BL21 CodonPlus Escherichia coli (Stratagene). Transformants were grown to an A600 of 1.0–2.0 at 37°C and induced with 1.0mM 1mM isopropyl ß-D-1 thiogalactopyranoside (IPTG) for 16 h at 20°C. The bacteria were resuspended in a buffer containing 50mM Tris (pH 8.0), 150mM NaCl, and 1mM dithiothreitol, disrupted by sonication, and centrifuged to remove debris. Recombinant proteins were affinity-purified with GSH-agarose (Sigma) and eluted with 10mM glutathione (Promega) in 50mM Tris (pH 8.5) and 150mM NaCl. GST-p50 mutants were created by PCR based site-directed mutagenesis using the Stratagene QuikChange II SDM kit, according to manufacturer's instructions and recombinant protein purified as above.

4.3.5 CRISPR/Cas9 genome editing

pSpCas9(BB)-2A-GFP (PX458) was a gift from Feng Zhang (Addgene plasmid # 48138). gRNA oligonucleotides were purchased from Eurofins Genomics (5'-CACC GACAAACTTACTTTGACCTGA-3', 3'-CTGTTTGAATGAAACTGGACTCAAA-5') and were cloned into pSpCas9(BB)-2A-GFP. ssODN was purchased from Integrated DNA Technologies (5'-AGAGGATTTCGTTTCCGTTATGTATGTGAAGGCCCATCCCATGGTGGACTACCTGGTGC CTCTAGTGAAAAGAACAAGAAAGCTTATCCTCAGGTCAAAGTAAGTTTGTGGTAGCTCTC CTTCTATTTGAATTCTGGAAATTTTGATTTCCTACGATTTCCAAGGAATT-3'). The ssODN contains synonymous codon changes designed to introduce the S80A mutation, a HindIII restriction site, and to abolish the PAM site. Prior to transfection, HEK293T cells were treated with 200ng/ml nocodazole (Sigma) for 16 h. Following treatment, cells were washed twice with media, trypsinised and cell density was determined. Cells were washed with Mg^{2+} and Ca^{2+} free PBS, and re-suspended in Neon buffer R (Invitrogen) at a density of $5x10^7$ cells/ml per transfection. 100µl of cells were co-transfected with pPGKpuro (a gift from

Rudolf Jaenisch, Addgene plasmid # 11349), pX458 Cas9/sgRNA vector and ssODN (1:2:2 ratio) by electroporation. Transfected cells were plated into a 10cm dish containing 10ml media and recovered overnight in culture. Transfected cells were then selected for ~3 days with 3µg/ml of puromycin. After selection, puromycin media was removed and cells washed and incubated in fresh media. Serial dilution of selected cells was performed to isolate single cell clones in 96-well plates. For clone screening, genomic DNA was extracted from cells using the DNeasy Blood & Tissue Kit (QIAGEN) according to the manufacturer's instruction. DNA was amplified by PCR using primers (F: 5'-ACCTGGCTTTTTAGCCATATCT-3'; R: 5'-TTCAGCTTAGGAGCGAAGGC-3') and HotStarTaq Master Mix Kit (QIAGEN) according to the manufacturer's instructions. Intitial screens were performed by HindIII (NEB) restriction digest of PCR products purified using QIAquick PCR Purification Kit (QIAGEN), according to the manufacturer's instructions. Gene editing of selected clones was confirmed by Sanger sequencing (GATC-Biotech).

4.3.6 Transcriptomic and bioinformatic analysis

For RNA-sequencing, total RNA was extracted from cells using RNeasy kits (QIAGEN), according to the manufacturer's instructions. Triplicate independently generated samples for each condition were sequenced to a read depth of 20million using Illumina NextSeq™ 500 platform. Single-end 75bp reads were aligned to human reference sequence iGenome NCBI GRCh38 using HISAT (Kim et al., 2015). Aligned reads were counted using HTseq-count (Anders et al., 2015). The above analyses were performed using the University of Glasgow Galaxy server. Differentially expressed genes were calculated and visualised using an in-house RNA-seq analysis pipeline which utilises DESeq2 (Love et al., 2014). Induced genes with adjusted p value <0.05 were included in this study. Heat maps of gene groups were produced using DESeq2 mean normalised counts, visualised using the online tool Morpheus (Broad Insitute). For real-time quantitative PCR, total RNA was extracted from cells using RNeasy kits (QIAGEN), according to the manufacturer's instructions. Isolated RNA was primed with random hexamer oligonucleotides and reverse transcribed using Primer Design precision nanoScript reverse transcription kit. PCR was performed with SYBR Green master mix and QuantiTect primers (QIAGEN). All data were

normalised to TBP. Gene expression changes were calculated using the $2^{-\Delta\Delta CT}$ method.

4.3.7 Motif analysis

Motif analysis was based on the GREAT approach (McLean et al., 2010), which incorporates both proximal and distal regulatory regions for enrichment analyses. Regulatory regions for genes were obtained from the Ensembl Regulatory Build (Zerbino et al., 2015). Genomic locations were obtained using the "fetch closest non-overlapping feature" tool. JASPAR Position Weight Matrix MA0105.1 was used to identify the best matching NFKB1 binding sites in regulatory regions using the "FIMO" tool. These analyses were performed using the University of Glasgow Galaxy server. Sequence logos were generated using WebLogo (Crooks et al., 2004). Transcription factor binding site shape analysis was performed using the TFBSshape online tool (Yang et al., 2014).

4.3.8 DNA affinity binding assay (DAPA)

5'-biotinylated and unlabelled single-stranded oligonucleotides containing a central κB site flanked either side a by 7-bp spacer were purchased from Eurofins Genomics (5'-AGTTGAGGGGNNTTTCCCAGGC-3'; 3'-TCAACTCCCCNNAAAGGGTCCG-5', where N represents -2 -1 κB site variation per assay). Oligonucleotides were annealed to create 5'-biotinylated, and unlabelled double-stranded duplexes. DAPA reactions were prepared by mixing 1.5µg of 5'-biotinylated double stranded oligo with 150µg of nuclear extract and 15µl streptavidin-agarose beads (Sigma) in a total of 500 µl DAPA buffer (10mM Tris HCl pH 7.5, 50mM NaCl, 1mM DTT, 5% glycerol, 1mM EDTA, 1mM NaF, 1mM Na_3VO_4, 1mM PMSF, 1µg/ml leupeptin, 2µg/ml aprotinin, 2µg/ml pepstatin). Unlabelled double stranded oligonucleotides were added in 10-fold excess to confirm binding specificity. Reactions were incubated at room temperature on a rotator for 1 hour. Beads were washed 3 times in 500µl DAPA buffer. To elute DNA-bound proteins, beads were resuspended in 20µl of 2X SDS sample buffer, and incubated at 95ºC for 5 mins. Eluates were resolved on SDS-PAGE gels and analysed by western blot.

4.3.9 Luciferase assays

Promoters containing 4 sequential copies of defined κB sites (5'-GGGAATTTCC-3', 5'-GGGACTTTCC-3', 5'-GGGGATTTCC-3', 5'-GGGGCTTTCC-3') were cloned into the pTAL-Luc vector (Clontech). WT and NFKB1^{S80A} HEK293T cells were co-transfected with 100ng pTAL-(4xκB) Luc vector and 10ng Renilla luciferase expression vector pRL-TK (Promega). 24 hours post transfection, cells were cultured with or without 10ng/ml TNFα for 8 hours before harvest. Luciferase activities of whole cell lysates were analysed using the Dual-Luciferase® Reporter Assay System (Promega). The ratio of firefly to Renilla luciferase activity was used to normalise for transfection efficiency across all samples.

4.3.10 Limited Proteolysis

Whole cell lysates were generated from cells suspended in RIPA buffer lacking protease inhibitors. Limited proteolysis was performed by adding varying ratios of trypsin (Sigma) to 50μg of whole cell lysate, and incubating the reaction mixture for 30 mins at 37°C. Proteolysis was terminated by adding 5X SDS sample buffer to the reaction and heating at 95°C for 5 mins. Samples were resolved by SDS/PAGE and analysed by western blot analysis.

4.4 Results

4.4.1 IKKβ phosphorylates NF-κB p50

The phosphorylation of NF-κB subunits is strongly linked to the activation of the NF-κB pathway. Indeed, in addition to IκBα, IKKβ also phosphorylates other components of the NF-κB pathway including p65 and p105 (Christian et al., 2016). To determine if IKKβ may also phosphorylate p50, we initially performed *in vitro* kinase assays using constitutively active IKKβ ($IKK\beta^{SSEE}$) immunoprecipitated from transiently transfected HEK293T cells and employing purified recombinant GST-p50 as substrate. These assays revealed that active IKKβ phosphorylates p50 *in vitro* (Figure 4.1A). Kinase assays incorporating immunoprecipitated kinase dead IKKβ ($IKK\beta^{K44M}$) and purified recombinant GST established that p50 phosphorylation was due to IKKβ kinase activity (Figure 4.1A). Similar results demonstrating the phosphorylation of p50 by IKKβ were obtained using *in vitro* kinase assays incorporating purified recombinant IKKβ (Figure 4.1B). Co-immunoprecipitation of p50 and constitutively active $IKK\beta^{SSEE}$ in transfected cells demonstrated that IKKβ and p50 interact (Figure 4.1C). Of note, p50 interacted with the active but not a kinase dead form of IKKβ indicating that IKKβ activity is required for interaction with p50 (Figure 4.1C). These data establish the p50 subunit of NF-κB as a novel substrate of the IKKβ kinase.

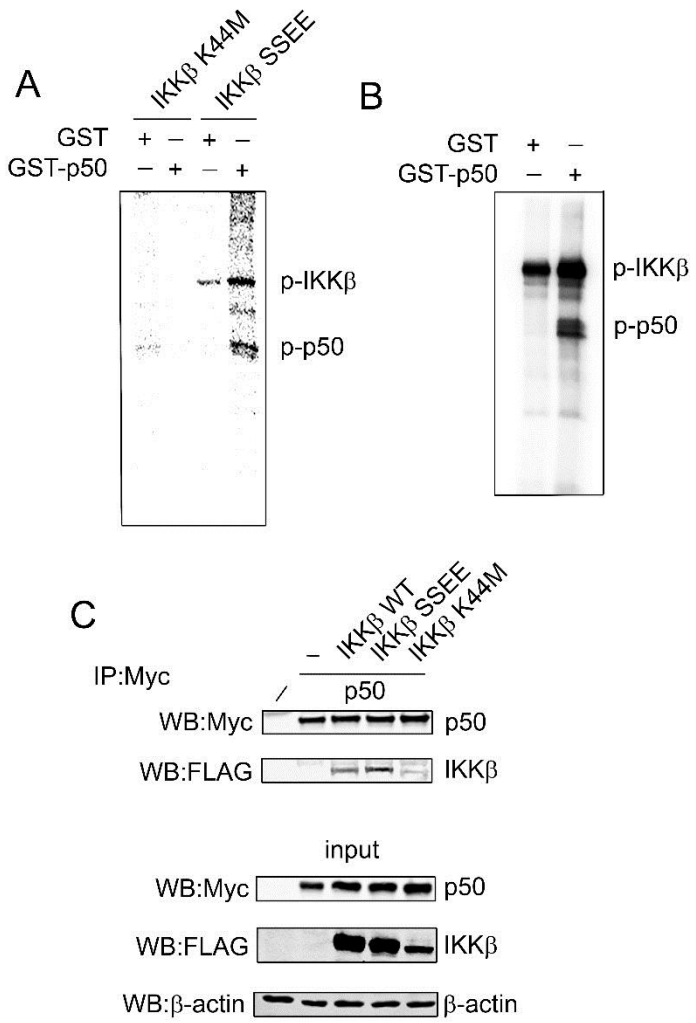


Figure 4.1: IKKβ phosphorylates NF-κB p50.
(A) HEK293T cells were transfected with FLAG-IKKβ^{K44M} (kinase dead) or FLAG-IKKβ^{SSEE} (constitutively active) as indicated. IKKβ was immunoprecipitated from whole cell lysates with anti-FLAG antibody and incorporated in a kinase assay using recombinant GST-p50 or recombinant GST, with [γ-^{32}P] ATP as a co-substrate. Samples were resolved by SDS-PAGE and visualised by autoradiography. (B) *In vitro* kinase assay employing recombinant IKKβ and recombinant GST-p50 or GST, with [γ-^{32}P] ATP as a co-substrate. Samples were resolved by SDS-PAGE and visualised by autoradiography. (C) HEK293T cells were co-transfected with Myc-p50 and FLAG-IKKβ^{SSEE} or FLAG-IKKβ^{K44M} as indicated. p50 was immunoprecipitated (IP) from whole cell lysates with anti-Myc antibody and analysed by western blot (WB) using the indicated antibodies.

4.4.2 Serine 80 is the IKKβ phosphorylation site of NF-κB p50

To identify the specific amino acids of p50 phosphorylated by IKKß we employed an *in vitro* kinase assay using a peptide array representing the entire amino acid sequence of p50. A series of 30 18-amino-acid-long peptides were SPOT synthesised on nitrocellulose, with each peptide overlapping by 3 residues to generate a p50 peptide array. The p50 array was subjected to *in vitro* kinase assays using recombinant IKKß as previously described (Kiely et al., 2009). This analysis identified four putative IKKß phosphorylation sites in p50; S73, S74, S80 and T315 (human p50 amino acid numbering used) (Figure 4.2A). To further test the putative sites of IKKß phosphorylation we next performed *in vitro* kinase assays incorporating recombinant IKKß and recombinant GST-p50 in which S73, S74, S80 and T315 are mutated to alanine. This analysis revealed that IKKß could still phosphorylate p50 when S73, S74, and T315 are mutated to alanine, but not when S80 is mutated to alanine (Figure 4.2B). These data identify S80 of p50 as the site of IKKß phosphorylation. Analysis of the available crystal structure of p50 homodimer bound to DNA revealed that S80 of p50 is located in an extended loop of the RHD and therefore available for phosphorylation (Figure 4.2C). To further confirm IKKß phosphorylation of p50 *in vivo*, we next generated an antibody raised against the S80 phosphorylation site of p50. This anti-phospho-S80 p50 antibody recognised IKKß dependent phosphorylation of p50 but not a $p50^{S80A}$ mutant in transiently transfected cells, demonstrating the phosphorylation of p50 S80 by IKKß in cells (Figure 4.2D). Unfortunately, this antibody was not of sufficient affinity to generate a specific signal in cell lysates of non-transfected cells and so was not of further use in investigating p50 S80 phosphorylation.

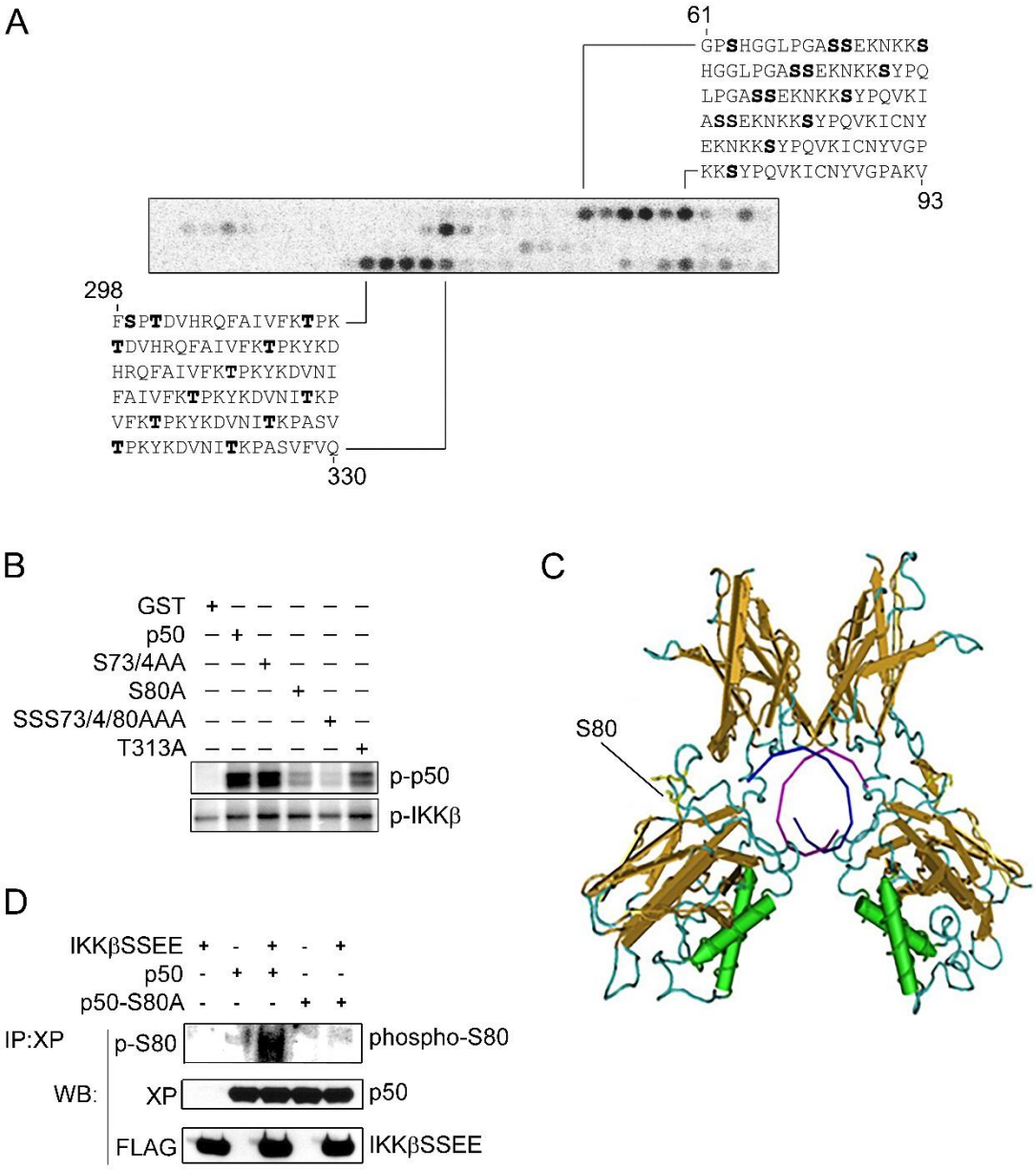


Figure 4.2: IKKβ phosphorylates NF-κB p50 at S80.
(A) *In vitro* IKKβ kinase assay using a peptide array of immobilised, overlapping 18-mer peptides the entire p50 sequence. Black spots represent IKKβ-phosphorylated peptides. Peptide sequences, amino acid number (mouse p50) and putative phosphosites (bold) are shown. (B) *In vitro* IKKβ kinase assay using recombinant GST, wild type and mutated recombinant GST p50 as indicated. (C) Image from the X ray crystal structure of p50 homodimer bound to DNA (1NFK) indicating the location of S80 in an extended loop of the Rel Homology Domain. (D) HEK293T cells were co-transfected with or without FLAG-IKKβ^{SSEE} and with either XP-p50 or XP-p50^{S80A} as indicated. p50 was immunoprecipitated (IP) from whole cell lysates with anti-XP antibody and analysed by western blot (WB) using anti-phospho-serine 80 (p-S80) antibody.

4.4.3 S80 phosphorylation is not required for NF-κB activation

Following the identification of S80 as a novel phosphorylation site on the p50 subunit, we next sought to investigate its role in regulating NF-κB activity. To achieve this we utilised CRISPR/Cas9 genome editing techniques to generate *NFKB1*^{S80A} knock-in HEK293T cells (Figure 4.3A). The p50 subunit of NF-κB is generated from the proteasomal processing of the p105 precursor which requires the IKKß mediated phosphorylation at the C-terminus of p105 (Heissmeyer et al., 1999, Heissmeyer et al., 2001, Salmeron et al., 2001, Lang et al., 2003). To determine whether S80 plays a role in the processing of p105 we analysed p105/p50 protein levels in whole cell lysates from wild type (WT) and cells by western blot. This revealed equivalent levels of p105/p50 in WT and *NFKB1*^{S80A} demonstrating that S80 is not required for the processing of p105 to p50 (Figure 4.3B). To determine the effect of S80 mutation on the activation of the NF-κB pathway we next stimulated WT and *NFKB1*^{S80A} cells with TNFα and assessed the phosphorylation and degradation of IκBα by western blot. This analysis demonstrated equivalent levels of IκBα phosphorylation and degradation in WT and *NFKB1*^{S80A} cells stimulated with TNFα (Figure 4.3C). Nuclear translocation of p50 and p65 in TNFα stimulated WT and *NFKB1*^{S80A} was assessed by immunoblot analysis of nuclear and cytoplasmic fractions. This demonstrated equivalent levels of NF-κB translocation to the nucleus following TNFα stimulation (Figure 4.3D). Together, these data demonstrate that upstream signalling and nuclear translocation of NF-κB following TNFα stimulation is unaffected by the p50 S80A mutation. To investigate the effect of S80 mutation on p50 interaction with p65 we performed an immunoprecipitation assay using anti-p50 antibody. Western blot analysis of immunoprecipitates showed that equivalent levels of p65 co-purified with p50 in both WT and *NFKB1*^{S80A} cells demonstrating that p50 S80A mutation does not alter the interaction of p50 and p65 (Figure 4.3E).

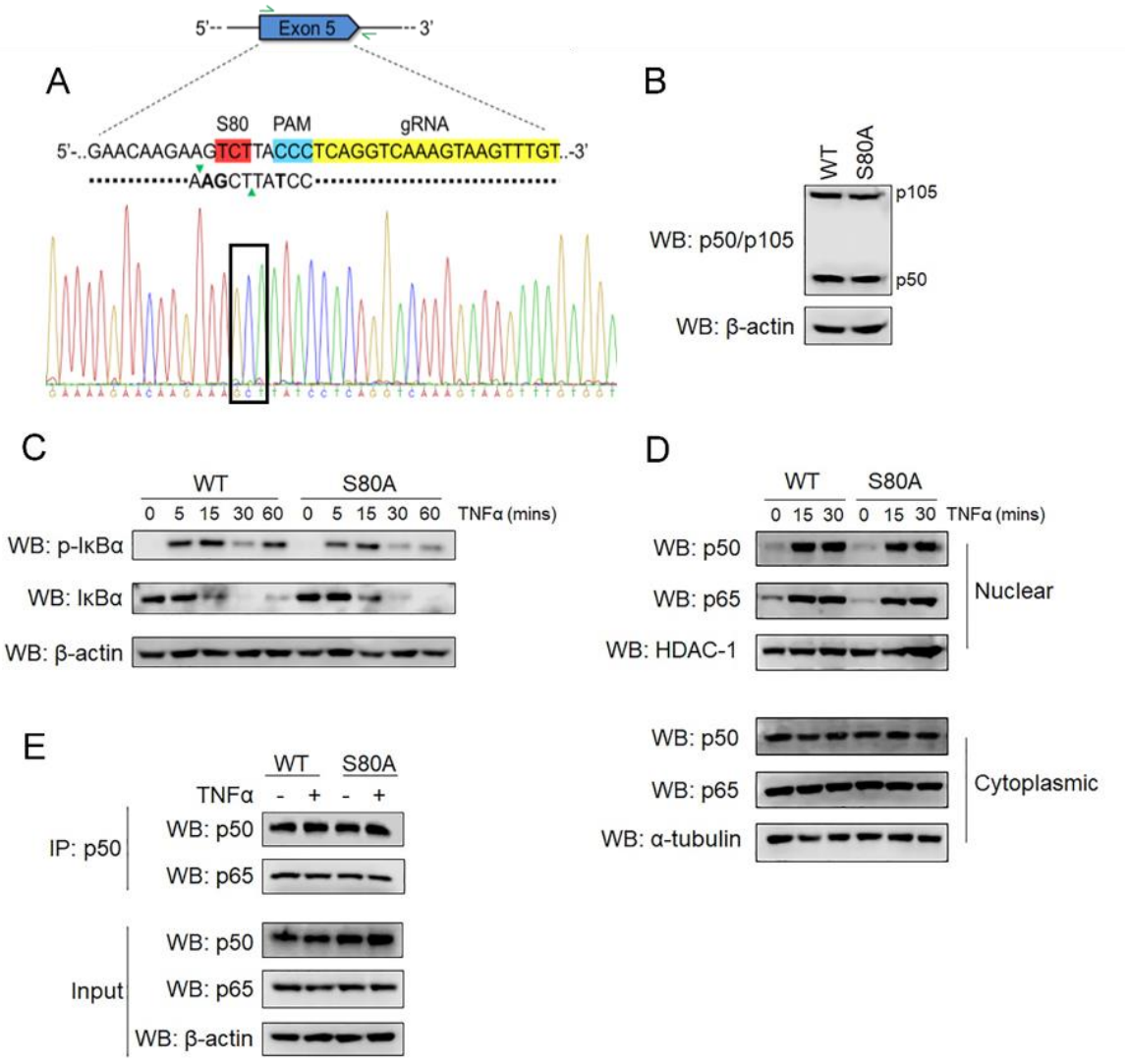


Figure 4.3: S80 phosphorylation is not required for NF-κB activation.
(A) Schematic of the CRISPR/Cas9 targeting strategy used to edit the *NFKB1* target locus, highlighting the position of the S80 codon (red); gRNA (yellow); PAM site (blue) and primers used for screening (green arrows). The ssODN template including homology arms (dotted line), HindIII restriction cut site (green triangles) and synonymous codon changes (bold) are shown. Also shown is DNA sequence chromatogram of the *NFKB1* target region confirming the S80A point mutation (black box). (B) Whole cell lysates extracted from WT and *NFKB1*^{S80A} HEK293Ts were analysed by western blot (WB) with the indicated antibodies. (C) WT and *NFKB1*^{S80A} HEK293Ts were stimulated with 10ng/ml TNFα for the indicated times prior to lysis. Whole cell extracts were analysed by western blot to detect levels of phosphorylated and total IκBα protein. (D) WT and *NFKB1*^{S80A} HEK293Ts were stimulated with 10ng/ml TNFα for the indicated times prior to lysis. Nuclear and cytoplasmic extracts were analysed by western blot using antibodies against p65 and p105/p50. (E) WT or *NFKB1*^{S80A} HEK293Ts were left untreated or were stimulated with 10ng/ml TNFα for 30 mins prior to lysis. p50 was immunoprecipitated (IP) from whole cell lysates with anti-p105/p50 antibody and analysed by western blot (WB) with anti-p105/p50 and anti-p65 antibodies as indicated.

4.4.4 S80 phosphorylation selectively regulates TNFα-induced gene expression

Site-specific phosphorylation of NF-κB subunits has previously been shown to regulate transcriptional activity (Christian et al., 2016). To determine the role of S80 phosphorylation in regulating NF-κB target gene expression, we next analysed TNFα-induced transcriptional responses in WT and *NFKB1*^{S80A} cells by RNA-seq. WT and *NFKB1*^{S80A} cells were untreated or treated with TNFα for 3 hours prior to RNA-seq analysis. This revealed distinct transcriptional profiles of differentially expressed genes in WT and *NFKB1*^{S80A} cells in response to TNFα (Figure 4.4A). In particular we identified two TNFα-inducible groups of genes composed predominantly of NF-κB target genes containing an identifiable NF-κB binding site in their promoter regions that were differentially regulated by WT and *NFKB1*^{S80A} cells (Figure 4.4B). The expression of genes encoding pro-inflammatory chemokines and cytokines including *TNF*, *IL8*, *CXCL2*, *CXCL1* and *CXCL10* was dramatically increased in *NFKB1*^{S80A} cells compared to WT cell following TNFα treatment. However, other NF-κB target genes predominantly encoding for intracellular signalling factors including *BCL3*, *MAP3K8*, *NFKB1* and *IRAK1* were expressed at equivalent levels in both WT and *NFKB1*^{S80A} cells following TNFα treatment. Analysis of selected genes from each group by qPCR confirmed the gene-selective regulation of transcription by p50 S80 phosphorylation (Figure 4.4C). Of note, although the expression levels of a number of TNF-inducible genes is higher in *NFKB1*^{S80A} cells than WT cells, there is a very high correlation between TNFα induced differentially expressed genes in WT and *NFKB1*^{S80A} cells (Figure 4.4D). This indicates that S80 phosphorylation regulates the levels of gene expression induced by TNFα treatment rather than the identity of the genes induced.

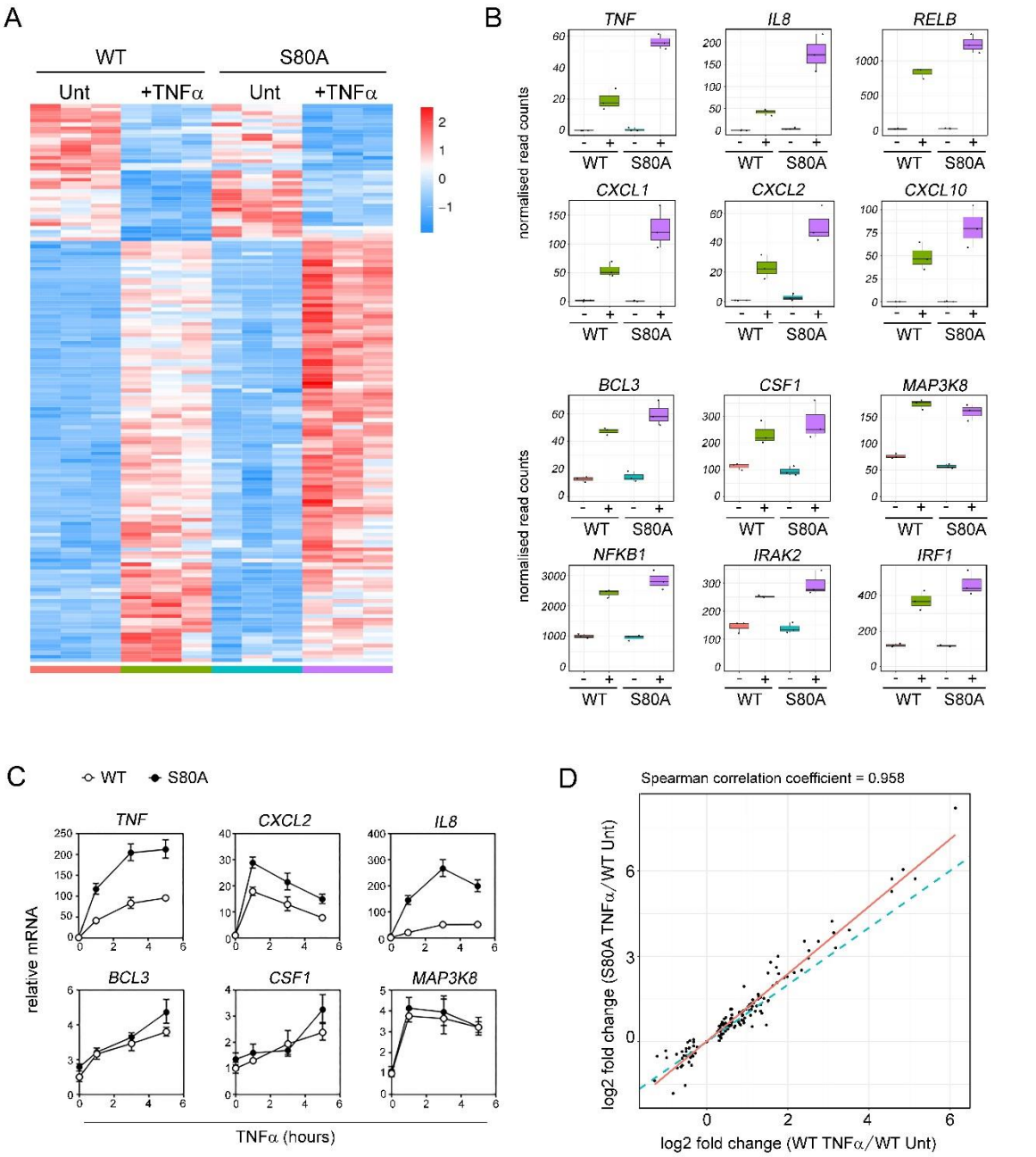


Figure 4.4: S80 phosphorylation selectively regulates TNFα-induced gene expression. (A) Triplicate samples of WT and *NFKB1*^{S80A} cells treated with TNFα for 3 hours (+TNFα) or untreated (Unt) were analysed by RNA-seq. The heat map displays differentially expressed genes (p.adj <0.05) scaled as per z-score. Genes were clustered using spearman distances and UPGMA agglomeration. (B) Box and whisker plots of gene expression level by RNA-seq for selected genes. Each dot represents a sample. Boxes show the 25th, 50th and 75 percentile with whiskers showing inter-quartile range. (C) WT and *NFKB1*^{S80A} HEK293Ts were stimulated with 10ng/ml TNFα for the indicated times prior to harvest and RNA extraction. Gene expression levels were determined by quantitative real-time PCR. Data are mean ± S.E of triplicate samples and are representative of three independent experiments. (D) Log2 fold change scatter plot of genes (dots) shown in (B). The blue dotted line indicates correlation the between WT TNFα/WT Unt v S80A 0hr/WT Unt fold changes (spearman correlation ~0.96). The red line represents linear regression.

4.4.5 Specific DNA-binding motifs associated with differential regulation of NF-κB target genes by S80

The gene selective effects of S80 mutation on TNFα-induced transcription suggested that phosphorylation of S80 may regulate the activity of p50 in a promoter specific manner. Binding site sequence preferences differ among NF-κB dimers (Siggers et al., 2011) and suggests a potential mechanism for the selective effect of S80 mutation on the transcription of a distinct set of target genes. To further explore this possibility, we performed transcription factor binding site analysis to search for the most enriched κB sites in TNFα-induced genes expressed higher in *NFKB1*^{S80A} cells than WT cells, and those expressed equally in both WT and *NFKB1*^{S80A} cells. The genomic region between the nearest upstream and downstream gene was used to search for both proximal and distal regulatory features (including promoters and enhancers) for each gene. Regulatory features were analysed for occurrences of a 10 base pair NFKB1 position weight matrix (JASPAR ID MA0105.1). This analysis identified distinct DNA binding motifs associated with each group of genes (Figure 4.5A and Figure 4.5B). Specifically, promoter and enhancer κB sites varied at the -1 and -2 positions (Figure 4.5A and Figure 4.5B). The κB sites of genes with increased expression in *NFKB1*^{S80A} cells relative to WT cells, contain an adenine (A) at the -2 position, and either an adenine or a cytosine (C) at the -1 position (Figure 4.5A). In contrast, the κB sites of genes where expression is unchanged between WT and *NFKB1*^{S80A} cells contain either an A or a guanine (G) at the -2 position while there is no enrichment of any particular nucleotide at the -1 position (Figure 4.5B). This analysis suggested that S80 phosphorylation may regulate p50 function in a DNA binding site sequence specific manner.

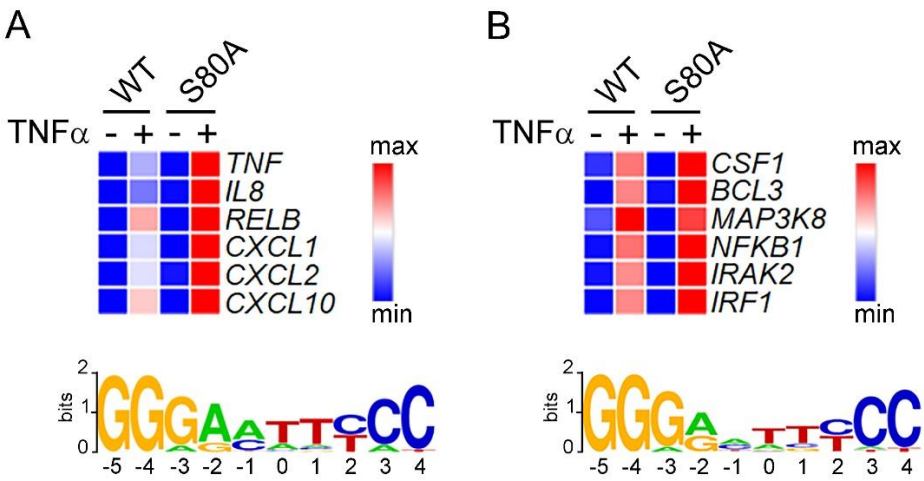


Figure 4.5: Specific κB sites are enriched in genes selectively regulated by S80 phosphorylation.
Gene expression values (scaled as per z-score) of selected genes (A) induced to a greater degree in *NFKB1*^{S80A} cells compared to WT cell, and (B) genes expressed equally between WT and *NFKB1*^{S80A} cells. Proximal and distal regulatory regions surrounding each gene were searched for occurrences of *NFKB1* motifs best matched to JASPAR position weight matrix MA0105.1. Sequence logos representing the most enriched κB site for each gene set are shown.

4.4.6 S80 phosphorylation reduces p50 affinity for -1A containing κB sites

To determine if S80 phosphorylation alters the binding of p50 to specific κB sequences we next carried out DNA affinity precipitation assays (DAPA) incorporating the κB sequences identified in our analysis of transcriptomic data (Figure 4.5). WT or *NFKB1*^{S80A} cells were left untreated or treated with TNFα prior to extraction of nuclear fractions. Nuclear extracts were incubated with 5'-biotinylated oligonucleotides representing κB sites with base pair substitutions at the -1 and -2 positions. The oligonucleotides were designed to represent the DNA-binding motifs identified in the gene sets identified by our transcriptomic analysis (Figure 4.5A and Figure 4.5B), and differed only in the base pair sequence at the -2 and -1 positions of a GGGA(-2)C(-1)TTTCC motif. Oligonucleotides representing κB sites from genes with enhanced TNFα inducible expression in *NFKB1*^{S80A} cells therefore contained an A at the -2 position and an A or C at the -1 position. Since the κB sites identified in genes with equivalent expression in both WT and *NFKB1*^{S80A} cells displayed heterogeneity at the -1 position we generated oligonucleotides with an A or C at the -1 position to reflect the previously reported prevalence of A and C at this position of κB sites (Zabel et al., 1991) and a G at the -2 position. Oligonucleotides and bound

protein were precipitated using streptavidin conjugated agarose beads. Specificity of protein DNA interaction was verified using control samples containing a 10X excess of non-biotinylated oligonucleotide. Precipitated proteins were resolved by SDS gel electrophoresis and analysed by western blot using antibodies against p50 and p65 and quantified using a digital chemiluminescence scanner. These assays demonstrated approximately 2 fold greater binding of $p50^{S80A}$ to the GGGA(-2)A(-1)TTTCC κB site when compared to WT p50 (Figure 4.6). Remarkably, there is also significantly greater binding of p65 to this κB site in *NFKB1*^{S80A} cells relative to WT cells, demonstrating that p50 S80 phosphorylation regulates the DNA binding of p50:p65 heterodimers. The requirement of an A at the -1 position for this effect was demonstrated by a single base change of -1A to -1C in the GGGA(-2)C(-1)TTTCC κB site which largely abolished the increased $p50^{S80A}$ binding and completely abolished the increased p65 binding seen with the GGGA(-2)A(-1)TTTCC site (Figure 4.6).

The significance of the -1A nucleotide to $p50^{S80A}$ binding was further demonstrated by the increased binding of $p50^{S80A}$ to a GGGG(-2)A(-1)TTTCC κB sites relative to WT p50 (Figure 4.6). Of note, the levels of p65 binding to this site were similar in both $p50^{S80A}$ and WT p50 expressing cells indicating that S80 phosphorylation may regulate the binding of p50 homodimers but not p50:p65 heterodimers to κB sites with this sequence. This suggested that binding to specific κB sequences may also be influenced by the composition of p50 containing dimers. An additional change from A to C in the -1 nucleotide position in the GGGG(-2)C(1)TTTCC site further reduces $p50^{S80A}$ binding relative to WT p50 as compared to GGGG(-2)A(-1)TTTCC κB sites (Figure 4.6). Together, these data show that the regulation of p50 binding by S80 is determined primarily by the identity of the -1 nucleotide which appears to modify the binding of both p50 homodimer and p50:p65 heterodimer complexes. Thus S80 phosphorylation reduces the binding of p50 containing NF-κB dimers to κB sites containing an A at the -1 position.

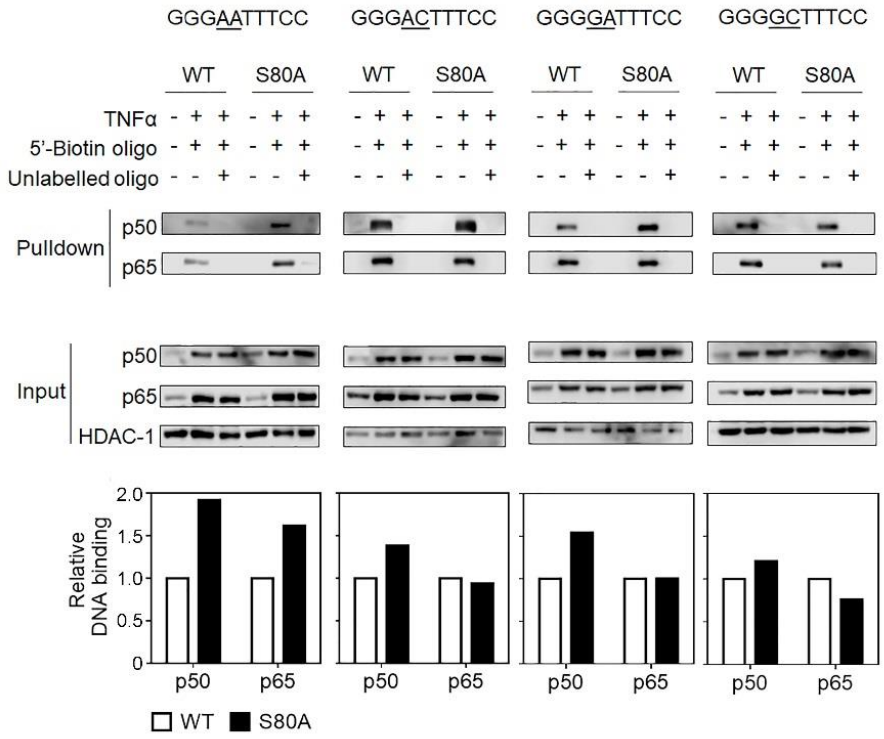


Figure 4.6: S80 phosphorylation regulates DNA binding affinity to -1A containing κB sites. WT and *NFKB1*^{S80A} HEK293T were left untreated or treated with 10ng/ml TNFα as indicated for 30 mins. Equal concentrations of nuclear lysates were incubated with 5'-biotinylated double stranded oligonucleotides, with or without a 10X excess of unlabelled competitor double stranded oligo indicated. The sequence of the double stranded oligos containing κB sites that vary at the -1 and -2 positions are indicated for each assay (underlined). Precipitated samples were analysed for p50 and p65 proteins by WB. Relative binding of p50 and p65 protein to DNA was quantified by normalising to input. Data shown are representative of 3 individual experiments.

4.4.7 S80 phosphorylation inhibits transcription from -1A containing κB sites

The transcriptomic and DNA binding analyses suggest that the increased transcription of pro-inflammatory genes in *NFKB1*^{S80A} cells results from increased binding of p50^{S80A} containing NF-κB dimers to promoter κB sites containing an A in the -1 position. We next sought to investigate whether increased DNA-binding of p50^{S80A} is sufficient to increase target gene transcription or whether the transcriptional outcome observed in TNFα treated *NFKB1*^{S80A} cells occurs only in the context of the gene promoter. We generated 4 different luciferase reporter constructs that contained four κB site repeats that vary at either the -2 or the -1 position immediately upstream of a minimal promoter and firefly luciferase reporter gene. The 4 reporter plasmids each contained a κB site sequence identical to each of the sequences employed in the DAPA experiments described

above (Figure 4.6). WT or *NFKB1*^{S80A} cells were transiently transfected with the different reporter plasmids along with a constitutive expression vector for Renilla luciferase to enable normalisation for transfection efficiency. Luciferase reporter activity was measured in untreated and TNFα-treated cells to determine the impact of p50^{S80A} on transcription driven by specific κB site sequences. TNFα-induced luciferase activity from the reporter containing the GGGA(-2)A(-1)TTTCC sequence was approximately 2 fold greater in *NFKB1*^{S80A} cells compared to WT cells (Figure 4.7), consistent with the increased DNA binding of p50^{S80A} to this sequence observed in the DAPA assays (Figure 4.6). Interestingly, TNFα induced reporter activity in both WT and *NFKB1*^{S80A} cells is highest in reporter plasmids containing the GGGA(-2)A(-1)TTTCC site compared to the other κB sites tested, suggesting that κB sites containing an A at the -1 and -2 position are more potent drivers of gene transcription. The importance of the -1 nucleotide is further highlighted by the observed overall decrease in luciferase activity in both WT and *NFKB1*^{S80A} cells when the -1A is changed to -1C (Figure 4.7).

The increased TNFα induced reporter activity from the A(-2)C(-1) κB site seen in *NFKB1*^{S80A} cells relative to WT cells (Figure 4.7) is also consistent with the observed increased expression levels in *NFKB1*^{S80A} cells of genes which contain either A(-2)A(-1) or A(-2)C(-1) κB sites (Figure 4.5A). The relative differences in TNFα induced reporter activity between WT and *NFKB1*^{S80A} cells are much less where the κB site contains a G(-2)A(-1) and G(-2)C(-1) sequence (Figure 4.7). This is consistent with the transcription profiles observed for genes that contain either G(-2)A(-1) or G(-2)C(-1) κB sites where expression is similar between WT and *NFKB1*^{S80A} cells (Figure 4.5B). Furthermore, DNA binding analysis also showed that p65 binding to these sites is either unaffected, or slightly reduced (G(-2)A(-1) and G(-2)C(-1), respectively) by S80 phosphorylation which correlates with the reporter activity (Figure 4.6). The reporter activity for all four κB sequences tested were equivalent between untreated WT and *NFKB1*^{S80A} cells, consistent with TNFα dependent activation of IKKβ as a requirement for p50 phosphorylation. Taken together, these data show that S80 phosphorylation regulates NF-κB transcriptional activity in a κB sequence specific manner and that binding site sequence differences are sufficient to determine transcriptional outcome following S80 phosphorylation. Specifically, S80 phosphorylation acts to

limit the NF-κB mediated transcription of genes containing a -1A κB site nucleotide.

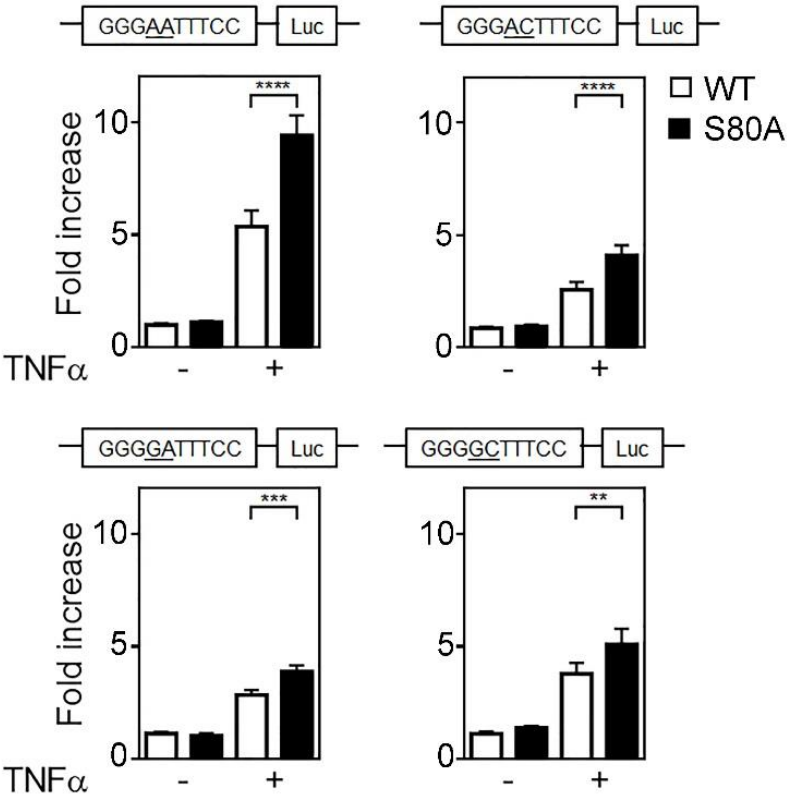


Figure 4.7: S80 phosphorylation regulates transcription in a κB sequence specific manner. WT and *NFKB1*^{S80A} HEK293T cells were transfected with pTAL-NFKB reporter constructs with four identical tandem κB sites that vary at the -1 and -2 positions as indicated (underlined). 24 hours post transfection, cells were either left untreated or treated with TNFα for 8 hours before luciferase activity was measured. The *Renilla* luciferase expression vector pRLTK was used as an internal control to normalise transfection efficiency across all samples. Reporter activity is represented as fold increase over untreated WT cells. Data shown are mean ± S.E. of quadruplicate samples. Statistical significance between treated WT and *NFKB1*^{S80A} cells was determined by Student's *t* test. **p≤0.01; ***p≤0.001; ****p≤0.0001.

4.5 Discussion

In this study, we have identified S80 as a novel phosphorylation site on the NF-κB p50 subunit. The phosphorylation of S80 by the IKKβ kinase also identifies p50 as a novel substrate for this kinase. Our data demonstrates that the phosphorylation of p50 at S80 selectively regulates TNFα-induced transcription by regulating the DNA binding of p50 at κB sites in a sequence specific manner. Thus, phosphorylation of S80 reduces p50 DNA binding to κB sites with a -1A, and thereby limits the expression of gene under the control of regulatory elements bearing this sequence. The regulation of p50 DNA binding by S80 phosphorylation occurs both in the context of p50:65 heterodimers and p50 homodimers, revealing the regulation of the transcriptional activity of the p65 subunit in *trans* through the modification of p50.

In addition to the central role of IKKβ as an activator of the NF-κB pathway through the phosphorylation of IκBα, IKKβ also phosphorylates a number of other components of the NF-κB pathway (Christian et al., 2016). This includes NF-κB p105, which is phosphorylated by IKKβ at the C-terminus to promote its limited proteasomal degradation to generate the p50 subunit of NF-κB. NF-κB p65 is also phosphorylated by IKKβ at S468 (Schwabe and Sakurai, 2005) and S536 (Schwabe and Sakurai, 2005, Sakurai et al., 1999, Sizemore et al., 2002, Haller et al., 2002, Yoboua et al., 2010) which serves to regulate p65 transcriptional activity.

Our identification of p50 as an IKKβ substrate places p50 alongside these other components of the NF-κB pathways as a regulatory target of IKKβ kinase activity. Of note, the regulation of sequence specific p50 DNA binding by IKKβ phosphorylation reveals a novel mechanism of IKKβ-mediated control of NF-κB activity. The phosphorylation of p65 at S468 and S536 by a number of other kinases in addition to IKKβ (Buss et al., 2004) suggests that S80 of p50 is also likely to be phosphorylated by other kinases. The future identification of additional S80 kinases will shed further light on the role of S80 phosphorylation in the regulation of NF-κB dependent transcriptional responses in the context of different cellular stimuli.

The NF-κB barcode hypothesis proposes that post-translational modifications of NF-κB subunits, either alone or in combination, generate distinct functional

states that direct transcription in a gene specific manner (Moreno et al., 2010). This hypothesis has largely been generated from studies of p65 phosphorylation which indicate that individual sites of p65 phosphorylation may regulate the expression of distinct subsets of NF-κB target genes (Moreno et al., 2010). The molecular basis for such gene specific effects of p65 phosphorylation are not clear is many cases. Although studies have demonstrated that S468 phosphorylation promotes the ubiquitination of p65 by enhancing interaction with an E3 ligase complex containing COMMD1, GCN5, Cullin2 and SOCS1 at certain promoters (Geng et al., 2009), it is not understood what directs this interaction at these specific promoters. In this study we establish that the gene-specific transcriptional effects of S80 phosphorylation are mediated by the differential binding of phosphorylated p50 with κB sites containing an A nucleotide at the -1 position. Importantly, p50 S80 phosphorylation primarily affects the binding of p50:p65 heterodimers at these sites to inhibit gene transcription. While our data does not identify S80 phosphorylation as a regulator of p50 homodimer function, further analysis will be required to determine whether cell or signal specific factors control dimer specific effects of S80 phosphorylation. Phosphorylation of p50 at S328 has previously been demonstrated to inhibit the binding of p50 to κB sites containing a C nucleotide at the -1 position (Crawley et al., 2013). This study, together with our data, identifies the -1 position of κB sites as a critical factor in determining the transcriptional consequences of p50 phosphorylation. These data also establish κB site sequence as an additional and important component to be considered in further developing the NF-κB barcode hypothesis.

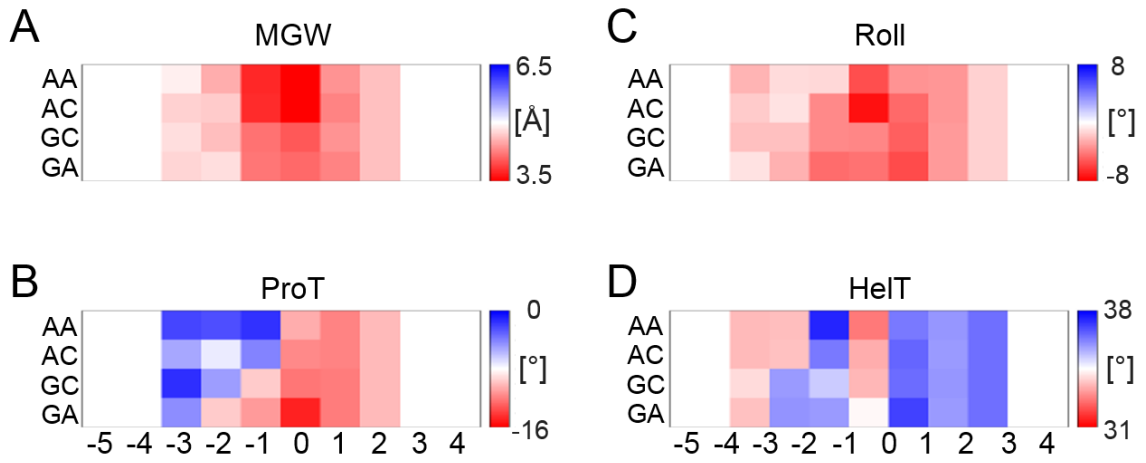
DNA binding sites can act as allosteric regulators of transcriptional regulators (Lefstin and Yamamoto, 1998). Distinct DNA conformations adopted by particular κB sequences provides a potential mechanism to explain why single nucleotide variations affect DNA binding of NF-κB dimers. In support of this an *in silico* transcription factor binding site shape analysis of the κB sites employed in our experiments predicts a unique conformation for each κB site (Yang et al., 2014) (Supplementary Figure 1). Previous studies have revealed that phosphorylation of p65 induces conformational changes that may influence the transcriptional outcome following DNA binding (Milanovic et al., 2014). Using similar approaches we assessed conformational differences between $p50^{WT}$ and $p50^{S80A}$ by limited

proteolytic digestion (Supplementary Figure 2). This analysis showed different sensitivities to digestion between $p50^{WT}$ and $p50^{S80A}$, indicating that S80 phosphorylation may induce a conformational change that could modify the binding of p50 dimers to κB sites containing an A nucleotide in the -1 position. Such conformational changes may modify the transactivating potential of NF-κB dimers by modifying DNA binding to specific sites, but also possibly by modifying interaction with other factors that in turn affect DNA binding.

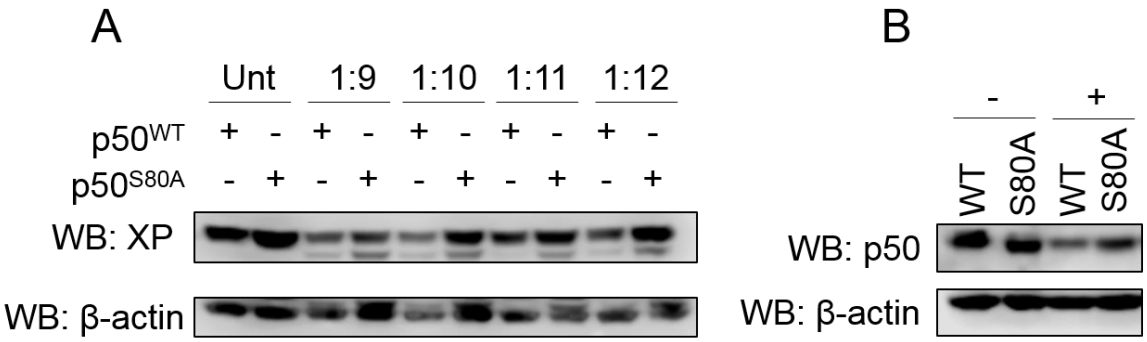
Our data demonstrate that the interaction of S80 phosphorylated p50 with κB sites containing an A at the -1 position is sufficient to inhibit DNA binding and gene transcription of associated genes and does not necessarily require the context of a promoter. Interestingly, our data also show that individual κB sequences have different capacities to drive transcription as measured by reporter assays incorporating κB sequences upstream of a minimal promoter. These analyses identified -2A, -1A κB sites as more potent drivers of transcription than other sites tested. Remarkably, the increased transcription from reporter vectors containing these κB sites is also reflected in TNFα stimulated cells where genes regulated by these sites are induced at significantly higher levels than genes regulated by other κB sites. Indeed, these κB sites appear to be enriched in the regulatory regions of genes encoding pro-inflammatory cytokines and chemokines, including TNFα, IL8 and the chemokines CXCL1 and CXCL2. The enrichment of specific κB sites in a functional class of genes provides strong evidence that NF-κB phosphorylation and κB sequences may establish regulatory networks to coordinate stimulus-specific transcriptional programmes.

In conclusion, this study describes a novel site of IKKβ phosphorylation of the p50 subunit that regulates TNFα induced transcriptional responses in a gene-selective manner. Our data demonstrates that the gene selective effect of S80 phosphorylation on transcription is encoded in the κB sequence, specifically by the -1 nucleotide position. The results of this study contributes further to our understanding of the regulation of transcriptional programmes by NF-κB. Future research may enable the prediction of transcriptional outcome based on an understanding of the phosphorylation status of NF-κB and DNA binding site sequence.

4.6 Supplementary data



Supplementary Figure 1: Transcription factor binding site shape analysis
Predicted DNA shape values for (A) minor groove width (MGW), (B) propeller twist (ProT), (C) roll and (D) helix twist (HelT) are displayed as heat maps. Each κB site analysed has the 10bp format GGG(-2)(-1)TTTCC, with nucleotide variations at the -2 and -1 position (indicated).



Supplementary Figure 2: S80 affects proteolysis of p50
(A) HEK293T cells were transfected with either $p50^{WT}$-XP or $p50^{S80A}$-XP plasmids as indicated. Whole cell lysates were either left untreated (Unt) or treated with varying trypsin:lysate (w/w) ratios as indicated. p50 protein degradation levels were analysed by western blot (WB) using anti-Xpress antibody. (B) Whole cell lysates extracted from TNFα treated (10ng/ml) WT or *NFKB1*^{S80A} cells were left untreated (-) or treated with trypsin at a 1:9 ratio of trypsin:lysate (w/w). p50 protein degradation levels were analysed by western blot (WB) using anti-p50 antibody.

Chapter 5

5 *NFKB1*^{S80A} THP-1 profiling analyses

5.1 Abstract

The NF-κB p50 subunit is a critical regulator of inflammatory gene expression, however, its regulation by phosphorylation is not well studied. *NFKB1* is one of the most highly expressed transcription factors in macrophages, and is central to regulating gene expression in macrophages. Therefore, understanding the mechanisms regulating p50 function is key to understanding essential inflammatory responses. In this chapter, *NFKB1*^{S80A} THP-1 cells generated using CRISPR/Cas9 genome editing were utilised to investigate the role of p50 S80 phosphorylation in a human macrophage model. It was found that S80 selectively regulates LPS-induced gene expression, in particular, the expression of pro-inflammatory genes. Transcription factor binding site analysis of LPS-induced gene promoters identified an enriched κB site within a distinct cluster of genes with elevated expression in *NFKB1*^{S80A} THP-1s compared to WT THP-1s. As a consequence, *NFKB1*^{S80A} THP-1 cells have an enhanced pro-inflammatory phenotype, supported by enrichment of gene sets related to inflammatory processes. In addition, the possibility of predicting the outcome of gene expression based on specific NF-κB binding sites was explored.

5.2 Introduction

The importance that the NF-κB p50 subunit has in regulating the immune response is clear from a number studies utilising *Nfkb1*$^{-/-}$ mouse models (described in section 1.4.1), and genome wide studies implicating *NFKB1* polymorphisms in human inflammatory diseases (Sun and Zhang, 2007). However, the particular role that p50 plays in the regulation of immune cells such as macrophages is becoming more apparent. A screen of human monocyte-derived macrophages treated with a number of different stimuli identified a broad spectrum of transcriptional profiles associated with macrophage activation (Xue et al., 2014). Within these transcriptional profiles, NFKB1 was identified as one of the most highly expressed transcription factors (Xue et al., 2014). In fact, it has been estimated that there are 30,000 p50 homodimers in the nuclei of resting bone marrow derived macrophages (BMDMs) (Cheng et al., 2011), suggesting that p50 plays an important role in regulating gene expression at basal level. Furthermore, p50 homodimers are important repressors of inflammatory gene expression, and their stability is crucial for limiting pro-inflammatory gene expression and establishing Toll-like Receptor tolerance in macrophages (Carmody et al., 2007, Yan et al., 2012). These studies implicate p50 as an important regulator of gene expression in macrophage. Therefore, identifying the mechanisms regulating p50 function in macrophages is essential to understanding transcriptional profiles controlling major processes of the immune response.

While studies of p50 function could be performed in mouse macrophage cell lines such as RAW264.7 or murine BMDM, it has been reported that there are substantial differences in TLR4-regulated gene expression between mice and humans (Schroder et al., 2012). Therefore, the transcriptional profiles of murine and human macrophages are not suitable for comparison. Although primary human macrophages could be cultured, their short *in vitro* life span and potential heterogeneity in inflammatory response limit the benefits of their use (Gordon et al., 2014). The most widely used cell line for investigating primary human macrophage function *in vitro* is the human leukaemia-derived monocytic cell line, THP-1. THP-1 cells can be differentiated into macrophage-like cells using stimuli such as phorbol 12-myristate-13-acetate (PMA), 1α, 25-dihydroxyvitamin D3 (vD3), or macrophage colony stimulating factor (M-CSF).

However, PMA-stimulated differentiation is generally preferred for generation of cells with similar phenotypic and functional characteristics of human peripheral blood mononuclear cell (PBMC) monocyte-derived macrophages (Chanput et al., 2014). Therefore, THP-1 cells are an appropriate model for studying human macrophage functions, mechanisms and signalling pathways. THP-1 cells have previously been used to study the LPS-induced NF-κB response, where expression profiles were similar to PBMC monocyte-derived macrophages, supporting use of THP-1s cells as a good model for inflammatory gene expression (Sharif et al., 2007). For this reason, *NFKB1*^{S80A} THP-1 cells were generated using CRISPR/Cas9 genome editing technology (Chapter 3) to be used as model to study S80 phosphorylation in human macrophage. While mutation of *NFKB1* S80 will provide insight into the role of NF-κB p50 phosphorylation in the regulation of inflammatory gene expression in human macrophages, it is important to consider how mutation of *NFKB1* might affect other signalling pathways in THP-1 cells. This is due to the ability of *NFKB1* to crosstalk with other transcription regulatory networks.

For example, ERK1/2 activation may be affected in THP-1 cells as a result of *NFKB1* mutation. While TNFα does not induce ERK1/2 phosphorylation in HEK293 cells (Bladh et al., 2009, Tu et al., 2014) ERK1/2 activation is induced by LPS in macrophages, regulated by the interaction of p105 with the kinase TPL-2. In macrophages, all detectable TPL-2 is associated with a pool of p105, inhibiting TPL-2 induced ERK1/2 activation (Beinke et al., 2003). Activated ERKs phosphorylate and activate a number of downstream targets, including transcription factors, which coordinate the regulation of a large number of genes involved in many biological processes (Roskoski, 2012). It has been shown in macrophages that following LPS-induced p105 processing, liberated TPL-2 specifically mediates activation of ERK1/2, and not other MAPKs such as p38 (Dumitru et al., 2000). Therefore, if p105 processing is impaired by S80A mutation, ERK1/2 activation may be defective, which in turn will affect expression of genes regulated by this pathway. The importance of NFKB1 in regulating ERK1/2 activation has been demonstrated using macrophages from *Nfkb1*$^{-/-}$ mice and *Nfkb1*^{SSAA} mice (Waterfield et al., 2003, Banerjee et al., 2006, Yang et al., 2012). *Nfkb1*^{SSAA} mice are defective in LPS-induced p105 processing due to mutation of IKKß target sites from serine to alanine residues. *Nfkb1*$^{-/-}$

macrophages are defective for TPL-2 activation of ERK1/2 signalling, which impairs production of IL-10 and COX2 (Waterfield et al., 2003, Banerjee et al., 2006). Similarly, *Nfkb1*^{SSAA} macrophages are defective for TPL-2 activation of ERK1/2, impairing expression of *Il6* and *Csf2* (Yang et al., 2012). In addition, it has been shown that NF-κB is capable of regulating a subset of interferon (IFN)-stimulated genes (Pfeffer, 2011). Moreover, p50 homodimers have been shown to binds to a subclass of IFN response elements (IRE) to repress IFN-inducible gene expression in macrophages (Cheng et al., 2011). Therefore, mutation of p50 S80 may have an effect on IFN gene expression.

The aims of this chapter were to:

1. Utilise CRISPR/Cas9 generated *NFKB1*^{S80A} knock-in THP-1 cells to investigate p50 S80 phosphorylation-driven transcriptional profiles in response to LPS.

2. Investigate mechanisms controlling LPS-induced selective gene expression.

3. Explore the phenotype of *NFKB1*^{S80A} knock-in THP-1 cells in an inflammatory context.

5.3 Results

5.3.1 NF-κB activation unaffected despite reduced p105/p50 protein levels in *NFKB1*^{S80A} THP-1 cells

Following the generation of *NFKB1*^{S80A} THP-1 cells (Figure 3.7, clone 10), the role of S80 phosphorylation in regulating NF-κB activity in human monocyte-derived macrophages was investigated. Immunoblot analysis of p105/p50 protein expression in whole cell lysates from undifferentiated wild type (WT) and *NFKB1*^{S80A} THP-1 cells revealed a slight reduction in both p105 and p50 protein levels in *NFKB1*^{S80A} cells (Figure 5.1A). This result was surprising, given that p105/p50 protein levels were equivalent between WT and *NFKB1*^{S80A} HEK293T cells (Chapter 4), suggesting a cell type-specific effect of the S80 mutation. To investigate the effect of S80 mutation on the activation of the NF-κB pathway, WT and *NFKB1*^{S80A} THP-1 cells were differentiated with 25ng/ml PMA for 3 days. Following a 4 day rest period, WT and *NFKB1*^{S80A} THP-1 cells were stimulated with 10ng/ml LPS, and IκBα phosphorylation and degradation by was assessed by immunoblot. This analysis demonstrated slightly lower levels of IκBα phosphorylation in *NFKB1*^{S80A} cells stimulated with LPS, compared to WT THP-1s (Figure 5.1B). However, the kinetics of IκBα phosphorylation are comparable between WT and in *NFKB1*^{S80A} THP-1s. Nuclear translocation of p50 and p65 in response to LPS was assessed in PMA-differentiated WT and *NFKB1*^{S80A} THP-1 cells. Immunoblot analysis of nuclear and cytoplasmic fractions demonstrated that p50 and p65 translocation to the nucleus following LPS stimulation is similar between WT and *NFKB1*^{S80A} THP-1 cells. Together, these data demonstrate that despite the slight reduction in p105/p50 protein levels in *NFKB1*^{S80A} THP-1 cells, the ability of p50 to translocate to the nucleus is not impaired.

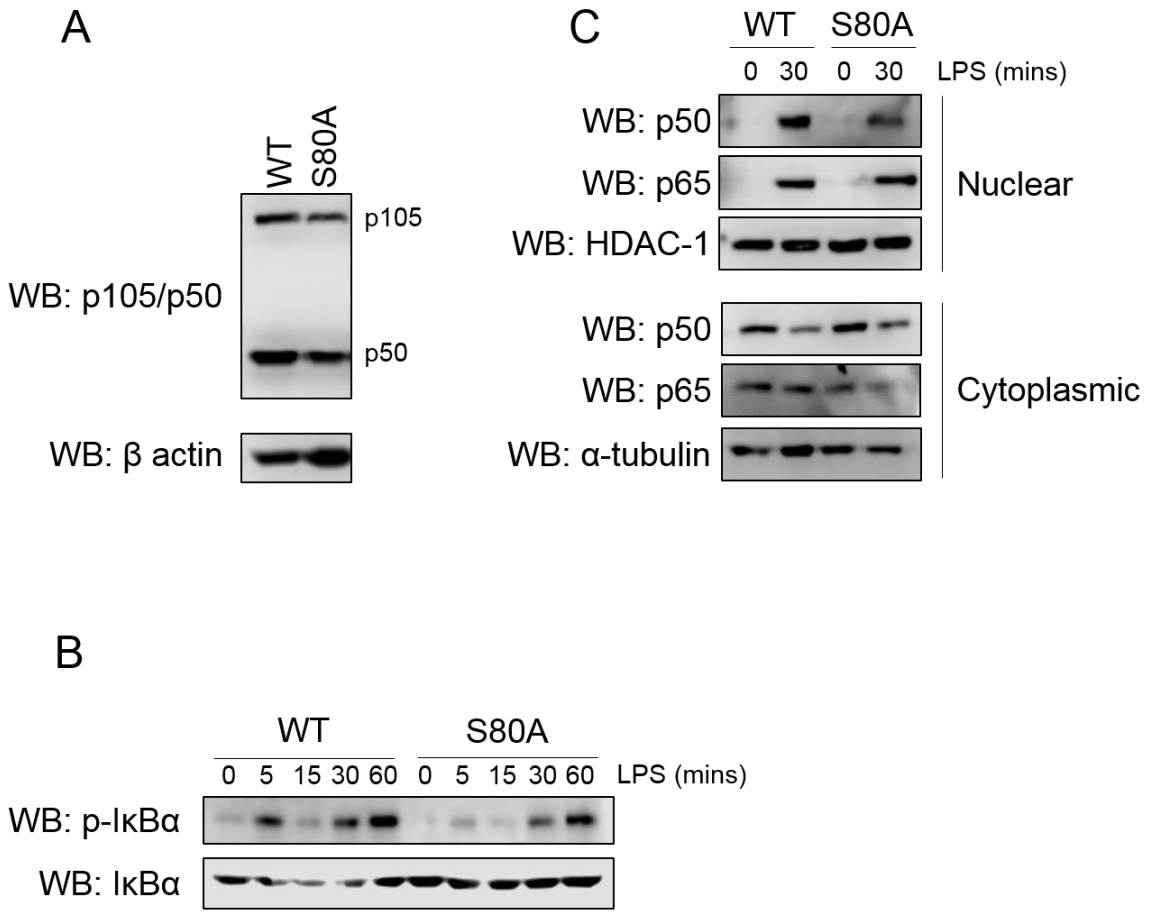


Figure 5.1: S80 phosphorylation is not required for NF-κB activation in THP-1 cells
(A) Whole cell lysates extracted from WT and *NFKB1*^{S80A} THP-1s were analysed by western blot (WB) to detect endogenous p105/p50 protein levels. (B) WT and *NFKB1*^{S80A} THP-1s were differentiated with 25ng/ml PMA for 3 days, and were stimulated with 10ng/ml LPS after 4 days rest for the indicated times prior to lysis. Whole cell extracts were analysed by WB to detect levels of phosphorylated and total IκBα protein. (C) WT or *NFKB1*^{S80A} THP-1s were differentiated with 25ng/ml PMA for 3 days, and were stimulated with 10ng/ml LPS after 4 days rest for the indicated times prior to lysis. Nuclear and cytoplasmic extracts were prepared and analysed for p50 and p65 proteins by WB.

5.3.2 ERK1/2 activation is increased in *NFKB1*^{S80A} THP-1 cells

Since p105 undergoes TLR-inducible processing to generate p50, the levels of p105/p50 protein in PMA-differentiated WT and *NFKB1*^{S80A} THP-1 cells were assessed in response to LPS treatment. Immunoblot analysis revealed a significant decrease of p50 protein levels in *NFKB1*^{S80A} THP-1 cells compared to WT THP-1 cells (Figure 5.2A). While it could be speculated that the S80A mutation may have an effect on LPS induced p105 processing, p50 protein levels in untreated *NFKB1*^{S80A} THP-1 cells are also reduced. The effect of impaired p105/p50 protein levels on the activation of the ERK1/2 MAPK pathway was next investigated. Immunoblot analysis revealed significantly increased levels of

phosphorylated ERK1/2 in both unstimulated and stimulated *NFKB1*^{S80A} THP-1 cells compared to WT THP-1 cells (Figure 5.2A). This effect was surprising, given that ERK1/2 activation has been shown to be dependent on p105 processing in murine macrophages (Yang et al., 2012). To investigate the effects of the S80A mutation on other MAPK signalling pathways, the p38 pathway was also examined (Figure 5.2A). Levels of phosphorylated p38 are comparable between WT and *NFKB1*^{S80A} THP-1 cells, suggesting that S80A mutation may specifically effect activation of ERK1/2, and not other MAPK pathways such as p38.

It has been shown that the period of rest following PMA exposure can influence some phenotypic properties and functional characteristics of THP-1 cells (Lund et al., 2016). In Figure 5.2A, THP-1 cells were stimulated with LPS immediately following PMA treatment. In order to rule out rest period as a factor that influences p105/p50 protein levels and upregulation of ERK1/2 signalling pathways, immunoblot analyses were repeated, however THP-1 cells were rested in culture for 4 days following PMA differentiation (Figure 5.2B). This analysis revealed a reduction in p105/p50 protein levels in *NFKB1*^{S80A} THP-1 cells compared to WT THP-1s (Figure 5.2B), consistent with the results observed in undifferentiated THP-1 cells (Figure 5.1A). However, the total level of p105/p50 protein in differentiated *NFKB1*^{S80A} THP-1 cells appears to be less compared to undifferentiated *NFKB1*^{S80A} THP-1 cells. The level of p50 protein in *NFKB1*^{S80A} THP-1 cells rested for 4 days is higher in comparison to *NFKB1*^{S80A} THP-1 cells that were not rested in culture (Figure 5.2B and Figure 5.2A). This analysis suggests that the S80 mutation affects endogenous p105/p50 protein levels in differentiated THP-1 cells, and that THP-1 differentiation protocol appears to have an effect on p105/p50 protein ratios.

This analysis also revealed that ERK1/2 activation is still increased in *NFKB1*^{S80A} THP-1 cells following a 4 day rest period (Figure 5.2B). However, the kinetics of ERK1/2 activation are different in both WT and *NFKB1*^{S80A} THP-1 cells. With 0 days rest, ERK1/2 activation is only visible in WT THP-1 cells after 60 min LPS stimulation, whereas in *NFKB1*^{S80A} THP-1 cells there is constitutive activation (Figure 5.2A). However, following 4 days rest, ERK1/2 activation in WT THP-1 cells peaks after 15 mins. In *NFKB1*^{S80A} THP-1 cells, ERK1/2 activation also peaks after 15 mins, but is sustained (Figure 5.2B). While levels of phosphorylated p38

appear reduced in *NFKB1*^{S80A} THP-1 cells, the kinetics of p38 activation are comparable to that of WT cells, and reflect the findings of Figure 5.2A. Lower phosphorylated p38 levels observed are perhaps due to reduced levels of total p38 protein (Figure 5.2B). Together, these data suggest that the shorter differentiation protocol appears to exaggerate differences in p105/p50 protein level and ERK1/2 activation observed in *NFKB1*^{S80A} THP-1 cells. Therefore, the 7 day differentiation protocol (3 days PMA differentiation followed by 4 days rest) was utilised for further analyses such as RNA-seq and qPCR described in the next section (5.3.3).

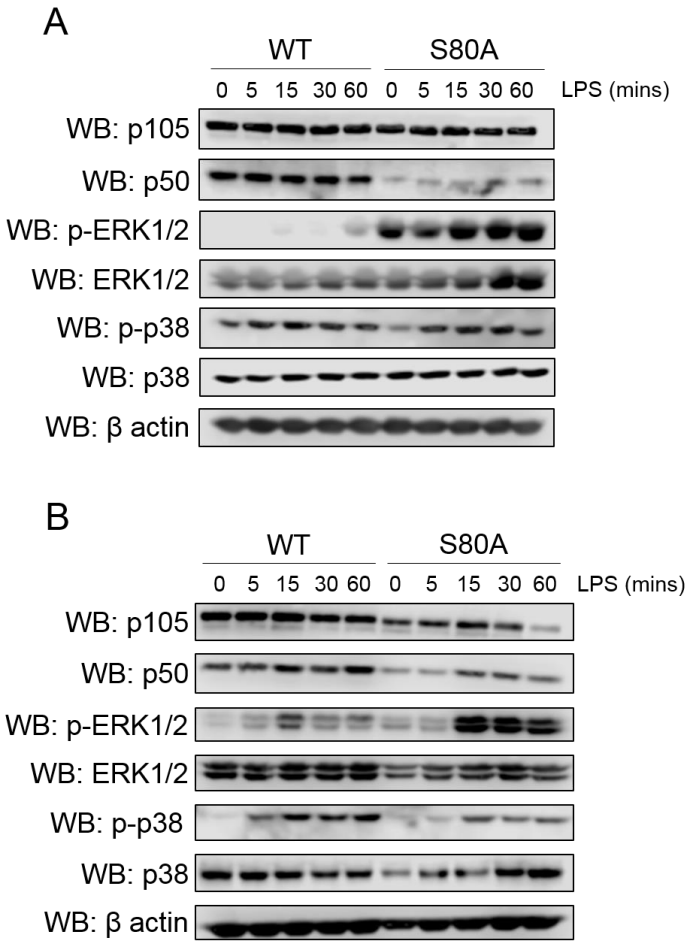


Figure 5.2: p105/p50 protein levels and ERK1/2 activation altered in differentiated *NFKB1*^{S80A} THP-1 cells

WT and *NFKB1*^{S80A} THP-1 cells differentiated with 25ng/ml PMA for 3 days and were either treated with 10ng/ml LPS after 0 days rest (A) or 4 days rest (B) for the indicated times prior to lysis. Whole cell lysates extracted from WT and *NFKB1*^{S80A} THP-1s were analysed by western blot (WB) to detect endogenous p105/p50 protein levels and levels of phosphorylated and total ERK1/2 and p38 protein.

5.3.3 Transcriptomic analysis of *NFKB1*^{S80A} THP-1 cells

5.3.3.1 RNA-seq analysis workflow

In order to investigate the role of S80 phosphorylation in regulating NF-κB target gene expression, LPS-induced transcriptional responses in WT and *NFKB1*^{S80A} THP-1 cells were analysed by RNA-seq. WT and *NFKB1*^{S80A} THP-1 cells were left untreated or treated with LPS for 3 hours prior to RNA-seq analysis. RNA-seq data was analysed using two different pipelines (Schematic shown in Figure 5.3A). Differentially expressed genes were identified by using two different methods. CuffDiff is a well-established method which calculates differential expression at transcript level. However, it has been reported that this method can produce slightly more false positive results compared to other methods such as DESeq2 (Rapaport et al., 2013, Seyednasrollah et al., 2015, Schurch et al., 2016). Therefore, a combination of both CuffDiff and DESeq2 methods were utilised to ensure robust analysis of RNA-seq data.

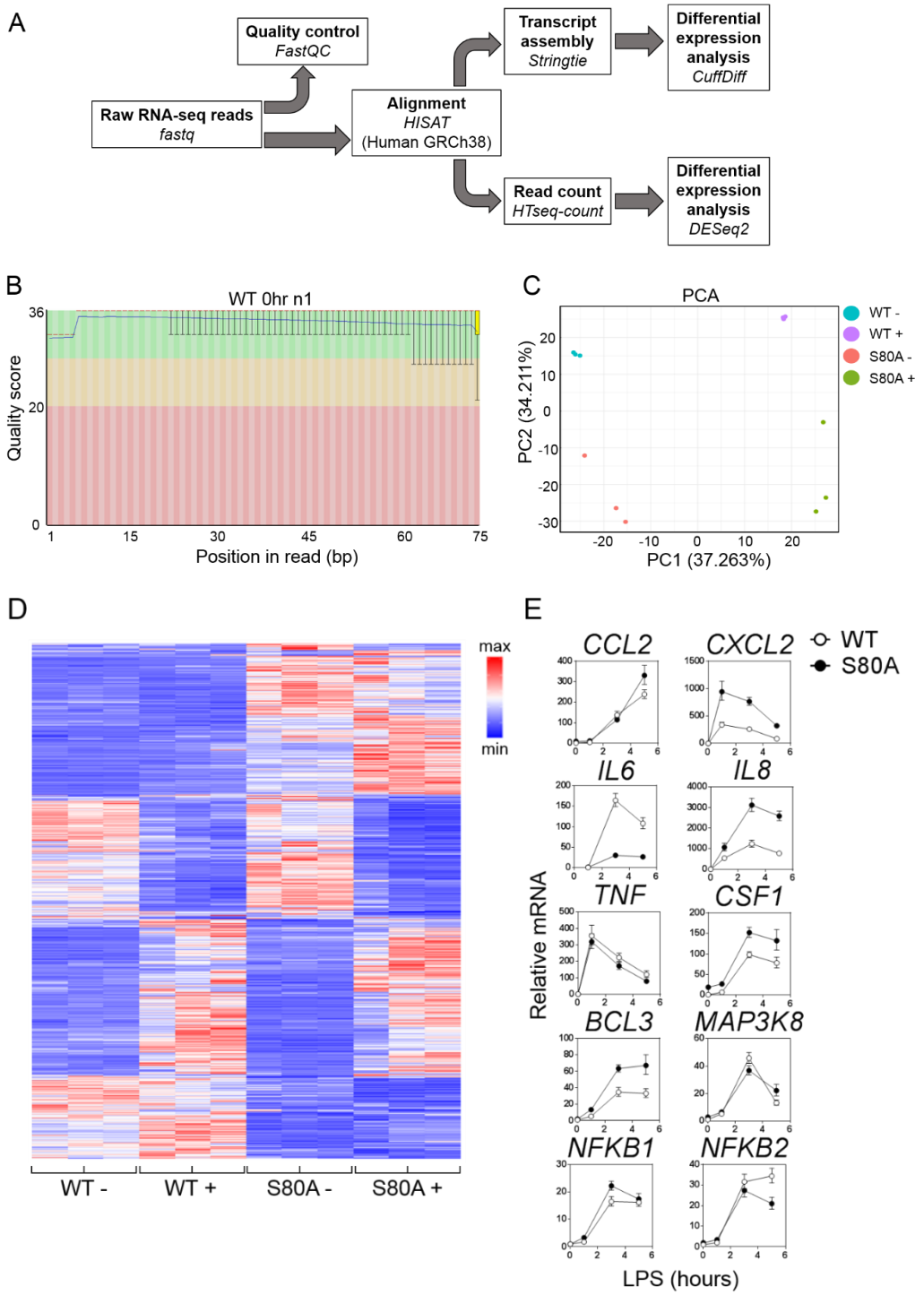


Figure 5.3: p50 S80 phosphorylation selectively regulates gene expression in THP-1s
(A) Workflow of RNA-seq analysis including CuffDiff and DESeq2 methods of calculating differentially expressed genes. (B) A representative per-base sequence quality plot generated by FastQC. The central red line represents the median value. The yellow box represents the inter-quartile range (25-75%). The upper and lower whiskers represent the 10% and 90% points. The blue line represents the mean quality. The background of the graph divides the y axis into quality of called: very good (green), reasonable (orange), and poor (red). (C) Principal component analysis (PCA) of DESeq2 analysed RNA-seq data, showing the first two components. Samples are represented by dots. The percentage of total variation explained by each component is given. (D) Triplicate samples of WT and *NFKB1*^{S80A} THP-1 cells treated with LPS for 3 hours (+) or untreated (-) were analysed by RNA-seq. The heat maps display differentially expressed genes calculated by DESeq2 (p.adj <0.01, replicates shown). Normalised read counts are scaled as per z-score. (E) WT and *NFKB1*^{S80A} THP-1s were stimulated with 10ng/ml LPS for the indicated times prior to harvest and RNA extraction. Gene expression levels were determined by quantitative real-time PCR. Data are mean ± S.E of triplicate samples and are representative of three independent experiments.

The FastQC tool was used to check the per-base sequence quality of raw RNA-seq data (Anders et al., 2015). All samples passed this quality control check, with high quality scores across all bases. A representative per-base sequence quality plot is shown in Figure 5.3B. Principal component analysis (PCA) was also performed to investigate the distribution of RNA-seq samples (Figure 5.3C). The PCA plot shows clear separation between each of the RNA-seq samples, and tight clustering of replicates, indicating little variation between each replicate.

5.3.3.2 S80 phosphorylation selectively regulates gene expression

RNA-seq analysis revealed distinct transcriptional profiles in both WT and *NFKB1*^{S80A} THP-1 cells (Figure 5.3D). Two distinct groups of differentially expressed genes were identified: LPS-independent and LPS-dependent. Within the LPS-independent group, there are genes expressed in WT THP-1 cells which are not expressed in *NFKB1*^{S80A} THP-1 cells. Similarly, there are genes expressed in *NFKB1*^{S80A} THP-1 cells which are not expressed in WT THP-1 cells. The transcriptional profiles of unstimulated cells suggests that S80 phosphorylation regulates the ability of p50 to control gene expression at basal level, potentially as part of a p50 homodimer. Within the LPS-dependent group, genes are both upregulated and downregulated in both WT and *NFKB1*^{S80A} THP-1 cells. Upon further inspection of LPS-inducible genes, the magnitude of gene expression differs between WT and *NFKB1*^{S80A} THP-1 cells (Figure 5.3). For example, while the expression of some genes remains equivalent between WT and *NFKB1*^{S80A} THP-1 cells, expression levels of other genes are upregulated to a greater extent in *NFKB1*^{S80A} THP-1 cells in comparison to WT THP-1s. This suggests a selective effect of the NFKB1 S80A mutation, and indicates that S80 phosphorylation might regulate p50 as part of a p50:p65 heterodimer in THP-1 cells.

To validate RNA-seq results, expression of selected genes with an identifiable κB site in their promoter region were independently measured by qPCR (Figure 5.3E). This analysis confirmed evidence of selective gene regulation identified by RNA-seq. Of note, the expression of pro-inflammatory chemokines and cytokines is regulated in a selective manner. Expression of *CCL2* and *TNFa* is similar between WT and *NFKB1*^{S80A} THP-1 cells, whereas expression of *CSF1*, *CXCL2* and *IL8* is much higher in *NFKB1*^{S80A} THP-1 cells compared to WT cells. In contrast, *IL6* expression is dramatically increased in WT THP-1 cells compared to *NFKB1*^{S80A}

THP-1 cells (Figure 5.3). These data suggest that S80 phosphorylation regulates LPS-induced transcriptional responses in a gene-selective manner. These results are consistent with those in Chapter 4, which analysed TNFα-induced gene expression in HEK293T cells. However, the expression pattern of *TNFα* differs between HEK293T cells and THP-1 cells, suggesting cell type or ligand specific expression patterns of certain genes. In addition, S80 phosphorylation appears to selectively regulate the expression of NF-κB modulators. Expression of *BCL3* is higher in *NFKB1*^{S80A} THP-1 cells than in WT THP-1s, whereas *MAP3K8*, *NFKB1*, and *NFKB2* expression is equivalent (Figure 5.3). Of note, *NFKB1* expression is slightly higher in *NFKB1*^{S80A} THP-1 cells, suggesting that differences in p105/p50 protein levels in *NFKB1*^{S80A} THP-1 cells may be controlled post-translationally. Together, these data indicate that S80 phosphorylation selectively regulates LPS-induced gene expression, including pro-inflammatory genes.

Although RNA-seq analysis identified a number of LPS-independent genes, these genes were not further analysed as part of this thesis. However, it is important to note that the unstimulated transcriptional profiles of WT and *NFKB1*^{S80A} THP-1 cells may provide insight into how S80 phosphorylation regulates expression of genes at basal level. Specifically, it suggests that S80 phosphorylation may regulate the activity of p50 as part of a p50 homodimer, thereby controlling gene expression in unstimulated macrophages. Due to the findings that p50 S80 phosphorylation predominantly regulates transcription through p50:p65 heterodimers to limit NF-κB driven pro-inflammatory gene expression (Chapter 4), further analyses in this chapter were focussed on LPS-induced genes.

5.3.4 Transcription factor and binding site enrichment analyses

5.3.4.1 NF-κB is the core regulator LPS-induced gene expression in *NFKB1*^{S80A} THP-1s

To verify that NF-κB is the core transcription factor responsible for LPS-induced transcriptional profiles in *NFKB1*^{S80A} THP-1 cells, upstream regulator enrichment analysis was performed. This analysis was used to identify the upstream transcriptional regulators that explain observed gene expression changes within LPS-induced WT THP-1 (Figure 5.4A and B) and LPS-induced *NFKB1*^{S80A} THP-1 (Figure 5.4D and E) datasets. Upstream regulator analysis is based on prior knowledge of expected effects between transcriptional regulators and their target genes, available from the TRRUST database (Han et al., 2018). Enriched upstream regulators are calculated based on significant overlap between dataset genes and known targets controlled by a transcriptional regulator (overlap p-value). An activation score for each transcriptional regulator is calculated based on comparing the predicted directionality against the observed directionality of expression of each gene controlled by the upstream regulator (z-score). This analysis revealed a similar enrichment profile between LPS-induced WT and *NFKB1*^{S80A} THP-1 cells (Figure 5.4C and Figure 5.4F). Importantly, NF-κB transcription factors such as RELA, NFKB1 and REL are significantly enriched within LPS-induced genes in both WT and *NFKB1*^{S80A} THP-1 cells. This suggests that NF-κB driven gene transcription is not impaired in *NFKB1*^{S80A} THP-1 cells, despite reduced p50 protein levels. In addition, it confirms that NF-κB is a strong contributor to the LPS-induced transcriptional profile observed *NFKB1*^{S80A} THP-1 cells. If the transcriptional profile of *NFKB1*^{S80A} THP-1 cells was predominantly caused by increased ERK1/2 activation, it would perhaps be expected to observe enrichment of more ERK1/2 activated transcription factors such as ELK1 (Roskoski, 2012). In addition, it might be expected that ERK1/2 activated transcription factors have a higher enrichment score than NF-κB transcription factors in *NFKB1*^{S80A} THP-1 cells. For example, JUN and SP1 are only slightly more enriched in *NFKB1*^{S80A} THP-1 cells in comparison to WT THP-1s (Figure 5.4F and Figure 5.4C). In addition, while SP1 is enriched, its activation is not significant in WT THP-1 cells or in *NFKB1*^{S80A} THP-1 cells. While the slightly higher enrichment of JUN and SP1 in *NFKB1*^{S80A} THP-1 cells may be an effect of increased ERK1/2 activation in *NFKB1*^{S80A} THP-1 cells, JUN and SP1 enrichment is

not more prominent than RELA and NFKB1. These data confirm that NF-κB is central to the LPS-induced transcriptional profiles observed in both WT and *NFKB1*^{S80A} THP-1 cells, and that NF-κB transcription is functional in *NFKB1*^{S80A} THP-1 cells.

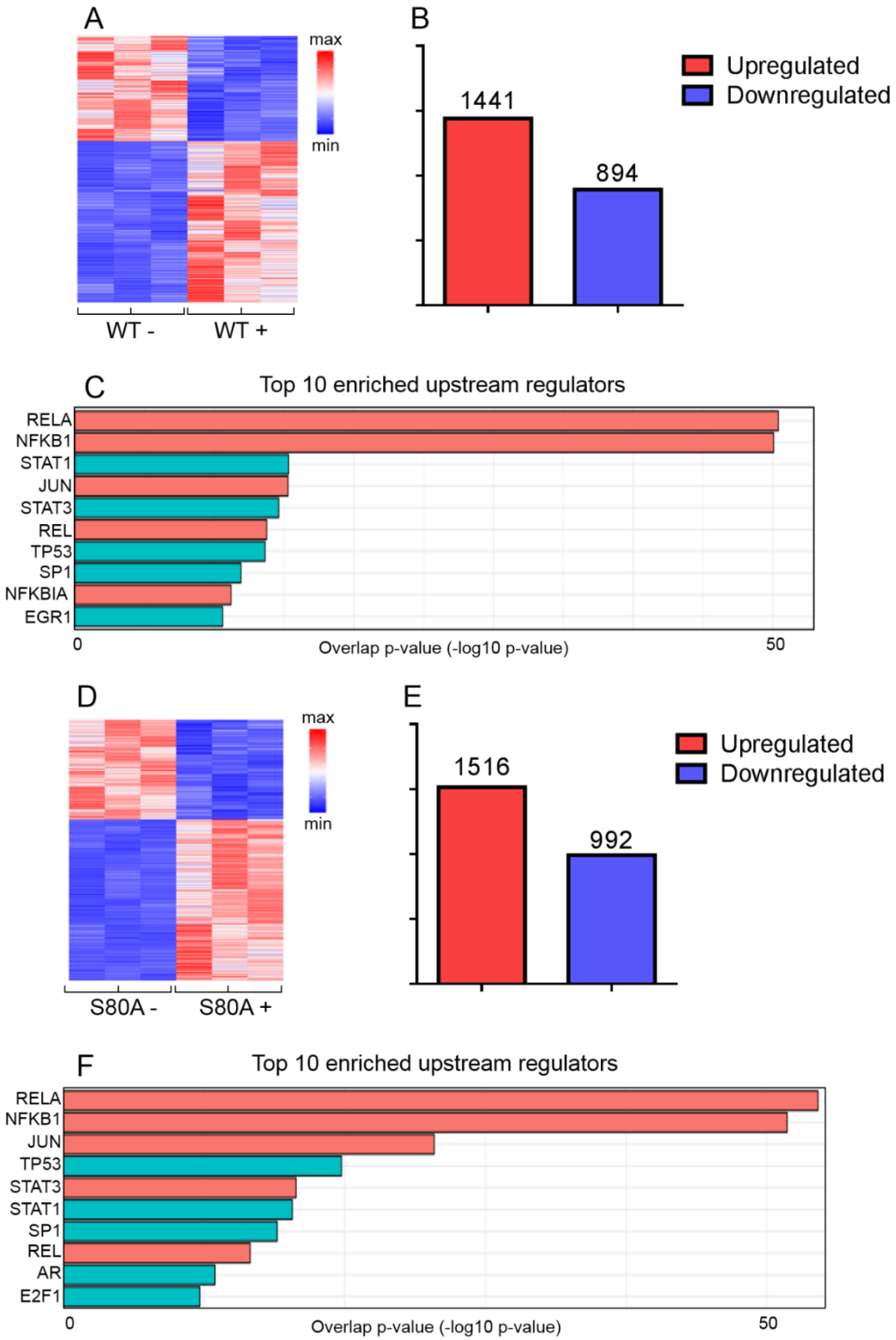


Figure 5.4: Upstream regulator enrichment analysis of LPS-induced genes
The heat maps display differentially expressed genes calculated by DESeq2 (p.adj <0.01, replicates shown). Normalised read counts are scaled as per z-score for LPS treated WT THP-1 cells (A) and LPS treated *NFKB1*^{S80A} THP-1 cells (D). The number and directionality of LPS-induced genes is indicated for WT THP-1 cells (B) and *NFKB1*^{S80A} THP-1 cells (E). Upstream regulator enrichment analysis of LPS-induced genes from WT THP-1 cells (C) and LPS-induced genes from *NFKB1*^{S80A} THP-1 cells (F). Upstream regulators were ranked by enrichment score (overlap p-value (–log10 p-value)) and by activation score (z-score). Red bars indicate both significant enrichment score (p-value <0.05) and significant activation score (significant z-score >2). Blue bars indicate significant enrichment (p-value <0.05), but not significant activation score (z-score <2).

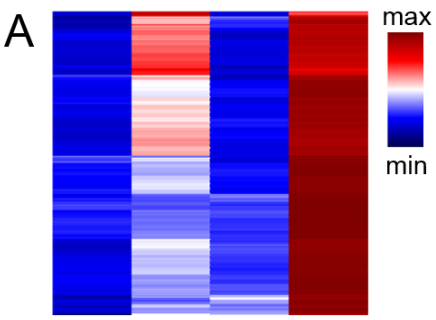
5.3.4.2 Specific κB site enrichment within a distinct cluster of LPS-induced genes

Having identified NF-κB transcription factors as enriched upstream regulators, enrichment of specific NF-κB binding sites was next explored. Specificity of NF-κB binding sites provides a mechanism to explain the potential selective effects of the S80 phosphorylation on LPS-induced transcription in THP-1 cells. Firstly, K means clustering of differentially expressed genes was performed to identify groups of genes which share similar expression profiles. This initial analysis revealed 20 clusters consisting of both LPS independent and LPS dependent genes (Figure 7.2). LPS-inducible gene clusters were combined based on the directionality of gene expression, creating groups of upregulated and downregulated genes. Genes were then categorised by magnitude of expression. This created 6 distinct groups of LPS-inducible genes, where expression was either up/downregulated to a greater extent in either WT or *NFKB1*^{S80A} THP-1 cells, or where gene expression was equivalent between WT and *NFKB1*^{S80A} THP-1 cells. In order to identify enriched NF-κB sites within each group of selectively regulated genes, transcription factor binding site analysis was performed using the *de novo* binding motif discovery program HOMER (Heinz et al., 2010). The promoters of genes belonging to each group were searched for *de novo* binding motifs +300bp/-50bp relative to the transcription start site. The most highly enriched DNA binding motif for each group of genes analysed is displayed in Figure 5.5 and Figure 5.6. In addition, the transcription factor best matched to the enriched *de novo* motif predicted by HOMER is shown.

Interestingly, interferon-sensitive response elements (ISRE) are enriched in the promoters of genes upregulated to a greater extent in WT THP-1 cells compared to *NFKB1*^{S80A} THP-1 cells (Figure 5.5B) and in the promoters of genes upregulated to an equivalent level in WT and *NFKB1*^{S80A} THP-1 (Figure 5.5C). Although LPS is not the best suited stimulus to investigate the induction of IFNs, these results indicate a potential role for S80 phosphorylation in the transcriptional control of IFN response. Of note, there are sequence differences in the enriched motifs in promoters of genes expressed higher in WT THP-1s compared to *NFKB1*^{S80A} THP-1s (Figure 5.5B) and between genes expressed equivalently (Figure 5.6C). These differences can be observed at the -6,-3,-2 and -1 positions of the binding site. While the promoters of genes expressed higher in WT THP-1s compared to

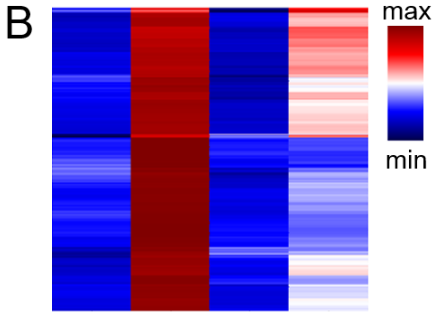
NFKB1^{S80A} THP-1s contain motifs that are ambiguous at these positions (-6G/C, -3T/G, -2G/C, -1A/G) (Figure 5.5B), the promoters of genes expressed equivalently between WT THP-1s and *NFKB1*^{S80A} THP-1s contain motifs that appear to be enriched with specific nucleotides at these positions (-6G, -3T, -2C, -1A) (Figure 5.5C). This suggests that regulation of p50 at these sites could be sequence specific. In addition, the promoters of genes upregulated equivalently between WT and *NFKB1*^{S80A} THP-1 cells are also enriched with an NF-κB binding site. However, the κB site has a high degree of degeneracy, and is not specifically enriched with particular bases at each position of the DNA binding site (Figure 5.5C).

Promoters of genes highly upregulated in *NFKB1*^{S80A} THP-1 cells compared to WT THP-1 cells are also enriched with an NF-κB binding site (Figure 5.5A). This *de novo* motif contains a specific 10 base pair κB site (consensus $G^{-5}G^{-4}G^{-3}A^{-2}A^{-1}T^{0}T^{1}T^{2}C^{3}C^{4}$). Expression of genes enriched with this κB site in their promoters are enhanced in *NFKB1*^{S80A} THP-1 cells, suggesting S80 phosphorylation regulates p50 activity in a κB site-specific manner. This is in contrast to the degenerate motif found to be enriched in promoters of genes expressed equivalently between WT and *NFKB1*^{S80A} THP-1 cells. Therefore, this analysis suggests that there are specific DNA-binding motifs associated with differential regulation of NF-κB target genes by S80. Of note, this 10bp κB site is identical to one of the most enriched motifs found in a gene cluster with the same expression profile in HEK293T cells (Chapter 4). The 10bp κB site $G^{-5}G^{-4}G^{-3}A^{-2}A^{-1}T^{0}T^{1}T^{2}C^{3}C^{4}$ was identified as being enriched in the promoters of gene highly expressed in *NFKB1*^{S80A} HEK293T cells in comparison to WT HEK239Ts (Chapter 4). This results validates the enrichment of the specific κB site $G^{-5}G^{-4}G^{-3}A^{-2}A^{-1}T^{0}T^{1}T^{2}C^{3}C^{4}$ within promoters of genes regulated in the same manner genes across two different cell types, and two different stimuli. In Chapter 4, S80 phosphorylation was shown to regulate the affinity of p50 to -1A containing κB sites, limiting gene transcription. The THP-1 κB site enrichment analysis suggests that the selective effect of the S80 mutation on NF-κB target gene expression in response to TNFα treatment in HEK293T cells may also regulate LPS-induced NF-κB target gene expression THP-1 cells.



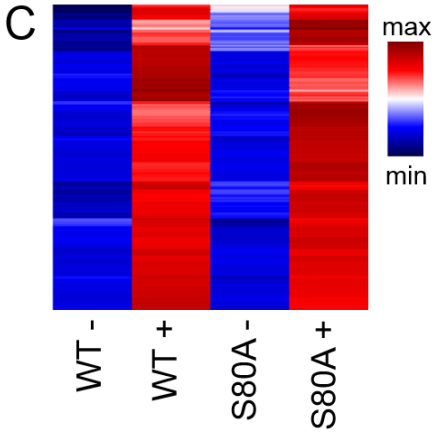
Expression upregulated in *NFKB1*^{S80A}

Rank	P-value	Best Match
1	1e-20	RELA



Expression upregulated in WT

Rank	P-value	Best Match
1	1e-14	ISRE



Expression level similar

Rank	P-value	Best Match
1	1e-14	ISRE

Rank	P-value	Best Match
2	1e-10	RELA

Figure 5.5: Transcription factor binding site analysis of LPS-induced upregulated genes
LPS-induced upregulated genes were categorised into 3 groups: (A) Expression upregulated to a greater extent in *NFKB1*^{S80A} THP-1s compared to WT THP-1s (B) Expression upregulated to a greater extent in WT THP-1s compared to *NFKB1*^{S80A} THP-1s. (C) Expression upregulated to similar level in WT and *NFKB1*^{S80A} THP-1s. The heat maps display the average FPKMs of LPS-induced genes calculated by CuffDiff (scaled as per z-score). Promoters of genes from each cluster were analysed using HOMER motif analysis software. Significantly enriched *de novo* motifs (p-value <0.001) and best matched transcription factors are shown.

Genes which are downregulated in response to LPS do not appear to have NF-κB binding sites enriched in their promoters (Figure 5.6). Promoters of genes downregulated to a similar extent in WT and *NFKB1*^{S80A} THP-1 cells, and to a greater extent in *NFKB1*^{S80A} THP-1 cells compared to WT THP-1s are enriched with motifs that are similar to those bound by zinc-finger containing transcriptional regulators GFI1B and ZFP161, respectively (Figure 5.6A and Figure 5.6C). Promoters of genes downregulated to a greater extent in WT THP-1 cells compared to *NFKB1*^{S80A} THP-1s are enriched with motifs similar to the ETS1 transcription factor site (Figure 5.6B).

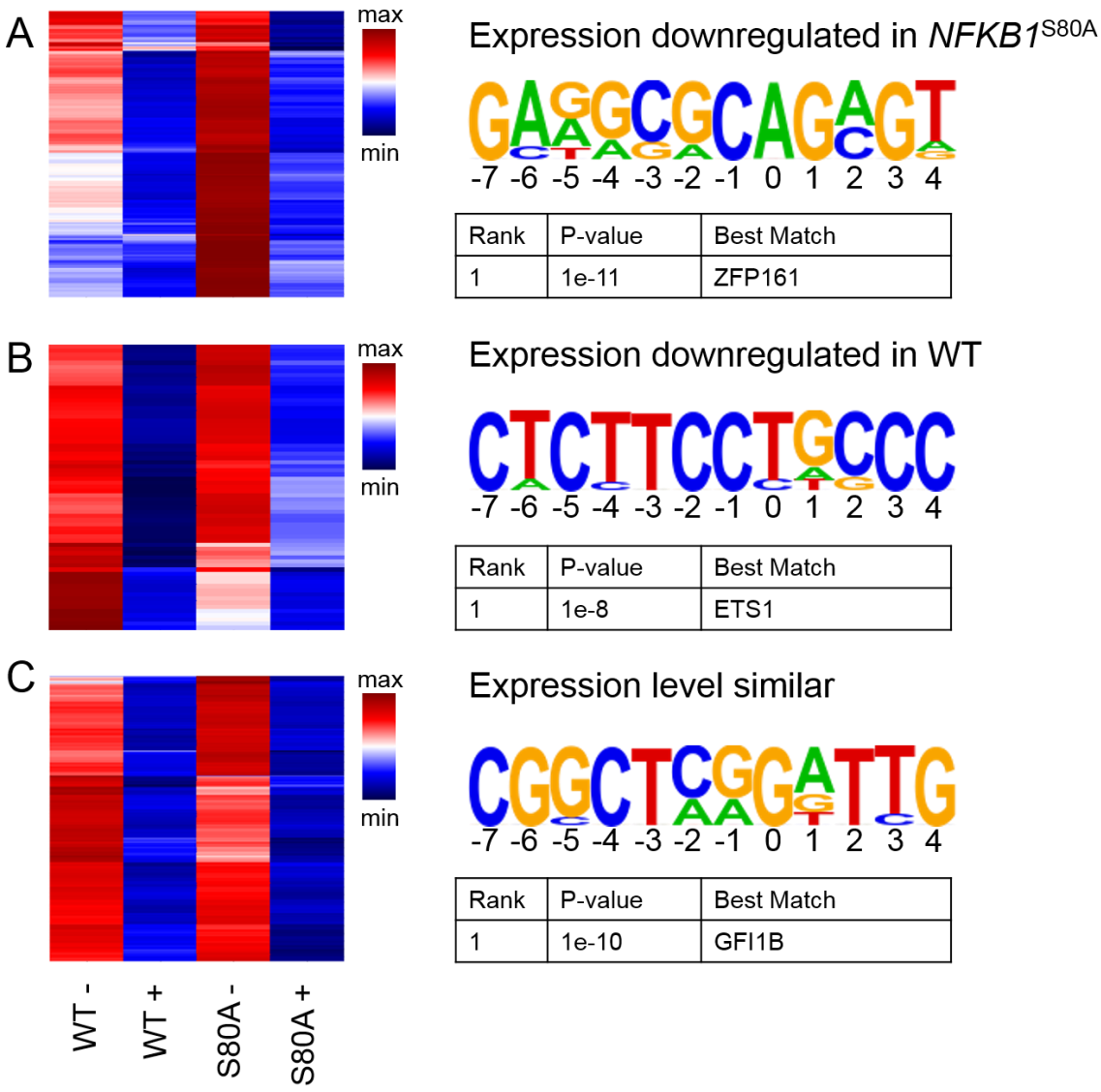


Figure 5.6: Transcription factor binding site analysis of LPS-induced downregulated genes
LPS-induced downregulated genes were categorised into 3 groups: (A) Expression downregulated to a greater extent in *NFKB1*^{S80A} THP-1s compared to WT THP-1s (B) Expression downregulated to a greater extent in WT THP-1s compared to *NFKB1*^{S80A} THP-1s. (C) Expression downregulated to similar level in WT and *NFKB1*^{S80A} THP-1s. The heat maps display the average FPKMs of LPS-induced genes calculated by CuffDiff (scaled as per z-score). Promoters of genes from each cluster were analysed using HOMER motif analysis software. Significantly enriched *de novo* motifs (p-value <0.001) and best matched transcription factors are shown.

5.3.5 Gene set enrichment analysis (GSEA)

5.3.5.1 Enhanced pro-inflammatory phenotype of *NFKB1*^{S80A} THP-1 cells

To further characterise the phenotype of the transcriptional profile of *NFKB1*^{S80A} THP-1 cells, gene set enrichment analysis was performed. Differentially expressed genes from LPS stimulated WT THP-1 cells vs *NFKB1*^{S80A} THP-1 cells (WT + vs S80A +) were ranked by fold change expression value. Genes expressed higher in *NFKB1*^{S80A} THP-1s compared to WT THP-1s feature at the top of the list, and genes expressed higher in WT THP-1s compared to *NFKB1*^{S80A} THP-1s feature at the bottom of the list. The ranked gene list was searched for enriched gene sets from the Molecular Signatures Data Base (MSigDB), using the GSEA online tool (available at http://software.broadinstitute.org/gsea/msigdb/index.jsp). GSEA calculates an enrichment score (ES) which reflects how overrepresented a gene set is at the top or bottom of a ranked list. A positive ES indicates gene set enrichment at the top of the ranked list (gene expression higher in *NFKB1*^{S80A} THP1s compared to WT THP-1s), and a negative ES indicates gene set enrichment at the bottom of the ranked list (gene expression higher in WT THP-1s compared to *NFKB1*^{S80A} THP-1s).

The ranked gene list was first searched for enrichment of the hallmark gene sets on MSigDB. The GSEA hallmark collection provides refined gene sets derived from multiple founder gene sets, providing robust enrichment analysis results (Liberzon et al., 2015). The top 3 significantly enriched hallmark gene sets for the top and bottom of the ranked gene list are shown in Figure 5.7A and Figure 5.7B, respectively.

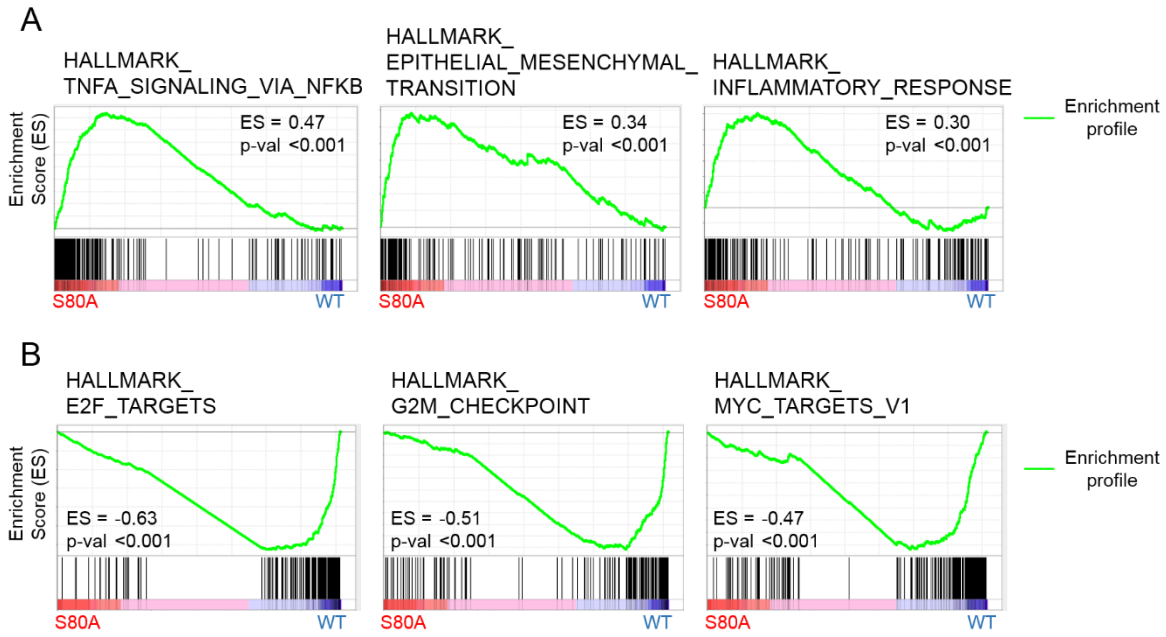


Figure 5.7: Enriched hallmark gene sets in LPS-treated WT vs *NFKB1*^{S80A} THP-1 cells
Gene set enrichment analysis of WT + vs S80A + ranked gene list (CuffDiff). Enrichment plots show the enrichment profile (green curve) across the ranked gene list. Black lines represent the genes which contribute to the enrichment profile. Peak enrichment scores (ES) and p-values are shown. (A) Top 3 significantly enriched hallmark gene sets at the top of the ranked gene list, indicated by positive enrichment score (expression higher in *NFKB1*^{S80A} THP-1s compared to WT THP-1s). (B) Top 3 significantly enriched hallmark gene sets at the bottom of the ranked gene list, indicated by negative enrichment score (expression higher in WT THP-1s compared to *NFKB1*^{S80A} THP-1s).

Significantly enriched gene sets at the top of the ranked list (gene expression higher in *NFKB1*^{S80A} THP-1s compared to WT THP-1s) include TNFα signalling via NF-κB and inflammatory response. Although THP-1 cells were stimulated with LPS, enrichment of these gene sets indicates that the induction of the inflammatory response in *NFKB1*^{S80A} THP-1s is coordinated by NF-κB gene expression. In addition, genes expressed higher in *NFKB1*^{S80A} THP-1s compared to WT THP-1s are enriched with the epithelial mesenchymal transition gene set. This hallmark collection is comprised of genes sets associated with reorganisation of the extracellular cellular matrix, and expression of matrix metalloproteinases and integrins (Liberzon et al., 2015). Expression of these genes could be associated with important processes performed by active macrophages such as migration, adhesion, phagocytosis and tissue repair (Wynn and Vannella, 2016). Enrichment of this gene sets suggests that S80 phosphorylation could regulate genes involved in these important inflammatory processes. Significantly enriched gene sets at the bottom of the ranked list (gene expression higher in WT THP-1s compared to *NFKB1*^{S80A} THP-1s) are related to

progression through the cell cycle and cell proliferation, indicated by the enrichment of E2F and MYC transcription factor targets, and genes involved in the G2/M checkpoint. Of note genes expressed higher in WT THP-1s compared to *NFKB1*^{S80A} THP-1s were enriched with the interferon gamma response and interferon alpha response hallmark gene sets (Figure 7.3). This indicates that genes expressed higher in WT THP-1s compared to *NFKB1*^{S80A} are involved in mediating a response to IFNγ. This is consistent with the transcription factor binding site enrichment analysis, which revealed enrichment of ISRE motifs in promoters of genes expressed to a greater extent in WT THP-1s compared to *NFKB1*^{S80A} THP-1s (Figure 5.7B).

In order to further explore the phenotype of *NFKB1*^{S80A} THP-1 cells, pathway analysis was performed using Kyoto Encyclopedia of Genes and Genomes (KEGG) gene sets available on MSigDB (Figure 5.8). Genes expressed higher in *NFKB1*^{S80A} THP-1s compared to WT THP-1s are associated with the lysosome pathway, indicating a potential increase in the phagocytic ability of *NFKB1*^{S80A} THP-1 cells (Figure 5.8A). In addition, enriched gene sets include focal adhesion and regulation of actin cytoskeleton, indicating that *NFKB1*^{S80A} THP-1 cells have the potential to be motile (Figure 5.8A). It is important to note that MAPK signalling was an enriched pathway associated with genes expressed higher in *NFKB1*^{S80A} THP-1s compared to WT THP-1s (Figure 7.3), which may be a result of increased ERK1/2 activation in this cell line. Enriched pathways associated with genes expressed higher in WT THP-1s compared to *NFKB1*^{S80A} THP-1 cells include spliceosome and ribosome, which are related with gene transcription and translation processes, respectively (Liberzon et al., 2015). Enrichment of the Parkinson's disease pathway relates to oxidative stress (Liberzon et al., 2015), which may be induced by LPS treatment. Overall, these analyses indicate that *NFKB1*^{S80A} THP-1 cells express genes that regulate inflammatory processes to a greater extent compared to WT THP-1 cells, indicating an enhanced pro-inflammatory phenotype of *NFKB1*^{S80A} THP-1 cells. In contrast, WT THP-1 cells express genes that regulate the cell cycle, suggests that WT THP-1 cells have a proliferative phenotype.

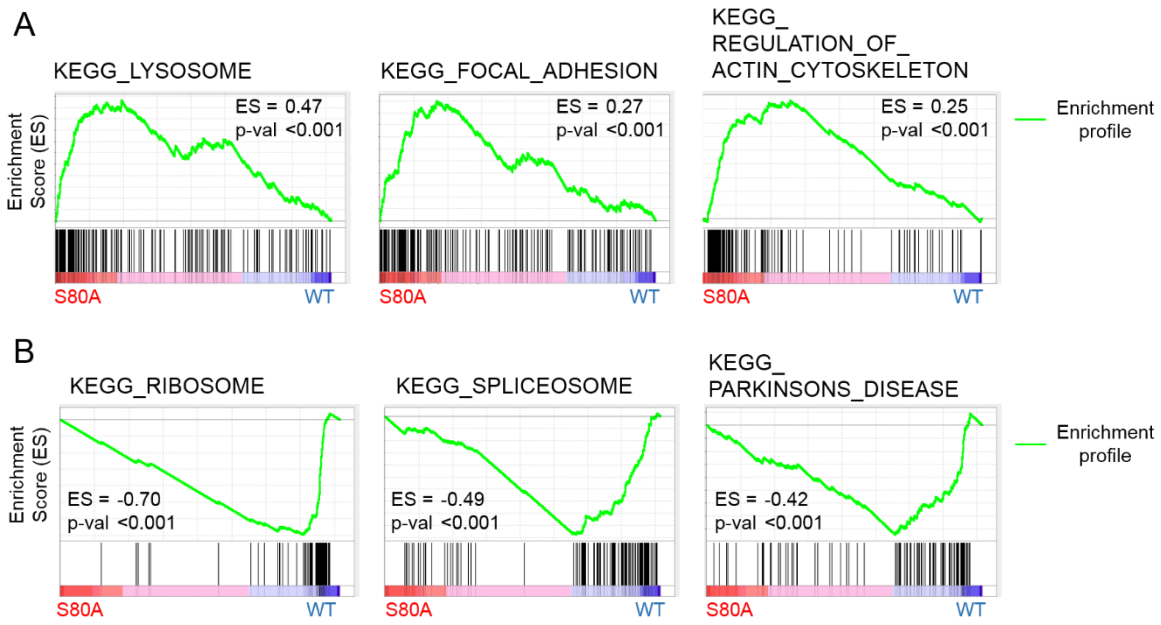


Figure 5.8: Enriched KEGG pathways in LPS-treated WT vs *NFKB1*^{S80A} THP-1 cells
Gene set enrichment analysis of WT + vs S80A + ranked gene list (CuffDiff). Enrichment plots show the enrichment profile (green curve) across the ranked gene list. Black lines represent the genes which contribute to the enrichment profile. Peak enrichment scores (ES) and p-values are shown. (A) Top 3 significantly enriched KEGG pathways at the top of the ranked gene list, indicated by positive enrichment score (expression higher in *NFKB1*^{S80A} THP-1s compared to WT THP-1s). (B) Top 3 significantly enriched KEGG pathways at the bottom of the ranked gene list, indicated by negative enrichment score (expression higher in WT THP-1s compared to *NFKB1*^{S80A} THP-1s).

5.3.5.2 Validation of A(-2)A(-1) κB enrichment within distinct sets of gene sets

In order to investigate a possible link between the enhanced pro-inflammatory phenotype of *NFKB1*^{S80A} THP-1 cells and increased expression of pro-inflammatory genes, the genes enriched in hallmark MSigDB gene sets were further explored. Genes expressed higher in *NFKB1*^{S80A} THP-1s compared to WT THP-1s cells that contributed to TNFα signalling via NF-κB and inflammatory response hallmark enrichment profiles were utilised for transcription factor binding site analysis. Using the GSEA-derived results provides a refined list of genes known to be specifically upregulated as a result of NF-κB signalling, and specifically involved in the activation of the inflammatory response, thus reducing potential noise from genes upregulated independently of NF-κB. Gene promoters were analysed using HOMER, which identified enriched *de novo* κB sites. The consensus κB site $G^{-5}G^{-4}G^{-3}A^{-2}A^{-1}T^{0}T^{1}T^{2}C^{3}C^{4}$ was enriched in the promoters of genes that contribute to the GSEA hallmark TNFα signalling via NF-

κB profile (Figure 5.9A). A similar κB site was enriched in the promoters of genes that contribute to the GSEA hallmark inflammatory response profile, with consensus sequence $G^{-5}G^{-4}G^{-3}A^{-2}A^{-1}W^{0}T^{1}T^{2}C^{3}C^{4}$ (W represents A or T) (Figure 5.9A). These κB site identified are extremely similar to the most enriched motif found in a much larger, less refined list of genes highly expressed in *NFKB1*^{S80A} THP-1 cells ($G^{-5}G^{-4}G^{-3}A^{-2}A^{-1}T^{0}T^{1}T^{2}C^{3}C^{4}$) (Figure 5.5A). These analyses demonstrate that the κB site $G^{-5}G^{-4}G^{-3}A^{-2}A^{-1}T^{0}T^{1}T^{2}C^{3}C^{4}$ is enriched in the promoters of distinct sets of genes. In particular, promoters containing A(-2)A(-1) κB sites regulate a number of pro-inflammatory chemokines and cytokines indicated on the heat maps (Figure 5.9A and Figure 5.9B). These results are also consistent with transcription factor binding site analysis performed in Chapter 4 with HEK293T cells, which also demonstrated that pro-inflammatory genes with A(-2)A(-1) κB sites are expressed higher in *NFKB1*^{S80A} HEK293T cells compared to WT HEK293T cells.

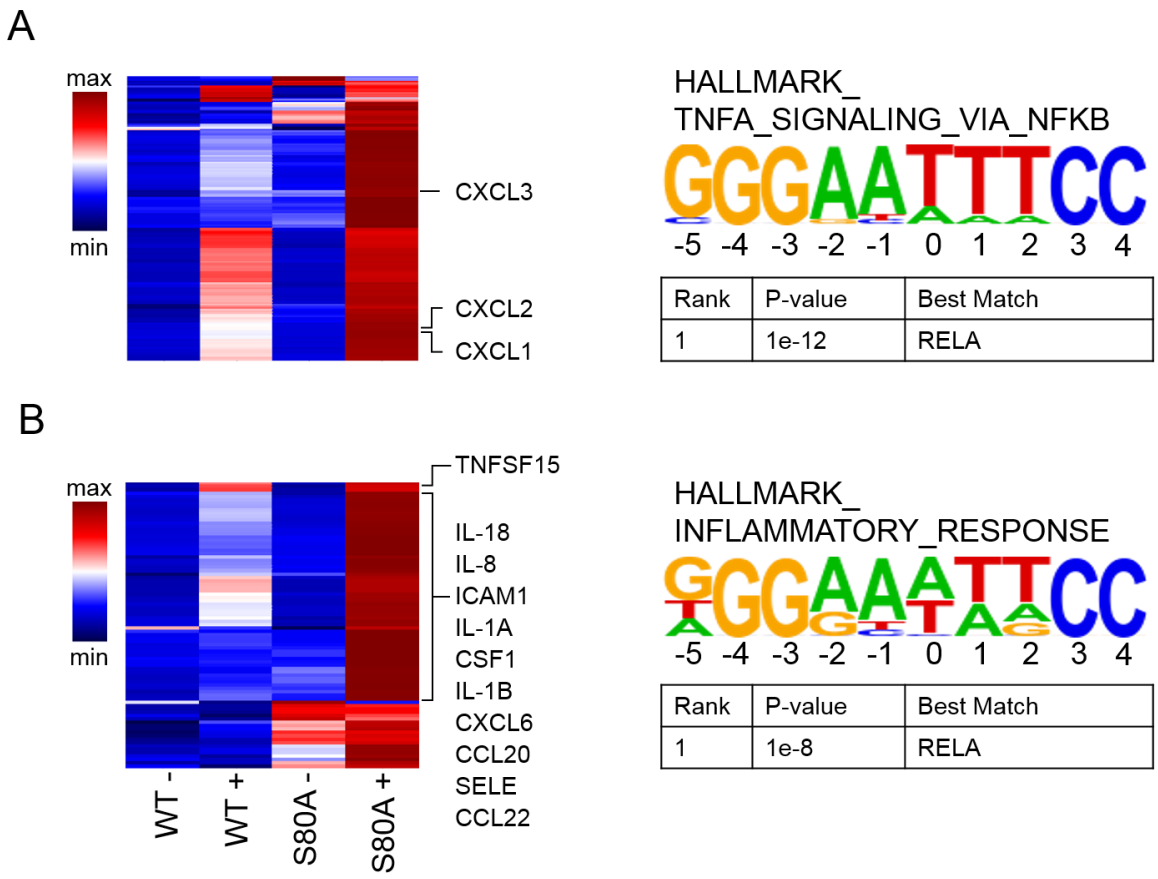


Figure 5.9: -A(-2)A(-1) κB site enrichment in GSEA refined gene sets
The heat maps display the average FPKMs calculated by CuffDiff (scaled as per z-score) of genes derived from (A) hallmark TNFα signalling via NF-κB and (B) hallmark inflammatory response GSEA enrichment profiles. Pro-inflammatory genes with an identifiable κB site are highlighted. HOMER promoter analysis of genes from (A) and (B) identifying significantly enriched *de novo* motifs (p-value <0.001) and best matched transcription factor.

5.3.6 κB site-based prediction of gene expression

The κB site $G^{-5}G^{-4}G^{-3}A^{-2}A^{-1}T^{0}T^{1}T^{2}C^{3}C^{4}$ has been consistently enriched in the promoters of genes with a distinct expression pattern. Genes containing this κB site are expressed much higher both *NFKB1*^{S80A} THP-1 (Figure 5.6A and Figure 5.9) and *NFKB1*^{S80A} HEK293T cells (Chapter 4) compared to WT cells. In particular, S80 mutation had the greatest effect on p50 DNA binding to κB sites containing an A at the -2 and -1 position compared to other the κB sites tested (Chapter 4). As a result, reporter activity from A(-2)A(-1) κB sites was increased the most compared to the other κB sites tested (Chapter 4). The presence of specific κB sites within distinct groups of genes indicates that S80 phosphorylation regulates p50 activity in a κB site-specific manner. Therefore, it might be possible to predict which genes are regulated by p50 phosphorylation based on κB site sequence. Searching for genes containing A(-2)A(-1) κB sites might provide insight into which NF-κB target are regulated by p50 S80 phosphorylation. In order to test this hypothesis, transcription factor binding site analysis was performed using the *de novo* motif finding program HOMER (Heinz et al., 2010). LPS-induced genes from WT THP-1 cells were searched for the following motif in their promoter region: $G^{-5}G^{-4}R^{-3}A^{-2}A^{-1}T^{0}Y^{1}Y^{2}Y^{3}Y^{4}$ (where R represents a purine and Y represents a pyrimidine) (Figure 5.10A). This κB site motif was designed to have a degree of degeneracy to include mismatches at particular bases, for example $G^{-5}G^{-4}G^{-3}A^{-2}A^{-1}T^{0}C^{1}T^{2}T^{3}C^{4}$. A total of 286 genes were identified as having A(-2)A(-1) κB sites, 236 of which were upregulated and 50 of which were downregulated. Since the mechanism by which S80 phosphorylation regulates transcription of genes containing A(-2)A(-1) κB sites was elucidated by analysing upregulated genes (Chapter 4), downregulated genes were not included in further analysis.

The expression of 150/236 genes containing A(-2)A(-1) κB sites was higher in *NFKB1*^{S80A} THP-1s compared to WT THP-1s (Figure 5.10B). Therefore, the transcriptional outcome of 64% of genes regulated by S80 phosphorylation was predicted based on κB site sequence alone. However, of the 236 genes analysed, 88 genes contained multiple κB sites. It is unknown how S80 phosphorylation selectively regulates the expression of genes multiple κB sites. To circumvent this, genes with only one A(-2)A(-1) κB site were analysed. Expression of 97/148 genes was higher in *NFKB1*^{S80A} THP-1s compared to WT THP-1s (66%) (Figure

5.10C). Removing multiple κB site containing genes only increased the prediction of transcriptional outcome by a small margin (2%). Overall, this analysis suggests that the expression of genes regulated by phosphorylated p50 could potentially be predicted by κB site sequence alone, to a certain extent.

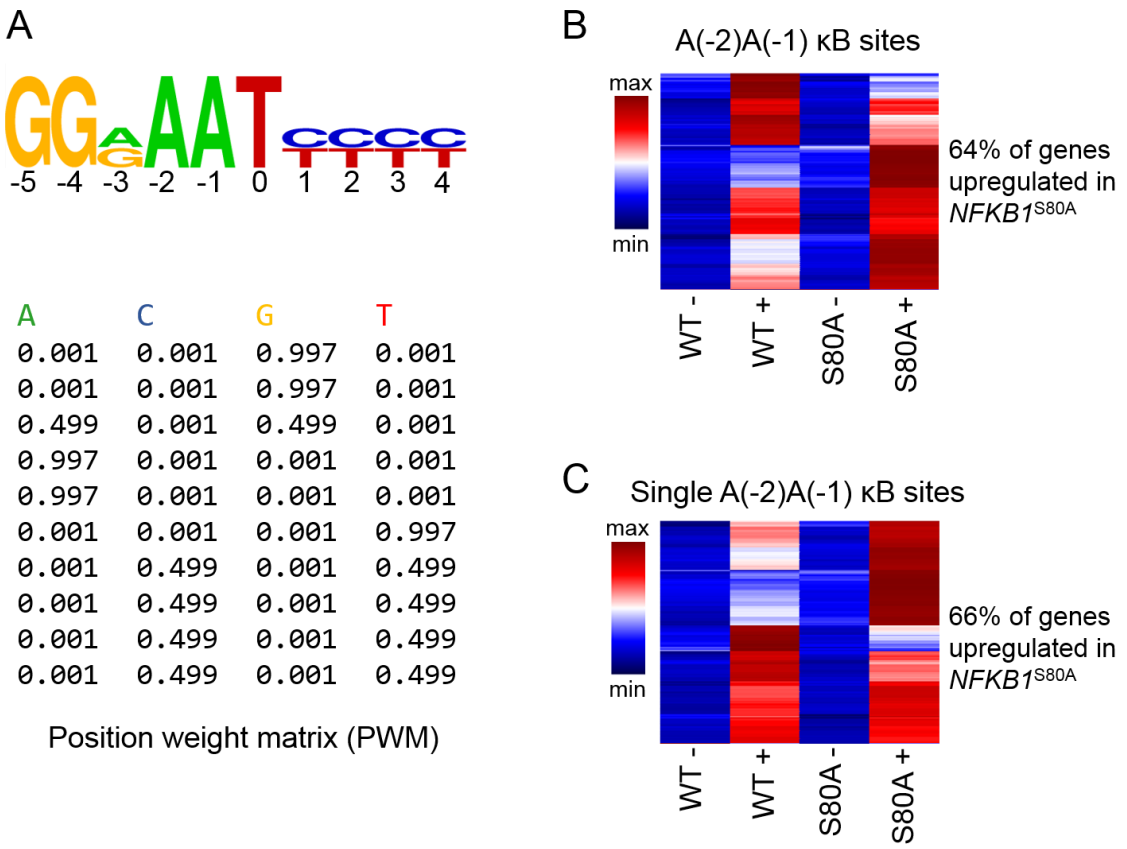


Figure 5.10: Prediction of p50 S80 regulated gene expression based on κB site sequence (A) Position weight matrix (PWM) of the motif used to search LPS-induced genes from WT THP-1 cells for A(-2)A(-1) κB sites using the HOMER motif analysis program. The heat maps display the average FPKMs calculated by CuffDiff (scaled as per z-score) of (B) genes containing A(-2)A(-1) κB sites and (C) genes containing single A(-2)A(-1) κB sites (C). The percentage of genes upregulated in *NFKB1*^{S80A} THP-1 cells are indicated.

5.4 Discussion

In this chapter, the role of p50 S80 phosphorylation in a human macrophage cell line model was investigated utilising CRISPR/Cas9 generated *NFKB1*^{S80A} THP-1 cells. Western blot analysis of *NFKB1*^{S80A} THP-1 cells revealed some surprising effects of the NFKB1 S80 mutation. The endogenous level of p105/p50 protein in *NFKB1*^{S80A} THP-1 cells is slightly reduced in comparison to WT THP-1 cells. This difference was not observed in HEK293T cells, where p105/p50 protein is expressed equivalently in WT and *NFKB1*^{S80A} HEK293T cells (Chapter 4). This suggests that reduction in protein levels could be a cell type-specific effect of the NFKB1 S80 mutation. The transcript levels of *NFKB1* are similar in WT and *NFKB1*^{S80A} THP-1 cells, suggesting a post-translational effect of the S80 mutation. It is possible that mutation of S80 may reduce p105/p50 translation or affect protein stability.

Previous studies utilising macrophages from *Nfkb1*$^{-/-}$ mice have established a link between p105 and ERK1/2 activation (Beinke et al., 2003, Waterfield et al., 2003, Dumitru et al., 2000, Banerjee et al., 2006). It has been reported that ERK1/2 activation is dependent on LPS-induced processing of p105, which liberates TPL-2 kinase, specifically mediating activation of ERK1/2. p105 also functions to maintain the integrity of TPL-2 since it is very unstable. Therefore, *Nfkb1*$^{-/-}$ macrophages are defective in LPS induced ERK activation due to the reduced steady state levels of TPL-2 (Waterfield et al., 2003, Beinke et al., 2003). Since *NFKB1*^{S80A} THP-1 cells have slightly less p105 protein compared to WT THP-1s, it might be expected that ERK1/2 activation is impaired in *NFKB1*^{S80A} THP-1s. However, phosphorylation of ERK1/2 is increased in *NFKB1*^{S80A} THP-1 cells compared to WT THP-1s, indicating that ERK1/2 activation is elevated in *NFKB1*^{S80A} THP-1 cells. This suggests that NFBK1 S80 mutation does not have a direct effect on ERK1/2 signalling through this mechanism.

It is possible that ERK1/2 activation could be upregulated by an intrinsic feedback loop. Perhaps a factor secreted by *NFKB1*^{S80A} THP-1s could upregulate and potentially sustain ERK1/2 signalling. For example, untreated and LPS-treated *NFKB1*^{S80A} THP-1 cells express more cytokines such as *CSF1* and *IL1β* compared to WT THP-1 cells. CSF1 (also known as M-CSF) stimulates macrophage development through activation of ERK1/2 pathways (Richardson et al., 2015).

IL1ß signals through interleukin-1 receptors (IL-1R), related to the TLR family. Both receptor families share the cytoplasmic TIR domain, and consequently utilise overlapping components for downstream signalling. IL1ß triggers signalling through the MyD88 dependent pathway (detailed in Chapter 1), which can activate components of the MAPK pathway, including ERK (Janssens and Beyaert, 2002). Increased pools of these cytokines in *NFKB1*^{S80A} THP-1 cells before and after LPS stimulation may explain the increase ERK1/2 signalling.

It is important to note that the phenotypes of *NFKB1*^{S80A} THP-1 cells appear to be exaggerated by the THP-1 differentiation protocol. In undifferentiated *NFKB1*^{S80A} THP-1 cells, p105/p50 protein level is slightly reduced compared to WT THP-1 cells. However, PMA-differentiated *NFKB1*^{S80A} THP-1 cells appear to have less p105/p50 protein levels compared to undifferentiated *NFKB1*^{S80A} THP-1 cells. To date, the role that NFKB1 plays in differentiation of THP-1 monocytes to macrophages is currently unknown. Furthermore, p50 protein levels were further affected depending on how many days *NFKB1*^{S80A} THP-1 cells were left to rest in culture following PMA differentiation. *NFKB1*^{S80A} THP-1 cells utilised immediately following PMA differentiation appeared to have drastically reduced p50 levels compared to *NFKB1*^{S80A} THP-1 cells rested for 4 days. In addition, ERK1/2 activation kinetics were affected by rest period following PMA differentiation. Inclusion of a 4 day rest period reduced constitutive activation of ERK1/2 observed in *NFKB1*^{S80A} THP-1 cells. It is necessary to allow differentiated THP-1 macrophages to recover in culture without PMA in order to restore a resting phenotype (Chanput et al., 2014). The level of p50 protein in *NFKB1*^{S80A} THP-1 cells appears to recover during the 4 day rest period following PMA treatment, further suggesting a role for NFKB1 in the differentiation of human macrophages, in particular p50. Given that p50 homodimers are important suppressors of gene expression by occupying the promoters of target genes in unstimulated macrophages (Carmody et al., 2007), perhaps p50 homodimers repress genes necessary for differentiation. The reduction in p50 protein during differentiation might be explained by degradation of p50 homodimers in order to remove repressive complexes from target gene promoters, allowing initiation of the transcriptional profile necessary for differentiation.

Despite reduced p105/p50 levels and increased ERK1/2 activation, NF-κB is the core regulator of LPS-induced gene expression in *NFKB1*^{S80A} THP-1 cells. RNA-seq analysis revealed a distinct transcriptional profile in *NFKB1*^{S80A} THP-1 cells, suggesting evidence of selective gene expression regulated by S80 phosphorylation. Mechanisms regulating selective gene expression were previously explored in Chapter 4, and it was shown that S80 phosphorylation regulates p50 activity in a κB site sequence specific manner. Therefore, it was speculated that S80 phosphorylation might regulate p50 activity in THP-1 cells by a similar mechanism. Transcription factor binding site analysis identified $G^{-5}G^{-4}G^{-3}A^{-2}A^{-1}T^{0}T^{1}T^{2}C^{3}C^{4}$ as an enriched κB site in the promoters of genes expressed much higher in *NFKB1*^{S80A} THP-1 cells compared to WT THP-1 cells. Of note, genes containing this κB site include chemokines and cytokines, suggesting that S80 phosphorylation regulates expression of pro-inflammatory genes in THP-1 cells. These data indicate that S80 phosphorylation regulates p50 activity in a κB site-specific manner as part of a p50:p65 heterodimer, consistent with the findings reported in Chapter 4. Consistent results across two cell types stimulated with different inflammatory stimuli indicates that regulation of p50 affinity towards specific κB sites by S80 phosphorylation is a conserved mechanism of selectively regulating gene expression.

Understanding the phosphorylation status of NF-κB and interaction with particular κB sites may enable the prediction of transcriptional outcome of genes regulated by a particular phosphorylation in other cell types. For example, genes containing a A(-2)A(-1) κB site are regulated by S80 phosphorylation which correlates with gene elevated expression in *NFKB1*^{S80A} HEK293T and THP-1 cells compared to WT cells. 66% of LPS-induced genes from WT THP-1 cells containing A(-2)A(-1) κB were expressed higher in *NFKB1*^{S80A} THP-1s compared to WT THP-1s. This suggests that transcriptional outcome of the majority of genes containing a particular κB site sequence can be predicted based on κB site alone. However, it is important to consider that κB site sequence is not the only determining factor for gene expression. Gene expression may be regulated by other mechanisms, different NF-κB dimers or modulated by co-factors. In addition not all κB site containing genes are regulated by NF-κB.

This was recently demonstrated by Tong et al. (2016) who revealed that the expression of a group of genes containing strong NF-κB promoter motifs (associated with RelA ChIP-seq peaks) was induced in *Rela*$^{-/-}$ fetal liver-derived macrophages, implicating other factors in their regulation (Tong et al., 2016).

Gene set enrichment analysis supported the concept that S80 phosphorylation regulates a pro-inflammatory transcriptional profile. Genes expressed higher in *NFKB1*^{S80A} THP-1 cells compared to WT THP-1 cells were more associated with inflammatory processes compared to WT. In contrast, genes expressed higher in WT THP-1 cells compared to *NFKB1*^{S80A} THP-1 cells were more associated with cell proliferation. This suggests that *NFKB1*^{S80A} THP-1 cells are more "active" and have an enhanced pro-inflammatory phenotype, correlating with elevated levels of pro-inflammatory genes. In particular, *NFKB1*^{S80A} THP-1 cells have a phenotype similar to that of an LPS-activated macrophage (Wu et al., 2009), whereby genes associated with processes involving motility and phagocytosis are expressed higher in comparison to WT THP-1 cells. However, to experimentally determine the enhanced pro-inflammatory phenotype of *NFKB1*^{S80A} THP-1 cells, functional assays such as migration assays or chemokine/cytokine secretion assays would have to be performed in order to confirm hypotheses generated from GSEA results.

In addition to κB sites, transcription factor binding site analysis of LPS-induced genes identified enrichment of DNA binding motifs belonging to transcription factors other than NF-κB. Enrichment of DNA binding motifs belonging to other transcription factors indicates that S80 phosphorylation of p50 may be required for interaction with proteins from other gene regulatory networks. In particular, p50 interaction (directly or indirectly) with other transcription factors may be required for downregulation of distinct sets of genes in response to LPS. Of note, ISRE motifs were identified as enriched DNA binding sites in LPS-induced transcriptional profiles of THP-1 cells. Given that p50 homodimers are able to bind to a subclass of IFN response elements (Cheng et al., 2011), it is possible that S80 may regulate p50 binding to ISRE motifs to regulate IFN gene expression in THP-1s. In particular, p50 binding to ISRE motifs may be regulated in a sequence specific manner, potentially revealing a previously unidentified consequence of p50 phosphorylation.

Chapter 6

6 General discussion

Phosphorylation of NF-κB is an important regulatory mechanism for the control of transcription. Despite the importance of NF-κB p50 in the regulation of inflammation, phosphorylation of p50 is not well studied. In this thesis, CRISPR/Cas9 genome editing techniques were utilised to generate *NFKB1*^{S80A} HEK293T and THP-1 cell lines in order to investigate the role of a novel p50 phosphorylation of S80, in the regulation of inflammatory gene expression. Transcriptomic analyses revealed that phosphorylation of p50 at S80 selectively regulates gene expression in response to inflammatory stimuli such as TNFα and LPS. The data shows that selective gene expression is regulated in a DNA sequence specific manner, due to the enrichment of specific κB sites within distinct subsets of NF-κB target genes including pro-inflammatory chemokines and cytokines. DNA binding analyses demonstrated that S80 regulates the binding of p50 to NF-κB binding sites in a sequence specific manner. Specifically, phosphorylation of S80 reduces the affinity of p50 for κB sites that have an adenine at the -1 position, and thereby limits the expression of genes under the control of regulatory elements containing -1A κB sites. The regulation of p50 DNA binding by S80 phosphorylation occurs both in the context of p50:65 heterodimers and p50 homodimers, revealing the regulation of the transcriptional activity of the p65 subunit in *trans* through the modification of p50.

The kinase responsible for the phosphorylation of S80 was discovered to be IKKβ, identifying NF-κB p50 as a novel substrate for this kinase. IKKβ is most well-known for its role as an activator of the canonical NF-κB pathway by phosphorylation of IκBα. However, IKKβ is capable of phosphorylating other components of the NF-κB pathway such as p105. To date, the majority of p105/p50 phosphorylation sites identified are located in the C-terminal region of p105, and are required for p105 processing to generate the p50 subunit. However, phosphorylation of the N-terminal region of p105 (i.e. p50) is much less well understood (described in 1.4.3.2). In addition, it has been previously reported that IKKβ phosphorylates p65 at S468 (Schwabe and Sakurai, 2005) and S536 (Schwabe and Sakurai, 2005, Sakurai et al., 1999, Sizemore et al., 2002, Haller et al., 2002, Yoboua et al., 2010), proposing a role for IKKβ-mediated

regulation of NF-κB subunit activity. However, the mechanisms regulating transcriptional activity were not elucidated. This thesis reveals the regulation of sequence specific p50 DNA binding by IKKβ phosphorylation as a novel mechanism of IKKβ-mediated control of NF-κB transcriptional activity. It is important to note that phosphorylation of particular residues of p65 are regulated by a number of kinases (Christian et al., 2016). In particular, S536 has been reported to be phosphorylated by 4 other kinases in addition to IKKβ (Buss et al., 2004). It is therefore possible that p50 S80 could be phosphorylated by other kinases, regulating transcriptional responses in the context of different cellular stimuli. It may also be possible that S80 phosphorylation occurs in the resting state. p65 S468 phosphorylation by GSK3β in unstimulated cells acts to negatively regulate p65 basal activity. This suggests that S80 could be phosphorylated by other kinases under resting conditions, regulating p50 activity at basal level. Identification of additional p50 S80 kinases will provide insight into the role of S80 phosphorylation in the regulation of NF-κB driven transcriptional control, depending on physiological context. In addition, the subcellular location of IKKβ kinase activity on p50 is not known. IKKβ phosphorylation of p65 at S468 and S536 appears to take place in the cytosol while it is bound to IκB (Schwabe and Sakurai, 2005). However, it has been reported that IKKβ is present in the nucleus (Tsuchiya et al., 2010, Anest et al., 2003). This suggests that IKKβ may phosphorylate p50 in the nucleus, although the status of its kinase activity in the nucleus is yet to be confirmed.

The ability of NF-κB subunits to be phosphorylated at multiple sites by different kinases, and at the same site by multiple kinases adds a layer of complexity to NF-κB transcriptional regulation. Combinatorial phosphorylation events could result in heterogeneous populations of phosphorylated NF-κB protein. Indeed, it has been demonstrated that phosphorylated p65 protein accumulates in distinct subcellular locations. S536 phosphorylated p65 is predominantly cytosolic, whereas S468 is localised to nuclear speckles (Moreno et al., 2010). This led to the proposal of the NF-κB barcode hypothesis which suggests that post-translational modifications of NF-κB subunits, alone or in combination, generate distinct functional states that direct transcription in a gene specific manner (Moreno et al., 2010). Rather than “one modification, one effect”, phosphorylation by multiple kinases may regulate NF-κB in different manners.

The barcode hypothesis provides an explanation as to how individual stimuli activating NF-κB leads to the specific induction of distinct sets of genes. However molecular mechanisms regulating NF-κB interaction with specific gene promoters remain unclear. This thesis established a mechanism for the selective regulation of gene expression by the p50 subunit. It was revealed that that gene-specific transcriptional effects of S80 phosphorylation are mediated by the differential binding of phosphorylated p50 with κB sites containing an A nucleotide at the -1 position (Chapter 4). In addition, p50 S80 phosphorylation primarily affects the binding of p50:p65 heterodimers at these sites to inhibit gene transcription from promoter constructs containing -1A κB sites. The importance of the -1 κB site position has been previously reported in the context of DNA damage. Phosphorylation of p50 S328 has been shown to attenuate the binding of p50 towards κB sites containing a -1C. This study together with this thesis identifies the -1 position of NF-κB binding sites as a critical regulator of p50-DNA binding. Therefore, κB site sequence can also be considered as an important factor contributing to the NF-κB barcode hypothesis.

The work of this thesis focussed on the impact that p50 S80 phosphorylation has on the regulating p50:p65 heterodimer driven transcription. However, it is important to consider the dimer specific effects of the S80 mutation. While this thesis does not fully explore S80 phosphorylation as a regulator of p50 homodimer function, there is data to suggest specific p50 homodimer effects. RNA-seq analysis of THP-1 cells revealed that expression of a number of genes were altered in the resting state in *NFKB1*^{S80A} cells (Chapter 5). This suggests a possible role for S80 phosphorylation in regulating the expression of genes at basal level, perhaps controlled by p50 homodimers. Previous studies have shown that there is a large number of p50 homodimers present in the nuclei of unstimulated macrophages (Cheng et al., 2011), and that the p50 subunit was exclusively present at NF-κB target gene promoters in unstimulated macrophages, indicating the presence of p50 homodimer binding (Carmody et al., 2007). Upstream regulator analysis of unstimulated THP-1 cells revealed NFKB1 as a significantly enriched transcription factor (Figure 7.4). Together, this indicates that p50 may also play a critical role in controlling the basal expression of NF-κB target genes, regulated by S80 phosphorylation. However, the mechanisms controlling the specificity of p50 homodimer affinity to specific

promoters is currently unknown. DNA binding analyses in this thesis revealed increased binding $p50^{S80A}$ to a specific κB site ($G^{-5}G^{-4}G^{-3}G^{-2}A^{-1}T^{0}T^{1}T^{2}C^{3}C^{4}$) relative to WT p50, however p65 binding to this site was similar in both $p50^{S80A}$ and WT p50 expressing HEK293T cells. This indicates that S80 phosphorylation may regulate the binding of p50 homodimers but not p50:p65 heterodimers to κB sites with this sequence. However, further analysis into the mechanisms of dimer specific phosphorylation is required.

Given that ubiquitin regulated stability of p50 homodimers has previously been shown to be critical for acting to threshold pro-inflammatory gene expression (Carmody et al., 2007, Cheng et al., 2011), it is important to consider ubiquitination and proteasomal degradation as a mechanism of p50 homodimer regulation, which may involve S80 phosphorylation. S80 phosphorylation may be involved in recruiting ubiquitination machinery to specific target gene promoters bound by p50 homodimers, triggering proteasomal degradation and thus clearing promoters. Phosphorylation-dependent ubiquitination of the p65 subunit, resulting in proteasomal degradation and promoter clearance, has been demonstrated (Lawrence et al., 2005, Mao et al., 2009). In particular, p65 phosphorylated at S468 by IKKβ was preferentially ubiquitinated by an E3 ligase complex containing COMMD1 (Mao et al., 2009).

Transcription factor binding site analysis revealed enrichment of both ISRE and NF-κB DNA binding motifs in the transcriptome of LPS-stimulated THP-1 cells (Chapter 5). In particular, genes containing specific NF-κB motifs are inhibited by S80 phosphorylation, whereas genes containing specific ISRE motifs are activated by S80 phosphorylation (Chapter 5). This further suggests how S80 phosphorylation might regulate p50 as part of a p50:p65 heterodimer and p50 homodimer respectively, regulating different functional outcomes. This analysis supports a previous study reporting that p50 homodimers bind to a subclass of IFN response elements to repress IFN-inducible gene expression in macrophages (Cheng et al., 2011). It is possible that S80 may regulate p50 homodimer binding to ISRE motifs to regulate IFN gene expression in THP-1s. In particular, p50 homodimer binding may be regulated in a sequence specific manner (Chapter 5). These data suggest that S80 phosphorylation modulates p50:p65 heterodimer activity to regulate pro-inflammatory gene expression, whereas S80

phosphorylation of p50 homodimers may regulate anti-viral gene expression. This indicates a potential new role of p50 phosphorylation in the regulation of anti-viral responses. In order to further investigate this, *NFKB1*^{S80A} THP-1 cells could be stimulated with type 1 IFNs such as IFN-α or IFN-ß to assess induction of anti-viral transcriptional programmes.

DNA binding sites can act as allosteric regulators of transcription factors (Lefstin and Yamamoto, 1998). Distinct DNA conformations adopted by particular κB sites provide a potential mechanism to explain why single nucleotide variations affect DNA binding of NF-κB dimers. Transcription factor binding site shape analysis of the κB sites investigated in this thesis were analysed for differences in DNA shape configurations that are important in protein-DNA recognition (Yang et al., 2014). This *in silico* analysis predicts a unique conformation for specific κB sites (Chapter 4). In particular, the $G^{-5}G^{-4}G^{-3}A^{-2}A^{-1}T^{0}T^{1}T^{2}C^{3}C^{4}$ κB site has a narrower minor groove width, and higher degree propeller twist at the -1 position, compared to other κB sequences analysed (Chapter 5). A narrow minor groove width increases the negative electrostatic potential of DNA, offering favourable conditions for interaction with positively charged proteins (Rohs et al., 2009). Higher degrees of propeller twist provide opportunity to form an additional hydrogen bond in the major groove, important for the recognition of specific DNA sequences by proteins (Rohs et al., 2009). Phosphorylation of S80 may lead to the addition of a negative charge to p50, reducing affinity to DNA with narrow grooves such as $G^{-5}G^{-4}G^{-3}A^{-2}A^{-1}T^{0}T^{1}T^{2}C^{3}C^{4}$ κB sites, predicted to have a high negative charge (Chapter 4).

Phosphorylation of S80 may also induce conformational change of p50, which could affect transcriptional outcome following DNA binding. Previous studies have revealed that phosphorylation of p65 induces conformational changes resulting in changes to ubiquitination status and differential protein interactions (Milanovic et al., 2014). Using similar approaches, conformational differences between p50^{WT} and p50^{S80A} were assessed by limited proteolytic digestion, which revealed different sensitivities to digestion between p50^{WT} and p50^{S80A} (Chapter 4). This indicates that S80 phosphorylation may induce a conformational change that could modify the binding of p50 dimers to -1A containing κB sites. Conformational changes may modify the transactivating potential of NF-κB

dimers by regulating DNA binding to specific κB sites, but also possibly by modifying interaction with other factors that in turn affect DNA binding. In addition, it has been reported that single nucleotide differences in κB site sequences are critical for the recruitment of specific co-activators required for NF-κB target gene expression (Leung et al., 2004). This study suggests that κB sites impart specific properties to the bound NF-κB dimer, thereby determining co-activator recruitment, highlighting the importance of κB site sequence for specific gene expression. This fits with the hypothesis that S80 phosphorylation could affect the interaction of p50 with other factors, ultimately altering DNA binding affinity to specific κB sites.

S80 phosphorylation of p50 is sufficient to inhibit DNA binding which limits gene transcription from κB sites containing an A at the -1 position. Of note, individual κB sequences have different capacities to drive transcription as measured by reporter assays (Chapter 4). These analyses identified A(-2)A(-1) κB sites as more potent drivers of transcription than other sites tested. Interestingly, the increased transcription from reporter vectors containing these κB sites is also reflected in both TNFα and LPS stimulated HEK293T and THP-1 cells respectively, where genes regulated by these sites are induced at significantly higher levels than genes regulated by other κB sites. These κB sites appear to be enriched in the regulatory regions of genes encoding pro-inflammatory cytokines such as TNFα and IL8 and the chemokines CXCL1 and CXCL2. Many NF-κB target genes contain multiple κB sites in their regulatory regions, which may contribute towards specific gene activation. Previous research on the cooperative interaction between two κB-sites has suggested that that one site is dominant over the other (Leung et al., 2004) which could relate to affinity of NF-κB towards certain sites. It will be interesting in the future to explore the cooperation of A(-2)A(-1) κb sites with other κB sites within the same reporter construct, exploring any dominant effects that A(-2)A(-1) κB sites might have.

The enrichment of specific κB sites in a functional class of genes provides strong evidence that NF-κB phosphorylation and κB sequences may establish regulatory networks to coordinate specific transcriptional programmes. Indeed, analyses of LPS-inducible genes in THP-1 cells revealed that A(-2)A(-1) κB sites are enriched within genes important for the induction of the inflammatory response, and are

associated with particular biological pathways important for active macrophage function (Chapter 5). These analyses suggest that the distinct transcriptional profiles driven by p50 S80 phosphorylation could potentially co-ordinate specific processes in macrophages. Further analysis of the physiological roles of S80 phosphorylation will be key to understanding how p50 regulates the immune response by acting as a critical regulator of gene expression in macrophages.

Identifying κB sites that correlate with specific transcriptional responses allows the potential for prediction of transcriptional outcome based on the phosphorylation status of p50. For example, by using a simple PWM to search for genes containing A(-2)A(-1) κB sites, the transcriptional outcome of 66% of analysed genes regulated by S80 phosphorylation was predicted based on κB site alone (Chapter 5). Future research may enable the outcome of gene transcription to be predicted based on knowledge of the phosphorylation status of NF-κB and DNA binding site sequence using a systems approach. For example, by integrating RNA-seq data with large-scale protein-DNA binding data generated from PBMs, a database identifying κB site binding preferences of specific phosphorylated forms of p50 correlated with distinct transcriptional outcome could be created. Combining high-throughput analyses could generate PWMs with strong predictive power. This may allow the functional consequences of site-specific p50 phosphorylation to be predicted across a range of different cell types, such as B-cells and T-cells, further investigating the contribution of p50 to immune cell function. Prediction of promoter binding with associated transcriptional outcome could also be extended to *in vivo* models. The functional and biological effects of p50 phosphorylation could be predicted and assessed in mouse models, elucidating the significance of p50 phosphorylation during development. Ultimately, this approach could be utilised to predict the consequences of modulating p50 phosphorylation in human inflammatory disease.

Chapter 7

7 Appendix

7.1 CRISPR/Cas9 targeting of *NFKB1*

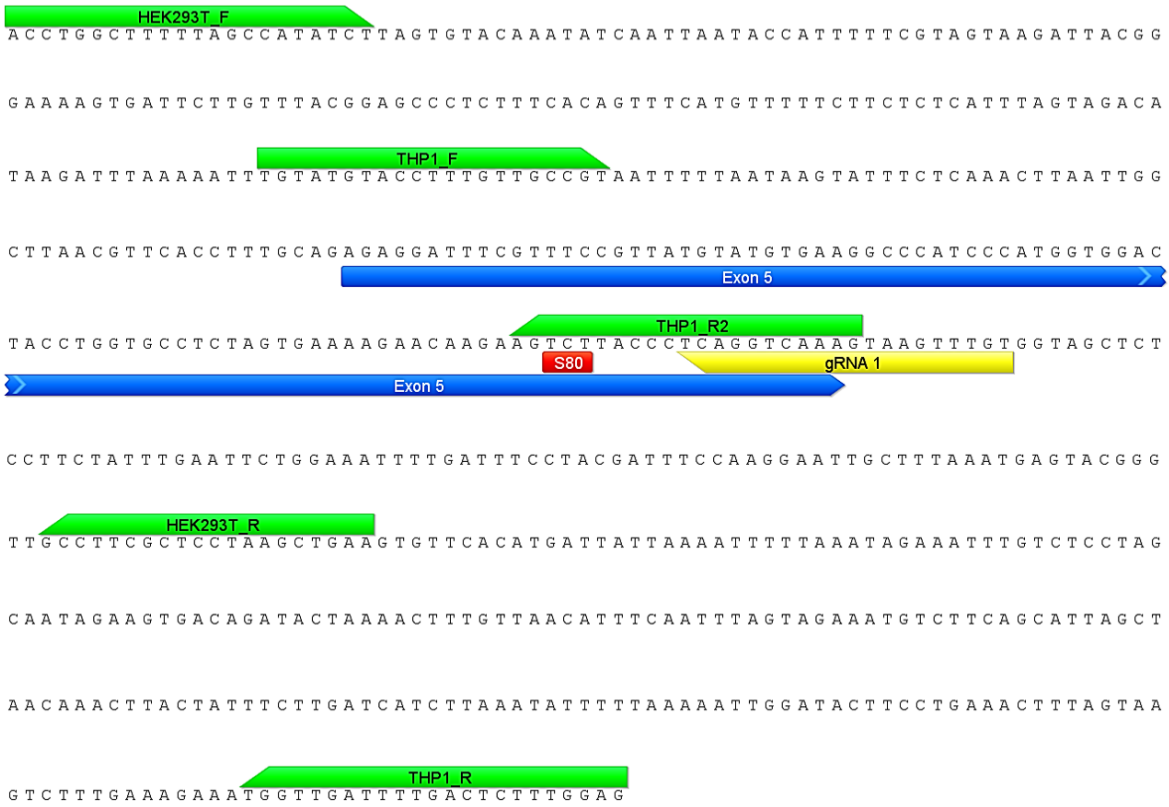


Figure 7.1: CRISPR/Cas9 targeting of *NFKB1*
Schematic representation of the human *NFKB1* gene highlighting the position of the gRNA utilised for CRISPR/Cas9 genome editing (yellow allow) relative to exon 5 (blue arrow) containing the S80 codon (red box). The sequence and binding sites of PCR primers designed for screening are shown (green arrows).

7.2 Clustering analysis of differentially expressed THP-1 genes

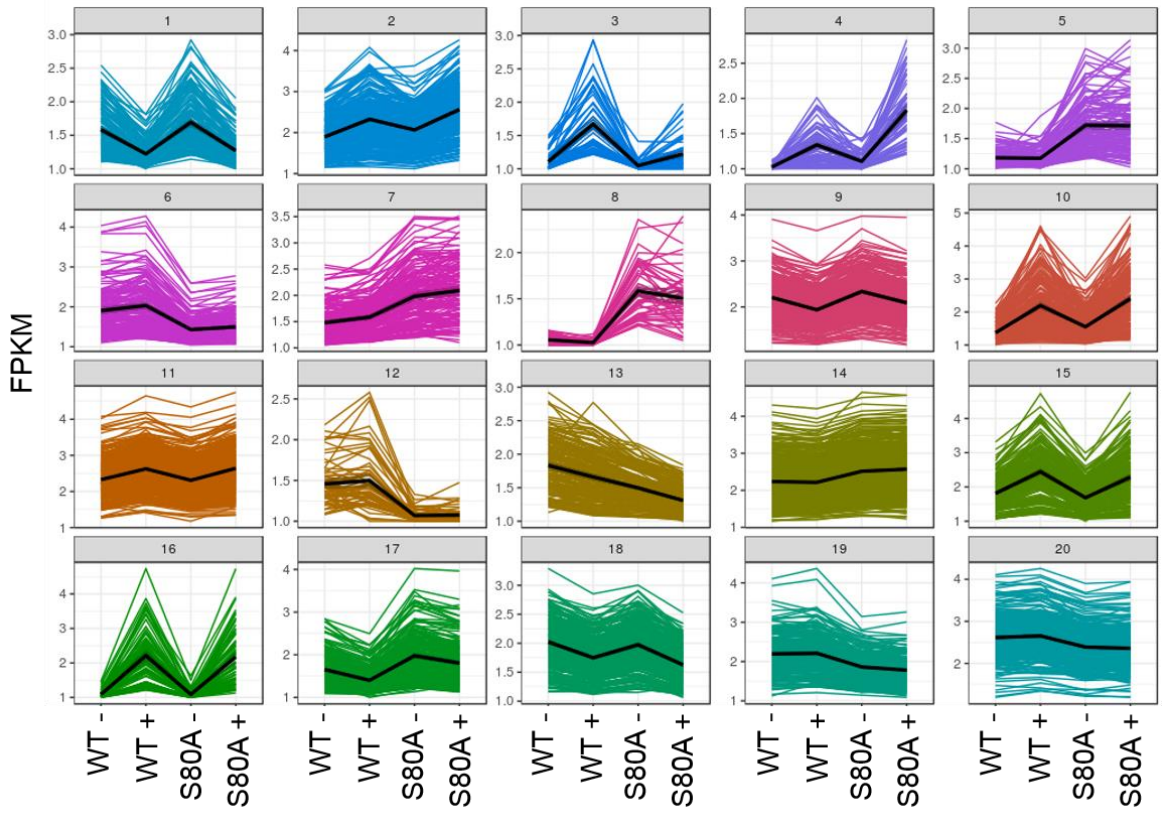


Figure 7.2: Clustering analysis of differentially expressed genes in THP-1 cells
Clustering analysis of differentially expressed genes (CuffDiff calculated: p.adj <0.05, fold induction ratio >2) was performed using the Bioconductor package CummeRbund with default options (K-means clustering with Jensen-Shannon distance), producing 20 clusters. Line plots display the average FPKMs of WT and *NFKB1*^{S80A} genes + or – LPS as indicated.

7.3 THP-1 clustering analysis

Upregulated: S80A			Upregulated: WT		Upregulated: Similar	
GADD45B	AMPD3	ATF3	VSNL1	MX1	MSX1	RP11-
NFKBIA	CPEB4	STARD4	ORM1	NOD2	FOSL1	806H10.4
PEAK1	KIF21A	FNDC3B	INHBE	MT1M	CNKSR3	MGC12916
IL1RAP	TMEM64	NFIL3	PTAFR	NBPF14	INO80D	SOCS3
CYLD	TNFRSF10D	EN2	ANKRD18DP	VILL	NABP1	LINC00936
SEPSECS-AS1	CYP7B1	SNN	LINC01136	LILRB2	MAML2	GPRC5B
SLC43A2	VAX2	TNFAIP1	KLF2	BATF3	FFAR2	CCL4
AHR	BCAR1	TBC1D10A	TMEM171	TLX3	CPXM1	SLC7A2
LOC102723779	CCSER2	FAM222A	EGR4	CFB	PPP1R15A	SSTR2
BTC	LOC101928487	PLXNC1	FAM167A	GRIN3A	MAP1B	CNP
NEURL3	ZBTB38	TECPR2	ALDH1A1	STAT4	TRIP10	C6orf62
LINC00346	GIMAP8	WT1	LOC286186	PTGIR	NRP2	SMCO4
HCAR3	RNF2	FEM1C	LOC102723959	NBPF10	ZHX2	LAP3
IDO2	KLHL5	PVR	HIST1H3I	NRG4	FLJ20021	SQRDL
HCAR2	CDC37L1	PIK3R5	NUDT9P1	ANKRD18A	MAP3K5	NUDCD1
IDO1	PRKCH	ABHD17C	TEX19	LOC344887	TPT1-AS1	ZBTB17
RP11-	PPP3CC	FOSL2	LOC100996717	HERC5	ZNFX1	SERPINB1
701P16.5	PDLIM5	HIF1A	MYBPC3	LINC00996	CD44	SFT2D2
CXCL2	KIF1B	ITGA6	CLUHP3	PSTPIP2	PANX1	UBE2L6
LAMP3	EMR2	PNRC1	SIM1	RP11-	STARD10	CPNE8
LOC102724012	PDE8A	LOC102723863	LOC102724608	330A16.1	PHLDA2	FBRS
LOC101928783	BCL2L11	RP1-249H1.4	NAV3	RSAD2	WWC2	GNB4
LOC101927369	KLF9	MALAT1	PLLP	CMPK2	XRN1	CHD2
TNIP3	KLF6	RFX2	CXCL14	HERC6	CEP85L	PRKD3
CXCL3	ARHGAP10	SPATA5L1	NOTCH2NL	FLVCR2	HIVEP3	NDEL1
BX255923.3	CDKN1A	KLF7	APOL3	TNFSF10	GPR84	ADPRHL2
BIRC3	LOC101927759	LCP2	CNTNAP1	LAMB1	FAM117A	SPATA2
AC058791.1	SLC4A7	KLF10	SYNGR3	HESX1	LOC100996740	UBE2Z
IL1A	USP9X	POGLUT1	TNFSF13B	LINC00528	RAPGEF2	TOP1P1
GCNT4	HDAC9	ZFC3H1	BMPER	MT2A	BLZF1	IFRD1
PLAT	BCL6	MARCKS	SFRP2	DAPP1	ADAM11	ATP2C1
CXCL6	ADAM17	NEAT1	ABCA6	SLC24A4	EXT1	ARID5A
IFNL1	IER3	ZNF674-AS1	PPP1R17	LHX2	NAB1	GMEB1
MB21D2	APLF	DAPK1	C6orf58	CKB	SMAD7	UBR4
TRAF1	C11orf96	RAB8B	FDCSP	BRCA2	C15orf37	BRWD3
CXCL1	BAMBI	TWISTNB	CLEC4D	RGL1	TICAM1	OAS1
KCNA3	CRIM1	IGSF8	MIR3945	MCOLN2	IL6ST	PPP1R15B
BDKRB2	PCDH9	USP43	ACSM1	MMP25	SLCO3A1	C1orf122
ELOVL7	S100A3	PDHX	OXTR	XAF1	TP53INP2	SBNO2
IL32	IL1R1	JDP2	BCL2L14	FAM201A	JUN	SEC24A
ITGB3	LINC01004	ABCA1	LOC102723429	PLAGL2	PRDM1	ADRB1
OCSTAMP	KLF5	MESDC1	CCM2L	IRF7	N4BP1	GTPBP1
LINC00622	MAMLD1	ZNF267	LAD1	MB21D1	JUNB	PHF11
MMP10	IL1B	TNFRSF10A	CCR7	SERPING1	BTG2	MLLT6
FGF18	NHSL1	LYSMD2	GPR31	LOC101927910	RRAS2	EPM2AIP1
LOC729040	DUSP5	BACH1	IL27	TRIM5	APBA3	BRPF3
PDGFB	LOC101928716	DUSP22	TEX41	FRY	WTAP	RAB20
ITGB8	LOC102724792	PPARD	AQP9	NPTX2	GLYCTK	METTL1
EFNB2	CSF1	TPRA1	ANKRD66	MX2	ANKRD28	EAF1
UNC13A	FAM49A	SYT4	HAPLN3	DMXL2	SKIL	PANDAR
COL5A3	SERPINE2	KCNN2	LOC102724443	KCND1	STK17A	USP13
CTHRC1	ZNF697	PNPLA8	PNPLA1	SAMD9L	HUWE1	C19orf66
CTGF	OR2B11	RASSF5	GPR37	ADA	PTGER4	STRIP2
NKX3-1	RPE65	SRFBP1	GAL3ST2	IL1RN	LOC401463	LYN
CYS1	ACKR3	MGLL	CGN	SMAD3	TP53BP2	ANP32D
EDN1	JHDM1D-AS1	PFKFB3	GBP5	C20orf112	TLR8	PCGF5
SEMA6D	MAFF	CDK6	SLAMF7	TTC39B	UXS1	TFEC
CAMK2A	RASGEF1B	TMEM140	AMOTL2	TRIM22	STC2	BATF
CACHD1	GLIS3	CREM	THBS1	DENND4A	BMF	PTRH1
ICAM5	ERRFI1	CLIC4	RHOH	MASTL	GCLM	ACSL5
TFPI2	LTB	B3GNT5	NFE2L3	CEP120	G0S2	MS4A7
FLT1	RP11-37B2.1	RUNX1	KCNJ2-AS1	PDGFRL	QPCT	TRADD
SPSB4	PGM2L1	JAK1	IFIT2	KCNJ2	ZFP36	NFKBIB
STEAP4	BCL2A1	RHOF	CHST2	FAM26F	PKIG	ATP11C
FGF2	HS3ST3A1	TMEM106A	FAM134B	HCK	NCF1	ETS2
MMP12	SMPD5	AEBP2	NBPF19	EPSTI1	RIN2	PTPN1
VEGFC	THAP2	FZD9	EGR1	SPATA13	APOL2	CNTLN
TMCC2	LOC100507403	NLRP3	MAP3K8	CASP10	RAB21	NCOA2
LEPREL1	LOC100132831	TCEB3	CXCL10	CD70	AC147651.4	CCNL1
GPRC5A	RP4-553F4.6	RCN1	NFKB2	PIK3AP1	MXD1	PVRL2
EVA1A	LOXL2	TLR2	GCH1	SIK3	NINJ1	CASP7
C2CD4B	DUSP1	ZC3H7B	OASL	MFSD2A	PML	NUB1

ADAMTS5	SVIL	RBM17	ZCCHC5	C5orf56	USP42	SLC11A2
PCDH17	LOC100288175	MED14	NTNG2	CYB5R2	TRAFD1	ZNF140
NTN1	FSTL3	YRDC	NCF1B	GMPR	AIDA	CEBPB
IL36G	RASSF8	STK40	CCL7	C10orf12	ERO1LB	SGTB
WNT1	ACVR2A	SLC30A7	SOCS1	JADE2	BTG3	ZBTB5
CYP7A1	LUCAT1	NFATC1	RRAD	LYRM4	DOT1L	EDEM1
TUBB2B	RGS16	LOC100507217	C1orf21	FBXO6	CRISPLD1	APOL1
CLDN1	TUBB2A	SQSTM1	WNT3	ISG15	CSNK1G3	TAP1
C2CD4A	SLC41A2	POLR3D	LFNG	VCAM1	LACTB	SLC25A28
SELE	RASL10A	BATF2	DST	CHI3L2	IKBKE	IFIT1
EFNA1	SFRP1	CCRL2	EIF3C	ASTL	RIPK1	MIRLET7BHG
PI3	ARRDC3	ZC3H12A	TOR1B	LINC00158	ZBTB42	CCL5
MEP1A	PLK3	CSRNP1	ZMIZ2	GBP1P1	METTL21B	TAF4B
INSM1	NFAT5	CD40	TMEM194A	MEOX2	PLEK	SLC7A11
VSTM1	CLIP2	FCAMR	DDIT4	CSF3	FAM177A1	HELZ2
CSF2	TCHH	MYO10	WARS	IL6	SRXN1	IFIH1
LRRC32	LOC101928054	CA13	JAK2	SHROOM3	TANK	APOL6
FRMD7	RIPK2	SNAI1	PTPN2	IL12B	CDYL2	C17orf96
C9orf135-AS1	TCF7L1	CCL22	NBN	AC008697.1	OSGIN1	IFIT3
C1QTNF1	BCAR3	FLJ46906	DDX60L	CLEC4E	BBC3	SLC2A6
INHBA	KIRREL2	CD83	IL12RB2	C1orf61	PRRG1	DDX60
LOC100288911	PPP4R1L	EHD1	MAP1LC3A	C11orf86	IL10RA	KMT2E-AS1
IL18R1	ABL2	LINC00884	RNF213	IL18RAP	ROBO1	TRAF3IP2
NUAK2	ZXDA	IL8	NAF1	TNFAIP6	TMEM39A	NFKB1
XIRP1	KITLG	ABTB2	SP110	CCL8	FAM129A	TBX3
PGLYRP2	USP12	LOC101928841	STAT1	CXCL9	RNF145	SOD2
IL2RA	IGFBP3	NOC3L	ADPRM	IL23A	NFKBIE	KYNU
N4BP3	ATP2B1	MPND	THUMPD3-AS1	LILRA6	GPR137B	DRAM1
CD69	GEM	TBX15	EYA3	MUC1	SLC31A2	TIFA
ANKRD26P3	OTUD1	IL18	STAT2	NCF1C	RBM7	IER5
RND1	STARD5	LGMN	TRIM21	GRAMD1A	ST8SIA4	F8A1
SRSF12	ICAM1	PRDM2	MYD88	FEZ1	CDC42EP3	LINC00477
CCRN4L	AREG	KPNA1	SIAH2	POU2F2	LIMK2	DENND5A
PPP2R3A	TTC9	CASP3	XBP1	SIX5	ANKLE2	DHX58
LINC00854	SEMA3C	PPP1R18	PARP9	NCR3LG1	LRRC3	CCDC71L
PIM1	CCL20	TAB2	OAS2	FAM95C	MED13	TRIM69
RDH10	BCL9L	RAP1B	IRF9	MTF1	IRF2	C7orf60
DSE	FMNL3	ARFGAP3	GTF2B	LOC730101	SSSCA1-AS1	FOXD4L1
PXDC1	PMAIP1	KCTD20	OAS3	PIM2	CBLN3	GPR132
BHLHE40	LOC101929796	PTP4A1	PARP12	TRIM36	ADTRP	ZC3HAV1
ELL2	IL3RA	NCOA4	PHACTR4	MCTP1	LOC102723996	TNIP1
DNAJB5	ETV3	UBE2E1	BCL11A	USP18	SLC16A9	SAMD9
CBR3	PDCD1LG2	SAT1	CD38	HIVEP1	DIXDC1	RELB
CD58	MEIS2	EIF5		IRG1	DDX58	IRF1
ZC3H12D	ACRC	MSTRG.12754			CCL2	NRIP3
LOC100132891	MAP2K3	RELA			PLEKHA4	SLC12A7
RICTOR	GBP3	STXBP5			STX11	GSAP
CLEC2D	RBM43	VIM			NR4A2	LONRF1
PLA2G4C	SNAPC1	DYNLT1			KIAA1644	VEGFA
LOC101928762	PLEKHF2	ST3GAL2			PNMA2	PIM3
PLAUR	LOC101927686	SATB2			C12orf61	DTX4
SLFN5	B4GALT5	DDX3X			DNAJB4	DCP1A
ITPRIPL2	LOC100505622	ARL8B			SLC35G1	IKZF1
AGO2	OPTN	FOXD4			TNFSF9	TWIST1
PLAU	ACSL1	WBP4			LOC101927402	IFI44
SPHK1	C4orf32	PITPNA			F3	HSPA6
KIAA0247	E2F7	DENND2D			REL	RP3-399L15.3
PDLIM4	RGPD2	ZNF189			NKX2-2	
RIN3	IRAK2	DDX21			CYP27B1	SLC4A5
IER2	IL7R	TRIM24			LINC01181	EMP2
OTUD4	LOC440934	CEBPD			ISG20	ZNF101
BIRC2	ICOSLG	AZIN1			MSC	TRAF3
B4GALT1	ICAM4	MARK3			SRC	DTX3L
SMURF2	DCUN1D3	TNIP2			CCL3L3	EIF2AK2
MICALL1	TNFSF15	PITPNB			KDM6B	PLSCR1
BCL2L1	IL15RA	SMS			GBP4	INSIG1
DENND3	ARL5B	FZD1			BCL3	LINC00641
STX12	LOC102724428	AAED1			HS3ST3B1	PNPT1
RAB12	ZBTB10	LITAF			GBP1	RAB39A
SPAG9	CDC42EP2	PIP5K1A			PELI1	MEFV
IFNGR2	GRHL1	NFE2L2			IFIT5	PARP14
DPYSL2	PPM1K	OGFRL1			LOC284648	SPACA6P
RAB5A	ECE1	ZNRF2			RNF19B	RTP4
EIF1B	CASP5	HMGCS1			ARAP2	LSS
LRP12	LOC101927902	ZCCHC2			TFAP2A	PI4K2B
N4BP2L1	TNFAIP3	DDX3Y			PDE4B	CAMK1G
ARHGAP31	WT1-AS	PTCHD3P1			TNF	CD80
IFFO2	GBP2	QKI			CD300E	CXCL11
RND3	RARRES1	ARID4B			SGPP2	MGAT4C

SAV1	DUSP16	DNAJB6		ANKRD33B	RNF144B
LOC374443	DUSP2	IFI16		GPX3	FOXF1
RFTN1	IL10	UBTD2		NR4A1	PTX3
FAM132B	CXCL12	CD82		PTGS2	SCN1B
MIR22HG	CCL3	SESTD1		PTGER2	SERPINB9
ZC3H12C	EBI3	CPD		TNFRSF9	HIVEP2
ZFY	OLR1	ARMCX1		CCL4L1	DLL4
TNFRSF10B	GPR157	NXPH4		RASL11A	ETV3L
PHLPP2	SYNPO2	TNFAIP8		HES4	CCDC147-AS1
TLR7	NAMPT	MST4			
ZBTB43	LINC00421	STAT5A			
JADE3	SDC4	NRARP			
GPC5	GRAMD3	TBC1D2B			
ADPRH	FAM171A1	MOB3C			

Table 7.1: Groups of upregulated genes utilised for transcription factor binding site analysis using HOMER

Downregulated: S80A		Downregulated: WT		Downregulated: similar	
GJC2C5orf55	ARHGAP30	FAM13A	ATP6V0E2	CD180	FRAT2
SLMO1	CD244	LOC102723993	DHX57	C11orf21	FAM174B
SIX3	MTMR12	SEMA3G	USP4	SIT1	MNT
NAP1L3	AAK1	KCNF1	TREML1	ZNF815P	PPP1R3E
ZNF836	ADCY9	LOC101928228	MRPL34	GPR65	DHRS4-AS1
PGPEP1	EVI2A	GPR55	C20orf194	RAB3A	PUS7L
MIR600HG	ACAD8	GPR82	SPIN1	CUX2	EHMT2
ARHGEF3	ZNF740	CEBPA-AS1	GIMAP4	TMEM169	NR2F6
CALHM2	ANKS1A	PPFIA4	KCNK13	EPHB3	PPP1R3G
AVPI1	HDAC5	FAM117B	OTUB2	LCNL1	ZNF589
CIPC	UAP1L1	GMCL1P1	NISCH	LOC102723672	UBL4A
FN3K	FRRS1	CEBPE	RTKN	RNF125	SLC45A4
SUFU	ZNF425	TMEM38A	PQLC2	SLC46A1	ZNF174
KLHL30	ESCO1	PIGR	SOCS6	TFAP4	MRPS26
USP2	VANGL1	CHN2	LOC102724791	C19orf35	LOC100507054
GREB1	TMEM187	SLC46A3	LOC100287896	LOC100996246	LEPREL4
OR2W3	CD1D	MTM1	LINC00094	PAQR7	LDLRAP1
LINC00920	TEX2	TMEM177	ANKRD39	LINGO3	DAG1
KLHDC8B	CFL2	MAPK14	ZADH2	TMEM37	RSAD1
GPRC5C	EPOR	IKBIP	DGCR2	ATOH8	ANKMY2
PDK4	PPCDC	KLHL18	RFPL1S	NLRC4	SASH3
LOC100507642	RGS14	NT5DC2	STX6	EEPD1	ZNF318
AC011899.9	SETD8	RAB11FIP1	TMEM9	FRAT1	ARMC7
CAMKK1	XYLT2	NAGK	ZFYVE21	MBLAC2	FAM203A
TLX1	SNX21	TRMT5	C6orf211	PDE7B	CCL28
SNORD3A	C1orf226	WDR91	VPS8	LRRC25	AKNA
LINC00663	TOB1	CPPED1	AKAP11	LOC101929536	RGS19
CARD10	TRAK1	GRAMD4	LGI2	CORO2A	FUT10
RAB11FIP4	LY9	FAM134A	ABI3	CAMSAP3	FAM53B
MTUS1	TBXA2R	EXOC3	CCDC132	HHEX	CEP250
GDPD1	MERTK	TLK1	MICALCL	DBP	ZFP64
PGF	GPRIN3	LEPROTL1	LOC100506985	SLC40A1	SETDB2
DLX3	PRR5L	UBXN2B	LAMB2P1	KCTD7	DIRAS1
RGS4	FAM102A	ARHGAP18	RASL11B	ZFP30	PLD6
ZNF280A	SESN1	LOC400548	CCDC170	ZNF395	MIEF2
VPS37D	ZCCHC24			B3GNT8	ZFP92
LOC101928230	RCBTB2			C3orf62	PDCD2L
PKDCC	PRKACB			PXMP4	GMCL1
MAGEE1	ZNF469			SOX12	ZNF786
DPF3	FILIP1L			RP11-426L16.8	C17orf59
C15orf52	PPP1R3F			ZNF33B	ZNF362
ID3	SLC38A1			HERC2P9	PP7080
TMEM86A	DKK2			MAP6D1	WDR92
KIF25-AS1	TGFBR2			LYL1	DGKZ
HOXB7	MTSS1L			ZNF10	ARHGAP35
MGC57346	PDLIM7			PARD6A	RGS12
CEMP1	ZNF281			C16orf54	AP3M2
FAM171A2	CAB39			ZNF74	BAIAP2-AS1
SNAI3	ST3GAL1			PDP2	PPAPDC3
LOC101927863	PIAS2			C11orf24	RNF166
LOC102724386	KANK2			ARV1	ZNF768
LOC101929666	LOC101928429			SPATC1L	PACSIN3
SH2D3C	DENND1B			FAM168A	ARID3B
P2RY1	LINC00984			RMND1	LOC101927070
USP51	MIOS			EVI2B	PNPLA3

RHOBTB2	FAM214B		SLC35C1	ZNF280B
HTR1B	ARRDC2		SCAMP5	WDR81
AC098823.3	TCEAL8		CXorf64	MTHFR
ARX	MADD		ZYG11B	THAP11
FAM212A	DLX4		GIT2	THRA
HOXA7	TUBA4A		ADCK2	POLRMT
NKX6-3	LRFN4		SH3BP5L	ZNF696
FGD3	MTA3		SMIM15	FITM2
THEM6	RBM38		INPP5D	ARL4C
WNT9A	MPI		WIPF1	CAMK2G
CDK5R1	KCNE3		XPC	TBC1D13
LMX1B	ARL11		GRK6	LRRC37BP1
SNX18	ZNF70		NAA40	ZNF616
IGIP	CCDC103		PEX6	ABHD8
KIAA1614	MAP10		LZTS2	ANKRD13D
LINC00324	ZC3HAV1L		TK2	TRIM32
MLNR	RNF168		APBB3	SLA
NHLRC4	GJD3		USP21	KRBA1
TNFSF14	GCNT1		MFSD5	LSM14B
ST3GAL5	ETV5		SLC7A6	ZFP62
CASP9	GPAM		TANGO2	SLITRK5
ACOT11	ZNF252P		CBR1	KLF16
PSMG3-AS1	IPCEF1		DOK2	ABHD6
TBL2	ETV5		FUT8	ALDH3A2
PER1	GABRG2		PRKAB1	HDHD3
PWWP2B	FLJ44313		COMMD9	ORAI3
TRAM2	PCDHB16		SLC25A38	MPV17L2
LRRC4	ADRB2		GDE1	CLEC3B
ATG16L1	C3orf58		KIAA1715	VPS26B
TRIM58	KLF4		C2orf43	RITA1
ZNF561	HPS3		SNAPC2	TMEM206
DUSP7	PLEKHA2		CSK	ZFP69
PCDHB8	PPP1R12B		IRX3	TMEM55A
MBP	XKR8		WDR55	MAGEF1
RAB11FIP5	MFSD9		ARHGEF7	AASDH
ZFPM1	NSD1		SEPHS2	GCLC
ZNF490	FADD		DBT	ASB13
CABIN1	PRR12		KIAA0930	KLF13
GIMAP6	RPRD2		NCF4	SIGLEC9
RAB3IL1	FAM84B		AGO4	GPATCH11
STOX2	NRROS		BAZ1B	MGME1
CEACAM21	AMIGO1		ICMT	HEATR3
ZBTB4	FBXL18		MLX	SFMBT2
RASGRP3	TNRC18		MCAT	NAT14
MAF	MAFG-AS1		C10orf91	TBC1D14
C9orf172	FLJ20464		RASSF7	PIP4K2B
CYTH4	SH2B2		FANCF	SDHAF1
FAM45B	DCLRE1A		PGBD2	SUOX
SLC35F6	NUDT18		KLHL35	MAP3K1
FAM219B	EEF2K		BCDIN3D	HLTF
GLTPD1	SWSAP1		RTN4R	TTC5
ZNF251	ZNF852		DOK3	NAT8L
GLCCI1	LOC101929523		TMEM223	FAM78A
EIF2AK4	BOLA1		ZNF439	LOC100289656
	ZNF780A		TCHP	OSBPL7
			WNT6	LOC93622
			MVK	RPP25

Table 7.2: Groups of downregulated genes utilised for transcription factor binding site analysis using HOMER

7.4 GSEA enrichment plots

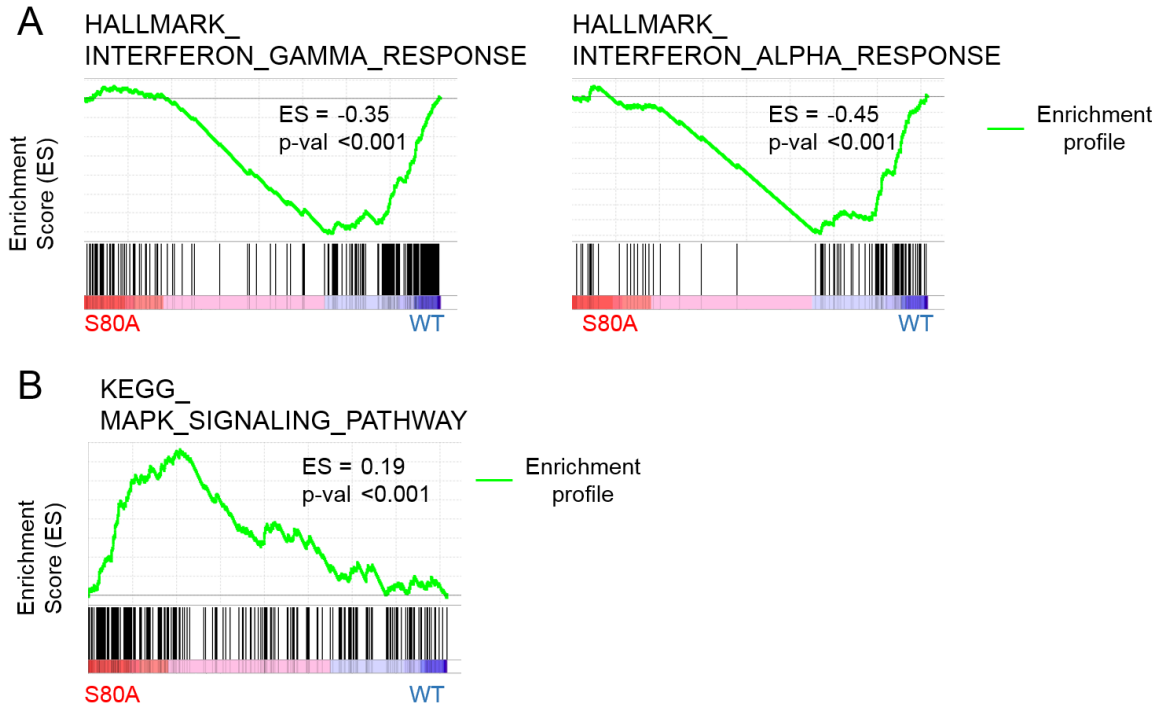


Figure 7.3: GSEA enrichment plots: LPS-treated WT vs *NFKB1*^{S80A}
Gene set enrichment analysis of WT + vs S80A + ranked gene list (CuffDiff). Enrichment plots show the enrichment profile (green curve) across the ranked gene list. Black lines represent the genes which contribute to the enrichment profile. Peak enrichment scores (ES) and p-values are shown. (A) 5th and 6th most significantly enriched hallmark gene sets at the bottom of the ranked gene list, indicated by negative enrichment score (expression higher in WT THP-1s compared to *NFKB1*^{S80A} THP-1s). (B) 6th most significantly enriched KEGG pathway at the top of the ranked gene list, indicated by positive enrichment score (expression higher in *NFKB1*^{S80A} THP-1s compared to WT THP-1s).

7.5 Upstream regulator analysis of untreated THP-1 cells

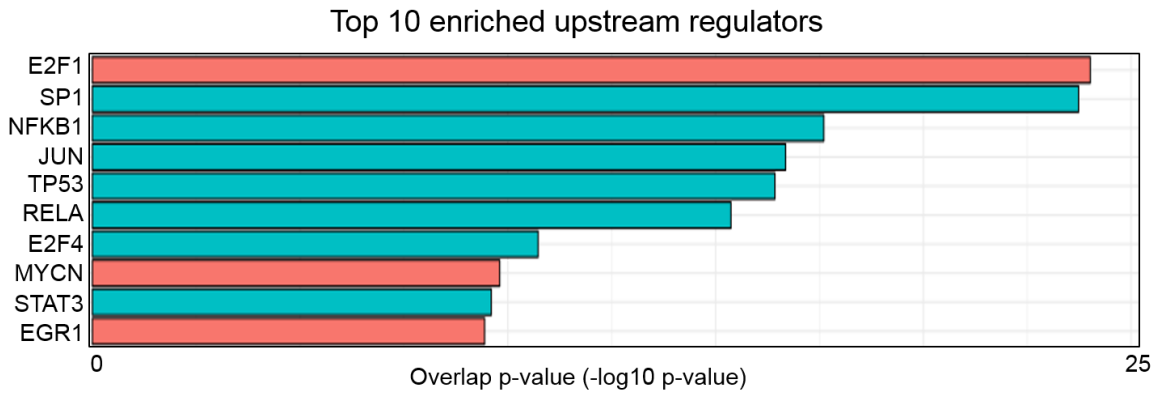


Figure 7.4: Upstream regulator enrichment analysis of untreated THP-1 cells
Upstream regulator enrichment analysis of DESeq2 calculated differentially expressed genes (p.adj <0.01) in unstimulated cells (WT – vs *NFKB1*^{S80A} –). Upstream regulators were ranked by enrichment score (overlap p-value (–log10 p-value)) and by activation score (z-score). Red bars indicate both significant enrichment score (p-value <0.05) and significant activation score (significant z-score >2). Blue bars indicate significant enrichment (p-value <0.05), but not significant activation score (z-score <2).

References

ANDERS, S., PYL, P. T. & HUBER, W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics,* 31**,** 166-169.

ANEST, V., HANSON, J. L., COGSWELL, P. C., STEINBRECHER, K. A., STRAHL, B. D. & BALDWIN, A. S. 2003. A nucleosomal function for IkappaB kinase-alpha in NF-kappaB-dependent gene expression. *Nature,* 423**,** 659-63.

ANRATHER, J., RACCHUMI, G. & IADECOLA, C. 2005. cis-acting, element-specific transcriptional activity of differentially phosphorylated nuclear factor-kappa B. *J Biol Chem,* 280**,** 244-52.

ARENZANA-SEISDEDOS, F., THOMPSON, J., RODRIGUEZ, M. S., BACHELERIE, F., THOMAS, D. & HAY, R. T. 1995. Inducible nuclear expression of newly synthesized I kappa B alpha negatively regulates DNA-binding and transcriptional activities of NF-kappa B. *Mol Cell Biol,* 15**,** 2689-96.

ARTIS, D., KANE, C. M., FIORE, J., ZAPH, C., SHAPIRA, S., JOYCE, K., MACDONALD, A., HUNTER, C., SCOTT, P. & PEARCE, E. J. 2005. Dendritic Cell-Intrinsic Expression of NF-κB1 Is Required to Promote Optimal Th2 Cell Differentiation. *The Journal of Immunology,* 174**,** 7154-7159.

ARTIS, D., SPEIRS, K., JOYCE, K., GOLDSCHMIDT, M., CAAMANO, J., HUNTER, C. A. & SCOTT, P. 2003. NF-kappa B1 is required for optimal CD4+ Th1 cell development and resistance to Leishmania major. *J Immunol,* 170**,** 1995-2003.

BAJWA, E. K., CREMER, P. C., GONG, M. N., ZHAI, R., SU, L., THOMPSON, B. T. & CHRISTIANI, D. C. 2011. An NFKB1 Promoter Insertion/Deletion Polymorphism Influences Risk and Outcome in Acute Respiratory Distress Syndrome among Caucasians. *PLoS ONE,* 6**,** e19469.

BANERJEE, A., GUGASYAN, R., MCMAHON, M. & GERONDAKIS, S. 2006. Diverse Toll-like receptors utilize Tpl2 to activate extracellular signal-regulated kinase (ERK) in hemopoietic cells. *Proc Natl Acad Sci U S A,* 103**,** 3274-9.

BARRANGOU, R., FREMAUX, C., DEVEAU, H., RICHARDS, M., BOYAVAL, P., MOINEAU, S., ROMERO, D. A. & HORVATH, P. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science,* 315**,** 1709-12.

BEINKE, S., DEKA, J., LANG, V., BELICH, M. P., WALKER, P. A., HOWELL, S., SMERDON, S. J., GAMBLIN, S. J. & LEY, S. C. 2003. NF-kappaB1 p105 negatively regulates TPL-2 MEK kinase activity. *Mol Cell Biol,* 23**,** 4739-52.

BELICH, M. P., SALMERON, A., JOHNSTON, L. H. & LEY, S. C. 1999. TPL-2 kinase regulates the proteolysis of the NF-kappaB-inhibitory protein NF-kappaB1 p105. *Nature,* 397**,** 363-8.

BIBIKOVA, M., CARROLL, D., SEGAL, D. J., TRAUTMAN, J. K., SMITH, J., KIM, Y.-G. & CHANDRASEGARAN, S. 2001. Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases. *Molecular and Cellular Biology,* 21**,** 289.

BLADH, L.-G., JOHANSSON-HAQUE, K., RAFTER, I., NILSSON, S. & OKRET, S. 2009. Inhibition of extracellular signal-regulated kinase (ERK) signaling participates in repression of nuclear factor (NF)-κB activity by glucocorticoids. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1793**,** 439-446.

BOLOTIN, A., QUINQUIS, B., SOROKIN, A. & EHRLICH, S. D. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology,* 151**,** 2551-61.

BOURS, V., FRANZOSO, G., AZARENKO, V., PARK, S., KANNO, T., BROWN, K. & SIEBENLIST, U. 1993. The oncoprotein Bcl-3 directly transactivates

through kappa B motifs via association with DNA-binding p50B homodimers. *Cell,* 72**,** 729-39.

BROUNS, S. J., JORE, M. M., LUNDGREN, M., WESTRA, E. R., SLIJKHUIS, R. J., SNIJDERS, A. P., DICKMAN, M. J., MAKAROVA, K. S., KOONIN, E. V. & VAN DER OOST, J. 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science,* 321**,** 960-4.

BUSS, H., DORRIE, A., SCHMITZ, M. L., HOFFMANN, E., RESCH, K. & KRACHT, M. 2004. Constitutive and interleukin-1-inducible phosphorylation of p65 NF-κB at serine 536 is mediated by multiple protein kinases including IκB kinase (IKK)-α, IKKβ, IKKε, TRAF family member-associated (TANK)-binding kinase 1 (TBK1), and an unknown kinase and couples p65 to TATA-binding protein-associated factor II31-mediated interleukin-8 transcription. *J Biol Chem,* 279**,** 55633-43.

CAI, X., CHIU, Y.-H. & CHEN, ZHIJIAN J. 2014. The cGAS-cGAMP-STING Pathway of Cytosolic DNA Sensing and Signaling. *Molecular Cell,* 54**,** 289-296.

CAMPBELL, I. K., VAN NIEUWENHUIJZE, A., SEGURA, E., O'DONNELL, K., COGHILL, E., HOMMEL, M., GERONDAKIS, S., VILLADANGOS, J. A. & WICKS, I. P. 2011. Differentiation of inflammatory dendritic cells is mediated by NF-kappaB1-dependent GM-CSF production in CD4 T cells. *J Immunol,* 186**,** 5468-77.

CARIAPPA, A., LIOU, H.-C., HORWITZ, B. H. & PILLAI, S. 2000. Nuclear Factor κb Is Required for the Development of Marginal Zone B Lymphocytes. *The Journal of Experimental Medicine,* 192**,** 1175-1182.

CARMODY, R. J., RUAN, Q., PALMER, S., HILLIARD, B. & CHEN, Y. H. 2007. Negative regulation of toll-like receptor signaling by NF-kappaB p50 ubiquitination blockade. *Science,* 317**,** 675-8.

CARTWRIGHT, T. N., WORRELL, J. C., MARCHETTI, L., DOWLING, C. M., KNOX, A., KIELY, P., MANN, J., MANN, D. A. & WILSON, C. L. 2018. HDAC1 interacts with the p50 NF-κB subunit via its nuclear localization sequence to constrain inflammatory gene expression. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*.

CARTY, M., GOODBODY, R., SCHRODER, M., STACK, J., MOYNAGH, P. N. & BOWIE, A. G. 2006. The human adaptor SARM negatively regulates adaptor protein TRIF-dependent Toll-like receptor signaling. *Nat Immunol,* 7**,** 1074-81.

CEN, H., ZHOU, M., LENG, R. X., WANG, W., FENG, C. C., LI, B. Z., ZHU, Y., YANG, X. K., YANG, M., ZHAI, Y., ZHANG, M., HU, L. F., LI, R., CHEN, G. M., CHEN, H., PAN, H. F., LI, X. P. & YE, D. Q. 2013. Genetic interaction between genes involved in NF-kappaB signaling pathway in systemic lupus erythematosus. *Mol Immunol,* 56**,** 643-8.

CHAN, F. K., CHUN, H. J., ZHENG, L., SIEGEL, R. M., BUI, K. L. & LENARDO, M. J. 2000. A domain in TNF receptors that mediates ligand-independent receptor assembly and signaling. *Science,* 288**,** 2351-4.

CHANPUT, W., MES, J. J. & WICHERS, H. J. 2014. THP-1 cell line: an in vitro cell model for immune modulation approach. *Int Immunopharmacol,* 23**,** 37-45.

CHARIOT, A. 2009. The NF-κB-independent functions of IKK subunits in immunity and cancer. *Trends in Cell Biology,* 19**,** 404-413.

CHEN, Y. Q., GHOSH, S. & GHOSH, G. 1998. A novel DNA recognition mode by the NF-kappa B p65 homodimer. *Nat Struct Biol,* 5**,** 67-73.

CHENG, C. S., FELDMAN, K. E., LEE, J., VERMA, S., HUANG, D. B., HUYNH, K., CHANG, M., PONOMARENKO, J. V., SUN, S. C., BENEDICT, C. A., GHOSH,

G. & HOFFMANN, A. 2011. The specificity of innate immune responses is enforced by repression of interferon response elements by NF-kappaB p50. *Sci Signal,* 4**,** ra11.

CHRISTIAN, F., SMITH, E. L. & CARMODY, R. J. 2016. The Regulation of NF-kappaB Subunits by Phosphorylation. *Cells,* 5**,** 1-19.

CHRISTIAN, M., CERMAK, T., DOYLE, E. L., SCHMIDT, C., ZHANG, F., HUMMEL, A., BOGDANOVE, A. J. & VOYTAS, D. F. 2010. Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics,* 186**,** 757-761.

COHEN, S., ACHBERT-WEINER, H. & CIECHANOVER, A. 2004. Dual effects of IkappaB kinase beta-mediated phosphorylation on p105 Fate: SCF(beta-TrCP)-dependent degradation and SCF(beta-TrCP)-independent processing. *Mol Cell Biol,* 24**,** 475-86.

COLLINS, P. E., KIELY, P. A. & CARMODY, R. J. 2014. Inhibition of transcription by B cell Leukemia 3 (Bcl-3) protein requires interaction with nuclear factor kappaB (NF-kappaB) p50. *J Biol Chem,* 289**,** 7059-67.

COLLINS, P. E., MITXITORENA, I. & CARMODY, R. J. 2016. The Ubiquitination of NF-kappaB Subunits in the Control of Transcription. *Cells,* 5.

CONG, L., RAN, F. A., COX, D., LIN, S., BARRETTO, R., HABIB, N., HSU, P. D., WU, X., JIANG, W., MARRAFFINI, L. A. & ZHANG, F. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science,* 339**,** 819-23.

COURTOIS, G. & GILMORE, T. D. 2006. Mutations in the NF-kappaB signaling pathway: implications for human disease. *Oncogene,* 25**,** 6831-43.

CRAMER, P., LARSON, C. J., VERDINE, G. L. & MÜLLER, C. W. 1997. Structure of the human NF-kappaB p52 homodimer-DNA complex at 2.1 A resolution. *The EMBO Journal,* 16**,** 7078-7090.

CRAWLEY, C. D., KANG, S., BERNAL, G. M., WAHLSTROM, J. S., VOCE, D. J., CAHILL, K. E., GAROFALO, A., RALEIGH, D. R., WEICHSELBAUM, R. R. & YAMINI, B. 2015. S-phase-dependent p50/NF-small ka, CyrillicB1 phosphorylation in response to ATR and replication stress acts to maintain genomic stability. *Cell Cycle,* 14**,** 566-76.

CRAWLEY, C. D., RALEIGH, D. R., KANG, S., VOCE, D. J., SCHMITT, A. M., WEICHSELBAUM, R. R. & YAMINI, B. 2013. DNA damage-induced cytotoxicity is mediated by the cooperative interaction of phospho-NF-kappaB p50 and a single nucleotide in the kappaB-site. *Nucleic Acids Res,* 41**,** 764-74.

CROOKS, G. E., HON, G., CHANDONIA, J. M. & BRENNER, S. E. 2004. WebLogo: a sequence logo generator. *Genome Res,* 14**,** 1188-90.

DECKERT-SCHLÜTER, M., BLUETHMANN, H., RANG, A., HOF, H. & SCHLÜTER, D. 1998. Crucial Role of TNF Receptor Type 1 (p55), But Not of TNF Receptor Type 2 (p75), in Murine Toxoplasmosis. *The Journal of Immunology,* 160**,** 3427.

DELHASE, M., HAYAKAWA, M., CHEN, Y. & KARIN, M. 1999. Positive and negative regulation of IkappaB kinase activity through IKKbeta subunit phosphorylation. *Science,* 284**,** 309-13.

DEMARCHI, F., BERTOLI, C., SANDY, P. & SCHNEIDER, C. 2003. Glycogen synthase kinase-3 beta regulates NF-kappa B1/p105 stability. *J Biol Chem,* 278**,** 39583-90.

DUMITRU, C. D., CECI, J. D., TSATSANIS, C., KONTOYIANNIS, D., STAMATAKIS, K., LIN, J. H., PATRIOTIS, C., JENKINS, N. A., COPELAND, N. G., KOLLIAS, G. & TSICHLIS, P. N. 2000. TNF-alpha induction by LPS is regulated posttranscriptionally via a Tpl2/ERK-dependent pathway. *Cell,* 103**,** 1071-83.

EA, C. K., DENG, L., XIA, Z. P., PINEDA, G. & CHEN, Z. J. 2006. Activation of IKK by TNFalpha requires site-specific ubiquitination of RIP1 and polyubiquitin binding by NEMO. *Mol Cell,* 22**,** 245-57.

ELSHARKAWY, A. M., OAKLEY, F., LIN, F., PACKHAM, G., MANN, D. A. & MANN, J. 2010. The NF-kappaB p50:p50:HDAC-1 repressor complex orchestrates transcriptional inhibition of multiple pro-inflammatory genes. *J Hepatol,* 53**,** 519-27.

FAN, C. M. & MANIATIS, T. 1991. Generation of p50 subunit of NF-kappa B by processing of p105 through an ATP-dependent pathway. *Nature,* 354**,** 395-8.

FIORINI, E., SCHMITZ, I., MARISSEN, W. E., OSBORN, S. L., TOUMA, M., SASADA, T., RECHE, P. A., TIBALDI, E. V., HUSSEY, R. E., KRUISBEEK, A. M., REINHERZ, E. L. & CLAYTON, L. K. 2002. Peptide-induced negative selection of thymocytes activates transcription of an NF-kappa B inhibitor. *Mol Cell,* 9**,** 637-48.

FRANK, R. 2002. The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports--principles and applications. *J Immunol Methods,* 267**,** 13-26.

FU, Y., FODEN, J. A., KHAYTER, C., MAEDER, M. L., REYON, D., JOUNG, J. K. & SANDER, J. D. 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnology,* 31**,** 822.

FUJIMOTO, K., YASUDA, H., SATO, Y. & YAMAMOTO, K. 1995. A role for phosphorylation in the proteolytic processing of the human NF-kappa B1 precursor. *Gene,* 165**,** 183-9.

FUJITA, T., NOLAN, G. P., LIOU, H. C., SCOTT, M. L. & BALTIMORE, D. 1993. The candidate proto-oncogene bcl-3 encodes a transcriptional coactivator that activates through NF-kappa B p50 homodimers. *Genes Dev,* 7**,** 1354-63.

GENG, H., WITTWER, T., DITTRICH-BREIHOLZ, O., KRACHT, M. & SCHMITZ, M. L. 2009. Phosphorylation of NF-kappaB p65 at Ser468 controls its COMMD1-dependent ubiquitination and target gene-specific proteasomal elimination. *EMBO Rep,* 10**,** 381-6.

GERONDAKIS, S., GRUMONT, R., GUGASYAN, R., WONG, L., ISOMURA, I., HO, W. & BANERJEE, A. 2006. Unravelling the complexities of the NF-kappaB signalling pathway using mouse knockout and transgenic models. *Oncogene,* 25**,** 6781-99.

GHOSH, G., VAN DUYNE, G., GHOSH, S. & SIGLER, P. B. 1995. Structure of NF-kappa B p50 homodimer bound to a kappa B site. *Nature,* 373**,** 303-10.

GHOSH, S. & KARIN, M. 2002. Missing Pieces in the NF-κB Puzzle. *Cell,* 109**,** S81-S96.

GORDON, S., PLÜDDEMANN, A. & MARTINEZ ESTRADA, F. 2014. Macrophage heterogeneity in tissues: phenotypic diversity and functions. *Immunological Reviews,* 262**,** 36-55.

GREENE, W. C. & CHEN, L. F. 2004. Regulation of NF-kappaB action by reversible acetylation. *Novartis Found Symp,* 259**,** 208-17; discussion 218-25.

GRUMONT, R. J., ROURKE, I. J., O'REILLY, L. A., STRASSER, A., MIYAKE, K., SHA, W. & GERONDAKIS, S. 1998. B lymphocytes differentially use the Rel and nuclear factor kappaB1 (NF-kappaB1) transcription factors to regulate cell cycle progression and apoptosis in quiescent and mitogen-activated cells. *J Exp Med,* 187**,** 663-74.

GUAN, H., HOU, S. & RICCIARDI, R. P. 2005. DNA binding of repressor nuclear factor-kappaB p50/p50 depends on phosphorylation of Ser337 by the protein kinase A catalytic subunit. *J Biol Chem,* 280**,** 9957-62.

HALLER, D., RUSSO, M. P., SARTOR, R. B. & JOBIN, C. 2002. IKK beta and phosphatidylinositol 3-kinase/Akt participate in non-pathogenic Gram-negative enteric bacteria-induced RelA phosphorylation and NF-kappa B activation in both primary and intestinal epithelial cell lines. *J Biol Chem,* 277**,** 38168-78.

HAN, H., CHO, J. W., LEE, S., YUN, A., KIM, H., BAE, D., YANG, S., KIM, C. Y., LEE, M., KIM, E., LEE, S., KANG, B., JEONG, D., KIM, Y., JEON, H. N., JUNG, H., NAM, S., CHUNG, M., KIM, J. H. & LEE, I. 2018. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res,* 46**,** D380-d386.

HAYDEN, M. S. & GHOSH, S. 2012. NF-kappaB, the first quarter-century: remarkable progress and outstanding questions. *Genes Dev,* 26**,** 203-34.

HEINZ, S., BENNER, C., SPANN, N., BERTOLINO, E., LIN, Y. C., LASLO, P., CHENG, J. X., MURRE, C., SINGH, H. & 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**,** 576-89.

HEISSMEYER, V., KRAPPMANN, D., HATADA, E. N. & SCHEIDEREIT, C. 2001. Shared pathways of IkappaB kinase-induced SCF(betaTrCP)-mediated ubiquitination and degradation for the NF-kappaB precursor p105 and IkappaBalpha. *Mol Cell Biol,* 21**,** 1024-35.

HEISSMEYER, V., KRAPPMANN, D., WULCZYN, F. G. & SCHEIDEREIT, C. 1999. NF-kappaB p105 is a target of IkappaB kinases and controls signal induction of Bcl-3-p50 complexes. *EMBO J,* 18**,** 4766-78.

HEYER, W. D., EHMSEN, K. T. & LIU, J. 2010. Regulation of homologous recombination in eukaryotes. *Annu Rev Genet,* 44**,** 113-39.

HINZ, M., ARSLAN, S. C. & SCHEIDEREIT, C. 2012. It takes two to tango: IkappaBs, the multifunctional partners of NF-kappaB. *Immunol Rev,* 246**,** 59-76.

HOCHRAINER, K., RACCHUMI, G. & ANRATHER, J. 2013. Site-specific phosphorylation of the p65 protein subunit mediates selective gene expression by differential NF-kappaB and RNA polymerase II promoter recruitment. *J Biol Chem,* 288**,** 285-93.

HOFFMANN, A., LEUNG, T. H. & BALTIMORE, D. 2003. Genetic analysis of NF-kappaB/Rel transcription factors defines functional specificities. *Embo j,* 22**,** 5530-9.

HOFFMANN, A., NATOLI, G. & GHOSH, G. 2006. Transcriptional regulation via the NF-κB signaling module. *Oncogene,* 25**,** 6706.

HONDA, K., TAKAOKA, A. & TANIGUCHI, T. 2006. Type I Inteferon Gene Induction by the Interferon Regulatory Factor Family of Transcription Factors. *Immunity,* 25**,** 349-360.

HOU, S., GUAN, H. & RICCIARDI, R. P. 2003. Phosphorylation of serine 337 of NF-kappaB p50 is critical for DNA binding. *J Biol Chem,* 278**,** 45994-8.

HSU, P. D., SCOTT, D. A., WEINSTEIN, J. A., RAN, F. A., KONERMANN, S., AGARWALA, V., LI, Y., FINE, E. J., WU, X., SHALEM, O., CRADICK, T. J., MARRAFFINI, L. A., BAO, G. & ZHANG, F. 2013. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology,* 31**,** 827.

HUANG, D., YANG, L., LIU, Y., ZHOU, Y., GUO, Y., PAN, M., WANG, Y., TAN, Y., ZHONG, H., HU, M., LU, W., JI, W., WANG, J., RAN, P., ZHONG, N., ZHOU, Y. & LU, J. 2013. Functional polymorphisms in NFkappaB1/IkappaBalpha predict risks of chronic obstructive pulmonary disease and lung cancer in Chinese. *Hum Genet,* 132**,** 451-60.

HUANG, D. B., CHEN, Y. Q., RUETSCHE, M., PHELPS, C. B. & GHOSH, G. 2001. X-ray crystal structure of proto-oncogene product c-Rel bound to the CD28 response element of IL-2. *Structure,* 9**,** 669-78.

HUANG, T. T., KUDO, N., YOSHIDA, M. & MIYAMOTO, S. 2000. A nuclear export signal in the N-terminal regulatory domain of IkappaBalpha controls cytoplasmic localization of inactive NF-kappaB/IkappaBalpha complexes. *Proc Natl Acad Sci U S A,* 97**,** 1014-9.

HUXFORD, T. & GHOSH, G. 2009. A structural guide to proteins of the NF-kappaB signaling module. *Cold Spring Harb Perspect Biol,* 1**,** a000075.

HUXFORD, T., HUANG, D. B., MALEK, S. & GHOSH, G. 1998. The crystal structure of the IkappaBalpha/NF-kappaB complex reveals mechanisms of NF-kappaB inactivation. *Cell,* 95**,** 759-70.

ISHIKAWA, H., CARRASCO, D., CLAUDIO, E., RYSECK, R. P. & BRAVO, R. 1997. Gastric hyperplasia and increased proliferative responses of lymphocytes in mice lacking the COOH-terminal ankyrin domain of NF-kappaB2. *J Exp Med,* 186**,** 999-1014.

ISHIMARU, N., KISHIMOTO, H., HAYASHI, Y. & SPRENT, J. 2006. Regulation of naive T cell function by the NF-kappaB2 pathway. *Nat Immunol,* 7**,** 763-72.

ISRAEL, A. 2010. The IKK complex, a central regulator of NF-kappaB activation. *Cold Spring Harb Perspect Biol,* 2**,** a000158.

JACOBS, M. D. & HARRISON, S. C. 1998. Structure of an IkappaBalpha/NF-kappaB complex. *Cell,* 95**,** 749-58.

JANSSENS, S. & BEYAERT, R. 2002. A universal role for MyD88 in TLR/IL-1R-mediated signaling. *Trends in Biochemical Sciences,* 27**,** 474-482.

JINEK, M., CHYLINSKI, K., FONFARA, I., HAUER, M., DOUDNA, J. A. & CHARPENTIER, E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science,* 337**,** 816-21.

JOHNSON, C., VAN ANTWERP, D. & HOPE, T. J. 1999. An N-terminal nuclear export signal is required for the nucleocytoplasmic shuttling of IkappaBalpha. *The EMBO Journal,* 18**,** 6682-6693.

JU, J., NAURA, A. S., ERRAMI, Y., ZERFAOUI, M., KIM, H., KIM, J. G., ABD ELMAGEED, Z. Y., ABDEL-MAGEED, A. B., GIARDINA, C., BEG, A. A., SMULSON, M. E. & BOULARES, A. H. 2010. Phosphorylation of p50 NF-kappaB at a single serine residue by DNA-dependent protein kinase is critical for VCAM-1 expression upon TNF treatment. *J Biol Chem,* 285**,** 41152-60.

KARBAN, A. S., OKAZAKI, T., PANHUYSEN, C. I., GALLEGOS, T., POTTER, J. J., BAILEY-WILSON, J. E., SILVERBERG, M. S., DUERR, R. H., CHO, J. H., GREGERSEN, P. K., WU, Y., ACHKAR, J. P., DASSOPOULOS, T., MEZEY, E., BAYLESS, T. M., NOUVET, F. J. & BRANT, S. R. 2004. Functional annotation of a novel NFKB1 promoter polymorphism that increases risk for ulcerative colitis. *Hum Mol Genet,* 13**,** 35-45.

KAWASAKI, T. & KAWAI, T. 2014. Toll-Like Receptor Signaling Pathways. *Frontiers in Immunology,* 5**,** 461.

KIELY, P. A., BAILLIE, G. S., BARRETT, R., BUCKLEY, D. A., ADAMS, D. R., HOUSLAY, M. D. & O'CONNOR, R. 2009. Phosphorylation of RACK1 on tyrosine 52 by c-Abl is required for insulin-like growth factor I-mediated regulation of focal adhesion kinase. *J Biol Chem,* 284**,** 20263-74.

KIM, D., LANGMEAD, B. & SALZBERG, S. L. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods,* 12**,** 357-362.

KIM, Y. G., CHA, J. & CHANDRASEGARAN, S. 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences,* 93**,** 1156.

KRAVTSOVA-IVANTSIV, Y., SHOMER, I., COHEN-KAPLAN, V., SNIJDER, B., SUPERTI-FURGA, G., GONEN, H., SOMMER, T., ZIV, T., ADMON, A., NARODITSKY, I., JBARA, M., BRIK, A., PIKARSKY, E., KWON, YONG T., DOWECK, I. & CIECHANOVER, A. 2015. KPC1-Mediated Ubiquitination and Proteasomal Processing of NF-κB1 p105 to p50 Restricts Tumor Growth. *Cell,* 161**,** 333-347.

KUNSCH, C., RUBEN, S. M. & ROSEN, C. A. 1992. Selection of optimal kappa B/Rel DNA-binding motifs: interaction of both subunits of NF-kappa B with DNA is required for transcriptional activation. *Mol Cell Biol,* 12**,** 4412-21.

KUWATA, H., MATSUMOTO, M., ATARASHI, K., MORISHITA, H., HIROTANI, T., KOGA, R. & TAKEDA, K. 2006. IkappaBNS inhibits induction of a subset of Toll-like receptor-dependent genes and limits inflammation. *Immunity,* 24**,** 41-51.

LANG, V., JANZEN, J., FISCHER, G. Z., SONEJI, Y., BEINKE, S., SALMERON, A., ALLEN, H., HAY, R. T., BEN-NERIAH, Y. & LEY, S. C. 2003. betaTrCP-mediated proteolysis of NF-kappaB1 p105 requires phosphorylation of p105 serines 927 and 932. *Mol Cell Biol,* 23**,** 402-13.

LAWRENCE, T., BEBIEN, M., LIU, G. Y., NIZET, V. & KARIN, M. 2005. IKKalpha limits macrophage NF-kappaB activation and contributes to the resolution of inflammation. *Nature,* 434**,** 1138-43.

LEE, J., MIRA-ARBIBE, L. & ULEVITCH, R. J. 2000. TAK1 regulates multiple protein kinase cascades activated by bacterial lipopolysaccharide. *J Leukoc Biol,* 68**,** 909-15.

LEFSTIN, J. A. & YAMAMOTO, K. R. 1998. Allosteric effects of DNA on transcriptional regulators. *Nature,* 392**,** 885-888.

LENARDO, M. J., FAN, C. M., MANIATIS, T. & BALTIMORE, D. 1989. The involvement of NF-kappa B in beta-interferon gene regulation reveals its role as widely inducible mediator of signal transduction. *Cell,* 57**,** 287-94.

LEUNG, T. H., HOFFMANN, A. & BALTIMORE, D. 2004. One nucleotide in a kappaB site can determine cofactor specificity for NF-kappaB dimers. *Cell,* 118**,** 453-64.

LI, J., YIN, Q. & WU, H. 2013. Structural Basis of Signal Transduction in the TNF Receptor Superfamily. *Advances in immunology,* 119**,** 135-153.

LI, T., HUANG, S., JIANG, W. Z., WRIGHT, D., SPALDING, M. H., WEEKS, D. P. & YANG, B. 2011. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Research,* 39**,** 359-372.

LIANG, X., POTTER, J., KUMAR, S., RAVINDER, N. & CHESNUT, J. D. 2017. Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA, Cas9 nuclease, and donor DNA. *Journal of Biotechnology,* 241**,** 136-146.

LIBERZON, A., BIRGER, C., THORVALDSDOTTIR, H., GHANDI, M., MESIROV, J. P. & TAMAYO, P. 2015. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst,* 1**,** 417-425.

LIBERZON, A., SUBRAMANIAN, A., PINCHBACK, R., THORVALDSDOTTIR, H., TAMAYO, P. & MESIROV, J. P. 2011. Molecular signatures database (MSigDB) 3.0. *Bioinformatics,* 27**,** 1739-40.

LIN, L., DEMARTINO, G. N. & GREENE, W. C. 1998. Cotranslational biogenesis of NF-kappaB p50 by the 26S proteasome. *Cell,* 92**,** 819-28.

LIN, L. & GHOSH, S. 1996. A glycine-rich region in NF-kappaB p105 functions as a processing signal for the generation of the p50 subunit. *Mol Cell Biol,* 16**,** 2248-54.

LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods,* 25**,** 402-8.

LOETSCHER, H., STUEBER, D., BANNER, D., MACKAY, F. & LESSLAUER, W. 1993. Human tumor necrosis factor alpha (TNF alpha) mutants with exclusive specificity for the 55-kDa or 75-kDa TNF receptors. *J Biol Chem,* 268**,** 26350-7.

LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology,* 15**,** 550-571.

LU, Y.-C., YEH, W.-C. & OHASHI, P. S. 2008. LPS/TLR4 signal transduction pathway. *Cytokine,* 42**,** 145-151.

LUND, M. E., TO, J., O'BRIEN, B. A. & DONNELLY, S. 2016. The choice of phorbol 12-myristate 13-acetate differentiation protocol influences the response of THP-1 macrophages to a pro-inflammatory stimulus. *J Immunol Methods,* 430**,** 64-70.

MALEK, S., CHEN, Y., HUXFORD, T. & GHOSH, G. 2001. IκBβ, but Not IκBα, Functions as a Classical Cytoplasmic Inhibitor of NF-κB Dimers by Masking Both NF-κB Nuclear Localization Sequences in Resting Cells. *Journal of Biological Chemistry,* 276**,** 45225-45235.

MALEK, S., HUANG, D.-B., HUXFORD, T., GHOSH, S. & GHOSH, G. 2003. X-ray Crystal Structure of an IκBβ·NF-κB p65 Homodimer Complex. *Journal of Biological Chemistry,* 278**,** 23094-23100.

MALI, P., YANG, L., ESVELT, K. M., AACH, J., GUELL, M., DICARLO, J. E., NORVILLE, J. E. & CHURCH, G. M. 2013. RNA-guided human genome engineering via Cas9. *Science,* 339**,** 823-6.

MAO, X., GLUCK, N., LI, D., MAINE, G. N., LI, H., ZAIDI, I. W., REPAKA, A., MAYO, M. W. & BURSTEIN, E. 2009. GCN5 is a required cofactor for a ubiquitin ligase that targets NF-κB/RelA. *Genes & Development,* 23**,** 849-861.

MARUYAMA, T., DOUGAN, S. K., TRUTTMANN, M. C., BILATE, A. M., INGRAM, J. R. & PLOEGH, H. L. 2015. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol,* 33**,** 538-42.

MCFARLANE, S. M., JUPP, O. J., COBBAN, H. J., HUNTER, I., ANDERSON, H. M., VANDENABEELE, P., NIXON, G. F. & MACEWAN, D. J. 2001. Stimulation of stress-activated but not mitogen-activated protein kinases by tumour necrosis factor receptor subtypes in airway smooth muscle. *Biochem Pharmacol,* 61**,** 749-59.

MCLEAN, C. Y., BRISTOR, D., HILLER, M., CLARKE, S. L., SCHAAR, B. T., LOWE, C. B., WENGER, A. M. & BEJERANO, G. 2010. GREAT improves functional interpretation of cis-regulatory regions. *Nature biotechnology,* 28**,** 495-501.

MEDZHITOV, R. & JANEWAY, C. A., JR. 1997. Innate immunity: the virtues of a nonclonal system of recognition. *Cell,* 91**,** 295-8.

MEDZHITOV, R., PRESTON-HURLBURT, P., KOPP, E., STADLEN, A., CHEN, C., GHOSH, S. & JANEWAY, C. A., JR. 1998. MyD88 is an adaptor protein in the hToll/IL-1 receptor family signaling pathways. *Mol Cell,* 2**,** 253-8.

MERCURIO, F., ZHU, H., MURRAY, B. W., SHEVCHENKO, A., BENNETT, B. L., LI, J., YOUNG, D. B., BARBOSA, M., MANN, M., MANNING, A. & RAO, A. 1997. IKK-1 and IKK-2: cytokine-activated IkappaB kinases essential for NF-kappaB activation. *Science,* 278**,** 860-6.

MILANOVIC, M., KRACHT, M. & SCHMITZ, M. L. 2014. The cytokine-induced conformational switch of nuclear factor kappaB p65 is mediated by p65 phosphorylation. *Biochem J,* 457**,** 401-13.

MITCHELL, J. P. & CARMODY, R. J. 2018. NF-kappaB and the Transcriptional Control of Inflammation. *Int Rev Cell Mol Biol,* 335**,** 41-84.

MIZGERD, J. P., LUPA, M. M., KOGAN, M. S., WARREN, H. B., KOBZIK, L. & TOPULOS, G. P. 2003. Nuclear factor-kappaB p50 limits inflammation and prevents lung injury during Escherichia coli pneumonia. *Am J Respir Crit Care Med,* 168**,** 810-7.

MOJICA, F. J., DIEZ-VILLASENOR, C., GARCIA-MARTINEZ, J. & SORIA, E. 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol,* 60**,** 174-82.

MORENO, R., SOBOTZIK, J. M., SCHULTZ, C. & SCHMITZ, M. L. 2010. Specification of the NF-κB transcriptional response by p65 phosphorylation and TNF-induced nuclear translocation of IKKε. *Nucleic Acids Res,* 38**,** 6029-44.

MÜLLER, C. W. & HARRISON, S. C. 1995. The structure of the NF-kappa B p50:DNA-complex: a starting point for analyzing the Rel family. *FEBS Lett,* 369**,** 113-7.

NAUDE, P. J., DEN BOER, J. A., LUITEN, P. G. & EISEL, U. L. 2011. Tumor necrosis factor receptor cross-talk. *Febs j,* 278**,** 888-98.

NAUMANN, M., WULCZYN, F. G. & SCHEIDEREIT, C. 1993. The NF-kappa B precursor p105 and the proto-oncogene product Bcl-3 are I kappa B molecules and control nuclear translocation of NF-kappa B. *Embo j,* 12**,** 213-22.

NOLAN, G. P., FUJITA, T., BHATIA, K., HUPPI, C., LIOU, H. C., SCOTT, M. L. & BALTIMORE, D. 1993. The bcl-3 proto-oncogene encodes a nuclear I kappa B-like molecule that preferentially interacts with NF-kappa B p50 and p52 in a phosphorylation-dependent manner. *Mol Cell Biol,* 13**,** 3557-66.

O'NEILL, L. A. & BOWIE, A. G. 2007. The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nat Rev Immunol,* 7**,** 353-64.

OAKLEY, F., MANN, J., NAILARD, S., SMART, D. E., MUNGALSINGH, N., CONSTANDINOU, C., ALI, S., WILSON, S. J., MILLWARD-SADLER, H., IREDALE, J. P. & MANN, D. A. 2005. Nuclear factor-kappaB1 (p50) limits the inflammatory and fibrogenic responses to chronic injury. *Am J Pathol,* 166**,** 695-708.

OECKINGHAUS, A. & GHOSH, S. 2009. The NF-kappaB family of transcription factors and its regulation. *Cold Spring Harb Perspect Biol,* 1**,** a000034.

OGANESYAN, G., SAHA, S. K., GUO, B., HE, J. Q., SHAHANGIAN, A., ZARNEGAR, B., PERRY, A. & CHENG, G. 2006. Critical role of TRAF3 in the Toll-like receptor-dependent and -independent antiviral response. *Nature,* 439**,** 208-11.

ORIAN, A., GONEN, H., BERCOVICH, B., FAJERMAN, I., EYTAN, E., ISRAEL, A., MERCURIO, F., IWAI, K., SCHWARTZ, A. L. & CIECHANOVER, A. 2000. SCF(beta)(-TrCP) ubiquitin ligase-mediated processing of NF-kappaB p105 requires phosphorylation of its C-terminus by IkappaB kinase. *EMBO J,* 19**,** 2580-91.

PANZER, U., STEINMETZ, O. M., TURNER, J. E., MEYER-SCHWESINGER, C., VON RUFFER, C., MEYER, T. N., ZAHNER, G., GOMEZ-GUERRERO, C., SCHMID, R. M., HELMCHEN, U., MOECKEL, G. W., WOLF, G., STAHL, R. A. & THAISS, F. 2009. Resolution of renal inflammation: a new role for NF-kappaB1 (p50) in inflammatory kidney diseases. *Am J Physiol Renal Physiol,* 297**,** F429-39.

PAQUET, D., KWART, D., CHEN, A., SPROUL, A., JACOB, S., TEO, S., OLSEN, K. M., GREGG, A., NOGGLE, S. & TESSIER-LAVIGNE, M. 2016. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature,* 533**,** 125-9.

PERTEA, M., PERTEA, G. M., ANTONESCU, C. M., CHANG, T.-C., MENDELL, J. T. & SALZBERG, S. L. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology,* 33**,** 290.

PFEFFER, L. M. 2011. The role of nuclear factor kappaB in the interferon response. *J Interferon Cytokine Res,* 31**,** 553-9.

PHELPS, C. B., SENGCHANTHALANGSY, L. L., MALEK, S. & GHOSH, G. 2000. Mechanism of kappa B DNA binding by Rel/NF-kappa B dimers. *J Biol Chem,* 275**,** 24392-9.

POHL, T., GUGASYAN, R., GRUMONT, R. J., STRASSER, A., METCALF, D., TARLINTON, D., SHA, W., BALTIMORE, D. & GERONDAKIS, S. 2002. The combined absence of NF-kappa B1 and c-Rel reveals that overlapping roles for these transcription factors in the B cell lineage are restricted to the activation and function of mature cells. *Proc Natl Acad Sci U S A,* 99**,** 4514-9.

POLTORAK, A., HE, X., SMIRNOVA, I., LIU, M. Y., VAN HUFFEL, C., DU, X., BIRDWELL, D., ALEJOS, E., SILVA, M., GALANOS, C., FREUDENBERG, M., RICCIARDI-CASTAGNOLI, P., LAYTON, B. & BEUTLER, B. 1998. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science,* 282**,** 2085-8.

POURCEL, C., SALVIGNOL, G. & VERGNAUD, G. 2005. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology,* 151**,** 653-63.

RAN, F. A., HSU, P. D., WRIGHT, J., AGARWALA, V., SCOTT, D. A. & ZHANG, F. 2013. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc,* 8**,** 2281-2308.

RAPAPORT, F., KHANIN, R., LIANG, Y., PIRUN, M., KREK, A., ZUMBO, P., MASON, C. E., SOCCI, N. D. & BETEL, D. 2013. Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. *Genome Biology,* 14**,** R95-R95.

RICE, N. R., MACKICHAN, M. L. & ISRAEL, A. 1992. The precursor of NF-kappa B p50 has I kappa B-like functions. *Cell,* 71**,** 243-53.

RICHARDSON, C. D., RAY, G. J., DEWITT, M. A., CURIE, G. L. & CORN, J. E. 2016. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat Biotechnol,* 34**,** 339-44.

RICHARDSON, E. T., SHUKLA, S., NAGY, N., BOOM, W. H., BECK, R. C., ZHOU, L., LANDRETH, G. E. & HARDING, C. V. 2015. ERK Signaling Is Essential for Macrophage Development. *PLoS ONE,* 10**,** e0140064.

RODRIGUEZ, M. S., MICHALOPOULOS, I., ARENZANA-SEISDEDOS, F. & HAY, R. T. 1995. Inducible degradation of I kappa B alpha in vitro and in vivo requires the acidic C-terminal domain of the protein. *Mol Cell Biol,* 15**,** 2413-9.

ROGERS, S., WELLS, R. & RECHSTEINER, M. 1986. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science,* 234**,** 364-8.

ROHS, R., WEST, S. M., SOSINSKY, A., LIU, P., MANN, R. S. & HONIG, B. 2009. The role of DNA shape in protein–DNA recognition. *Nature,* 461**,** 1248.

ROLOVA, T., PULI, L., MAGGA, J., DHUNGANA, H., KANNINEN, K., WOJCIEHOWSKI, S., SALMINEN, A., TANILA, H., KOISTINAHO, J. & MALM, T. 2014. Complex regulation of acute and chronic neuroinflammatory responses in mouse models deficient for nuclear factor kappa B p50 subunit. *Neurobiol Dis,* 64**,** 16-29.

ROSKOSKI, R., JR. 2012. ERK1/2 MAP kinases: structure, function, and regulation. *Pharmacol Res,* 66**,** 105-43.

SAKURAI, H., CHIBA, H., MIYOSHI, H., SUGITA, T. & TORIUMI, W. 1999. IkappaB kinases phosphorylate NF-kappaB p65 subunit on serine 536 in the transactivation domain. *J Biol Chem,* 274**,** 30353-6.

SALMERON, A., JANZEN, J., SONEJI, Y., BUMP, N., KAMENS, J., ALLEN, H. & LEY, S. C. 2001. Direct phosphorylation of NF-kappaB1 p105 by the IkappaB kinase complex on serine 927 is essential for signal-induced p105 proteolysis. *J Biol Chem,* 276**,** 22215-22.

SATO, S., SANJO, H., TAKEDA, K., NINOMIYA-TSUJI, J., YAMAMOTO, M., KAWAI, T., MATSUMOTO, K., TAKEUCHI, O. & AKIRA, S. 2005. Essential function for the kinase TAK1 in innate and adaptive immune responses. *Nat Immunol,* 6**,** 1087-95.

SAVINOVA, O. V., HOFFMANN, A. & GHOSH, G. 2009. The Nfkb1 and Nfkb2 proteins p105 and p100 function as the core of high-molecular-weight heterogeneous complexes. *Mol Cell,* 34**,** 591-602.

SCHMITT, A. M., CRAWLEY, C. D., KANG, S., RALEIGH, D. R., YU, X., WAHLSTROM, J. S., VOCE, D. J., DARGA, T. E., WEICHSELBAUM, R. R. & YAMINI, B. 2011. p50 (NF-kappaB1) is an effector protein in the cytotoxic response to DNA methylation damage. *Mol Cell,* 44**,** 785-96.

SCHRODER, K., IRVINE, K. M., TAYLOR, M. S., BOKIL, N. J., LE CAO, K.-A., MASTERMAN, K.-A., LABZIN, L. I., SEMPLE, C. A., KAPETANOVIC, R., FAIRBAIRN, L., AKALIN, A., FAULKNER, G. J., BAILLIE, J. K., GONGORA, M., DAUB, C. O., KAWAJI, H., MCLACHLAN, G. J., GOLDMAN, N., GRIMMOND, S. M., CARNINCI, P., SUZUKI, H., HAYASHIZAKI, Y., LENHARD, B., HUME, D. A. & SWEET, M. J. 2012. Conservation and divergence in Toll-like receptor 4-regulated gene expression in primary human versus mouse macrophages. *Proceedings of the National Academy of Sciences,* 109**,** E944.

SCHURCH, N. J., SCHOFIELD, P., GIERLIŃSKI, M., COLE, C., SHERSTNEV, A., SINGH, V., WROBEL, N., GHARBI, K., SIMPSON, G. G., OWEN-HUGHES, T., BLAXTER, M. & BARTON, G. J. 2016. How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *RNA,* 22**,** 839-851.

SCHWABE, R. F. & SAKURAI, H. 2005. IKKbeta phosphorylates p65 at S468 in transactivaton domain 2. *Faseb j,* 19**,** 1758-60.

SEN, R. & BALTIMORE, D. 1986. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. *Cell,* 46**,** 705-16.

SENFTLEBEN, U., CAO, Y., XIAO, G., GRETEN, F. R., KRAHN, G., BONIZZI, G., CHEN, Y., HU, Y., FONG, A., SUN, S. C. & KARIN, M. 2001. Activation by IKKalpha of a second, evolutionary conserved, NF-kappa B signaling pathway. *Science,* 293**,** 1495-9.

SEYEDNASROLLAH, F., LAIHO, A. & ELO, L. L. 2015. Comparison of software packages for detecting differential expression in RNA-seq studies. *Briefings in Bioinformatics,* 16**,** 59-70.

SHA, W. C., LIOU, H. C., TUOMANEN, E. I. & BALTIMORE, D. 1995. Targeted disruption of the p50 subunit of NF-kappa B leads to multifocal defects in immune responses. *Cell,* 80**,** 321-30.

SHARIF, O., BOLSHAKOV, V. N., RAINES, S., NEWHAM, P. & PERKINS, N. D. 2007. Transcriptional profiling of the LPS induced NF-kappaB response in macrophages. *BMC Immunol,* 8**,** 1.

SIGGERS, T., CHANG, A. B., TEIXEIRA, A., WONG, D., WILLIAMS, K. J., AHMED, B., RAGOUSSIS, J., UDALOVA, I. A., SMALE, S. T. & BULYK, M. L. 2011. Principles of dimer-specific gene regulation revealed by a comprehensive characterization of NF-kappaB family DNA binding. *Nat Immunol,* 13**,** 95-102.

SIZEMORE, N., LERNER, N., DOMBROWSKI, N., SAKURAI, H. & STARK, G. R. 2002. Distinct roles of the Ikappa B kinase alpha and beta subunits in liberating nuclear factor kappa B (NF-kappa B) from Ikappa B and in phosphorylating the p65 subunit of NF-kappa B. *J Biol Chem,* 277**,** 3863-9.

SNAPPER, C. M., ZELAZOWSKI, P., ROSAS, F. R., KEHRY, M. R., TIAN, M., BALTIMORE, D. & SHA, W. C. 1996. B cells from p50/NF-kappa B knockout mice have selective defects in proliferation, differentiation, germ-line CH transcription, and Ig class switching. *The Journal of Immunology,* 156**,** 183.

SUBRAMANIAN, A., TAMAYO, P., MOOTHA, V. K., MUKHERJEE, S., EBERT, B. L., GILLETTE, M. A., PAULOVICH, A., POMEROY, S. L., GOLUB, T. R., LANDER, E. S. & MESIROV, J. P. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A,* 102**,** 15545-50.

SUN, X. F. & ZHANG, H. 2007. NFKB and NFKBI polymorphisms in relation to susceptibility of tumour and other diseases. *Histol Histopathol,* 22**,** 1387-98.

TONG, A.-J., LIU, X., THOMAS, B. J., LISSNER, M. M., BAKER, M. R., SENAGOLAGE, M. D., ALLRED, A. L., BARISH, G. D. & SMALE, S. T. 2016. A Stringent Systems Approach Uncovers Gene-Specific Mechanisms Regulating Inflammation. *Cell,* 165**,** 165-179.

TRAPNELL, C., WILLIAMS, B. A., PERTEA, G., MORTAZAVI, A., KWAN, G., VAN BAREN, M. J., SALZBERG, S. L., WOLD, B. J. & PACHTER, L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol,* 28**,** 511-5.

TSUCHIYA, Y., ASANO, T., NAKAYAMA, K., KATO, T., JR., KARIN, M. & KAMATA, H. 2010. Nuclear IKKbeta is an adaptor protein for IkappaBalpha ubiquitination and degradation in UV-induced NF-kappaB activation. *Mol Cell,* 39**,** 570-82.

TU, Y.-C., HUANG, D.-Y., SHIAH, S.-G., WANG, J.-S. & LIN, W.-W. 2014. Regulation of c-Fos Gene Expression by NF-κB: A p65 Homodimer Binding Site in Mouse Embryonic Fibroblasts but Not Human HEK293 Cells. *PLOS ONE,* 8**,** e84062.

TUIJNENBURG, P., LANGO ALLEN, H., BURNS, S. O., GREENE, D., JANSEN, M. H., STAPLES, E., STEPHENS, J., CARSS, K. J., BIASCI, D., BAXENDALE, H., THOMAS, M., CHANDRA, A., KIANI-ALIKHAN, S., LONGHURST, H. J., SENEVIRATNE, S. L., OKSENHENDLER, E., SIMEONI, I., DE BREE, G. J.,

TOOL, A. T. J., VAN LEEUWEN, E. M. M., EBBERINK, E. H. T. M., MEIJER, A. B., TUNA, S., WHITEHORN, D., BROWN, M., TURRO, E., THRASHER, A. J., SMITH, K. G. C., THAVENTHIRAN, J. E., KUIJPERS, T. W., ADHYA, Z., ALACHKAR, H., ANANTHARACHAGAN, A., ANTROBUS, R., ARUMUGAKANI, G., BACCHELLI, C., BAXENDALE, H., BETHUNE, C., BIBI, S., BOARDMAN, B., BOOTH, C., BROWNING, M., BROWNLIE, M., BURNS, S., CHANDRA, A., CLIFFORD, H., COOPER, N., DAVIES, S., DEMPSTER, J., DEVLIN, L., DOFFINGER, R., DREWE, E., EDGAR, D., EGNER, W., EL-SHANAWANY, T., GASPAR, B., GHURYE, R., GILMOUR, K., GODDARD, S., GORDINS, P., GRIGORIADOU, S., HACKETT, S., HAGUE, R., HARPER, L., HAYMAN, G., HERWADKAR, A., HUGHES, S., HUISSOON, A., JOLLES, S., JONES, J., KELLEHER, P., KLEIN, N., KUIJPERS, T., KUMARARATNE, D., LAFFAN, J., ALLEN, H. L., LEAR, S., LONGHURST, H., LORENZO, L., MAIMARIS, J., MANSON, A., MCDERMOTT, E., MILLAR, H., MISTRY, A., MORRISSON, V., MURNG, S., NASIR, I., NEJENTSEV, S., NOORANI, S., OKSENHENDLER, E., PONSFORD, M., QASIM, W., QUINN, E., QUINTI, I., RICHTER, A., SAMARGHITEAN, C., SARGUR, R., SAVIC, S., SENEVIRATNE, S., SEWALL, C., et al. Loss-of-function nuclear factor κB subunit 1 (NFKB1) variants are the most common monogenic cause of common variable immunodeficiency in Europeans. *Journal of Allergy and Clinical Immunology*.

VAN DER POLL, T., JANSEN, P. M., VAN ZEE, K. J., WELBORN, M. B., 3RD, DE JONG, I., HACK, C. E., LOETSCHER, H., LESSLAUER, W., LOWRY, S. F. & MOLDAWER, L. L. 1996. Tumor necrosis factor-alpha induces activation of coagulation and fibrinolysis in baboons through an exclusive effect on the p55 receptor. *Blood,* 88**,** 922-7.

VONDERACH, M., BYRNE, D. P., BARRAN, P. E., EYERS, P. A. & EYERS, C. E. 2018. DNA Binding and Phosphorylation Regulate the Core Structure of the NF-κB p50 Transcription Factor. *Journal of The American Society for Mass Spectrometry***,** 1-11.

WAJANT, H. & SCHEURICH, P. 2011. TNFR1-induced activation of the classical NF-kappaB pathway. *Febs j,* 278**,** 862-76.

WATERFIELD, M. R., ZHANG, M., NORMAN, L. P. & SUN, S. C. 2003. NF-kappaB1/p105 regulates lipopolysaccharide-stimulated MAP kinase signaling by governing the stability and function of the Tpl2 kinase. *Mol Cell,* 11**,** 685-94.

WESSELLS, J., BAER, M., YOUNG, H. A., CLAUDIO, E., BROWN, K., SIEBENLIST, U. & JOHNSON, P. F. 2004. BCL-3 and NF-kappaB p50 attenuate lipopolysaccharide-induced inflammatory responses in macrophages. *J Biol Chem,* 279**,** 49995-50003.

WILSON, C. L., JURK, D., FULLARD, N., BANKS, P., PAGE, A., LULI, S., ELSHARKAWY, A. M., GIELING, R. G., CHAKRABORTY, J. B., FOX, C., RICHARDSON, C., CALLAGHAN, K., BLAIR, G. E., FOX, N., LAGNADO, A., PASSOS, J. F., MOORE, A. J., SMITH, G. R., TINIAKOS, D. G., MANN, J., OAKLEY, F. & MANN, D. A. 2015. NFkappaB1 is a suppressor of neutrophil-driven hepatocellular carcinoma. *Nat Commun,* 6**,** 6818.

WU, C. J., CONZE, D. B., LI, T., SRINIVASULA, S. M. & ASHWELL, J. D. 2006. Sensing of Lys 63-linked polyubiquitination by NEMO is a key event in NF-kappaB activation [corrected]. *Nat Cell Biol,* 8**,** 398-406.

WU, T.-T., CHEN, T.-L. & CHEN, R.-M. 2009. Lipopolysaccharide triggers macrophage activation of inflammatory cytokine expression, chemotaxis, phagocytosis, and oxidative ability via a toll-like receptor 4-dependent

pathway: Validated by RNA interference. *Toxicology Letters,* 191**,** 195-202.

WULCZYN, F. G., NAUMANN, M. & SCHEIDEREIT, C. 1992. Candidate proto-oncogene bcl-3 encodes a subunit-specific inhibitor of transcription factor NF-kappa B. *Nature,* 358**,** 597-9.

WYNN, T. A. & VANNELLA, K. M. 2016. Macrophages in Tissue Repair, Regeneration, and Fibrosis. *Immunity,* 44**,** 450-462.

XUE, J., SCHMIDT, SUSANNE V., SANDER, J., DRAFFEHN, A., KREBS, W., QUESTER, I., DE NARDO, D., GOHEL, TRUPTI D., EMDE, M., SCHMIDLEITHNER, L., GANESAN, H., NINO-CASTRO, A., MALLMANN, MICHAEL R., LABZIN, L., THEIS, H., KRAUT, M., BEYER, M., LATZ, E., FREEMAN, TOM C., ULAS, T. & SCHULTZE, JOACHIM L. 2014. Transcriptome-Based Network Analysis Reveals a Spectrum Model of Human Macrophage Activation. *Immunity,* 40**,** 274-288.

YAMAMOTO, M., YAMAZAKI, S., UEMATSU, S., SATO, S., HEMMI, H., HOSHINO, K., KAISHO, T., KUWATA, H., TAKEUCHI, O., TAKESHIGE, K., SAITOH, T., YAMAOKA, S., YAMAMOTO, N., YAMAMOTO, S., MUTA, T., TAKEDA, K. & AKIRA, S. 2004. Regulation of Toll/IL-1-receptor-mediated gene expression by the inducible nuclear protein IkappaBzeta. *Nature,* 430**,** 218-22.

YAMAUCHI, S., ITO, H. & MIYAJIMA, A. 2010. IκBη, a nuclear IκB protein, positively regulates the NF-κB–mediated expression of proinflammatory cytokines. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 11924-11929.

YAN, Q., CARMODY, R. J., QU, Z., RUAN, Q., JAGER, J., MULLICAN, S. E., LAZAR, M. A. & CHEN, Y. H. 2012. Nuclear factor-κB binding motifs specify Toll-like receptor-induced gene repression through an inducible repressosome. *Proceedings of the National Academy of Sciences,* 109**,** 14140-14145.

YANG, D., SCAVUZZO, M. A., CHMIELOWIEC, J., SHARP, R., BAJIC, A. & BOROWIAK, M. 2016. Enrichment of G2/M cell cycle phase in human pluripotent stem cells enhances HDR-mediated gene repair with customizable endonucleases. *Scientific Reports,* 6**,** 21264.

YANG, H. T., PAPOUTSOPOULOU, S., BELICH, M., BRENDER, C., JANZEN, J., GANTKE, T., HANDLEY, M. & LEY, S. C. 2012. Coordinate regulation of TPL-2 and NF-kappaB signaling in macrophages by NF-kappaB1 p105. *Mol Cell Biol,* 32**,** 3438-51.

YANG, L., ZHOU, T., DROR, I., MATHELIER, A., WASSERMAN, W. W., GORDÂN, R. & ROHS, R. 2014. TFBSshape: a motif database for DNA shape features of transcription factor binding sites. *Nucleic Acids Research,* 42**,** D148-D155.

YOBOUA, F., MARTEL, A., DUVAL, A., MUKAWERA, E. & GRANDVAUX, N. 2010. Respiratory syncytial virus-mediated NF-kappa B p65 phosphorylation at serine 536 is dependent on RIG-I, TRAF6, and IKK beta. *J Virol,* 84**,** 7267-77.

ZABEL, U., SCHRECK, R. & BAEUERLE, P. A. 1991. DNA binding of purified transcription factor NF-kappa B. Affinity, specificity, Zn2+ dependence, and differential half-site recognition. *J Biol Chem,* 266**,** 252-60.

ZERBINO, D. R., WILDER, S. P., JOHNSON, N., JUETTEMANN, T. & FLICEK, P. R. 2015. The Ensembl Regulatory Build. *Genome Biology,* 16**,** 56-64.

ZHONG, H., MAY, M. J., JIMI, E. & GHOSH, S. 2002. The phosphorylation status of nuclear NF-kappa B determines its association with CBP/p300 or HDAC-1. *Mol Cell,* 9**,** 625-36.