THE UNIVERSITY of EDINBURGH

Modelling and Analysis of Macrophage Activation Pathways

Sobia Raza

July 2011

Thesis submitted for the degree of
Doctor of Philosophy

College of Medicine and Veterinary Medicine,
University of Edinburgh

Thesis Word Count: 79,770

Declaration

The thesis presented is the work of the author except where stated otherwise by reference and/or acknowledgement. Any work presented, which has been conducted by (or in collaboration with) others is explicitly acknowledged at the end of each Chapter. No part of this work has been submitted in candidature for any other degree or qualification.

Name: ………………………………….   Date: …………………………………..

Contents

List of Tables and Figures

Figures

Chapter-1

Chapter-2

Chapter-5

Chapter-6

Tables

x

This page has been left intentionally blank

(i)   Abstract

Macrophages are present in virtually all tissues and account for approximately 10% of all body mass. Although classically credited as the scavenger cells of innate immune system, ridding a host of pathogenic material and cellular debris though their phagocytic function, macrophages also play a crucial role in embryogenesis, homeostasis, and inflammation. De-regulation of macrophage function is therefore implicated in the progression of many disease states including cancer, arthritis, and atherosclerosis to name just a few. The diverse range of activities of this cell can be attributed to its exceptional phenotypic plasticity i.e. it is capable of adapting its physiology depending on its environment; for instance in response to different types of pathogens, or specific cocktail of cytokines detected. This plasticity is exemplified by the macrophages capacity to adjust rapidly its transcriptional profile in response to a given stimulus. This includes interferons which are a group of cytokines capable of activating the macrophage by interacting with their cognate receptors on the cell. The different classes of interferons activate downstream signalling cascades, eventually leading to the expression (as well as repression) of hundreds of genes.

To begin to fully understand the properties of a dynamic cell such as the macrophage arguably requires a holistic appreciation of its constituents and their interactions. Systems biology investigations aim to escape from a gene-centric view of biological systems. As such this necessitates the development of better ways to order, display, mine and analyse biological information, from our knowledge of protein interactions and the systems they form, to the output of high throughput technologies. The primary objectives of this research were to further characterise the signalling mechanisms driving macrophages activation, especially in response to type-I and type-II interferons, as well as lipopolysaccharide (LPS), using a 'systems-level' approach to data analysis and modelling. In order to achieve this end I have explored and developed methods for the executing a 'systems-level' analysis. Specifically the questions addressed included: (a) How does one begin to formalise and model the

existing knowledge of signalling pathways in the macrophage? (b) What are the similarities and differences between the macrophage response to different types of interferon (namely interferon-β (IFN-β) and interferon-γ (IFN-γ))? (c) How is the macrophage transcriptome affected by siRNA targeting of key regulators of the interferon pathway? (d) To what extent does a model of macrophage signalling aid interpretation of the data generated from functional genomics screens?

There is general agreement amongst biologists about the need for high-quality pathway diagrams and a method to formalize the way biological pathways are depicted. In an effort to better understand the molecular networks that underpin macrophage activation an *in-silico* model or 'map' of relevant pathways was constructed by extracting information from published literature describing the interactions of individual constituents of this cell and the processes they modulate (Chapter-2). During its construction process many challenges of converting pathway knowledge into computationally-tractable yet 'understandable' diagrams, were to be addressed. The final model comprised 2,170 components connected by 2,553 edges, and is to date the most comprehensive formalised model of macrophage signalling. Nevertheless this still represents just a modest body of knowledge on the cell. Related to the pathway modelling efforts was the need for standardising the graphical depiction of biology in order to achieve these ends. The methods for implementing this and agreeing a 'standard' has been the subject of some debate. Described herein (in Chapter-3) is the development of one graphical notation system for biology the modified Edinburgh Pathway Notation (mEPN). By constructing the model of macrophage signalling it has been possible to test and extensively refine the original notation into an intuitive, yet flexible scheme capable of describing a range of biological concepts. The hope is that the mEPN development work will contribute to the on-going community effort to develop and agree a standard for depicting pathways and the published version will provide a coherent guide to those planning to construct pathway diagrams of their biological systems of interest.

With a desire to better understand the transcriptional response of primary mouse macrophages to interferon stimulation, genome wide expression profiling was performed and an explorative-network based method applied for analysing the data generated (Chapter-4). Although transcriptomics data pertaining to interferon stimulation of macrophages is not entirely novel, the network based analysis of it provided an alternative approach to visualise, mine and interpret the output. The analysis revealed overlap in the transcriptional targets of the two classes of interferon, as well as processes preferentially induced by either cytokine; for example MHC-Class II antigen processing and presentation by IFN-γ, and an anti-proliferative signature by IFN-β. To further investigate the contribution of individual proteins towards generating the type-I (IFN-β) response, short interfering RNA (siRNA) were employed to repress the expression of selected target genes. However in macrophages and other cells equipped with pathogen detection systems the act of siRNA trasfection can itself induce a type-I interferon response. It was therefore necessary to contend with this autocrine production of IFN-β and optimise an *in vitro* assay for studying the contribution of siRNA induced gene-knock downs to the interferon response (described in Chapter-5). The final assay design incorporated LPS stimulation of the macrophages, as a means of inducing IFN-β autonomously of the transfection induced type-I response. However genome-wide expression analysis indicated the targeted gene knock-downs did not perturb the LPS response in macrophages on this occasion. The optimisation process underscored the complexities of performing siRNA gene knockdown studies in primary macrophages. Furthermore a more thorough understanding of the transcriptional response of macrophages to stimulation by interferon or by LPS was required. Therefore the final investigations of this thesis (Chapter-6) explore the transcriptional changes over a 24 hour time-course of macrophage activation by IFN-β, IFN-γ, or LPS and the contribution of the macrophage pathway model in interpreting the response to the three stimuli.

Taken together the work described in this thesis highlight the advances to be made from a systems-based approach to visualisation, modelling and analysis of macrophage signalling.

Dedication

This thesis is dedicated to my wonderful parents Najeeb and Alveena Raza and my (late) Grandmother Umar Bibi, a truly inspirational woman.

(ii)   Acknowledgements

University of Edinburgh

The research presented in this thesis was performed in two different departments as I moved from the Division of Pathway Medicine (University of Edinburgh) to the Roslin Institute (University of Edinburgh) in 2008/2009. I therefore had the great opportunity to make contacts across two departments; meaning more colleagues, more expertise, and of course friends.

Very special thanks are due to my supervisor Prof. Tom Freeman; for always keeping his door open to my questions, for his patience, enthusiasm, insightful guidance with my work, and of course for reading and re-reading drafts of my thesis. Tom also provided me the opportunity to attend and demonstrate at the Wellcome Trust Advanced Course, an invaluable experience which I believe added to my overall doctoral training.

At the DPM I would particularly like to thank Prof. Peter Ghazal for his input during the early stages of my PhD. Gratitude is also due to other colleagues at the DPM, especially Dr. Garwin Sing and Paul Lacaze for their assistance with laboratory techniques and collaborations.

A great deal of appreciation is due to the group members of my wider-lab group at the Roslin Institute (this includes the groups of Prof. David Hume, Prof. Kim Summers, Dr. Geoff Faulkner, and Prof. Tom Freeman). I'm particularly grateful to Dr. Mark Barnett for his teaching of laboratory techniques and constant willingness to help with setting up of large and complex experimental screens, and for reading over the draft of the methods and materials. A great deal of thanks is due to Dr. David Sester, who was an invaluable source of macrophage knowledge in the lab. I appreciate Daves help with teaching of laboratory techniques, willingness to look at my cells, (and for introducing me to the works of George Carlin). In addition to Mark, and Dave I thank other lab colleagues (Valerie, Debby, Julie, Lynsey, Kenny, Ronan, Dario, Andru, Kenny, Geoff,

Ailsa) for both creating a great atmosphere in which to work, and answering my questions.

In addition to the wet-lab work, a significant portion of this thesis was spent performing computational analysis. This was facilitated by the development of supporting software by Athanasios Theocharidis. I'm also very grateful to Thanasis for his assistance with maintaining my computer hardware. Thanks are also due to Dr. Cristelle Roberts who has made available some of the work presented herein on the macrophages.com website. Thank you to Norrie Russell for assistance with the initial thesis binding, a challenging part of completing a thesis.

A great deal of gratitude is due to Prof. David Hume for sharing his insight, providing constructive input into the work presented, for reading a draft of this thesis and providing a constructive critique of it. Thanks also to David for nurturing what I have found to be a very stimulating and positive working environment, at the Roslin Institute. All other acknowledgements and specific contributions to the work presented in this thesis are outlined at the end of each Chapter.

Funding Sources and Others

My PhD stipend was awarded by the Biotechnology and Biological Sciences Research Council (BBSRC). The research presented here has been made possible with the financial support of a number of bodies; the BBSRC, INFOBIOMED EU FP6 programme, and the Wellcome Trust. The BBSRC also awarded me a three-month 'science-policy' placement, at the Scottish Parliament. The experience was a welcome addition to my overall doctoral training; special thanks are due to members of the Health, Environment and Europe research unit, particularly Tom Edwards and Graeme Cooke.

Personal

I am indebted to my incredible parents Najeeb and Alveena Raza for their constant support, encouragement, and endless patience. Thanks also to my younger sisters Yasrab, and Iman, and of course the feline family member Barney the cat. Edinburgh

has been a fantastic city to conduct a PhD, and I've had the opportunity to make wonderful friends from (nearly) all the different continents. Thanks to you all for your company and special thanks to those who have been supportive during last few months; Manuel, Camillo, and co.

(iii)   Abbreviations

2'5'OAS	2'-5'-Oligoadenylate Synthetases
API	Application Programming Interface
BMDM(s)	Bone Marrow Derived Macrophage(s)
CAGE	Cap Analysis Gene Expression
CSBE	Centre for Systems Biology Edinburgh
CSF-1	(Macrophage) Colony Stimulating Factor -1
DAMPs	Danger-associated Molecular Patterns
DAVID	Database for Annotation, Visualization and Integrated Discovery
DCs	Dendritic Cells
DPM	Division of Pathway Medicine
dsRNA	double stranded RNA
EGFP	Enhanced Green Fluorescent Protein
EMCV	Encephalomyocarditis Virus
EPE	Edinburgh Pathway Editor
EPN	Edinburgh Pathway Notation
FBS	Foetal Bovine Serum
GAS	IFN-γ activated sites
G-CFU	Granulocyte Colony Forming Unit
GEO	Gene Expression Omnibus
GM-CFU	Granulocytemacrophage Colony-forming-unit
GO	Gene Ontology
GOI	Gene(s) of Interest
GUI	Graphical User Interface
H / h	Hours
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
HGNC	Human Gene Nomenclature
HSC	Hematopoietic Stem Cells
IFN	Interferon

IFNB / IFN-β	Interferon Beta
IFNG / IFN-γ	Interferon Gamma
IRF	Interferon Regulatory Factor
ISG	Interferon Stimulated Genes
ISGF3	ISG Factor 3
ISRE	IFN-stimulated Response Elements
LBP	LPS Binding Protein
LPS	Lipopolysaccharide
LTA	Lipteichoic Acid
MCL	Markov Clustering
M-CFU	Macrophage Colony-forming Unit
M-CSF	Macrophage Colony Stimulating Factor
mEPN	modified Edinburgh Pathway Notation
MIAME	Minimum Information About a Microarray Experiment
MIM	Molecular Interaction Map
MPS	Mononuclear Phagocytic System
MS	Multiple Sclerosis
NLRs	NOD-like Receptors
NO	Nitric Oxide
NOD	Nucleotide-binding Oligomerization Domain
NT-siRNA	Non-Targeting siRNA
ODEs	Ordinary Differential Equations
PAMPs	Pathogen Associated Molecular Patterns
PBMC	Peripheral Blood Mononuclear Cells
PCR	Polymerase Chain Reaction
PDN	Process Description Notation
PRKR	Protein Kinase RNA-regulated
PRR	Pattern Recognition Receptors
RNAi	RNA-interference
SBGN	Systems Biology Graphical Notation

siRNA	short interfering RNA
ssRNA	single stranded RNA
shRNA	short hairpin RNA
SLE	Systemic Lupus Erythematosus
SNP	Single Nucleotide Polymorphisms
SOCS	Suppressor Of Cytokine Signalling
SPN	Signalling Petri Net
SUMO	Small Ubiquitin-like Modifier
TAMs	Tumour Associated Macrophages
TF	Transcription Factors
TLR	Toll-Like-Receptor
TYK2	Tyrosine Kinase 2
WTAC	Wellcome Trust Advanced Course

(iv)   Publications arising from this work

Articles in peer reviewed journals

- Raza S, McDerment N, Lacaze PA, Robertson K, Watterson S, Chen Y, Chisholm M, Eleftheriadis G, Monk S, O'Sullivan M, et al: **Construction of a large scale integrated map of macrophage pathogen recognition and effector systems.** *BMC Syst Biol* 2010, **4**:63.

- Freeman TC, Raza S, Theocharidis A, Ghazal P: **The mEPN scheme: an intuitive and flexible graphical system for rendering biological pathways.** *BMC Syst Biol* 2010, **4**:65.

- Hume DA, Summers KM, Raza S, Baillie JK, Freeman TC: **Functional clustering and lineage markers: insights into cellular differentiation and gene function from large-scale microarray studies of purified primary cell populations.** *Genomics* 2010, **95**:328-338.

- Summers KM, Raza S, van Nimwegen E, Freeman TC, Hume DA: **Co-expression of FBN1 with mesenchyme-specific genes in mouse cell lines: implications for phenotypic variability in Marfan syndrome.** *Eur J Hum Genet* 2010, **18**:1209-1215.

- Lacaze P, Raza S, Sing G, Page D, Forster T, Storm P, Craigon M, Awad T, Ghazal P, Freeman TC: **Combined genome-wide expression profiling and targeted RNA interference in primary mouse macrophages reveals perturbation of transcriptional networks associated with interferon signalling.** *BMC Genomics* 2009, **10**:372.

- Raza S, Robertson KA, Lacaze PA, Page D, Enright AJ, Ghazal P, Freeman TC: **A logic-based diagram of signalling pathways central to macrophage activation.** *BMC Syst Biol* 2008, **2**:36.

- Planas-Iglesias J, Guney E, Garcia-Garcia J, Robertson KA, Raza S, Freeman TC, Ghazal P, Oliva: **Extending signalling pathways with protein-interaction networks. Application to apoptosis.** BMC *Syst Biol* 2011, Submitted.

Abstracts

Oral Presentations

- **"A Logic Based Diagram of Signalling Central to Macrophage Activation".** 3rd annual conference for Synthetic Biology, Systems Biology and Bioinformatics. Held at Imperial College, London. April 2008.

- **"Modelling and Analysis of Macrophage Signalling Activity".** Poster and presentation to HRH the Princess Royal, Princess Anne. Held at The Roslin Institute, University of Edinburgh. November 2009.

Posters

- Raza S, Theocharidis A, Ghazal P, Freeman TC. **The Modified Edinburgh Pathway Notation (mEPN) Scheme: An Intuitive and Flexible Graphical System for Drawing Pathways.** Meeting in Functional Genomics and Systems Biology. Held at The Welcome Trust Genome Campus, Hinxton, Cambridgeshire. November 2009.

- Raza S, McDerment N, Lacaze P, Robertson KA, Watterson S, Ying C, Chisholm M, Eleftheriadis G, Monk S, O'Sullivan M, Turnbull A, Roy D, Theocharidis A, Ghazal P, Freeman TC. **Construction, Visualisation and Application of Logic-Based Process Diagrams of Macrophage Activation and Inflammatory Pathways.** Inflammation in Bacterial and Viral Infection: 2nd Annual symposium. Held at the Chancellors Building, University of Edinburgh. September 2009.

- Raza S, McDerment N, Lacaze P, Robertson KA, Watterson S, Ying C, Chisholm M, Eleftheriadis G, Monk S, O'Sullivan M, Turnbull A, Roy D, Theocharidis A, Ghazal P, Freeman TC. **Construction, Visualisation and Application of Logic-Based Process Diagrams of Macrophage Activation and Inflammatory Pathways.** 2nd Conference in Innate Immunity and the Pathogenesis of Rheumatic Diseases. Genoa, Italy. May 2009.

- Freeman TC, McDerment N, Raza S, Chisholm M, Ying C, Monk S, Turnbull A, O'Sullivan M, Eleftheriadis G, Theocharidis A, Robertson KA, Lacaze P, Ghazal P. **Construction, Visualisation and Application of Logic-Based Process Diagrams of Macrophage Pathways.** 4th EMBO Conference: From Functional Genomics to Systems Biology. Held at the European Molecular Biology Laboratory (EMBL) Heidelberg. November 2008.

- Raza S, Robertson KA, Lacaze PA, Page D, Enright AJ, Ghazal P, Freeman TC: **A Logic Based Diagram of Signalling Central to Macrophage Activation**. Centre

for Infectious Disease (CID) annual symposium. Held at The CID, Queens Medical Research Institute, University of Edinburgh. May 2008.

- Raza S, Robertson KA, Lacaze PA, Page D, Enright AJ, Ghazal P, Freeman TC: **A Logic Based Diagram of Signalling Central to Macrophage Activation**. 3rd Annual Conference in Pathway Medicine. Held at The Division of Pathway Medicine, University of Edinburgh. November 2007.

- Raza S, Robertson KA, Lacaze PA, Page D, Enright AJ, Ghazal P, Freeman TC: **A Logic Based Diagram of Signalling Central to Macrophage Activation**. Meeting in Functional Genomics and Systems Biology. Held at The Welcome Trust Genome Campus, Hinxton, Cambridgeshire. October 2007.

Chapter 1. Introduction

Mechnikov discovered phagocytosis after experimenting on the larvae of starfish. His theory was that certain white blood cells now called phagocytes (from the Greek *phagos* – to eat, *cyte* – cell) could engulf and destroy harmful bodies such as bacteria. In a pioneering experiment he observed these cells surrounding and attempting to devour a splinter he had introduced into the transparent body of a starfish larva. He proposed the role of these cells was to maintain integrity of the organism by protecting the animal from foreign invaders or clearing the body of unwanted cellular debris [1-2]. In vertebrates these phagocytes were analogous certain white blood cells. In drawing parallels between the phagocytes role in various species or settings, other important functions of these cells became apparent. In the tail of the tadpole muscle cells were 'eaten' at appropriate times of metamorphosis by the adjacent cells. Thus under certain developmental conditions, it appeared this cell was 'responsible' for defining organismal structures [2].
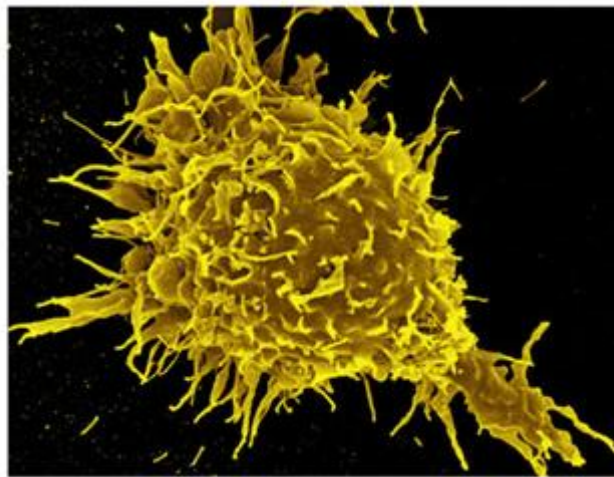
Figure 1.1: Scanning electron micrograph (30 μm × 25 μm) of a phagocytic macrophage; Taken from Rosenberger and Finley, 2003 [3].

The pioneering studies of Mechnikov earned him a Nobel prize in 1908. Since then a century of research into phagocytes, has shed light on their role in immunology (host defence), and other functions beyond simply *eating* [2], such as growth and

development. This progress has seen the characterisation of different phagocyte populations and their precursor cells; the key mammalian phogocytes being macrophages, dendritic cells, neutrophils and mast cells. This thesis explores the signalling events occurring in the macrophage.

The Macrophage

Macrophages are a heterogeneous population of antigen-presenting cells, varying in their anatomical location, phenotype, morphology, and specialised physiological function. The original definition and term *"macro-phage"* from the Greek "big-eater", is derived from the prodigious phagocytic properties of the cell. As the elegant studies of Mechnikov demonstrated, these phagocytic actions are not only required for the propagation of an innate immune response i.e. ridding a host of pathogenic material and cellular debris, but are also crucial for non-immunological trophic roles during development and homeostasis. Macrophages constitute ~10-15% of cells in most tissues and are found in every organ where they have a specialised function and are also recruited to sites of infection, injury, and inflammation [4-9].

Origin of Macrophages and the Mononuclear Phagocyte System

In the traditional view, macrophages are derived from pluripotent stem cells in the bone marrow which can develop into the macrophage precursors known as monocytes. Monocytes then enter the blood stream and under appropriate signals differentiate into macrophages. The mononuclear phagocytic system (MPS) is a classification scheme for defining macrophages, based on their bone marrow and monocytic origins. The MPS therefore comprises bone marrow progenitors, blood monocytes and resident tissue macrophages. In recent years the concept of the MPS and specialised cell lineages is being challenged in light of several developments including the transdifferentiation of MPS cells into other MPS populations, the existence of a separate embryonic phagocyte lineage and the local renewal of tissue macrophages as opposed to monocyte recruitment [10].

Macrophages originate from hematopoietic stem cells (HSC) found in both adult bone marrow and the developing foetus. However macrophage precursors found in the yolk sac may have distinct origins from the macrophage precursors found in adult bone marrow and in the foetal liver following the full initiation of hematopoiesis [11-12]. In the yolk sac and early hepatic hematopoiesis, primitive macrophages are thought to develop from their macrophage precursors without undergoing stages of monocytic cell development (as occurs in adult hematopoiesis) [13]. These primitive-foetal macrophages have the potential to proliferate and differentiate into resident macrophages in tissues in late ontogeny [13]. Later on in foetal hematopoiesis and in adult bone marrow, it is the progression through monocytic cell stages that eventually gives rise to monocyte-derived macrophages. Monocytes in the bone-marrow originate from a common myeloid progenitor. The hematopoetic stages in macrophage development are summarised in Figure 1.2 and include the development of the HSC into a progenitor of both macrophages and granulocytes i.e. the granulocytemacrophage colony-forming-unit (GM-CFU). The GM-CFU population can then commit to the macrophage colony-forming unit (M-CFU), or the granulocyte colony forming unit (G-CFU) group of cells. The M-CFU differentiate into monoblasts, pro-monocytes and monocyte cell stages, prior to becoming macrophages, in a process requiring the growth factor CSF-1 (also known as Macrophage Colony Stimulating Factor (M-CSF)). Monocytes migrate from the bone marrow into peripheral blood and then into tissues. Here they are thought to differentiate into resident macrophages or related cells (dendritic cells or osteoclasts) under the influence of appropriate growth factors. Other monocytes may migrate into tissues in response to infection/ inflammatory stimuli and differentiate into exudate macrophages. However it is now also appreciated that resident tissue macrophages possess proliferative capacity and can be replenished by self-renewal i.e. autonomously of bone-marrow recruited monocytes [10]. Moreover the local proliferation of macrophages, (as opposed to recruitment from peripheral blood) has also been demonstrated under certain inflammatory pathologies [14]. In general the developmental origins of macrophages, process of self-renewal, and function of tissue macrophage subsets are poorly understood [15].
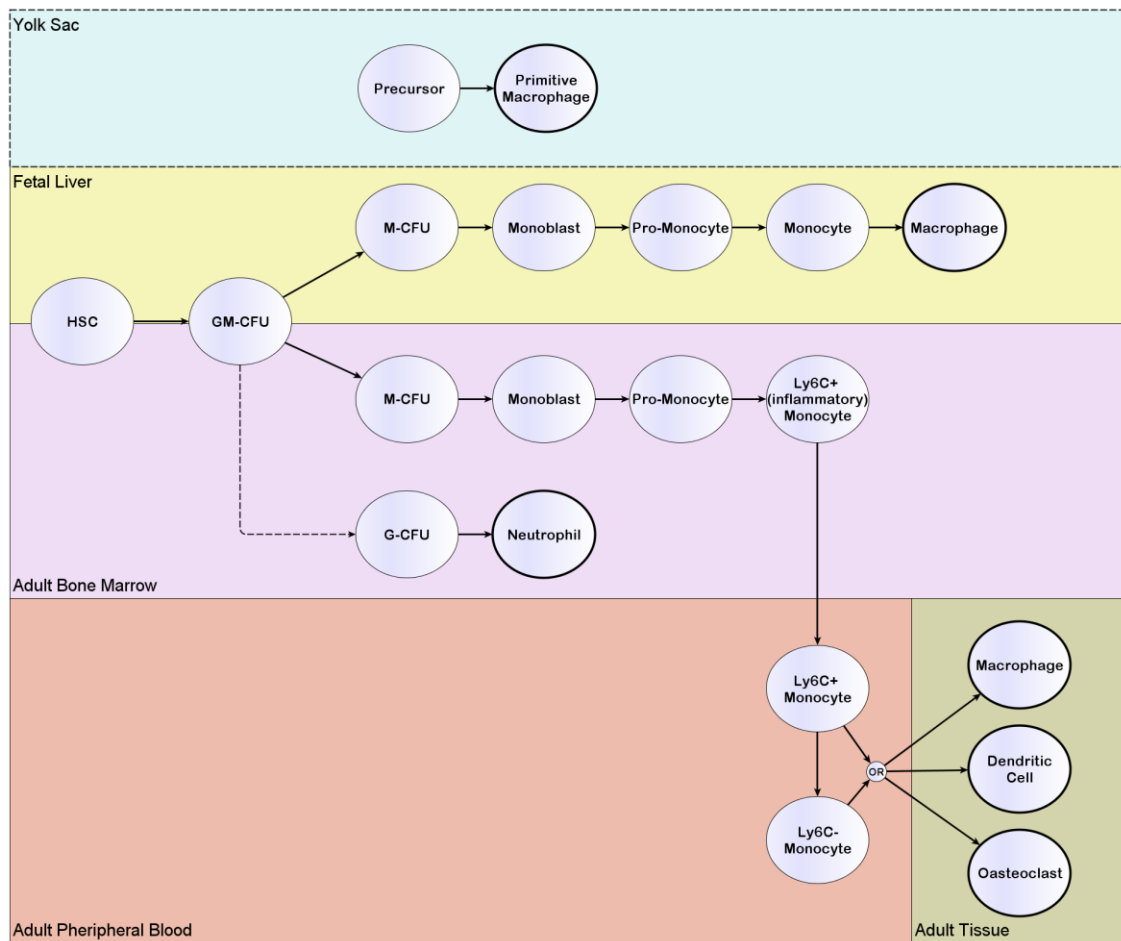
Figure 1.2: Mononuclear Phagocyte System (MPS); Taken and adapted from Gordon and Taylor, 2005 [11]. Hematopoietic stem cells (HSC) in the fetal liver or adult bone marrow develop into a progenitor of both macrophages and granulocytes; the granulocytemacrophage colony-forming-unit (GM-CFU). The GM-CFU population can commit to the macrophage colony-forming unit (M-CFU), or the granulocyte colony forming unit (G-CFU) group of cells. The M-CFU differentiate into monoblasts, pro-monocytes and monocyte cell stages, prior to becoming macrophages, in a process requiring the growth factor CSF-1. In mice Ly6C is a marker of the 'inflammatory' population of monocytes. The concept of the MPS has come under scrutiny following the discovery of a separate embryonic phagocyte lineage in the yolk sac.

The G-CFU population develop into neutrophilic granulocytes. Monocytic macrophages and neutrophilic granulocytes have therefore been traditionally viewed as distinct lineages. However there is a growing body of evidence suggesting a more interchangeable relationship, with the proposal that macrophages and granulocytes may interconvert [10]. This is based on the observations of highly overlapping transcriptional profiles of macrophages and granulocytes as well as the *in vitro* manipulation of granulocytes with various stimuli which induces them to adopt a macrophage-like phenotype [10, 16-17]. Macrophages and dendritic cells are also highly related cell populations; in fact the concept of distinct macrophage and dendritic cell lineages in peripheral tissue is currently being fervently debated [15, 18-20]. Dendritic cells (DCs) are generally considered the key initiators of the adaptive immune response, and are defined by their ability to activate naïve T-cells and the expression of particular markers, namely the integrin CD11c. Discrimination between macrophage and DC populations is blurred by a range of factors [10] e.g. the macrophage differentiation factor CSF-1 has also been shown to influence DC numbers (which express the CSF-1R) [21]; purified CD11c- negative macrophages have been shown to prime naïve T-cells *in vivo* [22]; the preferred DC marker CD11c is also expressed in some macrophage populations e.g. alveolar macrophages. In fact there seems to be no single unambiguous morphological or protein marker of macrophages or DCs, since conventional markers (such as F4/80, CD11c, CD11b and MHC class II) have turned out not to be specific [20]. As a result some argue that it is not possible to define macrophages and DCs as separate entities [10, 19-20]. Instead they are a "*continuum of progeny of a common precursor*" [20]. Whereas others suggest that there is ample evidence correlating the distinct functions of macrophages and DCs with different phenotypic markers [20]. More collaborations and open dialogue between 'macrophage' and 'DC' biologists is advocated to being to unravel the complexities in defining these cells [20].

Macrophage Growth Factors

Several growth factors are involved in the development of macrophages from monocyte populations. Perhaps the most crucial is CSF-1, also called M-CSF. CSF-1 regulates the survival, proliferation and differentiation of macrophages and their precursors. CSF-1 can also synergise with other factors including CSF-2 (GM-CSF) or IL-3 to mediate the proliferation of early haematopoietic progenitors. *In vitro* macrophage differentiation has also been demonstrated in the presence of CSF-2 and IL-3 in combination with CSF-1 [23-24]. The most commonly utilized method of differentiating macrophages *in vitro* is to culture progenitor cells obtained from bone marrow or blood in the presence of CSF-1 or in conditioned medium (containing CSF-1 secreted by a cell, such as from the L929 murine fibroblast cell line). Monocytes and macrophages can also be readily obtained from the peritoneal cavity or lungs.

The CSF-1 receptor, CSF1R, is expressed by all cells of the MPS. The transcription factor PU.1 is involved in the regulation of CSF1R expression. PU.1, CSF-1 or CSF1R deficiency are all associated with reduced macrophage numbers [25]. Mice with a targeted null mutation in Csf1r (Csf1r-/- mice) have a reduced life span (<5 weeks). Mice with a naturally occurring mutation in the *Csf1* gene (*Csf1$^{op/op}$*), known as osteopetrotic mice, display a range of developmental abnormalities attributed to reduced macrophage populations [25-26]. In particular, a lack of osteoclasts (bone remodelling macrophages) in Csf1$^{op/op}$ mice results in osteopetrosis. Csf1$^{op/op}$ rodents are commonly referred to as "toothless", due to lack of teeth. These phenotypic defects can however be reversed by systemic CSF1 administration [27]. Although CSF1R is likely to be the only receptor for CSF-1, CSF1R also binds IL-34 [28] as well as CSF-1; which may explain phenotypic differences in Csf1r-/- and Csf1op/op mice [26].

Monocyte Populations and Markers of Monocytes and Macrophages

Defining a macrophage has been a constant challenge in developmental biology [25]. The complexity of the macrophage is epitomised by the fact no single marker can define all populations. Similarly the progenitor cells giving rise to macrophages are exceptionally heterogeneous. Many efforts have been made to define subpopulations of monocytes based on the expression of surface markers. It is thought the developmental fates and phenotypic properties of monocyte subpopulations are defined by the pattern and extent of marker expression. Determining the functional roles of monocyte subsets in a physiological context remains a challenge [11].

Human monocyte populations are generally identified and defined by the expression of CD14, CD16 and CD64. Incidentally CD14 forms part of the receptor complex for lipopolysaccharide (LPS). $CD14^{hi}CD16^{-}$ monocytes are generally considered to be 'classic monocytes', since they resemble the original description of monocytes. The $CD14^{hi}CD16^{-}$ sub-population are highly phagocytic, produce significant amounts of cytokines and express CCR2 [11]. $CD14^{+}CD16^{+}$ monocytes tend to express higher levels of MHC class II molecules as well as the chemokine receptor CCR5 and are thought to be likely precursors for DCs. Other monocyte subsets are defined using CD64 (FcγRI); $CD14^{+}CD16^{+}CD64^{+}$ population are similar in their characteristics to $CD14^{hi}CD16^{-}$ monocytes, but distinct from $CD14^{+}CD16^{+}CD64^{-}$ population [11].

Murine monocytes can be identified by their expression of F4/80 and CD11b, and further subdivided based on the expression of CCR2, CD62L and CX_3CR1. Ly6C (or GR-1), a granulocyte surface antigen is also a marker of $CCR2^{+}$ monocytes and is widely used as marker for monocyte subsets in mice [11, 29]. Mouse monocytes that express CCR2 (Ly6C+) are considered to be the inflammatory subset, which differentiate into macrophages required for pathogen clearance and resolution of inflammation. CX_3CR1^{hi}/Ly6C- monocyte subsets are thought to differentiate into tissue resident macrophage and DC populations. Certain murine monocyte subsets correspond to human monocyte subsets ($CCR2^{+}CD62L^{+}CX3CR1^{low}Ly6C^{+}$ mouse monocytes are thought

to correspond to the classic human monocyte subset; whereas CCR2$^-$CD62L$^-$ CX3CR1^{hi}Ly6C$^-$ correspond to the human CD14$^+$CD16$^+$CD64$^-$ subset).

Macrophage Diversity and Biological Functions

The use of antigen markers (F4/80 in mouse and CD68 in mouse and human), has permitted the identification of the macrophages in every organ [30]. Development of mice expressing an enhanced green fluorescent protein (EGFP) linked to a promoter region of CSF1R has further permitted the identification and study of macrophage /macrophage-like populations in mouse tissues [4]. These animals are dubbed "macgreen mice" and images of fluorescent cell populations across various tissues are available on www.macrophages.com. Essentially monocytes adapt to their micro-environment and develop into the unique categories of macrophages found throughout the body. These resident tissue macrophages are distributed during development and throughout life. Resident macrophages are often stationed strategically, permitting their sentinel function, but also play a role in homeostasis, clearance of senescent cells, initiation of acute inflammation, remodelling and repair following inflammation and vascular changes [11]. This specialised function of macrophages is largely determined by their anatomical location and macrophages in these different setting often coined with specific names (Table 1.1).

Tissue	Macrophage name	Function
Bone	Osteoclast	Bone remodelling and providing a stem cell niche
	Bone marrow Macrophage	Erythropoiesis
Brain	Microglial cell	Neuronal survival and connectivity, and repair after injury
Epidermis	Langerhans cell	Immune surveillance
Eye	Macrophage	Vascular remodelling
Intestine	Crypt macrophage	Immune surveillance
Kidney	NA	Ductal development
Liver	Kupffer cell	Clearance of debris from blood and liver tissue regeneration after damage; liver development?
Lung	Alveolar Macrophage	Immune surveillance
Mammary Gland	Macrophage	Branching morphogenesis and ductal Development
Ovary	Macrophage	Steroid hormone production and ovulation
Pancreas	Macrophage	Islet development
Testis	Macrophage	Steroid hormone production; Leydig-cell development?
Uterus	Uterine DC	Angiogenesis and decidualization
	Uterine macrophage	Cervical ripening

Table 1.1: Macrophage diversity across tissues. Taken and adapted from Pollard, 2009 [25]. Macrophages are found in different tissues where they have specialised functions.

10

Each organ requires a specialised immune response shaped by the organs requirement for absolute or relative sterility [31]. For example the spleens response to bacteria is in contrast to the gut or colon response to bacteria [31]. Macrophages in the gut are highly phagocytic and bactericidal but produce relatively low levels of pro-inflammatory cytokines [32]. This phenotype is critical in maintaining the balance between the response against harmful pathogens and the induction of tolerance to commensal bacteria. Thus the tissue micro-environment is significant in defining the phenotypic properties of the macrophage. In addition to the organ specific responses, macrophages must elicit specialised response to different types of pathogens. Each scenario specific response witnesses an adaptation of the gene expression and secretory protein profiles of the macrophages. In essence exceptional plasticity and heterogeneity are hallmarks of macrophages.

Macrophages in Disease Pathologies

Given their heterogeneity and presence in every tissue, it is no surprise that macrophages play a role in almost all disease pathologies. The influence of macrophage signalling in a given disease can be both beneficial and detrimental to the host. Deregulation of macrophage signalling is seen to contribute to a number of chronic diseases e.g. cancer where macrophages recruited to the tumour micro-environment are known as "tumour associated macrophages" (TAMs). The role of TAMs in cancer in general is ambiguous and possibly cancer specific. Beneficial roles in the context of the host include detection, rejection and killing of cancer cells as well as clearance of apoptotic cells. At the same time TAMs are known to promote tumour progression and malignancy, for example by promoting angiogenesis [25, 33-35]. In fact, TAM density in human tumours correlates with poor prognosis in over 80% of cases [36].

Macrophages have also been attributed with the progression of acute and chronic symptoms of rheumatoid arthritis. These cells are activated and numerous in the inflamed synovial membrane (joint lining) [37-38] and there is a correlation between progression of joint destruction and extent of synovial macrophage infiltration [38]. Other examples of pathologies where macrophages are implicated include; atherosclerosis (where fat-laden macrophages, known as foam cells contribute to vascular occlusion); emphysema (which is associated with the uncontrolled activation of alveolar macrophages); and septic shock (where an over-zealous cytokine response is propagated by macrophages).

Development of targeted therapeutics requires an understanding of the signalling pathways underpinning macrophage activity and their deregulation in a given disease pathology. Although macrophages have been studied extensively *in vitro*, the challenge remains in translating this research into an *in vivo* understanding.

Macrophage Activation Pathways

Macrophage activation has been defined as the "*acquisition of competence to execute a complex function*" [39]. Gene Ontology (GO) describes the term macrophage activation as "*a change in morphology and behaviour of a macrophage resulting from exposure to a cytokine, chemokine, cellular ligand, or soluble factor*". Ultimately macrophage activation arises as results of the cells interactions with its surrounding environment. These environmental cues are detected by macrophage receptors.

Macrophages express an extensive repertoire of receptors which coupled with downstream signalling cascades mediate their activation and capacity to execute any number of diverse functions. The receptors may be expressed on the surfaces, in the cytosol, as well as vacuolar compartments e.g. endosomes. To avert unnecessary activation, inflammation, and/or damage to tissues, macrophages must distinguish between self and non-self. The detection of microbes or patterns associated with microbes is largely performed by pattern recognition receptors (PRRs). PRR categories include scavengers receptors, C-type lectins, Toll-like receptors, NOD-like receptors, RIG-like receptors, and others not strictly falling into those categories e.g. CD14 and AIM2 [40]. Between them the range of PRRs detect bacteria, virus, protozoa, fungi and their components (examples include; double and single stranded RNA, DNA, CpG DNA, flagellin, lipoprotein, envelope proteins, LPS, and lipteichoic acid (LTA)).

Macrophages also respond to endogenous stimuli e.g. cytokines, chemokines and growth factors, generated following infection/injury and homeostatic possesses. These stimuli are often produced by cells of innate immune system including the macrophages themselves and are crucial in defining macrophage activity. One group of cytokines central to the investigations in this thesis are known as "interferons", originally coined so due to their ability to *interfere* with viral replication.

Receptors for microbial or endogenous stimuli are coupled to downstream intracellular signal transduction pathways. Signalling pathways can broadly be defined as the series of interactions between cellular components, usually proteins and proteins complexes, which lead to the modulation of a given process. Some receptor initiated pathways are linked to specific downstream response pathways and others converge at common factors. The interactions underpinning some of the key signalling pathways activated in macrophages are discussed in greater detail in Chapter-2. The functional programmes regulated by the vast array of signalling pathways include (but are not limited to) growth, survival, apoptosis, migration, phagocytosis, antigen-presentation, remodelling, metabolic reprogramming and cytotoxicity [41]. The signalling pathways mediating these processes often do so by regulating the expression of specific cohorts of genes. In fact macrophage interaction with any given stimulus or combination of stimuli elicits a tailored transcriptional response. The customised transcriptional responses are necessary to provide functional specificity to a given response. For example different pathogens present different challenges for the host; therefore the macrophage response must be adaptable to deal with the specific challenges. LPS, an outer membrane component of gram negative bacteria, is commonly studied transcriptional activator of macrophages. LPS induces a complex transcriptional response, comprising multiple gene sets that encode a number of functional programmes [42]. The gene sets are often co-ordinately regulated by specific transcription factors [42].

Macrophage Activation States

Activated macrophages were originally defined as cells that secreted inflammatory mediators and killed intracellular pathogens. It is now appreciated that macrophages are a far more heterogeneous group of cells and their 'activated' status may in fact refer to a number of phenotypes required for performing distinct immunological functions. The terms M1 and M2 polarised macrophages were introduced to reflect the two extremes of the activated state which are analogous to T-helper cell polarisation (T_H1-T_H2) [43]. M1 polarised macrophages, otherwise known as "classically

activated" macrophages were originally described as being induced by the T_H1 cytokine IFN-γ in concert with microbial stimuli (e.g. LPS) or other cytokines (e.g. TNF-α, GM-CSF) [44]. Cytokine production characteristic of M1 activated macrophages includes high expression of the interleukins IL12, IL23, the T_H1 cell attracting chemokines CXCL9, CXCL10 and nitric oxide [45-46]. In contrast the "alternatively activated" M2 polarised phenotype is induced by the T_H2 cytokine IL-4 as well as IL-13 [47-49]. IL10, IL-1RA, CCL17, CCL22 and CCL24 are examples of cytokines preferentially induced in M2 polarised cells [46, 50]. M2 macrophages have been further classified into "M2-like" categories, which overlap with some but not all significant features of M2 macrophages (reviewed in [51]). For example in CMV infection of monocytes, the cells are biased towards a M1 phenotype but express what are described as typical M2 cytokines (Il1RA, IL10, CCL18), [52]. Crude categorisation of the physiological roles of the polarised macrophages places the M1 subtype as promoters of the T_H1 response, antigen presentation and as cells with intracellular microbiocidal as well as tumoricidal capacity [53]. In contrast alternatively activated macrophages are regarded as immunosuppressive cells, which promote the T_H2 response, tissue remodelling and parasite encapsulation and clearance [49]. Crucially the M1-M2 categories are not the only forms of macrophage polarisation but instead represent extremes across a spectrum of activation states. Even the advocates of M1-M2 polarisation (theory) of macrophage function emphasise this should be viewed as "an operationally useful, simplified, conceptual framework" for describing a continuum of diverse functional states [46]. Others have proposed grouping of macrophage populations based on their different homeostatic activities; host defence, wound healing and immune regulation [23].

There is limited understanding of the differential transcriptional cascades and secretory responses induced in macrophages following interaction with different stimuli (cells, cytokines, microbes). Therefore classifying macrophages based on their activation states remains a challenge. Moreover, the mechanism of transcriptional control and regulatory events that mediate the expression of functional gene sets is not well characterised. Ultimately a better understanding of gene expression patterns

in different macrophage populations may aid the development of therapeutics targeting this cell (e.g. gene therapy and anti-inflammatory drugs).

Interferon Signalling System

Biological Functions of Interferon Signalling

Interferons are a family of multifunctional cytokines that can modulate the transcription of subsets of genes in the target cells they stimulate. They are induced transiently *in vivo* and *in vitro* by viruses, microbial products, or other chemical and/or synthetic inducers. The interferon (IFN) family of cytokines are acknowledged as fundamental components of the innate immune system. The original definition of interferons was based on their ability to "interfere" with viral replication, although they are now also known to have anti-microbial, anti-proliferative and immunomodulatory effects. Furthermore, interferons are also being investigated for their role in the immuno-surveillance for malignant cells. Accordingly IFNs and components of the IFN system are exploited clinically for different therapeutic indications [54]. Recombinant interferon is used to treat hairy cell and chronic myelogenous leukaemias, and has shown to be effective in reducing tumour cell mass and/or malignancy of several other cancer types [54]. Interferons are also used to limit viral replication and control hepatitis B virus (HBV), hepatitis C virus (HCV), and other chronic viral infections (herpes zoster, HSV and cytomegalovirus infections) [54]. One type of IFN (known as IFN-β) is effective at reducing episodes of relapsing–remitting Multiple sclerosis (MS), an inflammatory disorder of the central nervous system which results in demyelination of axons. The therapeutic use of IFN in inflammatory disorders is somewhat a paradox given IFNs are immunostimulatory cytokines. However, viral infection is postulated to be a likely contributory factor in the pathogenesis of MS and early studies did indicate that MS patients secrete less IFN than controls following viral infection [55]. Defects in IFN production also contribute to the aetiology of systemic lupus erythematosus (SLE), a systemic autoimmune disease. Unabated production of IFN is common in SLE patients as is the up-regulation of interferon stimulated genes in

the peripheral blood mononuclear cells (PBMCs) of the patients [56-57]. The cellular actions of interferons are mediated via their regulation of subsets of genes (known as interferon stimulated genes (ISGs)). Different classes of IFNs have largely overlapping gene targets, but they also induce distinct sets of ISGs. The transcriptional responses induced by two classes of interferons (type-I and type-II) form the basis of the investigations in Chapters-4 and -6.

Classification of Interferons

There have been three types of interferons identified, classified according to receptor specificity and sequence homology. Type-I and type-II are by far the best characterised and studied interferons to date. The most recent class of interferon like molecules are the IFN-λ molecules, IFN-λ1, λ2, and λ3 also known as interleukins IL29, IL-28A and IL-28B, respectively. Table 1.2 summarizes the members of each class of interferon.

Type I Interferons	Type II Interferons	Other Interferons
IFN-α1, -α2, -α4, -α5, -α6, -α7, -α8, -α10, -α13, -α14, -α16, -α17, -α21 IFN-β IFN-δ *, IFN-ε, IFN-κ, IFN-τ **, IFN-ω IFN- ζ zeta? * described only for pigs ** described only for cattle	IFN-γ	IFN-λ1, λ2, λ3
Chromosomal Location Origin: 9 (human) 4 (mice)	**Chromosomal Location Origin:** 12 (human) 10 (mice)	**Chromosomal Location Origin:** 19 (human)
Induced by: Viruses, some intracellular bacteria, protozoans, other cytokines.	**Induced by:** Antigen stimulated T-cells, Natural Killer (NK) cells, Natural Killer T (NKT) cells	**Induced by:** Viruses
Functions: Antiviral, anti-proliferative, increase MHC class I expression.	**Functions:** Antiviral, increase MHC I and II expression, growth and maturation factor for some cells.	**Functions:** Anti-viral

Table 1.2: Summary of different classes of interferons. An overview of the three different classes of interferons including the different types of interferons in each class, their chromosomal location origin, induction and function.

There are two main classes of type-I interferons; IFN-α and IFN-β. In humans there are 13 known subtypes of IFN-α, and one class of IFN-β [58]. All type-I subtypes share considerable structural homology. In contrast to type-I interferons there is only one type-II IFN; IFN-γ (originally termed macrophage-activating factor) that is structurally unrelated to type-I interferons. Type-I interferons are synthesised in direct response to viral infection. Type-II interferons on the other hand are synthesised in response to recognition of infected cells by CTLs (CD4+ T helper 1 cells (Th1) and CD8+ T cytotoxic lymphocytes) and NK cells. There is now also evidence that other cells types such as B-cells, NKT cells and professional antigen presenting cells may also secrete IFN-γ [59]. All type-I IFNs bind a common cell surface receptor; the type-I IFN receptor and IFN-γ binds to a different receptor; the type-II IFN receptor.

Interferon-γ

Macrophages stimulated with IFN-γ induce direct anti-microbial and anti-tumour mechanisms in addition to up-regulating antigen processing and presentation pathways [59]. Additionally this cytokine can direct growth, maturation, and differentiation of many cell types as well as orchestrating leukocyte attraction. IFN-γ increases the expression of Fc receptors for IgG on macrophages and is the only type of interferon capable of efficiently up-regulating the MHC Class II expression on a variety of cells, and in turn promoting peptide-specific activation of CD4+ T cells [59-60]. IFN-γ is a key factor required for the 'classical' activation of macrophages.

The IFN-γ receptor is comprised of two ligand binding chains (IFNGR1) associated with two signal transducing (IFNGR2) chains. Both receptor chains lack intrinsic kinase / phosphatase activity and therefore associate with other signalling machinery for signal transduction. Arguably the best characterised association is with the JAK-STAT signalling cascade. The IFNGR1 intracellular domain contains binding motifs for JAK1 (Janus tyrosine kinase) and the intracellular region of IFNGR2 contains a binding motif for recruitment of JAK2. A phosphorylated IFNG-receptor complex can then transduce signal by phosphorylating the signal transducer and activator of transcription (STAT)1

on the tyrosine residue at position 701 (Tyr701) [61]. This phosphorylation then results in the formation of STAT1-STAT1 homodimers that translocate to the nucleus and bind IFN-**γ a**ctivated **s**ites (GAS) elements in the promotors of certain target genes. The transcription of type-II dependent genes is regulated by these GAS elements. Figure 1.3 provides a simplified schematic of type-I and type-II interferon signalling.
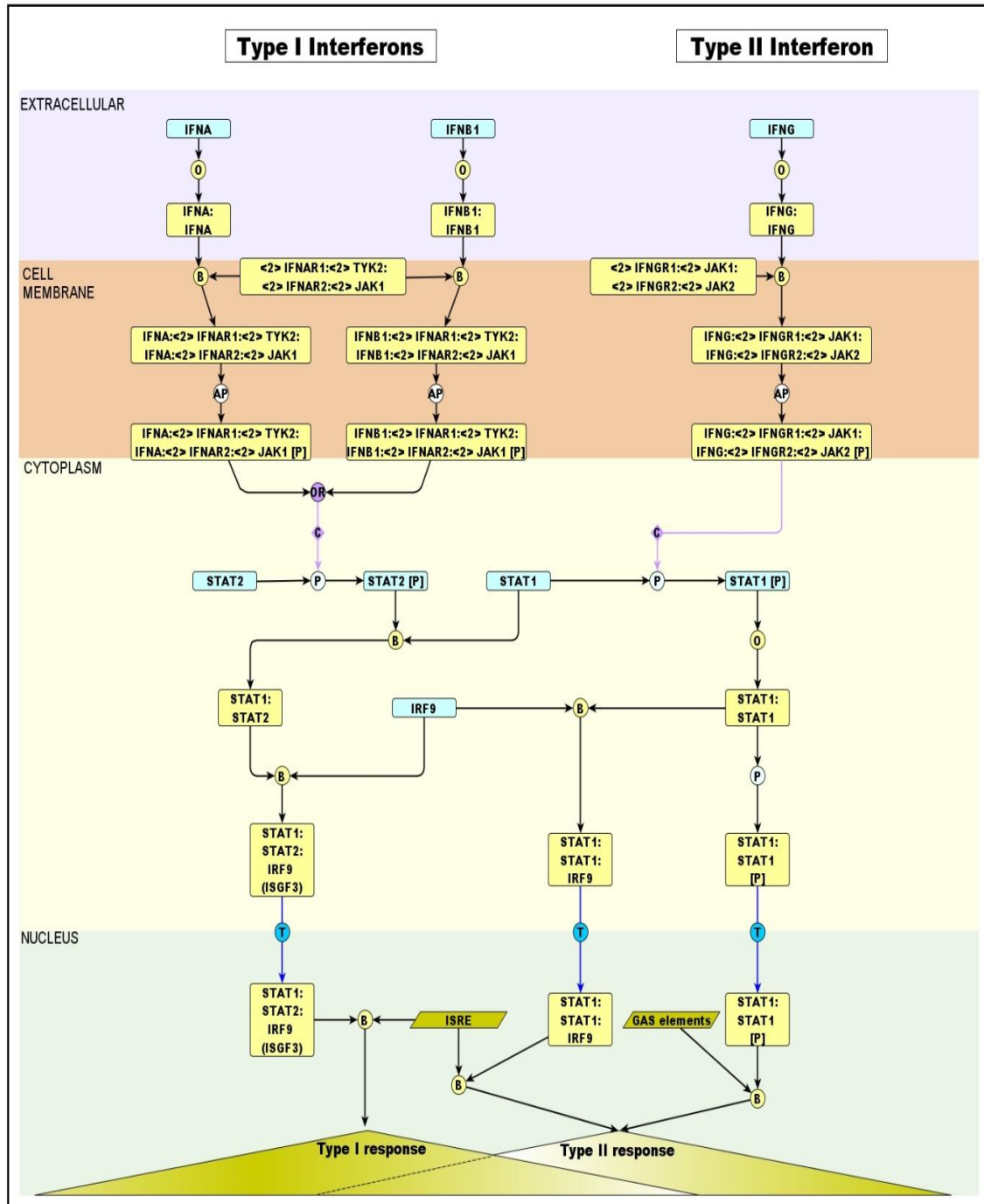
Figure 1.3: Simplified overview of signalling events following type-I or type-II interferon receptor activation. The pathway diagram is depicted using the modified Edinburgh Pathway Notation (a graphical notation for biology discussed in Chapter-3), and describes the series of events following type-I and type-II receptor engagement with their respective ligands. Type-I receptor activation leads to the auto-phosphorylation of the type-I receptor complex, which can then phosphorylate STAT2. Phosphorylated STAT2 can combine with STAT1, IRF9 to form the archetypal type-I transcription factor ISGF3. The activated type-II receptor complex phosphorylates STAT1, which is then capable of forming a STAT1 homodimer, or a complex of STAT1:STAT1:IRF9. The resultant type-I and type-II transcription factors bind to specific elements in the 'interferon stimulated' genes they target. Overlap between the transcriptional targets of type-I and type-II signalling exists.

Interferon-β

Type-I IFNs are secreted at low levels by almost all cell types. Viral infection is generally the classical stimuli for IFN-α or IFN-β production. Hematopoietic cells are the major producers of IFN-α, whereas fibroblasts are the major source of IFN-β. IFN-β is also produced by macrophages under appropriate stimuli [59].

As with the IFN-γ receptor, the type-I receptor also comprises multi-chain structures which are composed of at least two distinct subunits: IFNAR1 and IFNAR2. The IFNAR1 subunit is constitutively associated the tyrosine kinase 2 (TYK2), whereas IFNAR2 is associated with JAK1. The initial step is activation of the receptor associated JAK(1), which occurs in response to a ligand dependant rearrangement and dimerization of the receptor subunits. The activated receptor complex then phosphorylates STAT2 and from this a key type-I transcriptional complex ISG factor 3 (ISGF3) is formed. ISGF3 is composed of the phosphorylated (and activated) STAT1 and STAT2 along with the interferon regulatory factor (IRF)9. This complex binds specific elements known as IFN-stimulated response elements (ISREs) that are present in the promotors of some ISGs. As well as activation of the classical JAK-STAT pathway, there is evidence that activation of the IFN receptor associated JAKs may regulate (directly or indirectly) other downstream signalling cascades such as the mitogen activated protein kinase p38, and phosphatidylinositol 3-kinase signalling pathways [58].

The Post Genomics Era and its Technologies

The post-genomics era refers to the period following the sequencing of the human genome [62-63] a milestone in genomic research. This era has seen the has seen the emergence of novel functional genomics techniques and the significant improvement in existing technologies, not least; high-throughput DNA sequencing, DNA microarrays, high-throughput cell based assays, improvements in mass spectrometry, the discovery of RNA-interference (RNAi), cap analysis gene expression (CAGE) for predicting transcription start sites, and two-hybrid screening. Essentially the post-genomic era

has witnessed an exponential increase in available data facilitated by the increasing ease of gathering genomic and other high-through-put data. Consequently the fields of computational-biology and bioinformatics have also grown rapidly over the years. The premise of post-genomic research has been the transformation medicine and drug discovery. To achieve this end, the challenge lies in managing, interpreting and integrating the different data-sets to better understand the system of interest.

RNAi and gene expression microarrays are two investigative tools central to the studies described in this thesis and are therefore described in some detail below.

RNA-interference

RNA interference (RNAi) is a naturally occurring mechanism for gene regulation found in many eukaryotes [64]. The pathway used in this system is now exploited routinely in the biological sciences for investigating the role of various genes in a given cell or condition. The RNAi pathway involves small non-coding RNAs (e.g. exogenously introduced short interfering RNA (siRNA) or the endogenous microRNA (miRNA)) which associate with nuclease-containing regulatory complexes, then pair with complementary mRNA, and in doing so prevent their expression. The RNAi process begins with the processing of long double-stranded RNA (dsRNA) (that is endogenous to, or introduced into, the cell), into small RNA duplexes by a ribonuclease III (RNaseIII) enzyme known as Dicer (reviewed in [65-66]). These duplexes are then unwound, and one strand (guide strand) is preferentially loaded into a protein complex known as the RNA-induced silencing complex (RISC). The RISC complex complete with loaded single-stranded RNA (ssRNA), then directs the cleavage of messenger RNAs that contain sequence homologous to the ssRNA [65-66]. The endonuclease responsible for this cleavage has been identified as an Argonaute protein [67-69].

Systematic knockdown of protein function by RNAi used in combination with molecular and cellular phenotyping analysis is a powerful technique for elucidating gene function. Commonly the approach is used to explore gene function in *in vitro* cell culture and *in vivo* in model organisms. Synthetic siRNA or short hairpin-RNA (shRNA)

are used to mediate the RNAi pathway in experimental settings. shRNA (a short sequence of RNA that makes a tight hairpin turn), can also be synthesized in vivo from RNA polymerase III promoters [70-71]. shRNA are introduced into cells using a (viral) vector, allowing integration of the shRNA expression cassette into the host genome, and are therefore preferred in studies requiring long-term stable knockdown of targets [72]. siRNA tend to be easier to design and transfect than shRNA and so are generally preferred for screening experiments [72]. Moreover the dosing of silencing ssRNA is better controlled with the siRNA method [72].

Application of RNAi as a therapeutic is also being considered with great interest since it would potentially provide a powerful method for inhibiting any gene whose expression and protein product may contribute to disease [73-74]. Consequently RNAi has unveiled the opportunity to manipulate a vast number of disease targets that are otherwise currently intractable to traditional small-molecule and protein based intervention approaches. Macrophages are particularly attractive targets for (gene) therapeutic interventions given their central role in a wide variety of biological processes and pathologies.

There are however a number of obstacles to be addressed in order for siRNA use to be successful both *in vitro* as a laboratory tool and *in vivo* as a therapeutic [75]. These include challenges in siRNA delivery, immune activation and immune mediated toxicities, as well as non-immune off target effects. Initial studies in mammalian cells suggested that siRNA are specific and small enough to evade immune detection [64]. Others went on to report inadvertent effects such as activation of the interferon response at even low siRNA concentrations [76-77], and mRNA degradation mediated by partial sequence complementation [78-79]. A range of strategies have been deployed to design stable, specific synthetic siRNAs which circumvent immune activation and off target effects. These include computational screening of potential siRNA sequences to avoid use of known immuostimulatory sequences and those with a high potential for off target effects [80-81], as well as structural and chemical modifications to the siRNA [82-85].

Microarrays for Gene Expression Analysis

Microarrays permit the comprehensive and simultaneous analysis of nucleic acid sequences [86]. The most common applications of DNA microarrays are for measuring changes in gene expression level and for detecting or genotyping single nucleotide polymorphisms (SNPs). Arrays comprise of DNA probes which are generally immobilised onto a solid surface (e.g. glass slide, silicon chip or microscopic beads). Probes are generally either spotted cDNA or commonly oligonucleotides and are designed to interrogate specific gene or intergenic sequences (or polymorphisms) in a given organism. The probes for assaying gene expression i.e. mRNA transcript abundance are commonly designed to detect cDNA/cRNA generated from total RNA samples. The general procedure for expression analysis by microarrays involves the fluorescent labelling of samples, which are then hybridised to the microarrays under specific conditions. The arrays are washed to eliminate non-specific binding to probes and are scanned to determine fluorescent signal at all probes. The signal strength at each probe position corresponds to amount of complementary target bound. Raw signal data from multiple microarrays (to be compared) is subjected to a normalisation step to make adjustments for systematic/technical errors introduced when batch processing arrays e.g. differences in scanning or hybridisation.

The very first miniaturised microarray could assay 45 gene sequences of Arabidopsis thaliana [87]. Subsequent advances in microchip technology as well as the growing availability of sequence data for complex genomes, facilitated the development of more comprehensive microarrays. Current microarrays technology can interrogate all known and predicted genome sequences as well as individual exons. Multiple probes may be used to target a given sequence. For example the Affymetrix Human 1.0 ST array is comprised of 764,885 distinct probes targeting 28,869 well-annotated genes.

As microarray analysis developed, standards for recording information about microarray experiments were proposed [88]. Commercially available microarrays are now relatively inexpensive, robust and reproducible. Moreover, microarray data has

existed for over a decade; giving time for researchers to develop and debate analysis techniques to analyse this data. Indeed there is some debate as to whether microarrays will eventually become obsolete with the emergence of an alternative method for gene expression analysis; RNA-Seq (or Whole Transcriptome Shotgun Sequencing) [89]. RNA-Seq in principle can provide a more accurate measurement of transcript levels, makes no prior assumptions of sequences, and has become more affordable in recent years [90]. As with many high- throughput technologies, RNA-Seq analysis also faces a number of challenges, not least the computational management; from storage, processing, and optimal methods for data analysis.

Systems Biology

What is Systems Biology?

The traditional (or reductionist) approach to biology has centred on dissecting the properties of individual genes/proteins and their contribution to the operation of a process. Systems biology is a paradigm and discipline centred on the comprehensive quantitative analysis of all the components of a biological system and how their interactions influence the properties of a given system (or process). The holistic systems-level approach to modelling and analysis has been fostered by technological advances in high throughput technologies as well as breakthroughs in functional genomics techniques (discussed earlier). In fact, managing, mining and interpreting the mountains of available data arguably requires a systems level approach.

The 'system' in question can range from signalling networks, a cell, tissue, organ, organism etc. Systems are comprised of networks of interacting components. For example protein and gene regulatory networks are fundamental to the operation of a cell. The chief motivation and objective of systems biology is therefore to identify the components of a given system and how they interact with each other, in order to understand and/or predict emergent behaviour of the system; that is behaviour which

cannot be predicted from analysing individual components [91]. The premise of these endeavours is that it should eventually be possible to (i) predict the behaviour of a system in response to perturbation and (ii) redesign/perturb the systems-model to create new emergent properties [91]. If these points are eventually achieved they would revolutionise the fields of preventative medicine and drug discovery.

In practice systems biology often refers to efforts to both model systems both in terms of defining the interactions between the components of a system and computationally modelling its activity, as well as new approaches to interpreting and mining biological data derived from genome-wide studies.

Systems Level Analysis of Immune Signalling

A 'systems-level' approach is particularly suited for better understanding complex biological systems and processes (e.g. immune signalling, neurobiology). Although it is becoming increasingly evident that systems based challenges permeate all areas of biology and medicine. Emergence, robustness and modularity have been identified as three key concepts central to systems biology [91]. All of these concepts are relevant to our understanding of immune signalling and macrophage biology. Emergence relates to system properties which cannot be explained by the activity of individual components alone e.g. the activation status of a macrophage depends on the balance of activity of numerous proteins. Robustness refers to the maintenance of phenotypic stability in spite of perturbations. Robustness is a critical feature of the immune system which must be adaptable to environmental fluctuations and tolerant to errors, in order to contend with various challenges posed by microbes. Finally in biological systems, modularity describes networks of components e.g. proteins/genes that interact together in undertaking a common function [91]. Signalling pathways comprise networks of interacting proteins, and in the macrophage regulation of transcriptional networks is vital in defining the phenotypic properties of the cell.

There are now numerous systems-based approaches to studying the immune system and macrophage biology which have been employed (reviewed in [92]). These include

a number of initiatives to collate data and investigate immune signalling systems and pathways; examples including the INTERFEROME database of interferon gene targets [93], the Alliance for Cellular Signaling collaborations [94], the LIPID MAPS consortium [95], the Inflammation and Response to Injury Project [http://www.gluegrant.org/], the DC-ATLAS resource for interpreting high-throughput data generated from dendritic cells [96], and the Innate Immunity Project [www.innateimmunity-systemsbiology.org/], to name just a few.

The macrophage possesses a host of features which render it an attractive if not vital target for 'systems-level' analysis [97]. These features include; the substantial range of well characterised signalling pathways; the plasticity of the cell; the relative ease of obtaining and culturing primary macrophages (from human blood, mouse bone marrow and peritoneum); the availability of macrophage-like immortalised cell lines; and crucially the clinical significance and therapeutic appeal the macrophage holds [97]. Whilst the ability to perform quantitative and qualitative measurements on the cellular components of the macrophage has increased massively, as has knowledge on how they interact with each other. The challenge still remains in converting these observations into detailed systems-models. However, without such models we cannot hope to truly understand macrophages or indeed any other cell at a systems level. Detailed and well characterised models of the signalling pathways, disease processes active in the macrophage are still scarce. Other key challenges in deploying the systems-level approach include standardising datasets, and integrating different levels of data (e.g. transcription, protein expression, kinetic data) to build and inform system-models. Defining modularity, especially the transcriptional networks activated in macrophage and how these relate to the phenotypic properties of the cell is essential in understanding the heterogeneous behaviour of this cell.

Pathway Modelling

Interest in 'pathway biology' has never been greater as we struggle to comprehend cellular systems from a combination of targeted studies and the deluge of data flowing

from 'omics platforms. This is reflected by the escalating efforts to assemble pathway diagrams [98-102], develop standards for depicting pathways [103-106], software to support their construction [107-109] and exchange [110-112], and the development of approaches to model and predict pathway behaviour [113-115]. Whilst arguably there is no such thing as a pathway only one big integrated network of molecular interactions, it is still useful to think in terms of pathways as being connected modules of this network. As such a pathway may be considered to consist of a specific biological input or event that initiates a series of directional interactions between the components of a system leading to an appropriate shift in cellular activity. As we begin to appreciate the potential complexity of these molecular networks, there is increasing interest in modelling pathways in order to expand our understanding of biology from the traditional gene-centric view of life, to a systems level appreciation of biological function.

Much of this work describing pathways and their interactions remains locked in the literature where specific insights into pathway function are subject to the semantic irregularities that come with their description by different authors. Pathways are understood more generally by their description in reviews and diagrams produced on an *ad hoc* basis. Whilst such diagrams are clearly useful aids to understanding specific cellular processes, even at their best, they are not sufficient by themselves, relying on extensive textual descriptions to explain what is shown pictorially. Furthermore these pathways are rarely available as a cohesive network with pathway diagrams usually focusing only on a small part of a biological system which often reflects the curator's bias, such that the 'same' pathway described by different individuals may share little in common. In this sense pathway resources offer a fragmented view of systems with some proteins or metabolites being members of numerous pathways; the concept of pathway membership being a highly subjective division. Figure 1.4 shows the apoptosis pathway as curated by four different sources and underscores some of the issues with current pathway depiction. Whatever the source of these pathways and networks they generally suffer from graphically poor representation with ambiguity around the precise identity of what is being shown and the exact nature of their interaction.
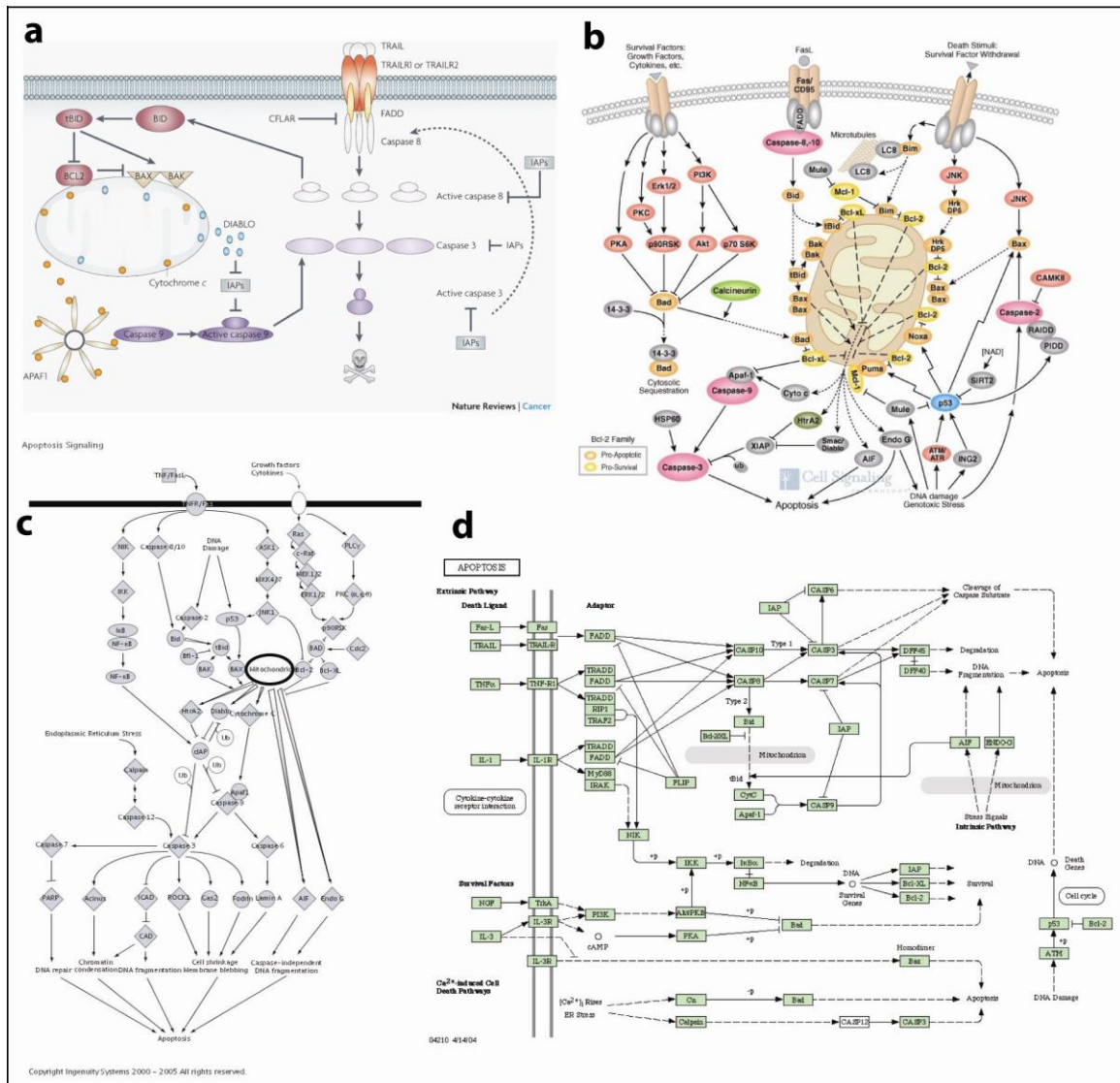
Figure 1.4: Apoptosis pathways curated and assembled by four different sources. The 'same' pathway may share little in common in terms of layout, pathway components, standard graphical notation and nomenclature systems (if any used) for pathway components. (a) Apoptosis Pathway as extracted from a review article; Johnstone *et. al* 2008, Nat Rev Cancer. (b) Mitochondrial Control of Apoptosis, taken from www.cellsignal.com/pathways/apoptosis-signaling. (c) Apoptosis Signalling as portrayed by Ingenuity Systems a commercial queriable pathway resource. (d) Apoptosis as taken from KEGG (Kyoto Encyclopaedia of Genes and Genomes) Pathways. In all cases these clearly highly simplified representations of the molecular interactions leading to apoptosis.

In order to escape the gene-centric view of biological systems, requires the development of better ways to order and display our knowledge of protein interactions and the systems they form. Moreover the creation and developments of standards is essential to enhance and permit the exchange of pathway knowledge. Until very recently biology has lacked standardised graphical notation schemes for illustrating pathway information [103, 116-118], and even now these standards are still being developed and are far from universally applied. Formalized diagrams (i.e. those constructed using a graphical language) act as a visual representation of the interactions between cellular components and provide a valuable resource for modelling network structure and the dependencies between components [119]. In addition, pathway models are an invaluable resource for interpreting the results of genomics studies [120-126], for performing computational modelling of biological processes [115, 127-130] and fundamentally important in defining the limits of our existing knowledge. Large integrated diagrams of metabolic pathways have been available for many years, for example Gerhard Michal's classic biochemical pathways wall chart first published by Boehringer-Mannheim in 1968. Such pathway diagrams are inevitably complex, but potentially liberate the user to explore the interconnectivity between what might be seen as separate pathways and get an overview of topology of the system as a whole. In contrast, the assembly of detailed diagrams of signalling pathways as integrated networks rather than a series of disconnected views has been little explored.

To gain a better understanding of a heterogeneous cell such as macrophage will require characterisation of its constituents and how they interact over time. As with many biological systems, certain macrophage pathways are very well characterized whereas little is known about many others. Even where pathway domain knowledge does exist however, it is generally fragmentary and subjective. Chapters-2 and -3 will explore means of creating a macrophage pathway resource and the *in silico* representation of the biological interactions underpinning the activity of this cell.

Networks in Biology

Visualisation and analysis of biological data as networks is now becoming a widespread and increasingly important approach in biological and medical research [131]. In classical graph theory, a network (or graph) consists of nodes connected by edges. For biological networks the nodes may represent a biological entity (for examples genes, proteins, organisms) and the edges denote a type of relationship or interaction (for example protein-protein binding, metabolic coupling, genetic origin as well as experimentally determined similarities). Biological relationships analysed using networks have included; sequence similarity, protein structure, protein interactions and evolutionary relationships [132-134]. Most biological networks tend not to be random, but follow a series of basic organizing principles in their structure and evolution. This also applies to other natural networks, as wells as technological and social systems, and distinguishes them from randomly linked networks. The structure of such "non-random" networks can provide insight, often overlooked by other methods of data representation.

Many biological networks show a high degree of clustering i.e. highly interlinked local regions in the network or topological modules [131]. A number of network-clustering tools, designed for identifying such modules have emerged over recent years [133, 135-137]. The clustering technique has performed exceptionally well for analysis of protein-protein similarity networks [133] and demonstrated that clustering nodes according their graph context, rather than iterative pair-wise clustering has great potential for the discovery of novel aspects of biological function.

Network Based Approach for Gene Expression Data Analysis

Traditionally statistical methods have taken precedence for the analysis of microarray transcriptomics data. These pair-wise comparison methods can include fold change/significance analysis between two or more conditions e.g. between treatment vs. controls. Whilst statistical methods are often viewed to be rigorous, there are

certain caveats to reliance purely on statistics for microarrays experiments, especially those with a large and varied number of treatment conditions. For example statistical methods ideally require multiple replicates of a condition to determine the significance threshold of a result. This can become inherently expensive and impractical for large scale screens or experiments with many conditions. There are also often limitations on the quantity of data that can be presented in an intuitive manner. In many cases the most changing transcripts (one condition vs. another) will be determined and much of the remaining data overlooked. Furthermore the individual sample by sample comparison approach does not take into account the relationship of all the data samples or arrays to each other. Correlation based methods whereby cohorts of genes related in their behaviour across a data set are identified, are also a commonly applied approach in expression analysis.

Although network analysis is proving to be valuable for the analysis of a range of biological data, its application for examining microarray gene expression data has until years recently been overlooked, possibly due to lack of supporting software tools. However the general quality, abundance and high-dimensional nature of expression-data make it compatible with network analysis [138]. Typically microarray expression data consists of many thousands of measurements of relative transcript abundance and depending on the study these measurements are derived from a few to several thousand biological samples [139]. The collection of expression values of any given transcript across the samples of interest is referred to as its *expression profile*. The network paradigm for the analysis of gene expression data is based on defining the level of correlation (degree of similarity) between expression profiles of different transcripts. The basic principles of generating and analysing gene-expression networks are described below.

Basic principles of correlation based network analysis for gene expression data

For the purpose of gene expression analysis in this thesis I have applied a network based approach to explore the microarray data generated by this project. Generation

of network graphs from gene expression data has used the Pearson correlation coefficient as a measure of similarity between expression profiles, where an expression profile is defined as the collection of numerical data values (expression values) of an individual probe or probe-set targeting a defined transcript, derived from the range of samples being investigated. Pair-wise Pearson correlation coefficients are calculated for every transcript represented on the array, and each calculation then defines the strength of a relationship between two transcripts or in essence provides a similarity score between their expression profiles between +1 (perfectly correlated) and -1 (negatively correlated). Correlation coefficients above a predefined threshold can be used to draw edges between genes (nodes) and generate a network graph of expression relationships. Figure 1.5 provides a simple worked example of a correlation based gene expression network.
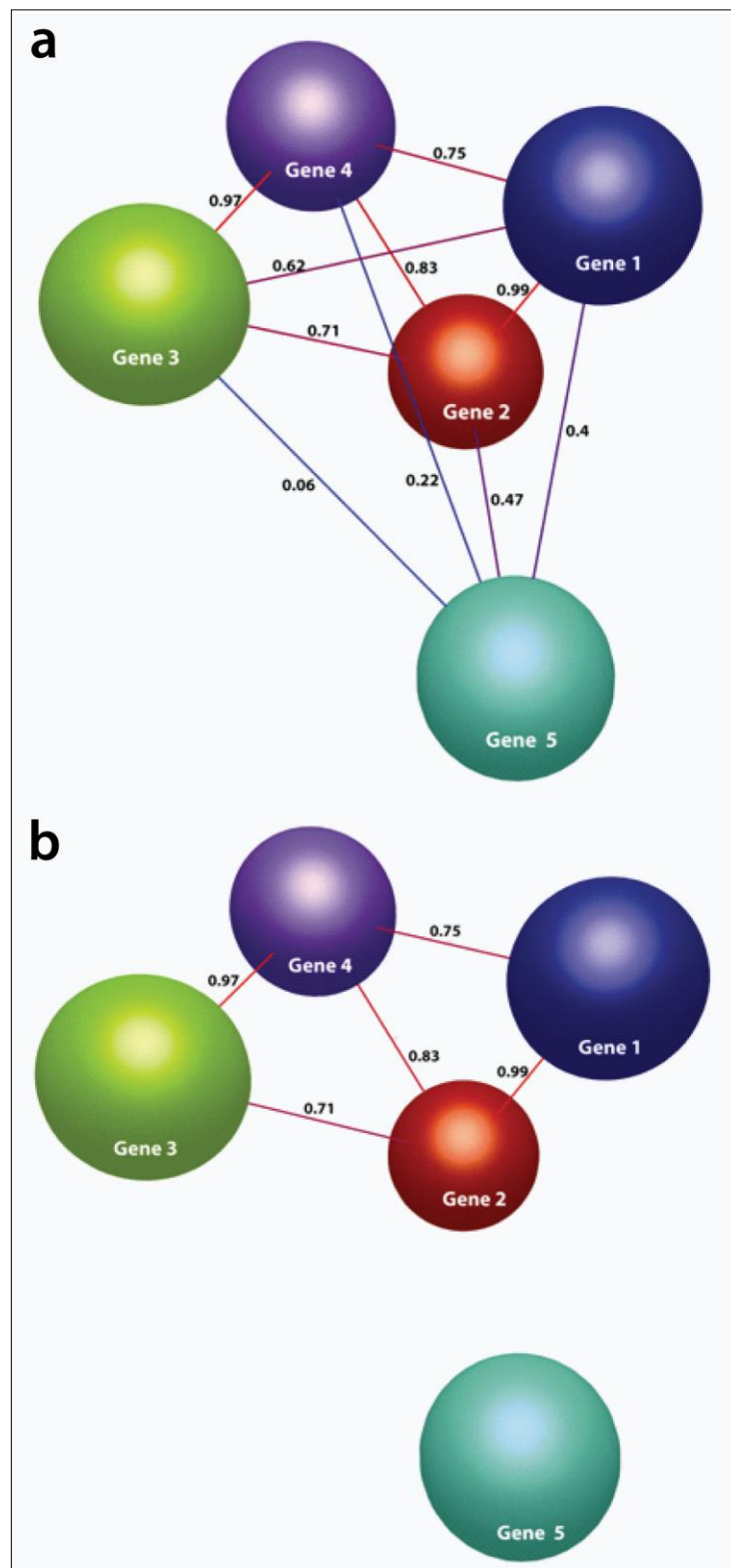
Figure 1.5: Example of a network graph consisting of nodes (genes) and edges (correlation between the genes). (a) The network graph has not been filtered so all Pearson correlation relationships are displayed, therefore all nodes (genes) form connections with each other. **(b)** The network graph has been filtered so only Pearson correlations above a threshold of 0.7 are displayed. Consequently the network is connected by fewer edges, defining only relationships above 0.7. In this scenario Gene 5 does not form part of the main network, since all of its relationships with other genes fall below the 0.7 Pearson correlation filter, indicating that its expression profile is poorly related to the other genes.

Expression data can be influenced by both biological variations across the samples and by technical artefacts. However generation of these network graphs from expression data makes no prior assumption of experimental variables, including design, normalisation method, microarray platform or even the questions being addressed by the study. Thus in essence this approach provides a truly unbiased view of the data.

BioLayout *Express*^{3D} is a software tool designed specifically for the visualization, clustering and analysis of large network graphs in two- and three-dimensional space derived primarily, but not exclusively, from biological data [138-139]. The tool had now been optimised so it can render graphs of approximately 45,000 nodes in size (connected by ~5,000,000 edges, although this is hardware dependant) and also has an inbuilt application for clustering graphs using the Markov clustering (MCL) algorithm [137]. The MCL algorithm is an unsupervised approach for sub-dividing graphs non-subjectively into discrete sets of genes sharing similarities in their expression, otherwise known as clusters (or graph modules). Clusters are defined on the basis of node connectivity and edge weight (strength of relationships). Consequently nodes (transcripts) contained within clusters are in close proximity in the graph and are highly related in their expression profiles. The MCL algorithm can be adjusted by use of an MCL-inflation value, to alter the stringency of the cluster membership and consequently the granularity of clustering. The higher the MCL inflation value is set, the stronger the correlation (in pattern of expression) between the cluster members. Although this also reduces the size of clusters within the graph as modules are sub-divided further based on their graph context.

Previous Use of Network Based Approach for Interpreting Transcriptomics Data

BioLayout *Express*^{3D} is a network analysis tool that has been specifically designed for the analysis of gene expression data and is especially suited for large and complex data sets. It has now been utilized in the analysis of a number of published studies

addressing a range of biological questions [140-144]. In order to explore the network based approach for analysis of transcriptomics data and to gain a better understanding of the transcriptional profile of macrophages I have contributed to the analysis of a number of gene expression datasets [145-147]. This includes a large data set assaying gene expression across 44 purified primary mouse cell populations or untransformed cells and two mouse organs; the spleen and mesenteric lymph nodes [145]. The approach allowed the concurrent analysis of 20,346 nodes (transcripts), and their relationships (944,650 edges) in one network graph (Figure 1.6), a scale that would otherwise be unattainable using other methods or network tools. From the network graph it was possible to extract a number of interesting observations; firstly that some patterns of expression i.e. clusters of transcripts were specific to particular cell types (Figure 1.6), some of which may prove to be markers for those cells. In other cases some of the clusters appeared to be associated not with cell lineage but with broad biological processes. This not only provided insight into the processes shared or specific to cohorts of cell types, but also demonstrated the potential to characterise as yet un-annotated proteins since the likely function or process in which they are involved can be inferred by their co-expression with genes encoding proteins of known function [145].
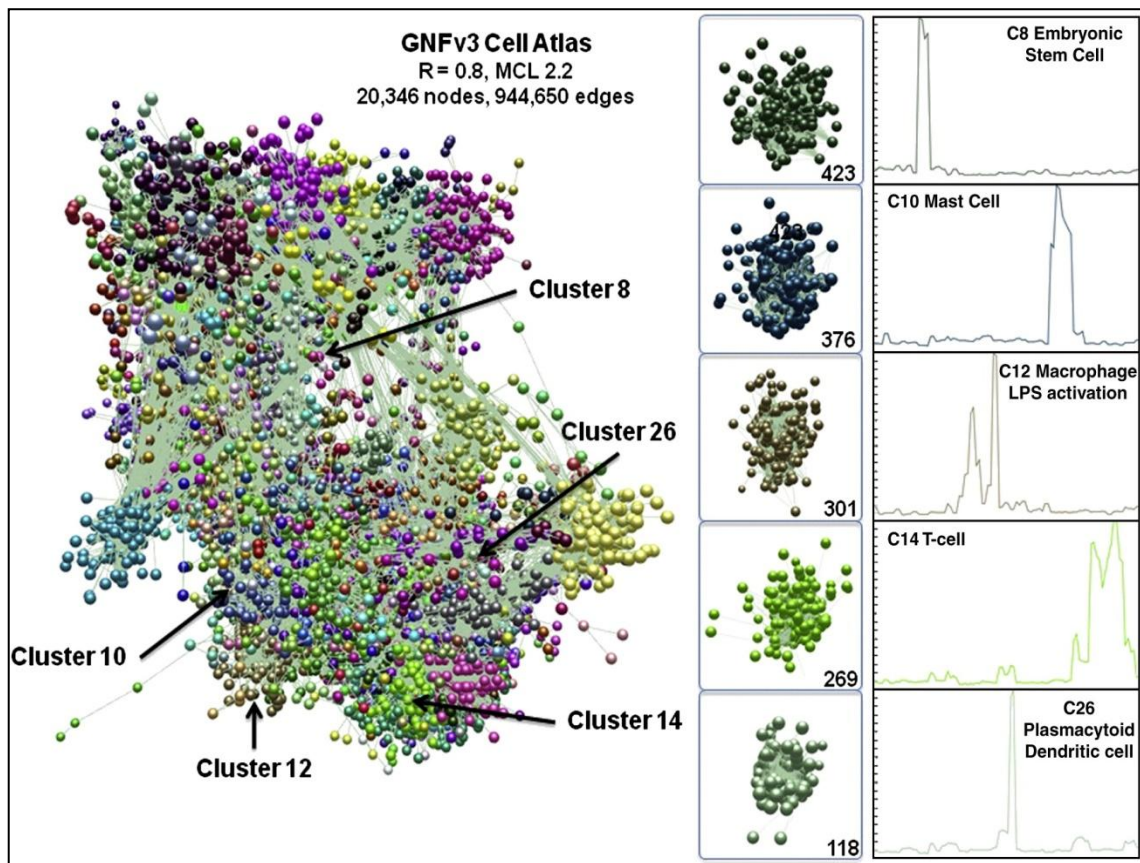
Figure 1.6: Network analysis of mouse transcriptomics atlas (**extracted from Hume *et al.,* 2010**). Samples of 44 mouse cell populations and 2 tissues (lymph node, spleen) were collected and analysed on Affymetrix MOE430 2.0 arrays in duplicate, see GEO dataset: GSE10246 (www.biogps.org). Results were normalised using the MAS5 algorithm and the tool BioLayout *Express*[3D] used to calculate pair-wise Pearson correlation coefficients for every transcript represented on the array. A network graph of 20,346 nodes (transcripts), connected by 944,650 edges was generated by filtering data to display only Pearson correlation relationships (between transcripts) of 0.8 or above. The resultant graph was then clustered using the Markov clustering algorithm with an inflation value of 2.2 resulting in 812 clusters containing > 4 nodes. Examples of these clusters, shown isolated from the main graph alongside the average expression profile of the transcripts that make up the cluster. The number of transcripts in each cluster is shown bottom right next to the graph of the cluster. Transcripts belonging to the five example clusters are largely expressed in specific cell types and/or under specific conditions i.e. LPS activation of macrophages.

The network based approach had also been used to assess genome wide changes in transcript abundance in response to targeted RNA-interference induced suppression of the expression of a number of key genes associated with IFN signalling in murine macrophages prior to stimulation with mouse IFN-γ (Figure 1.7) [146]. The original objectives of this investigation were to further our understanding of the mechanism by which certain ISG's contributed to the protective effect of IFNγ during viral infection by targeting the genes with siRNA. However the act of siRNA transfection itself induced a type-I IFN transcriptional response, thus even in the absence of IFN-γ treatment IFN response genes appeared to be regulated. Moreover one cluster of regulated genes represented transcriptional changes occurring regardless of IFN-γ treatment. This cluster-(d) was enriched for type-I response genes. Interestingly six siRNA treatments (targeting Ifnb1, Irf3, Irf5, Nfkb2, Stat1, Stat2) perturbed the transcriptional networks associated with IFN signalling (Figure 1.7). Some of these siRNA targets were known to act within the IFN pathway. The observations made by this study formed the basis of the follow up experiments described in Chapter-5.
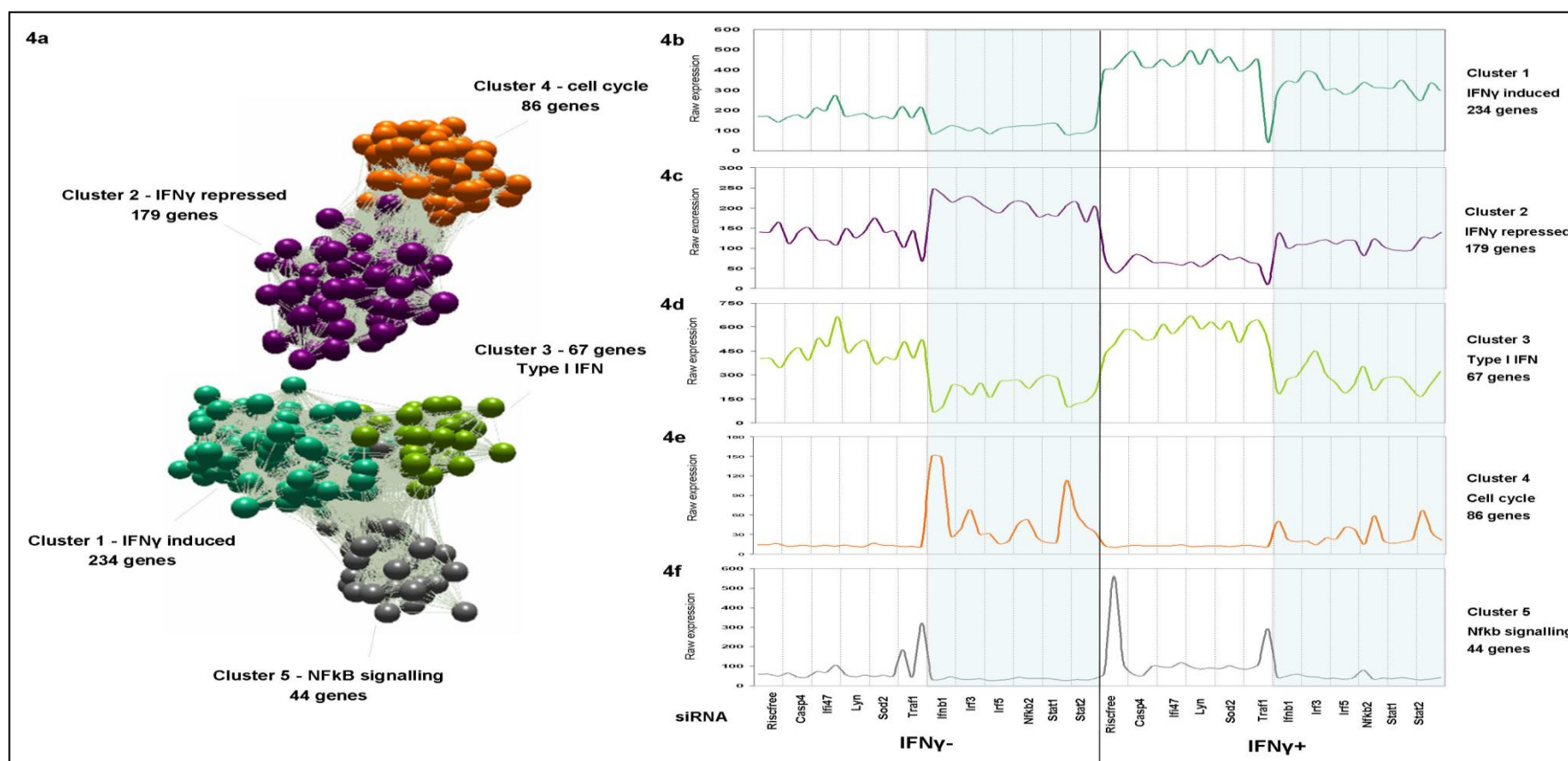
Figure 1.7: Clusters extracted from a transcriptional network of expression data from RNAi treated BMDMs, and the median expression profiles of transcripts within the clusters (Taken from Lacaze et al., 2009). A network graph filtered at a Pearson correlation r ≥ 0.9 was clustered using Markov clustering algorithm at an inflation value of 2.2). 4a: Five main clusters of co-expression emerged containing genes influenced most by siRNA & IFNγ treatment. A consistent disruption of transcriptional activity of mouse BMDM was observed using six particular siRNAs (shaded in blue) targeting Ifnb1, Irf3, Irf5, Nfkb2, Stat1 & Stat2 mRNAs. 4b: Cluster 1 – 234 genes whose expression is induced by IFNγ and repressed by the six active siRNAs. 4c: Cluster 2 – 179 genes repressed by IFNγ but de-repressed by six siRNAs. 4d: Cluster 3 – 67 genes whose expression is not influenced by IFNγ at 24 hour assay point but repressed by six siRNAs. Many of these are innate immune response genes 4e: Cluster 4 – 86 genes de-repressed by siRNAs, many of which have known functional association with cell cycle 4f: Cluster 5 – 44 genes enriched with annotation for NF-kB signalling.

Aims and Objectives

The overall objective of this thesis was to gain a better understanding of the signalling pathways underpinning macrophage activation (particularly in response to interferons) by applying a 'systems-levels' approach to biological modelling, data analysis and mining. Systems biology aims to escape the traditional gene-centric, reductionist view of biological investigations. This necessitates the development of better ways to order, display, mine, and analyse biological information, from the output of high-through-put 'omics technologies to our growing knowledge of protein interactions and the systems they form. Crucially these developments are necessary if we are to improve our understanding of complex and dynamic biological entities. The work described in this thesis is associated with three main objectives:

1. A wealth of literature describing the individual interactions of signalling pathways active in the macrophage currently exists. The challenge lies in capturing this information in a format with the potential for 'systems-level' analyses. Therefore an initial objective of this work was to develop a pathway resource of the signalling events active in the macrophage, particularly of receptor initiated pathways and type-I and type-II interferon signalling. Associated with this objective was the desire to explore and develop a graphical notation scheme for depicting biological concepts i.e. a language for drawing and exchanging pathway models.

2. Pathway models are a working hypothesis of how a system may operate over time and under given conditions. As such, models must undergo cycles of iterative refinement to become more accurate and detailed. In this case evaluation of the pathway resource created, exposed its limited coverage of transcriptional events. Thus a second objective of these studies was to explore the transcriptional networks associated with macrophage activation in response to type-I and type-II interferon, as well as LPS; three stimuli which are considered to prime macrophages towards the *M1*-type activation state. Traditionally statistical fold-change cut offs and lists of

the most differentially expressed genes are the approach applied for expression analyses. However the objective here was to explore the use of network-based analysis and presentation for transcriptomics data derived from macrophage activation studies.

3. Previous investigations by the group revealed that transcriptional networks associated with type-I interferon signalling, were perturbed using siRNA targeting genes which may act in the same pathway. In order to expand on these findings there was the desire to design and optimise a cell based assay to investigate the role of genes of interest in type-I interferon and LPS signalling.

Chapter 2. Construction of Large-Scale Diagrams of Macrophage Signalling and Effector Pathways
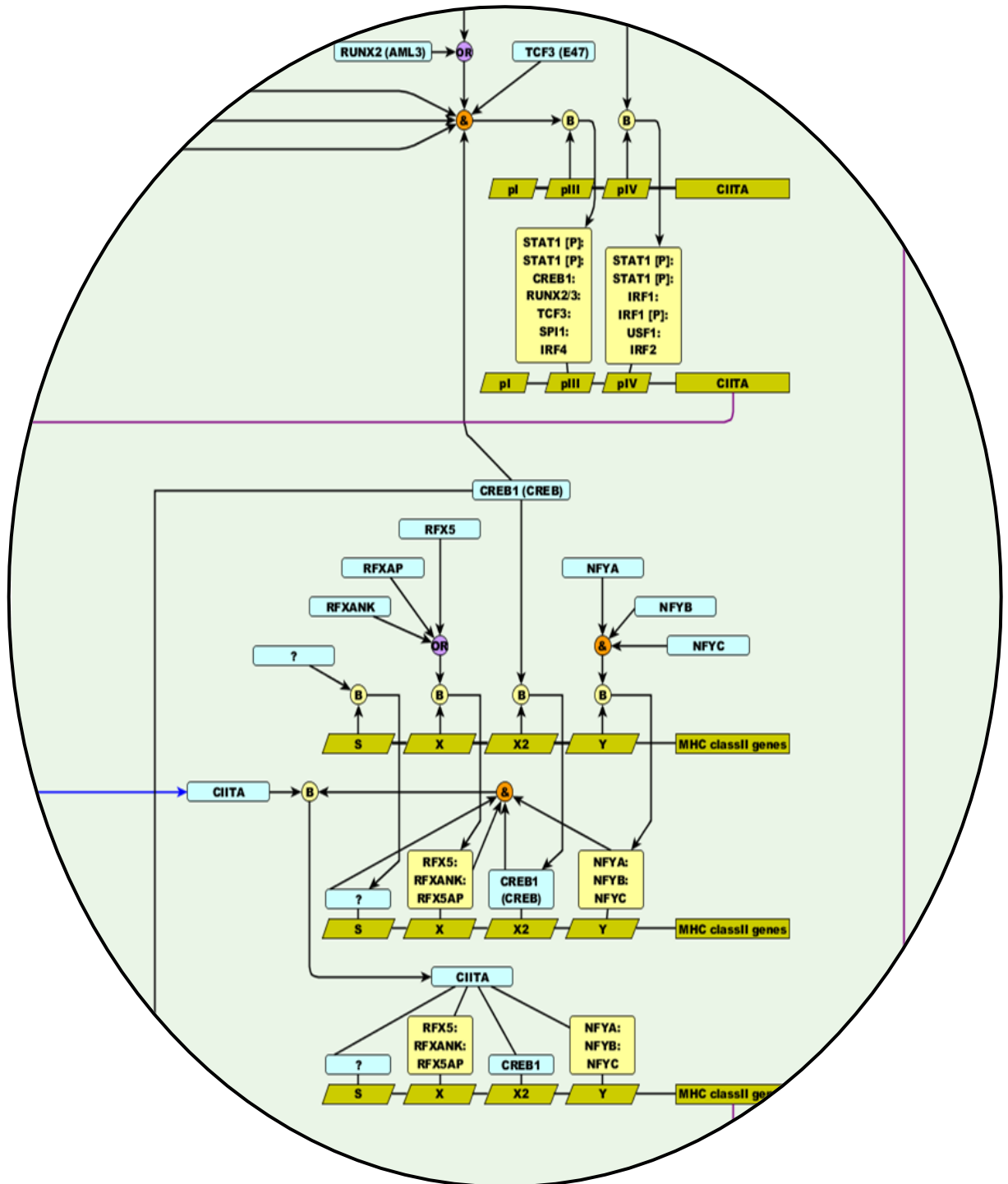
Introduction

This and the following chapter present my work on one of the central challenges in pathway biology: How does one construct clear concise pathway diagrams of the known interactions between cellular components that can be understood by and useful to a biologist? In our efforts to understand the signalling cascades fundamental to macrophage biology, I have endeavoured (as part of the group's efforts) to generate large integrated graphical models of these pathways.

In recognition of the importance of documenting pathway knowledge, many efforts have been made to collate pathway knowledge, together with information derived from large-scale interaction studies and literature mining, into public and commercial databases [148-157]. These offer searchable access to pathway diagrams and interaction data derived from a combination of manual and automated (text mining) extraction of primary literature, reviews and large-scale molecular interaction studies. Whilst invaluable and in many ways the best we have, a major problem with these efforts is that the information content of these diagrams is frequently limited and visualizations of these systems are of variable and often poor quality; Pathway components are often labelled using inconsistent nomenclature systems and depicted using a variety of shapes (glyphs) to illustrate component 'type' (see Chapter-3). Furthermore, notation schemes used for pathway diagrams to depict one molecule's interactions with another are not standard and are often limited in their ability to convey the exact nature of the relationship between components. Finally, pathway diagrams are generally highly subjective reflecting the curator's bias, such that two diagrams depicting the 'same' pathway may share little in common (see Chapter-1 Figure 1.4). Together these factors commonly result in uncertainty as to what exactly is being shown. The diagrams thereby fail to fulfil their basic purpose – to provide a comprehensive and unambiguous picture of what is known about a pathway. In an effort to address some of these issues, a number of groups have suggested notation schemes for drawing 'wiring diagrams' of cellular pathways [103-105, 158]. The evolution of one of these notation schemes, the modified Edinburgh Pathway Notation (mEPN) forms the basis of discussion in Chapter-3).

During the course of this thesis I have been engaged in constructing process diagrams [102, 159] of pathways important in regulating macrophage immune biology and known to be activated in these cells during infectious and inflammatory disease. The eventual aim being of these efforts would be a pathway resource that would serve to inform and aid the interpretation of wet-lab analyses of this cell and would also have the potential to be applied for computational modelling. The task of converting the literature into understandable, unambiguous, and useful pathways diagrams posed a range of challenges. Examples of these included; deciding on reliable sources for collecting pathway interaction data, deciding what constitutes a valid interaction, and how and what interaction data to store, how to arrange pathway components in a diagram, as well as choosing a suitable software for assembling the diagrams. In order to address these challenges and define a suitable protocol for constructing pathways, I was assigned the task of assembling a pathway diagram of apoptosis signalling (or programmed cell death). At the time of embarking on this task, the existing pathways of apoptosis were highly abstract and ambiguous (Chapter-1 Figure 1.4). Apoptosis is a carefully regulated process in the macrophage and is controlled by a diverse range of cell signals; both intrinsic (DNA damage, organelle stress) and extrinsic (cytokines, pathogens).

In constructing a model of apoptosis signalling it was necessary to establish some principle rules for assembling pathways. This stage of the pathway construction effort is described in **Phase 1** of the results section. Essentially the lessons learnt from this process laid the foundations for generating further pathways of interest, and in an effort to better understand the signalling cascades fundamental to macrophage activation, I then constructed an integrated framework diagram of macrophage signalling encompassing the TLR (Toll-Like-Receptor), IFN (interferon), NF-κB and apoptosis pathways. These signalling events are of central importance in defining the macrophages response to pathogens and do so in a highly inter-dependant manner [160]. The results of these labours are presented below (**Phase 2**) and in Raza *et al.,* 2008 [102]. The framework diagram comprised 295 nodes including 140 proteins and a

total of 272 interactions. Although this coverage was an improvement on other available macrophage pathways at the time this was still a very limited view of macrophage signalling. However the pathway proved to be a useful and a desire to expand this into a more comprehensive and powerful resource led the group to construct additional diagrams of pathways active in the macrophage including; antigen presentation, MAPKinase, non-TLR pathogen detection, to name just a few. These pathways and the framework diagram were then amalgamated to generate an integrated network of macrophage pathogen recognition and detection systems. These efforts are presented in **Phase 3** of the results and in Raza *et al.,* 2010 [159]. The current macrophage diagram is to our knowledge the most comprehensive network of signalling events in the macrophage constructed using a formalised graphical notation [159].

The pathways described were constructed using the Edinburgh Pathway Notation (EPN). In doing this another objective was to road test the Edinburgh Pathway Notation (EPN), (see Chapter-3) for its usability in portraying pathways. By constructing these graphical models which encompass a diverse range of biological pathways and concepts, it was found necessary to refine the Edinburgh Pathway Notation (EPN) scheme as previously proposed [102, 105]. The current 'modified' EPN (mEPN) scheme has been arrived at through extensive use and testing and is described in Chapter-3 and at http://www.mepn-pathway.org/ [116].

The macrophage pathway models presented here, explores some of the challenges associated with meeting the various demands of a pathway diagram. The hope is the integrated pathway will prove to be a useful resource for macrophage biologists, as well as an important contribution to the debate on pathway notation and depiction.

Results

Phase 1: Evolution of pathway construction methodology

When embarking on the task of pathway construction for the first time it was necessary to define a standard procedure for doing so. In fact the most time consuming aspect of the process was establishing the methodology of how to construct the pathway. Some of the crucial choices to be made were; deciding on the 'best' software for the task; how to arrange pathway components; how to improve the aesthetics of the pathway to make it more readable; and how to adapt the notation where needed to better reflect the biology of the interactions; how to record the interaction data of the pathways. The phases of the early pathway construction efforts are summarized in the workflow chart (Figure 2.1)

Pathway Construction Phases

(1) Establish means (literature search tools/ database/ software) for building pathways

⬇

(2) Explore notation and its rules for depicting interactions

⬇

(3) Construct pathway and refine notation and rules where necessary

⬇

(4) Improve aesthetics of pathway and assess usability

⬇

(5) Define set of rules for constructing pathway from experience in (1) - (4)

⬇

(6) Construct new pathways using refined notation and rules

Figure 2.1: A workflow diagram summarizing the early stage approach taken to assemble pathway diagrams. This process aids the establishment of rules for constructing further pathways.

Some time was spent in exploring software options for constructing pathways. Initially pathways were constructed using a relatively new academic software; Edinburgh Pathway Editor (EPE) (http://epe.sourceforge.net/SourceForge/EPE.html). However at the time of use the EPE was still in its early phase of development and lacked a user friendly GUI (graphical user interface), required extensive use and testing to resolve software 'bugs' and had significant stability issues. There were also limitations on the spectrum of graphics available and the editing capability (such as changing size and colours). Essentially it would take some time before this software is capable of supporting the scale (and quality) of pathways we were hoping to construct. An early draft of an apoptosis pathway generated in EPE (Figure 2.2a), exemplifies the issues with visualising diagrams constructed in this editor and using the early draft of the EPN. Efforts were made to improve the 'readability' of the pathway (Figure 2.2b) for example by increasing text to the maximal size permitted by the software, standardising component sizes, and adding colour to reflect the type of component. However even this version appeared cluttered and complicated and the software still limited the graphical rendering of biological concepts. One such example being, sub-cellular compartments could only be shown as blocks extending the width of the pathway. For this reason compartments such as the mitochondrion would be placed between the cytoplasm and nucleus. Therefore when components were shuttling between the cytoplasm and nucleus, edges defining this movement would also extend through the mitochondrion and in this way the translocation events of pathway components could easily be open to misinterpretation.
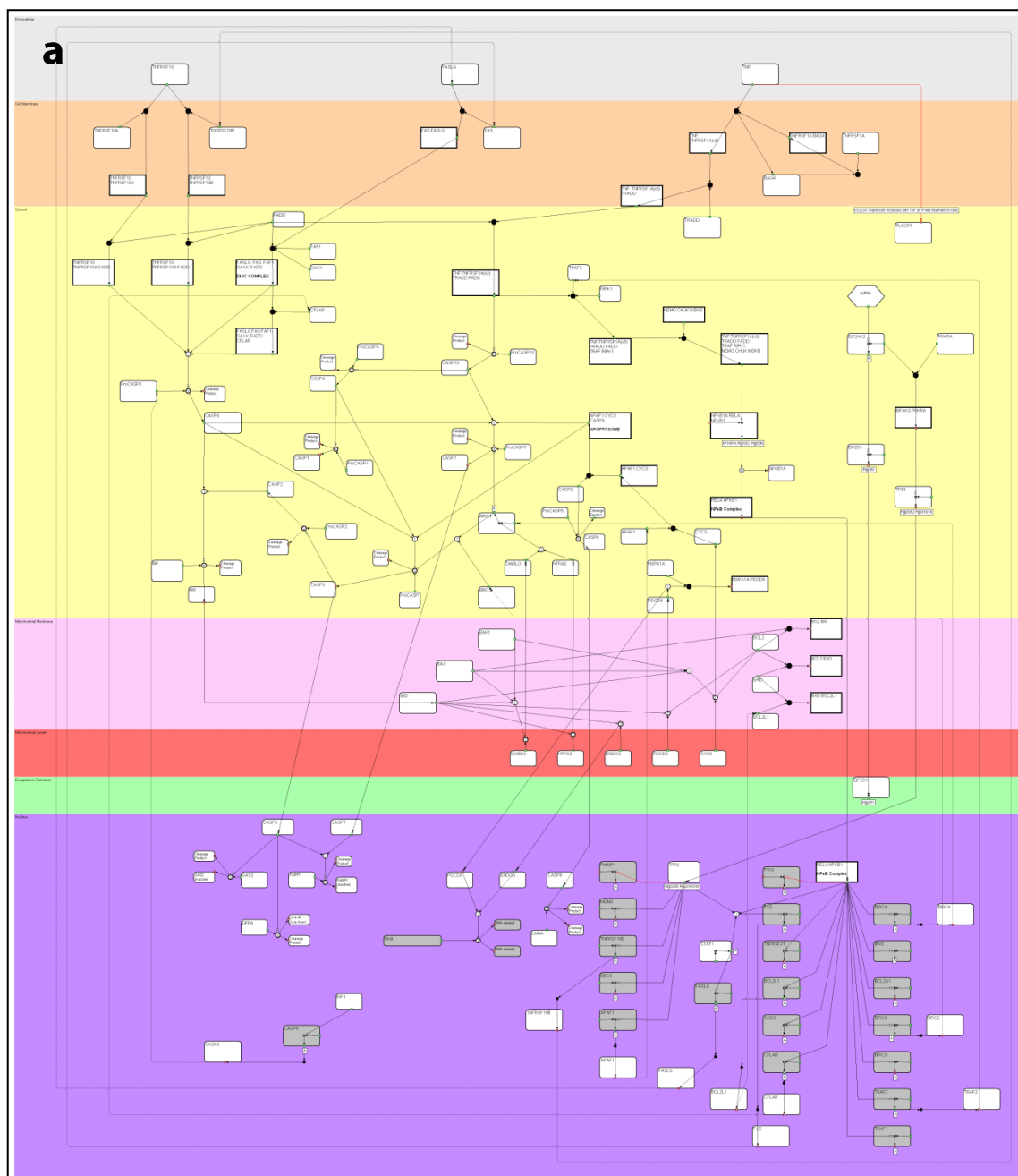
Figure 2.2: (a) Draft apoptosis pathway in Edinburgh Pathway Editor (EPE). Initial draft of the apoptosis pathway constructed using the original Edinburgh Pathway Notation (EPN) scheme and Edinburgh Pathway Editor software. The pathway runs from top (extracellular) to bottom (nuclear) with other sub-cellular compartments shown as the background layers. Text, symbols, and arrowheads were difficult to visualise in this version.
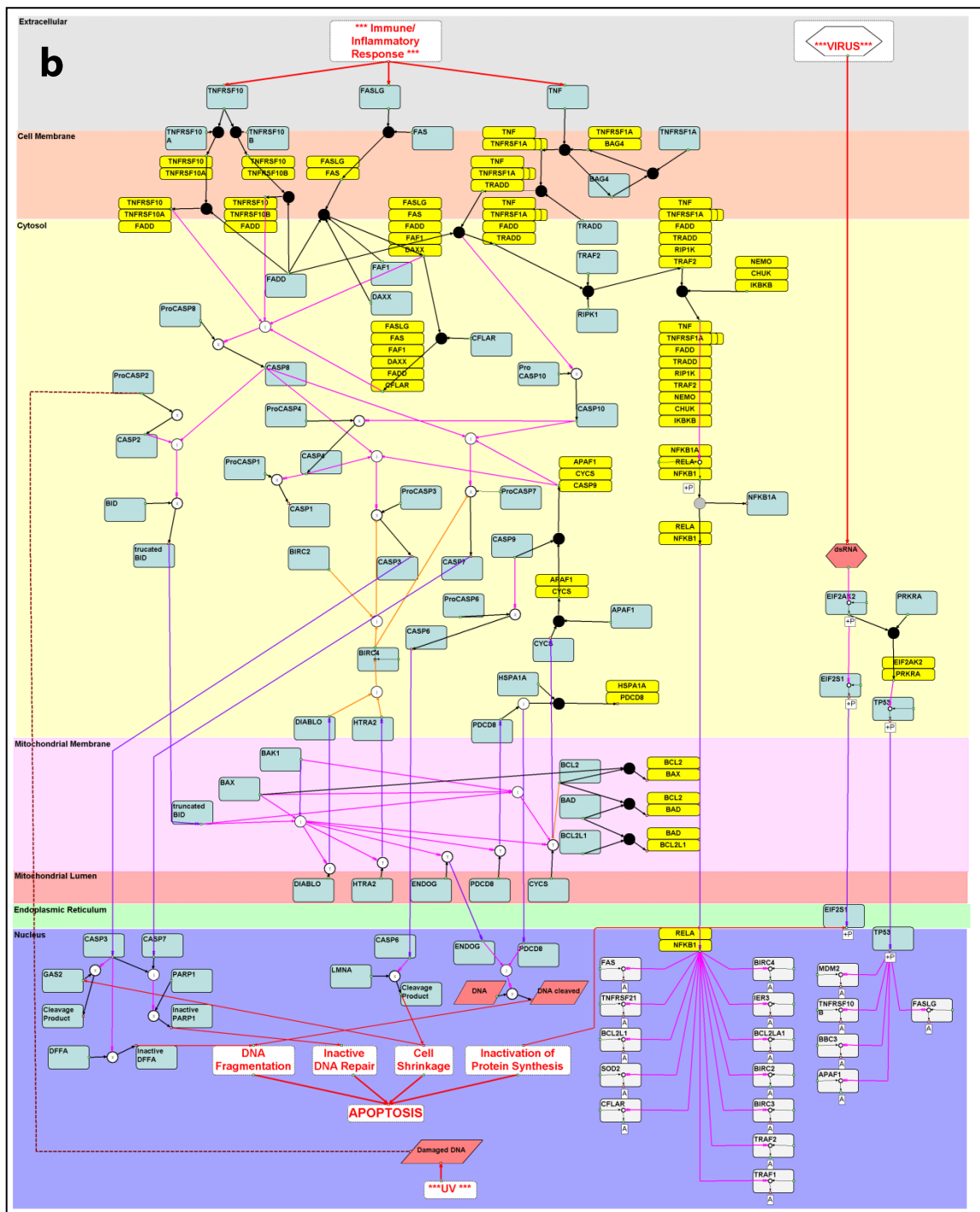
Figure 2.2: (b) An edited version of the draft apoptosis pathway in 2.2a. Text has been increased to maximal permitted size by the EPE software, as are notation symbols. Component sizes have been standardised and the use of colour is explored as a visual cue to assist with differentiation of components and interaction edges. Here proteins are coloured blue and protein complexes yellow. Additional textual annotation is added to add clarity as to where the end-point and process of apoptosis is executed.

Other network and pathway editing software including Cytoscape (www.cytoscape.org/) and Cell Designer (www.celldesigner.org/) were considered for the task but also suffered from a number of the following; limited graphics or restrictive support for notation schemes, limitations on the size of networks the programmes could support, unintuitive or complicated GUI's and finally at the time of starting the pathway construction (2007) these software did not support automated layout of networks and edge routing (a very useful functionality, especially for large networks). Microsoft PowerPoint was also explored although for improving visual aesthetics of the pathway and was not appropriate as a pathway editing tool. Figure 2.3 displays the apoptosis pathway illustrated using PowerPoint. This version demonstrated the pathways may benefit from running left to right rather than top to bottom and the display of compartments within compartments.
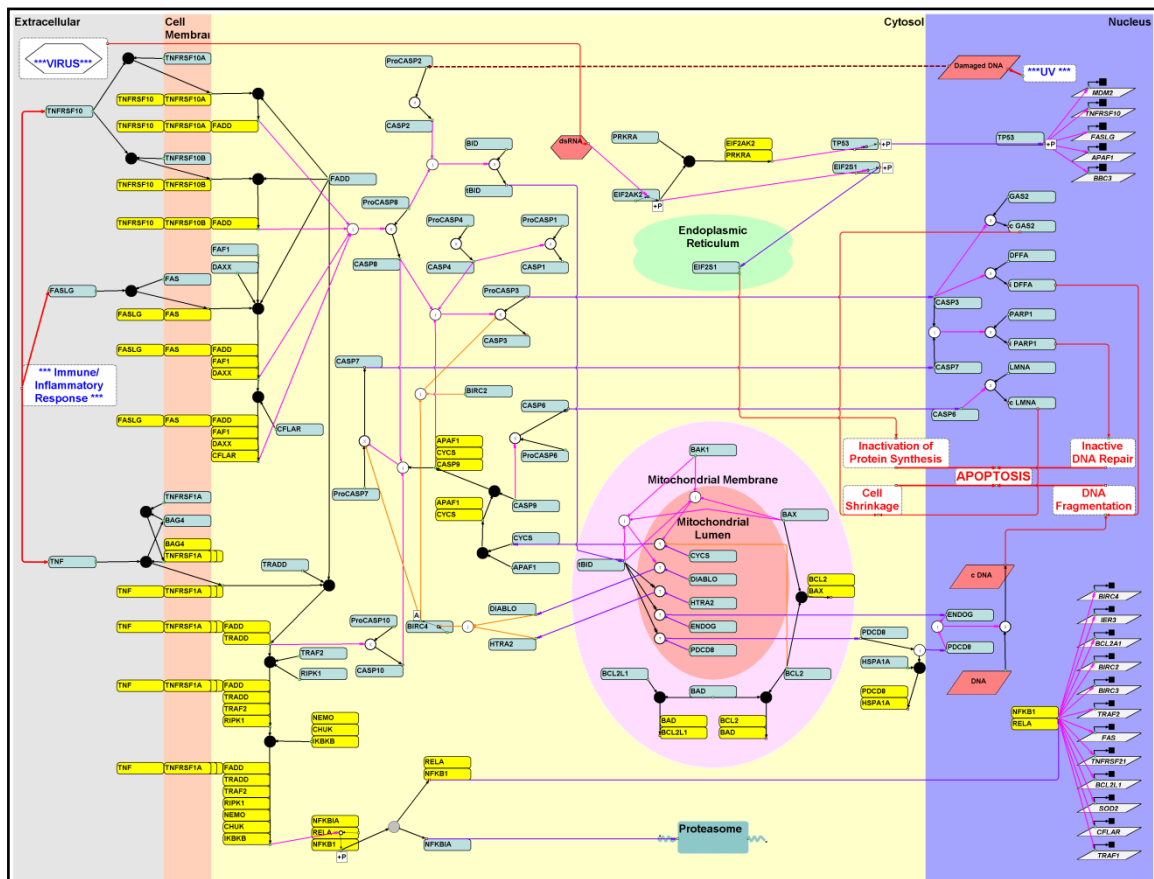
Figure 2.3: The apoptosis pathway assembled in Microsoft PowerPoint. The pathway runs from left (extracellular) to right (nucleus) with some compartments displayed as entities within other compartments, e.g. The mitochondrion and endoplasmic reticulum are placed within the cytoplasm. This arrangement allows for pathways to be more compact. Components can be aligned and arranged more precisely in this software than EPE. Although the visual aesthetics of the pathway have arguably improved in this version, PowerPoint is not a suitable pathway editing tool.

Eventually yEd (www.yworks.com), a Java2D generic graph editor was chosen for generating the pathway diagrams. The software was suited to the task for several reasons; it is available as a free download with no restrictions on its use, it is stable and 'bug' tested prior to release of new versions, has an intuitive interface with exceptional navigation, visualisation, multiple import/export possibilities and supports large graphs. Additional features which made yEd particularly appealing as a pathway editor included its integrated data storage, various network presentation possibilities and numerous automated layout algorithms to assist with the arrangement of a networks nodes and edges. An extract of the apoptosis pathway generated using yEd (Figure 2.4) exemplifies the progress made in creating readable pathways (cf., initial pathways Figure 2.2a and b) since the start of the task.
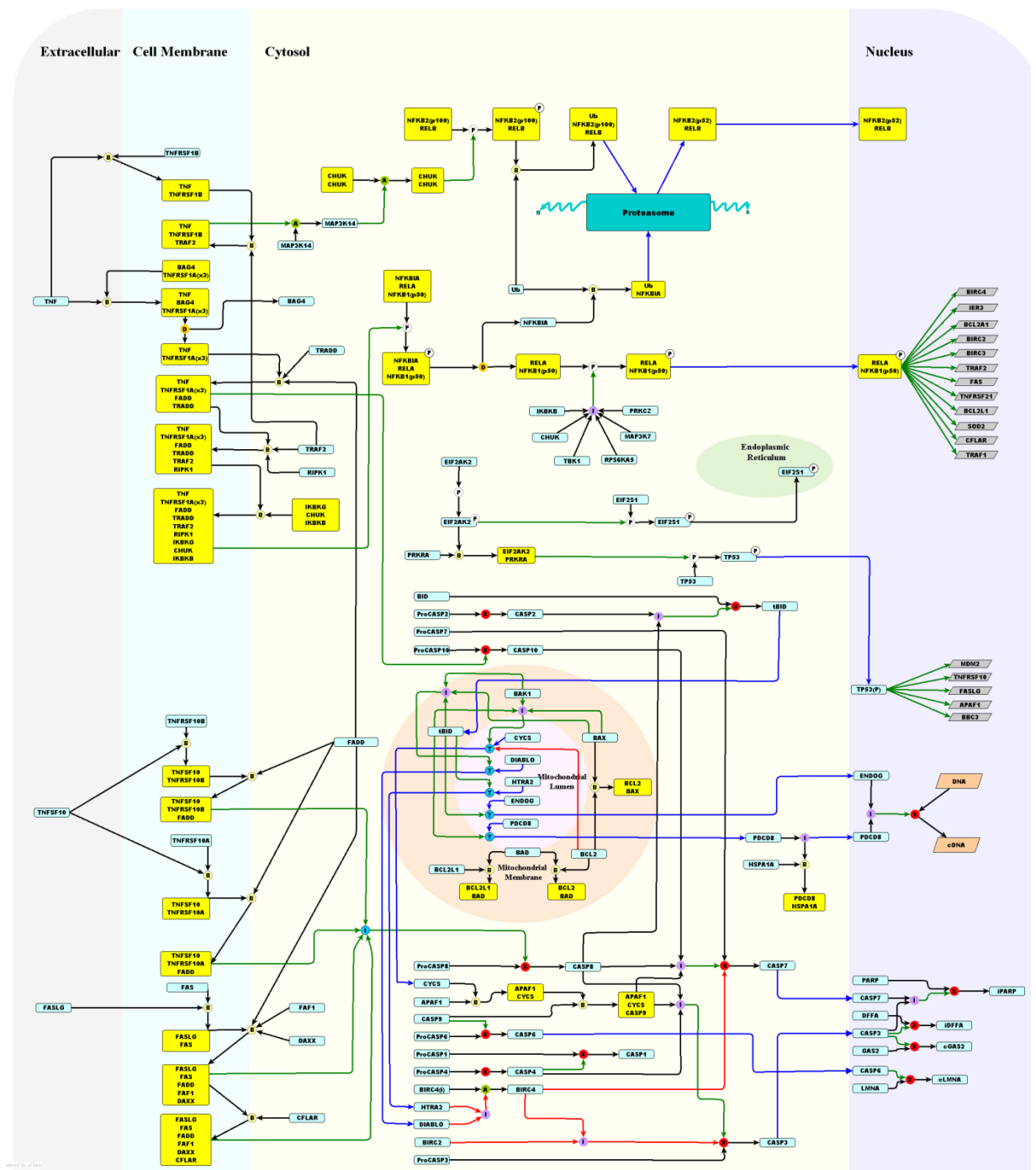
Figure 2.4: Apoptosis signalling pathway constructed using the yEd graph editor. The pathway is arranged to flow from left to right. Components are coloured according to type (protein, complex or gene) and arranged within the sub-cellular compartments in which they are active. yEd grap editor serves as a pathway editing tool and has exceptional graphics for producing clear diagrams. This pathway can be initiated by both intrinsic and extrinsic signals.

Apart from helping to establish the best choice of software the process of re-constructing the apoptosis pathway several times helped to uncover other obstacles with pathway construction that required addressing. Crucially the evolving pathways demonstrated how basic changes to the aesthetics of the pathway could drastically improve its readability. The process of data collection for the apoptosis pathway raised a number of questions as to where to obtain 'reliable' sources of information of collecting interactions of the pathway. Extensive group discussions explored which repositories would be an acceptable source for pathway interaction. With a plethora of interaction data now available, ranging from established published interactions, predicted interactions based on functional associations, hits from yeast-2-hybrid screens, it was key to define how 'inclusive' the pathways should be. Eventually it was agreed that only peer-reviewed published evidence, that explicitly cites an interaction would be used for the pathways. Although the pathways are designed to be a consensus of knowledge, if all sources of interaction data are included, integrity of the pathways is compromised and there would be the danger of creating an interaction network, rather than a pathways with a clear beginning and output.

Another issue arising whilst collecting interaction data, was that a given protein may have several alias names, for example NFKB1 is also referred to as p50 in the literature. In terms of presenting pathway components it was therefore deemed essential to identify whether a protein/gene has already been captured under another alias name. It soon became obvious to us that a standard nomenclature scheme should be adopted to label pathway components. HGNC were chosen for this, given the potential application of the pathway for overlaying results of transcriptomics data.

To summarize the main principles established during the task of constructing an apoptosis pathway were;

  (1) Pathways can be constructed to an acceptable standard using the freely available graphing tool yEd graph editor.

(2) To improve aesthetics of the pathway and in turn its readability; sizes of pathway components and text should be standardised.

(3) Used carefully, the addition of colour can enhance the 'readability' of the pathways

(4) Component layout should be performed manually or with the assistance of layout algorithms, however components can only be placed in their site of cellular activity, represented as predetermined areas (compartments) on the canvas.

(5) A conventional naming scheme should be adopted for labelling pathway components to avoid ambiguity of what is being represented; in this case HGNC was chosen.

(6) No interaction may be included within the pathway without published evidence. More than one paper may be used to support the same interaction (two or more is preferable).

(7) Evidence of an interaction between one component and another should be stored in an interaction table. Evidence to support an interaction should be derived from the primary literature (and reviews). This must include the interacting partners, the direction of the interaction is inferred by order HGNC1 -> HGNC2, the type of interaction (phosphorylation, cleavage), method, PubMed ID, site of specific change of state [P-Ser123].

These principles now form the basis for assembling the pathway diagrams within the group and are described in more detail in the methods section of this Chapter. The process of 'road-testing' apoptosis pathway construction was crucial in defining some standard methodology for the task. However not all issues with pathway construction were identified at this stage. In practise some issues would only come to surface with greater experience of building pathways; for example when trying to model new biological concepts or when creating larger pathway diagrams or when combining diagrams of multiple but associated pathways. The next stage of the pathway mapping efforts was to model other pathways of interest in macrophage biology and combine them into one diagram by identifying links between these pathways.

Phase 2. Construction of a framework diagram of macrophage activation pathways

The TLR (Toll Like Receptor), type-I and type-II IFN (interferon) signalling, NF-κB and apoptosis pathways are of central importance in defining the macrophages response to pathogens and do so in a highly inter-dependant manner [160]. In an effort to describe and consolidate knowledge of these signalling events, I set out to construct a pathway diagram based on published literature and using the EPN. Given the extensive cross talk between these pathways, they were combined into one integrated diagram. This initial framework map of macrophage activation [102] (Figure 2.5) a total of 295 nodes of which 140 are proteins, 99 complexes, 44 genes, and 12 other components (pathogens, DNA, RNA etc). A total of 272 interactions are described in the pathway map, of these 85 are binding events, 149 are various activation state modulations (67 activation of gene expression, 26 phosphorylation, 7 auto-phosphorylation, 1 dephosphoylation, 23 cleavage, 9 translocations and 16 activation by processes not defined). There are 10 inhibition reactions, 4 of these are inhibition of gene expression, 3 are inhibition of cleavage, and 1 is an inhibition of translocation. A total of 26 translocation events occur as well as 2 protein dissociations. 282 different references support the interactions shown on the pathway. In many circumstances the same paper may describe multiple interactions, for example Chaudhary *et al.,* (1997) report that both TNFRSF10A and TNFRSF10B recruit the protein FADD during apoptosis signalling [161].
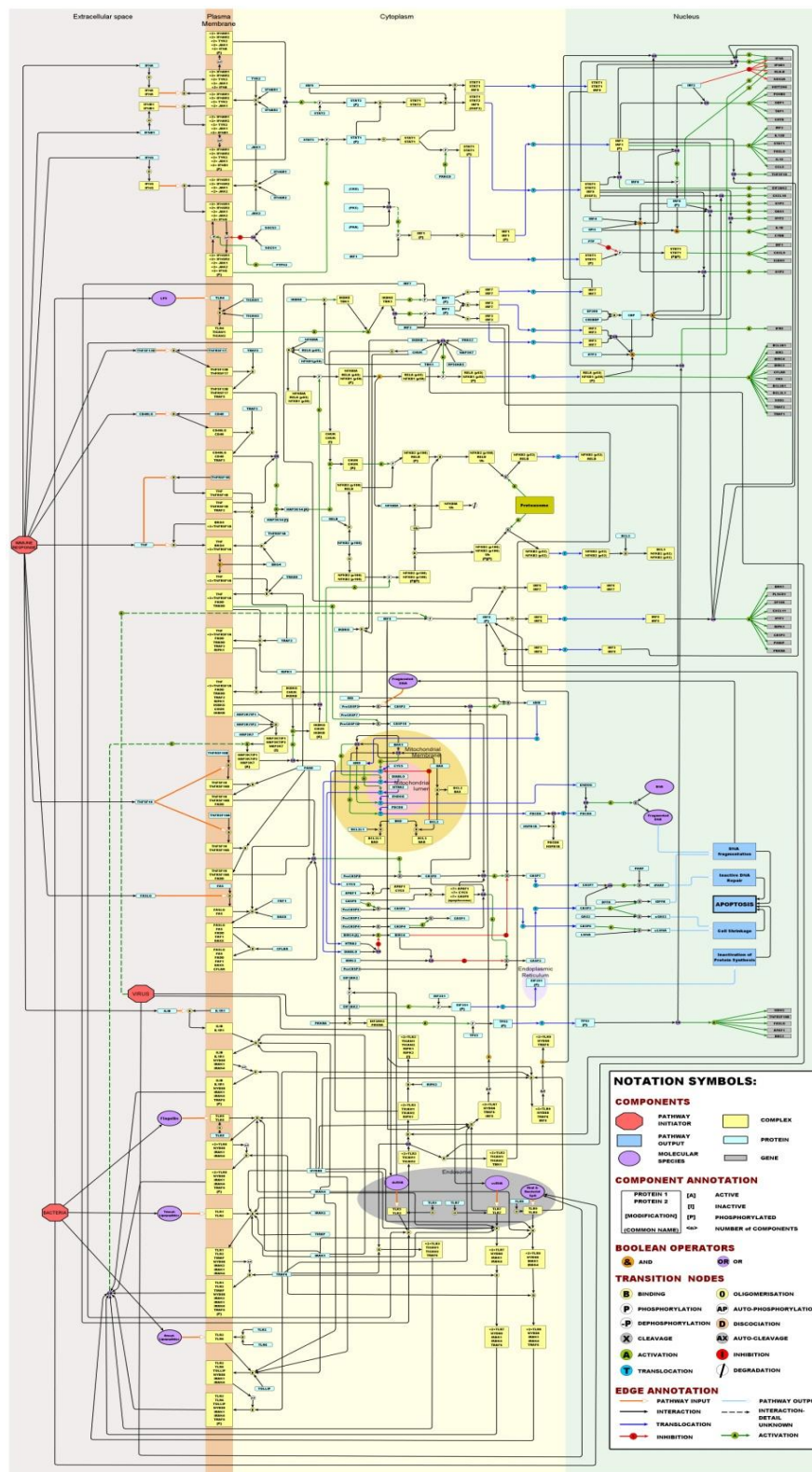
Figure 2.5: Framework integrated pathway map of signalling in the macrophage. The diagram includes the interferon signalling, NF-κB, apoptosis and toll-like receptor pathways, all represented as one integrated pathway due to their overlapping interactions. In general interactions of the interferon response pathway are in the top quarter of the map, with NF-κB directly below. Apoptosis is presented halfway down the map and toll-like receptor signalling is in the bottom quarter. 154 different protein or gene nodes are included in the pathway, along with 80 different complexes and 12 other molecular species (such as pathogens, DNA, RNA). The pathway diagram represents 272 different interactions.

The four signalling pathways were combined into one diagram since there is extensive crosstalk between them. Therefore, in order to check the integrity of the network each input (e.g. cytokine or pathogen molecule), was highlighted in turn and the logical flow of information from this input followed through the diagram. By following the flow of information from each pathway input, a different but expected output was observed, be that the activation of transcription or a process such as apoptosis (Figure 2.6). This suggests that although several signalling pathways have been integrated to form this diagram the specificity of connectivity has not been lost.
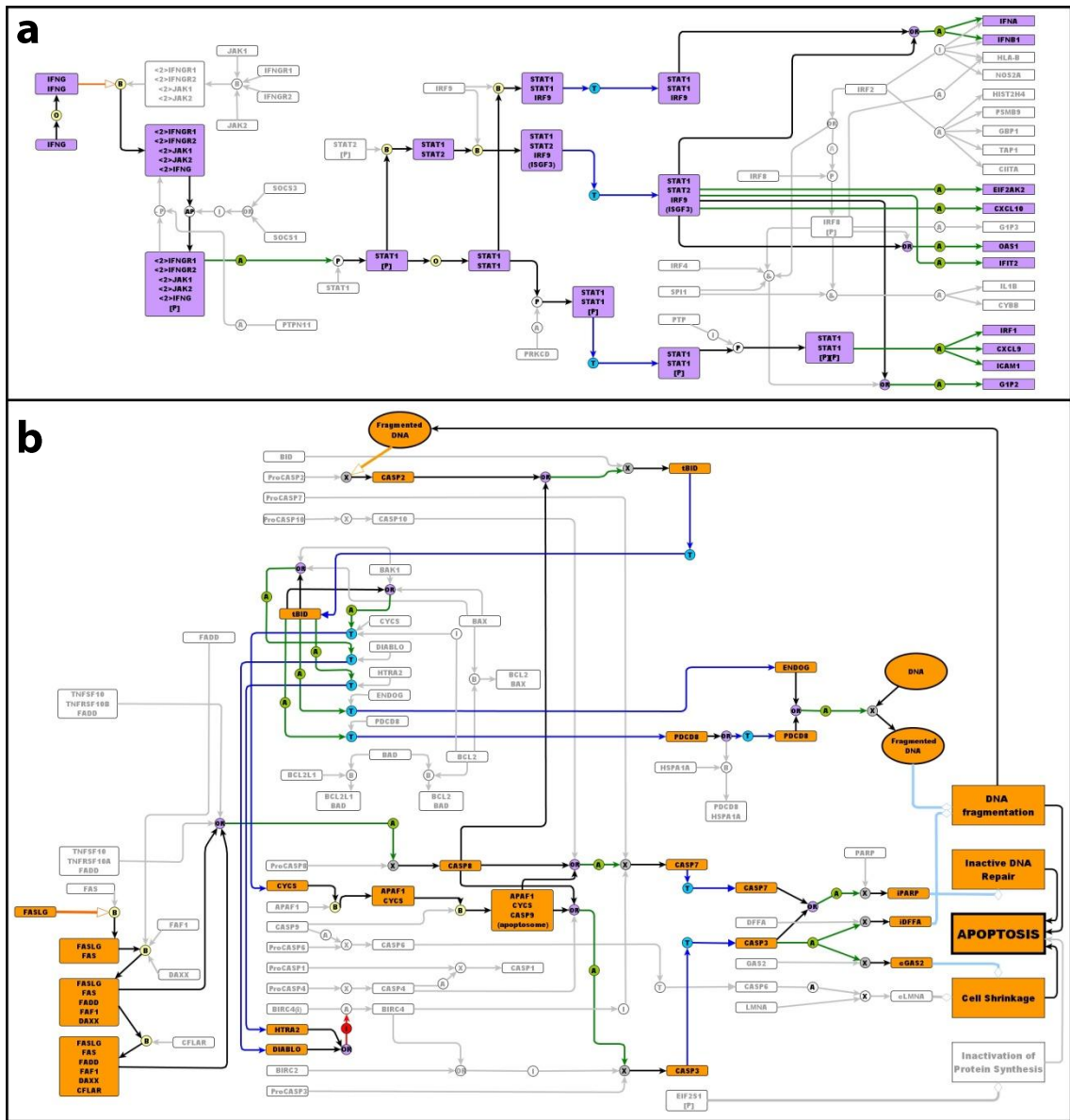
Figure 2.6: Follow through of signalling pathways stimulated by IFNG (a) and FASLG (b). The signalling events following the input signals of IFNG and FASLG have been highlighted on the entire map in lilac and orange, respectively. The nodes activated or directly affected by FASLG or IFN-gamma binding to their receptors are coloured and the interaction edges and gates are also highlighted. Nodes and edges not directly downstream of the FASLG or IFNG signalling are shown in grey. This figure demonstrates inputs into the pathway can clearly be followed to the expected outcome events. In the case of IFNG-input, gene transcription is the resulting event, and in the case of FASLG, apoptosis. Furthermore these examples clearly depict the interactions of the pathway can be followed logically and do not result in unexpected crosstalk.

Use of the Framework Pathway Diagram in the Interpretation of Transcriptomics Data

In order to explore the utility of this framework pathway diagram in the interpretation of transcriptomics data, the transcriptional events following the treatment of mouse bone marrow derived macrophages (BMDMs) with interferon-gamma (IFN-γ) were examined (This data set is discussed in more detail in Chapter-4). Using the network analysis tool BioLayout *Express*[3D] [162] a 3-D network of transcripts identified as being differentially expressed following IFN-γ stimulation was constructed. 1,491 transcripts were represented within the network. Of the 154 unique proteins/genes represented on the pathway map, 58 of were represented within the transcriptional network. All of the genes represented on the map were up-regulated in the data set. None of the transcripts down-regulated in response to IFN-γ were present on the map. Clusters of transcripts representing genes activated at different times following treatment were then further collated into 3 groups of up-regulated genes; genes activated at (1) 1-2 hours, (2) 2-4 hours and (3) 4-8 hours post-treatment. Genes that were activated and included in the set of mapped genes were then highlighted on the map and the possible downstream consequences (assuming *de novo* protein synthesis and activity following an increase in gene transcription) were highlighted (Figure 2.7). In this way it has been possible for the first time to interpret these transcriptional events in the context of the possible consequences of these observations.

Using this data overlay approach onto the pathway it was possible to extrapolate some interesting observations by visualizing the changes and the possible downstream effect of the changes. It was also possible to appreciate the connectivity and co-dependency of the changes over time. Although the framework pathway at this stage did show much promise of being a useful resource for data interpretation the group was acutely aware that in its current form the diagram covered only a limited snapshot of activity in the macrophage. For instance only a relatively small number of the genes shown to be transcriptionally regulated following IFN-γ treatment were present on the diagram. Furthermore none of the genes shown to be down-regulated by IFN-γ are shown in the diagram.
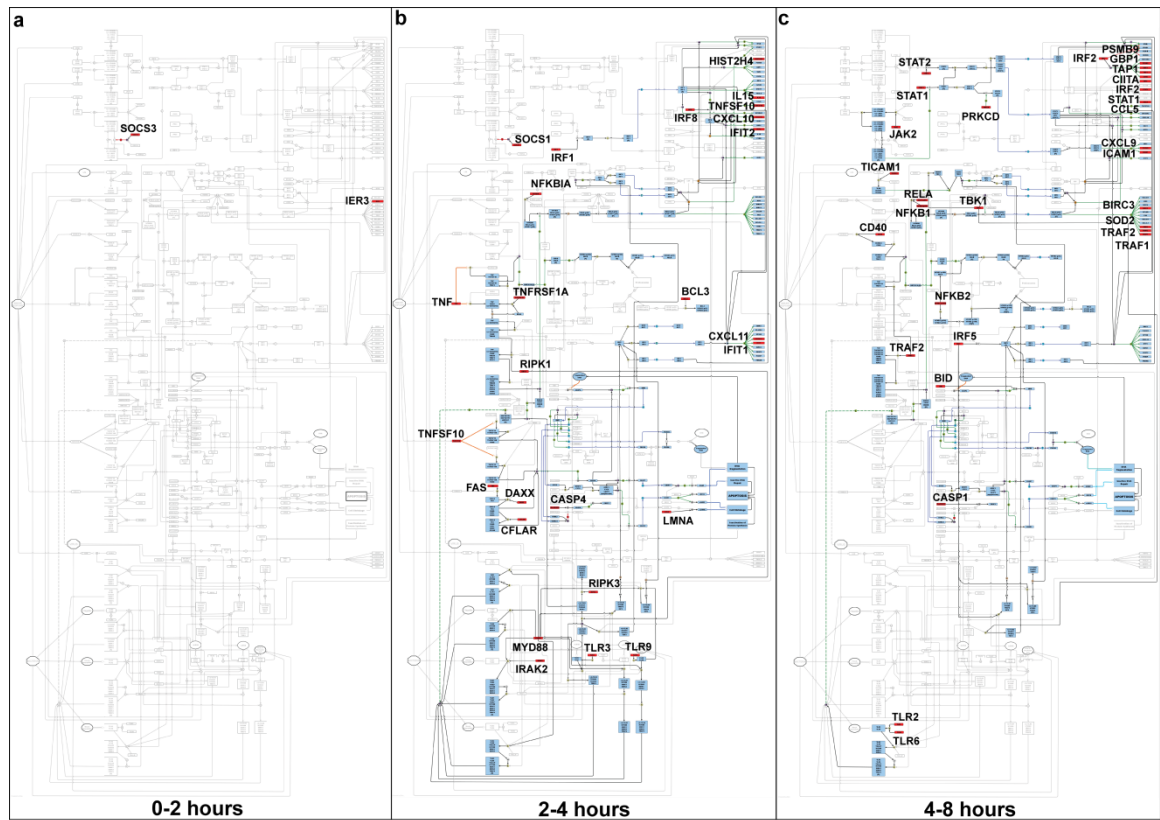
Figure 2.7: The integrated framework pathway diagram presented at (a) 1–2 hours, (b) 2–4 hours and (c) 4–8 hours post-IFN-γ treatment. Differentially expressed genes are highlighted in red and the possible consequential downstream events resulting from the changes, (assuming de novo protein synthesis) are highlighted in blue.

The exercise of constructing the framework pathway was essential in order to define a process for doing so and to explore how best to graphically depict complex biological systems. During Phase 2 of this exercise the resulting pathway proved to be a valuable resource to the group and for others with an interest in macrophage signalling events [146, 163]. It was possible to demonstrate that despite the integration of several pathways the diagram still maintained integrity as to the biology being portrayed (Figure 2.6). Furthermore, the pathway showed potential for being useful in the interpretation of transcriptomics data. However, it was apparent that the pathway provides just a snapshot of the processes active in the macrophage, as was evident from the IFN-γ-transcriptional response analysis, where only 58 of the 1,491 regulated transcript were represented on the map. As we explored more transcriptomics analysis performed on these cells it was apparent that a whole range of pathways regulated by different macrophage stimuli were absent from the diagram. It was also observed that the TLR representation was more simplistic than current knowledge, for example links with MAPKinase signalling were absent. Also pathways key to macrophage function, such as antigen presentation were not presented.

In order to obtain a more comprehensive resource on macrophage signalling and to assist with the interpretation of genomics data derived from this cell, additional pathways of interest were assigned to a cohort of MSc students registered for an MSc in Genomics and Pathway Biology at the Division of Pathway Medicine, who were presented with the task of assembling the interaction networks of the following; detailed views of the TLR pathways, other pathogen recognition systems (RIG's, NOD's NALPS's) , extensive view of NF-kappa-B signalling, MAP-kinase cascades, MHC antigen presentation and proteasome assembly. A workflow for constructing pathways was established from the lessons learnt from constructing the original framework diagram as summarized in the schematic Figure 2.8. The pathways were then combined along with the framework diagram into one integrated network. I supervised the reconstruction of the TLR system and related pathways, oversaw the integration of the separate pathways into one network diagram, and participated in regular discussions

with other pathway constructers to assess the progress of the pathways and addressed any issues that pertained to the graphical portrayal of specific biological concepts.
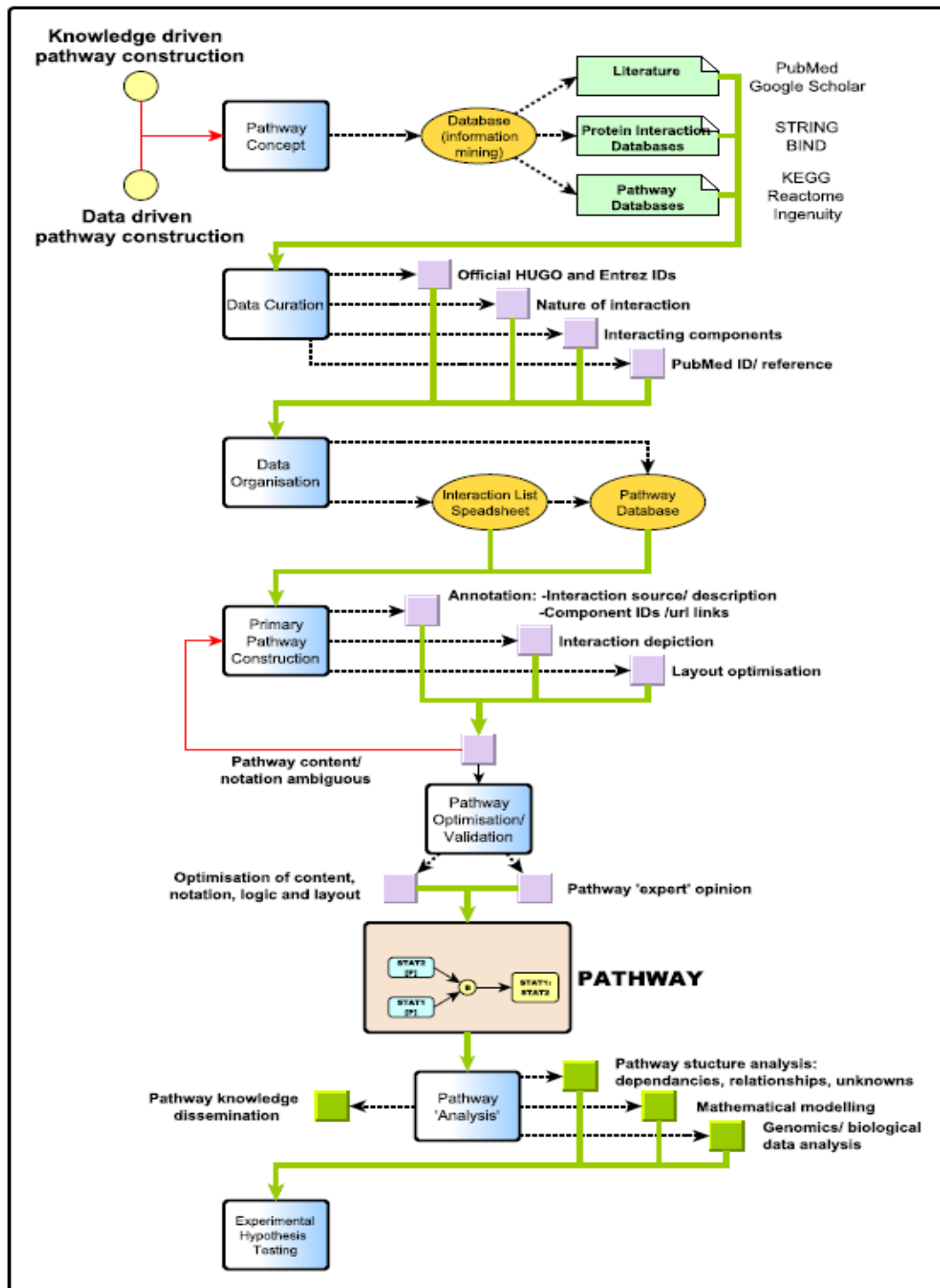
Figure 2.8: Pathway construction workflow. A workflow diagram summarizing the main stages of pathway assembly from concept to final diagram. Blue boxes portray the pathway construction phase. Each phase embodies a number of tasks (shown as lilac boxes or yellow-ellipses for data storage and processing), which when completed lead to progression towards the next stage of pathway construction (connected by green arrows). Red arrows indicate feedback to a previous construction phase. Lilac boxes describe the construction steps required pre-pathway assembly, whereas green boxes are linked to post-construction phases and describe the possible applications of the constructed and validated pathway diagram.

Phase 3. A large scale integrated map of macrophage pathogen recognition and effector systems

The work described in this section followed in the footsteps of earlier efforts to construct process notation diagrams of macrophage pathways [99-100, 159]. In the course of phase 3, solutions were sought to issues associated with the depiction of a variety of different biological systems, combining diagrams from multiple curators and the layout and integration of a large network model of these systems. Teams of 1-2 junior biologists (MSc students) were given a remit to describe a given pathway system using the mEPN scheme [116]. The signalling systems were chosen based on their significance to macrophage biology and/or the group's interest in interpreting the results of tanscriptomics studies. Macrophages are professional antigen presenting cells, however these pathways were missing entirely from the framework diagram. Therefore antigen processing and presentation was assigned as one remit. Closely linked to this pathway was unbiquitination and proteasome assembly, which also appeared to be regulated transcriptionally in response to a number of macrophage stimuli, however very few detailed pathways of these systems existed in the public domain. The framework diagram also lacked coverage of non-TLR-pathogen recognition systems, and the existing TLR system required further expansion, including links with other pathways such as MAPkinase signalling. The group also had a desire to contextualise gene expression results showing many members of the NF-κB family to be regulated, however the framework diagram only incorporated the most well characterised NF-κB members and pathways. Therefore an expansion of the NF-κB systems was undertaken and the results of this labour highlighted how simplistic the original views of this signalling system were. Once all individual pathways had been assembled by individual teams of curators, they underwent extensive editing in attempt to unify their notation usage, stylistic qualities and overall appearance. The original pathway diagram [102] was then used as a framework on to which the new pathways were connected. The final product of this labour provides what is to date the most comprehensive model of macrophage signalling in the public domain and one of the largest examples of a pathway constructed using a formalised graphical notation scheme.

Integrated Pathway Diagram

The pathway diagram presented here (Figure 2.9, also available at http://www.macrophages.com/macrophage-pathway-resources) is a consensus view of a number of pathway modules assembled based on the interpretation of the literature describing these systems. A given interaction between components of the pathway may be supported by evidence derived from one or more publications and a publication may provide evidence supporting more than one interaction. A total of 1,000 different interactions have been recorded, supported by 728 different original papers and reviews (the full list of interactions and supporting publications at http://www.macrophages.com/macrophage-pathway-resources). The network diagram is comprised of 2,170 nodes connected by 2,553 edges. The diameter of this network (maximum distance from one node to another) is 58 and there is an average node connectivity (number of inputs/outputs) of 2.37 (max 37). A detailed breakdown of the class (type) of the nodes that make up the diagram is shown in Figure 2.10. Briefly, 496 unique proteins are represented in the diagram many of which are shown to go on to be modified into different forms or bind together resulting in 412 different complexes. 81 genes are shown to be transcriptionally regulated by these pathways based on known associations between transcription factors and target genes. The interactions between these components are represented by 552 process nodes, 120 Boolean logic operators and 158 edge annotations. The pathway is drawn using the mEPN scheme, a full description of which can be found in Chapter-3 and on www.mepn-pathway.org and in Freeman *et al.,* 2010 [116]. The macrophage-related pathways are also available through www.macrophages.com.
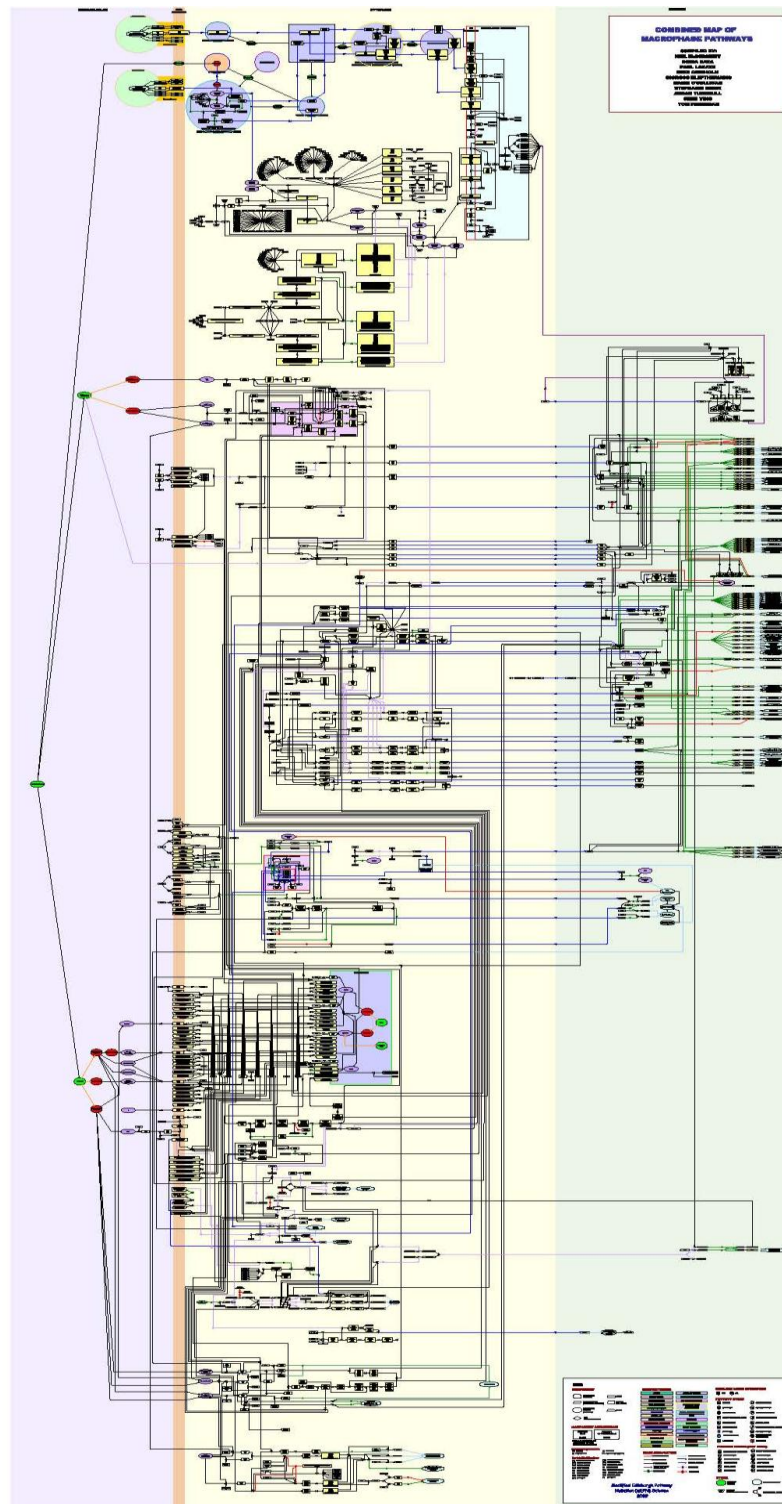
Figure 2.9: Integrated pathway diagram of innate immune and macrophage activation pathways.
The modified Edinburgh Pathway Notation (mEPN) scheme is used to describe the interactions of signalling pathways active in the macrophage. A total of 2,170 components in this network are connected by 2,553 edges. Components include 496 unique proteins, the complexes formed between them (412), 181 genes/ DNA/ promotor regions, in addition to other molecular species (e.g. pathogens, drugs, RNA) and the nodes representing the processes in which the components are involved. Components are arranged to reflect the location in which they are active and background colour is used to distinguish between different sub-cellular locations.

Description of the Biological Content of the Integrated Pathway Diagram

The pathway diagram presented (Figure 2.9 and http://www.macrophages.com/macrophage-pathway-resources) incorporates detailed views of some of the best characterized pathways associated with macrophage-specific biology, as well as some that are generic but in some way linked to the activity of these cells. Figure 2.10b shows the approximate location of the different pathway modules within the diagram. The size of the diagram requires it to be ideally viewed on a computer. Every effort was made to arrange modules so those with shared nodes and high connectivity are located in close proximity, however given the issues in depicting information on this scale, arriving at an 'ideal' arrangement of components is challenging. The interactions between components and limited views of the pathways they form have been described in detail in the literature used to construct this diagram (the interaction table, is available at http://www.macrophages.com/macrophage-pathway-resources). An overview of the pathway biology depicted by the diagram follows.

a   Pathway Node Count

Node Category	Node Type	Total
Component	Cell	3
	Compartment	21 (18 unique)
	Protein	576 (496 unique)
	Protein-Complex	412
	Gene	81
	DNA sequence/ Promoter	101
	Molecular Species/ Generic	45
	Simple-Molecule	2
	Drug	4
	Energy/ Molecular Transfer	6
	Module Pathway Input/Output	91
Boolean Logic	&	28
	OR	92
Process Node	Activation (A)	11
	Auto-Catalysis (AC)	1
	Auto-Phosphorylation (AP)	29
	Auto-Cleavage (AX)	8
	Binding (B)	222
	Dissociation (D)	45
	Inhibition (I)	2
	Oligomerisation (O)	25
	Phosphorylation (P)	63
	De-Phosphorylation (-P)	3
	Sumoylation (Su)	1
	Translocation (T)	67
	Transcription/Translation (TL)	3
	Ubiquitination (Ub)	26
	Cleavage (X)	33
	Unknown (?)	1
	Sink (Proteasomal Degradation) (/)	12
Edge Annotation	Catalysis	66
	Activation	77
	Inhibition	15
Grand Total		**2172**

1000 different interactions are displayed on the pathway. The interactions are supported by 2009 sources of evidence cited from 728 different peer-reviewed publications.

b

- MHC antigen presentation
- Ubiquitination
- Proteasome
- RIG receptors
- Interferon signalling
- NF-KB
- Apoptosis
- TLR signalling
- MAPKinase
- PI3Kinase
- β-catenin
- Non-TLR signalling
- Inflammasome

Figure 2.10 (a) Breakdown of node class in the integrated pathway diagram and (b) Key to the pathway layout and content of the integrated diagram. (a) Components, Boolean Logic, Process Nodes and Edge Annotation form the category of possible nodes. A detailed breakdown of the number of each type of node in each category is given. (b) The key reflects the approximate location of the different pathway modules depicted in the integrated diagram (Figure 2.9). Ideally pathways with high connectivity and sharing identical components are spatially located in close proximity.

71

Macrophage Pathogen Receptor Systems

Macrophages are equipped with a complex array of pattern recognition receptors (PRRs) that bind a varied assortment of pathogen-associated ligands. Perhaps the best studied of these are the membrane-associated toll-like receptors present on the cell's surface (TLRs 1/2/4/5/6/10) and lining their endosomal compartments (TLRs 3/7/8/9) [164-165]. The receptors commonly form complexes comprised of 6-8 protein subunits which undergo a series of phosphorylation, dissociation and binding events following engagement of the receptor with their respective ligand class. As with all such receptor-ligand interactions shown in the diagram, each successive stage in the formation and activation of the receptor complex is explicitly shown (Figure 2.11 a&b). TLR's are comprised of two functionally significant domains; one for recognizing specific pathogen associated molecular patterns (PAMPs) and one for recruiting signalling adaptor proteins following binding of an appropriate pathogen-derived ligand. The pathogen recognition domains of different TLR's are structurally highly variable [166], thereby allowing the recognition of diverse pathogen-derived molecular species ranging from viral double- and single-stranded RNA [167-168] bacterial flagellin [169-170], lipopeptides [171-172], lipopolysaccharides [173-174], and bacterial and viral CpG motifs [175-176]. In contrast, the internal domains tend to be more conserved, reflecting the ability of different TLR's to recruit the same adaptor proteins; in particular MYD88, IRAK4, IRAK1, TOLLIP, TIFA, and TRAF6 are common to most of the TLR receptor complexes. The use of common adaptor proteins by many of the TLR complexes represented a significant challenge in depiction, with many edges emanating out of each adaptor molecule. Much effort was therefore put into layout of this system so as to provide visual clarity. A comprehensive and systematic effort to depict TLR signalling has been reported elsewhere [100] however this was not used in the construction of the current view of TLR signalling. These receptors ultimately activate a number of downstream signalling pathways including the NF-κB, IRF (interferon regulatory factor) [177] and MAPKinase, ERK, and JNK signalling [178].

Also represented here are 9 non-TLR cytoplasmic PRRs, including the NOD (nucleotide-binding oligomerization domain)-Like Receptors and RNA helicases, whose role it is to

detect endogenous stress signals and intracellular pathogens. NOD-like receptors (NLRs) can be broadly divided into various classes, the NODS, NALPS and other types of NLRs, based on their protein domain structures. Bacterial associated PAMPs (e.g. flagellin and various classes of peptidoglycans) are recognised by NOD1, NOD2 and NLRC4 (IPAF1) which activate the NF-κB and MAPKinase pathways. NOD1 can also lead to the cleavage of IL1B into its active form. The NALPS (NLRP1, NLRP2, and NLRP3) detect a range of stress signals such as K^+ efflux, DNA, ATP or membrane damage, often collectively referred to as danger-associated molecular patterns (DAMPs). Once activated by DAMPs the receptors form oligomers with inflammatory caspases (CASP1/5) and in doing so activate the cleavage of the caspases [179]. The active complexes are known as 'inflammasomes' owing to their ability to cleave and activate interleukin proteins, and are crucial mediators of the inflammatory response. Although three NALP receptors have been depicted up to 14 different NALPS have been reported [180-181]. Finally, RNA helicases are responsible for the intracellular recognition of viral single stranded and double stranded RNA. The diagram shows DDX58 (RIG-1) and IFIH1 (MDA5) which recruit factors via their CARD domains and eventually initiate anti-viral gene expression by activation of the NF-κB system. The ZBP1 protein was recently characterized as a sensor of viral DNA [182] and activates the IRF3 transcriptional pathway. As such the PRR systems depicted represent a comprehensive view of these receptors and the signalling pathways they activate. However the diagram still lacks other known macrophage PRRs, including the surface mannose receptor, secreted receptors e.g. those belonging to the complement system and the recently described DNA receptor AIM2 [183-186].

Cytokine Activation Pathways

The diagram also describes a number of the main cytokine signalling systems active in macrophages. These include the interferon (type-I - IFNA/IFNB and type-II IFNG), interleukin 1B (IL1B), TNF, TNFSF10 (TRAIL), TNFSF13B, FASLG and CD40LG signalling pathways. In each case these have been depicted starting from their interaction with their receptor complexes (Figure 2.11a) through to the activation of their downstream signalling and effector pathways. Interestingly, the expression of a number of these

ligands is activated by PRR pathways (e.g. IL1B, IFNB) producing autocrine feed-forward loops. *IFNB* is perhaps one of the best studied genes in the whole genome in terms of its transcriptional regulation [165, 187-189]. It is also one of the primary targets for a number of the macrophage PRR activation pathways described above and we therefore constructed a detailed model of its regulation (Figure 2.11h). Part of the reason behind this was also to grapple with the issues with depicting transcriptional networks and we believe the solution arrived at should work for other systems. In the current diagram however, we have generally chosen not to depict the links between cytokine gene activation (or indeed between other genes) and their respective proteins (via translation). The depiction of these edges adds to the visual complexity to the diagram. For modelling purposes however these connections can be added.

Apoptosis (Programmed cell death)

A potential output of the innate immune response is to culminate in host cell suicide (apoptosis) thereby potentially limiting further reproduction of pathogenic organisms such as viruses. Two major routes of apoptosis execution have been identified, termed the intrinsic and extrinsic pathways. The intrinsic pathway is activated as a result of stress signals detected within the cell, for example, penetration of a viral pathogen into the cell or UV light induced DNA damage. Extrinsic apoptosis on the other hand is triggered by extracellular death-signalling ligands (FAS, TNFSF10 (TRAIL), TNF which are also members of the cytokine activation pathways) binding to the cell membrane receptors. Both intrinsic and extrinsic pathways activate a number of the caspase family of cysteine proteases. The initial caspases to be activated are categorized as initiators, (CASP2/4/6/8/9/10) and are capable of cleaving downstream executioner caspases, specifically CASP3 and CASP7. Caspases 3 and 7 initiate the series of events that directly lead to the morphological changes in a cell associated with apoptosis by the cleavage or inactivation of an array of molecules including, structural proteins, DNA repair proteins, and anti-apoptotic proteins.

Signal Transduction and Transcription Factor Networks

The NF-κB (nuclear factor kappa-light-chain-enhancer of activated B cells) family of transcription factors are pivotal in the regulation of a wide variety of biological processes [190-193]. This includes many aspects of the innate and adaptive immune response, as well as the regulation of a diverse range of stress-related stimuli [194]. 5 different NF-κB proteins (REL (c-Rel), RELB (REL), RELA (p65), NFKB2 (p100 or p52), NFKB1 (p105 or p50) have been identified and these form a variety of homo- or hetero-dimers resulting in an array of different NF-κB complexes. Previously [102] the framework diagram described the activation of two of the best-characterized NF-κB dimers; NFKB1:RELA (also known as p50-p65) and NFKB2 (p52):RELB, often referred to as the canonical and non-canonical pathways, respectively. However, it was soon realized that this diagram was a rather naïve view of the NF-κB system and the students were assigned to explore the literature on this system in greater detail. These efforts have resulted in the depiction of 14 different NF-κB dimers formed from combinations of the five NF-κB proteins. In addition to the 14 dimers, some NF-κB complexes form further complexes with a number of accessory proteins (NFKBIA/B/E/Z, BCL3, HMGA1, CREBBP, HDAC3, NCOR2) and together with their regulation by multiple phosphorylation events, give great diversity in the form and control of this important class of transcription factors. This goes someway to explaining the pleiotropic effects regulated by this system [195]. It is unlikely however that all the possible NF-κB systems depicted are active in the macrophage or indeed any other single cell type. NF-κB signalling is often cited in loose terms in literature with little reference to the exact NF-κB complex active in any given situation. The pathway diagram presented here demonstrates the complexity of this system and underscores the need for acknowledging the range of possibilities beyond the canonical and non-canonical NF-κB pathways.

As mentioned above, phosphorylation is a key element in the activation process of many of the NF-κB complexes. In the pathway, dimers of the core proteins may bind to NFKBIA, NFKBIB or NFKBIE, a group of I-kappa-B or NF-κB inhibitor proteins. When bound to their inhibitor, the NF-κB complexes are restricted to the cytoplasm. Upon

stimulation the I-kappa-B proteins are phosphorylated, leading to their ubiquitination and eventual degradation, and the release of the active NF-κB complex. Dissociation of the inhibitors exposes the nuclear localization domain on the NF-κB complex causing it to translocate to the nucleus where it can modulate transcriptional activity of target genes [196-197]. Other NF-κB complexes (which are not bound to inhibitors) are activated following cleavage into smaller DNA-binding subunits. This is induced by stimuli phosphorylating the complex leading to its ubiquitination and subsequent processing of one or more of their subunits into smaller DNA binding peptides e.g. NFKB2 is processed from p100 to p52. Currently there are 34 genes shown on the diagram as the transcriptional targets of NF-κB signalling (Figure 2.12g). In reality this is only a small percentage of the known NF-κB targets [198]. Furthermore, the representation of the transcriptional regulation of these genes is almost certainly a gross over simplification, as there are likely to be other transcription factors acting in concert with NF-κB to modulate gene expression.

The pathway diagram presented here also includes preliminary views of some of the kinase signalling pathways known to be associated with macrophage activation. These include the MAP2K1/2 (MEK1/2)-MAPK3/1 (ERK1/2) cascade known to be activated by TLR signalling and the MAPK8 (JNK)-JUN (AP1) and MAPK14 (p38) (Figure 2.12e) cascades activated by NOD1/2 and TLR signalling. These pathways are clearly important to macrophage activation and are known to influence a range of different processes such as cell differentiation, cell cycle, phagocytosis and apoptosis. Future efforts will need to integrate these pathways into the overall signalling network to improve the predictive value of the pathway diagrams.

Antigen Presentation and Related Pathways

Antigen presentation is not exclusive to macrophage biology but equally central to it. We have attempted to depict the MHC class I and II pathways from either the degradation of cellular proteins or the phagocytosis of pathogens, to the presentation of cellular/pathogen peptide antigens to CD8 cytotoxic T-cells or CD4 T-helper cells, respectively. In order to achieve this however, it was found necessary to construct

diagrams of the ubiquitination pathway (due to it role in tagging cellular and pathogen proteins with ubiquitin) and proteasome formation (due to the role of the proteasome in the digestion and/or processing of ubiquitinated proteins). These latter two systems are also crucial to many aspects of cell biology, being responsible for the activation and/or degradation of many cellular and pathogen proteins alike. Depiction of these systems proved to be challenging. In the first instance, whilst our view of phagocytosis is greatly simplified, we had to find ways to show the passage of key molecules in the antigen presentation pathways from their assembly in the endoplasmic reticulum (Figure 2.12f) to their transport to the phagosome via the golgi and intermediate endosomal compartments. This required us to show events both at the molecular level, as well as the transition of compartments in which they reside from one state to another. In this case nodes representing pathway modules have been used to link one compartment to another indicating a series of vesicular transitions and fusions; complex processes in their own right. Another challenge was the depiction of the ubiquitination pathway. In short, proteins are tagged for degradation or cleavage through their binding to E3 ligases. Each E3 ligase or E3 ligase complex binds specific protein targets. We have shown a number of classes of these molecules i.e. the HERC/HECT, ring finger, U-box and SCF E3 ligases, encompassing over 80 proteins in total (Figure 2.11c). However many more proteins are thought to be associated with this role, perhaps as many as 500 [199-200]. To add to this complexity there are 36 known E2 ligases and 6 E1 ligases which add further specificity to this system. Clearly it would be impossible to show each individual protein and their associated E3 ligase passing through the ubiquitin pathway (even if the details were known) and therefore it has been depicted as a generic process; proteins bind E3 ligases, which provide a scaffold for ubiquitin transfer from E2 ligases, resulting in the ubiquitinated-protein being presented subsequently to the proteasome for processing. When a protein in the pathway is ubiquitinated it is shown using a process node with the symbol Ub, which is essentially a short cut to showing this process. In the original framework pathway [102] the proteasome was depicted as a single node responsible for the cleavage and activation of ubiquitinated NF-κB complexes. However looking further into the nature of the proteasome, it was then appreciated that there is not just a

single proteasome but at least 5 specific proteasome complexes; the 26S, 11S capped and 19S+11S capped hybrid constitutive proteasomes, and 11S and 19S+11S capped hybrid immunoproteasomes [201-202]. In an effort to show something of the structure of these large barrel-like complexes, we chose to depict the layered arrangement of core subunits into four stacked rings (7 subunits per ring/layer), where appropriate capped with other subunits that form the regulatory particles (Figure 2.11d). By linking the proteasomes with the generic model of the ubiquitination pathway we have also attempted to show each proteasome's preference for the cleavage of a specific class of peptides. However in doing this, it is not possible to show whether the given output of a proteins ubiquitination and cleavage is an activated peptide, a peptide for antigen presentation or the complete destruction of the protein. Where a protein in a pathway is degraded the sink glyph (Ø) has been used as a short cut to indicate that the protein has been removed from the system by proteasomal degradation.
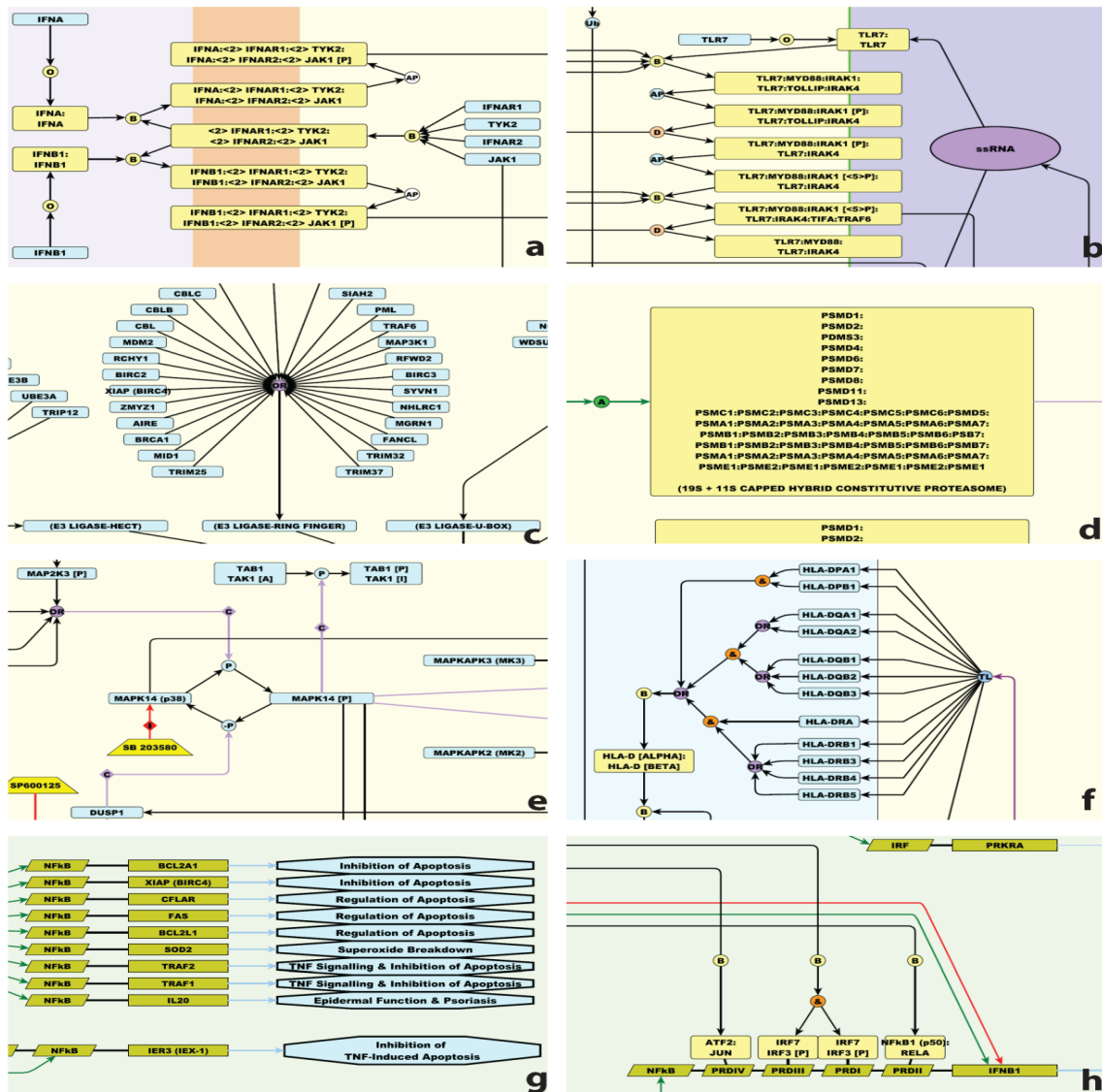
Figure 2.11 Snapshots from the integrated macrophage pathway diagram. (a) Activation of the interferon type-1 receptor through its interaction with interferon-α (IFNA) or interferon-β (IFNB1).   In each case binding of the ligand causes autophosphorylation of JAK1 which eventually leads to the type1-interferon response (not shown). (b) Activation of TLR7 by single stranded RNA in the endosome. This sequential multistep process involves binding events, autophosphorylations and dissociations steps. (c) E3 ligase system.   Up to 500 proteins may potentially function as E3 ligases and here the well documented members are shown. (d) Depiction of the proteasome. In some cases it is useful to lay out the subunits of a complex to reflect the complexes known structure. Represented here are the layers of the proteasome's barrel structure and cap. (e) Activation of MAPK14 (p38).   Phosphorylation of p38 is reversible; numerous kinases will phosphorylate p38. p38 is dephosphorylated by DUSP1 and inhibited by the specific inhibitor SB203580. (f) Combinatorial assembly of the MHC class 2 HLA-D (alpha/beta) complexes. The & and OR Boolean operators indicate the combinatorial assembly of HLA-D (alpha/beta) complex from different classes of MHC class 2 proteins. (g) Genes activated by NFKB1 (p50):RELA (p65) complex. A number of genes activated by the binding of the p50:p65 complex to known NFKB elements in their promoter. In each case the likely functional consequence of this activation is shown as a pathway output. (h) Regulation of IFNB1 expression. Shown are the known promoter elements and factors that bind to them leading to IFNB1 expression.

Compatibility of mEPN Pathways with Other Pathway Analysis Tools

The pathway model presented here is primarily designed to function as a computation resource. Its size as well as the fact that additional information is available through mouse-over or hyper-linked from it, means that it is best viewed on a computer. The diagram has been constructed using the freely available program yEd graph editor (yFiles software, Tubingen, Germany), a general purpose tool designed for the depiction of network-based diagrams. The standard file format used with this program is .graphml which is also supported by other network/pathway editing tools [203-206]. In this format pathways are available for editing and expansion or alternatively using the yEd editor, can be exported from the program in a number of image (.jpeg, .png, .pdf) and exchange formats (.tgf, .gml, .ygf, .xml, .html). The pathway is also available as .pdf and .html formats (see http://www.macrophages.com/macrophage-pathway-resources or http://www.biomedcentral.com/1752-0509/4/63). In order to enhance options for the display, analysis and integration of these pathways with other data types the group has recently implemented the import of .graphml files into Biolayout Express[3D], a network analysis tool [206]. This program supports a range of other network analysis features and is suited for working with small or large networks derived from other data sources. A 'layout' file can be generated such that the diagram can be viewed in this tool as either a conventional 2D or 3D network diagram, in both cases using the node co-ordinates from the .graphml file (although polylines are not supported). Alternatively the diagram can be viewed in 3D using a modified Fructerman-Rheingold organic layout algorithm [207]. A notation system consisting of 3-D shapes is applied in the 3-D view of the pathway [116] (also see Chapter-3 and www.mepn-pathway.org). With the ever increasing amounts of interaction data it becomes more evident that an extra dimension will be valuable for the visualisation of large pathways and eventually an *in silico* cell. The BioLayout *Express*[3D] interface also supports the follow through of connectivity in pathways such that the parent or children nodes of a given selection can be highlighted and selected nodes hidden or isolated.

Pathway diagrams are frequently used as an aid to the interpretation of experimental data e.g. gene expression analyses, proteomics screens whereby the results of these studies are be overlaid on top of pathways to provide context to the findings. To facilitate these analyses the group has recently implemented an "import class-sets" functionality into Biolayout *Express*[3D], allowing lists of genes of interest and/or annotations to be easily exported directly from the tool and identified on the pathway (demonstrated in Raza *et al.*, 2010 [159]).

In the ways described above, the diagram presented here represents a detailed consensus view of a range of pathway systems that are of interest to and the subject of ongoing research into macrophage biology. It has been designed to be easily accessible, distributable and can be modified by end users to suit their interests or knowledge-base. Finally software has been developed to facilitate the use of the pathway as a resource for pathway modelling and interpretation of genomics data.

Discussion

The studies of this Chapter set out to create a pathway resource describing the signalling events active in the macrophage. The pathways described here are of central importance to understanding macrophage biology and therefore innate immunity, and the diagrams provide a consensus view of these systems. This is not to say that the model is either viewed as complete or necessarily even correct, but only as a working model. Amongst the key signalling pathways yet to be incorporated into the macrophage model are the classical stress response, the hypoxia response, a number of PRRs, a range of cytokine and chemokine receptor activated pathways, and greater representation of transcriptional events and regulation. The pathway has been designed with the idea that it will need to be modified and expanded based on new publications, experimental observations or deeper insight into specific systems. All of the pathways depicted are reasonably well characterized and as such there is a relative abundance of information on them from a wide variety of sources. What was lacking prior to this work was a way of collating our understanding of these pathways and integrating this view with the abundance of data generated on these cells by ourselves and others. A key objective has therefore been the creation of a pathway diagram that graphically reflects the current view of a pathway system in a visibly intuitive manner. In so doing the hope was to create a resource for data integration, pathway modelling and hypothesis generation. In order to achieve these objectives it was found necessary to modify both the PDN [103] and EPN schemes [105] for pathway depiction (see Chapter-3). The original diagram [102] (Figure 2.5) acted as a framework for the current effort helping to highlight the many gaps in our understanding and together with developing interests in macrophage biology, helped to prioritize areas for future modelling. Modelling of the pathways continued to be based on labour-intensive curation of the literature. Post-graduate students were given an area of biology to examine, and all the resources for researching the literature and depicting their chosen pathway module. Regular debates on the progress of the pathway models were held, and through this process deficiencies were plugged in the graphical depiction of events, pathway content, notation, component labelling and the recording of the supporting information; a process which in itself was highly informative. An important

point is that the diagrams can be shared and understood by all those familiar with notation, and as a result all the work presented here has been subjected to form of internal critiquing. However, each new area of biology included in the current diagram has presented its own problems in layout and concept representation. As a result there has been subtle but almost constant re-evaluation of various aspects of the notation scheme and as it has been necessary to deal with new issues in the depiction of different systems. The mEPN scheme has matured to the point where little need to change the majority of the notation scheme presented is foreseen (see Chapter-2 and Freeman *et al.,* 2010 [116]), although clearly the modelling of other systems and ideas from others may present a case for further modifications.

Pathway diagrams are an established tool in our effort to interpret and explain results from functional genomics investigations. Overlay of results, usually from studies of the difference between one biological state and another, on top of pathway diagrams allows the investigator to visualize and link observations to defined pathways. BioLayout *Express*[3D], a network analysis tool developed within the group [138, 206], provides a powerful approach to visualize and analyze 'omics data from a variety of sources [138]. Recently implemented is the import of .graphml files into BioLayout *Express*[3D] and the tool now supports the visualization of pathway diagrams as 3D or 2D networks [206]. A parser automatically converts the mEPN notation into the equivalent 3D notation scheme and can use the diagrams original node co-ordinates to layout the pathway. Also implemented is the ability to export analyses from one dataset e.g. clustering of microarray gene expression data and import and overlay these analyses on to another network. As standard gene nomenclature is used in the assembly of this pathway it is possible to map directly between gene identifiers from data to genes/proteins in the pathway. In practice any number of lists with annotations can be imported as class-sets onto the pathway and one can envisage how this would facilitate the comparison of numerous data sets in the context of the macrophage pathway. Although the concept of data mapping onto pathways is not new and is supported by other pathway resources [122, 151, 208] these pathways suffer from a number of issues pertaining to the lack of standard graphical notation

used to depict them. Furthermore the nature of the pathway presented here (in terms of scale, detail, formalised notation, range of pathways covered and integrated nature of their presentation) presents a valuable additional resource for those interested in macrophage biology or any of the more generic pathways included. Clearly the better and more extensive the pathway diagrams are the easier it will be to provide a working hypothesis on the interpretation of data. Increasingly, it is now experimental data that is helping to refine existing pathway models and observations that are yet to be fully understood that are now driving the groups current modelling efforts.

The task of assembling this diagram has been time consuming and laborious involving 1,000's of hours of work. On the other hand, it summarizes the results of investigations that have taken many times that amount of time to perform and it is difficult to envisage how one could précis this body of work in any meaningful way other than as a diagram. To gain a systems level view of these pathways is to gain an insight into the molecular networks that regulate normal immune function and whose malfunction or manipulation underpins inflammatory and infectious disease. Greater understanding of the overall architecture of the immune system and its susceptibility to deregulation by pathogens and other disease causing agents, should ultimately lead to new strategies and targets for therapeutic intervention. Apart from summarizing decades of research, pathways depicted with formalized graphical notation schemes should aid the communication and comparison of biological data. During a thorough process of internal critiquing sections of the pathway were presented to others who were familiar with the notation scheme but not involved in constructing the pathway presented to them and asked to interpret the biology shown. This process ensured that the interactions of the pathway were not ambiguous in their depiction.

Another major incentive for generating pathways with standard notations is to permit the conversion of graphical models into computationally tractable ones, suitable for simulation analyses. For this purpose members of the group have been exploring the use of signalling Petri nets (SPN) [114] for modelling "flow" in the integrated pathway diagram. The approach is suited to large scale models and pathways drawn using the

mEPN system can easily be converted into a bipartite graph of places (nodes) and transitions connected by arcs (edges) that are required to support this approach.. The SPN algorithm uses stochastic flow simulations to distribute 'tokens' representing quantitative estimates of activity through a network graph over time using only the network structure to determine outcomes. The technique has the advantage of offering fast computational simulations on large networks (< 1 sec for ~100 node networks), can support concepts of co-dependency between components and requires no kinetic details for interactions. In this way it should be possible to estimate the dynamics of information flow through a network and the effects of perturbations on that flow. Having developed the comprehensive pathway resource the hope was it would serve as a dynamic tool to aid the groups' research in addition to being a useful point of reference for macrophage pathway knowledge. This is now being achieved with the application of the pathway for genomics data analysis, aid in informing wet-lab investigations and now with the potential for computational modelling.

The exercise of pathway construction has provided a resource for training, pathway modelling, literature/data interpretation, hypothesis generation and as such is now central to ongoing investigations of macrophage biology. Importantly however, the pathway model presented here also serves as a worked example of how pathways might be represented in a logical, unambiguous and biologist-friendly fashion, whatever the system of interest. What is arguably essential for the development of this resource is the support of the wider community in assembling and editing such diagrams. Such efforts are already underway [154-156] and are already providing a vital forum for debate on the known details of pathways in different cell systems. Ideally these efforts will result in detailed models of biological systems that can be shared and assimilated. However, in order to achieve this end pathway models clearly need to be assembled using standard rules and graphical languages. The hope is that this work will contribute to the ongoing community effort to develop such standards [158].

Conclusions

The formalised depiction of biological pathways is increasingly recognised as a crucial requirement for the exchange of pathway data, modelling of their activity and systems level interpretation of biological data. However, there are just a handful of worked examples of large pathway diagrams constructed using a formalised graphical modelling language. The model of macrophage signalling and effector pathways presented here is to our knowledge the most comprehensive pathway of its kind published to date. As such it offers a worked example of how large pathway and has also proved to be a testing ground for the mEPN system [116]. When presented in this manner the network reflects the extensive cross-talk between pathway modules and transcriptional networks and high degree of feedback and feed-forward control taking place.

Although a time consuming and laborious exercise, the act of converting literature derived knowledge into a formalised computational models is essential if we wish to truly gain a systems level understanding of any cellular system. The macrophage model presented here summarizes the results of years of investigations and has allowed the thorough testing of the notation system used to depict it. The hope is that this work will provide a useful resource for others interested in the macrophage and the pathways depicted, and will help contribute to the development of standard graphical depiction in biology.

Methods and Materials

Data mining, curation and organization

Ongoing analysis of macrophage-related datasets and an interest in consolidating knowledge of a number of signalling pathways directed the choice of pathways to be mapped. Public and propriety databases were initially used as resources for data mining, but ultimately all molecular interaction data was sourced from published literature. Manual curation of the literature was performed to firstly evaluate the quality of the evidence supporting an interaction and secondly, to extract the necessary and additional pieces of information required to 'understand' the pathway and construct an interaction diagram. Pathways have been drawn based on the groups long term desire to model pathways active in a human macrophage and therefore all components have been depicted using standard human gene nomenclature (HGNC). However, the understanding of the pathway components and the interactions between them, have been drawn largely from a consensus view of literature knowledge. As such the pathways presented here are based on data derived from a range of different cellular systems and mammalian species (human and mouse). The following details were captured in an interaction list spreadsheet: PubMed ID (of the paper citing the interaction); the names and official HUGO and Entrez IDs of the interacting components; the nature of the interaction, an extract from the interaction table can be found in Figure 2.12 (and the full list is available at http://www.macrophages.com/macrophage-pathway-resources).

Pathway construction

Phase 1 and phase 2 of this work was carried out largely by myself. In phase 3 of the work individual pathway diagrams focused on a specific area of biology were constructed by myself or by teams of 1 or 2 curators who were given a remit to describe a given pathway system using the mEPN scheme [116], over a 6 month period. Primary curators were junior biologists (MSc students) who were encouraged to use all information resources available to first build up an overall picture of these pathways prior to more detailed analyses and literature-based verification of

interactions. Great emphasis was also placed on the need to discuss and justify the information they were attempting to represent to others. Layout was assessed by several curators, as was pathway content and notation usage. Essentially, it was attempted to ensure that the graphical depiction of pathway/interactions was intelligible and unambiguous to another individual familiar with the notation scheme. Teams of curators were therefore encouraged to show and discuss their progress with other members of the group on a regular basis.

All pathways have been constructed as directional networks. Interactions between pathway components are drawn using the principles laid down by the mEPN scheme [116] and diagrams assembled according to the workflow described in Figure 2.9. The current mEPN scheme and a detailed description of the notation scheme and rules for its use are provided in Chapter-3 and Freeman *et al.*, 2010 [116] and www.mepn-pathway.org. Individual pathway maps were drawn using the freely available program yEd graph editor (yFiles software, Tubingen, Germany) and later the pathways were integrated using the same software. In order to make these diagrams an information-rich vehicle for conveying details about pathway components and the reactions between them, PubMed IDs supporting the interactions are stored on appropriate edges or nodes within the .graphml version of the diagram, as are URL-links to Entrez gene for each protein or gene component in the pathway and notes from the curators. Due to the nature of pathway construction edges are often moved, deleted then redrawn to optimise the pathway layout and for this reason annotation may also be linked to an appropriate process and/or edge annotation node.

Pathway optimization and integration

Following an initial development period, the focused diagrams went through extensive editing in attempt to unify their notation usage, stylistic qualities and overall appearance. All aesthetics of the pathways (component colours, text font, text size, edge thickness etc.) were standardized between the diagrams. The original pathway diagram [102] was then used as a framework on to which new pathways were joined. A central rule of the mEPN is that a particular component in a given state may only be

represented once in any sub-cellular compartment [102, 105, 116]. Thus when integrating the diagrams a crucial step was to identify, using the interaction and component lists, overlap between pathway members in the individual diagrams. Connections could then be built between the individual pathways based on shared pathway members and common interactions. For example, a number of the systems of interest feed into the NF-κB system including the TLR and non-TLR pathogen detection receptor signalling, TNF-receptor activation, apoptosis and MAPKinase signalling. If the representation of interactions differed between individual diagrams they were re-examined by going back to the literature. Furthermore, annotations and curators notes were moved and preliminary layouts optimized. Depicting this interconnectivity ultimately leads to numerous challenges in arranging the layout of the diagram. This was particularly acute when laying out the integrated diagram. A significant leap forward was made with the realization that however 'optimized' the layout of the diagram it was too large to be displayed in a readable format on a single page (as had always been the aim when working on a smaller scale when trying to produce a 'publication ready' layout). With this in mind it was possible to be more free with use of space and in the final layout, pathway 'modules' consisting of numbers of connected nodes involved in a similar system are separated out. This has the effect that more space is available to run tracks of parallel edges between modules and subsequent additions or editing are easier to perform. Following the integration of TLR system with the original diagram, the NF-κB, non-TLR and proteosome maps were added sequentially according to the same principles. The fully integrated map then underwent an extensive series of layout optimizations in order to bring visual clarity to final product.

Interacting Partner 1				Interacting Partner 2				Interaction Type	Interaction Location	NCBI-PubMed ID	S
Official Gene Symbol	Gene ID	Interactant Type	Interactant as on map	Official Gene Symbol	Gene ID	Interactant Type	Interactant as on map				
ATM	472	Protein	ATM[P]	IKBKG	4214	Protein	IKBKG[SU]	Binding	nucleus	16497931	
ATM	472	Protein	ATM[P]	IKBKG	4214	Protein	IKBKG[SU]	Binding	nucleus	16965765	
ATM	472	Protein	ATM[P]	MAP3K?		Protein	MAP3K?	Phosphorylation	nucleus	16410802	
ATM	472	Protein	ATM[P]	TP53	7157	Protein	TP53	Phosphorylation	nucleus	9407038	
ATM:ATM	472:472	Complex	ATM[P]:ATM[P]					Dissociation	nucleus	16622404	
ATM:CHUK:IKBKB:IKBKG	472:1147:3551:4214:23085	Complex	ATM[P]:CHUK:IKBKB:IKBKG[P][Ub]:ERC1(ELKS)	NFKBIA	4792	Complex	NFKB1(p50):RELA:NFKBIA	Phosphorylation	cytoplasm	16965765	
ATM:IKBKG	472:4214	Complex	ATM[P]:IKBKG[P][Ub]	CHUK	1147	Protein	CHUK	Binding	cytoplasm	16497931	
ATM:IKBKG	472:4214	Complex	ATM[P]:IKBKG[P][Ub]	ERC1	23085	Protein	ERC1(ELKS)	Binding	cytoplasm	15218148	
ATM:IKBKG	472:4214	Complex	ATM[P]:IKBKG[P][Ub]	ERC1	23085	Protein	ERC1(ELKS)	Binding	cytoplasm	16497931	
ATM:IKBKG	472:4214	Complex	ATM[P]:IKBKG[P][Ub]	IKBKB	3551	Protein	IKBKB	Binding	cytoplasm	16497931	
ATM:IKBKG	472:4214	Complex	ATM[P]:IKBKG[P][Ub]					Translocation	nucleustocytoplasm	16965765	
BCL2	596	Gene	BCL2	NFKB1(p50):NFKB1(p50)	N/A	Complex	NFKB1(p50):NFKB1(p50)	Activation	nucleus	14668329	
BCL2	596	Gene	BCL2	NFKB1(p50):NFKB1(p50)	N/A	Complex	NFKB1(p50):NFKB1(p50)	Activation	nucleus	11567031	
BCL2L1	598	Gene	BCL2L1	NFKB1(p50):RELB	N/A	Complex	NFKB1(p50):RELB	Activation	nucleus	17561400	
BCL3	602	Protein	BCL3	NFKB1(p50):NFKB1(p50)	N/A	Complex	NFKB1(p50):NFKB1(p50)	Binding	nucleus	17561400	
BCL3	602	Protein	BCL3	NFKB1(p50):NFKB1(p50)	N/A	Complex	NFKB1(p50):NFKB1(p50)	Binding	nucleus	16513645	
BCL3	602	Protein	BCL3	NFKB1(p50):NFKB1(p50)	N/A	Complex	NFKB1(p50):NFKB1(p50)	Binding	nucleus	11931770	
BCL3	602	Protein	BCL3	NFKB2(p52):NFKB2(p52)	N/A	Complex	NFKB2(p52):NFKB2(p52)	Binding	nucleus	15371334	
BCL3	602	Protein	BCL3	NFKB2(p52):NFKB2(p52)	N/A	Complex	NFKB2(p52):NFKB2(p52)	Binding	nucleus	17072322	
BCL3	602	Protein	BCL3	NFKB2(p52):NFKB2(p52)	N/A	Complex	NFKB2(p52):NFKB2(p52)	Binding	nucleus	15465827	
BCL3	602	Protein	BCL3	NFKB2(p52):NFKB2(p52)	N/A	Complex	NFKB2(p52):NFKB2(p52)	Binding	nucleus	14668329	
BCL3	602	Protein	BCL3	NFKB2(p52):NFKB2(p52)	N/A	Complex	NFKB2(p52):NFKB2(p52)	Binding	nucleus	16513645	
BCL3	602	Protein	BCL3	NFKB2(p52):NFKB2(p52)	N/A	Complex	NFKB2(p52):NFKB2(p52)	Binding	nucleus	9405619	
BCL3	602	Protein	BCL3	NFKB2(p52):NFKB2(p52)	N/A	Complex	NFKB2(p52):NFKB2(p52)	Binding	nucleus	17349209	
BCL3	602	Protein	BCL3	NFKB2(p52):NFKB2(p52)	N/A	Complex	NFKB2(p52):NFKB2(p52)	Binding	nucleus	10453354	
BIRC3	330	Gene	BIRC3	NCOR2(SMRT):NFKB1(p50):RELA(p65)	N/A	Complex	NCOR2(SMRT):NFKB1(p50):RELA(p65)	Activation	nucleus	16846590	
BIRC3	330	Gene	BIRC3	NCOR2(SMRT):NFKB1(p50):RELA(p65)	N/A	Complex	NCOR2(SMRT):NFKB1(p50):RELA(p65)	Activation	nucleus	16382138	
BIRC3	330	Gene	BIRC3	NFKB1(p50):NFKB1(p50)	N/A	Complex	NFKB1(p50):NFKB1(p50)	Inhibition	nucleus	16846590	
BIRC3	330	Gene	BIRC3	NFKB1(p50):NFKB1(p50)	N/A	Complex	NFKB1(p50):NFKB1(p50)	Inhibition	nucleus	16382138	
CCL19	6363	Gene	CCL19	NFKB2(p52):RELB	N/A	Complex	NFKB2(p52):RELB	Activation	nucleus	17561400	
CCL2	6347	Gene	CCL2(MCP-1)	NFKB1(p50):RELA(p65)	N/A	Complex	NFKB1(p50):RELA(p65)	Activation	nucleus	17561400	
CCL5	6352	Gene	CCL5	NFKB1(p50):RELA(p65)	N/A	Complex	NFKB1(p50):RELA(p65)	Activation	nucleus	15117956	
CDC37	11140	Protein	CDC37	HSP90AA1	3320	Protein	HSP90AA1	Binding	cytoplasm	15371334	
CDC37	11140	Protein	CDC37	HSP90AA1	3320	Protein	HSP90AA1	Binding	cytoplasm	17072322	
CDC37	11140	Protein	CDC37	HSP90AA1	3320	Protein	HSP90AA1	Binding	cytoplasm	17072322	
CDC37	11140	Protein	CDC37	HSP90AA1	3320	Protein	HSP90AA1	Binding	cytoplasm	11864612	
CDC37:HSP90AA1	N/A	Complex	CDC37:HSP90AA1	IKBKB:IKBKG:CHUK:ERC1	N/A	Complex	IKBKB:IKBKG:CHUK:ERC1	Binding		17072322	
CDC37:HSP90AA1	N/A	Complex	CDC37:HSP90AA1	IKBKB:IKBKG:CHUK:ERC1	N/A	Complex	IKBKB:IKBKG:CHUK:ERC1	Binding		16970925	
CHUK	1147	Protein	CHUK	CHUK	1147	Protein	CHUK	Binding	cytoplasm	15145317	
CHUK	1147	Protein	CHUK	IKBKG	8517	Protein	IKBKG	Binding	cytoplasm	11865184	
CHUK	1147	Protein	CHUK	IKBKG	8517	Protein	IKBKG	Binding	cytoplasm	12612076	
CHUK	1147	Protein	CHUK	IKBKG	8517	Protein	IKBKG	Binding	cytoplasm	10837071	
CHUK	1147	Protein	CHUK	NFKB1:NFKB1:HDAC3:NCOR2	4790:4790:8841:9612	Complex	NFKB1(p50):NFKB1(p50):HDAC3:NCOR2(SMRT):[?DNAseq.]	Phosphorylation	nucleus	15494311	
CHUK	1147	Protein	CHUK	NFKB1:NFKB1:HDAC3:NCOR2	4790:4790:8841:9612	Complex	NFKB1(p50):NFKB1(p50):HDAC3:NCOR2(SMRT):[?DNAseq.]	Phosphorylation	nucleus	16382138	
CHUK	1147	Protein	CHUK	NFKB1:RELA:NCOR2	4790:5970:9612	Complex	NFKB1(p50):RELA:NCOR2(SMRT):[?DNAseq.]	Phosphorylation	nucleus	15494311	

Figure 2.12: Extract from table of interaction data. For each interaction depicted on a pathway diagram it is crucial to keep a record of the supporting evidence for that interaction and unambiguous identifications for each interacting components. As the very minimum it is advisable to store the following information; Official Gene Symbol for both interactants, Gene IDs, the type of interaction, the appearance of the interactant as shown on the map (i.e. if a protein is interacting whilst it is in complex with other proteins then it's full complex name/ details are shown), the type of interaction (usually corresponding to the process node involved), the location of the interaction and the PubMed-ID references for each interaction. Some interactions have multiple references sources shown on the line below so new interactions are separated by a yellow line break. Other information, such as the technique used to identify the interaction, the cell type in which the interaction was identified and other supporting information can also be stored.

Pathway overlay of the transcriptional response analysis of mouse bone marrow derived macrophages to interferon-γ treatment

For details of cell culture, treatment, RNA preparation and microarray processing see Chapter-4. Genes that were considered transciptionally activated in response to IFN-γ treatment and included in the set of mapped pathway genes were manually highlighted on the map and the possible downstream consequences (assuming *de novo* protein synthesis and activity following an increase in gene transcription) were followed highlighted.

Chapter Contributions and Acknowledgements

I performed the work described in phase 1 and phase 2 of this Chapter with the guidance of Prof. Tom Freeman. This entailed the development and exploration of how to build pathway diagrams, the evolution of pathway construction process at the DPM (Division of Pathway Medicine), and generation of the first framework diagram of macrophage signalling at the DPM. During Phase 3 I was instrumental in the development of the mEPN scheme, oversaw the assembly of TLR signalling and integrated pathway diagrams, contributed to the development of the notation system, standardisation of pathway data collection. Neil McDerment played a major role in the integration of the final pathway diagram and together with Stephanie Monk, helped redefine the TLR signalling pathway; Paul A Lacaze constructed the initial interferon pathways and oversaw the construction of the non-TLR pathogen detection by George Eleftheriadis and Maire O'Sullivan; Dr. Kevin Robertson and Dr. Steven Watterson provided some useful comments and discussion on the development of the pathway notation and standardization of pathway data collection; Ying Chen and Michael Chisholm assembled the NF-κB pathway; Arran Turnbull assembled the antigen presentation, ubiquitin and proteasome pathways; Athanasios Theocharidis has been developing the program BioLayout *Express*[3D] to enhance its capabilities to support these pathways for visualization and data integration; Dr. Cristelle Roberts has been developing the macrophages.com website and has made the macrophage pathways available on this site; Prof. Peter Ghazal originally conceived the EPN scheme and

supported the current development; finally Prof. Tom Freeman oversaw the pathway construction, orchestrated the development of the mEPN scheme, has directed the development of improved computational resources for these pathways.

This page has been left intentionally blank

Chapter 3. Towards the Standardisation of the Graphical Representation of Biological Pathways: Development of the modified Edinburgh Pathway Notation (mEPN) Scheme
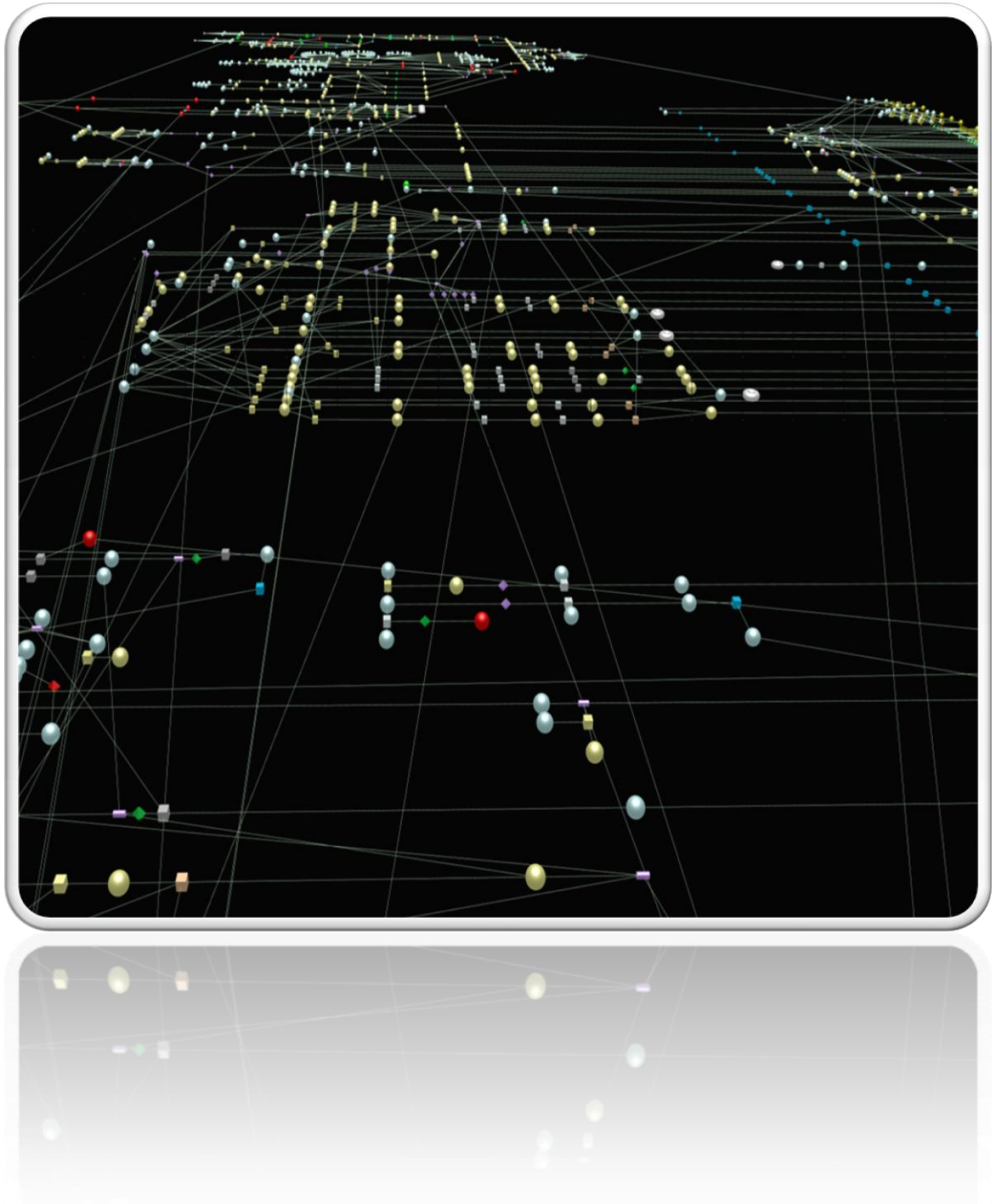
Introduction

Graphical Notation Schemes for Describing Pathways

Despite having a high ratio of graphical to textual information, biology has lacked standard graphical notations for illustrating pathway information. Standard notation for circuit diagrams has proven fundamental to the evolution of the electronics industry (http://www.iec.ch/; http://www.ansi.org) and in the design of computational network systems for example UML diagrams (http://www.uml.org/) diagrams. With the advent of analytical techniques able to perform genome-wide analysis of cell systems and the unprecedented increase in data to interpret, there is a pressing need in biology for formalised methods to depict molecular and cellular systems. Moreover there is also a need for comprehensive pathway models to assist with the interpretation of the vast quantities of medium to high-throughput-data now available.

In order to address these issues the groups of Kohn and Kitano began to devise new approaches to pathway notation using many ideas adopted from the electronics and computer industries [103-104, 209]. In particular the MIM (molecular interaction map) notation [209] a form of entity-relationship representation and the process description notation (PDN) [103], respectively. Examples of pathways that have been published using these notation systems include a molecular interaction map of macrophage signalling [210], Toll-Like-Receptor signalling [100], epidermal growth factor receptor signaling [101] and the RB/E2F pathway [98] which have been depicted using the PDN scheme; cell cycle control and DNA repair has been presented in the MIM notation [104]. However, in the course of our investigations it was found that the diagrams resulting from these elegant and pioneering efforts were not always easy to interpret and the notation system was a challenge to implement (Figure 3.1 provides examples of pathways constructed in the PDN and MIM schemes). Furthermore, it was found that the PDN did not support all of the concepts that are required to reflect the diversity of pathway components and the relationships between them. The original

Edinburgh Pathway Notation (EPN) scheme [105] was designed to allow the logical depiction of signalling pathways and is largely based on the original concepts of the PDN. The notation incorporated many of the ideas of the process PDN scheme but notably introduced the idea of using Boolean logic operators (AND/OR/NOT) nodes to represent co-dependencies between components. The basic objectives of the EPN were to create a notation scheme that was: a) flexible enough to allow the detailed representation of a diverse range of biological entities, interactions and pathway concepts; b) able to represent pathway knowledge in a semantically and visually unambiguous manner; c) able to the construct pathway diagrams that are understandable by a biologist; and d) able to produce diagrams that are sufficiently well defined that software tools can convert graphical models into formal models, suitable for analysis and simulation. Of primary importance to the EPN scheme is the desire to develop pathway maps that are 'readable' by a biologist. Since the pathway maps are primarily produced as a tool for communication it is critical that they are easily understandable and the notation can be applied and read by biologists with minimal training. Other objectives are that the notation should be computable, compact, show sub-cellular localization and be tolerable of incomplete knowledge. Whilst all of these objectives are valid, fulfilling them in practice is far from trivial and at the time of publication on the original EPN scheme there were few worked examples of large pathway diagrams depicted in standard notations, available in the public domain.
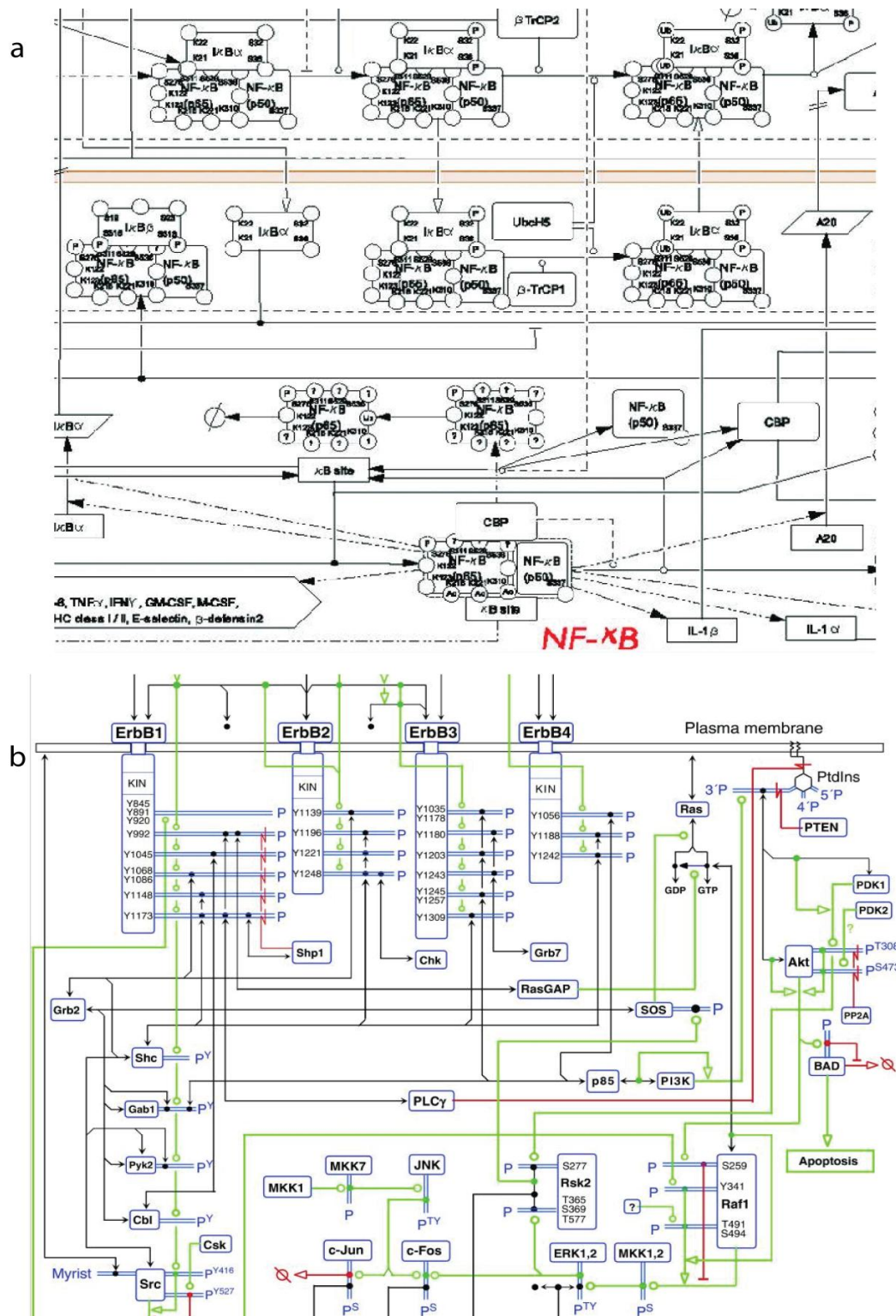
Figure 3.1: Examples of biological pathways depicted with graphical notations. Pathways depicted with existing notation schemes in 2006/2007 were found by our group to be complex to interpret and to difficult to implement at the time; this included their complicated depiction of protein-complexes; multiplicity of arrowheads; struggle of following pathway inputs to their outputs; and limited choice of software which could depict these edges and symbols **(a)** Extract from the Toll-like-receptor signalling pathway by Oda, et al. 2006 [100] depicted with the process description notation (PDN). **(b)** Extract from a map of epidermal growth factor receptor family signalling by Kohn, et al., 2006 [117] depicted using the molecular interaction map (MIM) notation.

Development of the modified Edinburgh Pathway Notation scheme

The original EPN was conceived and published in 2006 [105]. During early stages of "road-testing" of the original EPN it became evident that further refinements were necessary, if the scheme was to be able to fulfil its original objectives. Figure-3.2 shows a sketch created using the original EPN and original pathway editing software, the Edinburgh Pathway Editor (EPE) and Figure 3.3 displays the original EPN scheme as proposed in 2006. It became apparent that changes were needed to the aesthetics of the notation, choice and range of symbols, pathway syntax, as well as choice of supporting software if the notation was to fulfil its original objectives. Chapter 2 (Results → Phase 1) discusses how the appearance and readability of the pathways improved with change of supporting software. The improvement in aesthetics of the pathway was also heavily linked with changes made to the original notation scheme; from increasing size of the original symbols (see Figure 3.3), to amending symbols to appear more intuitive, addition of colour as a visual cue to assist with differentiating between components (for example complexes and proteins) and the eventual removal of numerous different arrow heads to depict interaction since these can easily become difficult to memorise. All in all changes were made where needed to improve the notation scheme to allow it to fulfil one of its fundamental objectives; produce diagrams that are easily understandable by a biologist. Hence in constructing pathway models of the apoptotic, NF-κB, interferon and toll-like receptor pathways [102], I have also been central to the group's effort to refine this notation scheme such that it was fit for purpose. This pathway construction exercise initiated the huge the jump forward from a theoretically useable notation scheme to one that was workable in practice. However as the group's pathway mapping efforts continued to develop it was necessary to further refine the EPN scheme from when deployed to construct a framework map of macrophage signalling. Changes were necessary in order to model a range of different pathways and biological concepts [211]. The objectives of the EPN as originally proposed remain preserved as do many of its original concepts [212]. However substantial modifications have been made to the notation system from the

introduction of new symbols, to changes in the aesthetics of the scheme and pathway syntax.
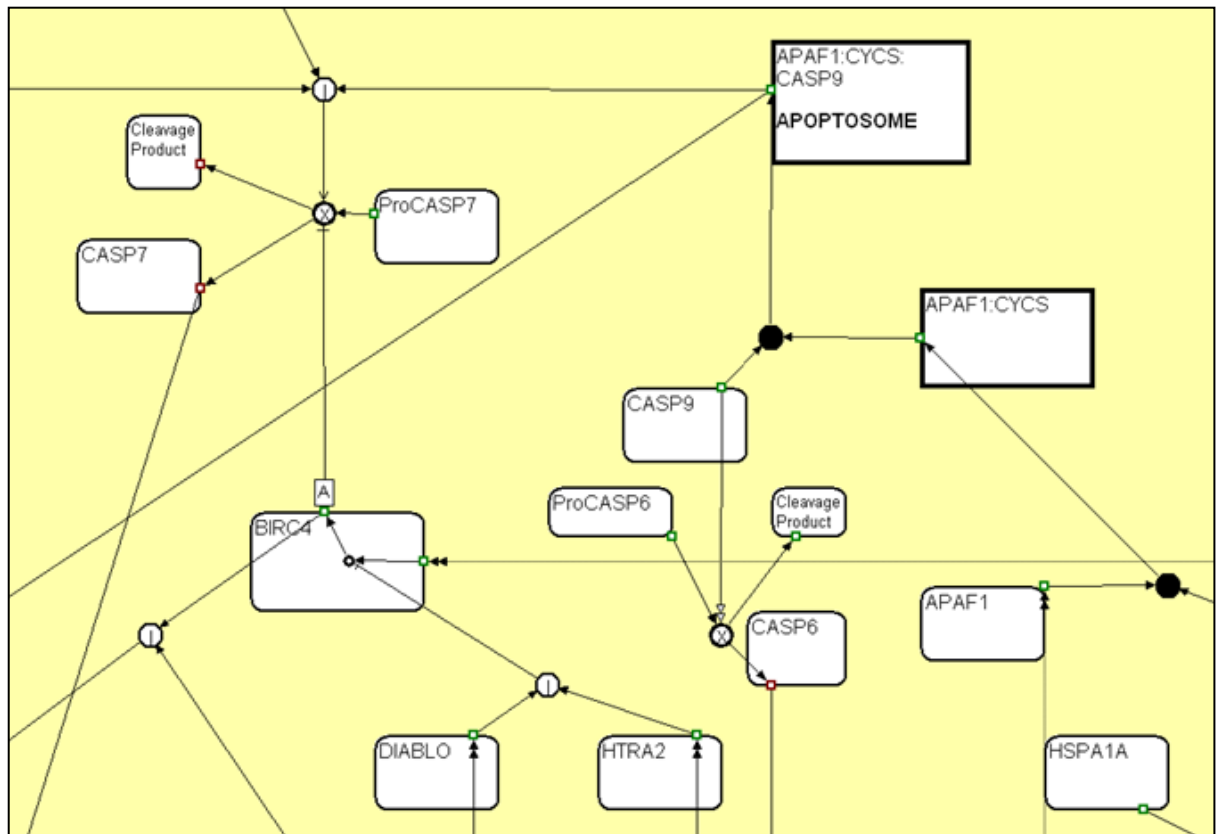
Figure 3.2: Extract from apoptosis pathway drawn using the original Edinburgh Pathway Notation (EPN) scheme. The sketch was drawn using the original EPN palette and supporting software Edinburgh Pathway Editor (EPE). Drawn during the early phases of notation testing the sketch displays some of the deficiencies of the scheme prior to optimization. Some of the original symbols were not intuitive, supporting software did not aid layout and aesthetics, and resulting diagrams were not easy to follow.

Symbol	Name	Description	Symbol	Name	Description
Complex Name	Complex	A protein complex. This can be a homo or hetero - dimer.		Cleavage	Cleavage of a protein into two or more proteins.
Protein Name	Protein	A monomeric protein.	□	State	The state of a complex or protein. If unlabeled the default or baseline state is indicated. The state name should be descriptive and unique to the particular species.
Gene name	Gene	A gene.		Activate	Initiates of a state transition.
─○→	State transition	An action that transforms a species from one state into another. Examples are a chemical modification or conformational change.		Inhibition	Causes a state transition not to occur.
	Complex formation	Two or more proteins (or complexes) aggregating to form a complex.	⊛	AND	Boolean AND gate.
	Dissociation	A complex breaking apart to form two or more proteins or complexes.	⊕	OR	Boolean OR gate.
──▶──	Protein expression	Transcription and translation of a gene to generate a gene product.	⊙	XOR	Boolean XOR gate.
─Ⓣ─▶▶	Translocation	Movement of a protein or complex from one subcellular location to another. This is a type of state transition and can be activated or inhibited in the same way as other state transitions.	⊙	NOT	Boolean NOT gate

Figure 3.3: The original EPN scheme. The original symbols and descriptions of the Ediburgh Pathway Notation as proposed by Moodie *et al.*, 2006.

100

This Chapter will describe the work which followed on from the initial publication of the EPN [105] and define a modified version of the EPN scheme which is aligned with the developing international standard, the Systems Biology Graphical Notation (SBGN), but has a number of important differences with this scheme as currently proposed. The mEPN graphical language has reached a sufficient level of maturity to now be formally described [212]. Arguably the mEPN has some important advantages over other proposed pathway notation schemes and attempts to address some of their shortcomings. As such we believe it is a positive contribution to the debate on standardizing pathway depiction.

Results

Definition of the modified Edinburgh Pathway Notation (mEPN) scheme

A *pathway* may be considered to be a directional network of molecular *interactions* between *components* of a biological system that act together to regulate a cellular event or process. In this context a *component* is any physical entity involved in a pathway e.g. a protein, protein complex, nucleic acid (DNA, RNA), molecule, etc. *Interactions* are generally the relationships between one component and another where one component influences the activity of another e.g. through its binding to, inhibition of, catalytic conversion of, etc. Interactions between cellular components thereby lead to a change in the status of the system. A *pathway notation scheme* is a collection of predefined symbols (shapes, lines, figures) that represent the constituent parts of a graphical system for depicting the components of a biological pathway, the interactions between them and the cellular compartments in which they occur. A scheme should also include rules for the use of these symbols in depicting information. *Glyphs* are stylized graphical symbols that impart information nonverbally and are used to portray different classes of biological entities e.g. protein, gene, pathogen etc. and the nature of the relationships between them. In network terminology all glyphs are nodes (vertices) of a specific type and the connectivity between them is defined by *edges* (lines/arcs). The entire set of glyphs employed in the mEPN scheme are shown in Figure 3.4 and their detailed description and rules for use have been published in a published specification document [116] available at http://www.biomedcentral.com/1752-0509/4/65/additional/ as well as at http://www.mepn-pathway.org/.

.
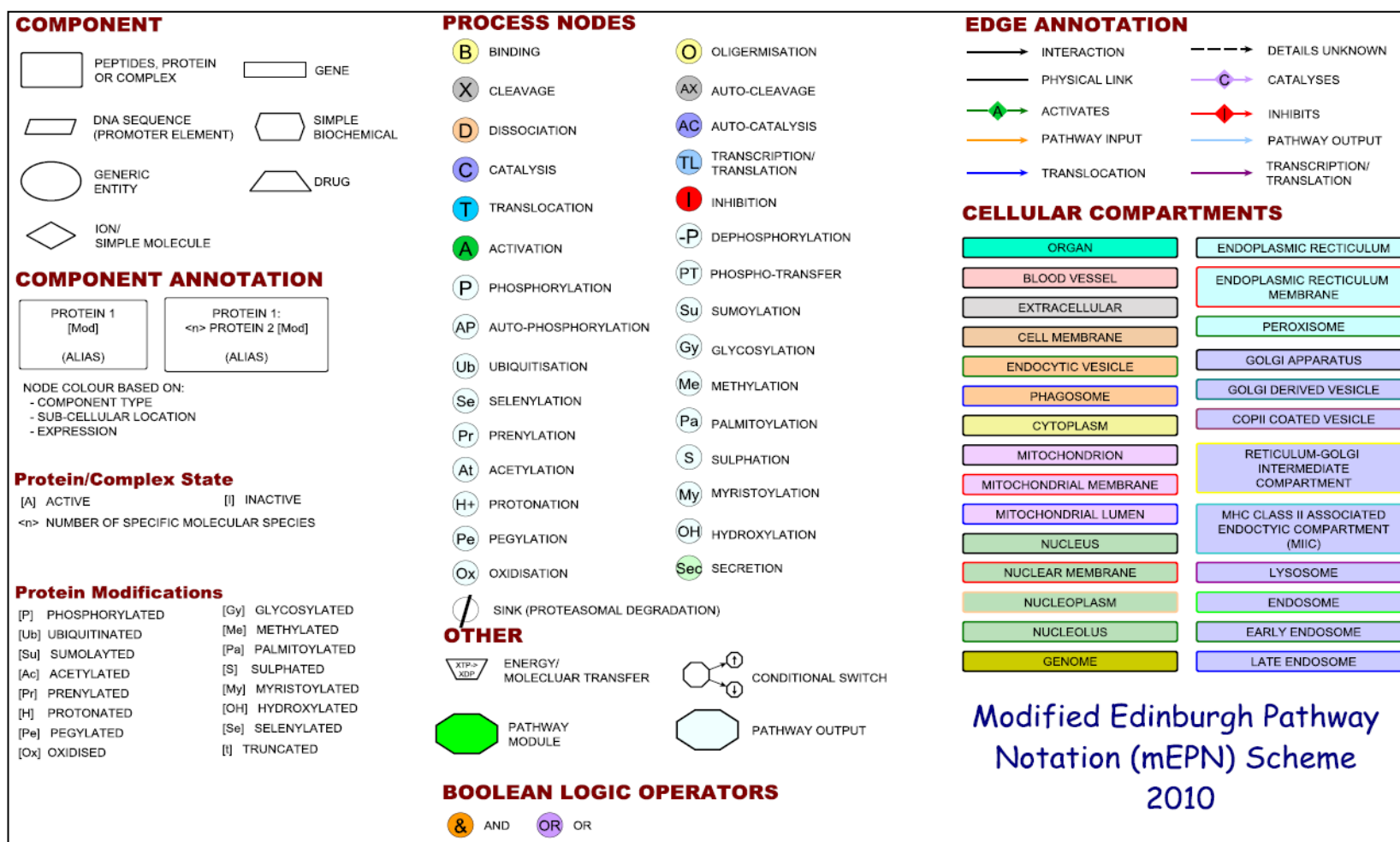
Figure 3.4: List of the glyphs used by the modified Edinburgh Pathway Notation (mEPN) scheme. Unique shapes and identifiers are used to distinguish between each element of the notation scheme. The notation scheme essentially consists of the following categories of nodes representing; cellular components, compartments, Boolean logic, edge annotations, reactions and processes.

Pathway Components

The mEPN uses a set of standard shapes to represent classes of molecules (components) from a rounded rectangle to represent proteins and protein complexes, to a diamond shaped glyph to represent simple ions and molecules e.g. Na^+, K^+, H_2O etc. Components play some role within the pathway and exist in one or a number of locations within a cell. An important rule of the mEPN is that a component may only be represented once in any given cellular compartment. Whilst this rule can potentially lead to a tangle of edges due to certain components possessing numerous connections to other components spread across the pathway, the benefits of the rule outweigh the issues in adhering to it. The number of edges leaving each node gives the reader an exact indication of a component's connections to other components and hence potential activity, without the need for scanning the entire diagram to find other instances where the component is described. A notable exception to this rule is in the depiction of small and ubiquitously present ions and molecules which may be represented numerous times and be involved in numerous processes. A component may however be shown more than once in a given cellular compartment if it changes from one state to another e.g. from an inactive form to an active form, in which case both forms are represented as separate components.

Component Annotation

Multiple names are often available to describe any given component e.g. a number of different names for the same protein may be in use in the literature at any one time. Likewise some common names may be used to describe more than one protein or complex. When non-standard nomenclature is adopted to name pathway components it therefore frequently leads to ambiguity as to the exact identity of what is being depicted. Use of standard nomenclature also assists in the comparison and overlay of experimental data with pathway models. The mEPN recommends the use of standard gene nomenclature systems e.g. HGNC or MGD to name human or mouse genes/proteins, respectively. These nomenclature systems now provide a near complete annotation of all human and mouse genes and their use in the naming of proteins provides a direct visual link between the identity of the gene and the

corresponding protein. Where other names (alias') are in common use these names may be shown as an addition to the label on the glyph representing the protein and are included as part of the node's label after the official gene symbol in rounded ( ) brackets. Protein complexes are named as a concatenation of the proteins belonging to the complex separated by a colon. Again if the complex is commonly referred to by a generic name this may be shown below the constituent parts. There are no strict rules as to the order in which the protein names are shown in the complex and are often shown in the order in which proteins join the complex, in the position they are likely to hold relative to other members of the complex (where known) or position relative to cellular compartments e.g. with receptor proteins in a membrane bound protein complex protruding into the extra-cellular space.

Where a specific protein is present multiple times within a complex, this may be represented by placing the number of times a protein is present within the complex in angle brackets < >. If the number of proteins in the complex is unknown this may be represented by <n>. The particular 'state' of an individual protein or a protein within a complex may be altered as a consequence of a particular process. This change in the component's state is marked using square [ ] brackets following the component's name; each modification being placed in separate brackets. This notation may be used to describe the whole range of protein modifications from phosphorylation [P], truncation [t], ubquitinisation [Ub] etc. An example of an annotated complex is shown in Figure 3.5. Where details of the site of modification are known this may be represented e.g. [P-L232] = phosphorylation at leucine 232. Alternatively the details of a particular modification may be placed as a note on the node visible only during 'mouse-over' or when viewing a node's properties. Where multiple sites are modified this may be shown using multiple brackets, each modification (state) being shown in separate brackets.
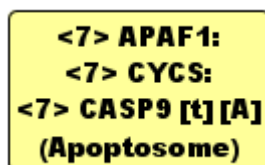
Figure 3.5: Example of a multimeric protein complex, the apoptosome depicted using the mEPN. The active apoptosome consists of 7 APAF1 proteins, 7 CYCS proteins, and 7 truncated CASP9 proteins.

Depiction of Interactions Between Components

Interactions are depicted by edges, sometimes referred to as lines or arcs, and signify the relationships between one component and another. Edges denote that an interaction occurs between components/processes in a pathway and convey the directionality of that interaction. The nature of an interaction is inferred through the use of edge annotation nodes, process nodes, and Boolean logic operators. Interaction edges may be coloured for visual emphasis but as with nodes, the definition of meaning is not reliant on colour. A number of edges contain an in-line annotation node to indicate the 'type' of interaction, as is sometimes depicted by the use of different arrowheads. An edge annotation is generally characterized as having only one input and one output, and functions to describe the type of activity implied by the line e.g. activation, inhibition, catalysis. Figure 3.6 provides an example of an interaction, as depicted using the mEPN.
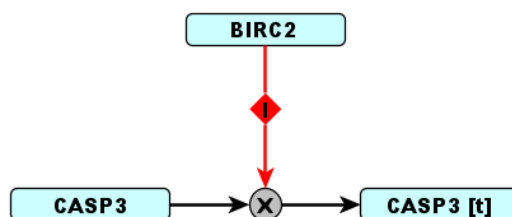
Figure 3.6: Depiction of a component interaction using the mEPN: BIRC2 inhibits the process of CASP3 activation by preventing its cleavage into the truncated form of the protein.

Depiction of Biological Processes

A process is a defined event occurring between components or to a component. A process node in the context of this notation system can be defined as a node that infers an action, transformation, transition or process. They impart information on the

type of process that is associated with transformation of a component from one state to another or movement in cellular location. They also act as junctions between components and as such may have multiple inputs or outputs to components. In the mEPN all process nodes are represented by a small circular glyph and the process they represent is defined by a one-to-three letter code. Colour is used as a visual clue for quick recognition of the nature of the process depicted and group processes into 'type' but again is not necessary for inferring meaning. There are currently 31 process nodes recorded under the mEPN. Different process nodes generally have different connections. For instance a 'binding' node will have multiple inputs and one output, the opposite is true for a dissociation node.

Boolean Logic Operators

Boolean logic operators define the dependencies between components of a system describing the relationship between multiple inputs into a process. An 'AND' operator is used when two or more components are required to bring about a process i.e. an event is dependent on more than one factor being present. In modelling flow through networks these act in a similar manner to 'bind' process nodes i.e. all inputs must be present before a product is formed or reaction proceeds. In contrast an 'OR' operator is used when one component or another may orchestrate the same change in another component. For instance multiple kinases e.g. MAP2K3, MAP2K6, MAP2K7 may catalyze the phosphorylation of p38 (MAPK14) and therefore shown connecting with p38 via an OR operator. OR operators have also occasionally been used to infer that a component(s) has potentially multiple out comes. The Boolean 'NOT' operator has not been included in the mEPN as it would seem self evident that if an event is not depicted it is not occurring or at least there is no recoded evidence that it is.

Other Nodes

There are a number of glyphs that represent concepts that do not sit neatly under the headings of being a component, a process or logic operator. These include:

Energy/molecular transfer nodes are used to represent simple co-reactions associated with or required to drive certain processes (e.g. ATP→ADP, GTP→GDP, NADPH→NADP$^+$). They are linked directly to the node representing the process in which they take part.

Conditional gates are used where there are potentially multiple fates of a component and the output is dependent on other factors such as the components concentration, time or is associated with a cellular state. These have been used to depict events such as the check point controls in cell cycle where the decision to go on to the next phase cell replication is under the control of a number of factors and two or more outcomes are possible. Another example is where cholesterol, depending on its intracellular concentration, may be either exported out of the cell or trigger the cholesterol synthesis pathway.

Pathway modules define complicated processes or events that are not otherwise fully described. Examples include signalling cascades, endocytosis, compartment fusion etc. They are a short-hand way of representing molecular events that are not known, not recorded or not shown.

Pathway outputs detail the cumulative output of series of interactions or function of an individual component at the 'end' of a pathway. Pathway outputs are shown in order to describe the significance of those interactions in the context of a biological process or with respect to the cell. The input lines leading into a pathway output node have been coloured light blue to emphasize the end of the pathway description.

Depiction of Cellular Compartments

A cellular compartment can be a region of the cell, an organelle or cellular structure, dedicated to particular processes and/or hosting certain sub-sets of components e.g. genes are found only in the nuclear compartment. Sub-cellular compartments are defined by a labelled background to the pathway and arranged with spatial reference to cell structure. Compartments are coloured differently for emphasis and to ease awareness the location of components. Similar or related compartments share the

same fill colour but have different coloured perimeters to define internal boundaries within a compartment e.g. membrane vs. lumen or to define the origin of compartments e.g. different classes of vesicles derived from the endoplasmic reticulum or plasma membrane.

mEPN Use of Colour

The mEPN scheme has been designed to function in the absence of colour and no aspect of it is dependent on colour for its full understanding, hence avoiding issues variable colour recognition capabilities between individuals and issues with a poor reproduction of figures. However colour is a powerful visual tool and has been used in the deployment of the mEPN for emphasis. A proposed colour scheme is described below although is open for adaptation to suit the end users needs or aesthetic tastes. Nodes may be coloured to differentiate between different node types e.g. between a protein, complex or gene, to denote their cellular location or expression/activity level.

IFNγ Activation of MHC Class II Gene Expression: A Worked Example of the mEPN in Use

In order to demonstrate the pathway notation system in action on a scale that can be viewed in this format, a small part of the pathway diagram in Chapter-2 has been extracted for discussion. Figure 3.7 depicts the activation of MHC class II genes by interferon-gamma (IFNG) as described in the literature and represented here using the mEPN scheme.
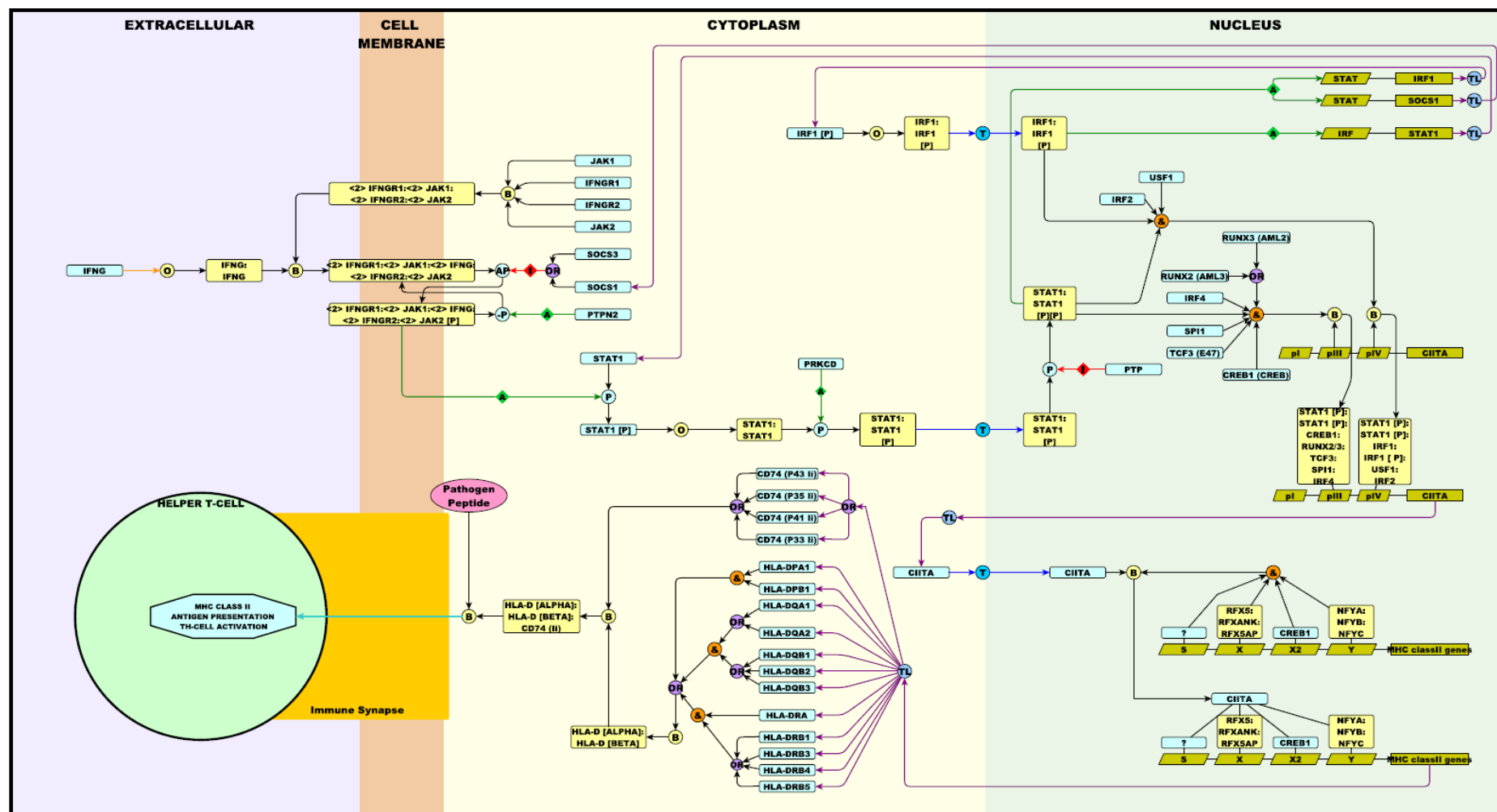
Figure 3.7: Graphical representation of the Interferon-gamma pathway leading to MHC Class II Antigen Presentation. Shown here are the known events between the release of IFNγ and the subsequent up regulation of MHC class 2 antigen presentation by macrophages using the mEPN scheme. See results for a full description of this pathway.

Going through these series of events (shown in Figure 3.7) in detail: IFNγ is secreted by T and NK cells upon activation [213-215] (not shown). It oligomerises to form a homodimer which then binds of to its receptor complex situated in the plasma membrane of macrophages [216]. This complex is formed from IFNGR1, IFNGR2, JAK1 and JAK2 [217-218], two copies of all proteins being present in the receptor complex. Binding of IFNγ causes the autophosphorylation of JAK2 [219] which in turn phosphorylates STAT1 [217]. The autophosphorylation of JAK2 can be inhibited by SOCS1 or SOCS3 [59, 220], and the activated complex dephosphorylated by PTPN2 [221-222]. STAT1 now activated, oligomerises, is further phosphorylated by PRKCD [223] and translocates to the nucleus where it directly activates gene expression by binding to STAT sites present in the promoters of numerous genes. Shown on the diagram are just two of these genes, *SOCS1* and *IRF1* [192, 224]. These form feedback inhibition and feed-forward activation loops, respectively. SOCS1 blocking further signal propagation through the inhibition of the IFNγ receptor complex (reviewed in [225] and IRF1 being necessary for the activation of *STAT1* expression as well as being a necessary component of the *CIITA* transcriptional initiation complex [148]. At least two complexes are reported to be necessary to activate the expression of *CIITA* (reviewed in [226], the first composed of STAT1, IRF1, USF1 and IRF2 which binds to the so called pIV element of the *CIITA*, the second is comprised of STAT1, CREB1, RUNX2/3, TCF3, SPI1 and IRF4 which binds to the pIII element of the gene. CIITA is a co-activator and the key missing element in the transcription of MHC class II genes. Once translated it binds to a preassembled transcription factor complex, including members of the RFX and NFY family of proteins and CREB1, thereby activating the expression of the MHC class II genes [226]. This class of genes includes *CD74, HLA-DPA/B, HLA-DQA/B, HLA-DRA/B* [225, 227] and through combinatorial assembly form a wide variety of complexes denoted here generically as CD74 (Ii):HLA-D (alpha):HLA-D (beta). It is this class of complexes that is shown in the main diagram to go on through a long series of steps to bind peptide antigen derived from the lysosomal degradation of pathogen proteins and present them to T-helper cells. As such this diagram serves as a graphical representation of the known pathway connecting IFNγ secretion to the activation of MHC class II antigen presentation.

Efforts in developing this notation scheme have now reached a point where little need to change the majority of the mEPN scheme as presented here is foreseen. Clearly the modelling of other systems and ideas from others however may in the future present a case for further modifications or refinements.

Visualization of Pathway Information in 3D Environments

The reliance of the mEPN scheme on the principles of network graphs and use of simple node shapes, labels, edges and colour to convey pathway information has presented the opportunity to examine the use of other environments in which to visualize pathways. Layout of pathways in 3D space begins to address the issue of scalability associated with visualizing very large pathway diagrams and offers a little explored environment to visualize and interact with pathway models. Hence devised for the first time is a 3D translation of mEPN scheme (Figure 3.8). The scheme is devised to reflect the colours and where possible glyphs used in the 2D mEPN process diagrams converting the 2D shapes into 3D objects. The proposed notation scheme is currently supported by the network visualization and analysis tool BioLayout *Express*^{3D} (see Figure 3.9) [138, 206]; http://www.biolayout.org/) which also currently supports the direct import of pathways as .graphml files, the main file type used by us to support our pathway modelling efforts. The potential of representing pathways in 3D environments is discussed below.
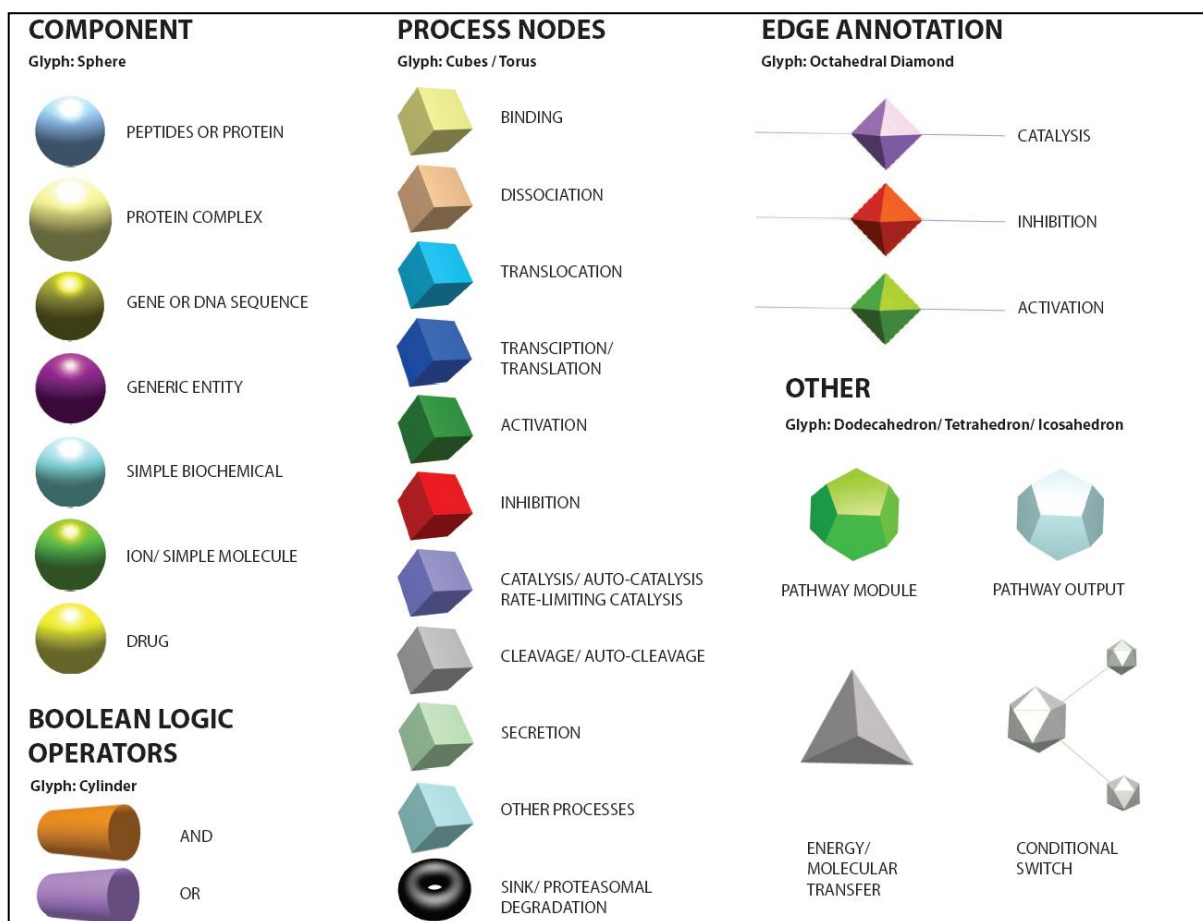
Figure 3.8: The mEPN³ᴰ scheme. Presented here is a conversion of the standard mEPN scheme into a series of shapes that can be used to depict the same pathway concepts in 3D environments.
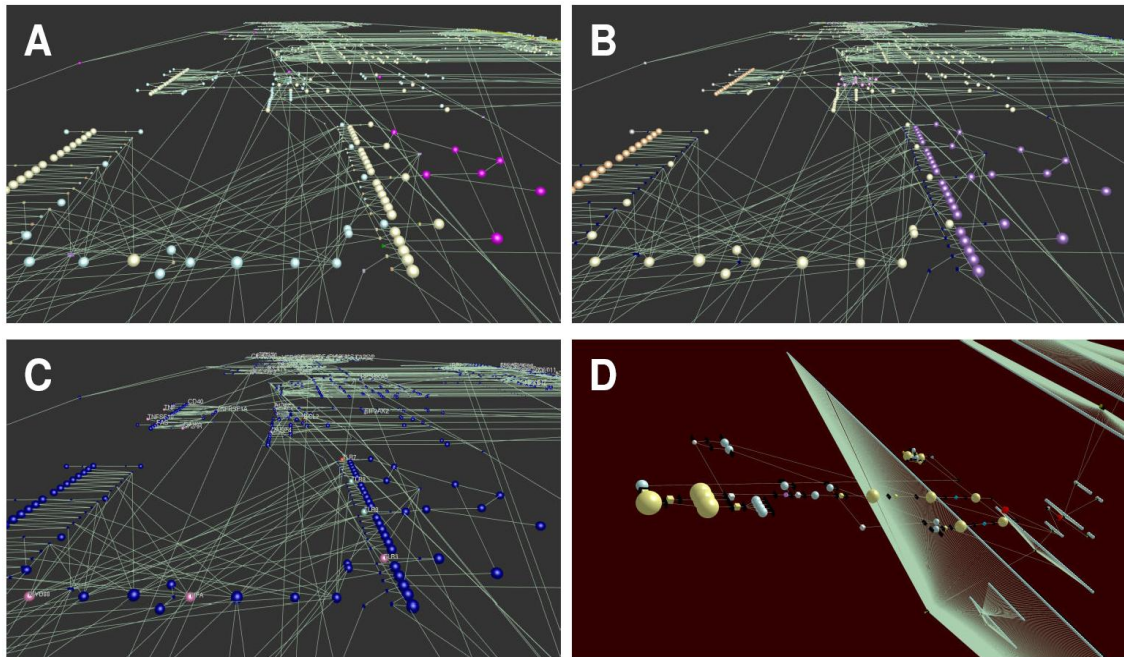
Figure 3.9: Pathway Representation in 3D Environment. Large macrophage activation pathway rendered in 3D environment where node shape, size and colour represents a components identity. **(a)** Nodes coloured according to type e.g. light blue - proteins, yellow - protein complexes, purple – generic molecular species. All process nodes are depicted as small cubes and coloured according to type. **(b)** Nodes coloured according to cellular location e.g. brown – plasma membrane, yellow – cytoplasm, purple – endosome, green nucleus. Process nodes/Boolean logic operators are shown as having no cellular location and are coloured dark blue (no class). **(c)** Nodes coloured according to overlay of data, in this case expression data. Colour of nodes represents co-expression cluster following stimulation of mouse macrophages with IFN-β **(d)** A representation of the interferon-beta signalling pathway and the transcriptional network it controls. The signalling network is represented using the mEPN^{3D} notation with the addition of transition nodes for use in modelling studies. Connected to it are clusters of genes up or down regulated by IFN-β which have been stacked in at different layers depending on the their time course of activation/repression.

Discussion

Models of pathways produced either as a graphical representation of known events or as a resource for mathematical modelling, are fundamental to our understanding the workings of biological systems. However the task of assimilating the large amounts of available data and representing this information in an intuitive manner remains a challenge. Accordingly there has been increasing interest in the biology community to develop approaches for representing biological pathways. The Molecular Interaction Map (MIM) and Process Description Notation (PDN) schemes were proposed by Kurt Kohn [104, 209] and Hiroaki Kitano (Kitano 2005), respectively, and their ideas laid the foundations for much of the work on pathway notation that has followed. The current mEPN scheme is the based on ideas from the PDN and original EPN schemes but importantly the experience of over four years of pathway construction, notation testing and discussions.

The objectives of the EPN as originally proposed remain preserved, as do many of the original concepts of the EPN and PDN schemes [103, 105]. However substantial modifications have been made to the notation system from the introduction of new symbols to changes in the aesthetics of the scheme and pathway syntax in order to achieve the schemes original objectives. Firstly was the desire to develop a notation system that was flexible enough to allow the detailed representation of diverse biological entities, interactions and pathway concepts. In this respect, the mEPN as described here has not only been used in the construction of the large macrophage pathway diagram [211] detailed in Chapter-2, which in its own right covers a diverse range of signalling and effector pathways, but also for the depiction of cholesterol metabolism and the cell cycle by other students (not shown). In all of these endeavours the mEPN scheme has been able to depict the literature-based understanding of these systems and where it was formerly unable to support a concept, it was modified to allow it to do so. Secondly was the need for a system for presenting pathway knowledge in a semantically and visually unambiguous manner. To some degree this is down to actually labelling components in a way that is

unambiguous. The use of standard gene nomenclature to label genes/protein components, together with a formalized system to describe modifications to them, goes some way to achieving this. This has meant in many cases that the literature which describes these systems using numerous different names for the same protein or complex must be de-convoluted. It means however that one component is unlikely ever to be represented more than once but with different names. It also facilitates use of the diagrams in the interpretation of experimentally derived data which is usually annotated using standard gene nomenclature. The third aim, which is related to the second, is that diagrams are as simple as possible to construct and are understandable by a biologist. To help ensure this to be the case all the work in creating the pathway diagrams has been performed by relatively junior biologists (myself as a PhD student and MSc students). Those constructing the pathways were encouraged to discuss their ideas and pathways with each other so as iron out areas where the information was not clearly depicted. For this to happen one must be able to communicate complicated biological concepts using the diagrams. The readability of a diagram is not only dependent on the notation system but also on its layout. Although a variety of automated layout algorithms exist for network graphs they do not perform as well as a human curator with an artistic eye for the task. Pathway layout is relatively trivial for small diagrams, but a long time has had to be spent on optimizing the layout all of our large pathways so that they are easily interpreted. Finally, pathway diagrams are central to efforts to computationally model the observed behaviour of biological systems [119]. The fourth objective has therefore been to develop the mEPN such that the semantics of the resulting network diagrams are sufficiently well defined that software tools can convert graphical models into formal models, suitable for analysis and simulation. Whilst the primary objective behind the efforts has been to create a graphical model of events, the group has been mindful to construct pathway diagrams as directional networks that could in principle support studies on the dynamics of these systems. In examining various approaches to pathway modelling some are clearly not scalable, such as those using ordinary differential equations (ODEs) that require interaction parameters to be known or computed. Other approaches do not

support the modelling of the co-dependencies between components of a pathway or give quantitative outputs (reviewed in [128, 227].

With the increasing interest in pathway science and depiction a community effort to develop a standard notation was formed, known as the Systems Biology Graphical Notation (SBGN)[228], and running concurrently with mEPN work has been this ongoing community effort to establish rules for best practice in pathway depiction. This effort aims to combine the strengths of the various proposed notation schemes and arrive to a consensus approach for representing biological pathways and only recently and a manuscript describing the SBGN Process Diagrams Level 1 specification was published [158, 228]. The mEPN scheme as described here aspires to many of the same goals as the SBGN and where possible we have tried to harmonize the mEPN scheme to the emerging SBGN specification. However, experience in building large-scale pathway models of a variety of biological systems has required the group to depict concepts not currently supported by the SBGN scheme. Furthermore, a lack of available pathway editing tools when this work began, as well as the scale of diagrams produced (see Chapter-2), have both played their part in determining the mEPN approach to pathway depiction. As a result there are a number of important differences that exist between the mEPN scheme described here and the SBGN scheme as currently proposed. Firstly, in common with the proposed SBGN scheme, the mEPN uses simple shapes to define the class of a component but only a labelling system to define the exact identity of components (nodes). The SBGN scheme proposes the use of circles overlaid on nodes to depict protein modifications. This has been found to be a considerable overhead to implement and can interfere the clarity of what is depicted rather than enhancing it. Furthermore the notation scheme is not supported by many of the general purpose network visualization tools e.g. yEd, Cytoscape, BioLayout *Express*[3D] [108, 229-230] in general use, requiring instead the use of dedicated pathway software. Secondly, the mEPN avoids the use of different styles of arrowheads to depict the nature of interactions (edges) which limits the vocabulary of edges and is a system that can be challenging to remember. Instead where appropriate, inline annotation nodes are used to depict the meaning of edges; these

carry a visual clue (a letter symbolizing the meaning of the edge e.g. A for activation, I for inhibition) and can potentially support a wider range of edge meanings. Again the use of a wide variety of arrowheads is not supported by many pathway/network editing software packages. Finally, using the mEPN one can explicitly state the nature of interactions by the use of labelled process nodes. In the proposed SBGN scheme process nodes are used but generally not as a means to convey the nature of interactions except in the case of protein binding (association) and dissociation. When pathways are large and the distance between interacting species may be great, having a visual clue as to the nature of interactions is very important. Whilst on these and other points the mEPN and SBGN schemes may differ, we are fully supportive of the SBGN's efforts to promote a common notation system and hope that current the work presented will contribute to the adoption of common notation schemes for pathway depiction. A full description of the differences between the SBGN level 1 notation and the mEPN as described here follows.

Comparison of the mEPN Scheme with the current SBGN Level 1 specification for the Depiction of Process Diagrams

The mEPN (modified Edinburgh Pathway Notation) and SBGN (Systems Biology Graphical Notation) schemes provide two similar but different ways to depict process diagrams. Each scheme is divided into a set of glyphs to depict different concepts (components, processes, relationships, cellular compartments) for the use of depicting what is known about a biological pathway as a network diagram. Both schemes were developed over roughly the same period of time the SBGN scheme by members of the SBGN community (which includes some of those involved with the mEPN); the mEPN scheme at the Division of Pathway Medicine, the Centre for Systems Biology Edinburgh (CSBE) and Roslin Institute, University of Edinburgh. Both schemes also aspire to fulfil many of the same goals.
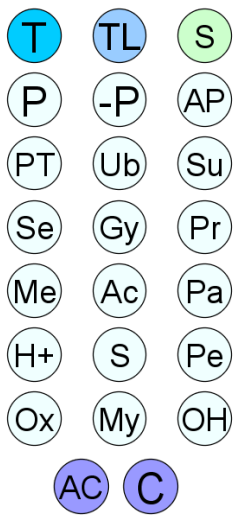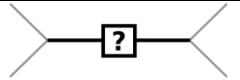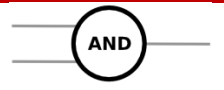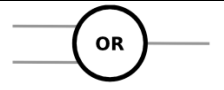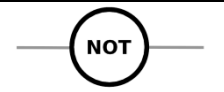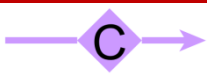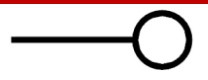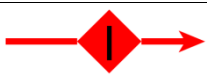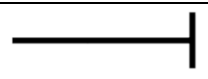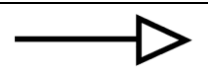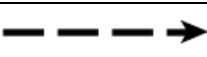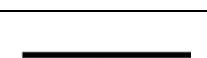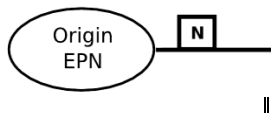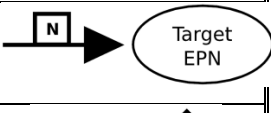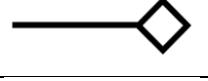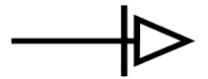
Although many of the concepts are named differently, many glyphs differ between the two notation systems they are similar enough to compare. Below is a comparison of the main features of both notation schemes and the glyphs used within them. Since

different naming conventions are used by the notation systems to describe glyphs, the most similar or conceptually equivalent nodes are compared below. Where a glyph exists in one notation scheme but not in the other a blank space can be found in the adjacent area of the table.

Full details of the SGBN scheme can be found at (Le Novère, N. et al., The Systems Biology Graphical Notation. <u>Nature Biotechnology</u> **27**: 735-741 (2009).

Concept	mEPN Node	mEPN Glyph	SBGN equivalent	SBGN Glyph
Pathway Components[i] (Entity Pool Nodes)	Peptide, protein, protein complex	LABEL	Macromolecule	LABEL
	Gene	LABEL	Nucleic acid feature	LABEL
	DNA sequence (e.g. promoter element)	LABEL		
	Simple biochemical	LABEL	Simple chemical	LABEL
	Ion/ simple molecule	LABEL	Simple chemical	LABEL
	Generic entity	LABEL	Unspecified entity	LABEL
	Drug	LABEL		
	Multimers (also complex)	<5> LABEL	Multimers	N:5 LABEL
Other	Pathway output[ii]		Observable	LABEL
	Pathway module		Submap	EPN or CN — A LABEL — B EPN or CN / C EPN or CN
	Energy/ molecular transfer[iii]	XTP-> XDP ↓ Ⓟ		XTP → XDP □ →
	Conditional switch	↑ / ↓		

			Tag	LABEL
			Perturbation	LABEL
Cellular Compartment [iv]	Cellular Compartment	LABEL COMPARTMENT	Container Node	e:INFO LABEL
Process Nodes (Transitions)	Binding	B	Association	
	Dissociation	D	Dissociation	
	Sink[v]	∅	Source/Sink	∅
	Other processes[vi]	A I	Transition	
		O X AX	Omitted process	

		T TL S P -P AP PT Ub Su Se Gy Pr Me Ac Pa H+ S Pe Ox My OH AC C		
			Uncertain process	⟩—[?]—⟨
Boolean Logic Operators	**AND**	&	**AND**	AND
	OR	OR	**OR**	OR
			NOT[vii]	NOT
Interaction Edges (Connecting Arcs)[viii]	**Catalyses**	—◆C◆→	**Catalysis**	—○
	Inhibits	—◆I◆→	**Inhibition**	—⊣
	Activates	—◆A◆→	**Stimulation**	—▷
	Details unknown	- - - →		
	Non-covalent or covalent bond	——		
			Consumption	Origin EPN —[N]——
			Production	—[N]▶ Target EPN
			Modulation	—◇
			Trigger	—|▷
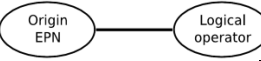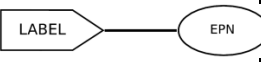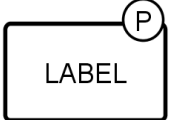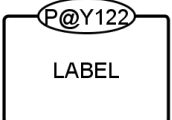
			Logic Arc	
			Equivalence Arc	
Component Annotation[ix] **(Auxiliary Items)**	Component ID	*Recommended standard gene ID*	Component ID	*No standard specified*
	Component Modifications (examples)	LABEL [P]	**State Variables (examples)**	
		LABEL [P@Y122]		
	Component State	LABEL [A]	Unit of information	

Table 3.1: Comparison of the mEPN and SBGN.

Notes

i **Pathway Interactant Depiction**

Nodes which are the biological constituents participating in a particular metabolic or signalling pathway are referred to as **Components** in the mEPN and **Entity Pool Nodes** are the equivalent in the SBGN.

ii **Comparisons of *observable* and *pathway output* glyphs**

Observable and *pathway output* are comparable nodes; *observable* is used to describe a process affected by, or a phenotype generated as a result of pathway signalling. *Pathway Output* is also used for this purpose and always at the end of a series of directional interactions to highlight the consequence of a set of interactions.

iii **Energy/Phospho Transfer Depiction**

SBGN use two nodes (*simple chemical)* each time to show the transfer of x-tri-phosphate → x-di-phosphate (e.g. ATP → ADP, GTP → GDP, whereas in the mEPN we have chosen to depict these reactions in one glyph which points to the process requiring energy/phosphate transfer. The use of a single glyph to depict energy/phospho-transfer was determined to be the most space efficient way to depict these co-reactions due their widespread occurrence in biology.

iv **Compartmentalisation**

In both notation schemes the glyphs used to contain nodes present in a given sub-cellular-compartment can take any geometry. The SBGN *container* nodes are also used for containing a *complex* or a *submap* (a node used to encapsulate processes within one glyph).

v **Sink Nodes**

Sink is an entity pool node in the SBGN and considered a process node in mEPN. Furthermore in mEPN the use of the *sink* node is restricted to defining the removal of a component from a system which in all cases to date has been by proteasomal degradation.

vi **Other Processes**

Other processes in the mEPN shown reading from left to right in the table are Activation (A), Inhibition (I), Oligomerisation (O), Cleavage (X), Auto-cleavage (AX), Catalysis (C), Auto-catalysis (AC), Translocation (T), Transcription/Translation (TL), Secretion (S), Phosphorylation (P), De-phosphorylation (-P), Auto-phosphorylation (AP), Phospho-transfer (PT), Ubiquitination (Ub), sumoylation (Su), selenylation (Se), glycosylation (Gy), prenylation (Pr), methylation (Me), acetylation (Ac), palmitoylation (Pa), protonation (H+), sulphatation (S), pegylation (Pe), oxidation (Ox), myristoylation (My), and hydroxylation (OH). Use of colour is optional. The nature of processes (transitions) are not generally defined under the current SBGN specification.

vii **Boolean Logic Operators**

Both notation systems make use of Boolean logic commands AND / OR, however extensive use of the mEPN has yet to find use of the NOT command for signalling pathways and is therefore currently not included from the mEPN notation. Something NOT doing something would seem to be obvious by its omission.

viii **Edges**

The lines connecting components and process nodes are referred to as **edges** in mEPN or **connecting arcs** in SBGN. The mEPN does not make use of different styles of arrowheads to depict the nature of interactions (edges) instead where appropriate an diamond-shaped inline annotation node carrying a visual clue (a letter symbolising the meaning of the edge e.g. A for activation, I for inhibition) is used to depict the meaning of edges.

ix **Node Annotations**

No naming conventions are currently recommended by SBGN. When non-standard nomenclature is adopted it frequently leads to ambiguity as to the exact identity as to what is being depicted as multiple component names are often available to describe a given component. Under mEPN we recommend the use of standard nomenclature systems for components e.g. HGNC or MGD conventions to name human or mouse genes/proteins, respectively. Use of standard nomenclature also assists in the comparison and overlay of experimental data with pathway models. Additional information about a component (modifications, states, numbers of given components within a complex) are referred to as **annotations** in the mEPN and **auxiliary items** according to the SBGN.

Other Differences

Cloning concept of SBGN

If a component (entity pool node) is duplicated on the map it is indicated by using a 'clone marker' (shading in the bottom third of the node) the purpose of this being to allow the reader a visual indication that the node has been duplicated elsewhere on the map. Whilst on a map of moderate size this maybe practical (allowing the reader to identify how many times the node is duplicated) on a larger scale and where multiple nodes are cloned it may become difficult to trace how many times the node is cloned.

In contrast to the SBGN, mEPN usage rules dictate that a component node (proteins, complexes, etc) representing a given entity may only be represented once in the in a given sub-cellular compartment (this rule does not apply to ubiquitous components or reactions e.g. simple ions, energy transfer reaction nodes). The trade off here is that so called 'hub' nodes (those with many connections) will have many edges emanating from them to other components in various different locations of the map. However, the number of edges leaving each node gives the reader an exact indication of its connections and hence activity in the map without the need for scanning the entire diagram to find cloned nodes. Furthermore, the SBGN has laid out rules as to which glyphs may be cloned and which then require clone markers, adding yet another set of rules that map constructers must learn. Although both notation systems do not provide a perfect solution to dealing with highly connected nodes arguably the mEPN rule (biological component can be shown only once in a given sub-cellular compartment) is a more practical resolution for readers and constructors of the diagram for the reasons discussed above and also since repetition of identical nodes consumes more space on the map.

3D Rendering of the mEPN

One advantage of the simple node and edge based approach to pathway element depiction in the mEPN is that it supports the notations conversion into other software

environments. In particular there is a growing use of technology developed in the gaming and animation industries to support the visualization of data in virtual 3D environments. BioLayout *Express*[3D], a network analysis tool developed at the Roslin Institute and EBI and employing the 3D graphics application programming interface (API) OpenGL, provides a powerful tool with which to visualize and analyze a variety of types of 'omics data as networks [138, 206]. Recently implemented is the import of .graphml files into BioLayout *Express*[3D] such that the pathway diagrams can be viewed in 3D as well as a 2D environment. In this environment node walks can also be performed to identify the parents or children of any given node or set of nodes, thereby allowing the connectivity between components in large pathway systems to be explored. Interestingly and at first a surprise, was the effect of translating a 2D pathway layout as described in the original .graphml node co-ordinates into a 3D environment. In this way diagrams may be rotated, zoomed in on and generally explored in an environment which is quite different to that of a 2D representation. In the 3D environment colour is a powerful device that can be used to further overlay visual information on to nodes (Figures 3.9.a c). Pathways can also be visualized using 3D organic layout algorithms (Figure 3.9. d.). This visualization of the pathways is engaging but is currently of limited utility. However, it is possible to imagine much larger models of pathway systems where the spatial layout of components in 3D space is based on a components cellular location and the visualization more closely approximating an *in silico* cell. With BioLayout *Express*[3D] now capable of supporting networks comprising of up to 45,000 node graphs there is considerable scope for building ever larger pathway models and further exploring the potential of 3D environments for pathway visualization and analysis.

Conclusions

There are significant efforts already underway to garner the support and interest of the wider biological community in assembling resources, information and pathway diagrams covering a broad spectrum of biology. Indeed, the need has never been greater for these resources. However, if they do not record pathways in a standardized way, integration of the results of these efforts will continue to be a considerable issue.

To this end the mEPN is fully supportive of the SGBN's effort to promote the principles of standard notation systems even if as a group we cannot fully support the proposed SBGN specification for process diagrams. The hope is that this work in its published form [212] and the pathway diagrams [102, 211](and Chapter-2) will act as a positive contribution to the debate about how best to graphically model pathway knowledge.

Chapter Contributions and Acknowledgements

This page has been left intentionally blank

Chapter 4. Analysis of the transcriptional networks induced by type I and type II interferons

Introduction

Transcriptional response of macrophages to interferons

IFNs induce or repress the expression of hundreds of genes known to mediate a wide range of biological responses. Transcriptional targets of the interferon response are collectively referred to as Interferon Stimulated Genes (ISGs). Presently around 2,000 [231] ISGs have been identified by transcript profiling. However these have been derived from a variety of studies based on the analysis of different cell types using different array platforms and only a few of the genes are functionally well defined. Some of the best studied ISGs [232] are known to play pivotal roles in host defence including; the double-stranded RNA-dependent kinase "protein kinase RNA-regulated" (PRKR) [233], a family of 2'-5'-Oligoadenylate Synthetases (2'5'OAS) that lead to the activation of RNase L and degredation of cellular RNA [234], and the Mx proteins (these possess GTP binding and GTPase activity) and have been shown to restrict growth of certain viruses [235-236]. However the anti-viral effects of IFNs can only be partially attributed to these genes since mice triply deficient for PRKR, RNase L and Mx genes retain a degree of responsiveness to the antiviral effects of IFNs [237]. Hence alternative antiviral pathways and other ISGs yet to be (fully) functionally characterised may yet play role as potent antiviral effectors.

Some ISG's are regulated by both type-I and type-II IFNs i.e. there is a degree of overlap in the transcriptional networks they activate, whereas others are selectively regulated. *IFITM1* for example is induced by all interferons whereas 2'-5'-Oligoadenylate Synthetase I is induced in response to type-I interferons IFN-α and IFN-β, and not the type-II IFN-γ [238]. *IRF1* expression on the other hand is preferentially induced by IFN-γ [238]. Moreover the activation of MHC Class II antigen presentation in macrophages is only efficiently achieved by IFN-γ regulation of the *CIITA* gene [239-240](depicted in Chapter-3 Figure 3.7). Selective regulation of ISGs by type-I or type-II IFN signalling could be attributed to the activation of different downstream signalling cascades, differential activation of transcription factors, and variation of promotor

elements in target genes; since some ISGs contain only ISRE's and others only GAS elements in their promotors, whereas others have both elements.

The JAK-STAT signalling cascades immediately downstream of the IFN type-I and type-II receptors is well characterised (and represented on our pathway model) however their links to other signalling pathways and the precise regulation of the transcriptional response is not well characterised or fully understood.

Functional relevance of type I and type II signalling convergence

The type-I and type-II signalling pathways cross-talk at multiple levels; sharing pathway components and overlapping in their transcriptional targets (see Chapter-1; Figure 1.3). IFN-γ primarily signals through STAT1:STAT1 homodimers, although it also activates to a lesser extent the archetypal type-I transcription factor ISGF3. ISGF3 is able to induce the expression type-I interferon and thereby amplify its response. Conversely type-I interferon can activate classical type-II signalling molecules e.g STAT1 and thereby modulate the transcription of ISGs with GAS elements. The cross-talk and convergence of these signalling pathways is biologically relevant, since *in vivo* cells are not generally exposed to single cytokines but rather a cytokine cocktail [59].

It has been suggested that the two interferon systems may have evolved to complement each other in overlapping but non-redundant activities in order to defend against a broad range of pathogens [241]. IFN-γ$^{-/-}$ and IFNGR1$^{-/-}$ mice appear to develop a normal immune system, however they show deficiencies in natural resistance to bacteria, parasitic and viral infections [241]. Some viruses appear to require both type-I and type-II pathways [241] and others require predominantly type-I or type-II interferon for efficient clearing [242]. Since IFNs illicit their response by activating a large transcriptional cascade, the differential regulation of genes by the two types of interferons warrants further analysis to advance our understanding of the overlap and differences in functionality of the different interferons.

Microarray experiments examining the transcriptional changes induced by interferons are available [238, 243-247] however those based on the response induced in macrophages (and specifically mouse BMDMs) are not as comprehensive as those investigating the LPS response. There are also limitations to some of the existing available microarray data. For example, studies conducted prior to 1995 are not genome wide; studies conducted before 2001 are not MIAME (m̲inimum i̲nformation a̲bout a m̲icroarray e̲xperiment) compliant and therefore unavailable as raw data files or poorly described and annotated. Other studies are not necessarily performed on robust platforms, for instance those conducted on spotted arrays [247]. Finally some experiments have a limited number of data-points (time-points analysed) and the overall experimental design is not conducive to exploring the temporal changes in transcription induced by interferons or contrasting the type-I/type-II response.

Results

In order to further our understanding of the transcriptional events in response to stimulation of mouse bone marrow derived macrophages (BMDMs) by type-I and type-II IFNs, microarray analysis was performed over a time-course of stimulation by these cytokines. IFNs are well known for their anti-proliferative effects. Murine BMDMs (when cultured in CSF-1) are actively in cell cycle; do generally not constitutively express IFN-β; and are MHC Class-II negative [11]. Thus this system includes a biology of IFNs not easily seen in other cellular systems. In separate experiments mouse BMDMs were stimulated with either 10 U/ml mouse IFN-γ or 10 U/ml mouse IFN-β, where one U/ml is defined as the concentration required to inhibit viral replication by 50%. In the case mouse IFN-γ and IFN-β biological activity is determined (by the manufacturers) by measuring the ability to induce cell resistance to infection by encephalomyocarditis virus (EMCV).

In both experiments, cells were harvested for 1, 2, 4, 8, and 24 h post-IFN treatment or pre-treatment (0 h). High quality RNA was processed for labelling and hybridisation to the Affymetrix Mouse Exon 1.0 ST arrays.

IFN-γ transcriptional network

12 Affymetrix Mouse Exon 1.0 ST Arrays were processed for this study although one array failed due to array scanning-equipment failure. This resulted in the loss of one of the two, 2-hour biological replicate sample. Remaining arrays were normalised using RMA. Statistical filters were applied as a method for removing noisy and potentially un-interesting data. I,678 transcript passed this filter and a network-graph of the data was generated by filtering edges at a Pearson correlation threshold of 0.9 . The resultant network graph (Figure 4.1) of 1,474 nodes connected by 26,617 edges was clustered using the graph-based clustering algorithm MCL [137] set at an inflation value of 2.2 resulting in 40 clusters with a membership of 6 or more nodes (transcripts).
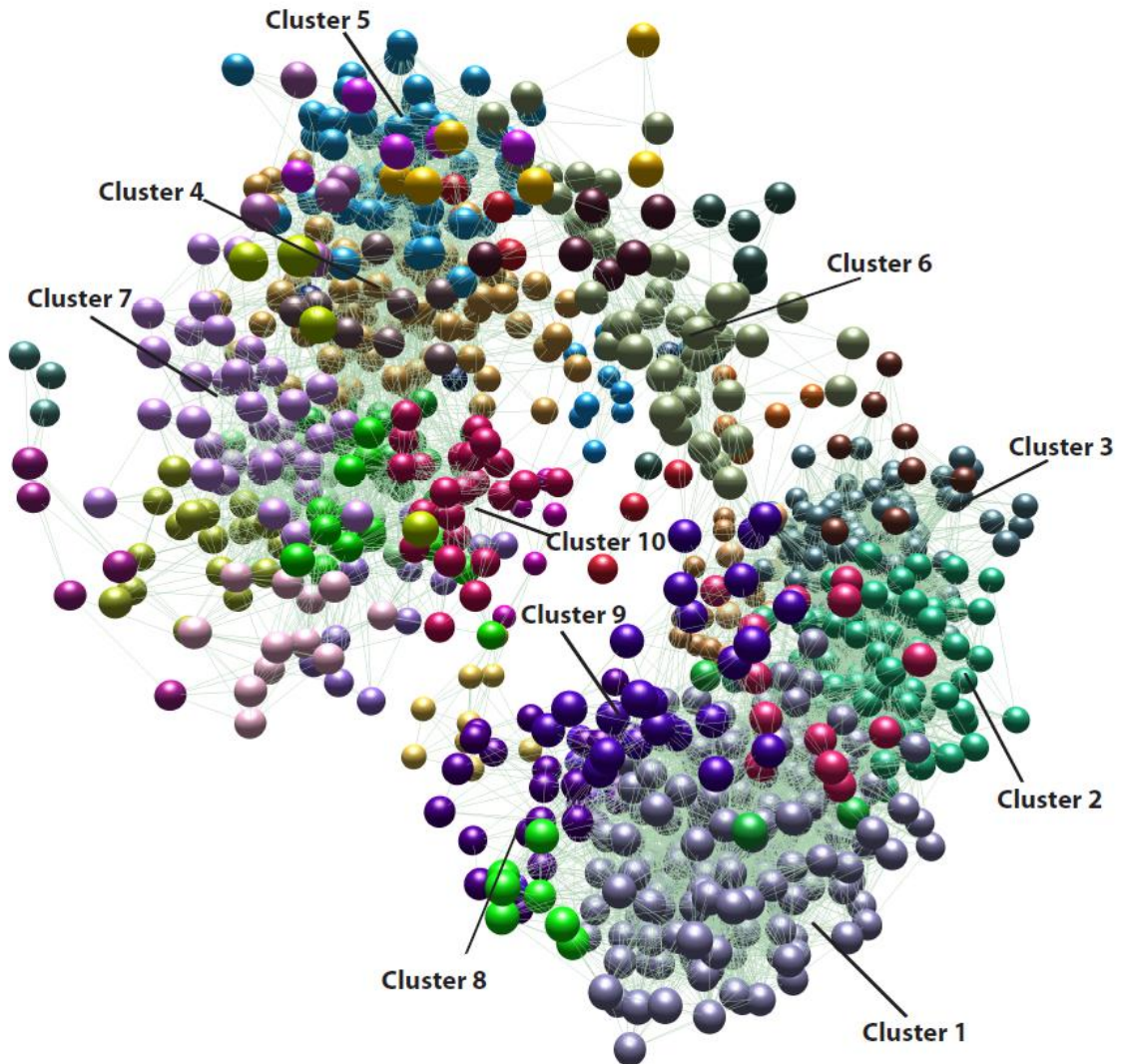
Figure 4.1: Transcriptional network formed from expression data of a time-course of IFN-γ stimulation of mouse BMDMs. Mouse BMDMs were treated for 1, 2, 4, 8, and 24 h with IFN-γ or were not treated (0 h). Gene expression across each time-point was measured on Affymetrix Mouse Exon ST1 Arrays and a network graph of the data generated using BioLayout *Express*[3D]. The network was filtered to display only relationships at or above a Pearson correlation threshold of 0.9, resulting in a graph 1,474 nodes connected by 26,617 edges. The resultant network was then clustered using the graph-based clustering algorithm MCL set at an inflation value of 2.2. Nodes (transcripts) belongings to same cluster share the same colour.

Description of clusters

Each cluster in the network represents a group of genes with a related pattern of expression over the time-course. Some clusters represent very similar patterns of expression, however form separate clusters due to subtle difference in their expression patterns. Individual clusters were assigned a description based on the expression pattern the cluster represented. Descriptions were based on a number of characteristics of the cluster, including;

(i)   the directionality of the change; where **Up** or **Down** defines up regulation or down regulation of transcripts within the cluster relative to time 0 h.

(ii)   The dynamics of the temporal change in expression; where **transient** implies the change in expression is reversible / not continuous/ occurs over given periods of time and then returns to basal levels. **Sustained** implies the change in expression is maintained over the time points studied.

(iii)   The time point(s) where maximal or minimal expression is reached in comparison to basal levels. This is denoted by the number(s) following the Up or Down description.

(iv)   The duration over which any change in expression is taking place; this is denoted by the numbers in brackets.

(v)   The original identifying cluster number is shown at the end of description.

For example the following description "**Ifng_transient_Up_8_(4-8)_C1**" would imply transcripts belonging to this cluster one are up-regulated in expression, reach maximal expression at 8 hours, however the increase in expression is already apparent from 4 hours and lasting until 8 hours. The same cluster naming scheme is adopted to describe clusters from the IFN-β time-course study.

Clusters were annotated based on their gene membership and over-representation of cohorts of functionally associated genes. Annotation was performed using the online Database for Annotation, Visualization and Integrated Discovery (DAVID) which

identifies enriched biological themes within gene lists [248-249]. For the purpose of annotating these datasets GO (Gene Ontology) annotations defining both biological processes and metabolic function were assessed using DAVID. The most over-represented terms (based on DAVID analysis of GO Ontology terms (GO FAT category)) were chosen to describe the biology of the clusters. Table 4.1 provides an overview of the major clusters of interest in the IFN-γ response data set, along with the number of transcripts within each cluster, examples of gene members and the most represented GO terms as determined by DAVID. The clusters are ordered in the table according to the temporal phase the transcripts are changing in expression; immediate early (1-2 hours), early (2-4 hours), mid (**4** > 8 hours), mid to late (4 < **8** hours) and late (8-24 hours).

Temporal Response	Cluster No.	No. of transcripts in cluster	Cluster Description	Example Gene Member	Processes/ Functions
Immediate Early response	16 ↑	13	**Ifng_transient_Up_1_C16**	Fos, Myc, Cxcl1	Positive regulation of transcription; positive regulation of gene expression; positive regulation of RNA metabolic process; negative regulation of apoptosis
	28 ↑	7	**Ifng_transient_Up_1_(1-2)_C28**	Ier3, Ccl3, Nfkbiz	Defence response; positive regulation of gene expression /transcription
	23 ↑	9	**Ifng_transient_Up_2_(1-2)_C23**	Fosl2, Socs3	Regulation of phosphorylation; regulation of intracellular signalling cascade
Early response	25 ↑	8	**Ifng_Early-sustained_Up_4_(1-4)_C25**	Ccl2, Ccl7,Ifi27	Response to organic substance (nitrogen); regulation of cell cycle, chemokine activity.
	11 ↑	21	**Ifng_transient_Up_2-4_C11**	Il17ra, Bach1	Response to wounding; regulation of phosphoinositide 3-kinase cascade; regulation of transcription
	38 ↓	7	**Ifng_transient_Down_2-4_C38**	Fgd3	guanyl-nucleotide exchange factor activity
Mid- response	22 ↓	10	**Ifng_transient_Down_4_(2-4)_C22**	Kns2	*Annotation Stats Unavailable*
	3 ↑	131	**Ifng_transient_Up_4_C3**	Map2k3, Foxp1, Stat5a	Positive regulation of cell proliferation/ differentiation; regulation of mesenchymal cell proliferation; regulation of transcription
	24 ↓	8	**Ifng_transient_Down_4_C24**	Oma1	Transcription factor binding
	36 ↓	7	**Ifng_transient_Down_4_C36**	Irak3	*Annotation Stats Unavailable*
	37 ↓	7	**Ifng_transient_Down_4_C37**	Ccdc128	Ribonucleotide binding
	2 ↑	175	**Ifng_transient_Up_4_(4-8)_C2**	Tnf, Daxx, Fas, Aqp9, Cxcl10	positive regulation of NF-κB transcription factor activity; immune response; regulation of apoptosis; regulation of transcription
	4 ↓	84	**Ifng_transient_Down_4-8_C4**	Mcm2, Pole, Cdc2a	DNA replication; DNA replication initiation, cell cycle, DNA repair, DNA packaging, DNA binding,
	5 ↓	62	**Ifng_transient_Down_4_(4-8)_C5**	Cdc6, Cep55, Ccne2	Cell cycle; cell division; mitotic cell cycle; M phase of cell cycle; organelle fission; chromosome segregation; DNA replication
	30 ↓	7	**Ifng_transient_Down_4_(4-8)_C30**	Fli1	Ion binding
	31 ↓	7	**Ifng_transient_Down_4_(4-8)_C31**	Mvd	ATP binding
	10 ↓	24	**Ifng_transient_Down_4-8_C10**	Gab3, Coro1a	Nucleotide receptor activity, G-protein coupled
	27 ↓	8	**Ifng_transient_Down_4-8_C27**	Dusp19, Ing4	*Annotation Stats Unavailable*
	26 ↓	8	**Ifng_4-min_8-max_Down_Up_4 & 8_C26**	Rab32, Prkcd	Cell activation; leukocyte activation
Mid-late response	1 ↑	224	**Ifng_transient_Up_8_(4-8)_C1**	Traf1, Traf2, Irf5, Nfkb1, Nfkb2, Nod2, Tlr2	(Innate) Immune Response; programmed Cell Death; Defence response;
	7 ↓	39	**Ifng_transient_Down_8_(4-8)_C7**	Ccdc5, Mxd4, Alox5	Oxidation reduction; fatty acid metabolic process; coenzyme binding
	12 ↓	21	**Ifng_transient_Down_8_(4-8)_C12**	Scamp5	Lipid localization / transport; fatty acid transport; vesicle mediated transport
	13 ↓	20	**Ifng_transient_Down_8_(4-8)_C13**	Scd2	Organic acid catabolic/ biosynthesis process; fatty acid metabolic process
	14 ↓	16	**Ifng_transient_Down_8_(4-8)_C14**	Aldoc	*Annotation Stats Unavailable*
	17 ↑	13	**Ifng_Up_8_(4-8)_C17**	Entpd1	Response to extracellular stimulus; protein maturation; protein processing
	8 ↑	27	**Ifng_transient_Up_8_C8**	Nfe2l1, Bag3	Negative regulation of apoptosis; regulation of RNA metabolic process
	15 ↓	15	**Ifng_transient_Down_8_C15**	Gas6	Protein transport; establishment of protein localization
Late Response	18 ↓	13	**Ifng_sustained_Down_8_(4-24)_C18**	Cd14	*Annotation Stats Unavailable*
	9 ↑	24	**Ifng_late-response_Up_24_(8-24)_C9**	Psmb9, Psma7	Proteolysis; peptidase activity; MHC class I protein binding; protein catabolic process
	6 ↑	47	**Ifng_late-response_Up_24_C6**	H2-Ea, Il2rg, Ccr5	Antigen processing and presentation of exogenous antigen; oxidation reduction; antigen processing and presentation of peptide antigen via MHC class II

Table 4.1: Description of clusters of co-ordinately expressed transcripts in response to IFN-γ stimulation of mouse BMDMs. Clusters are arranged according to the timing of response they reflect.

137

During the immediate early response (1-2 hours post IFN-γ) there were relatively few genes changing in expression as compared to the rest of the data. During this early phase a positive regulation of transcription was observed including the regulation of many transcription factors (TFs). These TFs may control the next waves of response observed in the data. Also during this immediate early phase there are indications of cytokine signalling regulation; for example the suppressor of cytokine signalling 3 (Socs3) known to control the extent and duration of the interferon response, increases very transiently in expression. Indications of an immune response are observed (chemokines/defence response) in a handful of changing transcripts during the early phase.

During the mid-response there is initially an increase in expression of genes contributing to a positive regulation of cell proliferation, however this closely followed by a strong signal to repress cell cycle (clusters-5 and 6). The mid-response stage also comprises the first major immune signalling response cluster (2) enriched for innate immune, apoptosis, and NF-kB signalling transcripts. However the largest cluster (1) containing the major interferon-stimulated-genes, appears from mid-late reaching maximal expression at 8 hours. Also during this mid-late phase fatty acid metabolism, lipid transport and metabolic processes are repressed. Interestingly and in contrast to cluster-2, there is an up-regulation during the mid-late phase of negative regulators of apoptosis. Finally during the late response genes involved in proteolysis and antigen processing and presentation are expressed.

IFN-γ Network Architecture

To better visualise the architecture of the IFN-γ transcriptional-response network, a hierarchical interaction network of the clusters and their relationships was generated. (Figure 4.2). With the aid of an automated layout algorithm the cluster relationships were arranged to flow from left to right and in this direction display the temporal changes in transcription (from early to late response). Up regulated components

(coloured red) form discreet clusters with distinct expression profiles and three of these are the largest in the data set. The down-regulated graph component (in green) on the other hand is made up of comparatively smaller clusters with more intricate connections. The largest portion of transcriptional changes occurred during the mid-period (taking place at 4 h post treatment). In comparison fewer genes were regulated at the early (1-2 h) or late stages (24 h). Furthermore only up-regulated clusters formed the immediate early response.

To better visualise the temporal changes in gene expression, the average expression profile of all transcripts within clusters across a given time phase (immediate early, early, mid, mid-late and late) was plotted across the 11 arrays (Figure 4.3). Average expression for up and down regulated clusters was calculated separately, and the average maximal/ minimal expression varies across the different time phases.
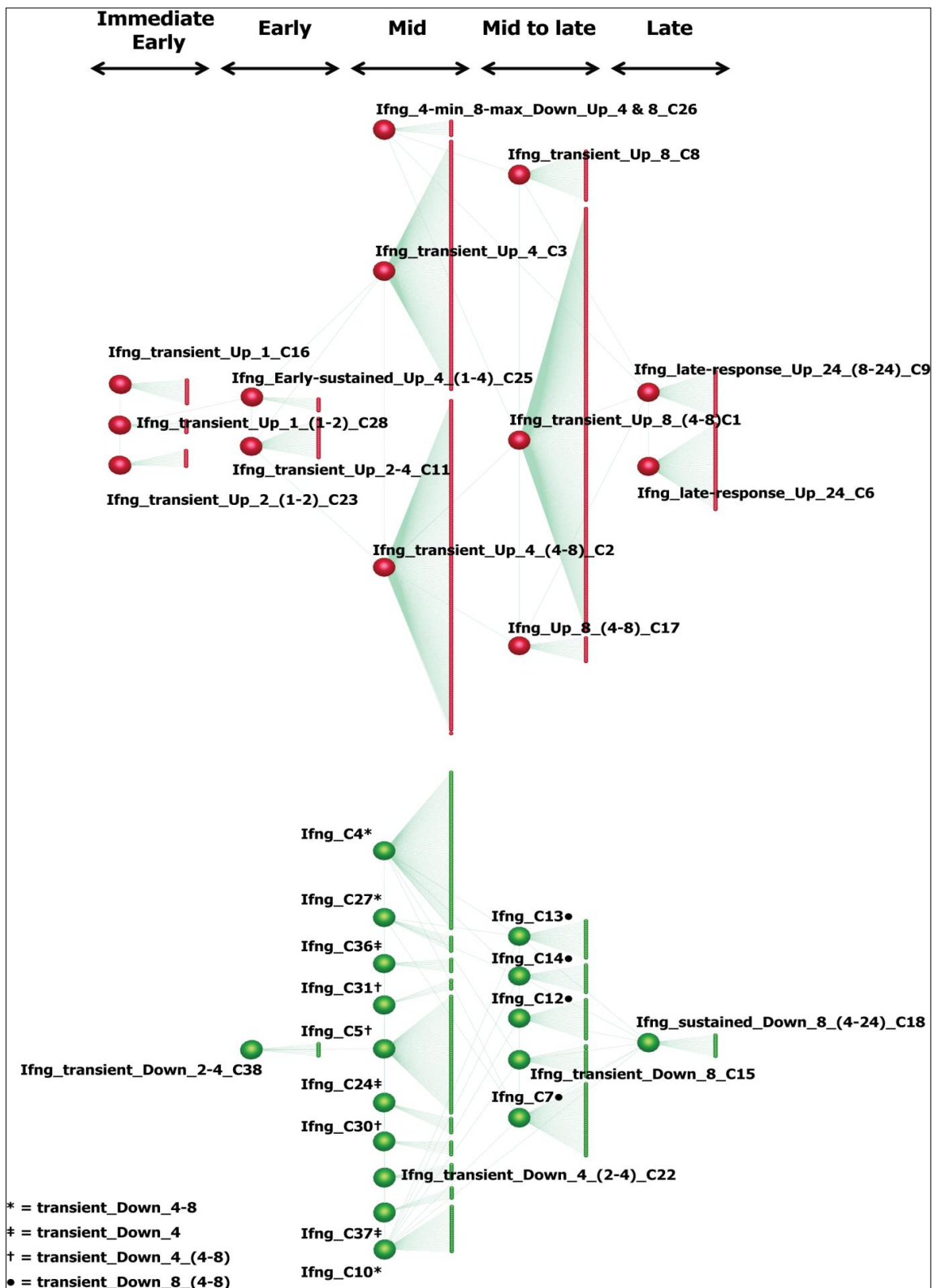
Figure 4.2: Hierarchical interaction network of the IFN-γ transcriptional response in mouse BMDMs. Clusters relationships are shown by edges connecting the large red or green spheres, representing up (red) or down (green) regulated components of the response. Cluster membership (and relative size) is shown by the smaller red or green spheres connected to the larger cluster node and clusters are arranged to follow the order of temporal changes in expression. Transcripts within immediate-early clusters change in expression ~ 1-2 hours post treatment, those in early clusters ~ 2-4 hours, mid ~ 4-8 (4>8) hours, mid to late ~ 4-8 (4<8) hours and late ~8-24 hours.

Ifn-γ transcriptional profiles

Up- response

Down- response

Expression

Time post stimulation (hours)

Time post stimulation (hours)
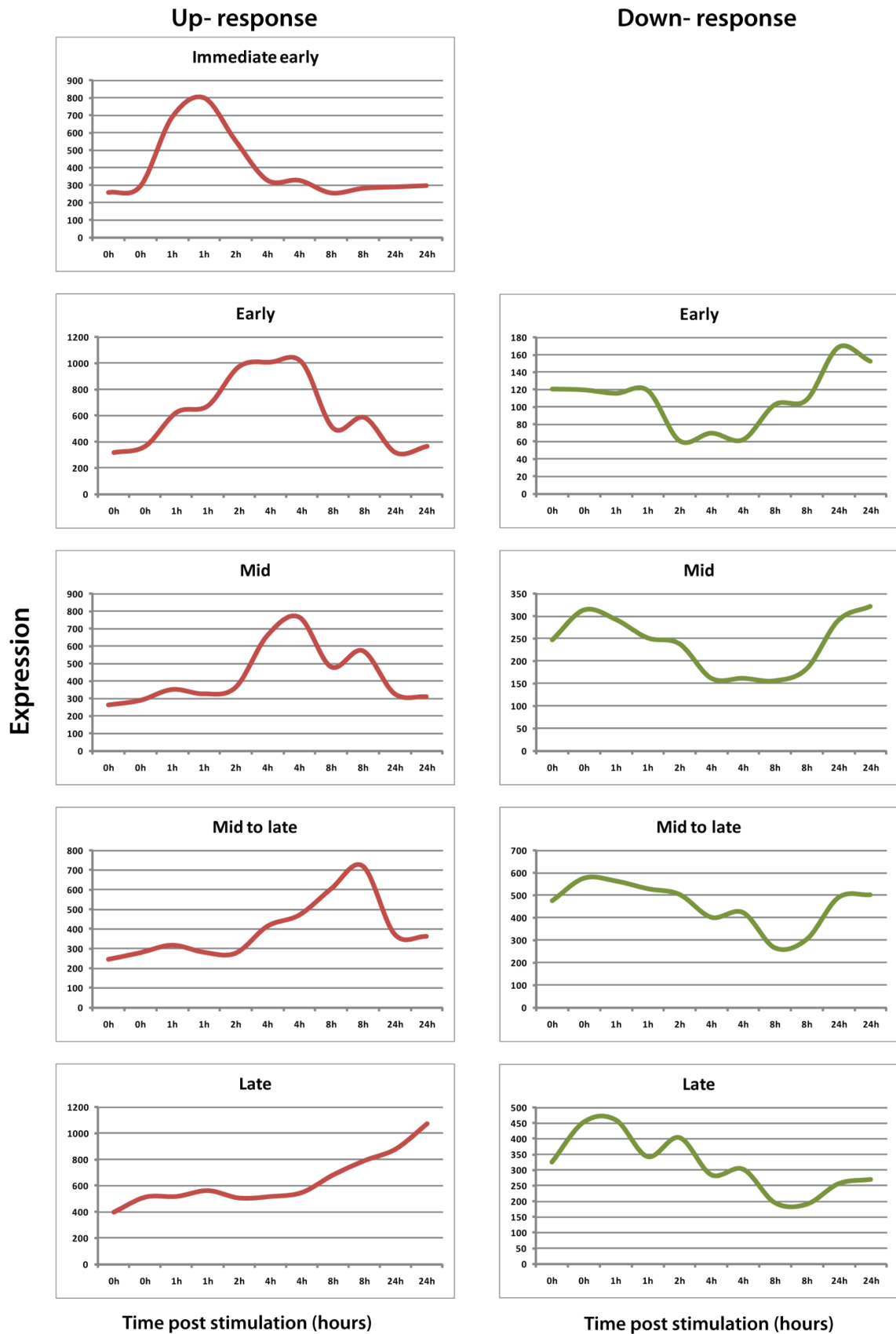
Figure 4.3: The average expression of transcripts stimulated (red) or suppressed (green) in response to IFN-γ treatment of BMDM across different temporal phases of the transcriptional response.

IFN-β transcriptional network

12 Affymetrix Mouse Exon ST Arrays were successfully processed for this study all passing the QC. As with the IFN-γ-analysis raw data was normalised using RMA and statistical filters were applied to the normalised data as a method for removing noisy and potentially un-interesting data. A network graph of the filtered data was generated resulting in a graph of 2,045 nodes, connected by 92,947 edges at a Pearson correlation threshold of 0.9 or above (Figure 4.4). The resultant graph was clustered using the graph-based clustering algorithm MCL set at an inflation value of 2.2 resulting in 33 clusters (with greater than 5 nodes). 18 of these clusters were considered to represent interesting patterns of expression. As with the IFN-γ network graph the IFN-β network also comprises two main graph components representing up-regulated transcripts or down regulated transcripts.
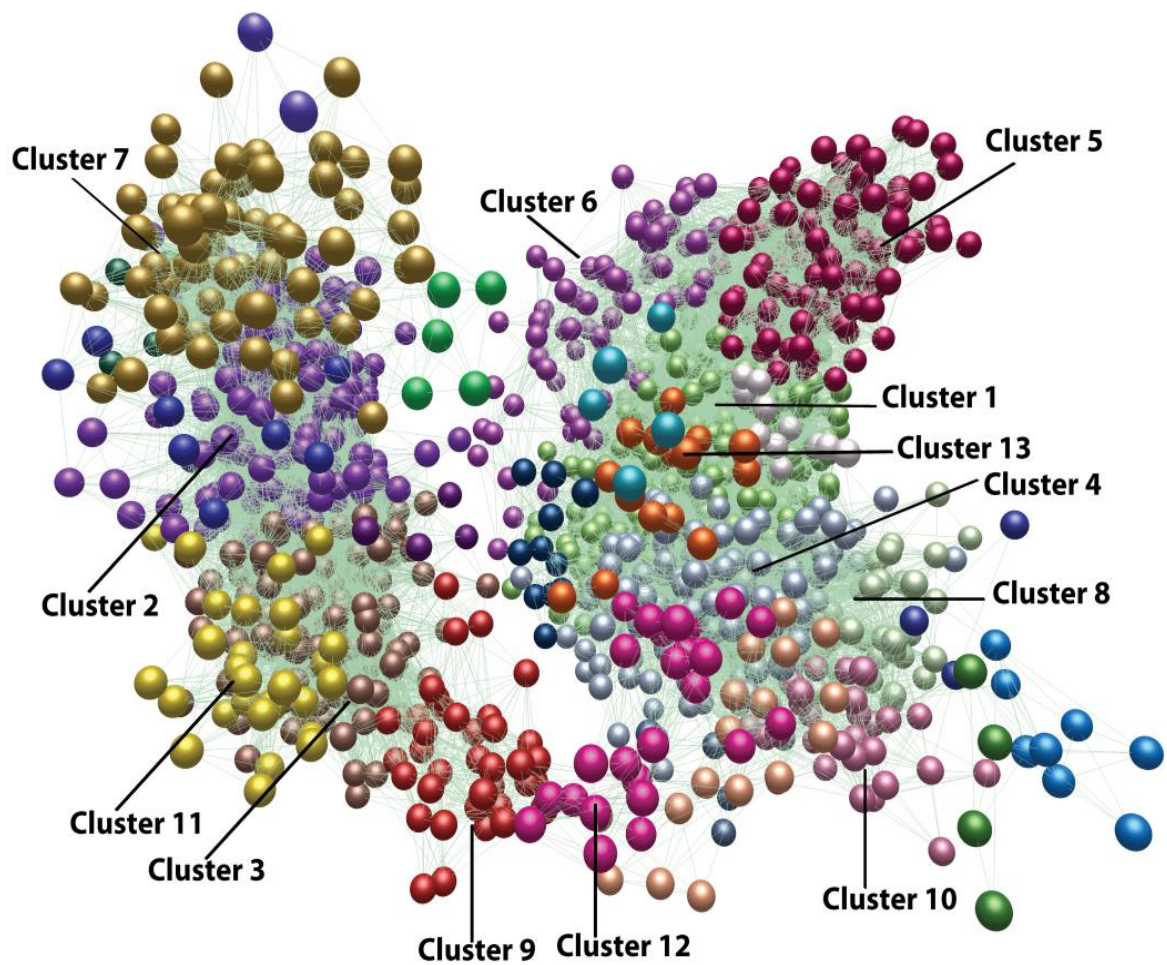
Figure 4.4: Transcriptional network formed from expression data of a time-course of IFN-β stimulation of mouse BMDMs. Mouse BMDMs were treated for 1, 2, 4, 8, and 24 hours with IFN-β or were not treated (0 hours). Gene expression across each time-point was measured on Affymetrix Mouse Exon 1.0 ST Arrays and a Network graph of the data generated using BioLayout *Express*[3D]. The network was filtered to display only relationships at or above a Pearson correlation threshold of 0.9, resulting in a graph 2,045 nodes connected by 92,947 edges. The resultant network was then clustered using the graph-based clustering algorithm MCL set at an inflation value of 2.2. Nodes (transcripts) belongings to same cluster share the same colour.

Using the same cluster description process as used with the IFN-γ dataset, the clusters were assigned a description based on the timing and duration of the expression pattern they represented. The IFN-β clusters were also annotated based on their gene membership and over-representation of cohorts of functionally associated genes using DAVID analysis of GO biological and metabolic process annotations. Table 4.2 summarizes the major clusters of interest in the IFN-β response data set arranged according to the temporal phase of the response they represent (early (1-2 h), early (2-4 h), mid (4-8 h), mid to late (4-8 h) and late (24 h)).

The immediate early response consisted of 16 transcripts increasing in their expression, including the suppressor of cytokine signalling, Socs3. The function of many of the genes represented within the immediate early clusters is currently poorly characterised and the only relevant annotation obtained from DAVID was 'receptor linked signal transduction'. Regulation of phosphorylation and transcription factor activity were the predominant signals during the early-phase of the IFN-β response, although this phase also saw the repression of genes related transcription-regulation and MAPKinase phosphatase activity. The mid-phase response consists of two of the largest clusters (clusters 1 and 2), both of similar size (379 and 376 transcripts respectively). Transcripts within cluster-2 reached maximal expression around 4 h post IFN-β stimulation and are enriched for many of the well-known interferon-response genes including; *Ifit1, Ifit2, Oasl1*. In contrast, cluster-1 represents the repression of cell-cycle related activity including DNA replication, DNA repair and mitosis. By the mid-to-late phase only 52 transcripts are represented in the Up-regulated portion of the network graph and 81 transcripts in the down-regulated side. The late response predominantly comprises the repression of additional cell cycle related genes (cluster-5), as well the induction of a handful of transcripts associated with T- and B-cell activation, potentially suggesting facilitation of adaptive immune response.

Temporal Response	Cluster No.	No. of transcripts in cluster	Cluster Description	Example Gene Member	Processes/ Functions
Immediate Early	18 ↑	10	**Ifnb_transient_Up_1-2_C18**	Socs3, Nfkbiz	*Annotation Stats Unavailable*
	30 ↑	6	**Ifnb_transient_Up_2_C30**	Olfr410 Olfr568 Trib3	Cell surface receptor linked signal transduction
Early	7 ↑	75	**Ifnb_transient_Up_2_(1-4)_C7**	Socs1,Irf1, Nod2	Regulation of phosphorylation, transcription factor activity, response to wounding hemopoiesis
	10 ↓	32	**Ifnb_transient_Down_2_(1-4)_C10**	Mapk7, Zfp52	Regulation of transcription; zinc ion binding; MAPkinase phosphatase activity.
	23 ↓	7	**Ifnb_transient_Down_2_(1-4)_C23**	Phf23 Zfp275 Zfp708	Zinc ion binding
	19 ↑	9	**Ifnb_sustained_Up_4_(1-8)_C19**	Rab9 Rtp4	Negative regulation of cell proliferation
Mid Response	2 ↑	376	**Ifnb_transient_Up_4_(1-4)_C2**	Casp7, Cd40, Ifit1, Irf5, Nod1	Immune response; defense response; cytokine activity; regulation of leukocyte activation; regulation of leukocyte proliferation; innate immune response; regulation of apoptosis
	3 ↑	208	**Ifnb_transient_Up_4_(2-4)_C3**	Stat1, Stat2, Gbp1, Psmb9	Immune response; ribonucleotide binding; antigen processing and presentation
	8 ↓	40	**Ifnb_transient_Down_4_(2-4)_C8**	Cdca7, Zfp60	Regulation of transcription, DNA binding, zinc ion binding, chromatin organization
	1 ↓	379	**Ifnb_transient_Down_4_C1**	Brca1, Cdk2, Mcm2, Mdm1	Cell cycle; DNA metabolic process; DNA replication; cellular response to stress; DNA repair; M phase; mitotic cell cycle
	4 ↓	197	**Ifnb_transient_Down_4_C4**	Dusp3, Mapkapk3,	Nucleotide receptor activity-G-protein coupled; dephosphorylation; transcription regulator activity; ,
Mid-late Response	16 ↓	13	**Ifnb_4-min_24-max_Down_Up_4_&_24_C16**	Cxcr3 Gas6	*Annotation Stats Unavailable*
	9 ↑	39	**Ifnb_sustained_Up_8_(1-8)_C9**	Ccl5, Tlr7	Immune response; inflammatory response; antigen processing and presentation; immune effector process; chemokine activity.
	6 ↓	81	**Ifnb_transient_Down_8_(4-8)_C6**	Pdk1, Rak3	Oxidation reduction; response to oxidative stress
Late Response	15 ↓	15	**Ifnb_sutained_Down_2-24_C15**	Ccdc18 Dus4l Fbxo5	Meiosis, M phase of Meiotic cell cycle, tRNA metabolic process
	5 ↓	173	**Ifnb_sustained_Down_24_(2-24)_C5**	Bub1, Kif22, Serpinb10	Cell cycle; cell division, M phase, mitotic cell cycle, nuclear division, organelle fission, DNA replication
	14 ↑	20	**Ifnb_sustained_Up_8-24_C14**	Fdg4, Tmem178	*Annotation Stats Unavailable*
	12 ↑	30	**Ifnb_sustained_Up_24_(8-24)_C12**	Ccr5, Cd28, Ly9	Positive regulation of interleukin-2 production, T cell selection, T cell proliferation, positive regulation of B cell activation.

Table 4.2: Description of clusters of co-ordinately expressed transcripts in response to IFN-β stimulation of mouse BMDMs. Clusters are arranges according to the timing of response they reflect.

IFN-β Network Architecture

In order to better visualise the structure of the transcriptional response network of mouse BMDMs to IFN-β stimulation a hierarchical interaction network of the clusters and their relationships was generated (Figure 4.5). As with the IFN-γ network (Figure 4.2) the clusters are arranged to follow the temporal response (moving from early to late) and the 'up' and 'down' regulated graph components are shown in red and green respectively. The overall sizes of the Up and Down graph components is comparable, suggesting the extent of transcriptional induction is similar to the extent of transcriptional repression. The majority of transcriptional changes occur during the mid-phase, and the immediate-early response is reserved to up-regulated transcripts. The average expression profile of transcripts within clusters across a given time phase (immediate early, early, mid, mid-late and late) is shown in Figure 4.6. The average expression for up and down regulated clusters was calculated separately, and the average maximal/minimal expression varies across the different time phases.
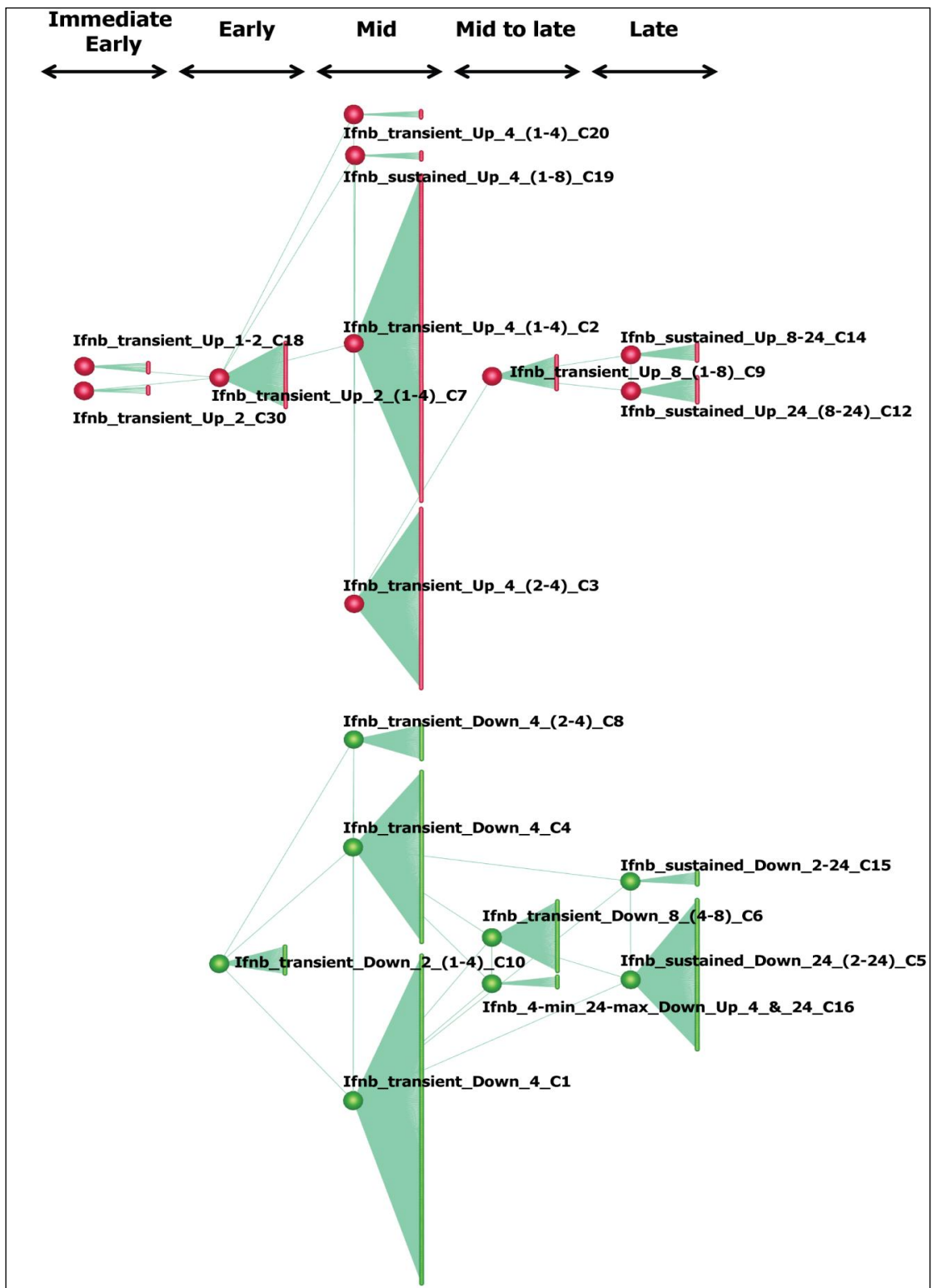
Figure 4.5: Hierarchical interaction network of the IFN-β transcriptional response in mouse BMDMs. Cluster relationships are shown by edges connecting the large red or green spheres, representing up (red) or down (green) regulated components of the response. Cluster membership (and relative size) is shown by the smaller red or green spheres connected to the larger cluster node and clusters are arranged to follow the order of temporal changes in expression. Transcripts within immediate-early clusters change in expression ~ 1-2 h post treatment, those in early clusters ~ 2-4 h, mid ~ 4-8 (4>8) h, mid to late ~ 4-8 (4<8) h and late ~8-24 h.

Ifn-β transcriptional profiles
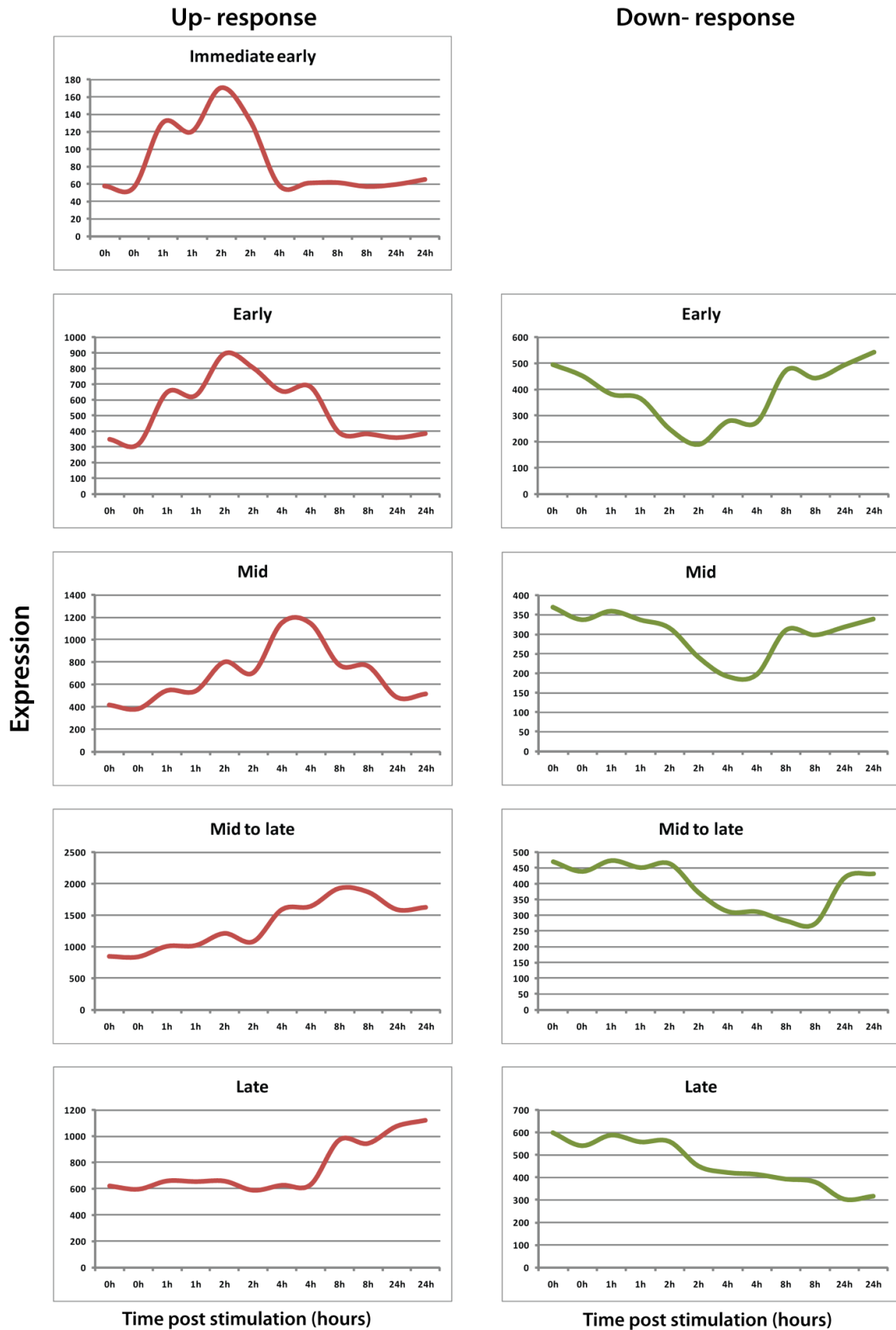
Up- response **Down- response**

Expression

Time post stimulation (hours) **Time post stimulation (hours)**

Figure 4.6: The average expression of transcripts stimulated (red) or suppressed (green) in response to IFN-β treatment of BMDM across different temporal phases of the transcriptional response.

148

Comparison of transcriptional responses of BMDMs to IFN-β and IFN-γ

The IFN-β and IFN-γ time-course experiments were originally conceived and analysed separately. Thus every aspect of the experimental set-up (from cell culture to microarray processing) was performed separately. However given our desire to understand the interplay between the type-I and type-II response [146], methods for comparing the two responses were explored. Ideally for comparative purposes the two experiments would have been set up and processed alongside each other. Nevertheless there was some legitimate reasons to believe that amalgamating the raw expression data from the two data sets and analysing these together would work, given the similarities of the experimental set up and the use of the same microarray platform. To explore whether the combined analysis of the two time-courses was viable the raw data sets were normalised together and some basic QC steps performed. A Pearson correlation matrix (Figure 4.7) of the normalised signal intensity across all 23 arrays (12 IFN-β and 11 IFN-γ) illustrated that the array samples correlate more strongly by the date of the experiment rather than by time post-treatment. In particular the 0 h IFN-β and 0 h IFN-γ pre-treatment samples do not appear closely correlated, as would be expected with un-treated samples of the same cell lineage, implying the concatenation of two datasets may not generate robust normalised data. This was further corroborated by plots of the average normalised expression of negative control probes on the arrays (Figure 4.8) where expression appeared elevated across the IFN-β arrays. Expression of control probes would be expected to be of similar intensity across all arrays but markedly different normalised values for negative control probes would further indicate fundamental differences between the two datasets.
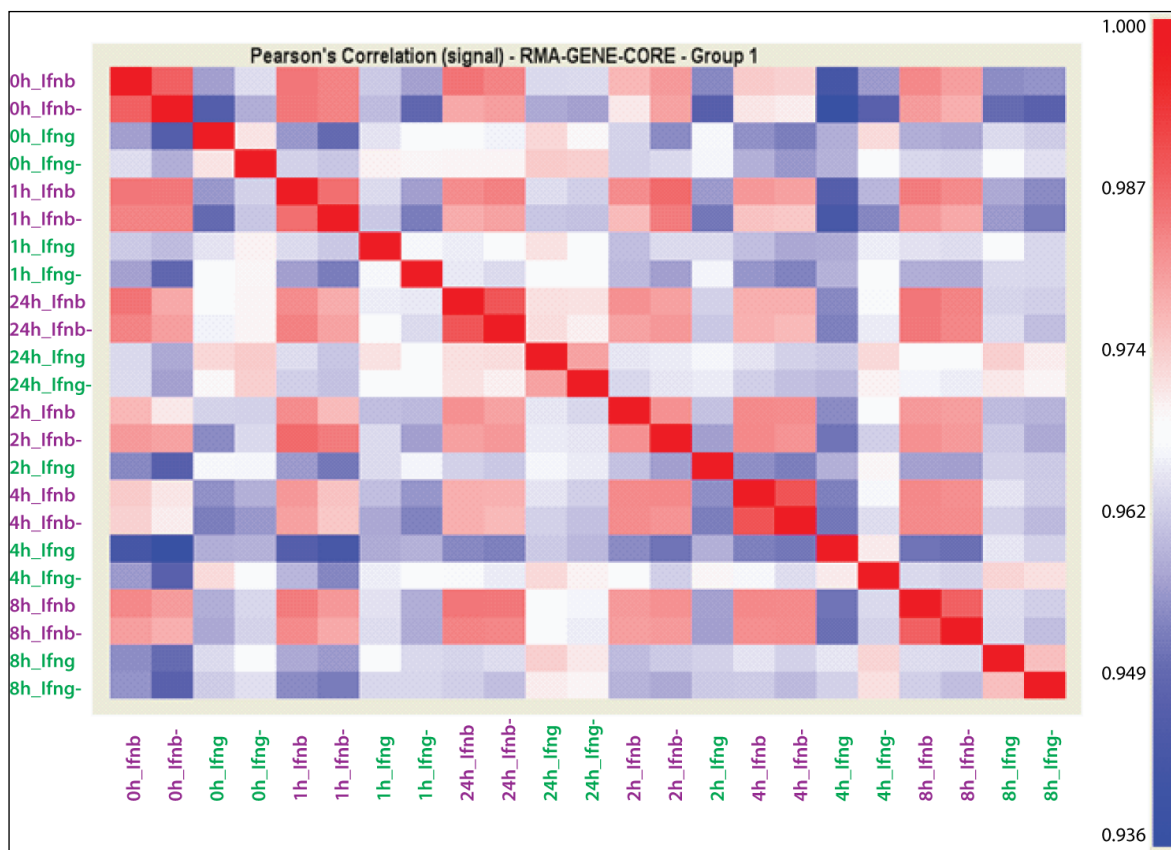
Figure 4.7: A Pearson correlation matrix of the normalised signal intensity across 23 microarrays (12 IFN-β and 11 IFN-γ). Raw data (CEL intensity files) from both the IFN-β and IFN-γ time-course experiments were normalised together using RMA. A Pearson correlation matrix of the signal intensity across the arrays was generated to gain an indication of which arrays were most correlated.
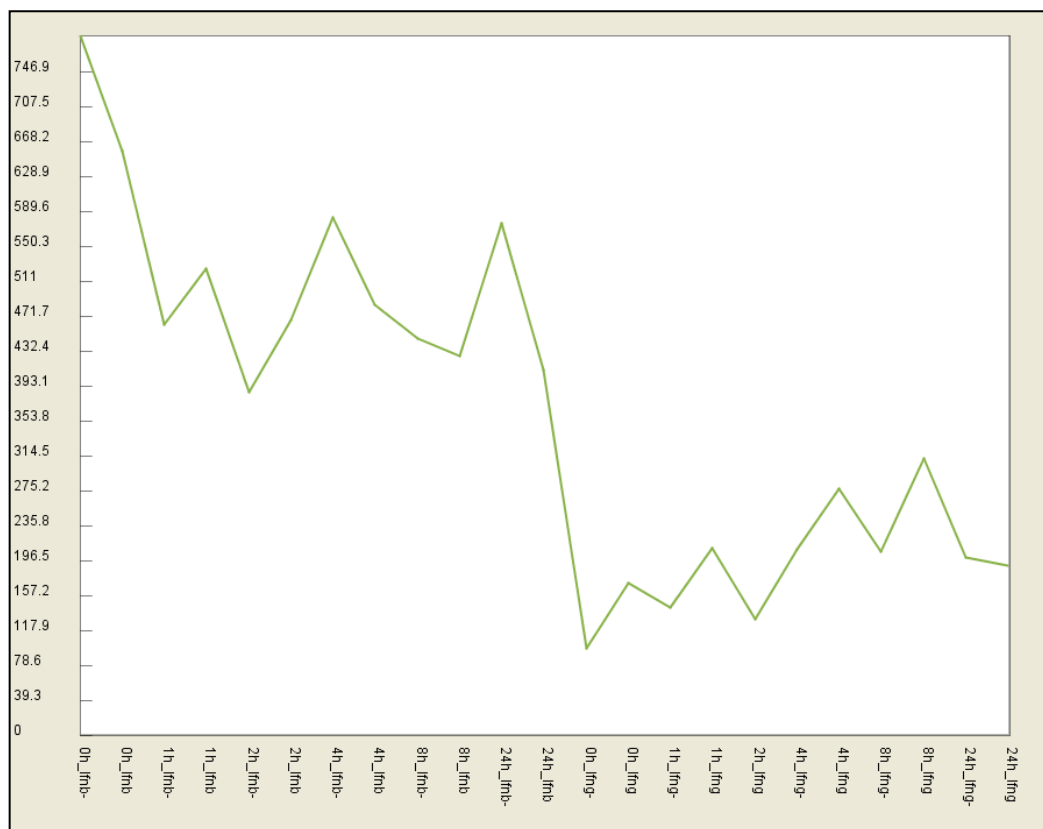
Figure 4.8: Average expression of Affymetrix negative control probes across 23 microarrays sampling IFN-β or IFN-γ treatment of macrophages. Raw data (CEL intensity files) from both the IFN-β an IFN-γ time-course experiments were normalised together using RMA. The average normalised expression of negative control probes on all 23 arrays was plotted.

The distinct differences of two sets of samples processed on different occasions impeded their combined analysis. However the network analyses of the individual data sets can still be utilized to compare and contrast the transcription response to these cytokines. Lists of co-ordinately expressed transcripts i.e clusters can be imported from one data-set and highlighted onto the network graph of another. Although absolute values of expression cannot be compared, it is possible to identify overlap in transcriptional activity and patterns of expression. 524 transcripts (matched by their unique gene-probe ID's) were found to be common between the IFN-β and IFN-γ networks. DAVID analysis of GO annotations revealed the most over-represented biological process terms for the overlapping transcripts were 'immune response', 'cell cycle' and 'intracellular signalling cascade'.

In order to better visualise areas of overlap between the type-I and type-II response, the transcripts common to both networks were highlighted on the IFN-γ network according to their membership in IFN-β clusters (Figure 4.9). Conversely the common-transcripts were also highlighted on the IFN-β network according to their membership in IFN-γ clusters. 71 transcripts from the IFN-β cluster-01 (Ifnb_transient_Down_4_C1) were represented exclusively in the down-component of the IFN-γ network. Over half (201) of the transcripts within IFN-β cluster-02 were present in (only) the up-response of the IFN-γ network and 69% of these were represented in the mid phase of the response. In contrast IFN-β_cluster-03 transcripts were predominantly present during the mid-to-late stage of the IFN-γ-Up response. Up to half of IFN-γ cluster-05 transcripts were common with transcripts from IFN-β cluster-05, both of which according to their GO annotations suggest repression of genes associated with the cell division and the mitotic phase of cell cycle. Four of the six transcripts (Axud1, Nfil3, Nfkbiz, Socs3) belonging to immediate early IFN-β-cluster-30 were also present in IFN-γ immediate early response. IFN-γ clusters with very little overlap of IFN-β cluster-transcripts included cluster-06, cluster-08 and cluster-16, all representing up-regulated transcripts. The DAVID based annotations for these three clusters suggest their transcripts are associated with positive regulation of transcription (cluster-16),

negative regulation of apoptosis (cluster-08), and antigen processing and presentation (cluster-06).

Figure 4.9: IFN-β transcript representation in the IFN-γ transcriptional network. Clusters of transcripts co-ordinately expressed in response to IFN-β stimulation of mouse BMDMs were imported from the IFN-β response network and highlighted on the IFN-γ network to reveal areas of overlap between the type-I and type-II response. 524 transcripts were common between the two data sets and these have been enlarged, elongated and coloured according to their IFN-β cluster membership. Transcripts that are not shared with the IFN-β network are shown as compressed, dark blue spheres. (See text for a description of the cluster overlap).

154

110 transcripts from IFN-γ cluster-01 and 122 transcripts from IFN-γ cluster 02 were present within the IFN-β network, and over 90% of these were within the mid-phase of the Up-response (see Figure 4.10 for illustration of cluster overlap). In total 57 members of IFN-γ cluster-03 appeared within the IFN-β network, the majority were within the IFN-β early or mid-phase Up response. 50 transcripts belonging to IFN-γ cluster-04 were scattered across the different phases of the IFN-β down response. 47 transcripts from IFN-γ cluster-05 were represented within the IFN-β network and 70% of these were within IFN-β cluster-05. Whereas IFN-y cluster-05 represents down regulation around the mid-phase, the IFN-β cluster-05 bears transcripts repressed during late stage of the IFN-β response. Other interesting areas of overlap include eight transcripts from the IFN-γ early response cluster-11 also present across the Up-early phase of the IFN-β response. IFN-β clusters with little overlap with IFN-y network transcripts were the down-response clusters; cluster-08, cluster-10 and cluster 15. DAVID based GO annotations for these clusters suggest transcripts belonging to either cluster-08 or cluster-10 are associated with regulation of transcription and zinc ion binding.

Figure 4.10: IFN-γ transcript representation in the IFN-β transcriptional network. Clusters of transcripts co-ordinately expressed in response to IFN-β stimulation of mouse BMDMs were imported from the IFN-γ response network and highlighted on the IFN-β network to reveal areas of overlap between the type-I and type-II response. 524 transcripts were common between the two data sets and these have been enlarged, elongated and coloured according to their IFN-γ cluster membership. Transcripts that are not shared with the IFN-γ network are shown as compressed, dark blue spheres. (See text for a description of the cluster overlap).

Discussion

The investigations of this Chapter set out to explore the pattern of transcriptional events in mouse BMDMs in response to type-I and type-II interferons. Specifically the objectives were twofold: (1) to gain a better understanding of the transcriptional events over a 24h time-course in response to IFN-β and IFN-γ respectively, and (2) to explore the application of network-based visualisation and analysis for the interpretation of the expression data generated.

The Affymetrix Exon Arrays are designed for two complementary levels of analysis; gene expression and alternative splicing. Multiple probes (approximately four) are designed to target each exon and these probes can also be summarised into an expression value of all transcripts from the same gene. A relatively new technology at the time, the arrays had many (17,000) probes un-annotated at the exon level. Furthermore the software support for analysing splicing events was still in its infancy. Exon-level analysis was explored for the data sets generated, but ultimately 'gene-level' analysis was used as few if any convincing alternative-splicing events could be detected over the time-course of investigation. Thus having considered these caveats and given the original interests and objectives were not to identify alternative splicing events, a 'gene-level' analysis was more apt for this purpose. The raw expression data was normalised and statistical filtering applied beforehand to remove those transcripts with little change in expression across the arrays. A network based explorative approach was performed to visualise and analyse the expression data.

Analysis of the individual data sets was performed first and overlap in the transcriptional response to the two cytokines was assessed. Attempts were also made to concatenate the raw data sets to analyse them together. In theory this was a feasible option since the two experiments were set up in similar formats; identical time-points, cell lineage, similar stimuli, and identical microarray platform. However in practice the data sets could not be analysed as one. This could be for several reasons technical reasons; the experiments were processed at different times, as were RNA

preparation, RNA labelling and microarray processing each factor potentially giving rise to systematic differences between the two data sets. Other biological reasons may have also influenced the differences in the two data-sets, for example the CSF-1 state of the cells. In the IFN-γ experiment cells were differentiated in L929-conditioned medium, whereas in the IFN-β experiment the cells were differentiated directly with recombinant CSF-1. CSF-1 was present throughout the time-course treatments in both experiments and is known to regulate a range of signalling pathways in macrophages [250]. The differences in cell culture techniques may therefore have influenced the concentration and action of CSF-1 on the cell and could be responsible for biological differences in the data sets.

Network visualisation and analysis

BioLayout *Express*[3D] was used to generate network graphs of the expression data derived from mouse BMDM stimulated with either IFN-β or IFN-γ over a time-course of 24 h. The network graphs generated from the data of both experiments (Figures 4.1 and 4.4) encompassed two main graph structures, one comprising nodes (or transcripts) whose expression was up-regulated in response to interferon stimulation and the other containing transcripts whose expression was repressed. The resulting networks were divided into clusters of co-ordinately expressed transcripts using the MCL algorithm and those representing interesting patterns of expression (i.e. not clusters representing co-ordinately regulated but not differentially expressed transcripts) were retained for further analysis. The clusters derived from the IFN-β time course data ranged from 379 nodes to 6 nodes and IFN-γ cluster membership ranged from 224 nodes to 7 nodes. Total combined IFN-β cluster membership was greater (1,711 transcripts) than total IFN-γ cluster membership (1,038 transcripts), suggesting the extent of transcriptional regulation in macrophages is greater in response to IFN-β compared to IFN-γ. Stimulation of human PBMCs with a number of cytokines over a 24-hour period also revealed the extent and scale of transcriptional changes are greater in response to type-I interferons (IFN-α, IFN-β, and IFN-ω) compared to the type-II interferon IFN-γ [246].

Methods to automatically organise networks (otherwise known as layout algorithms) can enable interesting relationships and structure within data to be seen more easily. Automated layouts are rarely perfect (discussed Chapter-2) and most are easier to interpret after subsequent manual node rearrangement. To better visualise the structure of the IFN-β an IFN-γ response networks, cluster relationships and cluster membership information were extracted from the 3D-network graphs (Figures 4.1 and 4.4) and with the aid of an automated layout algorithm in yED Graph editor [251] a hierarchical layout of the information was generated. With additional manual arrangement, clusters were arranged to flow according to the temporality of the transcriptional changes they represent, whereby early changes in transcription were placed at the left of the graph and later changes to the right. Arguably the hierarchical arrangement of the networks (Figures 4.2 and 4.5) facilitated the visualisation of the structure of the interferon response. The networks clearly illustrate, the majority of transcriptional changes take place during the mid-phase following IFN-β stimulation and across the mid and mid-to-late phase following IFN-γ stimulation. Genes of interest can be imported and highlighted onto to the (hierarchical) networks and their behaviour or annotation based on previous analyses can be contrasted to what is depicted in the existing network. This functionality was utilized to identify the overlap between the two networks and gain an appreciation how transcripts within IFN-β clusters behaved in the context of the IFN-γ network (and vice versa). Networks are well-defined mathematical objects and computational analysis of patterns within the network such as the use of the MCL algorithm here enables an automated and unbiased approach to data analysis and hypothesis generation. Nevertheless manual interpretation of the results is always necessary to ensure biological relevance.

The IFN-β and IFN-γ transcriptional response in mouse BMDMs

The networks generated (Figures 4.2 and 4.5) and functional annotation of clusters reveals both IFN-β and IFN-γ induce a complex transcriptional response, consisting of multiple gene sets encoding functional programmes controlling a number of different

processes. The structure and content of the type-I and type-II responses observed in this analysis concur with the previous studies [58, 238, 246]. IFN-β stimulation of macrophages resulted in the repression of a large number of genes associated with cell cycle and DNA replication around 4-8 h post treatment, including checkpoint-associated genes; *Chek1, Chek2*. Another cluster of genes associated with cell cycle were repressed during the late phase and included *Bub1, Bub1b, Mcm5, Mcm6*. The most changing transcripts, in terms of the up-regulated response, were present in the major mid and mid-to-late phase of the IFN-β network and were indicative of main immune-related signature. Typical interferon response genes, *Ifit1, Ifit2* were present within these clusters, as well as cytokines (*Il10, Il15, Tnf, Fas*), and apoptosis related machinery (*Daxx, Casp7, Casp12*). The largest proportion of changes in response to IFN-β treatment occurred during the mid-phase (4 h > 8 h), whereas in the IFN-γ response, the main body of changes were spread across the mid and mid-to-late phase (8 h > 4 h). This mid and mid-to late-phase of the IFN-γ up-regulated response was indicative of the activation macrophage anti-microbial signalling pathways and comprised; a number of genes encoding Toll-like-receptors (Tlr2, Tlr3, Tlr6), STAT proteins (Stat1, Stat2, Stat3, Stat5a and Stat5b), interferon regulatory factors (Irf5, and Irf8) as well as apoptosis related signalling (Tnf, Daxx, Fas, Traf1, Traf2). Ifng-cluster-01 (of the mid-to-late phase) also comprised the gene encoding the MHC Class II transactivator; Ciita (the pathway depicting its activity is described in detail in Chapter-3; Figure 3.7). As expected the late response to IFN-γ comprised the transcriptional activation of genes associated with MHC Class II antigen presentation (Ifng-cluster-6). The down regulated response induced by IFN-γ included the repression of cell cycle and DNA replication processes (Cdc6, Ccne2, Pole), as well as metabolic processes (clusters-7, 12 and 13).

In comparison to the main body of the interferon response the role and significance of the genes regulated during the immediate early response is less well understood. The IFN-β immediate early response consists of 16 nodes of which five are currently un-annotated and 29 transcripts make up the IFN-γ immediate early response (Figure 4.11). In both IFN-β and IFN-γ networks the immediate early response comprises only

transcripts whose expression was induced in response to interferon stimulation. *Socs3, Nfil3, Nfkbiz* and *Axud1* were present in the immediate early response to both cytokines. The role of the Socs (or <u>s</u>uppressor <u>o</u>f <u>c</u>ytokine <u>s</u>ignalling) proteins is well characterised as part of a classical negative feedback loop induced to regulate the extent and duration of the response to cytokines [252-253]. The role of the other shared immediate early transcripts in the interferon response is less well understood. One study suggests the protein encoded by the *Nfkbiz* gene may positively regulate IFN-γ production in KG-1 cells (a human myeloid cell line)[254]. Other immediate early transcripts common between the two networks (although not within immediate early clusters in the corresponding study) included *Trib3, Arid5b, Lrrc39, Ccl3* and *Ccl4*. Unique to the IFN-β network is *Prdm1*, which encodes a protein that is reported to act as a repressor of *Ifnb1* gene expression [255] thereby forming another level of feedback regulation. A number of IFN-γ induced transcripts, including *Dusp1, Gadd45G, Errfil* and *Ier3* have been implicated in some way in the cellular stress response [256-259]. Based on the limited annotation and studies related to these immediate early transcripts it is plausible their expression is induced in response to the stress of being exposed to the cytokines and well processes for regulating the extent of the activation by the cytokines e.g. *Socs3* and *Prdm1*.
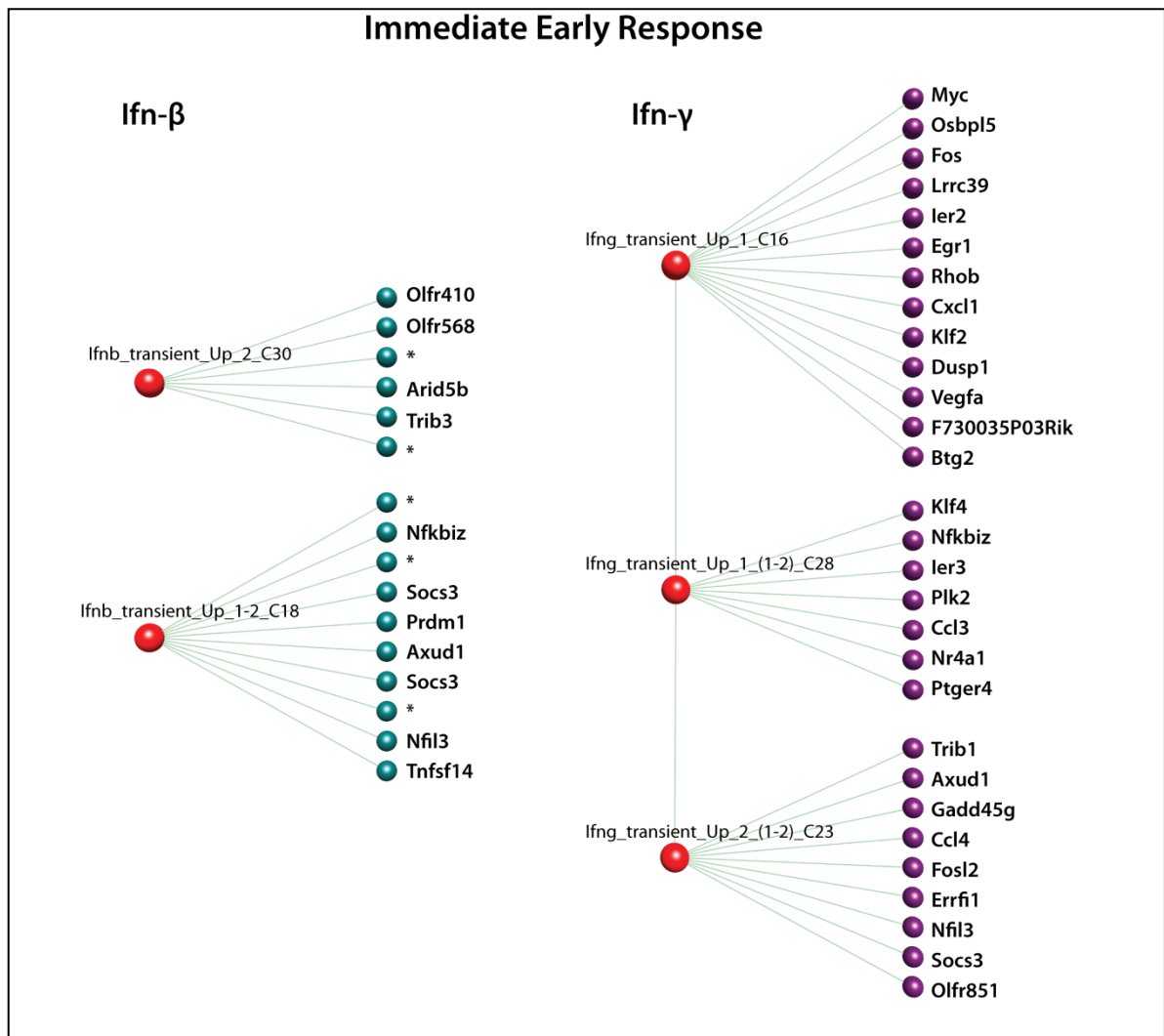
Figure 4.11: The IFN-β and IFN-γ immediate early transcriptional response. Transcripts belonging to the immediate early response clusters of the IFN-β and IFN-γ response. * denotes un-annotated probe sets.

The type-I and type-II transcriptional response are known to share a number of transcriptional targets and such overlap in response was observed during this analysis (Figures 4.9 and 4.10). The role of type-I and type-II interferons is generally acknowledged to be non-redundant since efficient clearing of some viruses requires both types of interferon, whereas others require only type-I or type-II interferon. Thus exploration of the overlap and variation in the transcriptional response of IFN-β and IFN-γ may provide insight into their functional role. 524 transcripts were common to both the IFN-β and IFN-γ networks (Figures 4.9 and 4.10). It is probable that there are other shared transcripts not identified on this occasion as they may not have met the original network filtering thresholds. The most common GO annotations for the overlapping transcripts were 'immune response', 'cell cycle' and 'intracellular signalling cascade'. A number of the shared ISGs are transcription factors (including *Stat1*, *Stat2*, *Stat3*, *Irf2*, and *Irf5*) and may drive the subsequent waves of transcription response following initial interferon stimulation. Network graphs highlighting the overlap between the IFN-β and IFN-γ data suggest the activity of the shared transcripts is on the whole consistent between the two responses. For example IFN-β cluster-02 and cluster-03 transcripts appear predominantly across the IFN-γ Up-response graph component around the mid- and mid-to-late phase. Although comparison of the data sets is somewhat confounded the overall response observed is still typical of the type-I and type-II response reported elsewhere [58, 238, 246]. For example the IFN-γ cluster-06 comprising transcripts associated with MHC class II antigen presentation and processing (a process regulated efficiently by type-II interferon [239-240]) shared very little overlap with the IFN-β network.

Conclusions & Further Work

The transcriptional changes observed in response to IFN-γ and IFN-β were fitting with the well known and documented activities of these cytokines, for examples; MHC Class II activation by IFN-γ, the anti-proliferative signature induced by the IFNs, as well as the classical anti-microbial response. The network based analysis of the gene-expression data illustrated clear structure in the transcriptional response generated by the two cytokines. Moreover the analyses also revealed a complex network of co-ordinately

expressed transcripts, many of aspects of which are currently poorly characterised and understood (e.g. the immediate early response genes). Future efforts may seek to unravel the significance of these genes in the type-I and type-II IFN responses.

The comparison of the IFN-β and IFN-γ response in mouse BMDM was limited by experimental differences in the generation and processing of samples for the two micro-array studies. For a more rigorous comparison the two time-course studies would be conducted on the same days on samples derived from the same culture of cells. Processing of the samples for hybridisation to microarrays (from RNA extraction to labelling) would also be conducted simultaneously as well as the all aspects of the microarray processing (for example use of scanners).

To further explore and validate the transcriptional responses observed, it would be valuable to follow up these findings using appropriate techniques. For example the IFN-β network analysis revealed a large response associated with the repression of cell-cycle related transcripts, suggesting cell division is being halted. FACS analysis could corroborate this suggestion as well as provide insight into the exact cell cycle phase the cells are in.

Methods and Materials

Generation of type I and type II interferon response data sets

During the course of these investigations I had two separate opportunities to run 12 Affymetrix Mouse Exon arrays as part of the [Wellcome Trust Advanced Course](#) (WTAC) series in Functional Genomics and Systems Biology, where I was participating as an assistant. Given the interest in better understanding the transcriptional response of macrophages to IFN, a time-course of IFN-γ stimulation of mouse BMDM was generated during the 2007 WTAC. Twelve microarrays were used to analyse 6 time-points in duplicate (0, 1, 2, 4, 8 and 24 hours); the final MIAME compliant dataset is available in Array Express (record [E-MEXP-1490](#)). During the 2009 WTAC the 12 microarrays were utilized to examine transcriptional events following IFN-β stimulation of mouse BMDMs at the same time points as the IFN-γ study and is available in the Gene Expression Omnibus (GEO accession: [GSE20403](#)).

During the interim period between the two courses I moved from the Division of Pathway Medicine (UoE) to the Roslin Institute (UoE) where a slightly different macrophage cell culture technique was adopted (see Methods) to that used in the previous lab. Thus the two time-course experiments were performed on different years and in different laboratories with slightly different cell culturing techniques. However, even if this compromised cross-comparison of experiments, the individual data sets generated were of high quality.

IFN-γ Study

Cell culture and treatment

Primary mouse bone marrow derived monocytes were prepared from male BALB/c mice 10–12 weeks old. Cells were washed, resuspended in DMEM-F12/10% FCS/L929 medium and counted before being plated in a 24-well plate at a concentration of 5 ×

10^5 cells/well. To differentiate the cells from monocytes into primary macrophages, cells were then incubated for 7 days in DMEM-F12 growth media supplemented with 10% L929 cell suspension releasing the MCP-1 macrophage stimulating factor, with media changes on days 3 and 5. On day 7 the growth medium was replaced with DMEM-12/10%FCS medium containing 10 U/ml recombinant mouse interferon gamma (Pierce-Thermofisher Scientific, Rockford US) and harvested 1, 2, 4 & 8 h following treatment or collected pre-treatment (0 h).

RNA extraction, QC and labelling for arrays

Total RNA was harvested from the cells using an RNeasy Plus kit (Qiagen) according to manufacturer's instructions. RNA was quantified and quality controlled using a NanoDrop spectrophotometer (Nano-Drop Technologies) and BioAnalyser 2100 (Agilent). Replicate 150 ng samples of total RNA derived from two separate wells per time point were labelled using the Affymetrix whole transcript labelling protocol and hybridized for 16 h at 45°C to Affymetrix mouse exon 1.0 ST arrays. They were then washed and scanned according to manufacturer's recommendations.

Data processing and network analysis

Data (ArrayExpress Ac. No: E-MEXP-1490) was normalized using the RMA package within the Affymetrix Expression Console software and annotated. Transcripts which might be considered to be differentially expressed were identified using either the Empirical Bayes function within Bioconductor (http://www.bioconductor.org/) or using the annova function within GeneSpring (Agilent Technologies, Stockport, Cheshire) with a 1.6 fold cut-off. In total 1,678 transcripts were identified by one or both of these filters. The data corresponding to this list was then loaded into the network visualization tool BioLayout *Express*[23] using a Pearson correlation cut-off of 0.9 to filter edges. The resultant network graph (Figure 4.1) of 1,474 nodes was clustered using the graph-based clustering algorithm MCL [137] set at an inflation value of 2.2 resulting in 40 clusters with a membership of 6 or more nodes (transcripts).

IFN-β Study

Cell culture and treatment

Bone marrow derived macrophages (BMDM) were prepared from femurs of 8-9 week old male BALB/c mice. RPMI 1640 medium (Sigma-Aldrich, Gillingham, UK) supplemented with 10% heat inactivated foetal bovine serum (FBS) (Sigma-Aldrich Gillingham, UK), 25 U/ml penicillin (Invitrogen, Paisley, UK), 25 µg/ml streptomycin (Invitrogen, Paisley, UK), and 2 mM L-glutamine (Invitrogen) (complete medium) was used for culture of the BMDM. Briefly, bone marrow cells were cultured for 6 days in complete medium in the presence of 10,000 U/ml CSF1 (a gift from Professor David Hume) on 10 cm square bacteriological plastic plates, with a re-supplement of CSF1 on day 5. On day 6 cells were harvested, counted, re-suspended in complete medium with 10,000 U/ml CSF1 and plated out onto 6-well tissue culture plates at a density of 1 x 10^6 cells per well and cultured for a further 24 h. Cells were maintained in a 37°C incubator containing 5% CO2. On day 7 cells were treated with 10 U/ml recombinant mouse interferon-beta (IFN-β) (PBL Interferon Source, New Jersey, USA) and harvested 1, 2, 4, 8 and 24 h following treatment or collected pre-treatment (0 h). RNA extraction, QC and labelling for arrays was performed as for the IFN-γ study with the exception that 300 ng of RNA was used instead of 150 ng.

Data processing and network analysis

Data (GEO accession no GSE20403) was normalized using the RMA package within the Affymetrix Expression Console software and annotated. Transcripts which might be considered to be differentially expressed at each time point compared to 0 h were identified using the Empirical Bayes function within the Bioconductor (http://www.bioconductor.org/) package of the R statistical programme using a 1.5 fold cut-off and a p-value of 0.05. According to this analysis 2,300 transcripts were considered to be differentially expressed. Data corresponding to these transcripts was then loaded into the network visualization tool BioLayout Express3D [138, 206] using a Pearson correlation cut-off of 0.9 to filter edges. The resultant network graph (Figure 4.4) of 2,045 nodes (connected by 92,947 edges) was clustered using the graph-based

clustering algorithm MCL [137] set at an inflation value of 2.2 resulting in 33 clusters (with greater than 5 nodes). Manual inspection of the clusters revealed 21 clusters of interest i.e. their expression profile was consistent with the genes being regulated by IFN-β. Resulting clusters represent patterns of co-expression amongst transcripts in the network graph. The clustered data was then exported as "class-sets" for overlaying onto other datasets (including the IFN-γ response network) as well as the macrophage pathway diagram.

Chapter Contributions and Acknowledgements

Both the IFN-γ and IFN-β studies were designed under the supervision of Prof. Tom Freeman. I participated in the cell culture, treatment and RNA extraction and QC for the IFN-γ study alongside Dr. David Page (DPM). I performed the cell culture, treatment, RNA extraction and QC for IFN-β time-course study. Processing of RNA for hybridisation to the Affymetrix microarrays was performed by participants of the Wellcome Trust Advance Course (2007 and 2009) and the entire laboratory process was overseen and supervised by myself. Statistical filtering of both datasets was supervised by Dr. Anton Enright (European Bioinformatics Institute) and Dr. Cei Abreu-Goodger (European Bioinformatics Institute). The network analysis described within this Chapter has been performed by me under the supervision of Prof. Tom Freeman. Thanks are due to both Tom and Prof. David Hume for their useful comments on this Chapter.

This page has been left intentionally blank

Chapter 5. siRNA Targeting of Key Regulators of the Interferon Pathway

Introduction

Chapters 2 & 3 described the assembly of an *in silico* model of receptor-initiated signalling pathways in the macrophage, including type-I and type-II interferon signalling cascades. The analyses included in Chapter-4 attempted to characterise the transcriptional response of BMDMs to IFN-β and IFN-γ using a network based approach to analyse microarray data. Leading on from these studies (in Chapters-2 to 4) and other investigations within the group [146] was the desire to study the contribution of individual pathway components to the interferon response. Based on previous investigations we found the siRNA induced knockdown of certain genes, perturbed transcriptional networks associated with type-I interferon signalling [146] and using the pathway model it was possible to hypothesise a rational explanation for these observations. The investigations described in this Chapter-(5) aimed to reproduce our previous findings [146] and extend these investigations to other genes known or hypothesised to modulate type-I interferon signalling. In order to begin to perform such an analysis it was first necessary to optimise an *in vitro* assay for using siRNA-based screening in mouse primary BMDMs. Therefore the major focus of the work described here was to define and optimise an *in vitro* screening assay to study the effect of gene knockdowns on the interferon response. The secondary focus was to then use (and test) the optimised screening protocol to the study the role of genes of interest in type-I interferon signalling.

Modulation of type-I interferon signalling by targeted siRNA induced gene knockdown - Prior work

Previous data generated by our group suggested targeted knockdown of certain genes by siRNA in mouse BMDMs perturbs siRNA-lipofection induced type-I interferon signalling [146]. Specifically the siRNA's resulting in this marked shift in the macrophage transcriptome, targeted the Ifnb1, Irf3, Irf5, Stat1, Stat2 and Nfkb2 transcripts. Transfection of other siRNA's (in this case targeting the Casp4, Ifi47, Lyn, Sod2, Traf1 transcripts), resulted in the activation of transcriptional networks typically

associated with interferon signalling to a similar degree to control non-targeting siRNA (Chapter-1; Figure 1.7). A simplified schematic describing the expected outcome (in terms of ISG expression) to treatment with non-targeting siRNA or with siRNA targeting a gene which acts as a positive regulator of type-I signalling is shown in Figure 5.1.

Certain macrophage PRRs (TLR3, TLR7, TRL8, RIG-like-receptors) are thought to be sensitive to synthetic siRNA resulting in activation of the immune system in mammalian cells [260-263]. Engagement of these receptors with viral RNA or in this case the synthetic siRNA leads to the eventual assembly and activation of the IFN-β enhancesome. Four (Irf3, Ifnb1, Stat1, Stat2) of the six gene targets which perturbed transcriptional networks associated with type-I IFN signalling [146] are known to act within the same (IFN-β) pathway [189, 264-265].

The role of the Irf5 and Nfkb2 transcription factors in the IFN response is less well characterised and/or established, particularly at the pathway level. There is evidence to implicate these proteins in the immune and IFN response [266-271]. The most compelling of this evidence suggests high IRF5 expression (in mice and humans) is characteristic of the M1 phenotype of macrophages in which it directly activates the transcription of certain genes encoding inflammatory mediators [272].
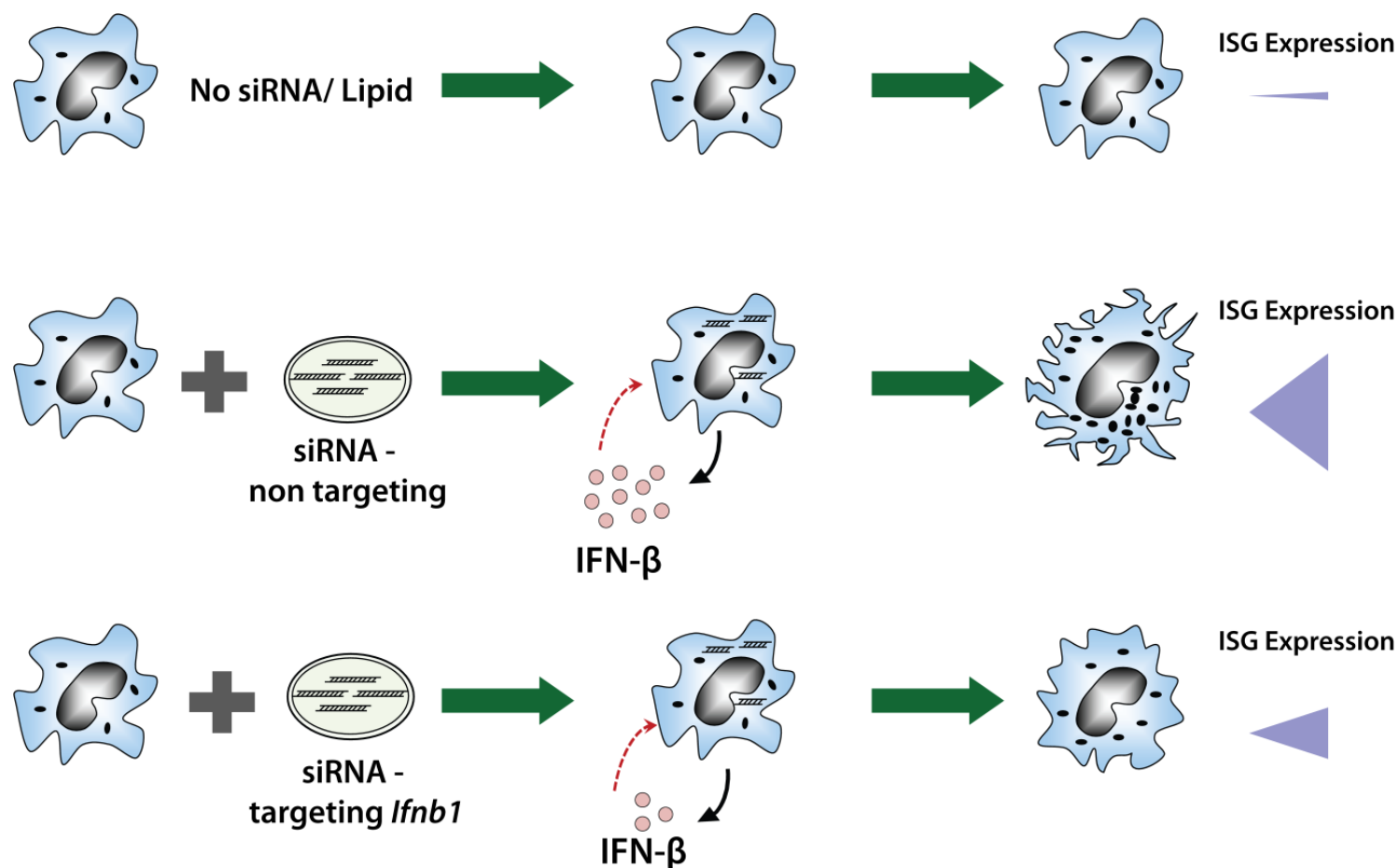
Figure 5.1: Predicted response of primary mouse macrophages to siRNA tranfection using a lipid delivery vector. Macrophages not exposed to siRNA/lipid are expected to express low/ basal levels of interferon stimulated genes. Transfection of a siRNA-lipid complex is expected to induce autocrine IFN-β production in the macrophage, as the siRNA/ lipid are detected by pattern recognition receptors. If the siRNA transfected into the macrophage happens to target a key component of IFN-β/ type-I signalling (such as the *Ifnb1* gene itself), it would be expected that level of autocrine IFN-β production and subsequent ISG expression would be lower relative to cells transfected with control-Non-Targeting siRNA or siRNA targeting genes not critical in the propagation of the interferon response.

Further studies are required to understand how and at what level the Irf5 and Nfkb2 proteins are acting in the (type-I) interferon pathway. Other members of the group are currently investigating the role of Nfkb2 in macrophage signalling to follow up the observed effects of siRNA/ type-I response study [146] and to gain a better understanding of its action on macrophage signalling. Preceding these investigations was the desire to develop methods of studying the effects of siRNA mediated knockdown of the Nfkb2 gene, as well as other genes, with the objectives of:

(i)   Investigating whether it was possible to replicate the observations of the previous siRNA screen [146] (under new laboratory settings).

(ii)   Exploring whether other genes (not considered during the original study [146]) might also be implicated in mediating the type-I response in mouse macrophages.

(iii)   Optimising an experimental assay for screening genes by siRNA knockdown and thereby assessing their contribution to the type-I response or other macrophage activation pathways.

Requirement for a new *in vitro* assay to study the role of other GOI in the type-I interferon response and a preliminary assay design

The experimental setup of the previous study was not specifically designed to assay the effects of the gene knockdowns in the siRNA-induced type-I response. In fact the original study was designed to understand the role of the genes of interest in the context of the type-II interferon response (see Chapter-1 and [146]). The discovery that siRNA's targeting the six genes (described above) perturbed transcriptional networks associated with type-I interferon signalling was to some extent serendipitous. For the investigations in this Chapter a new *in vitro* assay was designed to test the role of genes in the type-I response using siRNA. One of the major reasons for the design of a new assay, as opposed to keeping entirely with the methodology of the previous screen was that the system of studying type-I response perturbation using a method (siRNA lipofection) that induces the signalling of interest may complicate the analysis of the role of the GOI. For example, it is possible the act of

siRNA transfection may induce *Ifnb1* expression before the siRNA has had sufficient time to knockdown the GOI, therefore masking any influence the GOI may have in *Ifnb1* regulation. Alternatively by allowing sufficient time for gene KD, then using another (exogenous) stimulator of *Ifnb1* expression in macrophages it may be possible to better attribute the contribution of the GOI to *Ifnb1* production and/or subsequent type-I response. An exogenous stimulator of IFN-β signalling in macrophages is Lipopolysaccharide (LPS). LPS is the major component of the outer membrane of Gram-negative bacteria. It causes monocytes, macrophages and neutrophils to up-regulate phagocytic functions and to release inflammatory mediators, including interleukins (IL-1, IL-6, IL-12), TNF, nitric oxide (NO) and interferons [273-274]. In macrophages, LPS stimulation induces the rapid transcription of IFN-β mRNA and protein secretion [275]. In turn IFN-β signalling forms a key portion of the LPS transcriptional response and LPS induced lethality [276-277]. Studying the effect of GOI knockdown in the context of the LPS response could be valuable not only because LPS serves as a stimulator of *Ifnb1* expression, but also to further explore the contribution of IFN-β signalling in the LPS response. Thus another motivation to develop a novel *in vitro* assay was the interest in studying the role of the GOIs in the context of the LPS (induced-type-I) response.

Figure 5.2 describes how LPS stimulation of macrophages could be incorporated to study the role of the GOI in the type-I interferon signalling response. The basic design of the assay was firstly to knockdown the genes using siRNA and measure any consequential change in *Ifnb1* and type-I gene expression, and secondly challenge the cells with LPS to determine variability in *Ifnb1* and type-I gene expression induction. Given this analysis was based on the desire to replicate and expand on observations in the previous siRNA screen [146] some consistency between the previous screen were maintained, including the supply of siRNA used (Dharmacon (Thermo Scientific) ON-TARGET *plus* SMARTpool), and the choice of transfection method and reagent (Lipofectamine 2000). Previous data from the group suggested BMDMs are sensitive to Lipofectamine, but less so than Lipofectamine in combination with siRNA [146]. Others have optimised siRNA delivery by electroporation in primary BMDMs and suggested

this method does not affect macrophage function and phenotype [278]. However the study of lipid based siRNA delivery to macrophages warrants investigation in itself since it is argued to be the most tractable form of siRNA delivery in *in vivo* settings where electroporation is not viable. Furthermore synthetic non-viral, lipid-based vectors are considered the most feasible method for delivering DNA/RNA therapeutic agents since they are regarded to be safer than their viral counterparts and improvements in their chemistry and formulation are beginning to position them as viable alternatives to viral vectors [279].

Knock-down Genes of Interest (GOI) by RNA interference

Stimulate macrophages with LPS

Measure *Ifnb1* expression and type I response gene expression

↓ *Ifnb1* and/or type I gene expression in GOI KD samples **<** controls

≈ *Ifnb1* and/or type I gene expression in GOI KD samples **≈** controls

↑ *Ifnb1* and/or type I gene expression in GOI KD samples **>** controls

GOI potential +ve regulator of IFN-β expression/ response pathways

GOI potential -ve regulator of IFN-β expression/ response pathways

GOI potentially not involved in regulation of IFN-β expression/ response pathways

Figure 5.2: Preliminary work flow for investigating the role of genes of interest (GOI) in the IFN-β signalling pathway. GOI knocked down in mouse bone marrow derived macrophages using siRNA, and then challenged with LPS. Subsequent *Ifnb1* and type-I gene expression was measured, and some of the possible outcomes are shown. If a GOI knockdown results in either a reduced or augmented type-I response and *Ifnb1* expression compared to controls, it could be plausible the GOI is either a positive or negative regulator respectively, of type-I response signalling. Other outcomes (not shown) are possible; a GOI knockdown may repress *Ifnb1* expression but, enhance type-I expression or contrariwise.

Questions to be addressed in the design of an assay to study the role of genes of interest in the type I response in BMDMs

Despite the advances in siRNA design to improve potency and reduce off-target effects, siRNA effectiveness can still vary across different cell types [280]. Furthermore conditions for optimal siRNA efficacy may need to be altered depending experimental design factors, for example the scale of throughput (screening in 6/24/96/384 well plates), method of siRNA delivery, other treatments the cells are exposed to, and the cellular environment [281]. Ultimately each experimental scenario requires careful optimisation to find a balance between siRNA induced knockdown potency and off-target effects. Some of the specific questions that were to be addressed during the assay optimisation included:

- What is the 'ideal' siRNA concentration to use for screens in mouse BMDMs?

To explore this point the following parameters were investigated:

o Uptake of increasing concentrations of siRNA in mouse BMDMs

o Knockdown efficiency of increasing concentrations of siRNA targeting a gene of interest

o Assessment of type-I response gene induction with increasing concentrations of siRNA

o Changes in cell morphology with increasing concentrations of siRNA combined with LPS treatment.

- What concentration of LPS should be used in the in vitro assay?

To explore this point the following parameters were investigated:

o Comparison of type-I response induced by LPS, IFN-β and NT-siRNA

o Analysis of macrophage response to combined siRNA and LPS treatment

o Analysis of macrophage response to 5 ng/ml LPS and 20 nM siRNA (concentrations which were chosen for the in vitro assay). The individual response and combined response to these stimuli was studied.

- What is the ideal siRNA-treatment time for obtaining optimal gene knockdown and type-I repression following LPS stimulation?

To explore this point the following parameters were investigated:

o Analysis of protein level knockdown over time

o Analysis of gene-level knockdown and efficacy of downstream response 24 h or 48 h post-siRNA transfection

In addressing the above points the optimal time for measuring *Ifnb1* expression in response to siRNA transfection as well as following LPS stimulation in mouse BMDMs was determined. Finally using the parameters established in the optimisation process (siRNA concentration, LPS dose, time of siRNA treatment), an *in vitro* assay for studying the role of GOI in the type-I and LPS response using siRNA was finalised. To test the assay as well as the potential role of the GOI in LPS signalling, a siRNA screen targeted a number of genes was performed and type-I signalling levels assayed by QPCR as well as genome-wide transcriptional profiling using siRNA.

Results

Optimisation of an *in vitro* assay for screening the role of gene knockdowns in the interferon-β response and LPS response

(1) Determination of 'ideal' siRNA concentration to use for screens in mouse BMDMs

Previous data generated within the group relied on a final concentration of 20 nM siRNA when transfecting primary mouse BMDMs [146]. This concentration was not previously optimized. I therefore examined the dose-dependence of siRNA actions in mouse BMDMs. 5, 20, 50 and 100 nM final concentrations of siRNA were transfected (using Lipofectamine 2000) into mouse BMDMs and cell morphology, siRNA transfection efficiency, target gene knockdown levels and immune activation were assessed.

(1.1)   Uptake of increasing concentrations of siRNA in mouse BMDMs

Fluorescently labelled siRNA, siGLO Green, is a RISC-independent non-targeting control intended as a qualitative indicator of delivery, especially lipid-mediated transfection. In order to gauge the extent and dynamics of lipid-based siRNA delivery, the uptake of 20 nM or 50 nM final concentration of siGLO in mouse BMDMs was monitored over time by confocal microscopy and cell images taken at 4, 7 and 24 h post transfection. This process was first performed at a cell seeding density suited for confocal imagery, (that is 50,000 cells per glass slide placed in a well of 24-well tissue culture plate) and was then repeated closer to a seeding density used in 24-well tissue culture experiment.

At 50,000 cells per slide and challenged with 50 nM final concentration siRNA, almost every cell contained the fluorescently labelled siRNA (Figure 5.3a&b). At 4 h post-transfection, cells contained granular clusters of green fluorescence, as well as staining overlapping with the nuclear dye DAPI. By 7 h this staining was more diffuse and even more so by 24 h. Furthermore by 24 h patches of staining were observed in areas not

overlapping with cells. Subsequent observations suggested this may be residual stain from cellular debris arising from cell death (Figure 5.3.a&b). When 50,000 cells were treated at a final siGLO concentration of 20 nM, every cell appeared to overlap with fluorescent staining. Although at 4 h there appeared to be fewer granules of fluorescence compared to the 50 nM siRNA treatment, by 24 h there were large and clearly visible pools of staining overlapping or surrounding the DAPI staining (Figure 5.4). Hence even with 20 nM siRNA every cell appeared to be transfected by 24 h of treatment. To determine if the extent of transfection was changed at a seeding density closer to that normally used in a well of 24 well plate, 163,000 cells per glass slide were exposed to a final concentration of 20 nM siRNA. Over the time-course of treatment the intensity of the staining increased and by 24 h the fluorescence overlapped or surrounded all nuclear staining, suggesting that even at a higher seeding density the siRNA is transfected into most if not every cell (Figure 5.5).
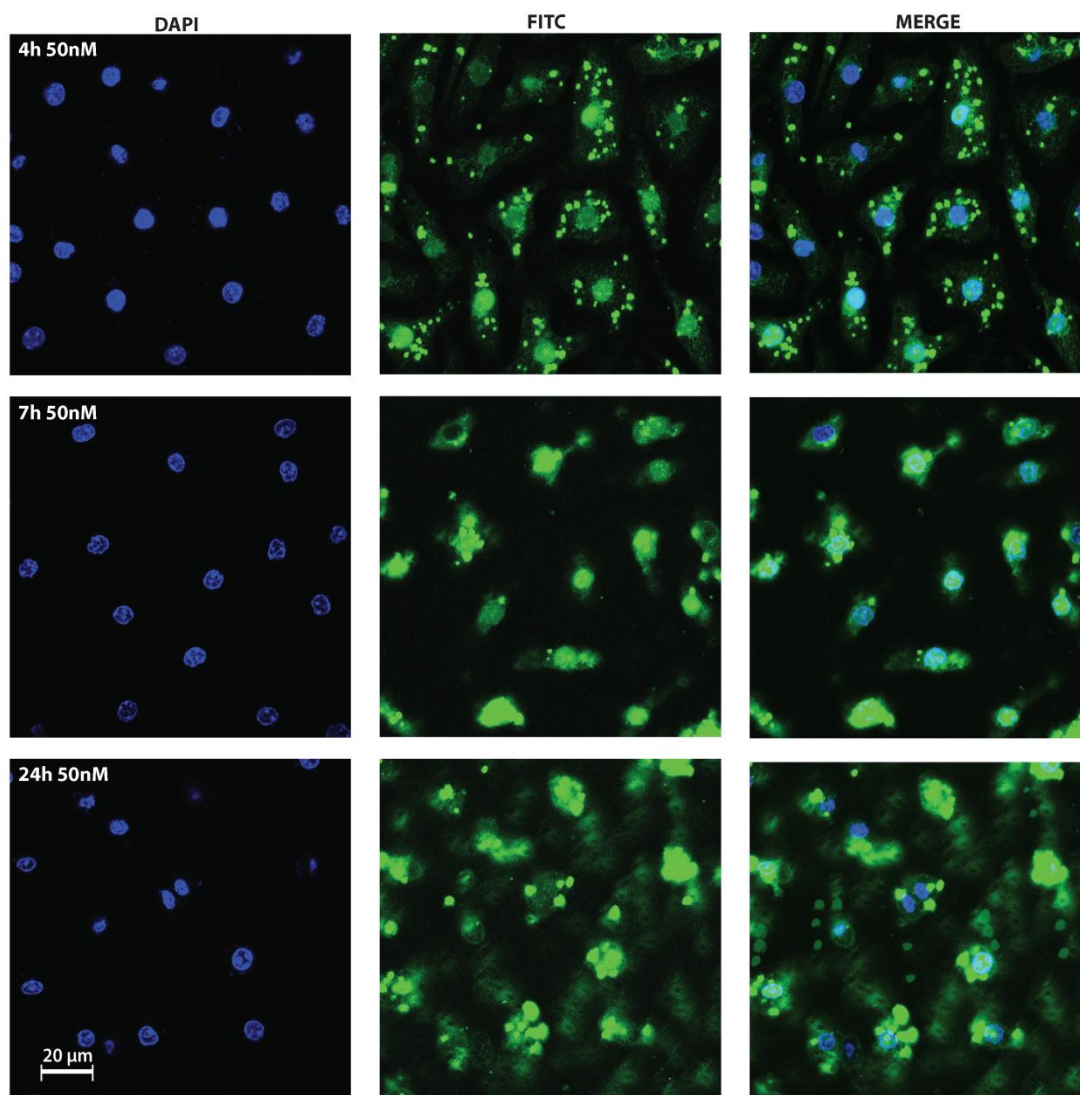
Figure 5.3. (a) Time-course of uptake of 50 nM fluorescently labelled siRNA (siGLO) in mouse bone marrow derived macrophages. Mouse BMDMs were seeded at a density of 50,000 cells per glass cover slip placed in a well of a 24-well tissue culture plate. Lipofectamine 2000 was used to transfect the BMDMs with siGLO (fluorescent siRNA) at a final concentration of 50 nM and cells were fixed at 4, 7 and 24 hours post-transfection. Nuclei were stained with DAPI. siGLO is chemically labelled with a 6-FAM fluorophore. Cell images were captured using confocal microscopy. No fluorescence was visible in the FITC channel of control (untreated) cells.
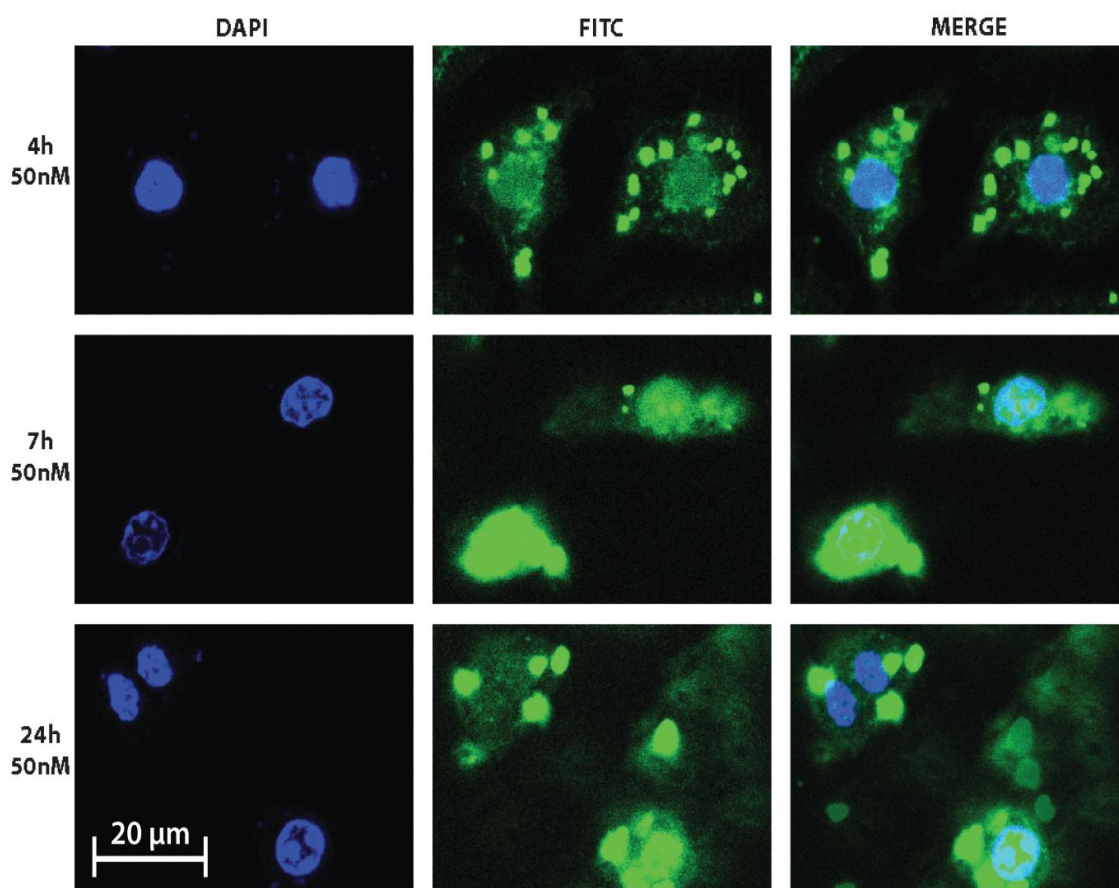
Figure 5.3. (b) Detailed cell images extracted from figure 5.3a showing uptake of 50 nM fluorescently labelled siRNA (siGLO) in mouse bone marrow derived macrophages over time. Mouse BMDMs were seeded at a density of 50,000 cells per glass slide place in a well of a 24-well tissue culture plate. Lipofectamine 2000 was used to transfect the BMDMs with siGLO (fluorescent siRNA) at a final concentration of 50 nM and cells were fixed at 4, 7 and 24 hours post transfection. Nuclei were stained with DAPI. siGLO is chemically labelled with a 6-FAM fluorophore. Cell images were captured using confocal microscopy. No fluorescence was visible in the FITC channel of control (untreated) cells.
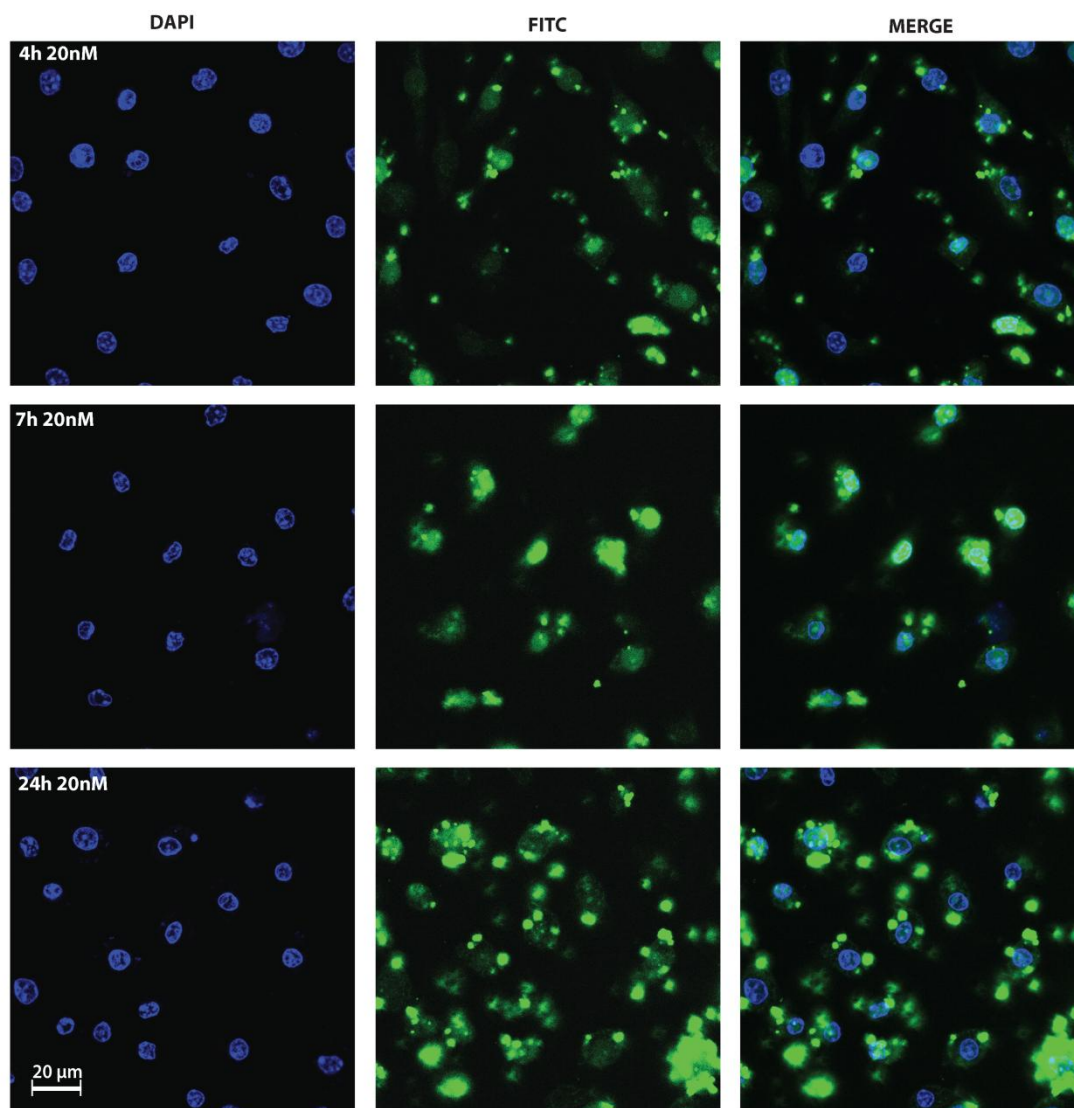
Figure 5.4: Time-course of uptake of 20 nM fluorescently labelled siRNA (siGLO) in mouse bone marrow derived macrophages. Mouse BMDMs were seeded at a density of 50,000 cells per glass slide place in a well of a 24-well tissue culture plate. Lipofectamine 2000 was used to transfect the BMDMs with siGLO (fluorescent siRNA) at a final concentration of 20 nM and cells were fixed at 4, 7 and 24 hours post transfection. Nuclei were stained with DAPI. siGLO is chemically labelled with a 6-FAM fluorophore. Cell images were captured using confocal microscopy. No fluorescence was visible in the FITC channel of control (untreated) cells.
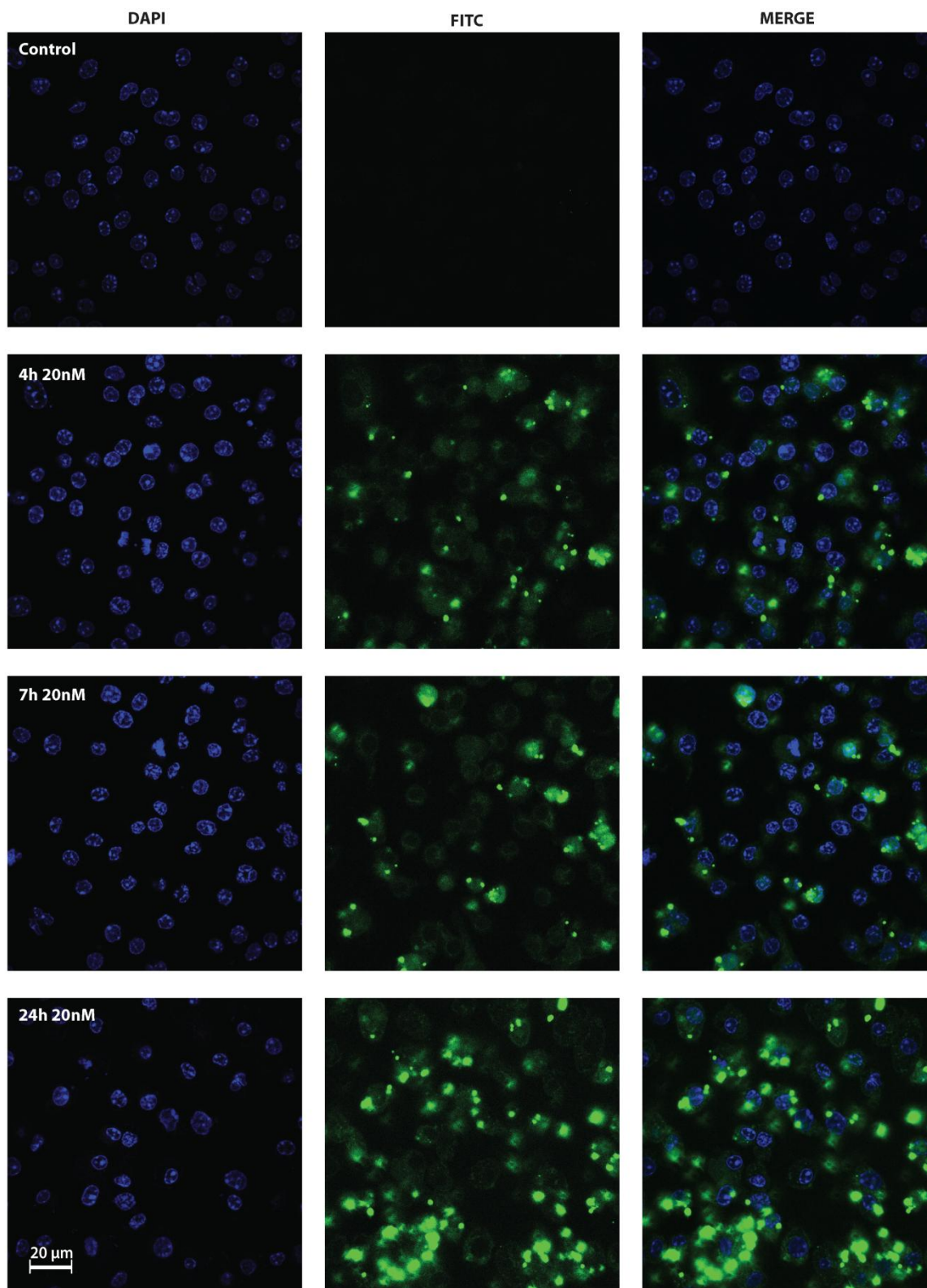
Figure 5.5: Time-course of uptake of 20 nM fluorescently labelled siRNA (siGLO) in mouse bone marrow derived macrophages. Mouse BMDMs were seeded at a density of 163,000 cells per glass slide placed in a well of a 24-well tissue culture plate. Lipofectamine 2000 was used to transfect the BMDMs with siGLO (fluorescent siRNA) at a final concentration of 20 nM and cells were fixed at 4, 7 and 24 hours post transfection. Nuclei were stained with DAPI. siGLO is chemically labelled with a 6-FAM fluorophore. Cell images were captured using confocal microscopy. No fluorescence was visible in the FITC channel of control (untreated) cells.

1.2   Knockdown efficiency of increasing concentrations of siRNA targeting gene of interest

Having established that lipofection results in an efficient uptake of siRNA into macrophages at concentrations previously used, was the desire to then determine how the extent of gene knockdown would vary with differing siRNA concentrations. It might be expected that increasing concentrations of siRNA may achieve higher levels of gene KD, nevertheless any increase in efficiency of gene or protein knockdown should be evaluated alongside other variable factors including; the increase in cost of reagents, extent of immune activation, changes in cell morphology and cell toxicity.

siRNA targeting either *Nfkb2* or *Irf7* genes was transfected into mouse BMDMs at 5, 20, 50 or 100 nM final concentration for 24 h and expression of the target genes assayed by QPCR. To control for any transfection induced changes in expression of the genes, non-targeting (NT) siRNA and a negative control designed to lack identity with known gene targets, were also transfected at the same concentrations into mouse BMDMs for 24 h. The data generated (Figure 5.6) several observations of interest; (1) *Nfkb2* expression was suppressed beyond basal levels in samples treated with siRNA targeting Nfkb2, however *Irf7* expression, even if lower than NT-siRNA treated samples, was slightly elevated compared to basal levels; (2) regardless of the siRNA concentration used the expression levels of the GOI in the target-siRNA (*Nfkb2* or *Irf7*) treated cells was similar (3) however increasing concentrations of NT-siRNA resulted in an almost exponential increase in GOI expression. Thus in terms of knockdown efficiency, comparing target gene expression at 100 nM NT siRNA treated cells vs. 100 nM siRNA targeting the GOI, would appear to give the largest percentage knockdown. However, when transfected with 100 nM NT-siRNA the target genes are induced well beyond their basal expression level and this may have consequential downstream effects on macrophage signalling.

Nfkb2 protein expression was also induced by tranfection of NT-siRNA, and to a greater extent with higher doses of siRNA (≥50 nM) (Figure 5.7), substantiating the

observations at the transcriptional level. In cells transfected with increasing concentrations of siRNA targeting the Nfkb2 gene, protein expression was supressed relative to cells transfected with the corresponding concentrations of NT-siRNA.

(a)

Irf7 Expression

(b)

Nfkb2 Expression
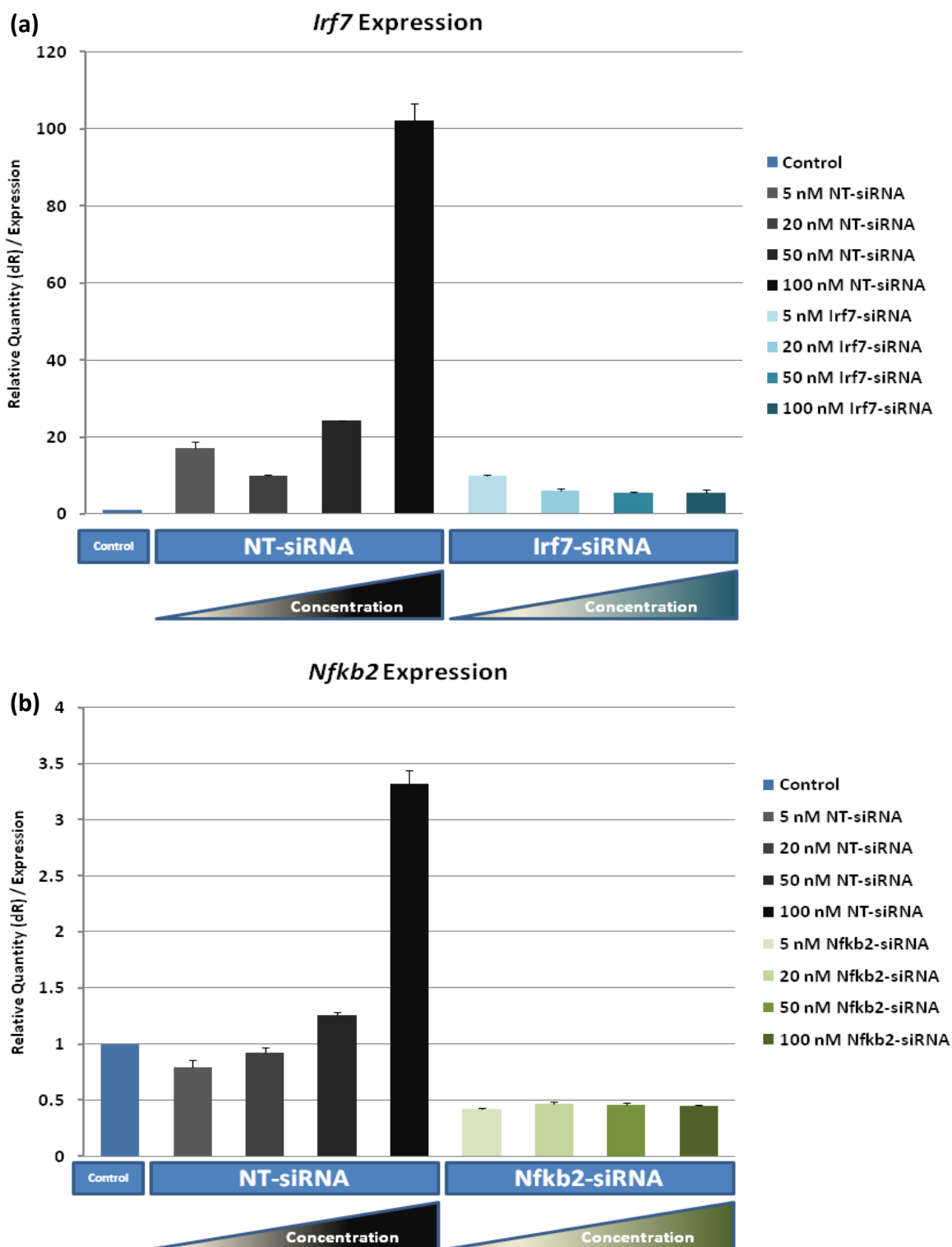
Figure 5.6: Assessment of gene knockdown efficiency 24 hours post treatment with increasing concentrations of siRNA. (a) Mouse BMDMs were transfected with four different concentrations of siRNA targeting the *Nfkb2* gene or four different concentrations of non-targeting siRNA, or cells were untreated (control). Following 24 hours of treatment *Nfkb2* expression was assayed in all samples. **(b)** Mouse BMDMs were transfected with four different concentrations of siRNA targeting the *Irf7* gene or four different concentrations of non-targeting siRNA, or cells were untreated (control). Following 24hours of treatment *Irf7* expression was assayed in all samples.
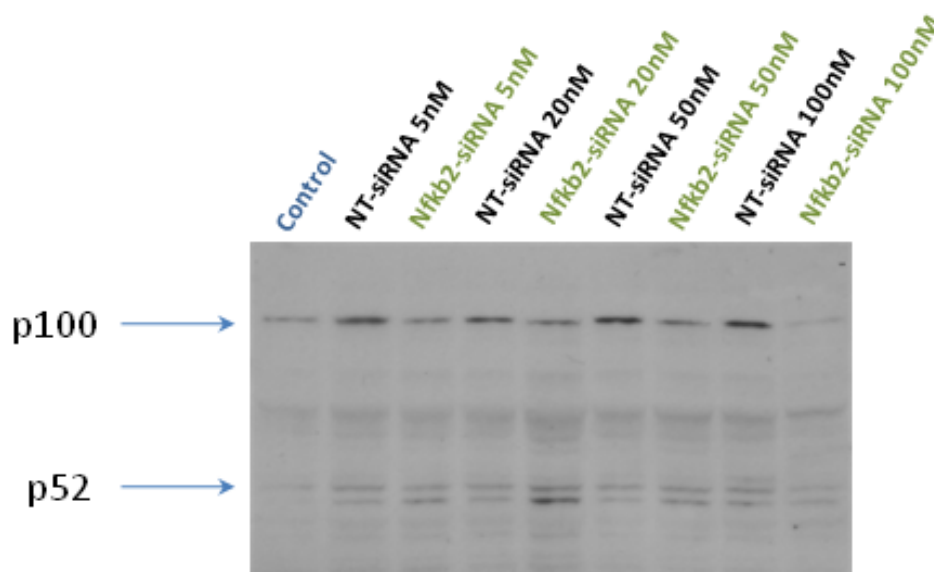
Figure 5.7: Assessment of Nfkb2 protein knockdown efficiency 48 hours post treatment with increasing concentrations of siRNA targeting the *Nfkb2* gene. Mouse BMDMs were transfected with four different concentrations of siRNA targeting the *Nfkb2* gene or four different concentrations of non-targeting siRNA, or cells were untreated (control). Following 48 hours following siRNA transfection cell were harvested and Western blot analysis performed on protein extracts from the cells using an Nfkb2 antibody targeting the p100 and p52 isoforms of the Nfkb2 protein.

1.3   Assessment of type-I response gene induction with increasing concentrations of siRNA

The gene knockdown results (Figure 5.6) suggested similar levels of knockdown are achievable with a range of siRNA concentrations, however GOI expression (Irf7, and Nfkb2) is induced by transfecting increasing concentrations of NT-siRNA. To better understand how type-I response signalling could be affected by increasing concentrations of siRNA, expression of *Oasl1* and *Ifit1* was measured in the 24 h gene knockdown samples (Figure 5.8). Measurement of these classical type-I response genes across the screen of siRNA concentrations suggests different patterns of concentration dependant expression. An exponential increase of *Oasl1* and *Ifit1* was observed in response to increasing concentrations of NT-siRNA; the response genes were induced to similar levels in samples treated with increasing concentrations of siRNA targeting the *Irf7* gene, with the exception of 100 nM Irf7-siRNA, and finally; decreasing expression of *Oasl1* and *Ifit1* were observed with increasing concentration of siRNA targeting the *Nfkb2* gene, with the exception of 100 nM Nfkb2-siRNA.

(a)

(b)
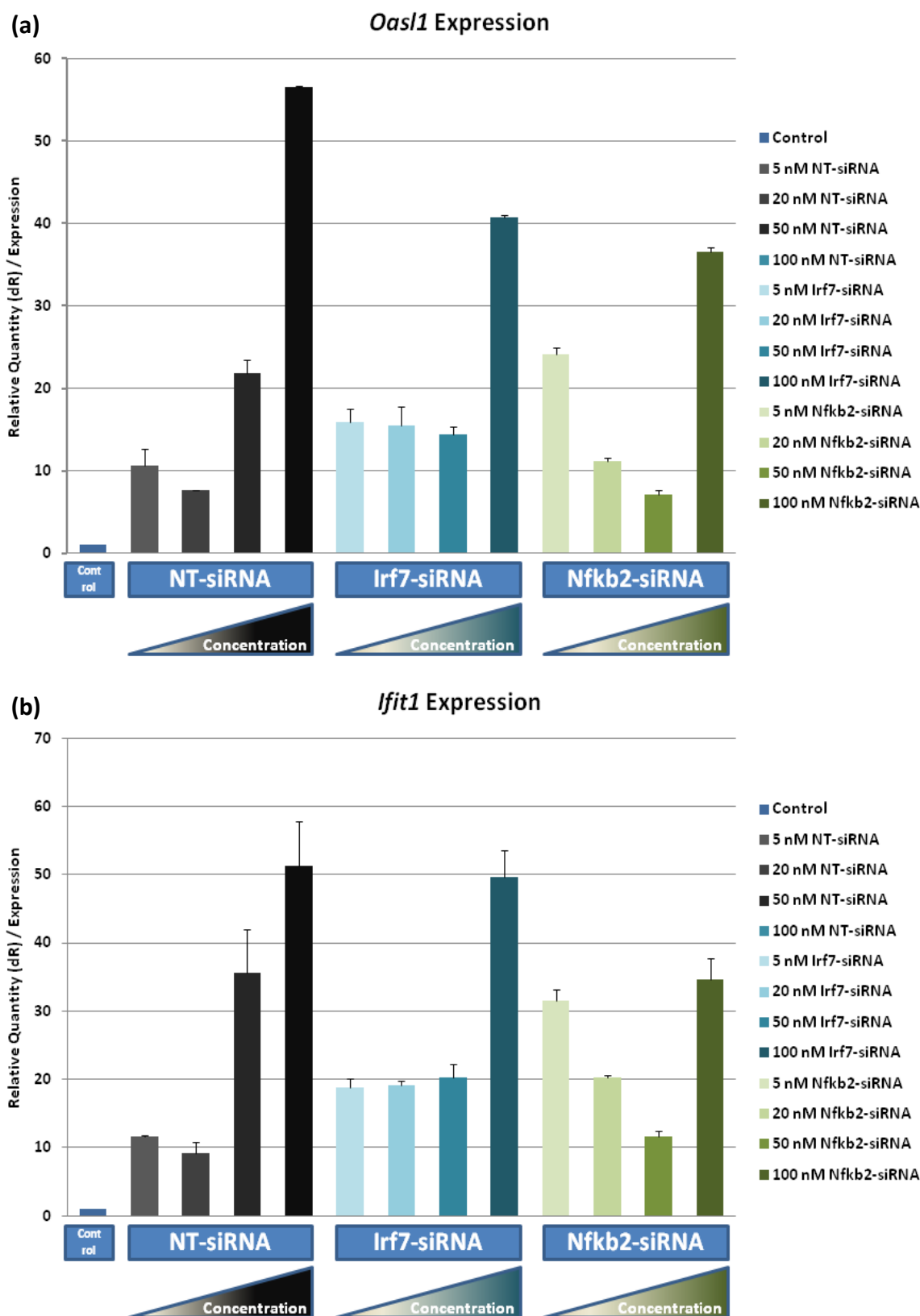
Figure 5.8: Type-I response gene expression in response to increasing concentrations of siRNA transfected into mouse BMDMs. (a) expression of *Oasl1* or **(b)** *Ifit1* was measured by QPCR in samples treated with four different concentrations of siRNA targeting either *Irf7* gene or *Nfkb2* gene or non-targeting for 24 hours.

1.4   Changes in cell morphology with increasing concentrations of siRNA combined with   LPS treatment.

A key aspect of the assay design was to incorporate LPS stimulation of the macrophages in order to better understand how siRNA induced knockdown of the GOI may perturb the (interferon component of the) LPS response. Consequently the cells would be exposed to two waves of immune activation, firstly by the process of siRNA lipofection and secondly by LPS stimulation. Results suggest different concentrations of siRNA stimulate type-I response to varying extents (Figure 5.8), although the effects of subsequent LPS stimulation of the macrophages is unknown. In order to better understand the effects of the double stimulation of macrophages, cell morphology was monitored in BMDMs pre-treated with NT-siRNA at four different concentrations and then stimulated with modest concentrations of LPS (5ng/mL). Distinct changes in cell morphology in response to LPS stimulation were observed in cells treated pre-treated for 24 h with 50 nM or 100 nM NT-siRNA; the cells appeared highly granular, irregular in shape, with evidence of apoptotic bodies (Figure 5.9). These observations suggested the phenotype of the LPS stimulated macrophages pre-treated with 50 nM to 100 nM siRNA are considerably different from those stimulated with LPS alone or LPS in combination with lower doses of siRNA. The combination of results (Figures 5.3 – 5.9) indicate 50 – 100 nM does not enhance target gene knockdown levels, induces stronger immune signalling than lower concentrations, and in combination with LPS drives the phenotype of the cells to one indicative of cell death.

Variability between the actions of 5 and 20 nM final concentration of siRNA was less distinguishable, and arguably the use of 5 nM concentration could have been investigated further. Given the effect of siRNA is transient and may wear off over time, and also since the previous siRNA screen was performed using 20 nM, this concentration was chosen over any lower dose for the remainder of the assay optimisation.
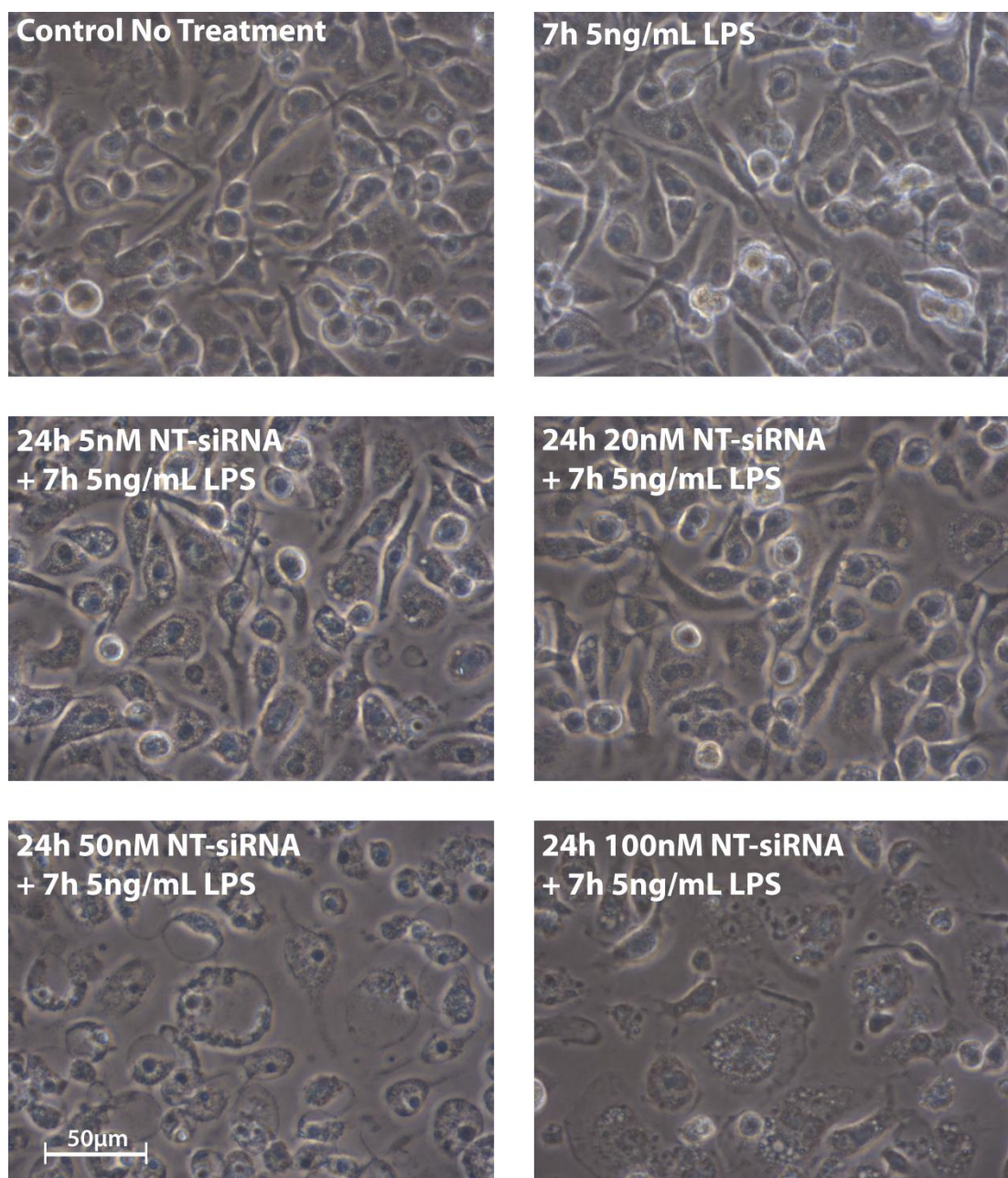
Figure 5.9: Changes in cell morphology with increasing concentration of siRNA and in combination with LPS. Mouse BMDMs were transfected with 5nM, 20 nM, 50 nM or 100 nM non-targeting siRNA for 24 hours and subsequently challenged with 5ng/mL of LPS (Salmonella Minnesota rough strain) for 7 hours. Cell morphology was assessed using a Nikon Diaphot inverted microscope and images captured at x400 magnification.

(2)   Analysis of LPS response in mouse bone marrow derived macrophages

To assess the ideal concentration of LPS to use following siRNA treatment of BMDMs the LPS response in these cells was examined. Specific questions explored included; how the LPS response compares with IFN-β stimulation of BMDMs; the concentration of LPS required to induce IFN-β mRNA detectable by QPCR; and the effects of stimulating macrophages that have already undergone siRNA lipofection with LPS.

2.1   Comparison of type-I response induced by LPS, IFN-β and NT-siRNA

To compare the extent of (type-I) activation of macrophages by 100 ng/ml LPS with IFN-β at a concentration (10 U/ml) previously studied (Chapter-4) and the intended final concentration of siRNA to be used (20 nM), BMDMs were treated with the three stimuli for 24 h and type-I gene expression assessed by QPCR (Figure 5.10). At this time-point the expression of the *Oasl1* response gene was 13 to 30 fold more potent in 100 ng/ml LPS treated cells compared to IFN-β treated or NT-siRNA treated cells respectively. To further compare type-I response activation following IFN-β or LPS treatment, *Oasl1* expression was assayed at 24 h post treatment in response to a range of concentrations of IFN-β and LPS. The most comparable induction of *Oasl1* was achieved at 100 U/ml Ifn-β and 10 ng/ml LPS (Figure 5.11).
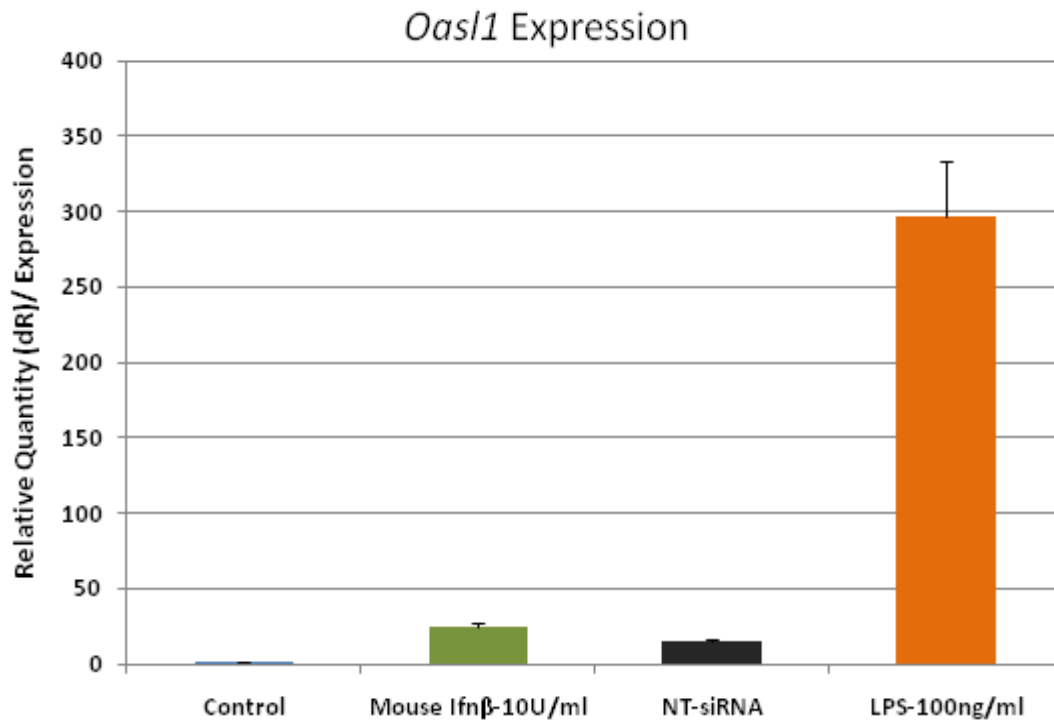
Figure 5.10: *Oasl1* **gene expression in mouse bone marrow derived macrophages 24 hours following stimulation with mouse Ifn-β or non-targeting siRNA or LPS.** Mouse BMDMs were treated with either 10 U/ml mouse Ifn-β, or 100 ng/ml LPS, or with 20 nM final concentration siRNA for 24 hours, or were untreated (control). Expression of the type-I response gene *Oasl1* was measured by QPCR and is shown relative to its basal expression (control).
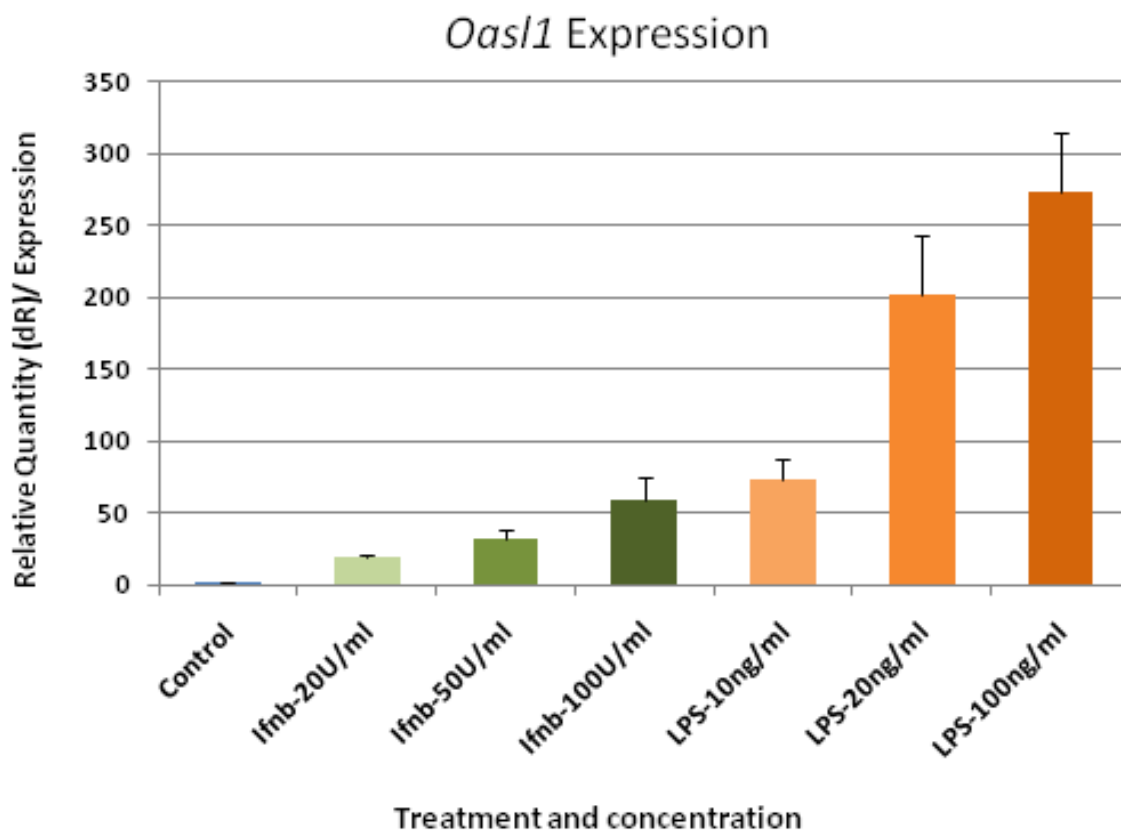
Figure 5.11: *Oasl1* **gene expression in response to increasing concentrations of mouse IFN-β or LPS stimulation of mouse bone marrow derived macrophages.** Mouse BMDMs were stimulated with either IFN-β at concentrations of 20 U/ml, 50 U/ml or 100 U/ml, or with LPS at doses of 10 ng/ml, 20 ng/ml, or 100 ng/ml for 24 hours, or were untreated (control). The relative expression of the type-I response gene *Oasl1* was measured by QPRC from RNA extracted from the 24 hour treated samples.

In the context of this assay, LPS would be used to stimulate IFN-β production and type-I response gene expression. Although the contribution of IFN-β signalling in the overall LPS response is acknowledged, few studies have dissected the precise differences between LPS and IFN-β response. To grasp some of the differences in the expression and dynamics of gene induction by LPS and IFN-β, the expression of a selection of interferon response genes was measured by semi-quantitative PCR over time in response to stimulation by a range of LPS or IFN-β doses (Figure 5.13). Also performed in parallel for comparative purposes was the stimulation of BMDMs with 20 nM NT-siRNA. Finally, in order to determine the extent of IFN-β production in response to varying doses of LPS and in response to the NT-siRNA , *Ifnb1* gene expression was measured by QPCR in the early time-point samples (1 and 2 h) (Figure 5.12).

In agreement with previous results *Ifnb1* expression was induced in a dose responsive manner following LPS stimulation (Figure 5.12). When contrasting equivalent LPS doses, *Ifnb1* expression was in general twofold higher 2 h post LPS stimulation compared to 1 hour post stimulation. At 1 hour and 2 h post-NT-siRNA transfection, *Ifnb1* gene expression was not detectable. As expected *Ifnb1* expression was not induced in BMDMs stimulated with IFN-β.
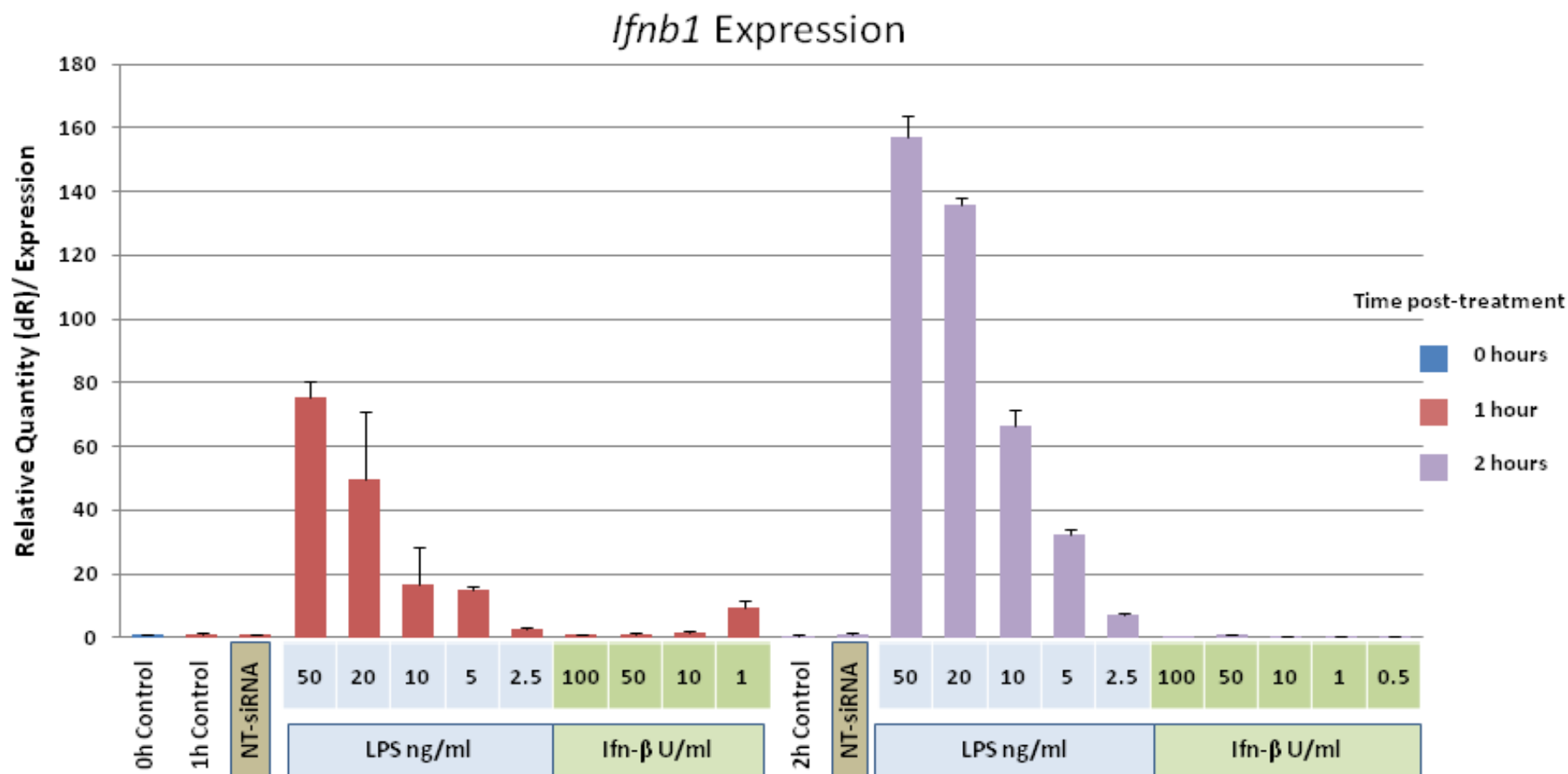
Figure 5.12: *Ifnb1* gene expression in mouse bone marrow derived macrophages following treatment with non-targeting siRNA, or varying doses of LPS, or varying concentrations mouse IFN-β cytokine. Mouse BMDMs were stimulated with any of; 20 nM NT-siRNA; LPS at 50ng/ml, 20ng/ml, 10ng/ml, 5ng/ml or 2.5ng/ml; or mouse IFN-β at 100U/ml, 50U/ml, 10U/ml, 1U/ml, or 0.5U/ml for 1 hour or 2 hours. Expression of the *Ifnb1* gene was measured by QPCR and is shown relative to *Ifnb1* expression in 0 hour (untreated) control sample.

Of the response genes assayed by semi-quantitative PCR, *Tnf* was preferentially induced by LPS, *Il1b* expression was exclusive to LPS, and both genes were induced as early as 1 hour post-stimulation (Figure 5.13). *Mx2* and *Socs1* genes were induced earlier by IFN-β and later by LPS, possibly reflecting the time-lag between LPS stimulation and autocrine IFN-β production. *Oasl1* was induced marginally earlier in response to IFN-β (1 h) compared to LPS treatment (2 h). However by 7 h both LPS stimulated cells and cells treated at 100 U/ml to 10 U/ml expressed comparable levels of *Oasl1*. *Irf5* was induced by both LPS and IFN-β. The response genes analysed were not induced at 1 hour or 2 hour post NT-siRNA transfection, however by 7 h *Irf1*, *Irf5*, *Tnf*, *Mx2*, *Socs1* and *Oasl1* were expressed in cells transfected with NT-siRNA. Since *Socs1*, *Mx2* are expressed early in response to IFN-β stimulation, this might suggest that IFN-β is produced some hours after siRNA transfection and may explain the absence of *Ifnb1* expression at 1 h and 2 h post transfection.
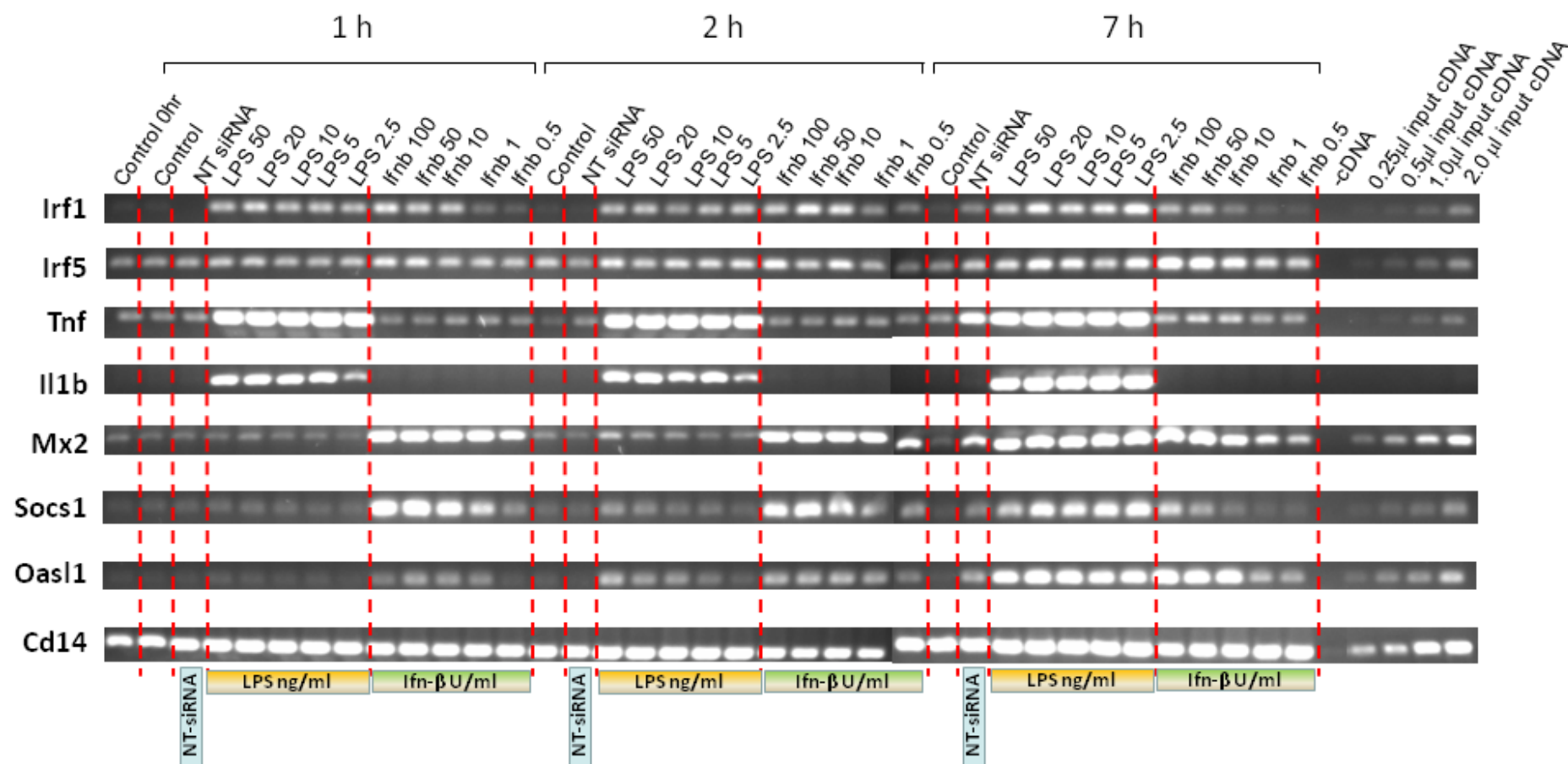
Figure 5.13: Time-course of expression of selected immune response genes in mouse bone marrow derived macrophages following treatment with non-targeting siRNA, or varying doses of LPS, or varying concentrations mouse IFN-β cytokine. Mouse BMDMs were stimulated with any of; 20 nM NT-siRNA; LPS at 50, 20, 10, 5 or 2.5 ng/ml; or mouse IFN-β at 100, 50, 10, 1, or 0.5 U/ml for 1 hour, 2 hours or 7 hours. Expression of the selected immune response gene was measured by semi-quantitative PCR. Expression of the macrophage cell surface marker (CD14), serves as an indicator of cDNA loading.

2.2 Analysis of macrophage response to combined siRNA and LPS treatment

Clearly detectable levels of *Ifnb1* and type-I response expression was induced at the LPS doses examined (50 to 2.5 ng/ml) (Figure 5.12). However the effects of combining LPS and siRNA treatment on *Ifnb1* and type-I expression were unknown. Preliminary experiments, suggested treating BMDMs with 20 ng/ml LPS after they have undergone prior siRNA transfection is toxic to the cells (data not shown) and resulted in apoptosis and reduced yields of RNA, compared to cell cultures treated with siRNA or LPS alone. It was possible that using lower concentrations of LPS could circumvent this issue of toxicity, given the correlation between LPS dose and type-I response induction. Therefore the effect of combining siRNA transfection with subsequent LPS treatment was examined using lower concentrations of LPS (10 to 0.02 ng/ml) than previously employed (Figure 5.14). *Ifnb1* expression was measured at 2 and 7 h post-LPS treatment in BMDMs with and without 24 h prior NT-siRNA transfection. Also in parallel with the LPS treatment additional BMDMs were transfected with NT-siRNA and cells harvested at 2 or 7 h post-treatment. Once again a dose responsive induction of *Ifnb1* expression was observed in response to increasing doses of LPS (Figure 5.14). Furthermore the results indicate *Ifnb1* expression in response to LPS stimulation is amplified in cells pre-treated with siRNA. In cells treated with NT-siNA alone, *Ifnb1* expression was clearly evident at 7 h post-transfection. This suggests the optimal time to detect *Ifnb1* expression in response to siRNA transfection is some h later than the transfection time and not 1 or 2 h post treatment as is suited to detecting *Ifnb1* expression in response to pathogen challenge. The results also showed slightly elevated levels of *Ifnb1* expression in cells treated with NT-siRNA for 24 h compared to control untreated. On this occasion the combination of lower doses of LPS with prior siRNA treatment did not appear to induce cell death and consistent yields of RNA were extracted from all samples.
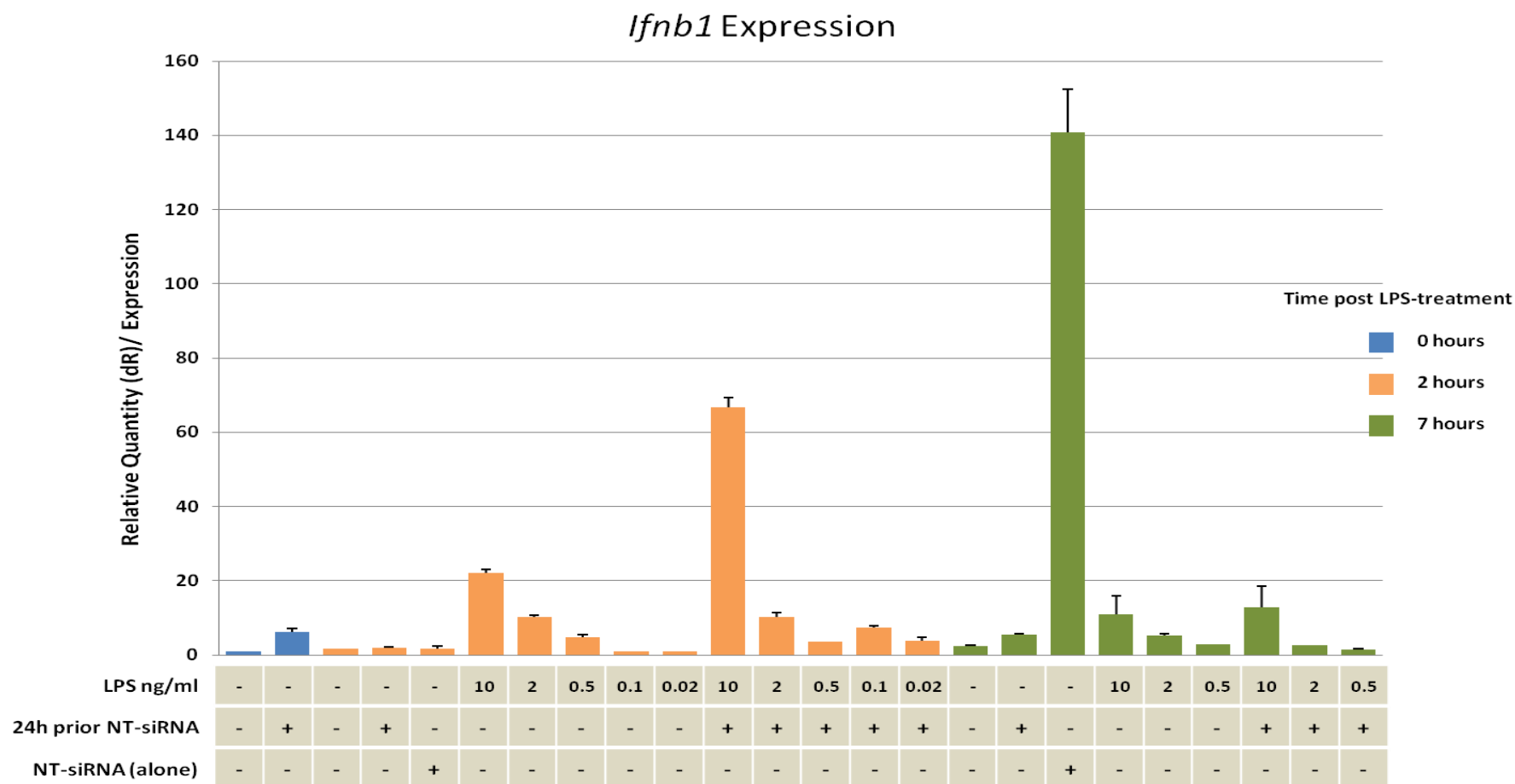
Figure 5.14: *Ifnb1* **expression in mouse bone marrow derived macrophages stimulated with varying concentrations of LPS, in the presence and absence of prior non-targeting–siRNA transfection.** Mouse BMDMs were transfected with 20 nM final concentration NT-siRNA or were untreated. 24 hours later both cells that had been transfected with siRNA or left untreated were further challenged with varying doses of LPS (10ng/ml, 2ng/ml, 05.ng/ml, 1ng/ml and 0.02 ng/ml) for 2 hours and 7 hours. Running parallel with the LPS treatment, additional mouse BMDMs were transfected with NT-siRNA for 2 hours and 7 hours. *Ifnb1* expression was determined by QPCR and is shown relative to its expression in the 0 hours control sample.

2.2.1 Analysis of macrophage response to 5 ng/ml LPS and 20 nM siRNA (i.e. concentrations chosen for the in vitro assay)

A dose of LPS (5ng/ml) falling within the range of low doses examined (10 – 0.02 ng/ml) was chosen as a potential concentration to be used in the final assay design. To compare the dynamics of *Ifnb1* induction in response to 5ng/ml LPS or 20 nM siRNA, BMDMs were treated with LPS or siRNA and the expression of *Ifnb1* assayed by QPCR over a time-course. Following 5ng/ml LPS treatment maximal *Ifnb1* expression was reached 2-4 h post treatment (Figure 5.15). In response to 20 nM NT-siRNA *Ifnb1* expression was reached around 6-8 h, corroborating earlier results (Figures 5.13 and 5.14).
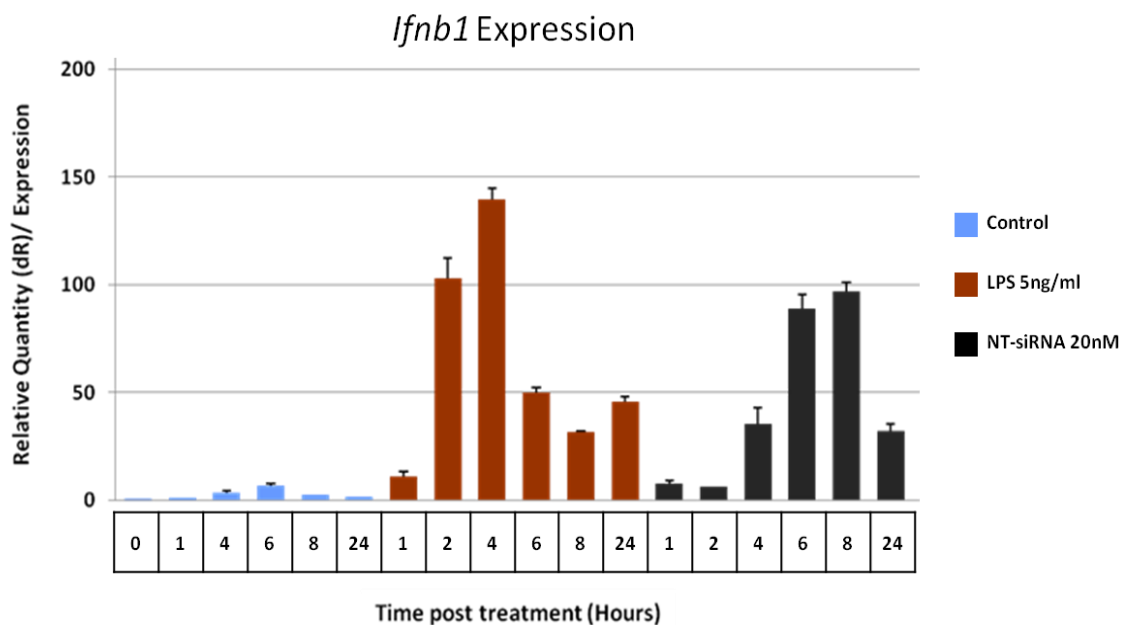
Figure 5.15: Time-course of *Ifnb1* expression in bone marrow derived macrophages following stimulation by LPS, or non-targeting siRNA. Mouse BMDMs were stimulated with 5ng/ml LPS, or 20 nM NT-siRNA, and the cells harvested at 1, 2, 4, 6, 8, and 24 hours post stimulation. Untreated (control) cells were also harvested at the same time-points and at time 0 hours. *Ifnb1* gene expression was measured in the samples by QPCR and is shown relative to its expression at 0 hours (control).

To further validate the use of 5ng/ml LPS, BMDMs were transfected with siRNA targeting the Irf3 gene, or the Nfkb2 gene or with non-targeting-siRNA, and after 24 h cells were treated with 5ng/ml LPS (Figure 5.16). *Ifnb1* expression was assayed at 0, 1, 2, 4, 6, and 8 h post-LPS treatment, where 0 h equates to the 24 h time-point of siRNA treatment. Irf3 is a transcription factor known to bind to the *Ifnb1* enhancesome [282], and could serve as a positive control when monitoring the effects of targeted gene knockdown on *Ifnb1* expression and the type-I response. *Ifnb1* expression was suppressed below basal levels in cells treated with Irf3 siRNA alone for 24 h (i.e. 0 h post LPS treatment) (Figure 5.16). *Ifnb1* expression was also suppressed in cells treated with siRNA targeting the Nfkb2 gene (for 24 h), relative to cells treated with NT-siRNA for 24 h (Figure 5.16). Furthermore for every comparable time-point post-LPS treatment, *Ifnb1* expression was lower in those cells treated with siRNA targeting either Irf3 or Nfkb2 expression compared to those treated with NT-siRNA. In contrast to earlier observations (Figure 5.13) combining siRNA transfection with subsequent LPS treatment resulted in lower levels of *Ifnb1* expression, compared to LPS treatment alone (Figure 5.16). The data also suggested maximal *Ifnb1* expression is obtained around 2 h post-LPS treatment (with or without prior siRNA transfection). Finally, the cells appeared viable in response to 5ng/ml LPS treatment regardless of prior siRNA transfection and the phenotype of the cells was not altered compared to those untreated or treated only with LPS (data not shown).

Based on the range of observations, 5ng/ml was chosen as the concentration of LPS to stimulate the BMDMs, following siRNA mediated knockdown of the GOI. The results also suggested that ~7 h post-siRNA transfection was the optimal time to detect siRNA lipofection induced *Ifnb1* expression. Whereas LPS induced *Ifnb1* expression was best detected ~2 h post-stimulation.
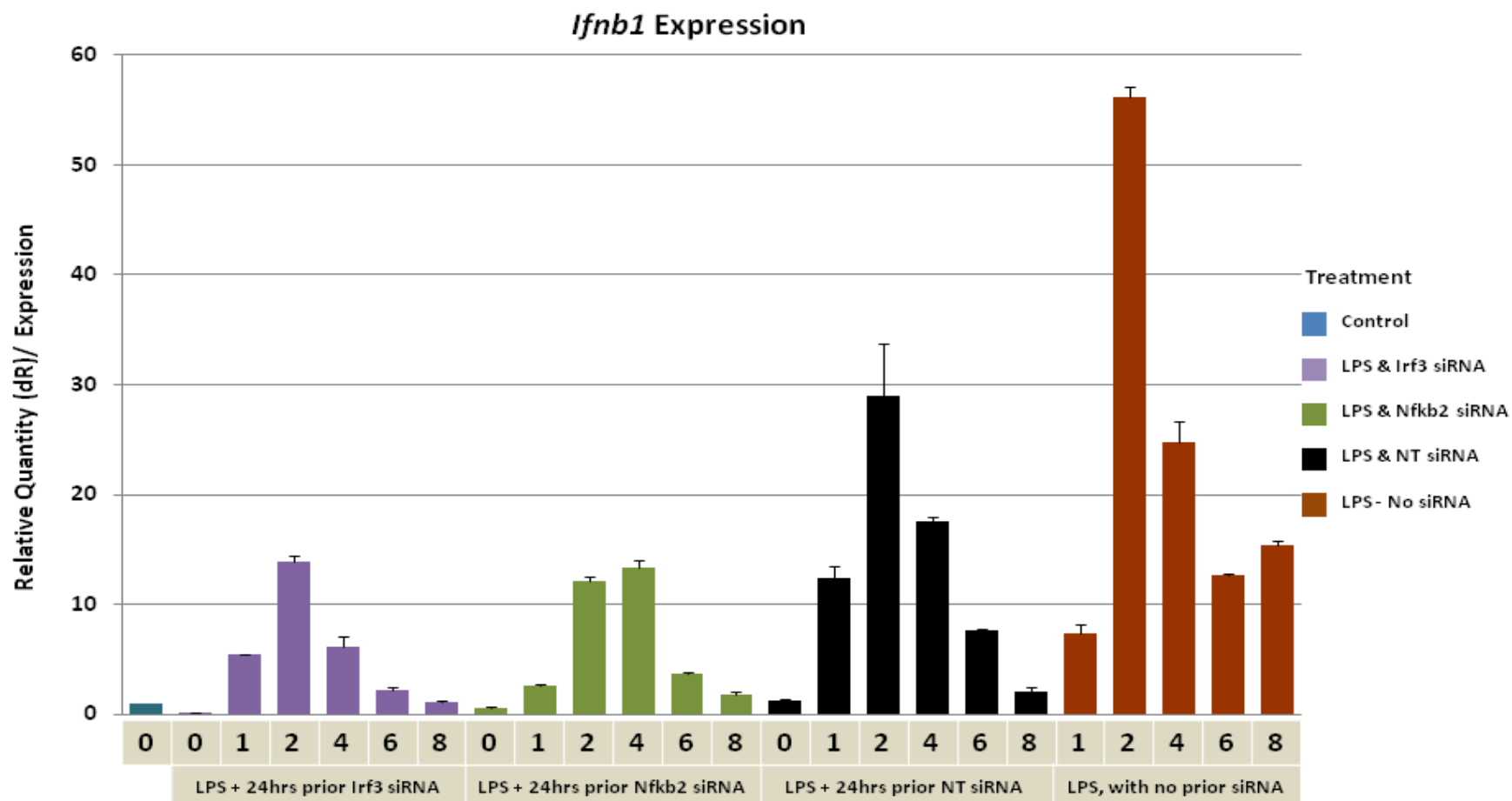
Figure 5.16: Time-course of *Ifnb1* expression in bone marrow derived macrophages following stimulation by LPS, with and without prior siRNA transfection targeting the *Irf3* or *Nfkb2* genes or non-targeting. Mouse BMDMs were transfected with siRNA targeting the *Irf3* or *Nfkb2* genes or with non-targeting siRNA. 24 hours post siRNA lipofection, the cells were stimulated with 5ng/ml LPS, and the cells harvested at 1, 2, 4, 6, 8, and 24 hours post stimulation. Untreated (control) cells were also harvested at the same time-points and at time 0 hours. *Ifnb1* gene expression was measured in the samples by QPCR and is shown relative to its expression at 0 hours (control).

(3) Determination of ideal siRNA treatment time for obtaining optimal gene knockdown and type-I repression following LPS stimulation.

3.1 Analysis of protein level knockdown over time

Typical time-points for detecting optimal knockdown of target genes by siRNA at the transcriptional level (as recommended by the manufacturer (Dharmacon Thermo Scientific) range from 24 to 48 h. Knockdown at the protein level is typically observed around 48 h. To understand the dynamics of protein expression as well as suppression by siRNA, BMDMs were stimulated with NT-siRNA, LPS, or siRNA targeting Nfkb2 for 4, 8, 24, and 48 h and Western blot analysis performed on protein extracts. Nfkb2 protein expression was induced by LPS to a greater extent than NT-siRNA transfection (Figure 5.17a). Expression of the p100 and p52 isoforms peaked at 24 h post LPS or NT-siRNA treatment. Knockdown of Nfkb2 protein was apparent at 24 h post transfection with siRNA targeting Nfkb2, and was further reduced at 48 h (Figure 5.17a). Successful knockdown of *Irf5* (relative to NT-siRNA transfected samples) at the protein level, was also observed 48 h following tranfection of siRNA targeting the *Irf5* (Figure 5.17b).
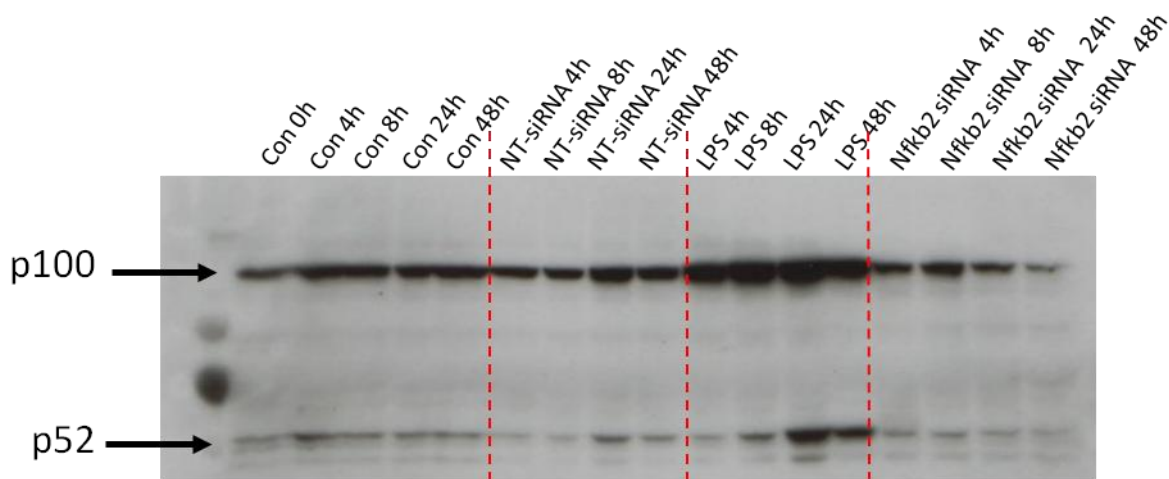
Figure 5.17: (a) Time-course of Nfkb2 p100 and p52 protein subunit expression in mouse BMDMs following treatment with NT-siRNA, Nfkb2 siRNA or LPS (20ng/ml). Mouse BMDMs were treated with 20 nM NT-siRNA, or 20 nM of siRNA targeting Nfkb2, or 20ng/ml LPS. Following 4, 8, 24 or 48 hours or treatment cells were harvested and western blot analysis performed on protein extracts from the cells using an Nfkb2 antibody targeting the p100 and p52 isoforms of the Nfkb2 protein. 20ng/ml LPS was the most potent inducer of Nfkb2 p100 and p52 subunit expression. NT-siRNA also induced p100 and p52 expression following 24 hours of treatment. Nfkb2-siRNA was effective at knocking down p100 and p52 expression by 24 and 48 hours of treatment.
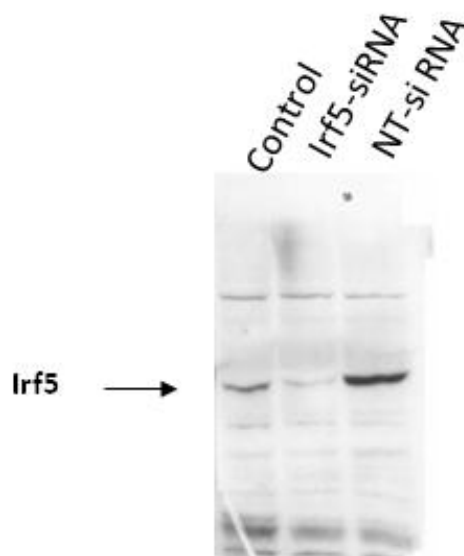
Figure 5.17: (b) Expression of Irf5 protein in BMDMs following treatment with siRNA targeting the *Ifr5* gene or NT-siRNA. Mouse BMDMs were transfected with 20 nM siRNA targeting the Irf5 gene or NT-siRNA. 48 hours post transfection cells were harvested and western blot analysis performed on protein extracts from the cells using an Irf5 antibody. Ifn5 expression was induced in cells treated with NT-sRNA, but repressed (relative to control-untreated) in cells treated with Irf5 targeting siRNA.

3.2   Analysis of gene-level knockdown and efficacy of downstream response 24 h or 48 h post siRNA transfection

Up until this stage all experiments exploring the role of GOIs in the type-I/LPS response, were based on 24 h of prior treatment with siRNA. The efficacy of GOI knock-down achieved at different time-points post-siRNA transfection had not been explored. Given knockdown at the protein level is apparent later than that at gene level there could be variability in the effect of GOI knockdown in the LPS response as monitored after different times of treatment with siRNA. Therefore to examine the efficacy of gene knockdown over time as well as determine the optimal time post-siRNA to obtain the most efficacious response to GOI knock-down in the context of the LPS response, mouse BMDMs were transfected with siRNA targeting Nfkb2, and 24 or 48 h post transfection cells were challenged with 5ng/ml LPS. At both 24 h and 48 h post siRNA transfection *Nfkb2* expression was knocked down at the message level (~ 40 – 50%) relative to NT-siRNA tranfected cells (Figure 5.18). In response to subsequent LPS stimulation, Nfkb2 expression followed a similar pattern of expression regardless of whether cells had been treated for 24 h or 48 h of prior siRNA (5.19). In both cases Nfkb2 expression was induced by LPS stimulation alone and in combination with siRNA, and to a lesser extent in cells treated with prior siRNA targeting the Nfkb2 transcript.

Type-I expression was also assayed in response to LPS stimulation in cells with 24 or 48 h prior siRNA targeting *Nfkb2*, *Ifnb1*, or *Sod2* transcripts or non-targeting siRNA. *Ifit1* expression was induced in a similar fashion and extent in response to pre-treatment with the different siRNAs at both 24 and 48 h (Figure 5.20). LPS challenge in cells pre-treated with NT-siRNA and siRNA targeting Sod2 induced *Ifit1* expression to a greater extent than siRNA targeting Nfkb2 or Ifnb1, regardless of the duration of prior-siRNA treatment.

Taken together the data suggested there was not a major difference at the message level (in terms of knockdown efficiency and type-1 response) of treating with siRNA for 24 or 48 h. Considering this and the fact our previous investigations [146] were based on 24 h of treatment with siRNA, this duration was chosen for the final assay design.
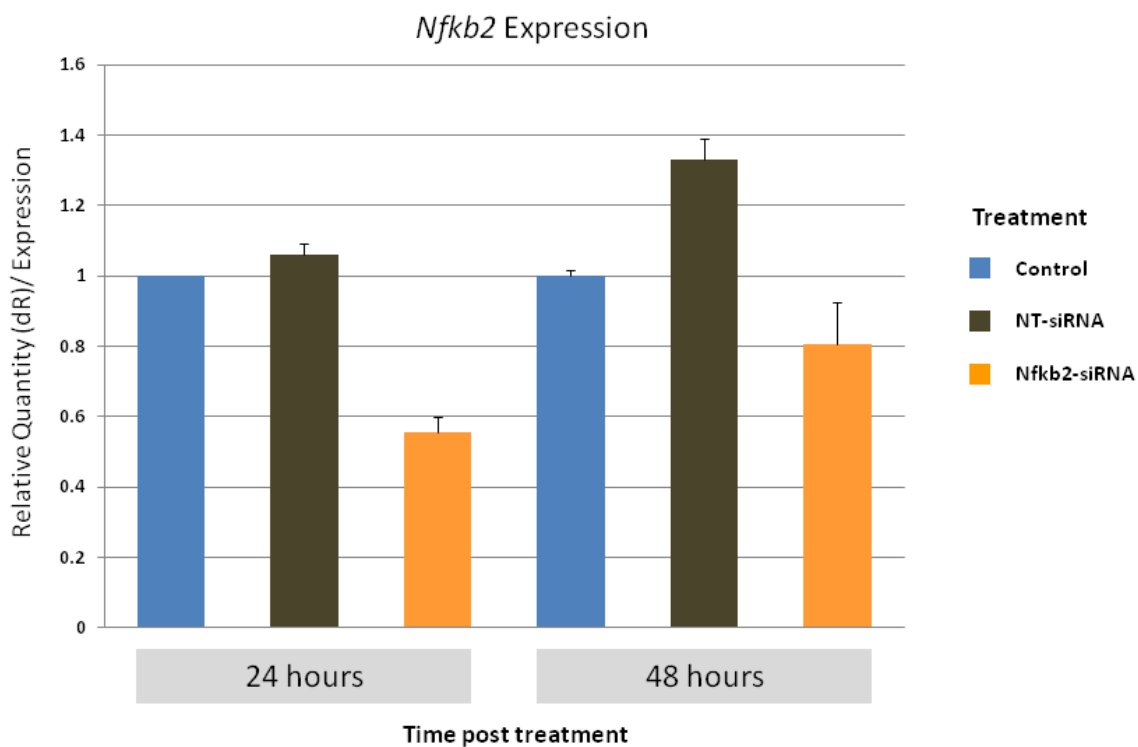
Figure 5.18: Nfkb2 expression in mouse bone marrow derived macrophages 24 hours or 48 hours post transfection with siRNA targeting the Nfkb2 gene. Mouse BMDMs were transfected with siRNA targeting the *Nfkb2* gene or non-targeting siRNA for 24 or 48 hours, or cells were left untreated for the duration (control). Expression of the Nfkb2 gene was then assessed by QPCR in all samples.

211
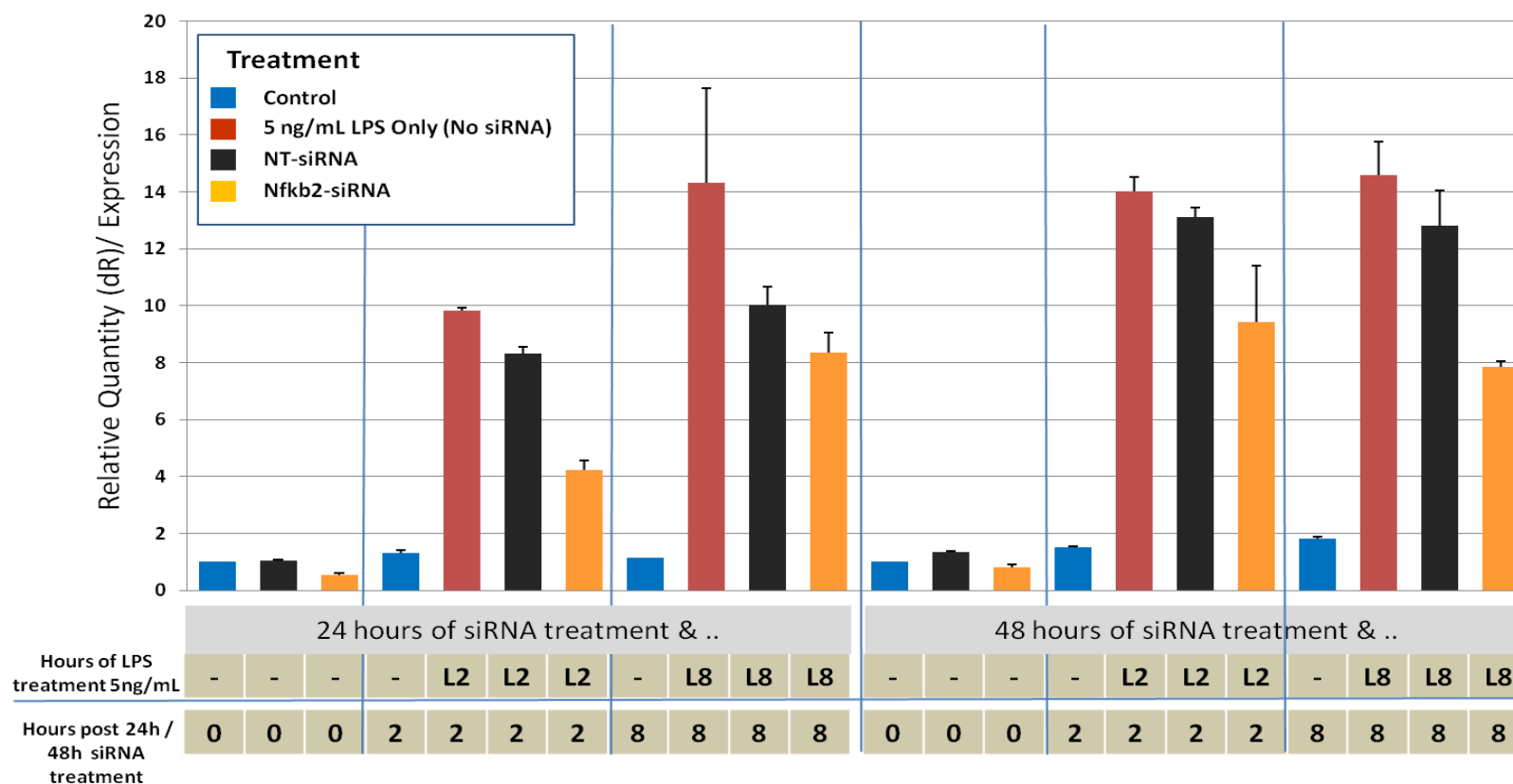
5.19: Nfkb2 expression in mouse bone marrow derived macrophages stimulated with LPS following 24 hours or 48 hours of treatment with siRNA targeting the Nfkb2 gene. Mouse BMDMs were transfected with siRNA targeting the Nfkb2 gene or non-targeting siRNA for 24 or 48 hours, or cells were left untreated for the duration (control). Subsequently cells were stimulated with 5ng/ml LPS for 2 hours or 8 hours. Expression of the Nfkb2 gene was then assessed by QPCR in all samples, having undergone siRNA transfection and/or LPS stimulation, or no treatment (control).
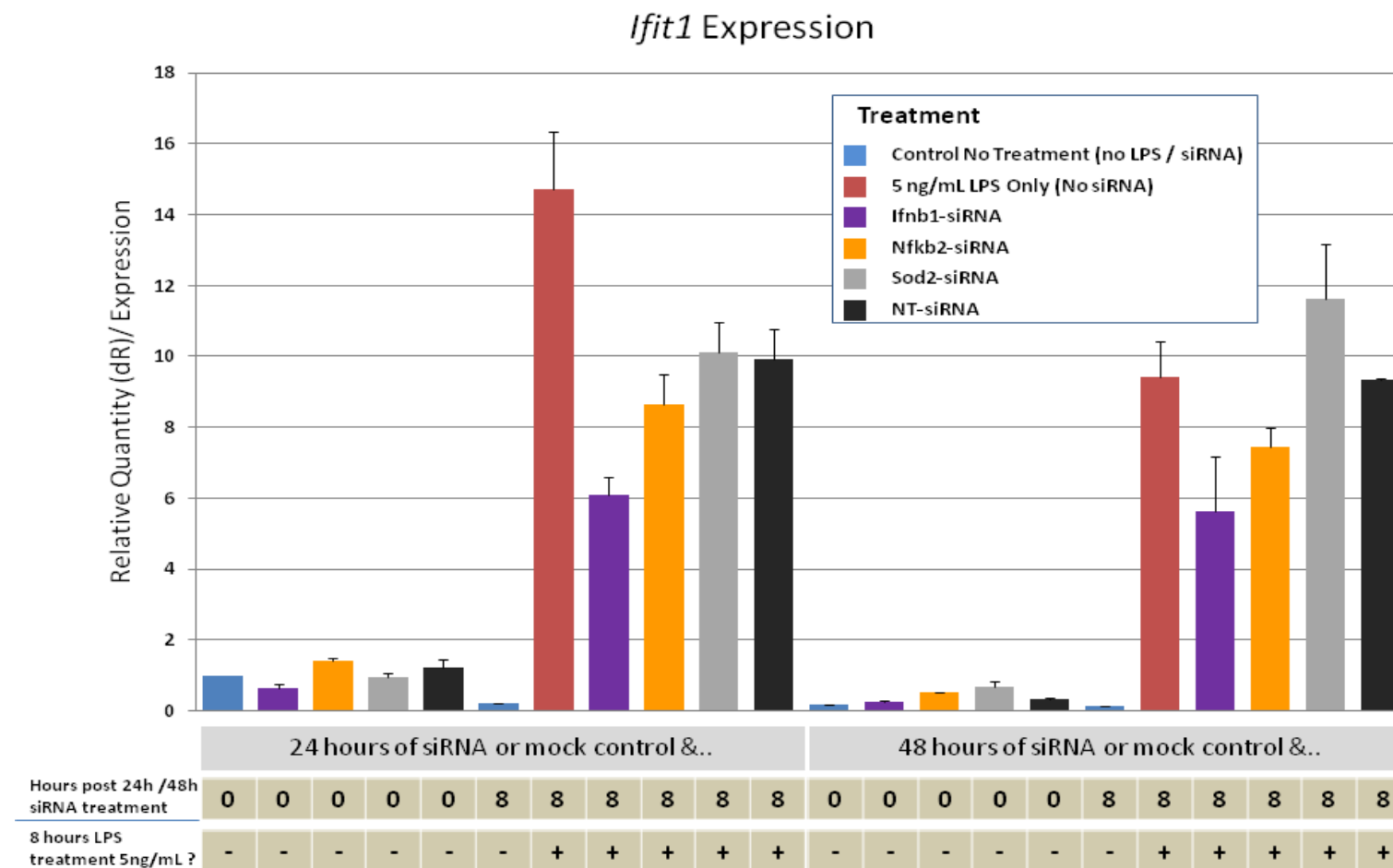
Figure 5.20: ***Ifit1*expression in mouse bone marrow derived macrophages in response to 2 hours of LPS stimulation, with and without 24 hours or 48 hours of treatment with siRNA targeting the *Ifnb1* or Nfkb2 *or Sod2* genes.** Mouse BMDMs were transfected with siRNA targeting the *Nfkb2* gene, or the *Ifnb1*, or *Sod2* gene or with non-targeting siRNA. Cells were also left untreated (control). Following 24 hours or 48 hours incubation with the siRNA the cells were stimulated with 5ng/ml LPS for 8 hours. Expression of the *Ifit1* gene was then assessed by QPCR in all samples, having undergone siRNA transfection and/or LPS stimulation, or no treatment (control).

Optimised work-flow for studying the role of Genes of Interest in the macrophage response to type-I interferons signalling stimulation

Based on the series of questions addressed in results sections 1-3, the work-flow for assessing the role of GOI in the type-1 and LPS response was finalised as illustrated in Figure 5.21. Mouse BMDMs were to be cultured from femurs of 7-10 week old male BALB/c mice with recombinant CSF-1. Although the assay optimisation process was performed on macrophages derived from the C57BL/6 strain of mice, the decision was taken to use BALB/c macrophages for the *in vitro* screen. The reasons for this change in mouse strain is discussed in greater detail and justified in the Discussion section. The siRNA and LPS induced effects on macrophages derived from the two mice strains were examined (data not shown). Similar levels of gene knockdown were achieved and similar patterns of type-I expression were observed in both strains. The major difference was the extent of type-I gene expression, which was higher in macrophages derived from C57BL/6 and has been observed by others [283].

Following seven days of differentiation the cells were to be transfected with 20 nM final concentration of siRNA for a total of 24 h. However a portion of cells would be harvested at 7 h post transfection to monitor levels of *Ifnb1* gene expression in response to the transfection process, since this was found to be the optimal time to monitor *Ifnb1* expression following siRNA lipofection. Cells would also be harvested at 24 h post siRNA transfection in order to assay type-I response stimulation. The remaining transfected cells would be stimulated with 5 ng/ml LPS and cells harvested at 2 h and 7 h post LPS to assay *Ifnb1* expression and ISG expression respectively.
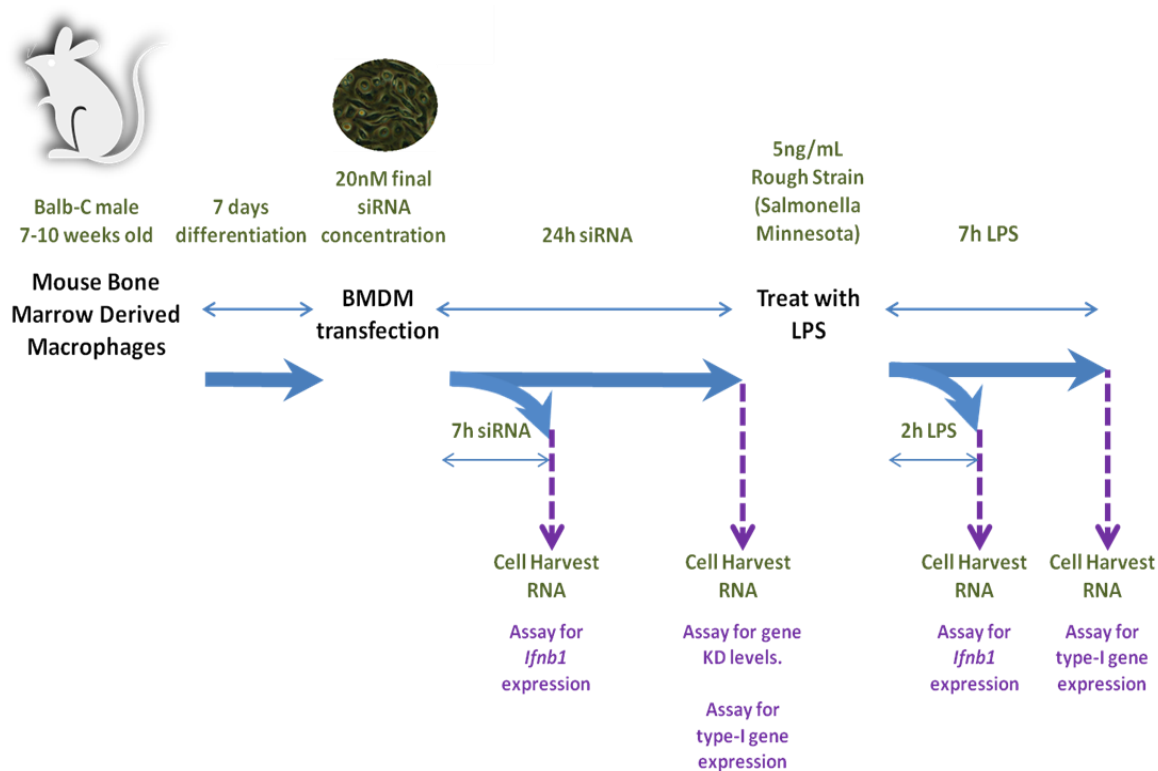
Figure 5.21: Finalised work flow for studying the role of gene of interests in the type-I response. Bone marrow derived macrophages are cultured by obtaining cells from 7-10 week old BALB/c male mice and differentiating using CSF-1. Following 7 days of differentiation cells are transfected with siRNA at a final concentration of 20 nM (siRNA) for 24hours before proceeding to the next stage of the experiment. However samples of siRNA treatment alone are taken at 7 hours and 24 hours post transfection to obtain RNA for assaying *Ifnb1* production and type-I response activation respectively. The 24 hour samples can also be used to determine target gene knockdown. Remaining siRNA transfected samples are treated with 5ng/ml concentration of LPS and samples harvested at 2hours and 7hours post treatment to obtain RNA for assaying *Ifnb1* production and type-I response activation respectively.

(4) Screen of siRNAs targeting genes of interest and their effect on the macrophage response to siRNA lipofection and LPS stimulation

Having established and optimised a workflow for testing the effect of siRNA mediated gene knockdown in mouse BMDMs, the desire was to now (a) test the assay on a medium-throughput scale and (b) test the contribution of genes of interest (GOI) in type-I interferon signalling. The rationale behind the choice of genes selected is discussed below.

Genes of interest (GOI) potentially implicated in IFN-β signalling

To validate and extend the observations of previous investigation [146], genes that are either known or could potentially play a role in IFN-β signalling were chosen for analsysis in the in vitro assay (Table 5.1). Some of the selected targets have well established roles in IFN-β signalling (Ifnb1, Irf3, Nfkb1, Rela and Irf7) [282] and were therefore expected to repress LPS induced type-I signalling. Conversely non-targeting siRNA, or siRNA targeting the Sod2 gene, would not be expected to repress the transfection or LPS induced type-I response; therefore if other gene knock-downs generate similar levels of ISG expression it may suggest they also are not crucial for type-I signalling. To validate previous observations, Irf5 and Nfkb2 were also selected for the *in vitro* screens. According to the macrophage pathway model (Chapter-2), Nfkb2 can form a homodimer with itself or form heterodimers with Rel, Rela, Relb, as well other complexes along with Bcl3 and Nfkb1. Thus the transcriptional partners of Nfkb2 were also chosen as genes of interest for further investigation. Finally Nfkbia and Socs3 which serve as inhibitory proteins of NF-κB and cytokine signalling respectively were chosen to expore whether their knockdown would result in exaggerated ISG expression (compared to NT-siRNA transfected cells).

	siRNA targets	Group/ Description	Predicted Effect of knockdown on transfection induced Type-1 response
1	Ifnb1	Interferon / Cytokine	Repress
2	Irf3	Interferon regulatory factor	Repress
3	Irf5	Interferon regulatory factor	Repress*
4	Irf7	Interferon regulatory factor	Repress
5	Nfкb1	Nfкb transcription factor	Repress
6	Nfkb2	Nfкb transcription factor	Repress*
7	Rela	Nfкb transcription factor	Repress
8	Relb	Nfкb transcription factor	?
9	Rel	Nfкb transcription factor	?
10	Nfkbia	Inhibitor of Nfкb complex	Exacerbate
11	Bcl3	Transcriptional co-activator (assoc NfкB)	?
12	Socs3	Suppressor of cytokine signalling	Exacerbate
13	Sod2	Superoxide dismutase. Mitochondrial protein	Induce type-I response (No Effect)
14	Non-Targeting	N/A (siRNA that does not target any known/ predicted sequence)	Induce type-I response (No Effect)

Table 5.1. Predicted effects of targeting different genes with siRNA in mouse bone marrow derived macrophages. 13 genes of Interest (GOI) and the control Non-Targeting siRNA (NT-siRNA) to be transfected into primary bone marrow derived macrophages. Knock-down of genes associated with positive regulation of IFN-β message production or the type-I response is expected to repress type-I signalling relative to transfection of NT-siRNA or siRNA targeting of a negative control gene. Knockdown of genes associated with the negative regulation of cytokine activation pathways are expected to exacerbate the type-I signalling.

(4.1) Screen of *Ifnb1* and type-I gene expression by QPCR following targeted gene knockdown of GOI in mouse BMDMs

Two independent screens targeting the selected GOI were conducted. *Ifnb1* and type-I expression in response to siRNA lipofection are shown in Figure 5.22 and 5.23 respectively and the targeted genes are ordered according to least-to-most-immunostimulatory response (from left to right). The scale of *Ifnb1* expression in response to siRNA transfection was ≈ 25 fold different between the two screens. Furthermore the order of least-to-most stimulatory gene targets was variable between the two screens; for example siRNA targeting Relb stimulated the greatest *Ifnb1* expression in one screen (5.22a), and was the least stimulatory gene target in the other screen (5.22b). Surprisingly in one screen (b) siRNA targeting Ifnb1 resulted in one the highest levels of *Ifnb1* expression 7 h post transfection.

Bcl3 and Relb were the gene targets resulting in the lowest levels *Oasl1* expression 24 h post-siRNA transfection in both screens (Figure 5.23). Cells treated with NT-siRNA or siRNA targeting Sod2 were amongst the samples with the highest levels of Oasl1 expression. Contrary to the predicted response, targeting the Socs3 gene did not exacerbate the transfection induced type-I response (Figure 5.23), or indeed *Ifnb1* expression 2 h following LPS stimulation (Figure 5.24). Socs3, Blc3, Nfkbia and Irf3 were amongst the gene targets that resulted in the lowest levels of *Ifnb1* expression in response to LPS stimulation in both screens. Cells treated with siRNA targeting the Rela gene had one of the highest inductions of Ifnb1 expression following LPS stimulation.

Screen- (a)

Ifnb1 Expression
7 hours post siRNA transfection
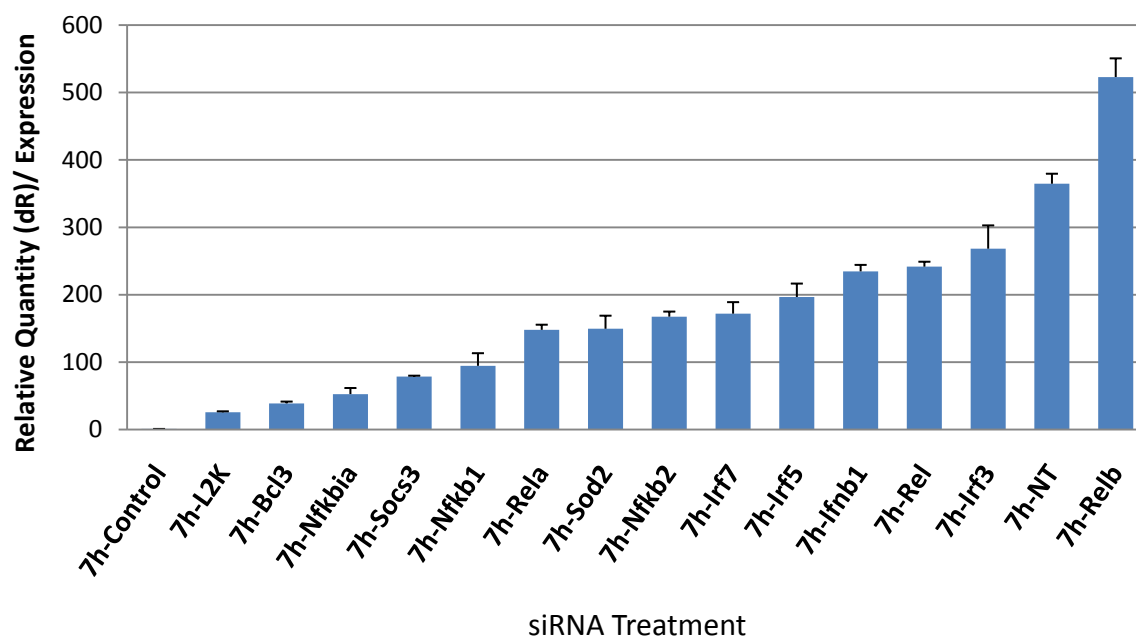
Screen- (b)

Ifnb1 Expression
7 hours post siRNA transfection
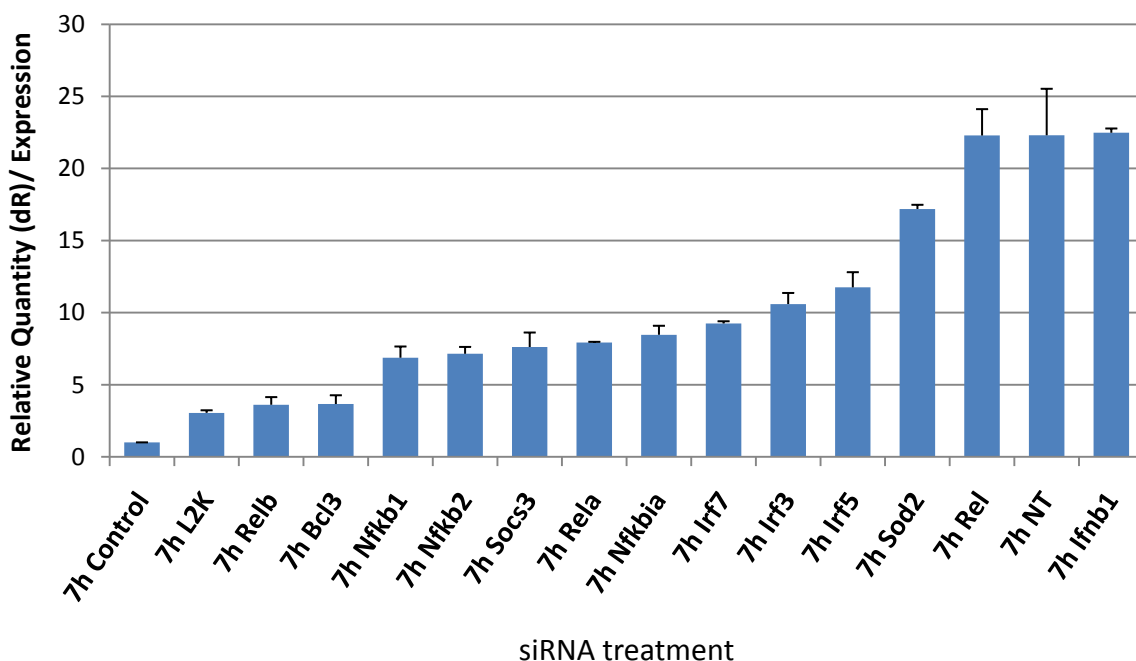
Figure 5.22: *Ifnb1* **gene expression in two independent screens of targeted RNA-interference in mouse BMDMs.** BMDMs were transfected with siRNA using lipofecamine 2000 (L2K), targeting any of 13 different genes (as identified in the graphs), NT-siRNA, L2K alone or were untreated (control). Following 7 hours of treatment cells were harvested, and RNA extracted to determine *Ifnb1* expression by QPCR. Two independent screens were performed (a and b). The bar graph is arranged to place to most stimulatory inducer of Ifnb1 to the right of the graph.

Screen- (a)

Oasl1 Expression
24 hours post siRNA transfection

Screen- (b)

Oasl1 Expression
24 hours post siRNA transfection
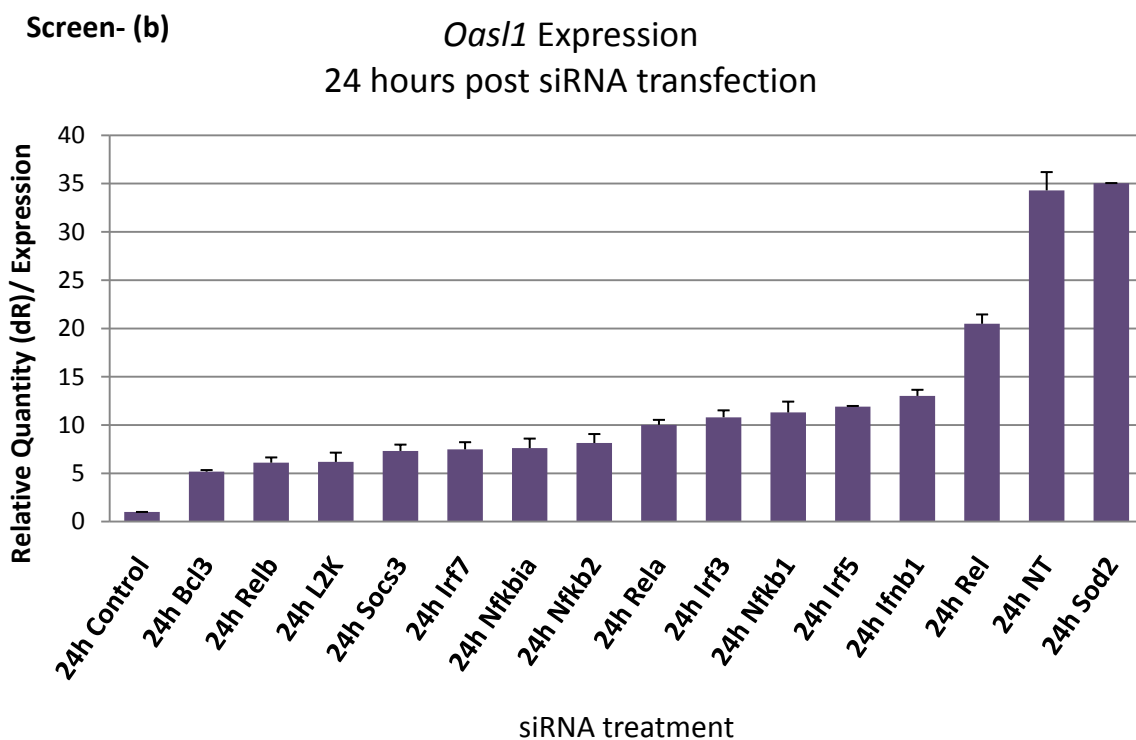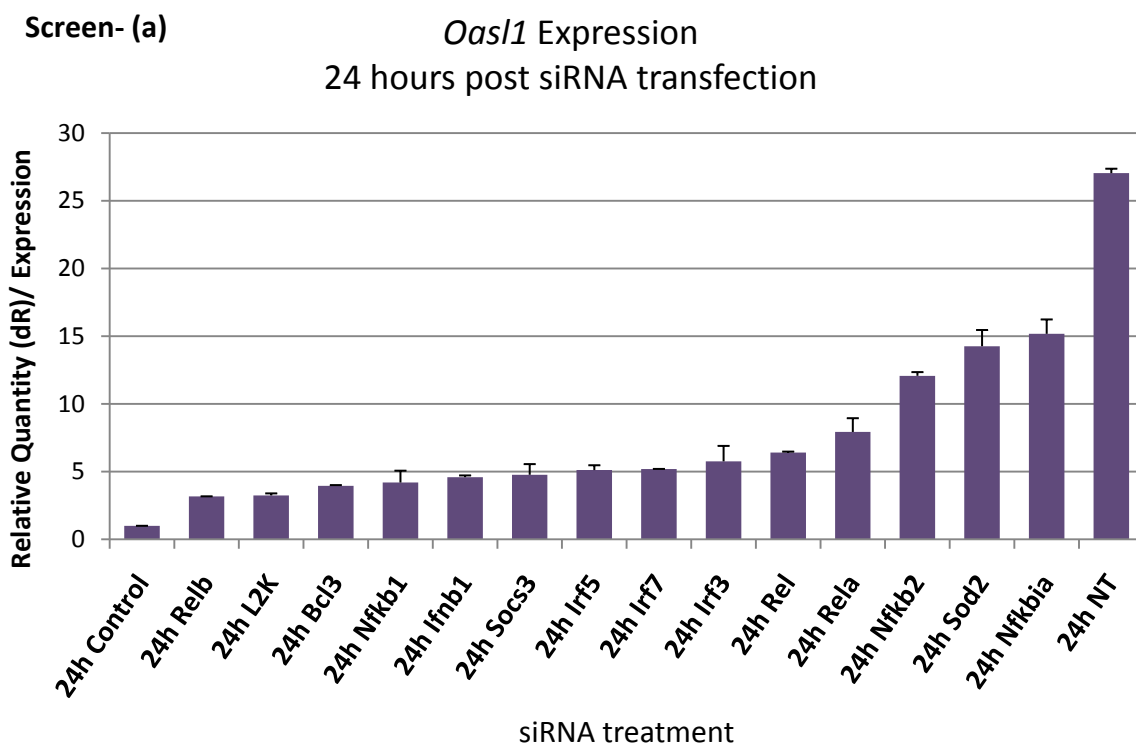
Figure 5.23: *Oasl1* gene expression in two independent screens of targeted RNA-interference in mouse BMDMs. BMDMs were transfected with siRNA using lipofecamine 2000 (L2K), targeting any of 13 different genes (as identified in the graphs), NT-siRNA, L2K alone or were untreated (control). Following 24 hours of treatment cells were harvested, and RNA extracted to determine *Oasl1* expression by QPCR. Two independent screens were performed (a and b). The bar graph is arranged to place to most stimulatory inducer of Oasl1 to the right of the graph.

Screen- (a)

Ifnb1 Expression
2 hours post LPS treatment with 24 hours prior siRNA

Screen- (b)

Ifnb1 Expression
2 hours post LPS treatment with 24 hours prior siRNA
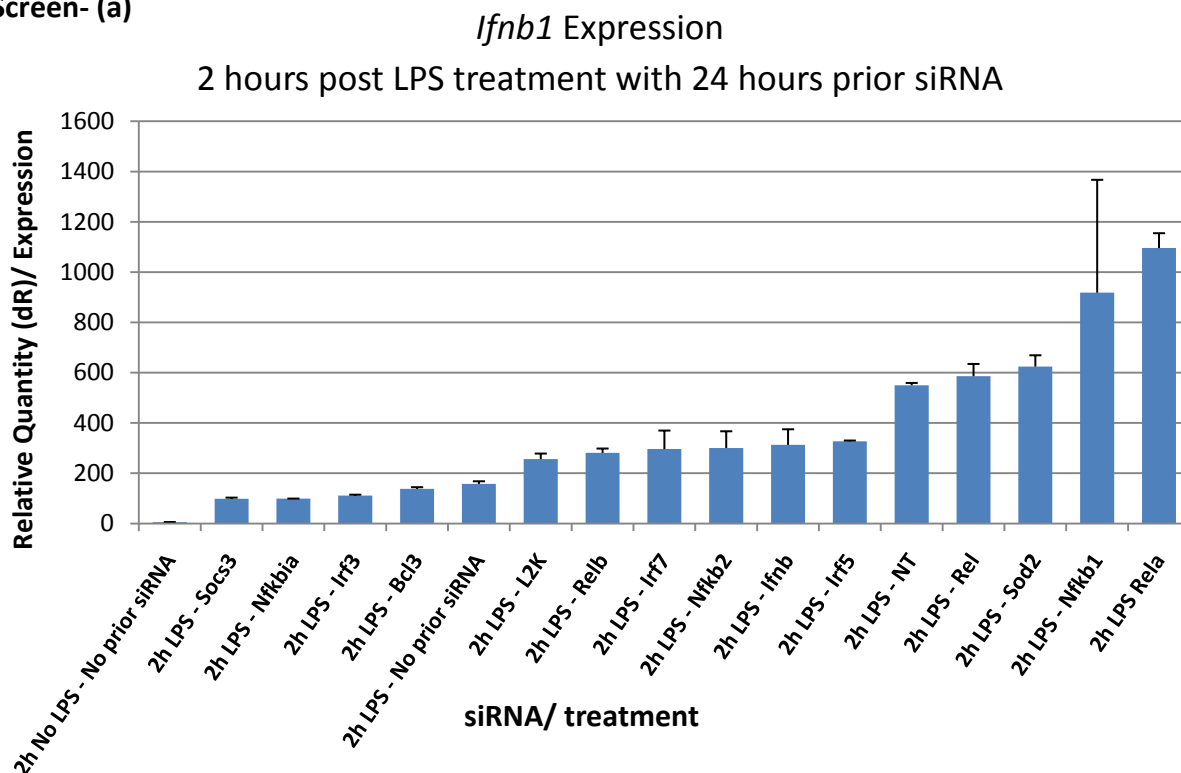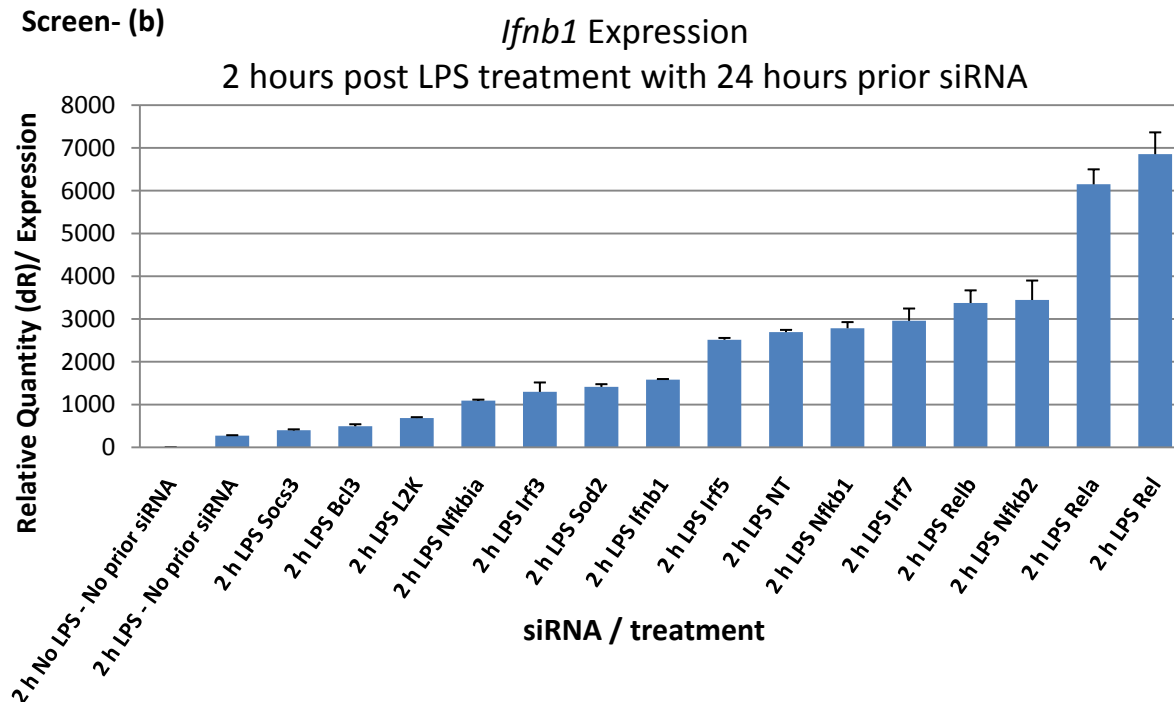
Figure 5.24: *Ifnb1* gene expression in response to LPS stimulation in two independent screens of targeted RNA-interference in mouse BMDMs. BMDMs were transfected with 20 nM siRNA using lipofecamine 2000 (L2K), targeting any of 13 different genes (as identified in the graphs), NT-siRNA, L2K alone or were untreated (control - No prior siRNA). 24 hours post siRNA treatment cells were treated with 5ng/ml LPS, and 2 hours post LPS treatment the cells were harvested and RNA extracted to determine *Ifnb1* expression by QPCR. Two independent screens were performed (a and b). The bar graph is arranged to place to most stimulatory inducer of Ifnb1 to the right of the graph.

The QPCR results of the two screens were on the whole inconclusive as to the effect of individual gene knock-downs on IFN-β signalling. Interpreting the contribution of the targeted genes was complicated by both the lack of reproducibility between screens, and contradictory action of gene-KDs within the screen. However some of the siRNA targets were more consistent in their action between screens (including Socs3, Bcl3, Rela) than others. Ultimately it was difficult to abstract a convincing argument as to what effect the gene knockdowns have on the type-I and LPS response, based on QPCR data for the two genes (*Ifnb1* and *Oasl1*) alone. Therefore to further elucidate if and how the GOI knockdowns had influenced the macrophage transcriptome in the context of the LPS response, genome wide expression was assayed by microarrays in selected samples from screen-b.

4.2   Genome-wide microarray analysis of the LPS response following targeted knockdown of genes of interest

Eight siRNA treatments (from screen-b) were chosen for follow up analysis by microarrays; NT-siRNA (would serve as a control); Nfkb2 and Irf5 were chosen based on their activity in previous screens [146]. Soc3, Bcl3 and Rela were selected given their activity in the QPCR screens was generally more consistent than the other targeted genes; namely cells treated with siRNA targeting Socs3 or Bcl3 generally expressed lower levels of type-I genes (*Ifnb1* and *Oasl1*), whereas those targeting Rela expressed *Ifnb1* to a greater extent than many of the other genes targeted following LPS stimulation in both screen-a and screen-b. Rela along with Nfkb1 are known to bind to the IFN-β enhancesome so logically it might be expected that their knock-down should repress type-I induction. Relb was also chosen for further analysis.

In addition to the siRNA targeted samples, a timecourse of (NT)-siRNA treatment of mouse BMDMs was also examined alongside the data. The samples examined by microarray analysis are summarised in Table 5.2.

Stimulus	Concentration	Pre-treatment?	Time-points (hours)
None	N/A	No	2h post experiment start
LPS	5 ng/ml	No	2h post LPS
LPS	5 ng/ml	24 hours of 20 nM **Non-targeting** siRNA	2h post LPS
LPS	5 ng/ml	24 hours of 20 nM siRNA targeting **Nfkb2**	2h post LPS
LPS	5 ng/ml	24 hours of 20 nM siRNA targeting **Irf5**	2h post LPS
LPS	5 ng/ml	24 hours of 20 nM siRNA targeting **Socs3**	2h post LPS
LPS	5 ng/ml	24 hours of 20 nM siRNA targeting **Relb**	2h post LPS
LPS	5 ng/ml	24 hours of 20 nM siRNA targeting **Bcl3**	2h post LPS
LPS	5 ng/ml	24 hours of 20 nM siRNA targeting **Rela**	2h post LPS
LPS	5 ng/ml	24 hours of 20 nM siRNA targeting **Nfkb1**	2h post LPS
None	N/A	No	7h post experiment start
LPS	5 ng/ml	No	7h post LPS
LPS	5 ng/ml	24 hours of 20 nM **Non-targeting** siRNA	7h post LPS
LPS	5 ng/ml	24 hours of 20 nM siRNA targeting **Nfkb2**	7h post LPS
LPS	5 ng/ml	24 hours of 20 nM siRNA targeting **Irf5**	7h post LPS
LPS	5 ng/ml	24 hours of 20 nM siRNA targeting **Socs3**	7h post LPS
LPS	5 ng/ml	24 hours of 20 nM siRNA targeting **Relb**	7h post LPS
LPS	5 ng/ml	24 hours of 20 nM siRNA targeting **Bcl3**	7h post LPS
LPS	5 ng/ml	24 hours of 20 nM siRNA targeting **Rela**	7h post LPS
LPS	5 ng/ml	24 hours of 20 nM siRNA targeting **Nfkb1**	7h post LPS
NT-siRNA	20 nM	No	0 h (pre-treatment) *
NT-siRNA	20 nM	No	1 h post transfection *
NT-siRNA	20 nM	No	2 h post transfection *
NT-siRNA	20 nM	No	4 h post transfection *
NT-siRNA	20 nM	No	8 h post transfection *
NT-siRNA	20 nM	No	24 h post transfection *

Table 5.2: Description of macrophage samples chosen for follow up analysis by genome wide expression profiling using microarrays. Mouse BMDMs were treated with 5 ng/ml LPS for either 2 or 7 hours following 24 hours of prior targeted siRNA treatment. In a separate experiment a time course of NT-siRNA treatment was performed in mouse BMDMs; with samples taken at 1, 2, 4, 8, 24 hours post-treatment or pre-treatment (0 hours). Two biological replicates were run for samples marked with an*

32 RNA samples were processed for hybridisation to Affymetrix Mouse Gene 1.1. ST Arrays. These arrays were obtained as part of a 96-Array plate format, which enables the simultaneous high-throughput profiling of 96 samples, using the same content as the individual Mouse Gene 1.1 ST cartridge arrays. The remaining 64 arrays on the 96-Array plate were used to process samples from experiments exploring other questions of interest to this thesis and are discussed accordingly in other Chapters (6).

A network graph of the normalised data for the 32 microarrays pertaining to the investigations into targeted siRNA knockdown in mouse BMDMs was generated in BioLayout *Express*[3D], by filtering for nodes with relationships across the arrays at a Pearson correlation of 0.85 or above. This resulted in a graph of 12,619 nodes, connected by 505,590 edges, which was was then clustered at an MCL inflation value of 2.2. After filtering to remove un-interesting clusters and those representing technical artefacts, the remaining graph comprised 4,674 nodes connected by 433,761 edges (Figure 5.25). There were four main structures evident in the filtered graph; one representing up-regulated components, another, down-regulated components, and the remaining two were specific to changes during the NT-siRNA time-course.

To determine whether the targeted siRNA treatment was effective at suppressing GOI expression following LPS challenge, GOI expression in the knockdown samples was compared to the untreated controls, LPS-only treated and LPS treated with prior-NT-siRNA (Figure 5.26). All seven GOI's were induced by LPS treatment alone (at 2 and 7 h) and LPS treatment in combination with 24 h prior NT-siRNA. Relative to their expression in the LPS-with-prior-NT-siRNA samples, *Nfkb2*, *Socs3*, and *Relb* were all knocked down when treated with their specific targeting siRNA. Nfkb1 and Bcl3 were knocked down to their basal (control-untreated) levels of expression, and Rela and Irf5 were further knocked down beyond their basal levels of expression. Therefore at the message-level all seven GOIs were convincingly repressed when targeted with the corresponding siRNA.
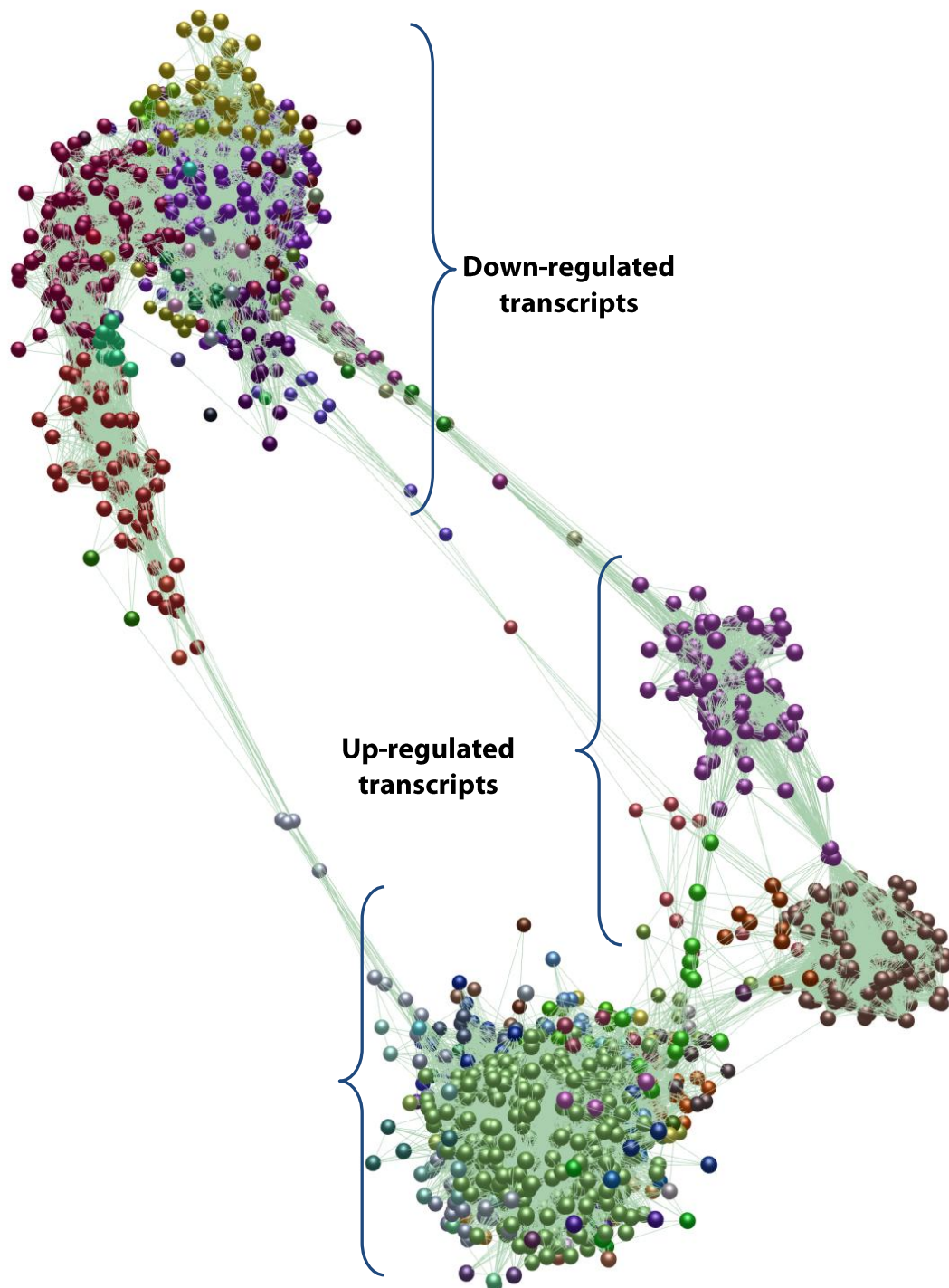
Figure 5.25: A network graph of the transcriptional changes occurring in response to siRNA treatment of mouse BMDMs A network graph of the normalised data for the 32 microarrays pertaining to the investigations into targeted siRNA knockdown in mouse BMDMs was generated in BioLayout Express3D, by filtering for nodes with relationships across the arrays at a Pearson Correlation of 0.85 or above. The resultant graph of 12,619 nodes connected by 505,590 edges was clustered at an MCL inflation value of 2.2 to identify groups of co-ordinately expressed transcripts. Un-interesting clusters and those representing technical artefacts were removed from the graph leaving 4,674 nodes connected by 433,761 edges (above). This filtered graph comprised three main components; transcripts up-regulated in response to LPS/ NT-siRNA treatment, transcripts down-regulated in response to LPS or NT-siRNA treatment and transcripts specifically regulated in the NT-siRNA timecourse.
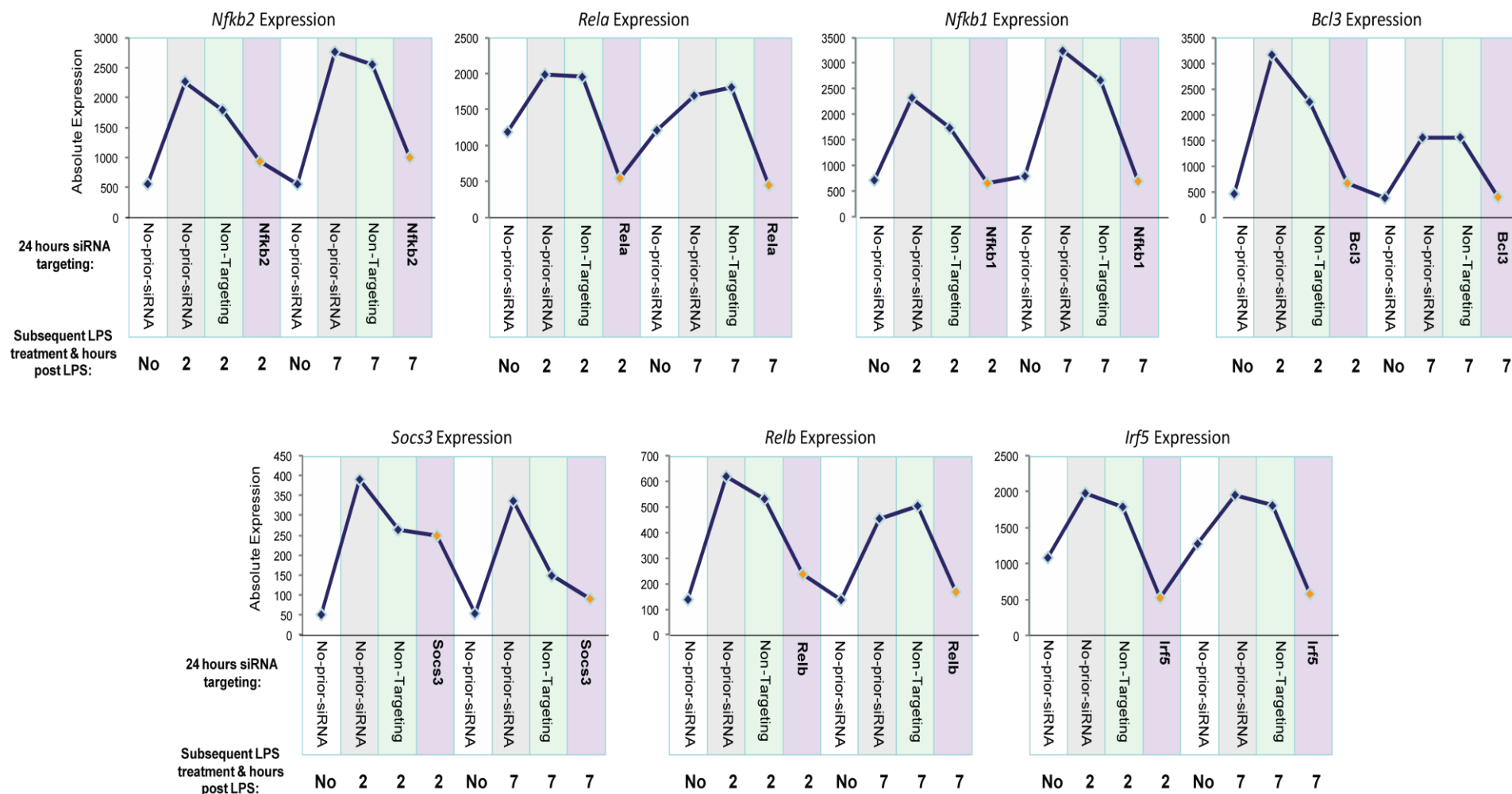
Figure 5.26: Transcriptional expression and repression of genes of interest in response to LPS stimulation of BMDMS with/without prior targeted siRNA treatment. Mouse BMDMs were treated with specific targeting siRNA or control NT-siRNA for 24 hours, and subsequently treated with 5ng/ml LPS for 2 and 7 hours. Expression levels of the targeted genes (as measured by microarrays) is shown above for seven genes of interest in cells with no-treatment, LPS-treated only, NT-siRNA treated with subsequent LPS, and those treated with the specific targeting siRNA followed by LPS.

To explore the main transcriptional signatures associated with the data set as well as establish whether any of the targeted gene KDs had perturbed this signalling, the expression patterns associated with the clusters in the filtered network graph were examined. Despite clear indication of target gene knockdown (Figure 5.26), there was no evidence of transcriptional signatures associated with any of the specific knock-downs; instead the majority of clusters were an outcome of LPS stimulation of macrophages and unaffected by the suppressed expression of any of the GOIs. The average expression of co-ordinately expressed transcripts was plotted across the siRNA-KD data set (Figure 5.27) for the major clusters within the data set which are also summarised in Table 5.3.

Cluster-ID	No of Transcripts	Expression Pattern	Go Terms/ associated processes
Cluster-a	1,248	Repressed 7 h post LPS treatment. Repressed at 4 h in NT-siRNA time-course	Cell-cycle progression (e.g Cdc6, Cdca2/3/4/5/7/8), cyclins, (e.g Ccna2, Ccnb1, Ccnb2, Ccnd1, Ccne1/2), DNA polymerase subunits (Pola1, Pola2, Pole, Pole2).
Cluster-b	1,029	Up-regulated 7 h post-LPS, and at 8 h in the NT-siRNA time-course	Regulation of immune signalling, leukocyte activation and cytokine production. (e.g Cxcl3/9/10/11, Ifit1/2/3), Il6, Il7, Il12a, Il15, Oasl1, Oasl1a, Oasl2, Oas2, Tlr1, Tlr3, Tlr6, Tlr8, Myd88)
Cluster-c	351	Represses 7 h post LPS treatment. Repressed 8 h in the NT-siRNA time-course	Membrane organisation and endocytosis
Cluster-d	223	Upregulated 2 and 7 h post LPS and from 4-8 hours in the NT-siRNA timecourse.	Cytokine production and regulation (e.g. Ccl2, Cxcl1, Cxcl2, Cxcl16, Irf1, Il1a). Apoptosis signalling (Cflar, Fas, Tnf).
Cluster-e	138	Upregulated 8h in NT-siRNA timecouse only.	Response to virus exposure, regulation of kinase activity.
Cluster-f	100	Upregulated 2 h post-LPS and 4-8 hours in NT-siRNA timecourse	Signalling cascades, and lymphocyte activation.

Table 5.3: Overview of the main clusters of co-ordinatley expressed genes in the screen of siRNA knockdown effects on macrophage signalling.

If knockdown of certain genes specifically perturb type-I signalling it was possible this effect was obscured if the ISGs are members of clusters with transcripts whose expression was not affected by the gene KDs. Therefore to further verify the GOI-knock-downs had not had an effect, the expression of genes regulated in response to IFN-β stimulation of macrophages, (as determined in Chapter-4) was examined in this array study. 1,639 of the IFN-β study genes mapped back to the data-set here. Cohorts of genes which were regulated across different temporal phases of the IFN-β response (as determined in Chapter-4) were identified and their average expression in this array study was plotted (Figure 5.28). However the expression of these IFN-β response genes was consistent across all of the LPS treated samples, regardless of any prior targeted gene KD. The data would therefore suggest on this occasion the gene knockdowns have not perturbed transcriptional networks associated with type-I and LPS signalling.
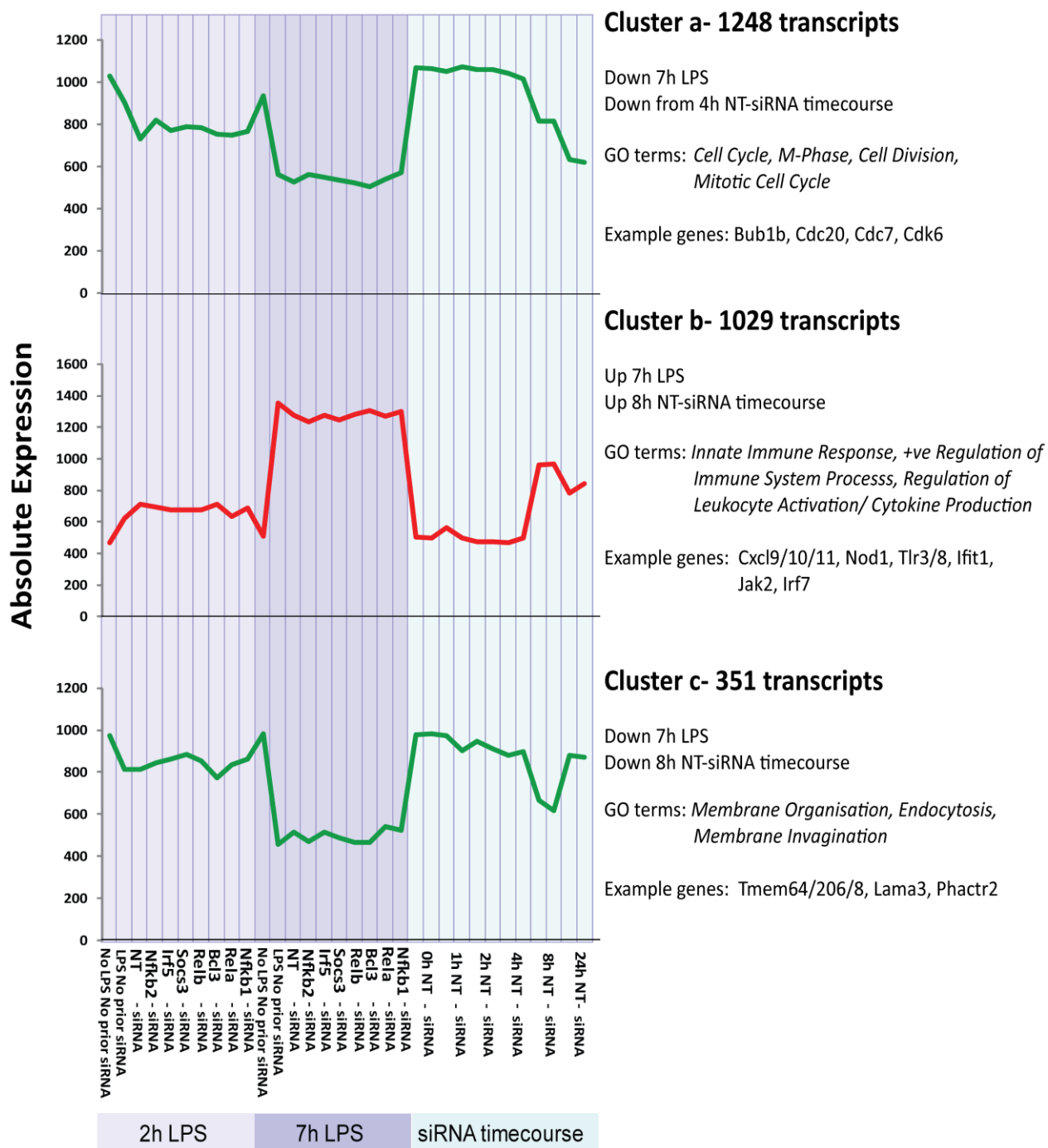
Figure 5.27: Average expression profile of cohorts co-ordinately expressed transcripts (clusters), generated from a network analysis of transcriptional changes in BMDM stimulated with LPS or NT-siRNA. BMDMs were transfected with 20 nM siRNA using lipofectamine 2000 (L2K), targeting any of Nfkb2, Irf5, Socs3, Relb, Bcl3, Rela, Nfkb1 or Non-targeting-siRNA (control). 24 hours post siRNA treatment cells were treated with 5ng/ml LPS, and were then harvested at both 2 and 7 hours post treatment for microarray-expression analysis on RNA extracts. In a separate experiment a timecourse of NT-siRNA treatment of macrophages was performed. Network based analysis was of the entire data-set was performed and clusters of co-ordinately expressed transcripts identified. Six examples of clusters representing changes associated with LPS and/NTsiRNA treatment are shown. Data points are not continuous, but represented as such to ease interpretation.

Figure 5.27 (continued): Average expression profile of cohorts co-ordinately expressed transcripts (clusters), generated from a network analysis of transcriptional changes in BMDM stimulated with LPS or NT-siRNA. BMDMs were transfected with 20 nM siRNA using lipofectamine 2000 (L2K), targeting any of Nfkb2, Irf5, Socs3, Relb, Bcl3, Rela, Nfkb1 or Non-targeting-siRNA (control). 24 hours post siRNA treatment cells were treated with 5ng/ml LPS, and were then harvested at both 2 and 7 hours post treatment for microarray-expression analysis on RNA extracts. In a separate experiment a timecourse of NT-siRNA treatment of macrophages was performed. Network based analysis was of the entire data-set was performed and clusters of co-ordinately expressed transcripts identified. Six examples of clusters representing changes associated with LPS and/NTsiRNA treatment are shown. Data points are not continuous, but represented as such to ease interpretation.

Figure 5.28: Expression of IFN-β regulated transcripts (as determined in Chapter-4) in the siRNA-LPS-screen and NT-siNRA time-course study. The average expression of cohorts of transcripts regulated at specific temporal phases in response to IFN-β stimulation of macrophage was plotted based on their expression values in the current analysis described in Figure 5.27.
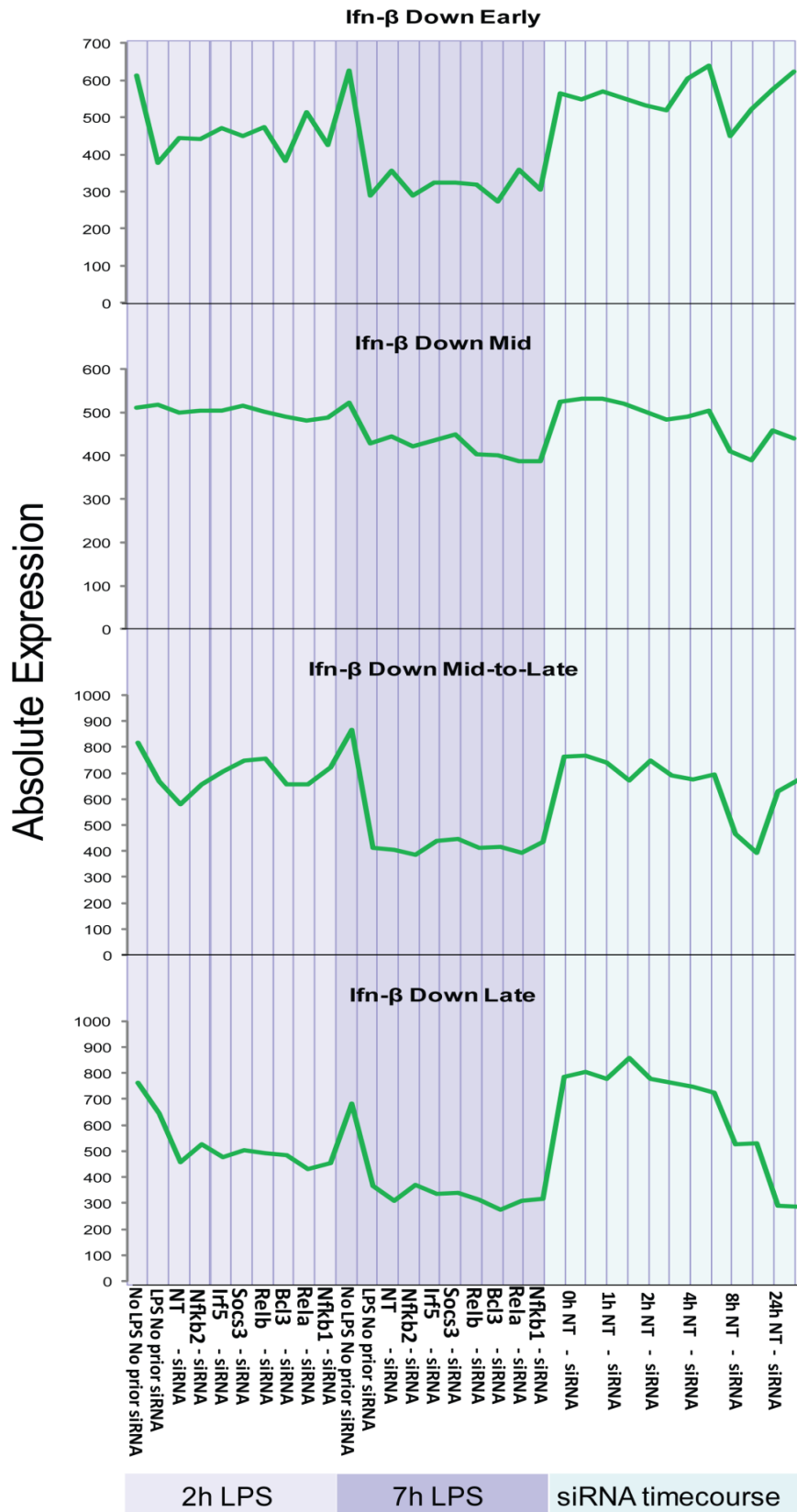
Figure 5.28: (continued) Expression of IFN-β regulated transcripts (as determined in Chapter-4) in the siRNA-LPS-screen and NT-siNRA time-course study. The average expression of cohorts of transcripts regulated at specific temporal phases in response to IFN-β stimulation of macrophage was plotted based on their expression values in the current analysis described in Figure 5.27.

Discussion

The investigations of this Chapter set out to optimise an *in vitro* assay to test the role of selected genes in the type-I response using targeted RNA-interference, QPRC and genome-wide expression profiling. Study of the type-I response using siRNAs is inherently complicated by the fact IFN-β and type-I response signalling is induced by the transfection of the siRNA itself. To contend with this, as well as better understand the macrophage response to stimulation with siRNA and/or LPS, a host of fundamental questions pertaining to the use of siRNA and LPS in mouse macrophages were examined, including;

- The dose dependant effects of siRNA.

- The dose dependant effects of LPS alone and in combination with prior siRNA treatment.

- The timing of siRNA treatment to achieve efficient gene knockdown as well the most efficacious downstream response to knockdown.

This optimisation process highlighted the issues surrounding the use of siRNA as a study tool in macrophages, and helped inform the design of an assay to study the roles of the selected GOI in the type-I and LPS response. The assay incorporated 5 ng/ml of LPS treatment in mouse BMDMs which had been transfected with siRNA targeting the selected GOI 24 h earlier. Network analysis of genome-wide transcriptional data revealed there were no patterns of expression specifically associated with any of the prior gene knockdowns. These results would suggest that the LPS response was not sensitive to knockdown of any of the GOI.

Factors explored during the assay optimisation process but not discussed as part of these results included the possibility of siRNA sequence specific effects on type-I responsiveness and the effect of using macrophages derived from different strains of mice (BALB/c and C57BL/6). Assay optimisation screens were at first performed on macrophages derived from C57BL/6 mice strain. The strain was later changed to the BALB/c since (i) the original study forming the basis of these investigations was performed in BALB/c derived BMDMs [146], (ii) other data generated in this thesis

(Chapters 4 and 6) is based on BALB/c derived BMDMs, (iii) C57BL/6 macrophages are more sensitive to DNA induced death and are hyper-responsive to LPS compared with BALB/c macrophages [283]. With respect to the latter point, during the course of these investigations it was found that C57BL/6 macrophages on some occasions but not others would be prone to siRNA-induced cell death (data not shown). In contrast viability of BALB/c macrophages following siRNA treatment was more consistent from week to week. BALB/c unlike C57BL/6, do not possess basal levels autocrine interferon signalling [283] this factor along with the greater sensitivity of C57BL/6 to nucleic acid induced cell death may have contributed to the observed differences. The effect of mice strain on the downstream response to siRNA could have been further explored. In essence this would form a study in itself, over and above the optimisation process given the vast number of variables that could be investigated.

Also explored but not discussed in the context of these results was the presence and variability in possible immunostimulatory sequences across the panel of siRNAs used in these investigations. It is possible the differences in the abundance of these sequences across the different siRNA's may account for variability in their ability to induce a type-I response. The sequences implicated in siRNA immunostimulatory capacity include; GUCCUUCAA [284], UGUGU [285], UGGC [85], and GU [286]. Others have also found a correlation between the uradine content of the U-rich strand and the immunostimulatory activity of the siRNA duplex as well as the number of CG/ GC clamps interspersed along the length of the siRNA [287]. Attributing the contribution of individual sequences to the immunostimulatory capacity of the siRNAs is inherently complex. Ultimately the genome-wide transcriptional analysis revealed that the expression of type-I response genes following LPS stimulation was uniform across all samples regardless of the target of the siRNA (i.e. sequence of the siRNA duplexes) (Figure 5.27).

Effect of increasing concentrations of siRNA in BMDMs

Whilst lipofection of 50 to 100 nM has been recommended for stable knockdown in RAW cells [288], the data here indicated these concentrations were toxic to primary

macrophages and distorted their phenotype considerably from the control-untreated state, and more so when subsequently treated with LPS (Figures 5.9).

Uptake of fluorescently labelled siRNA (siGLO) was observed as early as 4 h post-transfection at both 20 and 50 nM concentrations and by 24 h post-transfection the siGLO staining was indicative of almost all cells having been successfully transfected (Figures 5.3-5.5). siGLO is intended as a qualitative indicator of siRNA transfection efficiency and in this respect it was apparent the cells were being transfected. Counterstaining with a cytoskeletal dye (e.g. actin) may have provided a more accurate indication of uptake of siRNA and the dynamics of uptake of siGLO (for example numbers of transfected vs. non-transfected cells).

Efficient transfection does not necessarily correspond with efficient gene knock-down, in every cell type, as a cells proliferative rate [289] and capacity for degrading (exogenously introduced) nucleic acid may influence the potency of the siRNA KD. Hence also examined was the level of target gene repression by siRNA at a range of final concentrations. For the two genes tested (Irf7 and Nfkb2), their message level of expression was suppressed to the same extent regardless of the concentration of targeting siRNA used (Figure 5.6). However expression of *Irf7* and *Nfkb2* was induced in an exponential manner by non-targeting siRNA, as was expression of *Oasl1* and *Ifit1* (type-I response genes). Interestingly the dynamics of *Oasl1* and *Ifit1* expression with increasing concentrations of siRNA varied across the three siRNAs tested (non-targeting, Irf7 or Nfkb2-siRNA) (Figure 5.8). The results here indicate increasing the concentration of the siRNA (targeting key IFN-β signalling genes) does not necessarily correlate with enhanced repression of the type-I response. Ultimately there are likely a number of factors that influence the type-I induction for a given siRNA at a given concentration. These could be the kinetics of target gene repression, as well as that of protein knockdown, half-life and turnover.

Others have also observed a dose responsive increase in *Oasl1* and *Ifit1* expression when transfecting RCC1 renal cell carcinoma cells with 10, 20, 50 and 100 nM of siRNA

targeting GAPDH. The study suggests that the IFN system is activated in response to all concentrations of siRNA tested, however some ISGs are only induced in a concentration-dependent manner [77]. The results here (Figure 5.8) suggested that although type-I responsiveness increases in a dose dependant manner to NT-siRNA, this does not necessarily hold true for other siRNAs, possibly given their known or potential role in interferon signalling.

Examination of cell morphology suggested the phenotype of the LPS stimulated macrophages pre-treated with 50 nM to 100 nM siRNA are considerably distorted from those stimulated with LPS alone or LPS in combination with lower doses of siRNA (Figure 5.9). Further examination, for example an annexin-V apoptosis assay would provide a better indication of the extent of apoptosis the macrophages were undergoing.

Ultimately a balance had to be achieved between using enough siRNA to achieve knockdown at the transcriptional and protein level, yet not causing toxicity and overriding the effect of any gene knockdown due to high levels of siRNA-dose dependant type-I expression. 20 nM siRNA (final) concentration was chosen for use in the screens in BMDMs for the remainder of the investigations. Knock-down of targets genes was achieved at both the message and protein level with this concentration of siRNA.

Macrophage response to LPS and combining siRNA transfection with LPS treatment

Type-I expression in response to a range of LPS doses was examined. Essentially a balance had to be drawn between using enough LPS to efficiently induce detectable levels of *Ifnb1* and type-I expression yet avoiding causing toxicity to the cells as observed previously (due to the combinatorial use siRNA and LPS) (Figure 5.9).

Initial investigations comparing the potency of type-I induction in response to IFN-β and LPS indicated a 12-fold difference between the extent of *Oasl1* induction in

response to 10 U/ml IFN-β or 100ng/ml LPS (Figure 5.10). Given the extent of difference in type-I response activation by 100 ng/ml LPS and 10 U/ml IFN-β, the following investigations set out to determine how the type-I response varied in response to different concentrations of IFN-β and LPS. Macrophages induced *Oasl1* expression in a dose responsive manner in response to increasing concentrations of IFN-β and increasing doses of LPS respectively, 24 h post-treatment (Figure 5.11). The results indicated the extent of *Oasl1* induction was far greater with doses of LPS used (10, 20, or 100 ng/ml), compared to doses of IFN-β examined (20, 50 or 100 U/ml) (Figure 5.11).

Given the potency of type-I induction at 100 ng/ml LPS, a lower range of LPS doses were examined (50, 20, 10, 5, and 2.5 ng/ml). A dose dependant induction of *Ifnb1* expression was observed following 1 and 2 h of LPS treatment (alone) (Figure 5.12). By 7 h post-LPS treatment type-I response gene induction was evident at all doses of LPS examined (as measured using semi-quantitative PCR) (Figure 5.13). Attempts were also made here to define the overlap in the transcriptional targets of LPS and IFN-β, however this analysis was very limited by the number of genes which were analysed by semi-quantitative PCR (Figure 5.13). Amongst the genes tested the results did demonstrate *Il1b* expression (and to a lesser extent *Tnf* expression) was exclusive to LPS treatment. There was also a time-lag between the induction of the interferon inducible genes in the LPS treated samples, possibly due to time taken for autocrine IFN-β signalling to initiate in LPS treated cells. Ultimately a better understanding of the overlap and differences in the LPS and IFN-β, may have aided the assay design.

Although the investigations of this Chapter did not set out to specifically explore macrophage responsiveness to LPS following siRNA-lipofection, a better understanding of the cells response under these circumstances would have been valuable given the design of the assay. During the optimisation process, it was found that treatment of BMDMs with 20 ng/ml LPS after 24 h pre-exposure to siRNA, resulted in cell toxicity (data not shown). However at lower doses of LPS treatment, the cells were viable and in some experiments (Figure 5.14) generated higher levels of *Ifnb1* in cells pre-treated

with NT-siRNA, compared to those treated with LPS alone. On other occasions *Ifnb1* expression was higher in cells which had not undergone prior siRNA transfection (Figure 5.16). Eventually a dose of 5ng/ml LPS was chosen for the assay, since detectable levels of *Ifnb1* expression were induced at this concentration and cell viability did not appear compromised when 5 ng/ml LPS was used in combination with 24 h prior siRNA. In a preliminary screen, *Ifnb1* expression was lower at all measured time-points following 5 ng/ml LPS treatment in cells which had been pre-treated with siRNA targeting either Irf3 or Nfkb2 expression compared to those pre-treated with NT-siRNA (Figure 5.16). This observation was fitting with the original hypothesis that siRNA mediated knock-down of genes known to positively regulate type-I signalling will repress the LPS induced type-I response. Of the two genes tested, Irf3 is known to form part of the *Ifnb1* enhancesome and Nfkb2 is suspected to play a role in type-I signalling regulation based on our previous observations [146].

The use of siRNA to investigate components of LPS signalling in primary and immortalised macrophages is not novel [290-291]. However the possibility that pre-treatment with siRNA may in-itself modulate responsiveness to LPS in immune cells has broadly been overlooked. Pre-exposure to immunological stimuli can make macrophages hyper-responsive or tolerant to subsequent LPS treatment. For example macrophages primed with IFN-γ display enhanced sensitivity to endotoxin, and *in vivo* IFN-γ pre-treatment followed by even subclinical endotoxin exposure is toxic [292-293]. Viral infection induced type-I interferon also heightens sensitivity to LPS challenge as measured by increased lethality in vivo and augmented TNF-α serum levels [294]. In contrast, pre-exposure to low doses of LPS results in tolerance to subsequent LPS stimulation as determined by down-regulation of cytokine production in vivo [295-296] and in primary macrophage cultures [297-300]. The mechanism of endotoxin tolerance has been linked to the MyD88-dependent signalling, up-regulation of negative regulators of the TLR pathway, and transcriptional re-programming whereby proinflammatory cytokine expression is suppressed and anti-inflammatory cytokines are over-expressed [301].

The collection of data here strongly suggests siRNA trasfection in macrophages results in their activation by IFN-β stimulation, therefore the cells are type-I interferon primed prior to any subsequent LPS treatment. In RAW264 macrophages, the sustained activation response to different TLR agonists has been found to be perfectly additive [302]; so that the response to activation of TLR4 by LPS was additive with that of TLR2/6 (but not with TLR4). Similarly LPS stimulation of macrophages which have been exposed to a lipid-based transfection reagent and siRNA, is likely to result in the activation of different TLR's, the consequence of which is unclear.

Determination of ideal siRNA treatment time for optimal KD and type-I repression following LPS stimulation

Given knockdown at the protein level is apparent later than that at gene level there could be variability in the effect of GOI knockdown in the LPS response as monitored after different times of treatment with siRNA. For example one possibility is that by 48 h post transfection the siRNA may have had more time to take effect at the protein level. Thus if the targeted GOI's activity at the protein level is key to the LPS or type-I response then this could be better observed after 48 h following siRNA transfection. Conversely since the effect of siRNA is transient, it is possible the potency of the effects of knocking down a GOI may wear off over time following siRNA transfection.

Protein level analysis of Nfkb2 expression indicated the peak induction of Nfkb2 by NT-siRNA is achieved at 24 h post treatment (Figure 5.17a). Nfkb2 protein expression was repressed from 24 h post treatment in samples treated with siRNA targeting the Nfkb2 gene, and to a greater extent by 48 h. At the message level, Nfkb2 expression was knocked-down to a similar extent at both 24 and 48 h post siRNA transfection (Figure 5.18). Furthermore type-I response gene expression following LPS challenge presented a similar pattern of expression in cells pre-treated with siRNA for either 24 or 48 h (Figure 5.19). Thus the siRNA treatment time of 24 h was maintained for the remainder of the investigations. Knockdown analysis of GOI in the genome-wide microarray

screen revealed effective message level knockdown of the targets at both 2 and 7 h post LPS, when cells had been treated with 24 h prior siRNA (Figure 5.26).

Studies of targeted gene knockdown in mouse BMDMs

The purpose of the siRNA optimisation process was to aid the design an *in vitro* assay to study the role of GOI in the type-I response. LPS stimulation of macrophages induces the potent expression of IFN-β message and protein [275]. Consequently type-I signalling forms a significant portion of the response to LPS [277, 303-304]. Therefore to measure the effect of gene knock-downs LPS was chosen as a stimulus of type-I signalling.

BMDMs were transfected with siRNA targeting specific GOI, and 24 h later were treated with 5ng/ml LPS. Analysis of *Ifnb1* expression demonstrated a great deal of variability in the level of *Ifnb1* induction (following both lipid-transfection of the siRNA and the subsequent LPS stimulation), across the different knockdown samples. The magnitude of *Ifnb1* induction also varied between the two screens performed. Furthermore reproducibility within and between screens was for many of the genes tested inconsistent, with respect to *Ifnb1* expression (Figures 5.22 & 5.24). The siRNA targets which were more consistent in their action between the two screens included Socs3, Bcl3, and Rela. The discrepancies witnessed during these screens, and to some extent throughout the optimization process could be attributed to the innate variability of primary macrophage cell cultures. The variability in *Ifnb1* expression observed between the two screens could partly be attributed to the very transient expression of *Ifnb1,* the peak induction of which may have been induced at slightly different times between the two screens and also between target-genes in a given screen. QPCR, semi-quantitative PCR and microarray data all suggest *Ifnb1* is maximally expressed ~6-8 hours post-siRNA transfection (Figures 5.15, 5.16). Microarray analysis of a timecourse of NT-siRNA treatment in mouse BMDMs revealed the predominant transcriptional activity occurs at 8 h and later following transfection, with negligible changes at the earlier time-points analysed (1, 2 and 4 h) (Figure 5.27). These results underscore the importance of sample timing when trying to deduce levels of *Ifnb1*

expression in response to siRNA transfection. Moreover reproducibility of results in primary macrophage cultures is often unwieldy, as epitomised by the findings of the RNAi screen in BMDMs [146] where three biological replicates would yield variable outputs in terms of gene expression. To garner a more convincing view of the effects of the gene knockdowns on the IFN-β signaling, genome-wide analysis of the downstream response was performed.

Efficient gene knockdown of the selected GOIs was observed relative to NT-siRNA (Figure 5.26). However the network based analysis of the results showed that the gene KDs did not perturb transcriptional signalling networks associated with the LPS and type-I interferon response (Figures 5.27, 5.28). One possibility is that some of the tested genes might be dispensable for eliciting the type-I response. Some of the GOIs have been shown elsewhere to play a role in type-I signalling. Interestingly in a recently published study Krausgruber et al, demonstrated high IRF5 expression in M1-type human and mouse macrophages and argues that IRF5 directly activates the transcription of certain M1-characteristic imflammatory mediators [272]. In contrast to the experiments here, the authors differentiated the mouse BMDMs with GM-CSF rather than M-CSF (CSF-1) and found cells differentiated with GM-CSF had higher expression of IRF5 protein than M-CSF-derived BMDMs. Furthermore in human macrophages inhibition of endogenous IRF5 via RNA-interference resulted in lower mRNA expression for a number of LPS induced inflammatory mediators [272].

Knockdown of members of the classical/canonical NfκB transcription factor, Nfkb1 and Rela also failed to distort type-I or LPS signalling based on the microarray analysis in this Chapter. Nfkb1:Rela are established to bind to the *Ifnb1* enhancesome, but some have reported these proteins to be largely dispensable for the RIG-like-receptor triggered IFN-β induction [305-306]. Others found that Rela controls autocrine IFN-β and basal ISG expression [307] but is only required for a small subset of inducible ISGs . Despite this the authors argue Rela is critical in the interferon response as the absence of Rela results in the delayed induction of IFN-β and subsequent ISG expression as well as increased susceptibility to viral infection [307]. It could therefore be possible that

although the knockdown of the GOI tested in the screen were not critical enough to abrogate IFN-β and ISG expression they may have influenced the timing of their onset; a factor not explored in these investigations.

The screens in this Chapter suggest the LPS response is not sensitive to changes in expression of any of the GOIs. Such resilience to perturbation has been described as 'robustness' in biological networks [308], a property allowing a given system to maintain its function despite internal and/or external assaults. IFN-β signalling is one of several pathways activated in response to LPS treatment of macrophages; the cells are also stimulated by the autocrine production of TNF-α, TGF-β, a number of interleukins as well as the differentiation factors CSF-1 and CSF-2 (reviewed in [309]). Activation of the downstream MAPKinase, PKC (protein kinase C), and heterotrimeric G-protein pathways are all recognised in response to LPS stimulation [309]. It is therefore plausible that activation of these various pathways contributed to the resistance of the overall transcriptional network to the gene knockdowns. A better understanding of the overlap and differences in the transcriptional response of macrophages to LPS and IFN-β is required to further elucidate why the gene knockdowns were ineffective at perturbing the transcriptional signatures associated with LPS and type-I signalling.

Conclusions and Further Work

The work described in this Chapter has highlighted the issue that the optimisation and use of siRNA in macrophages is complex, requires many considerations, and is often confounded given the highly attuned pattern recognition machinery of these cells. At the same time exploring the use siRNA in macrophages remains imperative given its power as a tool for functional genomics screening and also since macrophages are attractive targets for siRNA-based therapeutics.

One of the predominant issues during these investigations was the lack of reproducibility between experiments, which most likely could be attributed to the

plasticity and adaptability of primary macrophages. Arguably reproducibility may have been improved if the optimization process was performed in a macrophage-like cell lines (e.g. RAW macrophage). However the phenotype of immortalized macrophage-like myeloid cell lines is often very different from primary tissue-derived macrophages. Thus to understand the immunobiology of macrophages one needs to analyse primary cells.

There are a range of factors which could be further explored to improve the assay design including; testing of other lipid-based transfection vectors to determine if there was a difference in the immunostimulatory capacity of different transfection reagents; the inclusion of cell toxicity assays (for example annexin V apoptosis assays) would have given a more accurate indication of the levels of apoptosis induced by different variables (e.g. siRNA concentration; siRNA combination with LPS treatment; target gene of the siRNA). In the assays described in this Chapter, IFN-β expression was measured at the message level by QPCR, however mRNA expression does not always correspond with protein expression therefore it would have been desirable to also measure IFN-β expression by ELISA's.

One possibility why the genome wide transcriptional profiling following the siRNA screens did not reveal a difference if type-I responsiveness could be because the LPS stimulation compensated for the lack of individual signalling components (siRNA target genes). Therefore a better understanding and comparison of the LPS and IFN-β induced transcriptional responses may have better informed the assay design.

Methods and Materials

Cell culture

Bone marrow derived macrophages were prepared from the femurs of 7-10 week old male BALB/c or C57BL/6, in the presence of CSF-1. Full cell culture technique is as described in Chapter-4 Methods and Materials – (IFN-β study). For Western Blot analysis cells were plated out in 6-well tissue culture plates on day six of differentiation at a density of 1,000,000 cells / well. For all other experiments cells were seeded in 24-well tissue culture plates, on day six of differentiation at a density of 200,000-210,000 cells/ well, which is equivalent to the seeding density (cell/ over given surface area) of 6-well plates.

LPS treatment

LPS from Salmonella Minnesota (Sigma, Poole, UK) was used at a range of concentrations as indicated throughout the text of the Results section in cell culture experiments, and diluted in culture medium.

siRNA transfections

siRNAs (SMARTpools, Thermo Scientific, MA, USA) were purchased at a 5 nmol scale and redissolved in 1× siRNA buffer (Thermo Fisher Inc, MA, USA) to a final concentration of 10 µM. Each siRNA pool comprises a mixture of four siRNA duplexes targeting different sequences of the same mRNA, combined into one reagent. The SMARTpool siRNAs are selected (by the manufacturers) using a weighted algorithm which incorporates a number of criteria (including sequence specific and thermodynamic parameters) and is thought to improve identification of potent, functional siRNAs [310]. The *ON-TARGET* plus design of these reagents incorporates modifications to both the sense and anti-sense strands of the siRNA duplexes to destabilise off-target activity and enhance target specificity [311-312].

Details of the siRNA used in this Chapter are summarised in Table 5.4

Murine Target Gene	Thermo Scientific SMARTpools ID
Rela	ON-TARGETplus SMARTpool L-040776-00-0005 NM_009045
Nfkb1	ON-TARGETplus SMARTpool L-047764-00-0005 NM_008689
Rel	ON-TARGETplus SMARTpool L-047122-00-0005 NM_009044
Socs3	ON-TARGETplus SMARTpool L-040626-01-0005 NM_007707
Irf5	ON-TARGETplus SMARTpool L-041093-01-0005 NM_012057
Sod2	ON-TARGETplus SMARTpool L-062893-00-0005 NM_013671
Nfkb2	ON-TARGETplus SMARTpool L-046030-01-0005 NM_019408
Relb	ON-TARGETplus SMARTpool L-040784-01-0005 NM_009046
Bcl3	ON-TARGETplus SMARTpool L-045102-01-0005 NM_033601
Irf3	ON-TARGETplus SMARTpool L-041095-00-0005 NM_016849
Ifnb1	ON-TARGETplus SMARTpool L-043699-00-0005 NM_010510
Irf7	ON-TARGETplus SMARTpool L-041094-00-0005 NM_016850
Nfkbia	ON-TARGETplus SMARTpool L-044170-01-0005 NM_010907
Non-Targeting	ON-TARGETplus Non-targeting pool D-001810-10-05

Table 5.4: Details of siRNA SMARTpool purchased to target murine genes.

To transfect at a final concentration of 20 nM in a 6-well format the redissolved siRNA (10 µM) was diluted in nuclease free water to working concentration of 2 µM. For each treatment-well 10 µl of siRNA SMARTpool was combined with 90 µl of Optimem Reduced Serum Medium (Invitrogen, Paisley, UK) solution while 5 µl of Lipofectamine 2000 (Invitrogen, Paisley, UK) was mixed with 95 µl Optimem. Following incubation for 5 min, the siRNA mix was added to the L2K mix and incubated for a further 30 min, after which 800 µl of complete medium (+CSF-1) but lacking antibiotics was added to the siRNA:L2K complexes. The growth medium on the cells was removed and replaced with 1000 µl of the siRNA:L2K liposome medium. Cells were then incubated for a given period of time depending on the study (as denoted in the results section) at 37°C, 5% CO_2.

To transfect at a final concentration of 20 nM in a 24-well format the redissolved siRNA (10 µM) was diluted in nuclease free water to a working concentration 2 µM. For each treatment-well 5 µl of siRNA SMARTpool was combined with 45 µl of Optimem (Invitrogen, Paisley, UK) solution while 2.5 µl of Lipofectamine 2000 (L2K, Invitrogen, Paisley, UK) was mixed with 47.5 µl Optimem. Following incubation for 5 min, the siRNA mix was added to the L2K mix and incubated for a further 30 min, after which

400 µl of complete medium (+CSF-1) but lacking antibiotics was added to the siRNA:L2K complexes. The growth medium on the cells was removed and replaced with 500 µl of the siRNA:L2K liposome medium. Incubation conditions are as described above. Each biological sample was generated by pooling cells from three separate wells.

For siRNA experiments performed at final concentrations other 20nM, the working concentration of the siRNA pools was adjusted and all other conditions remained the same. For example to treat at a final concentration of 50nM the redissolved siRNA (10 µM) was diluted in nuclease free water to a working concentration 5 µM and the standard protocol followed as above.

RNA extraction, quantification and quality control

Procedures for RNA extraction, quantification, and QC are as described in Chapter-4 or Chapter-6 Methods. RNA integrity was screened across all (RNA-based) experiments described to ensure only high-quality RNA was used for QPCR and microarray procedures.

mRNA analysis by genome wide Microarray profiling

Full details of microarray processing are described in Chapter-6. The 32 RNA samples analysed in this Chapter were processed for hybridisation to Affymetrix Mouse Gene 1.1. ST Arrays. These Arrays were obtained as part of a 96-Array plate format, which enables the simultaneous high-throughput profiling of 96 samples, using the same content as the individual Mouse Gene 1.1 ST cartridge arrays. The remaining 64 arrays on the 96-Array plate were used to process samples from experiments exploring other questions of interest to this thesis and are discussed accordingly in other Chapters (6).

A network graph of the normalised data expression data was generated in BioLayout *Express*[3D], by filtering for nodes with relationships across the arrays at a Pearson Correlation of 0.85 or above. This resulted in a graph of 12619 nodes, connected by

505590 edges. The graph was then clustered at an MCL inflation value of 2.2 and the expression profiles of the clusters were analysed to: (1) identify and remove un-interesting clusters i.e. those representing technical artefacts or unchanged expression profiles, and (2) to identify clusters of interest i.e. those representing changes related treatment parameters. The filtered graph comprised 4674 nodes connected by 433761 edges (Figure 5.25).

mRNA analysis by QPCR

Methodology for QPCR has been described previously [146, 313]. The following Taqman Primer probe sets (below) were purchased from Applied Biosystems, Warrington, UK.

Gene	Assay ID
Ifnb1	Mm00439546_s1
Nfkb2	Mm00479807_m1
Irf5	Mm00496477_m1
Oasl1	Mm00455081_m1
Ccl5	Mm01302427_m1
Nfkb1	Mm00476379_m1
Rela	Mm00501346_m1
Gapdh	Mm03302249_g1
Relb	Mm00485664_m1
Bcl3	Mm00504306_m1
Irf7	Mm00516788_m1
Sod2	Mm00449726_m1
Ifit1	Mm00515153_m1
Irf3	Mm01203177_m1
Rel	Mm00485657_m1
Nfkbia	Mm00477798_m1
Socs3	Mm01249143_g1

Prior to proceeding with expression analysis, primers were tested on range of concentrations of murine BMDM RNA to determine PCR efficiency. Only an efficiency of 97% or > was accepted. An example, of primer efficiency test is shown below for the TaqMan primer designed for Gapdh detection.

Standard Curve
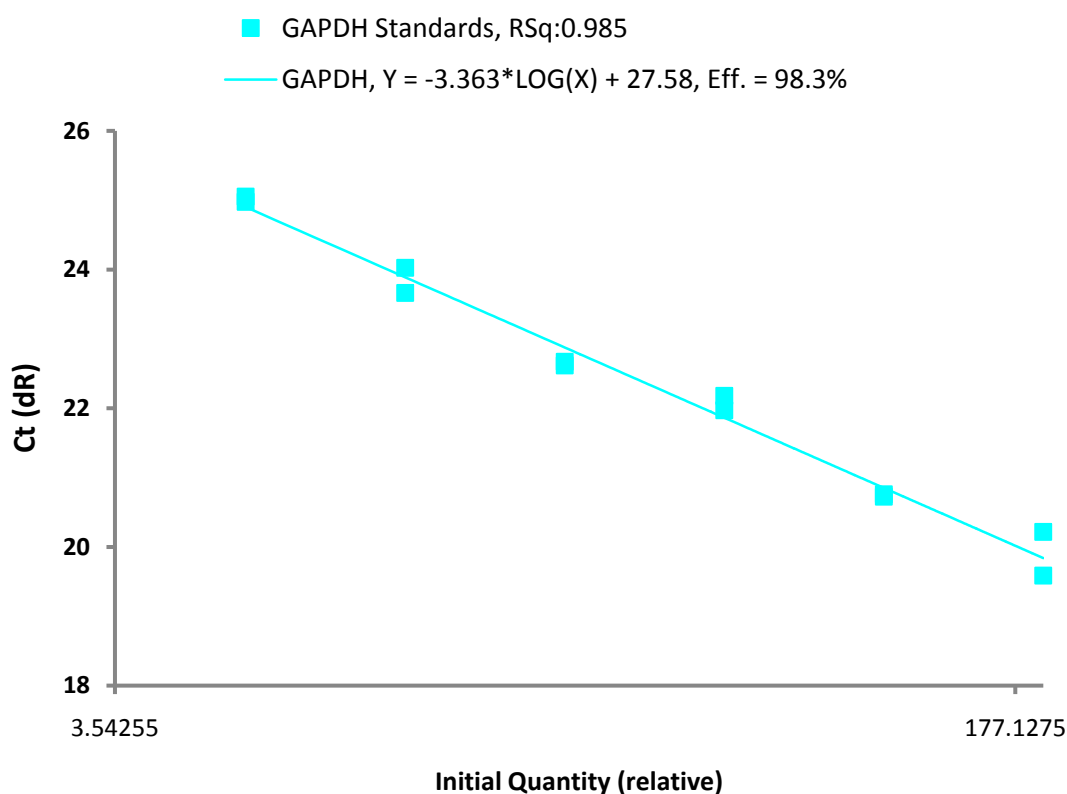
Figure 5.29: Standard Curve for the Taqman *Gapdh* Primer-probe. Ct values for Gapdh were determined at six different concentrations of input RNA (200, 100, 50, 25, 12.5, and 6.25 ng), and each input RNA was tested in duplicate. The RSq value (between 0 and 1) is an indicator of the quality of the fit of the standard curve to the Standard data points plotted. The closer the value is to 1, the better the fit of the line. The slope of the curve is directly related to the average amplification efficiency throughout the cycling reaction. PCR Efficiency (calculated from the slope of the curve) corresponds to the proportion of template molecules that are doubled every cycle.

Individual reactions (for each RNA sample against a given primer probe set) were performed in 20 µl volumes using MicroAmp Optical 96-well reaction plates and MicroAmp Optical Caps (Applied Biosystems). For each reaction 2 µl of RNA (≈100 ng total RNA) was added to 10 µl of 2× Brilliant® II QRT-PCR Master Mix, 1-Step (Agilent Technologies, Stockport, UK), 1 µl of a Taqman primer/probe set for the gene of interest at the recommended concentration, 6.9 µl of nuclease free double-distilled H$_2$0, and 0.1 RT Enzyme (Agilent Technologies, Stockport, UK). At least two technical replicates were included for each given reaction.

The reaction mixture were then incubated at the following conditions on the MXPRO3000P instrument (Stratagene, CA, USA): an initial step at 50°C for 30 min, followed by 95°C for 10 min after which the samples were then subject to 40 cycles under Taqman standard conditions of 95°C for 30 s and then a combined annealing and primer extension phase at 60°C for 1 min and a short denaturation at 72°C for 30 s. Stratagene MXPro software (Stratagene, CA, USA) was then used to analyze the data. Threshold determinations were automatically performed by the instrument for each reaction. Gapdh was used as a loading control, for normalisation purposes. Nomalised CT values were exported into Microsoft Excel and relative quantification of marker gene mRNA expression was calculated with the comparative CT method. Standard error is calculated for technical PCR replicates.

mRNA expression analysis by Semi-Quantitative RT-PCR

Primer Design

PCR primers of 20 bp in length with a 45-60 % GC content and TM of 57.3 – 61.4 °C were selected using Primer Designer (Scientific and Educational Software 3.0) to amplify products of 107-228 bp in length. TM was calculated by the primer manufacturer (Eurofins MWG Operon) using the following formula: TM [°C] = 69.3 + [41(nG + nC) / s − (650 / s)], where n = number of nucleosides of type X and s = number of all nucleosides per sequence. Primers were pre-screened to determine the optimal conditions for specific cDNA amplification on cDNA derived from mouse macrophage RNA. They were tested using a range of PCR cycles (between 25 and 30) at 55°C annealing temperature under standard assay conditions (see below), to ensure only a single band of the predicted size was generated. If the latter was not achieved then primers were redesigned.

Semi-Quantitative (Cresol Red and sucrose) RT-PCR

First strand cDNA was generated from 5 μg of RNA using random hexamers as primers in a final reaction volume of 30 μl. cDNA synthesis was performed in the presence and

absence of superscript reverse transcriptase (Invitrogen, Paisley, UK) and RNA. PCR amplification of cDNA equivalent of 10 ng of total RNA was carried out in 20 µl reactions containing the cDNA in a 4 µl volume and; 1x Reaction/PCR buffer (3.5 mM MgCl$_2$), 12.46% sucrose, 0.1 mM cresol red (Sigma, Gillingham, uk (#114480)), 12 mM beta-mercaptoethanol, 0.5 mM dNTPs (Invitrogen, Paisley UK (#10297018)), 0.6 U Taq DNA polymerase (Invitrogen, Paisley UK (#18038-026)), and primers were used at 100 ng/reaction. The inclusion of cresol red and sucrose in the reaction mixture allows direct loading of the product onto an agarose gel without addition of loading buffer. Amplifications were carried out on DNA Engine Tetrad PTC-225 Peltier Thermal cycler (Tetrad, MJ Research, US). This process involved an initial 2 min denaturing step (92 °C), after which each PCR cycle consisted of 30 sec denaturing (92 °C), 90 sec annealing (55 or 60 °C), and 60 sec elongation (72 °C). For each given primer pair-set an appropriate number of PCR cycles were chosen to allow termination of amplification in the linear phase of the PCR reaction. Otherwise all assays were conducted under identical conditions and only the number of cycles or annealing temperature varied. After the final PCR cycle, the reaction was held for 10 min at 72 °C. The PCR products were then separated on a 2.5% agarose gel, stained with SYBR safe DNA gel stain (Invitrogen, Paisley UK) in 1 X TBE and imaged on Syngene G:Box gel documentation system.

The primer sequences and optimized number of PCR cycles for the genes analyzed by semi-quantitative RT-PCR in this Chapter is shown below:

Gene	Primers (forward and reverse 5' -3')	Size of amplified fragment	Annealing temperature	Number of cycles for linear phase detection
Irf1	For-GAGGAACCAGAGATTGACAG Rev-AGCAGGCTGTCCATCCACAT	107	55°C	27
Irf5	For-GAGAAGAATGGCCTGATGTC Rev-GATGCTGTCTGCCGACCAAG	125	55°C	25
Tnf	For-GGACAGTGACCTGGACTGTG Rev-GAGGCAACCTGACCACTCTC	127	55°C	Less than 25
Il1b	For-GAAAGCTCTCCACCTCAATG Rev-GTATTGCTTGGGATCCACAC	193	55°C	25
Mx2	For-CTGGATTGTGATTCAGGGAC Rev-GCTAAATGGTGGGCAAGAAG	228	55°C	27
Socs1	For-TGGTTGTAGCAGCTTGTGTC Rev-AATGAAGCCAGAGACCCTCC	118	55°C	Less than 25
Oasl1	For-TGGCAGAAGGCTACAGATGG Rev-GCACGGTCACCTGGATATCG	138	55°C	25
Cd14	For-TACAGCTGCAAGGACTAGAC Rev-TCCAGCCTGTTGTAACTGAG	175	55°C	Less than 25

Table 5.5 Primer sequences used in semi-quantitative PCR analysis.

siGlo uptake analysis by confocal microscopy

Cell culture for the purpose of preparing slides for confocal imaging was performed as described in the cell culture section with the following exceptions: sterilized (in 70% EtOH), dry, round glass coverslips were placed into wells of 24-well tissue culture plates. On day 6 of macrophage differentiation, cells were harvested from the 10 cm square bacteriological plates and counted. 100 μL of cell suspension containing 50,000 or 163,000 cells (as stated in the text of the results sections) was added directly onto each glass coverslip and after 20 minutes (to allow the cells to adhere to the glass slides) 400 μL of culture medium (containing CSF-1) was added to the wells.

Transfection of fluorescent siRNA (siGLO) did not differ from the standard siRNA transfection protocol (See siRNA transfection). (siGLO siRNA is chemically labelled with a 6-FAM fluorophore, which is visible/ excited under the FITC channel). At appropriate times post siGLO-transfection (see results) the culture medium was removed from the glass slides of BMDM which were then fixed in 4% PFA for 5 min at room temperature

and rinsed with 0.5ml PBS (with 0.1%Triton X-100). The cells were then treated with a blocking solution containing 1% goat serum in PBS for 10 min. The glass coverslips were mounted using Vectorshield HardSet Mounting Medium with DAPI at 1.5 μg/ml (Vector Laboratories, Peterborough, UK) onto frost-free slides.

Fluorescence images were captured on a Nikon EC-1 confocal scanning laser microscope (Nikon Instruments, Surrey, UK). 2-D optical sections were acquired using Nikon EZ-C1 software (Nikon Instruments, Surrey, UK) with sequential acquisition (Frame Channel Mode) to give separate image files for each channel with minimal spectral overlap. The following stains and laser/filter combinations were used: DAPI nuclear stain (excitation 405 nm, emission BandPass460/50nm), 6-FAM fluorophore (siGLO), excitation 488nm, emission BandPass530/30nm).

Protein analysis by western blotting

BMDMs were washed with PBS and resuspended in a whole cell lysis buffer (50 mM HEPES, pH 7.5, 1% Triton X-100, 50 mM NaCl containing protease inhibitors (Roche Complete #11 697 498 001), phosphatase inhibitor cocktails I and II(Sigma P2850 and P5726)) at appropriate times (see results) post treatment or siRNA transfection. Cell lysates were centrifuged at 13000 rpm for ~ 1 min (at 4 °C) and the supernatants collected. Protein concentrations were determined by Lowry assays using the the Bio-Rad Dc Protein assay (Bio-Rad Laboratories) and the NanoDrop Spectrophotometer as per manufacturer's instructions. Western Blot analysis was performed by loading 10 μg of each protein sample onto a lane of a 10% polyacrylamide gel with a 4% stacking gel and separating proteins by SDS-PAGE. The proteins were then transferred to nitrocellulose membranes, and stained with Amido Black (Sigma, Gillingham, UK) to confirm equal loading of protein. Blots were then incubated in 1X TBS-triton for an hour at room temperature with gentle agitation followed by incubation with the primary antibody (anti-Nfkb2 (Cell Signalling Technology #4882) at 1:1000; or anti-Irf5 (Cell Signalling Technology #4950) at 1:100) in immunomedium (DMEM (Invitrogen, Paisley, UK)) overnight at 4 °C. The following day the membrane was washed three times for 10 min each in TBS-Triton, before being placed in secondary antibodies in

immunomedium (anti-rabbit HRP at 1:7000 (Sigma, Gillingham, UK) (A9169)), coupled to horse raddish peroxidase for 1-2 h at room temperature. Proteins were visualized using ECL reagents (Pierce, Rockford, IL) and Amersham Hyperfilm ECL (GE Healthcare)

Chapter Contributions and Acknowledgements

This page has been left intentionally blank

Chapter 6. Analysis of the transcriptional networks of mouse bone marrow derived macrophages in response to stimulation with a number of M1 phenotype activators
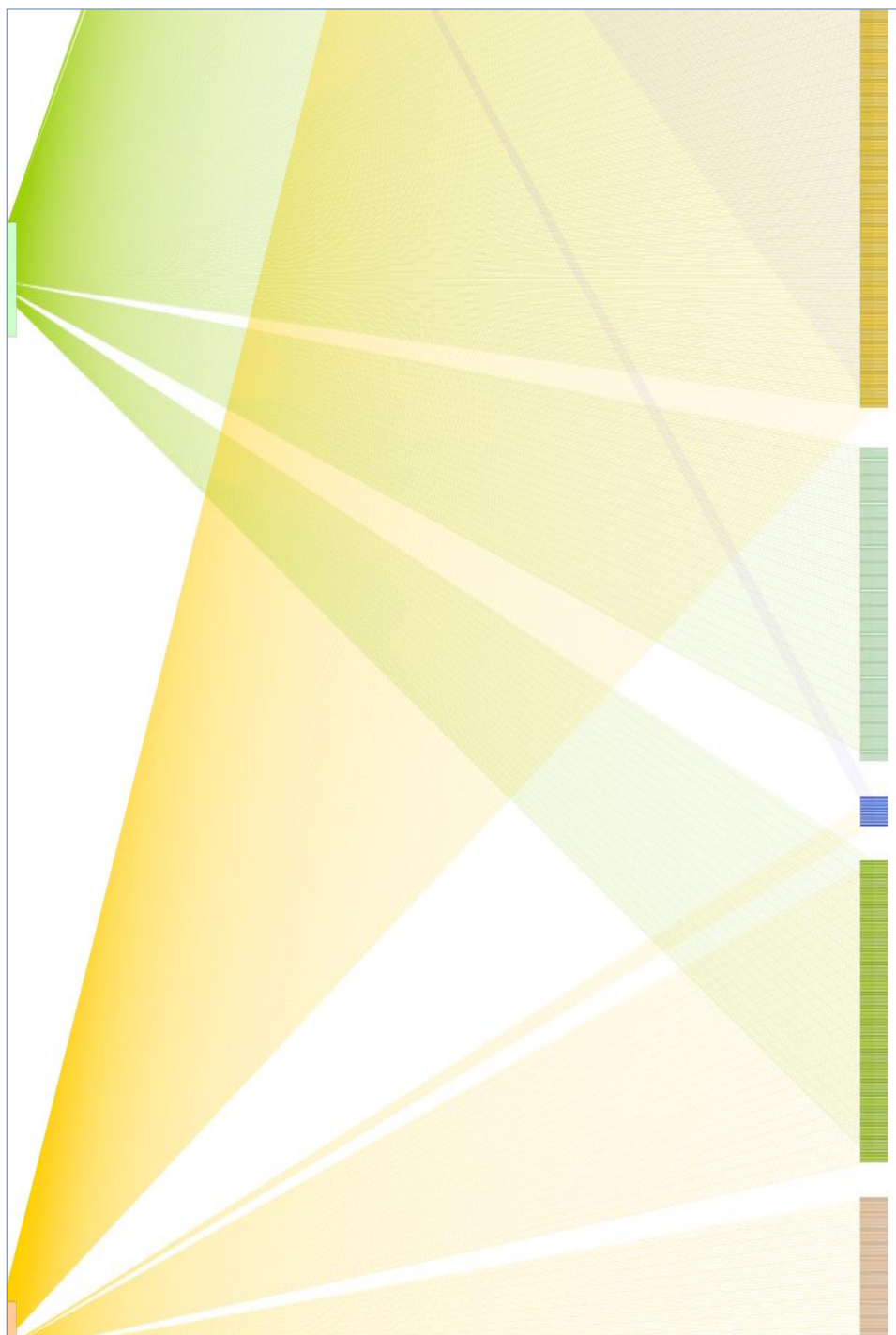
Introduction

The primary objectives of the work described in this Chapter were to examine the transcriptional signatures arising in response to three stimuli known to activate or contribute towards the M1 polarisation of macrophages; IFN-γ, IFN-β, and LPS. Previous attempts in this thesis to compare the type-I and type-II interferon response were confounded by the fact the two treatments (IFN-γ, IFN-β) and microarray processing were performed on separate occasions. As described in Chapter-4, one of the key interests in comparing and contrasting the transcriptional response to type-I and type-II interferon lies in the fact that at the functional level these cytokines have evolved to complement each other in overlapping but non-redundant activities [241]. A recent clinically relevant example of this non-redundancy was demonstrated in monocytes and macrophages cultured from patients with complete IFN-γR deficiency [314]. These patients are prone to severe infections with even weakly virulent Mycobacteria and type-I interferon has been proposed as a treatment. However the authors demonstrated IFN-α was deficient at priming for certain LPS inducible cytokines (such as IL-23 and TNF), as well as killing of M. smegmatis [314]. Thus distinguishing what is common and unique in the transcriptional signatures of the two classes of interferons may aid the understanding of why one class of interferon cannot substitute for the lack of another in response to certain pathogens.

As reflected in its original name, "macrophage activating factor", one the key functions of IFN-γ is to sensitise macrophages to activation by pathogen challenge. Indeed the classical activation of macrophages arises from IFN-γ stimulation in combination with TLR ligation [44, 315]. The TLR4 agonist LPS, is arguably one of the best studied and most potent activators of macrophages [11, 316], and the priming of macrophages with IFN-γ is known to have a significant effect on the cells response to LPS (reviewed in [317]). IFN-γ and LPS signals synergise at a number of levels, from signal recognition to target gene regulation [317]. For example IFN-γ positively regulates TLR signalling components (including TLR2, TLR4, MYD88 and the CD14 co-receptor), thereby amplifying sensitivity to TLR signals [317]. In contrast LPS induces the SOCS (suppressor

of cytokine signalling) proteins, known to negatively regulate interferon-receptor activation [318]. Signals also converge at key transcription factors, such as LPS mediated phosphorylation of STAT1, a crucial cytosolic factor for IFN-γ-dependant gene regulation [317]. IFN-γ priming of RAW264.7 cells and monocytes have shown to potentiate NF-κB (NFKB1:RELA) activation in response to subsequent LPS stimulation [319].

Also acknowledged in the literature is the extent of crosstalk and synergy between LPS and type-I interferon signalling. In macrophages, LPS stimulation induces the rapid transcription of IFN-β mRNA and protein secretion [275]. In turn IFN-β signalling forms a key portion of the LPS transcriptional response and LPS-induced lethality [276-277]. In Chapter-5 LPS-induced IFN-β production formed the basis of the assay design to study the role of selected genes of interest in type-I signalling. However these investigations underscored the need to better understand the overlap between the type-I and LPS response. For example, are all IFN-β regulated transcripts also regulated by LPS?

The pathway construction efforts of Chapter-2 depicted the signalling pathways activated in response to these three stimuli (IFN-γ, IFN-β and LPS). A simplified schematic extracted from the larger integrated pathway (Figure 6.1) also illustrates the receptor activation and downstream signalling. As discussed in previous Chapters, although type-I and type-II interferons are structurally distinct and bind to different receptors, the signalling pathways employed by the two cytokines are inter-related (Figure 6.1), and this is reflected in their overlapping target genes (Chapter-4). There is also overlap in the transcriptional targets of interferon and LPS signalling, given the induction of type-I interferon by LPS stimulation. Recognition of LPS in mammalian cells involves a series of interactions with several proteins including the LBP (LPS binding protein), CD14, LY96 (MD-2) and TLR4 (reviewed in [320-321]) (Figure 6.1). TLR4 signal transduction is often divided into the MYD88 dependant and MYD88-independant pathways. The latter pathway relies on TICAM1 an adaptor protein which plays a key role in the activation of the IRF3 transcription factor and subsequent type-I

interferon regulation. The MYD88 dependant pathway is the main protagonist in the activation of the MAPKinase and NF-κB signalling pathways which result in pro-inflammatory cytokine expression.
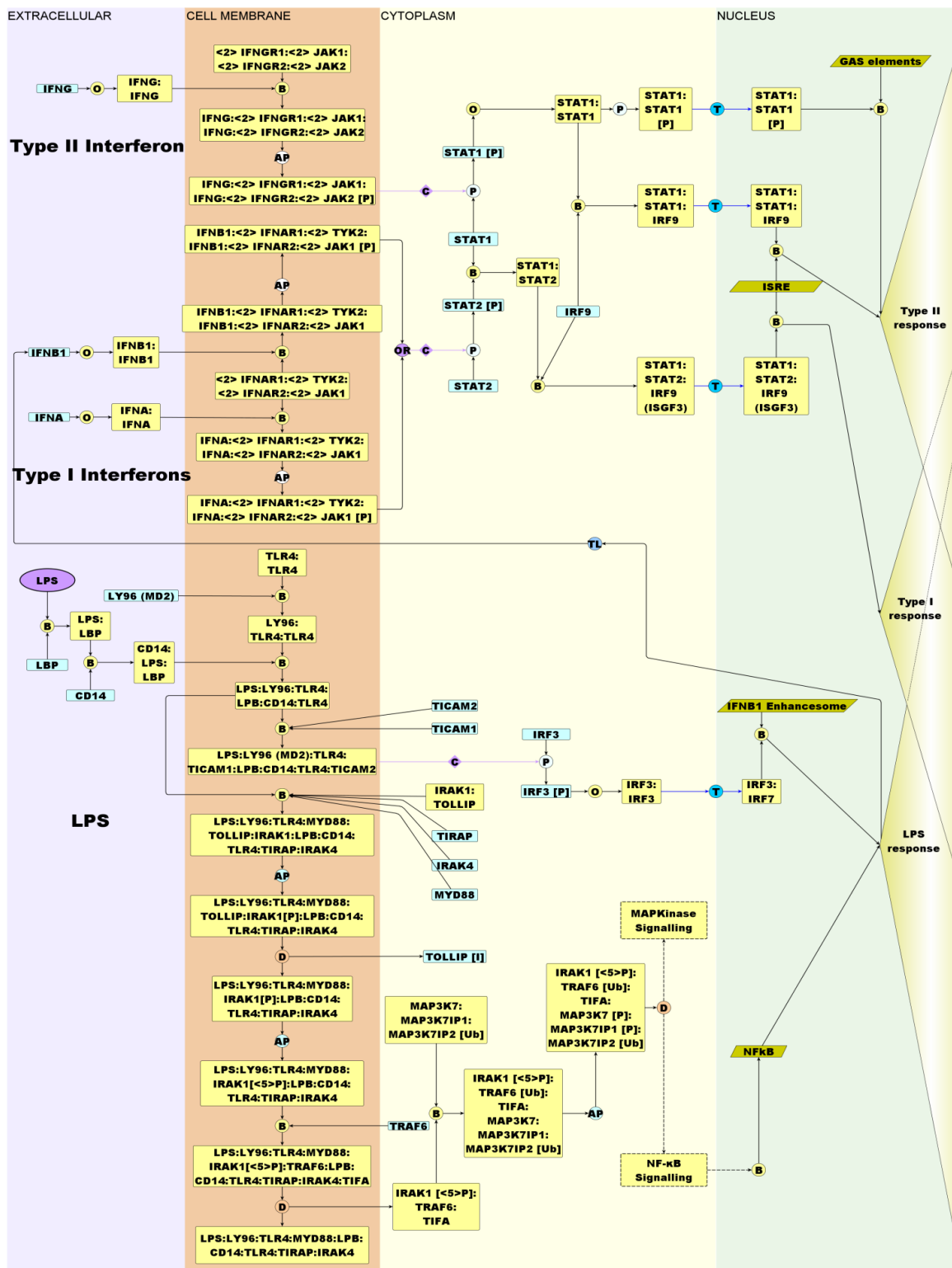
Figure 6.1: Type-I interferon, Type-II interferon and LPS activation of their respective receptor complexes. Type-I (IFN-α/ IFN-β) and type-II interferon (IFN-γ) bind to their respective receptor complexes. Both signalling pathways share the transcription factor STAT1 and overlap I in the transcriptional response induced. TLR4 signalling is commonly divided into the MYD88-dependent and independent pathways; whereby the independent pathway leads to IRF3 activation subsequent regulation of IFN-β enhansesome, and the dependent pathway activates MAPKinase and NFκB signalling cascades. Type-I interferon, type-II interferon and LPS signalling overlap in their transcriptional targets, and are also expected to induce unique sets of genes.

Therefore based on the existing literature and the pathway models of these signalling systems it would be expected that IFN-β, IFN-γ and LPS overlap in their transcriptional profiles. However given there are also key differences in the downstream signalling, there are likely to be unique sets of genes induced in response to the three inputs. The transcriptional response to LPS might be expected to be broader ranging than that of interferon alone. This assumption is based on the fact that LPS stimulation of TLR4 will activate type-I interferon signalling (via the MYD88 independent pathway) as well additional signalling systems through via the MYD88-dependant pathway. Moreover, given this autocrine type-I signalling, one might predict that the transcriptional networks induced by LPS signalling, overlap to a greater extent with those induced by IFN-β signalling than those induced by IFN-γ. Unravelling the individual transcriptional responses to these cytokines may go some way to increasing our understanding of how they contribute to the activated (M1) macrophage phenotype and indeed what we understand by this label. Furthermore dissecting the transcriptional response would considerably improve the transcriptional coverage of the pathway models, where currently the regulation of only a handful of target genes is depicted. Recent searches of microarray data repositories (Array Express and GEO) show there are no data-sets submitted where all three stimuli are compared over a time-course series in primary mouse (bone marrow derived) macrophages. Furthermore, datasets where the transcriptional responses to at least two of three stimuli of interest are measured across multiple time-points in murine BMDMs are not available. The analysis described in this Chapter therefore represents a valuable insight into the transcriptional networks associated with factors known to contribute towards 'classical' activation of macrophages (IFN-β, IFN-γ, and LPS). The work presented will contribute in part to a more detailed, focused analysis for publication purposes. There are however several observations of interest to be made in this analyses.

Results

Analysis of the transcriptional response to IFN-β, IFN-γ or LPS in BMDMs

In order to analyse the transcriptional changes induced by three different stimuli (IFN-β, IFN-γ and LPS) in mouse BMDMs, transcriptional profiling using microarrays was performed over a series of time-points pre- and post-treatment. Cells were challenged with 10 U/ml IFN-β, or 10 U/ml IFN-γ, or 5 ng/ml LPS for 1, 2, 4, 8, and 24 h, or were harvested pre-treatment (0 h). Two biological samples were generated for each time-point per treatment regime. A description of the samples for this analysis is provided in Table 6.1. High quality RNA was processed for labelling and hybridisation to Affymetrix Mouse Gene 1.1. ST Arrays, which were obtained as part of 96-array plate (see Methods).

Sample No	Stimulus	Concentration	Time-points (hours)
1 & 2	None	N/A	0h (pre-treatment/ experiment start)
3 & 4	IFN-β	10 U/ml	1 h post IFN-β
5 & 6	IFN-β	10 U/ml	2 h post IFN-β
7 & 8	IFN-β	10 U/ml	4 h post IFN-β
9 & 10	IFN-β	10 U/ml	8 h post IFN-β
11 & 12	IFN-β	10 U/ml	24 h post IFN-β
13 & 14	IFN-γ	10 U/ml	1 h post IFN-γ
15 & 16	IFN-γ	10 U/ml	2 h post IFN-γ
17 & 18	IFN-γ	10 U/ml	4 h post IFN-γ
19 & 20	IFN-γ	10 U/ml	8 h post IFN-γ
21 & 22	IFN-γ	10 U/ml	24 h post IFN-γ
23 & 24	LPS	5 ng/ml	1 h post LPS
25 & 26	LPS	5 ng/ml	2 h post LPS
27 & 28	LPS	5 ng/ml	4 h post LPS
29 & 30	LPS	5 ng/ml	8 h post LPS
31 & 32	LPS	5 ng/ml	24 h post LPS

Table 6.1: The treatment regimes of 32 BMDM samples generated for genome wide transcriptional analysis. Mouse BMDMs were treated with 10 U/ml IFN-β or, 10 U/ml IFN-γ or 5 ng/ml LPS, over a time-course. All treatments were performed simultaneously and on macrophages prepared from the same culture.

A network graph of the normalised data for the 32 samples was generated in BioLayout *Express*[3D], by filtering for transcripts that shared a Pearson Correlation of 0.85 or above. This resulted in a graph of 11,258 nodes, connected by 270,601 edges (not shown). The graph was then clustered at an MCL inflation value of 2.2, resulting in over 600 clusters with ≥3 nodes. 9,122 transcripts were represented within clusters. This initial network graph was then used to identify clusters representing cohorts of genes that did not exhibit a profile of expression that varied with treatment (such as those unchanging across all arrays or technical artefacts), and transcripts within these clusters were removed from the graph. As might be expected many transcripts were relatively evenly expressed across all treatments, and after filtering only 3,747 transcripts remained in clusters representing regulated genes.

A further network graph of the transcriptional data pertaining only to the 3,747 transcripts was generated, (again using a Pearson Correlation threshold of 0.85). Again two major graph components were identifiable; one representing up-regulated transcripts, the other down-regulate transcripts. The graph of 3,747 nodes connected by 172,688 edges was clustered at an MCL inflation value of 2.2 (Figure 6.2), resulting in over 70 clusters with ≥3 nodes. The clusters were then inspected for patterns of expression related to the three treatment regimes and annotated accordingly. The annotation of clusters was based on two predominant factors:

(i)   The directionality of the change; where **Up** or **Down** defines up or down regulation of expression of transcripts within the cluster.

(ii)   The specificity of the change with respect to a given treatment regime. Clusters annotated "**Specific**" represent transcripts which are only regulated in a given treatment. Whereas clusters annotated "**Preferential**" are regulated in all treatments, but to a greater extent in the denoted treatment regime.

For example cluster-1 "Down Preferentially in LPS" indicates the transcripts within this cluster are down regulated in all treatments but more so in those cells treated with LPS. On the other hand cluster-2 "Up Specifically in LPS" implies the induction of transcripts within this cluster is observed only in cells stimulated with LPS. A handful

of clusters represented changes occurring in all three treatments to a similar extent and these are annotated "Up/Down Similar in All Treatments".
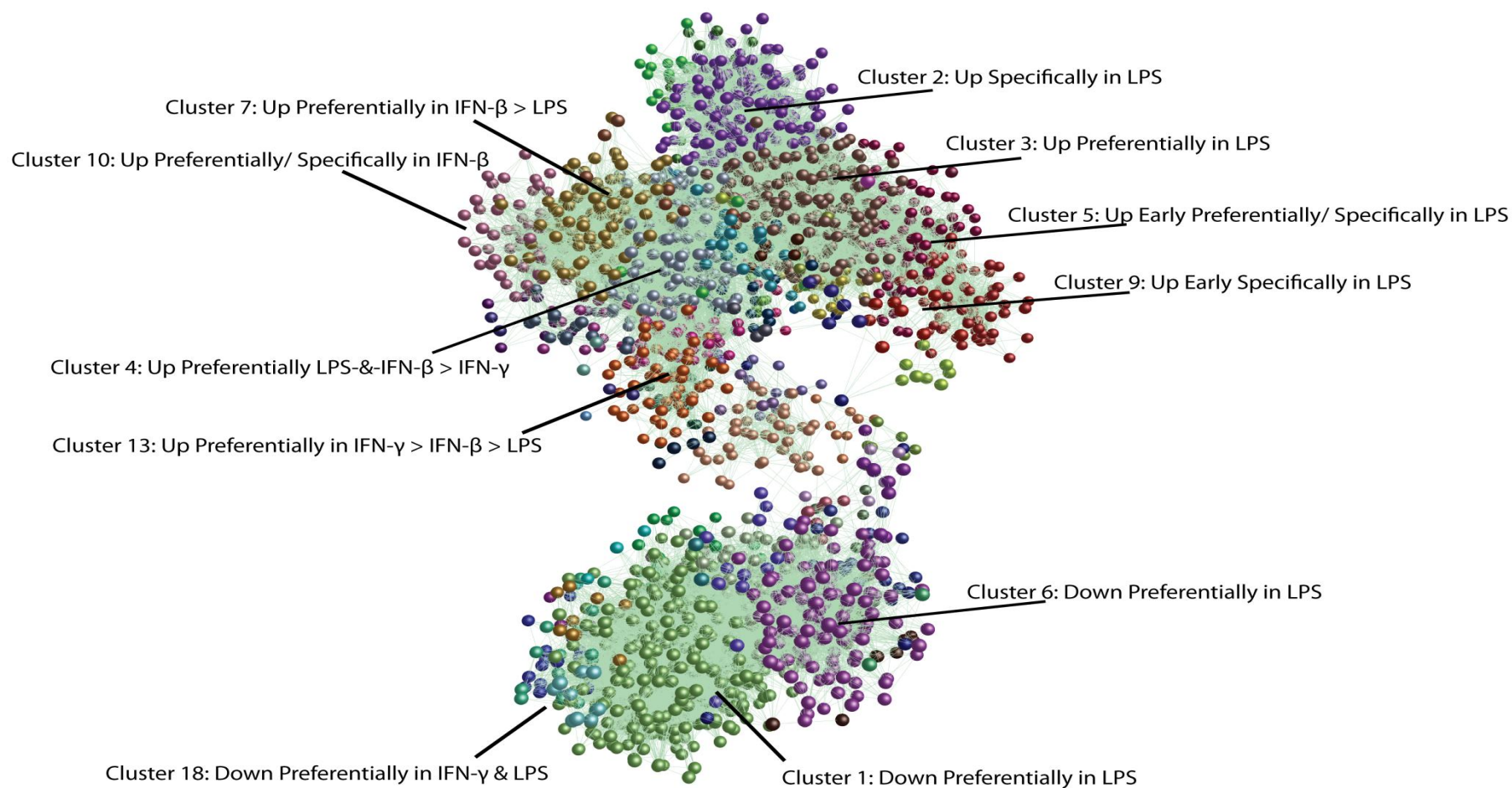
Figure 6.2: Network graph of transcriptional changes occurring in mouse BMDMs in response to IFN-β, IFN-γ or LPS challenge. The network graph of transcriptional data pertaining to the treatment of BMDMs with IFN-β, IFN-γ or LPS over a time-course comprising 3,747 nodes (transcripts), connected by 172,688 edges at a Pearson correlation of 0.85 or above. The graph clustered at an MCL inflation value of 2.2 comprises over 70 clusters with ≥3 nodes. Nodes (transcripts) within the same cluster share the same colour. The upper component of the graph represents transcripts up-regulated in response to treatment, whereas the lower component represents transcripts down-regulated in response to treatment.

To further appreciate the structure of the transcriptional network generated (in Figure 6.2) in response to the three treatments, a collapsed-cluster network graph was generated; whereby the nodes making up this graph represent the different clusters and the edges denote the relationships between the clusters. The size of the nodes shown in the graph is proportional to the cluster membership i.e. number of transcripts. The collapsed-cluster graph was arranged in 2-dimentions to reflect the 3-dimentional (un-collapsed) network graph (Figure 6.3). The graph illustrates that only a few clusters represent transcriptional networks that are activated specifically in response to a given treatment. The majority of clusters represented transcriptional changes occurring preferentially in one treatment over the others. Furthermore the expression of most transcripts was preferentially altered in response to LPS treatment.

As in previous Chapters the clusters generated from this analysis were annotated based on their gene membership and over-representation of cohorts of functionally associated genes using the Database for Annotation, Visualization and Integrated Discovery (DAVID) [248-249]. The most over-represented terms (based on DAVID analysis of GO Ontology terms (GO FAT category)) were chosen to describe the biology of the clusters. Tables 6.2(a-e) provide an overview of all clusters (with ≥ 10 nodes); this includes the number of transcripts within each cluster, examples of gene members and the associated GO terms. The expression profiles of selected clusters of interest are also shown in Figures 6.4(a-d). The cluster profiles are sorted according to treatment centric changes (e.g. changing specifically/ preferentially in LPS/ IFN-β/ IFN-γ), and the maximal/ minimal time of gene induction or repression, respectively.
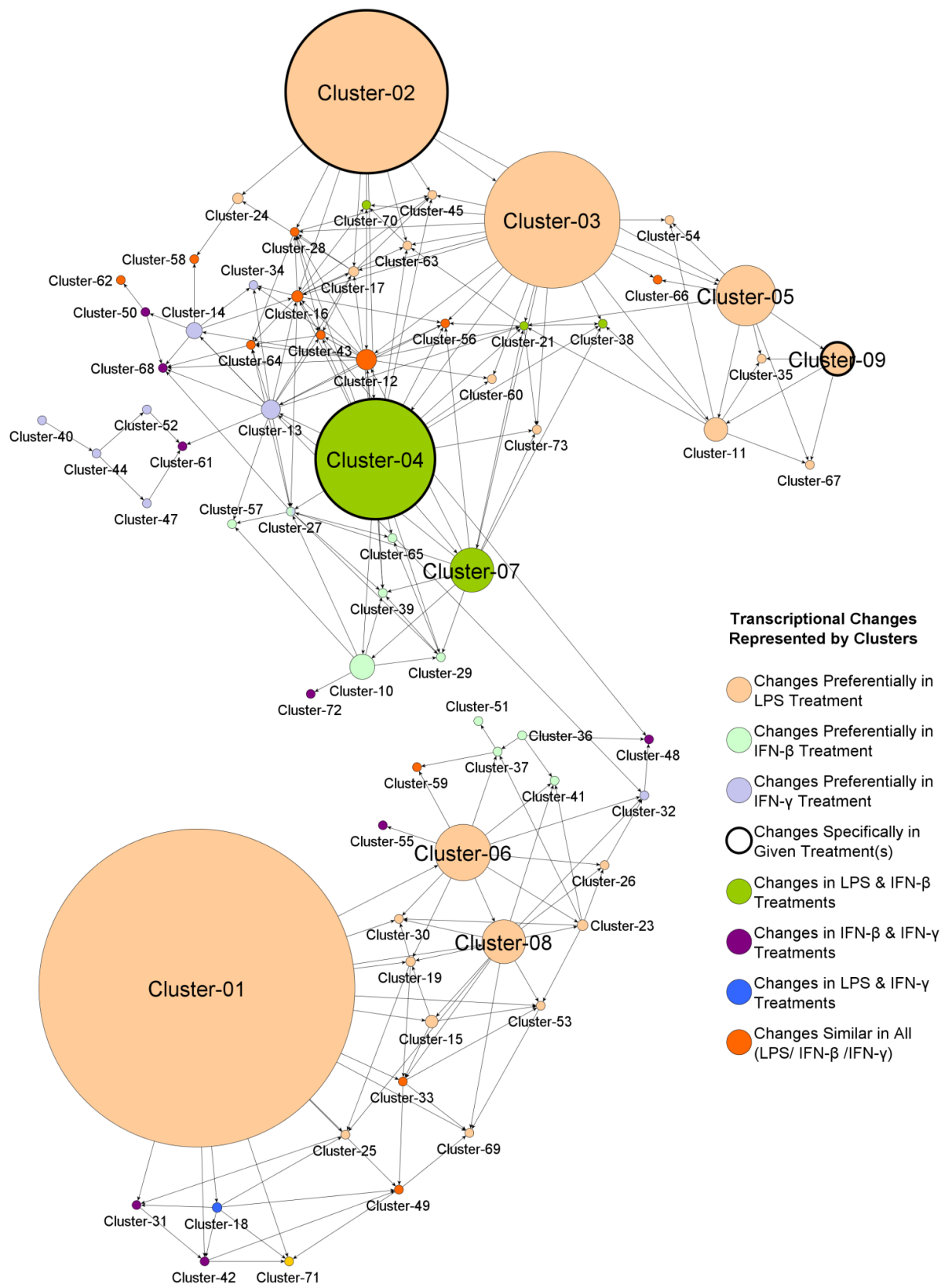
Figure 6.3: Collapsed cluster relationship network of the transcriptional changes occurring in mouse BMDMs in response to IFN-β, IFN-γ or LPS challenge based on the transcriptional network shown in Figure 6.2. Nodes represent individual clusters and are sized proportionally to their transcript membership and are coloured according to the treatment specific pattern of expression they represent. Edges denote relationships between clusters.

LPS Centric Clusters –Average Expression Profile
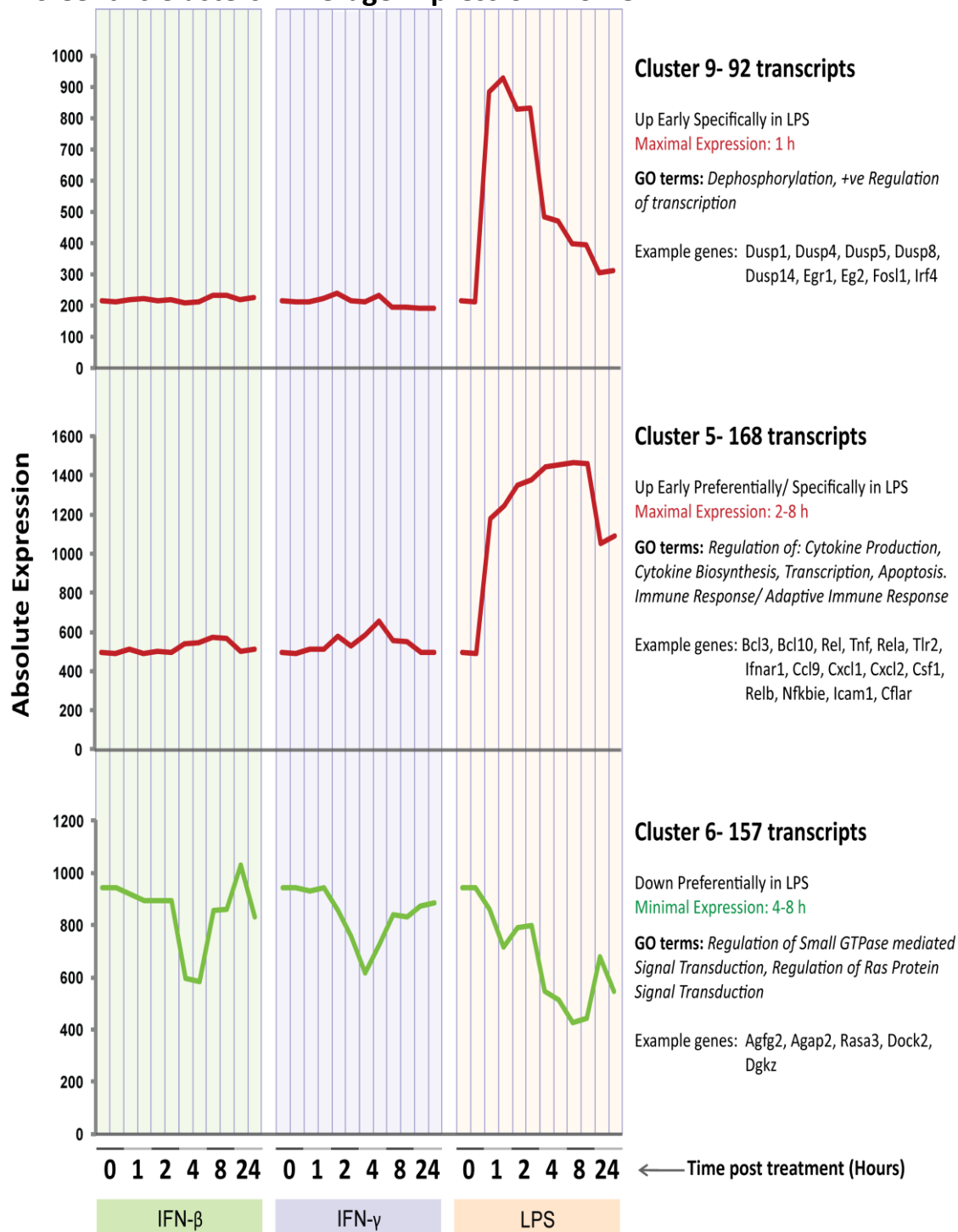
Figure 6.4: (a) Average expression profiles across clusters associated with preferential or specific changes in LPS treated mouse BMDMs. Mean expression profiles for each sample (array) are calculated from the expression levels of all transcripts within the given clusters. Expression levels are plotted across the different time-points sampled and the three different treatment regimes (IFN-β, IFN-γ and LPS). GO terms associated with genes in each cluster are shown to the right of each plot.

LPS Centric Clusters –Average Expression Profile

Cluster 3- 380 transcripts

Up Preferentially in LPS
Maximal Expression: 8 h

GO terms: *Regulation of: Phosphorylation, Lymphocyte Activation, T Cell Activation, Leukocyte Activation, Cytokine Activity, Cell activation, Apoptosis, Mononuclear Cell Proliferation*

Example genes: Il1b, Il6, Il18, Il12a, Nod2, Tlr6, Ccl17, Ccl22, Cxcl16, Cxcl3, Traf1, Traf2, Traf5, Casp7

Cluster 1- 886 transcripts

Down Preferentially in LPS
Minimal Expression: 24 h

GO terms: *Cell Cycle, M Phase, M Phase of Mitotic Cell Cycle, Mitosis, Nuclear Division*

Example genes: Bub1, Bub1b, Cdk1, Cdc20, Cdc25b, Cdc7, Cenpe, Cenpf, Chek1, Ccna2, Ccnf, Mcm2

Cluster 2- 454 transcripts

Up Specifically in LPS
Maximal Expression: 24 h

GO terms: *Response to Wounding, Immune/ Inflammatory Response, Oxidation Reduction, +ve Regulation of Immune System Process, Macrophage Activation*

Example genes: Ccl5, Cxcl5, Aox1, Nos2, Prdx5, Cd38, Cd5, Tlr1

Figure 6.4: ((a) continued from previous page) Average expression profiles across clusters associated with preferential or specific changes in LPS treated mouse BMDMs. Mean expression profiles for each sample (array) are calculated from the expression levels of all transcripts within the given clusters. Expression levels are plotted across the different time-points sampled and the three different treatment regimes (IFN-β, IFN-γ and LPS). GO terms associated with genes in each cluster are shown to the right of each plot.

IFN-β Centric Clusters –Average Expression Profile

Cluster 7- 123 transcripts

Up Preferentially in IFN-β > LPS
Maximal Expression: 4 h

GO terms: *Regulation of Transcription, Macro-molecule Catabolic Process, Chromatin Modification*

Example genes: Daxx, TCf4, Zfp213, Socs7, Usp12, Acvr1, Dll1, Foxf1a, Nr3c1

Cluster 10- 70 transcripts

Up Preferentially/ Specifically in IFN-β
Maximal Expression: 4 h

GO terms: *--Unavailable--*

Example genes: Il10, Casp2, Tnfsf8

Cluster 27- 18 transcripts

Up Preferentially in IFN-β > IFN-γ
Maximal Expression: 4-8 h

GO terms: *Acute Inflammatory Response, Innate Immune Response*

Example genes: Fcgr1, Masp1, Treml2

Cluster 36- 11 transcripts

Up Late Preferentially/ Specifically in IFN-β
Maximal Expression: 24 h

GO terms: *Protein Amino Acid N-linked Glycosylation, Sodium Ion Binding*

Example genes: Cd4, Cdkn1c, Slc5a3, Spry3
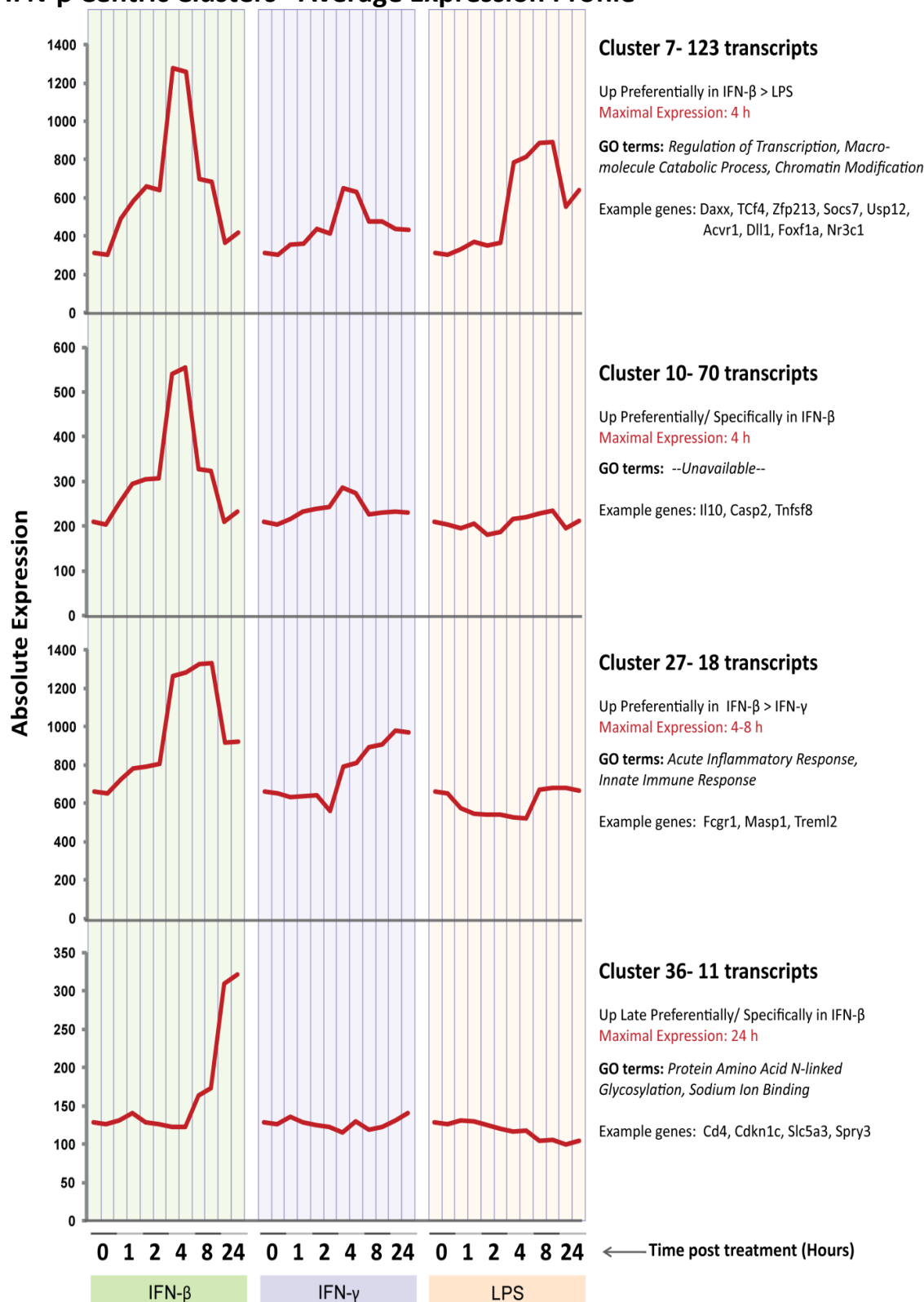
Figure 6.4: (b) Average expression profiles across clusters associated with preferential or specific changes in IFN-β treated mouse BMDMs. Mean expression profiles for each sample (array) are calculated from the expression levels of all transcripts within the given clusters. Expression levels are plotted across the different time-points sampled and the three different treatment regimes (IFN-β, IFN-γ and LPS). GO terms associated with genes in each cluster are shown to the right of each plot.

IFN-γ Centric Clusters –Average Expression Profile

Cluster 13- 52 transcripts

Up Preferentially in IFN-γ > IFN-β > LPS
Maximal Expression: ~4-8 h

GO terms: *GTP Binding, Immune Response, Response to Cytokine Stimulus*

Example genes: Cxcl9, Gbp4, Iigp1

Cluster 14- 45 transcripts

Up Preferentially/ Specifically in IFN-γ
Maximal Expression: 24 h

GO terms: *Immune Response, Adaptive Immune Response, Immunoglobulin Mediated Immune Response, Lymphocyte Mediated Immunity, Complement Activation*

Example genes: Il2rb1, H2-Dma, H2-Eb1, Ly6a, Ciita, C4b

Cluster 34- 12 transcripts

Up Preferentially in IFN-γ > LPS
Maximal Expression: 24 h

GO terms: *Immune Response, Antigen Processing & Presentation of Peptide Antigen via MHC class II*

Example genes: Cd74, H2-Aa, H2-Ab1, H2-Ea
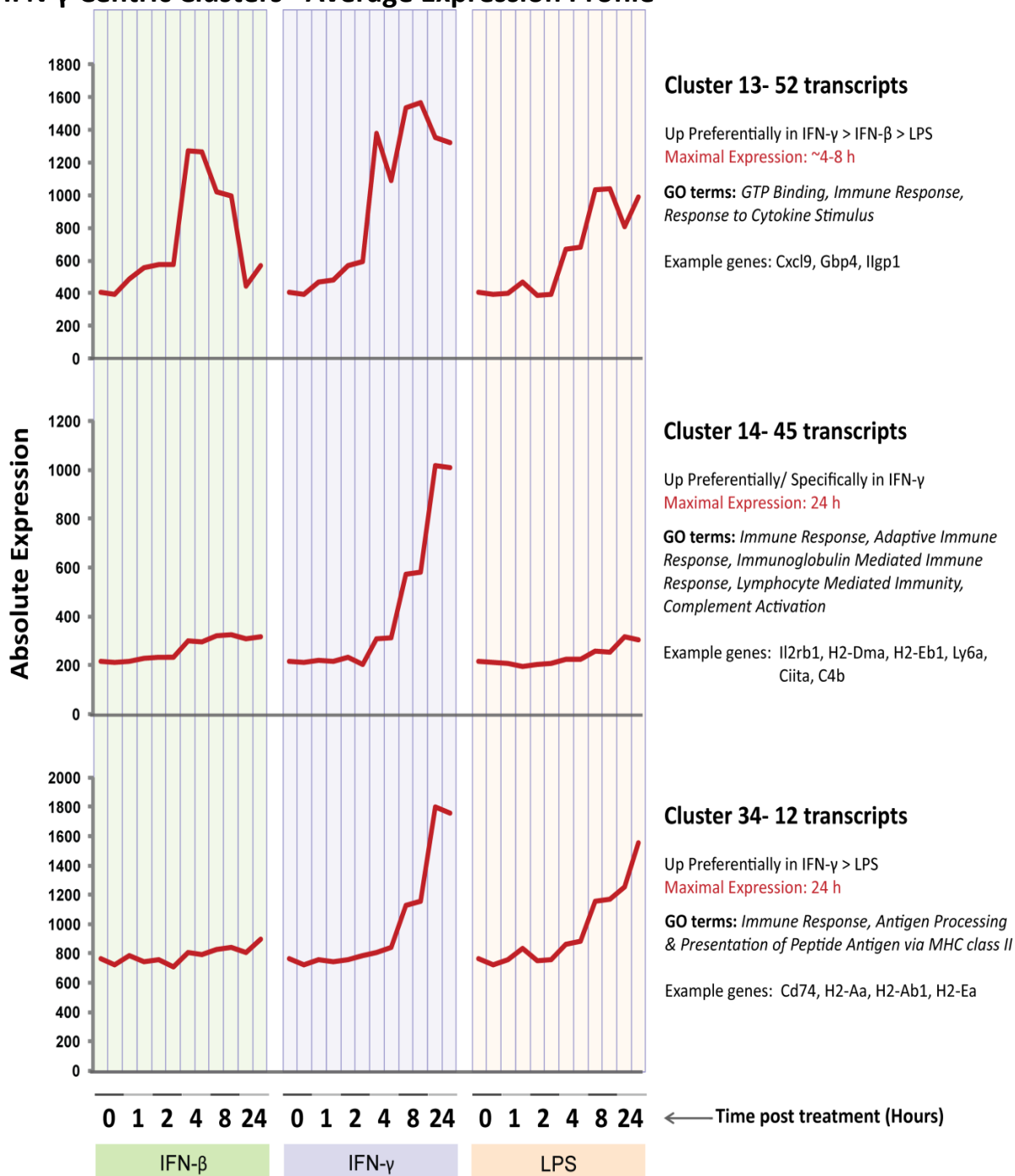
Figure 6.4: (c) Average expression profiles across clusters associated with preferential or specific changes in IFN-γ treated mouse BMDMs. Mean expression profiles for each sample (array) are calculated from the expression levels of all transcripts within the given clusters. Expression levels are plotted across the different time-points sampled and the three different treatment regimes (IFN-β, IFN-γ and LPS). GO terms associated with genes in each cluster are shown to the right of each plot.

Other Clusters –Average Expression Profile

Cluster 4- 335 transcripts

Up Preferentially LPS-&-IFN-β > IFN-γ
Maximal Expression: ~4-8 h

GO terms: *Cellular Macromolecule Catabolic Process, Response to Virus, ISG15-protein conjugation, Immune Response*

Example genes: Oas2, Casp3, Pmepa1, Psma4, Usp25, Ifih1, Tlr3, Tlr8, Usp18, Uba7, Ube2l6

Cluster 12- 57 transcripts

Up Similar in All Treatments
Maximal Expression: ~4-8 h

GO terms: *GTPase Activity, Immune Response, Transcription*

Example genes: Gbp9, Gnb4, Gbp2, Gbp6, Irgm2, Gbp1, Irf5, Stat31

Cluster 16- 32 transcripts

Up Similar in All Treatments
Maximal Expression: ~8-24 h

GO terms: *Antigen Processing & Presentation, Transmembrane Transport*

Example genes: Psme1, Psmb8, Tap2

Figure 6.4: (d) Average expression profile across clusters associated with changes across all three treatments studied in mouse BMDMs. Mean expression profiles for each sample (array) are calculated from the expression levels of all transcripts within the given clusters. Expression levels are plotted across the different time-points sampled and the three different treatment regimes (IFN-β, IFN-γ and LPS). GO terms associated with genes in each cluster are shown to the right of each plot.

LPS Centric Clusters – Down Regulated Response

Cluster-No-ID	No of Transcripts	Cluster- Description	Up or Down *	Example Genes	Enriched GO Annotation
1	886	Down Preferentially in LPS		Bub1, Bub1b, Cdk1, Cdc20, Cdc25b, Cdc7, Cenpe, Cenpf, Chek1, Ccna2, Ccnb2, Ccnf, Mcm2,	Cell Cycle, M Phase, M Phase of Mitotic Cell Cycle, Mitosis, Nuclear Division
6	157	Down Preferentially in LPS		Agfg2, Agap2, Rasa3, Dock2, Dgkz	Regulation of Small GTPase mediated Signal Transduction, Regulation of Ras Protein Signal Transduction
8	122	Down Preferentially in LPS		Cdt1, Sigirr, E2f2, Cdc6, Cdt1, Ccne1, Cdk2,	-ve Regulation of Macromoecule Metabolic Process, -ve Regulation of Cellular Biosynthetic Process, Cell Cycle, -ve Regulation of Gene Expression
15	35	Down Preferentially in LPS		Decr1, Nqo2, Aldh9a1, Dhrs7, Gpd1l	Oxidation Reduction
19	27	Down Preferentially in LPS		Adcy3, Nme3, Padi2,	Nitrogen Compound Biosynthetic Process, Purine Nucleotide Biosynthetic Process
23	22	Down Preferentially in LPS		Nfam1, Card9, Eif2ak3,	+ve Regulation of Binding, +ve Regulation of Signal Transduction
25	19	Down Preferentially in LPS > IFN-γ > IFN-β		Atp2a3, Cacna1d, Slc24a3, L1cam	Calcium Ion Transport, Cell Adhesion
26	18	Down Preferentially in LPS		Kdm1b, Dhrs7b, Spr	Oxidation Reduction
30	15	Down Preferentially in LPS		Plau, Rassf2	--Unavailable--
31	13	Down Preferentially in LPS & IFN-β		Sf1, Hnrnpa1, Ckap2l	RNA Splicing, mRNA Processing

Table 6.2: (a) Description of clusters associated with down-regulated transcriptional changes occurring preferentially/ specifically in LPS treatment.

LPS Centric Clusters – Up Regulated Response

Cluster-No-ID	No of Transcripts	Cluster-Description	Up or Down *	Example Genes	Enriched GO Annotation
2	454	Up Specifically in LPS		Ccl5, Cxcl5, Aox1, Nos2, Prdx5, Cd38, Cd5, Tlr1	Response to Wounding, Immune Response, Inflammatory Response, Oxidation Reduction, +ve Regulation of Immune System Process, Macrophage Activation
3	380	Up Preferentially in LPS		Il1b, Il6, Il18, Il12a/b, Nod2, Tlr6, Ccl17, Ccl22, Cxcl16, Cxcl3, Traf1/2/5, Casp7	Regulation of: Phosphorylation, Phosphate/Phosphorus Metabolic Process, Lymphocyte Activation, T Cell Activation, Leukocyte Activation, Cytokine Activity, Cell activation, Apoptosis, Mononuclear Cell Proliferation
5	168	Up Early Preferentially/ Specifically in LPS		Bcl3, Bcl10, Rel, Tnf, Rela, Tlr2, Ifnar1, Ccl9, Cxcl1, Cxcl2, Csf1, Relb, Nfkbie, Icam1, Cflar,	Regulation of Cytokine Production, Immune Response, +ve Regulation of Multicellular Organismal Process, Regulation of Transcription, +ve Regulation of Cytokine Biosynthesis, Adaptive Immune Response, Regulation of apoptosis
9	92	Up Early Specifically in LPS		Dusp1, Dusp4, Dusp5, Dusp8, Dusp14, Egr1, Eg2, Fosl1, Irf4	Dephosphorylation, +ve Regulation of transcription,
11	66	Up Preferentially/ Specifically in LPS		Atp2b4, Slc22a21, Slc22a5, Slc39a14, Stat5, Tnfsf15	(Cation) Transport, Intracellular Signalling Cascade
17	28	Up Preferentially in LPS		C2, Irf7, Bst2	Immune Response, Immunoglobulin Mediated Imme Response, B Cell/ Lymphocyte Mediated Immunity
21	24	Up Preferentially in LPS > IFN-β > IFN-γ		Il15, Casp4, Zfp800	--Unavailable--
24	20	Up Late Preferentially in LPS		Ddit3, Derl1, Alkbh2, Aifm2	Cellular Response to Stress, Cellular Response to Unfolded Protein, +ve Regulation of Apoptosis
35	11	Up Preferentially in LPS		Phldb1, Plscr1	--Unavailable--

Table 6.2: (b) Description of clusters associated with up-regulated transcriptional changes occurring preferentially/ specifically in LPS treatment.

IFN-β Centric Clusters

Cluster-No-ID	No of Transcripts	Cluster-Description	Up or Down *	Example Genes	Enriched GO Annotation
37	11	Down & Up Preferentially in IFN-β		Gba2, Gpr162, Map3k12	--Unavailable--
7	123	Up Preferentially in IFN-β > LPS		Daxx, TCf4, Zfp213, Socs7, Usp12, Acvr1, Dll1, Foxf1a, Nr3c1	Regulation of Transcription, Macromolecule Catabolic Process, Determination of Left/Right Symmetry, Chromatin Modification
10	70	Up Preferentially/ Specifically in IFN-β		Il10, Casp2, Tnfsf8,	--Unavailable--
27	18	Up Preferentially in IFN-β > IFN-γ > LPS		Fcgr1, Masp1, Treml2	Acute Inflammatory Response, Innate Immune Response
29	16	Up Preferentially in IFN-β		Abcb1a, Prnp	Response to Metal Ion
36	11	Up Late Preferentially/ Specifically in IFN-β		Cd4, Cdkn1c, Slc5a3, Spry3	Protein Amino Acid N-linked Glycosylation, Sodium Ion Binding
39	11	Up Preferentially in IFN-β		Pou3f1, Dennd1b	--Unavailable--

Table 6.2: (c) Description of clusters associated with transcriptional changes occurring preferentially/ specifically in IFN-β treatment.

IFN-γ Centric Clusters

Cluster-No-ID	No of Transcripts	Cluster-Description	Up or Down *	Example Genes	Enriched GO Annotation
13	53	Up Preferentially in IFN-γ > IFN-β > LPS		Cxcl9, Gbp4, Ilgp1	GTP Binding, Immune Response, Response to Cytokine Stimulus
14	45	Up Preferentially/ Specifically in IFN-γ		Il2rb1, H2-Dma, H2-Eb1, Ly6a, Ciita, C4b	Immune/Adaptive Immune Response, Immunoglobulin Mediated Immune Response, Lymphocyte Mediated Immunity, Complement Activation
32	13	Up Preferentially in IFN-γ		Gm2a, Pafah1b3, Scn3b	Lipid Catabolic Process
34	12	Up Preferentially in IFN-γ > LPS (not IFN-β)		Cd74, H2-Aa, H2-Ab1, H2-Ea	Immune Response, Antigen Processing & Presentation of Peptide Antigen via MHC class II
40	11	Up Preferentially in IFN-γ		Ccr1, Adora3, Sbno2, Gpr146	Myeloid Cell Activation During Immune Response
44	10	Up Preferentially in IFN-γ		Selp, Traf3ip2, Ppard	+ve Regulation of Protein Kinase Cascade

Table 6.2: (d) Description of clusters associated with transcriptional changes occurring preferentially/ specifically in IFN-γ treatment.

Other Clusters

Cluster-No-ID	No of Transcripts	Cluster-Description	Up or Down *	Example Genes	Enriched GO Annotation
18	27	Down Preferentially in IFN-γ & LPS		Cd9, Fn1, Clstn1, Itgax	Cell Adhesion, Biological Adhesion, Cell Junction Assembly
33	13	Down in all treatments		Ipo11, Psmg2, Chst11	Protein Complex Assembly
41	10	Down & Up in IFN-β_Down in IFN-γ & LPS		Vash2, Gab3	--Unavailable--
4	335	Up Preferentially LPS-&-IFN-β > IFN-γ		Oas2, Casp3, Pmepa1, Psma4, Usp18, Usp25, Ifih1, Tlr3, Tlr8, Usp18, Uba7, Ube2l6	Cellular Macromolecule Catabolic Process, Response to Virus, ISG15-protein conjugation, Immune Response
12	57	Up Similar in all treatments		Gbp9, Gnb4, Gbp2, Gbp6, Irgm2, Gbp1, Irf5, Stat3	GTPase Activity, Immune Response, Transcription
16	32	Up Similar in all treatments		Psme1, Psmb8, Tap2	Antigen Processing & Presentation, Transmembrane Transport
28	17	Up Similar in all treatments		Atxn7l3, Kdm2a, Mll5, Msl1	Chromatin Modification/Organisation
43	10	Up Late Similar in all treatments		Tapbp, H2-T23, H2-T24	Antigen Processing & Presentation

Table 6.2: (e) Description of clusters associated with transcriptional changes in response to any of IFN-β, IFN-γ or LPS.

To better visualise the overlap and distinction in the transcriptional response to IFN-β, IFN-γ, and LPS, network diagrams displaying the targets of each treatment (as determined in this analysis) were generated. Separate networks were generated for the up and down regulated gene targets. Connections (edges) were defined between the treatment-type (IFN-β, IFN-γ, and LPS) and the genes regulated in response to a given treatment. Therefore if a particular gene was regulated in response to all three treatments there would be thee input edges into that gene. Connections from treatment-type to genes were captured regardless of the extent of transcriptional induction/ repression, i.e. genes preferentially induced/repressed in LPS and also induced/repressed in response to IFN-β and IFN-γ (but to a lesser extent) are connected to all three treatment-types.

The up-regulated response comprised transcriptional targets common to all three treatments (948 transcripts), targets shared between two treatments only (IFN-γ and IFN-β: 33, IFN-γ and LPS: 11, LPS and IFN-β: 111 transcripts), and genes unique to the individual treatments (IFN-γ: 50, IFN-β: 115, and LPS: 727 transcripts). In contrast to the inducible response all the down-regulated targets within clusters were common to all three stimuli.

(a) Up-Regulated Gene Targets

(b) Down-Regulated Gene Targets

→ IFN-γ targets

→ IFN-γ & IFN-β targets

IFN-γ

IFN-β

LPS

IFN-γ , IFN-β, LPS targets

→ IFN-β targets

→ IFN-γ & LPS targets

→ IFN-β & LPS targets

LPS targets

IFN-γ

IFN-β

LPS
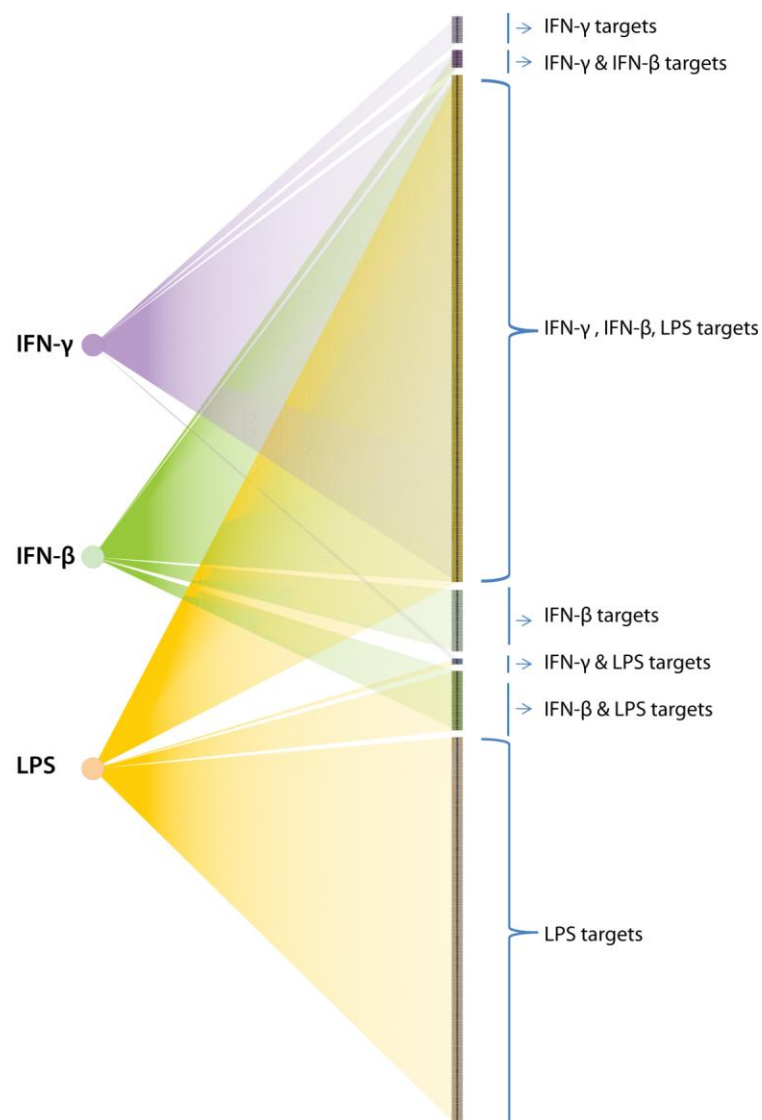
IFN-γ, IFN-β, LPS targets
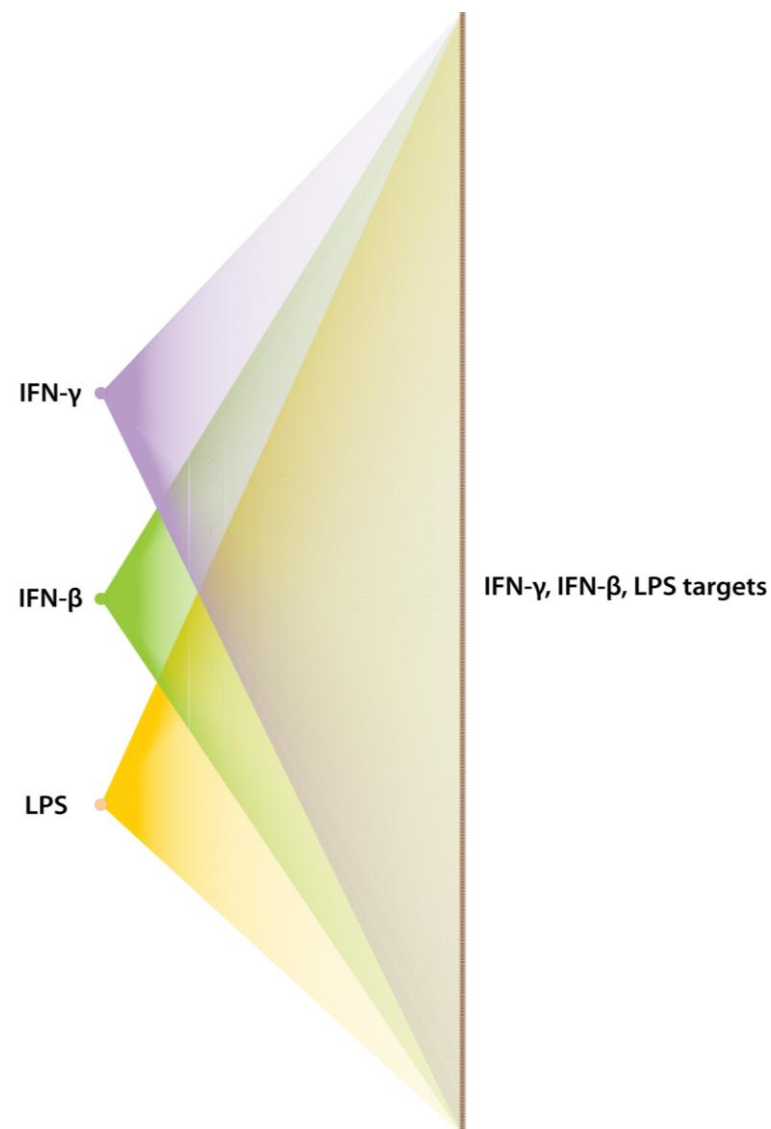
Figure 6.5: Overlap in the up and down regulated transcriptional targets of IFN-β, IFN-γ, and LPS. (a) Overlap and divergences in the transcriptional networks induced by IFN-β, IFN-γ and LPS in mouse BMDMs. (b) Overlap in the transcriptional networks repressed by IFN-β, IFN-γ and LPS in mouse BMDMs.

279

Discussion

This final results Chapter presents a high-level analysis of the macrophage response to three stimuli which are considered to prime cells towards what is often described as an M1 phenotype [44]. IFN-β, IFN-γ and LPS activated signalling pathways converge at a number of levels and elicit an anti-microbial response by regulating the expression of hundreds of genes. "Classical" activation of macrophages is considered to be attained by IFN-γ exposure in concert with a microbial stimulus such as LPS. Arguably LPS is by far the most extensively studied microbial activator of macrophage signalling and transcriptional cascades. LPS stimulation of macrophages induces the expression a number of cytokines including IFN-β which then act in an autocrine manner to contribute to the LPS transcriptional response [276-277]. As discussed in previous Chapters (1 & 4) the type-I interferon, IFN-β, and the type-II interferon, IFN-γ are structurally unrelated and bind to different receptors however they do share downstream signalling machinery and overlap in their transcriptional response. The variance in the type-I and type-II transcriptional signatures are potentially fundamental in the non-redundant activities of these cytokines; i.e. where one interferon cannot substitute for the lack of another as demonstrated in both experimental and clinical states of infection [241-242, 314].

Despite the acknowledged signal convergence of these three stimuli and their interrelated activation of the macrophage transcriptome, high-quality detailed transcriptional datasets comparing and contrasting the actions of IFN-β, IFN-γ and LPS are not currently available. In an attempt to begin to understand and identify the common and distinct patterns of gene expression the analyses of this chapter have attempted to delineate the transcriptional events in response to each of the three stimuli of interest in mouse BMDMs. It is acknowledged that the individual reposes to these stimuli have been studied in great deal in their own right [97, 238, 243-244, 246-247, 302, 309, 322-326]. Thus this analysis did not attempt to describe all characteristics of the macrophage response to endotoxin or to interferons. Instead the

focus was on identifying patterns of expression unique and overlapping in the three treatments.

Overall Transcriptional Network Structure in BMDMs Following IFN-β, IFN-γ, or LPS Stimulation

For this analysis primary mouse BMDMs were treated with 10 U/ml IFN-β or, 10 U/ml IFN-γ or, 5 ng/ml LPS over a time-course. The time-points analysed in this study (0, 1, 2, 4, 8 and 24 h) had been effective at distinguishing temporal classes of inducible and repressible genes in the previous analysis Chapter-4.

A filtered network graph of transcripts regulated in this data-set comprised 3,747 nodes connected by 172,688 edges. As with the time-course study of Chapter-4 the network graph had two main components; one comprising up-regulated transcripts and the other down regulated transcripts. The up-regulated graph component had two predominant sections; changes which were preferential/specific to LPS treatment and changes which were preferential/specific to other treatments. Whilst the up-regulated response comprised a number of clusters representing changes specific to given treatments (e.g. cluster-2, cluster-4, and cluster-9), no such clusters existed in the down regulated response. Therefore all the down-regulated targets in this analysis were common to all three treatments (Figure 6.5). However 1,313 of the 1,423 down-regulated transcripts were repressed to a greater extent in LPS treatment, compared to the interferon treatments.

Similar to the down-regulated response, the inducible transcriptional response was much broader and potent (in terms of fold-change induction) in LPS treated macrophages compared to those treated with either IFN-β or IFN-γ. 727 transcripts belonged to clusters which were specific to changes in LPS treatment (e.g. cluster-2 and cluster-9) or highly specific to LPS treatment (those annotated "Preferential/Specific"). It was hypothesised that LPS would induce a broader transcriptional response since TLR-4 activation stimulates a number of downstream pathways in addition to autocrine IFN-β signalling. Indeed, this was the observed

outcome and many of the LPS-specific genes were related to signalling pathways activated directly downstream of TLR-4 ligation (i.e. MAPKinase signalling and NF-κB). 948 transcripts overlapped between all three treatments however 50% of these transcripts were preferentially induced by LPS stimulation. Amongst the overlap specific to two treatments, LPS and IFN-β had more (inducible) transcripts in common (111) than LPS and IFN-γ (11 transcripts). Both IFN-γ and IFN-β induced small subsets of transcripts unique to each treatment (50 and 115 respectively) and not overlapping with LPS. Therefore as predicted each treatment type (IFN-β, IFN-γ and LPS) induced unique sets of genes, which might be attributed to the differences in the signalling pathways employed by these stimuli. However, a large number of the transcriptional targets of these three stimuli were overlapping, suggesting a non-specific general macrophage response to interferon and LPS. Discussed herein are some of the key observations in the differential and overlapping transcriptional response to IFN-β, IFN-γ or LPS in mouse BMDMs.

LPS Specific and Centric Transcriptional Changes

The most distinguishable "early" induced set of transcripts belonged to cluster-9-representing changes occurring specifically in LPS treated macrophages (Figure 6.4a). The 92 transcripts within this cluster were induced transiently from 1 hour post-LPS treatment. Amongst the genes represented within this cluster, there were five dual specificity phosphatases (*Dusp1, Dusp4, Dusp5, Dusp8,* and *Dusp14*), regulators of the NF-κB system (*Nfkbia, Nfkbid*, and *Nfkbiz*), a number of transcripts representing microRNA's (*Mir17, Mir18, Mir19a, Mir19b-1, Mir221*, *Mir222*, and *Mir92-1*), and the interferon regulatory factor *Irf4*. The changes observed in this cluster are fitting with the literature and highly indicative of negative feedback control of the TLR-4 activated pathways. For example IRF4 has been demonstrated as negative regulator of TLR signalling and proinflammatory cytokine production [290, 327]. Moreover IRF4$^{-/-}$ mice display increased sensitivity to LPS-induced shock and exaggerated TNF-α and IL-6 production [290].

As established in the literature and the pathway diagrams (Chapter-2), the MAPKinase signal transduction pathways acts downstream of TLR-4 receptor activation and this signalling cascade is involved in cytokine production. Many DUSP proteins (often referred to as MAPKinase phosphatases), are responsible for dephosphorylating threonine and tyrosine residues on MAPKkinase proteins and in doing so control the duration and intensity of MAPKinase signalling [328]. For example DUSP-1 has been shown to be a key negative regulator of the inflammatory response by regulating the p38 and JNK (Jun N-terminal protein kinase) MAPKinase pathways and thereby pro- and anti-inflammatory cytokine production [329-332] (Figure 6.6).
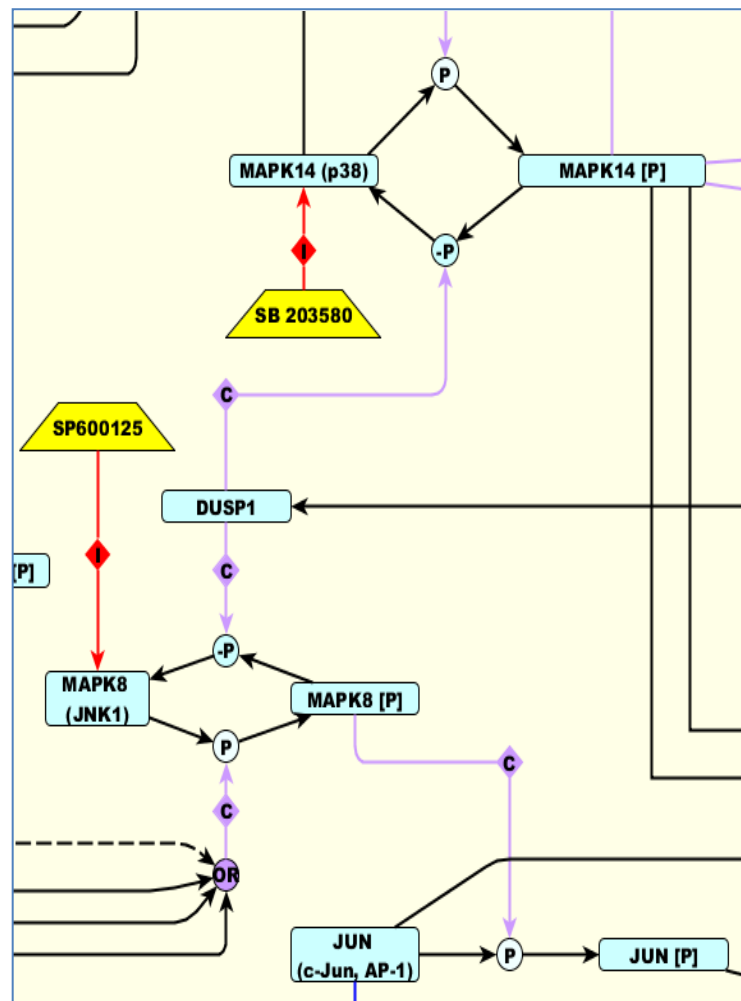
Figure 6.6: Extract from the integrated pathway diagram illustrating the action of the MAPKinase phosphatase, DUSP1. DUSP1 catalyses the de-phosphorylation of both MAPK8 and MAPK14 proteins.

NF-κB-signalling is also activated in response to TLR4 receptor engagement and evidence of negative regulation of NF-κB was also observed in cluster-9. Both NFKBIA and NFKBIZ are represented on the pathway maps as inhibitors of the NFKB1:RELA (p50:p65) complex. NFKBID currently not represented on the macrophage pathway model (Chapter-2), has not been studied as extensively as the other NF-κB inhibitor proteins. However there is some evidence to suggest a role for NFKBID as a negative regulator of a subset of TLR-dependant genes through inhibition of the NFKB1:RELA (p50:p65) transcription factor [333]. Interestingly, NFKBIA was a gene of interest in the siRNA screens of Chapter-5. It was expected that knock-down of the murine *Nfkbia* gene might perturb transcriptional networks associated with type-I signalling and the type-I component of the LPS response. This was not the case however, and one explanation could be that the transcriptional networks associated with LPS signalling compensated for the lack of *Nfkbia* expression. As the data here demonstrates the magnitude of change in expression of most of the regulated transcripts is far higher in LPS treated cells compared to IFN(-β or –γ) treated.

Seven microRNAs were also induced early (from 1 h) in response to LPS stimulation. microRNAs have a well established role in viral infections [334-335] and have now also been suggested as belonging to the first line of anti-bacterial defence. A number of microRNAs have been shown to mediate the LPS response in both RAW264.7 and primary murine macrophages [336-338]. The microRNAs within this cluster-9 are potentially interesting candidates for further investigation into their functional role in the macrophage response to LPS. One plausible prediction, based on function of other genes within this cluster could be that these microRNAs are involved in the (negative) regulation of signalling pathways downstream of TLR-receptor activation.

The gene content of cluster-5 was comparable to some extent with that of cluster-9 (Figure 6.4a). Cluster-5 also contained inhibitors of NF-κB signalling, (*Nfkbib*, *Nfkbie*), a number of transcripts encoding microRNAs (*Mir146, Mir155, Mir191, Mir425*) and cytokines (*Ccl9, Cxcl1, Cxcl2*, and *Tnf*). However in contrast to cluster-9 where genes were transiently expressed, cluster-5 represented transcripts whose expression was sustained from 1 h to 8 h. Differences in the degradation rate of mRNA has been

proposed to determine some of the temporal changes in RNA levels in mammalian cells [339-342]. In LPS treated dendritic cells changes in transcription rates are shown to determine the majority of temporal changes in RNA levels, however it is changes in degradation rates that tend to shape sharp 'peaked' responses [342]. It is therefore possible that higher mRNA degradation rates are a feature of genes in clusters representing (very) transient expression (e.g. cluster-9). Many of the cluster-5 transcripts were also induced by IFN-β/γ (4-8 h) although not to the same magnitude as LPS. Whilst cluster-9 was highly indicative of (the negative) regulation of TLR4 activated pathways, cluster-5 comprised factors required to execute these pathways; For example members of the NF-κB transcription factor family: *Bcl3, Rel, Rela, Relb*, and type-I interferon signalling (*Ifnb1* and the type-I receptor component *Ifnar1*). Other genes of interest in this cluster included the major macrophage growth and differentiation factor Csf-1 and eight transcripts encoding zinc finger proteins. LPS is known to induce CSF-1 in macrophages [309, 343], and a number of zinc finger proteins have been found to regulate pro-inflammatory activation in macrophages [344-346].

Cluster-3 comprised 380 transcripts up-regulated preferentially in LPS treated macrophages compared to IFN-β or IFN-γ treated macrophages (Figure 6.4a). Maximal expression of cluster-3 transcripts was reached at 8 h post-treatment. GO annotation analysis of these transcripts revealed an over-representation of terms associated with *regulation of phosphorylation*, *lymphocyte activation*, and *cytokine activity*. Indeed, the gene content of this cluster was highly representative of an anti-microbial response which included chemokines (Ccl17, Ccl22, Ccl24, Cxcl1, Cxcl16, Cxcl3), interleukins (Il12a, Il12b, Il17rd, Il18, Il1a, Il12, Il6) and interleukin receptor subunits (Il17rd, Il20rb, Il2ra). Also within this cluster were genes encoding other NF-κB family members (*Nfkb1*, *Nfkb2*) and three MAPKinase signalling components Map3k10, Mapkapk2 and Mapkbp1. Mapkapk2 (MK2), as represented on the pathway diagram is phosphorylated by MAPK14 and is suggested to play a role in the regulation of mRNA stability and pathogen phagocytosis [347] (Figure 6.7).
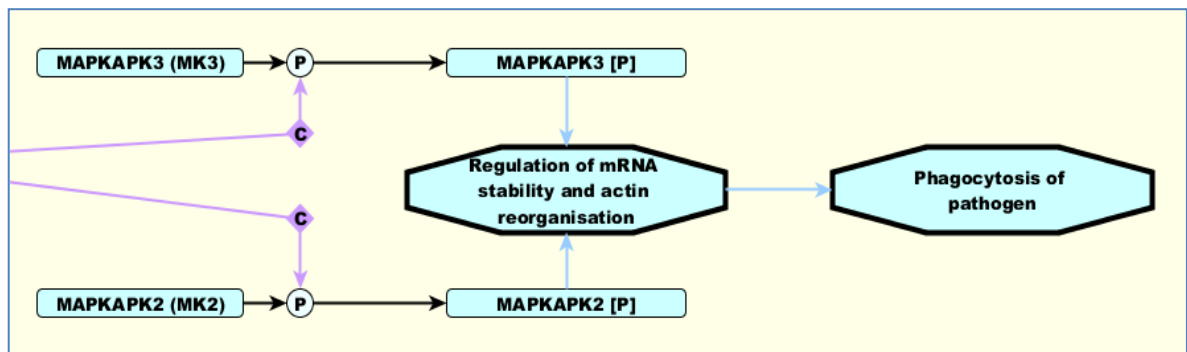
Figure 6.7: Extract from the integrated pathway diagram depicting the action of MAPKAPK2 protein. Phosphorylated MAPKAPK2 (and MAPKAPK3) play a role in the regulation of MRNA stability and actin reorganisation.

Cluster-2 (454 transcripts) represented transcriptional changes occurring specifically in LPS treatment, reaching maximal expression at 24 h post-treatment. GO annotation for this cluster was enriched for terms such as "positive regulation of immune system process" and "macrophage activation". Concurrent with this annotation was the presence of the *Nos2* (nitric oxide synthase) gene, one of the most characteristic inducible markers of (M1) (mouse) macrophage activation [48, 348]. The chemokine Ccl5, and chemokine receptor Ccr7, (present within cluster-2) have been reported as being expressed in M1-polarised macrophages [46]. Also present within this cluster were 15 *solute carrier family member* transcripts and the janus kinases *Jak1* and *Jak2*. The latter two genes are required for type-I and type-II interferon receptor signal transduction. The solute carrier family proteins are predominantly involved in the transport of divalent cations and small organic molecules. They have been associated with immune and inflammatory disease susceptibility [349-350] although a specific role in macrophage activation has yet to be elucidated. Others have found classically activated macrophages are characterised by increased expression of certain solute carrier family members; namely SLC21A15 and SLC31A2 [50] the latter of which was present in cluster-2. This cluster therefore presents a number of other solute carriers that are potential markers of macrophage activation by LPS.

Overlapping Transcriptional Changes

Universal to all three treatments was the repression of genes associated with cell cycle (cluster-1). These genes were repressed to a greater extent in LPS treatment compared to cells treated with IFN-β or IFN-γ (Figure 6.4a). Specific categories of genes included those associated with cell-cycle progression (e.g. Cdc20, Cdc25b, Cdc25c, Cdc45, Cdc7), cyclins (e.g. Ccna2, Ccnb1, Ccnb2, Ccne2), centromere proteins (Cenpa, Cenpe, Cenpf, Cenph, Cenpi), DNA polymerase subunits (e.g. Pola1, Pold1, Pold2, Pole, Pole2) and 49 transcripts encoding histones. Histones constitute half the mass of chromatin, and play a crucial role in DNA packaging, efficient replication and segregation of chromosomes [351].

Cell cycle arrest by LPS has also been linked with the down-modulation of the CSF-1 receptor and some genes are induced as a consequence of the ablation of CSF1R signalling [352]. The gene encoding this receptor (CSF-1R) was not present within clusters in these analyses; however its expression profile revealed it was repressed specifically in LPS treatment and not in response to IFN-β or IFN-γ (Figure 6.8). This might suggest CSF-1R repression is a cell cycle arrest mechanism employed in response to LPS stimulation, but is independent of the anti-mitotic action of IFN-β. This data may also indicate there is possibly a (small) subset of genes repressed in LPS treatment but not by IFN-γ or IFN-β, however such genes have not met the Pearson filtering threshold set for this analysis.
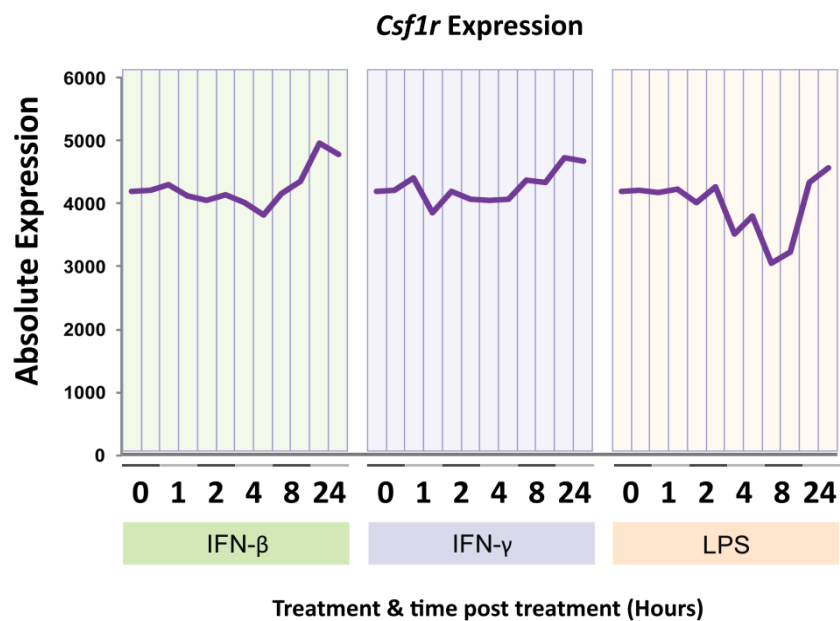
Figure 6.8: *Csf1r* **expression in mouse BMDMs following treatment with any of IFN-β, IFN-γ, or LPS over 24 hours.** Expression of the transcript encoding the Csf1r protein, plotted across three treatment regimes from samples representing pre-treatment (0 h) to specific time-points post treatment (1,2,4,8, and 24h).

Transcripts within cluster-12 and cluster-16 were expressed to a similar extent in all three treatments (Figure 6.4d). The clusters differed marginally in the temporal patterns of their average expression profiles. Present within these clusters were the transcription factors Stat1, Stat3 and Irf5. Irf5 was a gene of interest in the siRNA screens of Chapter-5 and was pursued based on group findings that its siRNA induced repression perturbed transcriptional signatures associated with IFN-β signalling [146]. Recently high Irf5 expression has been suggested as being characteristic of M1 macrophages, which encourage a T helper type 1 (TH1)-TH17 response [353]. The authors of the study demonstrated Irf5 was induced in the presence of GM-CSF [353]. This dataset shows Irf5 mRNA can also be induced by IFN-γ, IFN-β and LPS, further supporting its role in contributing towards an M1-type macrophage. Stat1 which is known to be required for executing both type-I and type-II interferon signalling as well as LPS induced gene expression [354] was induced to an almost identical magnitude and duration in each treatment (Figure 6.9). GO annotation analysis of cluster 16, indicated an over-representation of terms associated with antigen processing and presentation and included genes encoding proteasome subunits (Psme1, Psmb8, Psmb10) and a peptide transporter associated with antigen processing (Tap2).
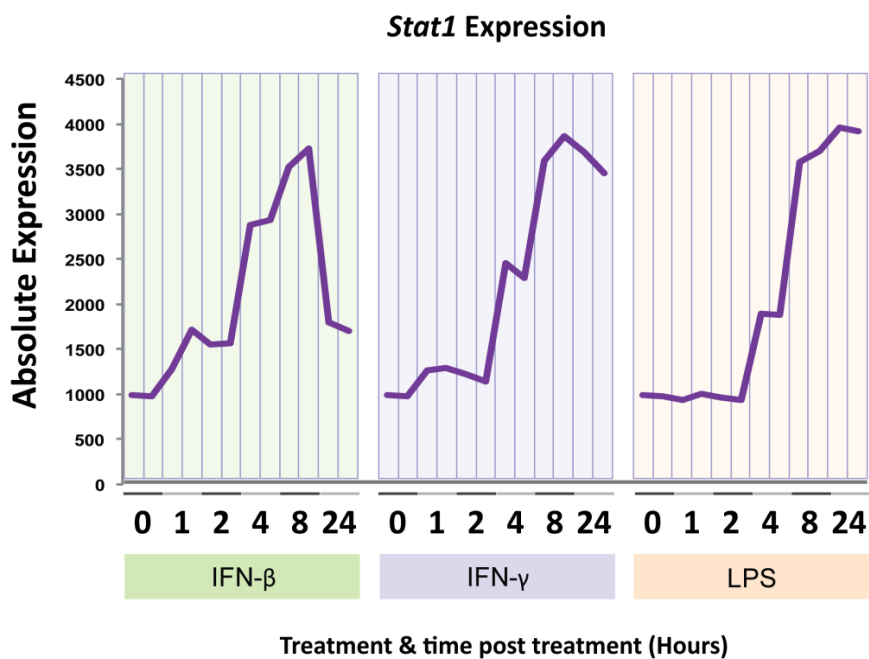
Figure 6.9: *Stat1* **expression in mouse BMDMs following treatment with any of IFN-β, IFN-γ, or LPS over 24 hours.** Expression of the transcript encoding the Stat1 protein, plotted across three treatment regimes from samples representing pre-treatment (0 h) to specific time-points post treatment (1,2,4,8, and 24h).

Cluster-4 comprised 335 transcripts whose expression was preferentially induced in IFN-β and LPS treatment to a greater extent than IFN-γ treated cells (Figure 6.4d). The gene content of this cluster was highly characteristic of a 'classical' type-I interferon response and included a number of interferon inducible proteins (*Ifi202b*, *Ifi204*, *Ifi205*, *Ifih1*, Ifi35, *Ifi44*, *Ifit2*, *Ifit3*), oligoadenylate synthetases (Oas1a, Oas1b, Oas1g, Oas2, Oasl1, Oasl2), the GTPase Mx1 and the viral-RNA detecting endosomal toll-like-receptors, Tlr3 and Tlr8. Based on the expression profiles and gene content, it is tenable that other genes in cluster-4 are regulated via type-I interferon signalling; this included 40 as of yet functionally un-annotated RIKEN cDNA transcripts. Cluster-7 was related to cluster-4 in terms of its expression pattern. Cluster-7 transcripts were preferentially expressed in IFN-β treatment, followed by LPS-treatment and to a lesser extent IFN-γ treatment (Figure 6.4b). The 123 cluster-7 members included the intracellular pattern recognition receptors; Nod1 and Aim2 and the interferon inducible *Ifit1*, and *Mx2*. Nod1 is a member of family of intracellular proteins that mediate host recognition of bacterial peptidoglycan [355-357], and AIM2 is sensor of cytoplasmic double stranded DNA (dsDNA) [358]. Thus the gene content of cluster-7 would suggest a priming of the intracellular pathogen detection systems by IFN-β.

As expected the response to IFN-β and LPS was more overlapping than that of LPS and IFN-γ. There were however a handful of transcripts (in cluster-34) expressed specifically in IFN-γ or LPS treatment but not IFN-β treatment. The most convincing of these LPS/IFN-γ specific genes were MHC Class II antigens (Cd74, H2-Aa, H2-Ab1, H2-Ea) or related proteins (Ctsh (cathepsin H) which encodes a lysosomal cysteine proteinase required for degradation of lysosomal proteins). In macrophages MHC Class II expression is only efficiently induced by IFN-γ, whereas in other cells (such as B-cells) its expression is constitutive. Others have demonstrated LPS increases the expression of MHC II molecules in dendritic cells and B cells [359-360] and this is brought about by enhancing MHC Class II transcription independently of CIITA [361]. In macrophages LPS may enhance or inhibit IFN-γ induced MHC class II depending on the sequence of treatment; simultaneous IFN-γ and LPS treatment is inhibitory, whereas LPS added after IFN-γ augments class II expression [362]. This data-set presents a number of MHC

class II candidates induced (at the message level at least) by LPS in mouse BMDMs independently of IFN-γ.

Transcriptional Changes Specific or Centric to Either IFN-β or IFN-γ Treatment

In addition to cluster-34, MHC Class II genes were also found in cluster 14 (45 transcripts) (Figure 6.4c). This cluster represented transcripts whose expression was most specific to IFN-γ treated cells and included the MHC class II related transcripts; H2-DMa, H2-DMb2, H2-Eb1, Ciita, as well as transcripts encoding complement components; C1qb, C1qc, C1qa, C4b. The full spectrum of 'classical' activation of macrophages is thought to be induced by IFN-γ in concert with a microbial stimulus, such as LPS. The data here shows, apart from a handful of transcripts (including those in cluster 14), LPS is capable of regulating the same targets as IFN-γ. Thus the major contribution of IFN-γ to the classically activated macrophage would appear to be the capacity to efficiently induce MHC Class II antigen presentation and the complement system, which is less recognised.

Cluster-10 and cluster-36 comprised transcripts whose expression was mostly restricted to IFN-β treatment. However the role of the genes (within these clusters) in the IFN-β response is not well characterised, as well the reasons to why they are expressed in IFN-β treatment but not following LPS stimulation. LPS induces IFN-β mRNA and protein expression, and therefore IFN-β signals in an autocrine manner to stimulate the cells and forms a significant portion of the LPS response, as has been previously studied [275-277]. However the existence of genes induced by IFN-β but not by LPS, (i.e. IFN-β targets which are repressed by the actions of LPS), have not been previously studied to date. This repression of cytokine inducible genes could potentially be an interesting aspect of the macrophage response to microbe.

Functional annotation of these IFN-β clusters was poor. Following further inspection of the expression profiles of genes within these clusters, a refined list of most IFN-β specific transcripts was generated (Table 6.3). Amongst these genes was *Casp2*, which

was induced as early as 1 h by IFN-β, yet repressed in response to LPS treatment (Figure 6.10). The LPS response is known to possess an anti-apoptotic component [363][, and the repression of *Casp2* could be an example of an anti-apoptotic mechanism employed in macrophages in response to endotoxin.
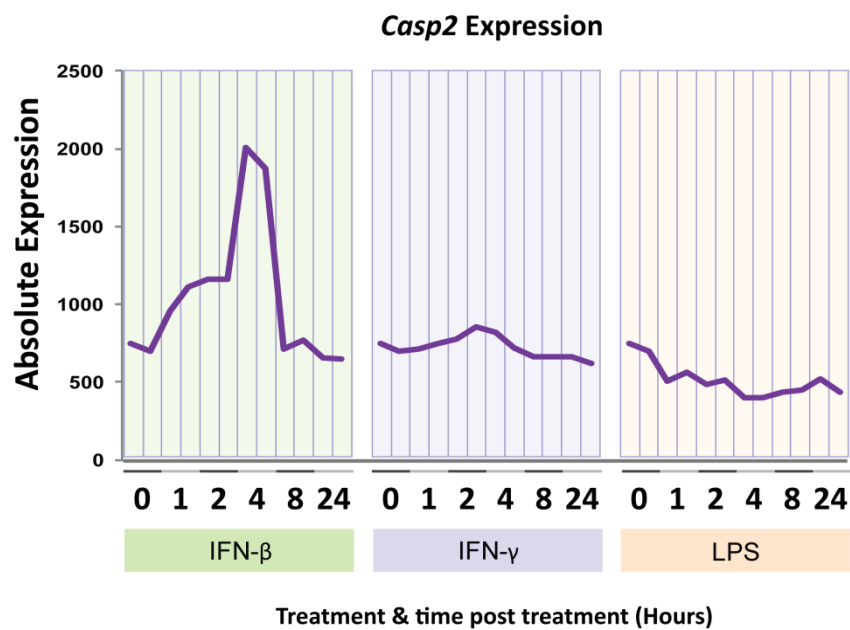
Figure 6.10: *Casp2* **expression in mouse BMDMs following treatment with any of IFN-β, IFN-γ, or LPS over 24 hours.** Expression of the transcript encoding the Casp2 protein, plotted across three treatment regimes from samples representing pre-treatment (0 h) to specific time-points post treatment (1,2,4,8, and 24h).

Gene	Description
Abcb1a	ATP-binding cassette, sub-family B (MDR/TAP), member 1A
Casp2	caspase 2
Ccdc141	coiled-coil domain containing 141
Ccdc39	coiled-coil domain containing 39
Cd4	CD4 antigen
Cdkn1c	cyclin-dependent kinase inhibitor 1C (P57)
Dio2	deiodinase, iodothyronine, type II
Eif1	eukaryotic translation initiation factor 1
Fgfbp3	fibroblast growth factor binding protein 3
Gna14	guanine nucleotide binding protein, alpha 14
Gprc5b	G protein-coupled receptor, family C, group 5, member B
Grap2	GRB2-related adaptor protein 2
Il10	interleukin 10
Klrg2	killer cell lectin-like receptor subfamily G, member 2
Lrrc14b	leucine rich repeat containing 14B
Nsmaf	neutral sphingomyelinase (N-SMase) activation associated factor
Tnfsf8	tumor necrosis factor (ligand) superfamily, member 8

Table 6.3: Gene expressed specifically in IFN-β of mouse BMDMs in a comparison of the transactional response to IFN-β, IFN-γ and LPS. A filtered list of genes expressed exclusively in response to IFN-β treatment of BMDMs, and not in response to IFN-γ or LPS.

Conclusions and Further Work

The analyses of this Chapter set out to explore and better understand the overlap and distinctions in the transcriptional response generated in mouse BMDMs following stimulation with IFN-γ or IFN-β or LPS. The individual response to these three stimuli have been characterised previously and many of the major transcriptional changes found here confirmed the findings of previous studies [97, 238, 243-244, 246-247, 302, 309, 322-326]. However to date no attempts have been made to thoroughly compare and contrast the response to these stimuli over time in primary macrophages. Therefore this data-set provides a valuable source of information to address this question. In these analyses it was found the treatments overlapped to a large degree in their transcriptional profiles but also induced unique sets of genes. LPS induced the greatest proportion of 'unique' transcripts in this comparison and shared a greater deal of overlap with IFN-β treatment, rather than IFN-γ. Surprisingly there were a group of genes induced in IFN-β treatment, yet not induced by LPS (and in some cases even repressed by LPS treatment). The response to LPS is generally assumed to comprise all aspects of the IFN-β response, thus the identification these "IFN-β-exclusive" genes presents potentially interesting candidates for further study into a feature of the LPS response previously unexplored.

The analyses presented in this Chapter are not exhaustive and represent the beginnings of attempts to delineate the shared and unique transcriptional response to the three stimuli. The clusters of co-ordinately expressed transcripts identified here could be further filtered to obtain a more refined list of transcripts preferentially or exclusively expressed in any given treatment or in two or more treatments. The study could also be complemented by sequence analysis of the promoter regions of the clusters of regulated genes to explore whether any particular transcriptional-factor binding sites were over-represented within these groups. It is also important to consider that the findings of this analysis may not fully correlate across species or indeed inter-species in other mice strain. Moreover some aspects of the transcriptional profiles generated in this study may well be specific to the doses studied (10 U/ml IFN-

β/γ and 5 ng/ml LPS) on this occasion. Data in the preceding Chapter-(5), demonstrated the dose-dependant induction of certain LPS inducible genes. In a related analysis (also processed on this 96-array plate but) not discussed as part of thesis the macrophage response to three different doses of LPS was studied. The data revealed cohorts of genes which were clearly induced or repressed in a dose dependant manner, as well as genes unaffected by LPS dose (i.e. induced/repressed to the same extent regardless of treatment dose) (unpublished, data). This underscores another level of complexity when making comparisons across the treatment types, as genes found to be preferentially changing in one treatment over others may only hold true under the specific conditions studied.

The analyses here, as with that of Chapter-4 highlights the fact that the full spectrum of macrophage activation is far from fully captured on our integrated pathway diagram. At the same time the function of many of the regulated components with respect to their role in macrophage activation has yet to be fully elucidated. For example many transcripts encoding microRNAs were found to be in clusters in this analysis, but are not currently represented on the macrophage pathways. However characterisation of the microRNAs themselves as well as their function in macrophage signalling has been an area of intense activity over recent years. This analysis also incorporates many regulated transcripts which are as of yet functionally un-annotated (e.g. Riken cDNA transcripts) but their expression is strongly co-ordinated with genes of know function and/or within clusters of genes contributing to a particular aspect of the macrophage activation response. Thus as with other analyses [143, 145, 364-366] there is the potential here to characterise as of yet un-annotated proteins since the likely function or process in which they are involved can be inferred by their co-expression with genes encoding proteins of known function.

Finally, one of the interests in exploring the overlap between these three stimuli is based on the fact the signalling pathways they activate are either inter-related or synergise at a number of levels. As a means of extending this analyses and further characterising this synergy it would be interesting to compare the individual responses

to these stimuli with transcriptional data relating to the combinatorial stimulation of macrophage with the treatments.

Methods and Materials

Cell culture and treatment

Cell culture is as described previously in Chapter-4 (for the IFN-β study), briefly; bone marrow derived macrophages (BMDM) were prepared from femurs of 7-8 week old male BALB/c mice, by differentiating bone marrow progenitors using the differentiation factor CSF-1. On day six of differentiation cells were harvested from 10 cm square bacteriological plates and seeded into 24-well tissue culture plates at a density of 200,000 – 210,000 cells/ well. 24 hours later, (on day seven) cells were treated with one of; 10 U/ml recombinant mouse interferon-beta (IFN-β) (PBL Interferon Source, New Jersey, USA), 10 U/ml interferon-gamma (IFN-γ) (Perbio Science, Northumberland, UK), or 5 ng/ml LPS (from Salmonella minnesota Re595 (Sigma, Gillingham, UK)) and harvested 1, 2, 4, 8 and 24 h following treatment or collected pre-treatment (0 h). All treatments were performed in the presence of CSF-1 since CSF-1 is constitutively present *in vivo*. Moreover CSF-1 is itself induced upon macrophage activation with LPS and has been shown to enhance the activation of some genes by LPS [367]. It has therefore been argued that in vitro studies of macrophage activation should be performed in the presence of CSF-1 [283].

RNA extraction, QC and labelling for arrays

Total RNA was harvested from the cells using an RNeasy Plus kit (Qiagen, Crawley, UK) according to manufacturer's instructions. RNA was quantified and quality controlled using a NanoDrop spectrophotometer (Nano-Drop Technologies, Delaware, USA) and BioAnalyser 2100 (Agilent, California, USA) to determine RNA purity and integrity. Replicate 250 ng samples of total RNA derived from two separate wells per time point were first processed using the Ambion WT (whole transcript) Expression Kit (Ambion)

to generate amplified and biotinylated sense strand DNA targets from the entire genome without bias. The senses strand DNA samples were then labelled, and hybridized to the Affymetrix® Mouse Gene 1.1 ST (obtained as part of a high-throughput 96-array plate) using the Affymetrix GeneChip WT terminal labelling and hybridisation kit (Affymetrix) and according to manufacturer's protocol. The Affymetrix® Mouse Gene 1.1 ST Array Plate enables the parallel processing and expression profiling of 96 samples. The individual arrays interrogate more than 28,000 well-annotated genes with more than 770,000 distinct probes. Hybridisation, washing and scanning of the 96 arrays was performed using the Affymetrix GeneTitan instrument; this parallel and high-throughput processing reduces the chances of introducing technical variability that may well arise when processing the same number of samples of separate occasions.

Data processing and network analysis

Data (to be submitted to the GEO repository) was normalized using the RMA package within the Affymetrix Expression Console software and annotated. Network analysis of the normalised expression data was performed using BioLayout Express according to the network analysis principles described in Chapter-1 and as described in the Methods of Chapter-4. The parameters employed in this analysis included an initial filtering step of expression data relating to all the probes (transcripts) on the array) at a 0.85 Pearson correlation cut-off threshold. This network correlation graph of probes falling within this threshold (11,258 nodes, connected by 270,601 edges) was then filtered at an MCL inflation value of 2.2, to cluster the graph (and resulting in over 600 clusters with at least ≥3 nodes). Clusters related to technical artefacts or patterns of expression unchanging across the 32 arrays were eliminated. A further network graph relating only to the data from probes within 'interesting' clusters of the filtered graph (3,747 nodes) was generated (by filtering relationships at a Pearson correlation of 0.85, and clustering the consequential graph at an MCL inflation value of 2.2). The clusters were inspected for patterns of expression associated with treatment over the time-courses and gene lists associated with clusters were exported for GO annotation

analysis (Biological Processes Level-FAT) using the DAVID (Database for Annotation, Visualization and Integrated Discovery) tool.

Chapter Contributions and Acknowledgements

This page has been left intentionally blank

Chapter 7. Conclusions

Overview of main findings, challenges and future work

The work described in this thesis explores the use of systems-level modelling and analysis methods to the study of macrophage activation and associated pathways. Specifically, the work outlined in Chapters 2 and 3 aimed to address the challenge of accurately modelling the large number of known interacting components and biological processes regulated in this cell. Secondly, the interest was in investigating the transcriptional signatures generated in response to IFNs and what we understand by the classical-activation of macrophages (Chapters 4 and 6). Moreover was the desire to examine the role of genes of interest in the type-I IFN response pathway, by targeting their expression with siRNA. The exceptional plasticity and highly attuned pattern recognition systems of primary macrophages can complicate the use and reproducibility of functional genomics screening using siRNA, and therefore attempts were made to address this issue during the development of a cell-based assay (Chapter-5).

Chapter-2 described the construction of pathway models of macrophage signalling. An initial framework map of macrophage activation comprising 272 interactions was published in 2008. During its construction process many challenges of converting pathway knowledge into computationally-tractable yet 'understandable' diagrams, were addressed. The lessons learnt and rules established from constructing the framework map then provided a more robust agenda for assembling further pathways of interest, and the subsequent integrated pathway resource included a total of 496 unique proteins, 412 protein complexes, 81 genes, and 101 DNA sequence/promoter regions. In total the network comprised 2,170 nodes connected by 2,553 edges. This model of macrophage signalling is to our knowledge the most comprehensive pathway of its kind published to date. The work summarizes years of investigations, and brings

together thousands of published findings into one navigable, searchable, visual aid. The pathway has proven to be a valuable resource in the interpretation of functional genomics data. At the same time it is appreciated the model still only captures a modest number of changes occurring in the macrophage, as was evident from the transcriptional analyses performed in Chapters 4 and 6. Future efforts will need to integrate the transcriptional data as well as other signalling pathways known to be active in the macrophage, into the overall signalling network to improve the predictive value of the pathway diagrams. The longer term eventual aim will be to expand the pathway resource to create a more comprehensive *in silico* cell model of macrophage signalling with the predictive power to demonstrate the transcriptional networks activated in response to combinations of stimuli (pathogens and/or cytokines) as encountered *in vivo*.

The desire to perform pathway analysis on the outputs of functional genomics screens and high-dimensional data has arguably never been greater. However the overhead of creating and maintaining pathway resources is often a thankless and onerous task. Recently (21 May 2011) it was announced the popular and previously free pathway repository KEGG pathways is to move away from its open-source/free-software roots and change to a paid-subscription system. Therefore in the interests of maintaining the availability of high-quality free-pathway resources to enhance academic research, the construction of such in-house pathways as described in Chapter-2 as well as keeping up the debate on how pathway information should be depicted and exchanged (Chapter-3) is more important than ever.

Chapter-3 describes the development of the modified Edinburgh Pathway Notation (mEPN) scheme a graphical notation system for biology originally devised a number of years ago and now through use has been refined extensively. This development has been primarily driven by the attempts to produce process diagrams for a diverse range of biological pathways (described in Chapter-2). In addition to the pathways described in Chapter-2 the mEPN notation scheme has now been deployed to depict cell-cycle signalling as well as lipid-metabolism (unpublished). Through thorough testing and

refinement the mEPN has been found to be suitable to describe a number of biological concepts. The next challenge for both the mEPN and in a more general sense for systems biology is how we begin to model (*in silico*) the cellular interactions underpinning a biological process (*in vivo*) and their deregulation in disease states. With respect to the macrophage, its behaviour, phenotypic and genotypic properties vary depending on its surrounding cellular milieu. Thus to better understand the highly plastic nature of macrophages requires an understanding of its interactions with other cells in a given setting. Indeed, systems biology involves the study of biological processes at multiple levels of abstraction e.g. from molecules, to cells, to tissues, to organs and so on. The challenge remains in bridging the insights acquired from multiple levels and devising models to reflect multi-level concepts and observations. One of the key incentives for generating pathways with standard notations was to permit the conversion of graphical models into computationally tractable ones, suitable for simulation analyses. Efforts are now well advanced in the Group to convert the mEPN notation into language suitable of stochastic flow modelling (signalling Petri nets (SPN) [114]). This has required further adaptation of the notation scheme and development of supporting software (BioLayout *Express*³ᴰ) to visualise the output. Initial tests have been promising and the SPN approach is particularly well suited to large networks such as those generated in Chapter-2.

Thus we are now closer to a pathway-model system where one can begin to perform computational predictions about pathway behaviour and the signalling response to pathway perturbation. In the latter case this could be how macrophage signalling responds to perturbation induced by siRNA gene knockdown. The potential for computational flow-modelling would not have been possible without first the development of the mEPN and the construction of extensive pathway models on which to perform/test the simulations. The longer term but ultimate objective in systems-based therapeutics is to exploit the predictive power of these *in silico* models to identify novel drug targets.

The investigations of Chapter-4 set out to analyse the type-I and type-II interferon response in macrophages. Previous attempts to analyse the transcriptional response to these cytokines have predominantly relied on statistical fold-change cut offs. In this work the application of network-based explorations of correlation matrices were used to visualise and interpret the transcriptional events over a 24 h time-course. The IFN-β and IFN-γ time-course experiments were originally conceived and analysed separately. Later in the course of these investigations, the type-I and type-II time-course experiments were repeated under controlled conditions to allow for a more accurate comparison (Chapter-6). The clustering approach used to analyse the expression data provides a powerful method for identifying correlation amongst genes and from which the function of currently un-annotated genes may be inferred. The next challenge lies in unravelling the causality of the regulatory relationships in the transcriptional networks. This would require identifying negative and positive feedback loops in the response networks. Some feedback loop are very well characterised (e.g. *CIITA* regulation of MHC-Class II expression), whereas many other causal relationships remain to be identified.

The studies of Chapter-5 set out to optimise an *in vitro* cell-based assay for investigating the role of selected genes in the type-I interferon and LPS response. The genes of interest were targeted with siRNA. Genome-wide transcriptional analysis revealed that knockdown of genes of interest did not perturb transcriptional networks associated with the LPS response, including the expression of type-I response genes. On one level these observations may well have be indicative of the robustness of macrophage signalling networks, especially in response to a potent microbial stimulus (LPS). Future work could explore whether this robustness is impaired with combinatorial gene knockdowns (i.e. two or more genes targeted).

The entire assay optimisation process also highlighted the complexities of using siRNA in primary macrophages. For example, increasing doses of different siRNAs generated different patterns of lipofection-induced type-I gene expression. siRNA pre-treatment would on some occasions enhance responsiveness to LPS and on other occasions

decrease responsiveness. The lack of reproducibility observed between investigations might be attributed to the innate plasticity of macrophages or variability in BMDMs prepared on different days. Ultimately targeting a heterogeneous cell type like the macrophage can generate a spectrum of responses depending on the collective state of the cells. Given their roles in numerous diseases, macrophages are attractive targets for siRNA based therapeutics. The observations of this Chapter mirror the obstacles faced in utilizing siRNA in the clinical setting.

The final investigations described in this thesis set out to explore the transcriptional networks associated with the BMDM response to type-I interferon, type-II interferon and LPS stimulation. The investigations of Chapter-4 also attempted to contrast the type-I and type-II response in primary BMDMs, however this comparison was limited by the fact the experiments and microarrays were processed on separate occasions. Nevertheless Chapter-4 demonstrated the power of the network-based analysis approach for mining high-throughput transciptomics data, and this approach was again utilized in the analyses of Chapter-6. The overlapping and unique transcriptional responses to the three stimuli (LPS, IFN-β, and IFN-γ) were identified. Generally the inducible transcriptional response was much broader and potent (in terms of fold-change induction) in LPS treated macrophages compared to those treated with either IFN-β or IFN-γ. Moreover under these experimental conditions 727 transcripts were specifically regulated in LPS treatment. 50 transcripts were unique to IFN-γ treatment and 115 transcripts were unique to IFN-β treatment. In concert with previous findings, MHC Class II related transcripts were among those specifically expressed in response to IFN-γ, as well as members of the complement system. The role of genes expressed specifically in IFN-β, is less well characterised. Interestingly some genes were induced by IFN-β, yet repressed by LPS treatment. The identification of "IFN-β-exclusive" genes presents potentially interesting candidates for further study into a feature of the LPS response previously unexplored. The output of the transcriptional screens in Chapters 4 and 6, underscored the fact that the full spectrum of macrophage activation has yet to be fully captured on the integrated pathway diagram. The next steps in developing the macrophage pathway model will be its integration with transcriptional networks

identified in Chapters 4 and 6. Further work would also seek to build transcription factor networks, as a means of integrating the transcriptional data generated with the pathway model.

Overall each of the aspects of the systems biology paradigm of *model-manipulate-measure and mine*, have been explored in some sense in these investigations. The model in this case was the generation of *in silico* macrophage signalling pathways. The pathway resource then informed the stages of manipulation; by cytokine treatment or by gene knockdown using siRNA. Changes occurring in the macrophage were measured using appropriate techniques. Central to these investigations was measurement of transcriptional changes using microarray technology. The resulting data was then mined, and its interpretation can now inform further improvements and development of the macrophage pathways. For example developments could include the incorporation of feed-forward and negative feedback loops arising is response to specific treatments. Even with the integration of the additional information, it's imperative to acknowledge the *in silico* model of macrophage signalling, is purely a *model*. Iterative cycles of development will continue to enhance the value of the model to inform further investigations. Eventually the hope is that such models will have the predictive power to identify targets for therapeutic intervention in diseases underpinned by macrophage activity.

This page has been left intentionally blank

Bibliography

1.   Metchnikoff E, Gourko H, Williamson DI, Tauber AI: *The evolutionary-biology papers of Elie Metchnikoff.* Dordrecht ; Boston: Kluwer Academic Publishers; 2000.
2.   Tauber AI: **Metchnikoff and the phagocytosis theory.** *Nat Rev Mol Cell Biol* 2003, **4:**897-901.
3.   Rosenberger CM, Finlay BB: **Phagocyte sabotage: disruption of macrophage signalling by bacterial pathogens.** *Nat Rev Mol Cell Biol* 2003, **4:**385-396.
4.   Sasmono RT, Oceandy D, Pollard JW, Tong W, Pavli P, Wainwright BJ, Ostrowski MC, Himes SR, Hume DA: **A macrophage colony-stimulating factor receptor-green fluorescent protein transgene is expressed throughout the mononuclear phagocyte system of the mouse.** *Blood* 2003, **101:**1155-1163.
5.   Hume DA, Perry VH, Gordon S: **The mononuclear phagocyte system of the mouse defined by immunohistochemical localisation of antigen F4/80: macrophages associated with epithelia.** *Anat Rec* 1984, **210:**503-512.
6.   Hume DA, Gordon S: **Mononuclear phagocyte system of the mouse defined by immunohistochemical localization of antigen F4/80. Identification of resident macrophages in renal medullary and cortical interstitium and the juxtaglomerular complex.** *J Exp Med* 1983, **157:**1704-1709.
7.   Hume DA, Halpin D, Charlton H, Gordon S: **The mononuclear phagocyte system of the mouse defined by immunohistochemical localization of antigen F4/80: macrophages of endocrine organs.** *Proc Natl Acad Sci U S A* 1984, **81:**4174-4177.
8.   Hume DA, Loutit JF, Gordon S: **The mononuclear phagocyte system of the mouse defined by immunohistochemical localization of antigen F4/80: macrophages of bone and associated connective tissue.** *J Cell Sci* 1984, **66:**189-194.
9.   Hume DA, Robinson AP, MacPherson GG, Gordon S: **The mononuclear phagocyte system of the mouse defined by immunohistochemical localization of antigen F4/80. Relationship between macrophages, Langerhans cells, reticular cells, and dendritic cells in lymphoid and hematopoietic organs.** *J Exp Med* 1983, **158:**1522-1536.
10.   Hume DA: **The mononuclear phagocyte system.** *Curr Opin Immunol* 2006, **18:**49-53.
11.   Gordon S, Taylor PR: **Monocyte and macrophage heterogeneity.** *Nat Rev Immunol* 2005, **5:**953-964.
12.   Lichanska AM, Hume DA: **Origins and functions of phagocytes in the embryo.** *Exp Hematol* 2000, **28:**601-611.
13.   Takahashi K, Naito M, Takeya M: **Development and heterogeneity of macrophages and their related cells through their differentiation pathways.** *Pathol Int* 1996, **46:**473-485.
14.   Jenkins SJ, Ruckerl D, Cook PC, Jones LH, Finkelman FD, van Rooijen N, MacDonald AS, Allen JE: **Local macrophage proliferation, rather than recruitment from the blood, is a signature of TH2 inflammation.** *Science* 2011, **332:**1284-1288.
15.   Geissmann F, Manz MG, Jung S, Sieweke MH, Merad M, Ley K: **Development of monocytes, macrophages, and dendritic cells.** *Science* 2010, **327:**656-661.
16.   Araki H, Katayama N, Yamashita Y, Mano H, Fujieda A, Usui E, Mitani H, Ohishi K, Nishii K, Masuya M, et al: **Reprogramming of human postmitotic neutrophils into macrophages by growth factors.** *Blood* 2004, **103:**2973-2980.
17.   Lindemann SW, Yost CC, Denis MM, McIntyre TM, Weyrich AS, Zimmerman GA: **Neutrophils alter the inflammatory milieu by signal-dependent translation of constitutive messenger RNAs.** *Proc Natl Acad Sci U S A* 2004, **101:**7076-7081.

18.   Hume DA: **Applications of myeloid-specific promoters in transgenic mice support in vivo imaging and functional genomics but do not support the concept of distinct macrophage and dendritic cell lineages or roles in immunity.** *J Leukoc Biol* 2011, **89:**525-538.

19.   Hume DA: **Macrophages as APC and the dendritic cell myth.** *J Immunol* 2008, **181:**5829-5835.

20.   Geissmann F, Gordon S, Hume DA, Mowat AM, Randolph GJ: **Unravelling mononuclear phagocyte heterogeneity.** *Nat Rev Immunol* 2010, **10:**453-460.

21.   MacDonald KP, Rowe V, Bofinger HM, Thomas R, Sasmono T, Hume DA, Hill GR: **The colony-stimulating factor 1 receptor is expressed on dendritic cells during differentiation and regulates their expansion.** *J Immunol* 2005, **175:**1399-1405.

22.   Pozzi LA, Maciaszek JW, Rock KL: **Both dendritic cells and macrophages can stimulate naive CD8 T cells in vivo to proliferate, develop effector function, and differentiate into memory cells.** *J Immunol* 2005, **175:**2071-2081.

23.   Mosser DM, Edwards JP: **Exploring the full spectrum of macrophage activation.** *Nat Rev Immunol* 2008, **8:**958-969.

24.   Fortier AH, Falk LA: **Isolation of murine macrophages.** *Curr Protoc Immunol* 2001, **Chapter 14:**Unit 14 11.

25.   Pollard JW: **Trophic macrophages in development and disease.** *Nat Rev Immunol* 2009, **9:**259-270.

26.   Pollard JW, Dominguez MG, Mocci S, Cohen PE, Stanley ER: **Effect of the colony-stimulating factor-1 null mutation, osteopetrotic (csfm(op)), on the distribution of macrophages in the male mouse reproductive tract.** *Biol Reprod* 1997, **56:**1290-1300.

27.   Marks SC, Jr., Wojtowicz A, Szperl M, Urbanowska E, MacKay CA, Wiktor-Jedrzejczak W, Stanley ER, Aukerman SL: **Administration of colony stimulating factor-1 corrects some macrophage, dental, and skeletal defects in an osteopetrotic mutation (toothless, tl) in the rat.** *Bone* 1992, **13:**89-93.

28.   Lin H, Lee E, Hestir K, Leo C, Huang M, Bosch E, Halenbeck R, Wu G, Zhou A, Behrens D, et al: **Discovery of a cytokine and its receptor by functional screening of the extracellular proteome.** *Science* 2008, **320:**807-811.

29.   Geissmann F, Jung S, Littman DR: **Blood monocytes consist of two principal subsets with distinct migratory properties.** *Immunity* 2003, **19:**71-82.

30.   Gordon S: **The macrophage: past, present and future.** *Eur J Immunol* 2007, **37 Suppl 1:**S9-17.

31.   Raz E: **Organ-specific regulation of innate immunity.** *Nat Immunol* 2007, **8:**3-4.

32.   Smythies LE, Sellers M, Clements RH, Mosteller-Barnum M, Meng G, Benjamin WH, Orenstein JM, Smith PD: **Human intestinal macrophages display profound inflammatory anergy despite avid phagocytic and bacteriocidal activity.** *J Clin Invest* 2005, **115:**66-75.

33.   Mantovani A, Bottazzi B, Sozzani S, Peri G, Allavena P, Dong QG, Vecchi A, Colotta F: **Cytokine regulation of tumour-associated macrophages.** *Res Immunol* 1993, **144:**280-283; discussion 294-288.

34.   Mantovani A, Schioppa T, Porta C, Allavena P, Sica A: **Role of tumor-associated macrophages in tumor progression and invasion.** *Cancer Metastasis Rev* 2006, **25:**315-322.

35.   Bingle L, Lewis CE, Corke KP, Reed MW, Brown NJ: **Macrophages promote angiogenesis in human breast tumour spheroids in vivo.** *Br J Cancer* 2006, **94:**101-107.

36.   Bingle L, Brown NJ, Lewis CE: **The role of tumour-associated macrophages in tumour progression: implications for new anticancer therapies.** *J Pathol* 2002, **196:**254-265.

37.   Kinne RW, Brauer R, Stuhlmuller B, Palombo-Kinne E, Burmester GR: **Macrophages in rheumatoid arthritis.** *Arthritis Res* 2000, **2:**189-202.

38.   Mulherin D, Fitzgerald O, Bresnihan B: **Synovial tissue macrophage populations and articular damage in rheumatoid arthritis.** *Arthritis Rheum* 1996, **39:**115-124.

39.   Adams DO: **Molecular interactions in macrophage activation.** *Immunol Today* 1989, **10:**33-35.

40.   Pluddemann A, Mukhopadhyay S, Gordon S: **Innate immunity to intracellular pathogens: macrophage receptors and responses to microbial entry.** *Immunol Rev* 2011, **240:**11-24.

41.   Taylor PR, Martinez-Pomares L, Stacey M, Lin HH, Brown GD, Gordon S: **Macrophage receptors and immune recognition.** *Annu Rev Immunol* 2005, **23:**901-944.

42.   Medzhitov R, Horng T: **Transcriptional control of the inflammatory response.** *Nat Rev Immunol* 2009, **9:**692-703.

43.   Mills CD, Kincaid K, Alt JM, Heilman MJ, Hill AM: **M-1/M-2 macrophages and the Th1/Th2 paradigm.** *J Immunol* 2000, **164:**6166-6173.

44.   Mosser DM: **The many faces of macrophage activation.** *J Leukoc Biol* 2003, **73:**209-212.

45.   Mantovani A, Sica A, Locati M: **Macrophage polarization comes of age.** *Immunity* 2005, **23:**344-346.

46.   Mantovani A, Sica A, Sozzani S, Allavena P, Vecchi A, Locati M: **The chemokine system in diverse forms of macrophage activation and polarization.** *Trends Immunol* 2004, **25:**677-686.

47.   Stein M, Keshav S, Harris N, Gordon S: **Interleukin 4 potently enhances murine macrophage mannose receptor activity: a marker of alternative immunologic macrophage activation.** *J Exp Med* 1992, **176:**287-292.

48.   Gordon S: **Alternative activation of macrophages.** *Nat Rev Immunol* 2003, **3:**23-35.

49.   Gordon S, Martinez FO: **Alternative activation of macrophages: mechanism and functions.** *Immunity* 2010, **32:**593-604.

50.   Martinez FO, Gordon S, Locati M, Mantovani A: **Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: new molecules and patterns of gene expression.** *J Immunol* 2006, **177:**7303-7311.

51.   Biswas SK, Mantovani A: **Macrophage plasticity and interaction with lymphocyte subsets: cancer as a paradigm.** *Nat Immunol* 2010, **11:**889-896.

52.   Chan G, Bivins-Smith ER, Smith MS, Smith PM, Yurochko AD: **Transcriptome analysis reveals human cytomegalovirus reprograms monocyte differentiation toward an M1 macrophage.** *J Immunol* 2008, **181:**698-711.

53.   Martinez FO, Sica A, Mantovani A, Locati M: **Macrophage activation and polarization.** *Front Biosci* 2008, **13:**453-461.

54.   Borden EC, Sen GC, Uze G, Silverman RH, Ransohoff RM, Foster GR, Stark GR: **Interferons at age 50: past, current and future impact on biomedicine.** *Nat Rev Drug Discov* 2007, **6:**975-990.

55.   Jacobs L, O'Malley J, Freeman A, Ekes R: **Intrathecal interferon reduces exacerbations of multiple sclerosis.** *Science* 1981, **214:**1026-1028.

56.   Banchereau J, Pascual V: **Type I interferon in systemic lupus erythematosus and other autoimmune diseases.** *Immunity* 2006, **25:**383-392.

57.   Bennett L, Palucka AK, Arce E, Cantrell V, Borvak J, Banchereau J, Pascual V: **Interferon and granulopoiesis signatures in systemic lupus erythematosus blood.** *J Exp Med* 2003, **197:**711-723.

58.   Platanias LC: **Mechanisms of type-I- and type-II-interferon-mediated signalling.** *Nat Rev Immunol* 2005, **5:**375-386.

59. Schroder K, Hertzog PJ, Ravasi T, Hume DA: **Interferon-gamma: an overview of signals, mechanisms and functions.** *J Leukoc Biol* 2004, **75:**163-189.

60. Mach B, Steimle V, Martinez-Soria E, Reith W: **Regulation of MHC class II genes: lessons from a disease.** *Annu Rev Immunol* 1996, **14:**301-331.

61. Briscoe J, Rogers NC, Witthuhn BA, Watling D, Harpur AG, Wilks AF, Stark GR, Ihle JN, Kerr IM: **Kinase-negative mutants of JAK1 can sustain interferon-gamma-inducible gene expression but not an antiviral state.** *EMBO J* 1996, **15:**799-809.

62. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, et al: **Initial sequencing and analysis of the human genome.** *Nature* 2001, **409:**860-921.

63. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, et al: **The sequence of the human genome.** *Science* 2001, **291:**1304-1351.

64. Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T: **Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells.** *Nature* 2001, **411:**494-498.

65. Siomi H, Siomi MC: **On the road to reading the RNA-interference code.** *Nature* 2009, **457:**396-404.

66. Meister G, Tuschl T: **Mechanisms of gene silencing by double-stranded RNA.** *Nature* 2004, **431:**343-349.

67. Hutvagner G, Simard MJ: **Argonaute proteins: key players in RNA silencing.** *Nat Rev Mol Cell Biol* 2008, **9:**22-32.

68. Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song JJ, Hammond SM, Joshua-Tor L, Hannon GJ: **Argonaute2 is the catalytic engine of mammalian RNAi.** *Science* 2004, **305:**1437-1441.

69. Song JJ, Smith SK, Hannon GJ, Joshua-Tor L: **Crystal structure of Argonaute and its implications for RISC slicer activity.** *Science* 2004, **305:**1434-1437.

70. Paddison PJ, Caudy AA, Bernstein E, Hannon GJ, Conklin DS: **Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells.** *Genes Dev* 2002, **16:**948-958.

71. Paddison PJ, Caudy AA, Sachidanandam R, Hannon GJ: **Short hairpin activated gene silencing in mammalian cells.** *Methods Mol Biol* 2004, **265:**85-100.

72. Rao DD, Vorhies JS, Senzer N, Nemunaitis J: **siRNA vs. shRNA: similarities and differences.** *Adv Drug Deliv Rev* 2009, **61:**746-759.

73. Bumcrot D, Manoharan M, Koteliansky V, Sah DW: **RNAi therapeutics: a potential new class of pharmaceutical drugs.** *Nat Chem Biol* 2006, **2:**711-719.

74. Robinson R: **RNAi therapeutics: how likely, how soon?** *PLoS Biol* 2004, **2:**E28.

75. Aagaard L, Rossi JJ: **RNAi therapeutics: principles, prospects and challenges.** *Adv Drug Deliv Rev* 2007, **59:**75-86.

76. Sioud M, Sorensen DR: **Cationic liposome-mediated delivery of siRNAs in adult mice.** *Biochem Biophys Res Commun* 2003, **312:**1220-1225.

77. Sledz CA, Holko M, de Veer MJ, Silverman RH, Williams BR: **Activation of the interferon system by short-interfering RNAs.** *Nat Cell Biol* 2003, **5:**834-839.

78. Jackson AL, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, Li B, Cavet G, Linsley PS: **Expression profiling reveals off-target gene regulation by RNAi.** *Nat Biotechnol* 2003, **21:**635-637.

79. Semizarov D, Frost L, Sarthy A, Kroeger P, Halbert DN, Fesik SW: **Specificity of short interfering RNA determined through gene expression signatures.** *Proc Natl Acad Sci U S A* 2003, **100:**6347-6352.

80. Patzel V: **In silico selection of active siRNA.** *Drug Discov Today* 2007, **12:**139-148.

81. Tilesi F, Fradiani P, Socci V, Willems D, Ascenzioni F: **Design and validation of siRNAs and shRNAs.** *Curr Opin Mol Ther* 2009, **11:**156-164.

82. Marques JT, Devosse T, Wang D, Zamanian-Daryoush M, Serbinowski P, Hartmann R, Fujita T, Behlke MA, Williams BR: **A structural basis for discriminating between self and nonself double-stranded RNAs in mammalian cells.** *Nat Biotechnol* 2006, **24:**559-565.

83. Sioud M: **On the delivery of small interfering RNAs into mammalian cells.** *Expert Opin Drug Deliv* 2005, **2:**639-651.

84. Birmingham A, Anderson E, Sullivan K, Reynolds A, Boese Q, Leake D, Karpilow J, Khvorova A: **A protocol for designing siRNAs with high functionality and specificity.** *Nat Protoc* 2007, **2:**2068-2078.

85. Fedorov Y, Anderson EM, Birmingham A, Reynolds A, Karpilow J, Robinson K, Leake D, Marshall WS, Khvorova A: **Off-target effects by siRNA can induce toxic phenotype.** *RNA* 2006, **12:**1188-1196.

86. Bilitewski U: **DNA microarrays: an introduction to the technology.** *Methods Mol Biol* 2009, **509:**1-14.

87. Schena M, Shalon D, Davis RW, Brown PO: **Quantitative monitoring of gene expression patterns with a complementary DNA microarray.** *Science* 1995, **270:**467-470.

88. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, et al: **Minimum information about a microarray experiment (MIAME)-toward standards for microarray data.** *Nat Genet* 2001, **29:**365-371.

89. Shendure J: **The beginning of the end for microarrays?** *Nat Methods* 2008, **5:**585-587.

90. Wang Z, Gerstein M, Snyder M: **RNA-Seq: a revolutionary tool for transcriptomics.** *Nat Rev Genet* 2009, **10:**57-63.

91. Aderem A: **Systems biology: its practice and challenges.** *Cell* 2005, **121:**511-513.

92. Gardy JL, Lynn DJ, Brinkman FS, Hancock RE: **Enabling a systems biology approach to immunology: focus on innate immunity.** *Trends Immunol* 2009, **30:**249-262.

93. Hertzog P, Forster S, Samarajiwa S: **Systems biology of interferon responses.** *J Interferon Cytokine Res* 2011, **31:**5-11.

94. Gilman AG, Simon MI, Bourne HR, Harris BA, Long R, Ross EM, Stull JT, Taussig R, Arkin AP, Cobb MH, et al: **Overview of the Alliance for Cellular Signaling.** *Nature* 2002, **420:**703-706.

95. Fahy E, Sud M, Cotter D, Subramaniam S: **LIPID MAPS online tools for lipid research.** *Nucleic Acids Res* 2007, **35:**W606-612.

96. Cavalieri D, Rivero D, Beltrame L, Buschow SI, Calura E, Rizzetto L, Gessani S, Gauzzi MC, Reith W, Baur A, et al: **DC-ATLAS: a systems biology resource to dissect receptor specific signal transduction in dendritic cells.** *Immunome Res* 2010, **6:**10.

97. Ravasi T, Wells CA, Hume DA: **Systems biology of transcription control in macrophages.** *Bioessays* 2007, **29:**1215-1226.

98. Calzone L, Gelay A, Zinovyev A, Radvanyi F, Barillot E: **A comprehensive modular map of molecular interactions in RB/E2F pathway.** *Mol Syst Biol* 2008, **4:**173.

99. Oda K, Kimura T, Matsuoka Y, Funahashi A, M. M, Kitano H: **Molecular Interaction Map of a Macrophage.** In *The Alliance for Cellular Signaling (AfCS) Research Reports*, vol. 2; 2004.

100. Oda K, Kitano H: **A comprehensive map of the toll-like receptor signaling network.** *Mol Syst Biol* 2006, **2:**2006 0015.

101. Oda K, Matsuoka Y, Funahashi A, Kitano H: **A comprehensive pathway map of epidermal growth factor receptor signaling.** *Mol Syst Biol* 2005, **1:**2005 0010.

102. Raza S, Robertson KA, Lacaze PA, Page D, Enright AJ, Ghazal P, Freeman TC: **A logic-based diagram of signalling pathways central to macrophage activation.** *BMC Syst Biol* 2008, **2:**36.

103. Kitano H, Funahashi A, Matsuoka Y, Oda K: **Using process diagrams for the graphical representation of biological networks.** *Nat Biotechnol* 2005, **23:**961-966.

104. Kohn KW: **Molecular interaction map of the mammalian cell cycle control and DNA repair systems.** *Mol Biol Cell* 1999, **10:**2703-2734.

105. Moodie SL, Sorokin A, Goryanin I, Ghazal P: **A Graphical Notation to Describe the Logical Interactions of Biological Pathways.** *Journal of Integrative Bioinformatics* 2006, **3:**11.

106. Novere NL, Hucka M, Mi H, Moodie S, Schreiber F, Sorokin A, Demir E, Wegner K, Aladjem MI, Wimalaratne SM, et al: **The systems biology graphical notation.** *Nat Biotechnol* 2009, **27:**735-741.

107. Funahashi A, Matsuoka Y, Jouraku A, Morohashi M, Kikuchi N, Kitano H: **CellDesigner 3.5: A Versatile Modeling Tool for Biochemical Networks.** *Proceedings of the IEEE* 2008, **96:**1254-1265.

108. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: a software environment for integrated models of biomolecular interaction networks.** *Genome Res* 2003, **13:**2498-2504.

109. Brown CT, Rust AG, Clarke PJ, Pan Z, Schilstra MJ, De Buysscher T, Griffin G, Wold BJ, Cameron RA, Davidson EH, Bolouri H: **New computational approaches for analysis of cis-regulatory networks.** *Dev Biol* 2002, **246:**86-102.

110. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, et al: **The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models.** *Bioinformatics* 2003, **19:**524-531.

111. Lloyd CM, Halstead MD, Nielsen PF: **CellML: its future, present and past.** *Prog Biophys Mol Biol* 2004, **85:**433-450.

112. Luciano JS: **PAX of mind for pathway researchers.** *Drug Discov Today* 2005, **10:**937-942.

113. Gormley P, Li K, Irwin GW: **Modelling molecular interaction pathways using a two-stage identification algorithm.** *Syst Synth Biol* 2007, **1:**145-160.

114. Ruths D, Muller M, Tseng JT, Nakhleh L, Ram PT: **The signaling petri net-based simulator: a non-parametric strategy for characterizing the dynamics of cell-specific signaling networks.** *PLoS Comput Biol* 2008, **4:**e1000005.

115. Watterson S, Marshall S, Ghazal P: **Logic models of pathway biology.** *Drug Discov Today* 2008, **13:**447-456.

116. Freeman TC, Raza S, Theocharidis A, Ghazal P: **The mEPN scheme: an intuitive and flexible graphical system for rendering biological pathways.** *BMC Syst Biol* 2010, **4:**65.

117. Kohn KW, Aladjem MI, Kim S, Weinstein JN, Pommier Y: **Depicting combinatorial complexity with the molecular interaction map notation.** *Mol Syst Biol* 2006, **2:**51.

118. Le Novere N, Hucka M, Mi H, Moodie S, Schreiber F, Sorokin A, Demir E, Wegner K, Aladjem MI, Wimalaratne SM, et al: **The Systems Biology Graphical Notation.** *Nat Biotechnol* 2009, **27:**735-741.

119. Kwiatkowska MZ, Heath JK: **Biological pathways as communicating computer systems.** *J Cell Sci* 2009, **122:**2793-2800.

120. Pandey R, Guru RK, Mount DW: **Pathway Miner: extracting gene association networks from molecular pathways for predicting the biological significance of gene expression microarray data.** *Bioinformatics* 2004, **20:**2156-2158.

121. Dahlquist KD, Salomonis N, Vranizan K, Lawlor SC, Conklin BR: **GenMAPP, a new tool for viewing and analyzing microarray data on biological pathways.** *Nat Genet* 2002, **31:**19-20.

122. Arakawa K, Kono N, Yamada Y, Mori H, Tomita M: **KEGG-based pathway visualization tool for complex omics data.** *In Silico Biol* 2005, **5:**419-423.

123. Ekins S, Nikolsky Y, Bugrim A, Kirillov E, Nikolskaya T: **Pathway mapping tools for analysis of high content data.** *Methods Mol Biol* 2007, **356:**319-350.

124. Cavalieri D, De Filippo C: **Bioinformatic methods for integrating whole-genome expression results into cellular networks.** *Drug Discov Today* 2005, **10:**727-734.

125. Babur O, Colak R, Demir E, Dogrusoz U: **PATIKAmad: putting microarray data into pathway context.** *Proteomics* 2008, **8:**2196-2198.

126. Antonov AV, Dietmann S, Mewes HW: **KEGG spider: interpretation of genomics data in the context of the global gene metabolic network.** *Genome Biol* 2008, **9:**R179.

127. van Riel NA: **Dynamic modelling and analysis of biochemical networks: mechanism-based models and model-based experiments.** *Brief Bioinform* 2006, **7:**364-374.

128. Eungdamrong NJ, Iyengar R: **Modeling cell signaling networks.** *Biol Cell* 2004, **96:**355-362.

129. Saffrey P, Orton R: **Version control of pathway models using XML patches.** *BMC Syst Biol* 2009, **3:**34.

130. Taubner C, Mathiak B, Kupfer A, Fleischer N, Eckstein S: **Modelling and simulation of the TLR4 pathway with coloured petri nets.** *Conf Proc IEEE Eng Med Biol Soc* 2006, **1:**2009-2012.

131. Barabasi AL, Gulbahce N, Loscalzo J: **Network medicine: a network-based approach to human disease.** *Nat Rev Genet* 2011, **12:**56-68.

132. Enright AJ, Kunin V, Ouzounis CA: **Protein families and TRIBES in genome sequence space.** *Nucleic Acids Res* 2003, **31:**4632-4638.

133. Enright AJ, Van Dongen S, Ouzounis CA: **An efficient algorithm for large-scale detection of protein families.** *Nucleic Acids Res* 2002, **30:**1575-1584.

134. Li L, Stoeckert CJ, Jr., Roos DS: **OrthoMCL: identification of ortholog groups for eukaryotic genomes.** *Genome Res* 2003, **13:**2178-2189.

135. Girvan M, Newman ME: **Community structure in social and biological networks.** *Proc Natl Acad Sci U S A* 2002, **99:**7821-7826.

136. Palla G, Derenyi I, Farkas I, Vicsek T: **Uncovering the overlapping community structure of complex networks in nature and society.** *Nature* 2005, **435:**814-818.

137. van Dongen S: **Graph Clustering by Flow Simulation.** *PhD thesis.* University of Utrecht, 2000.

138. Freeman TC, Goldovsky L, Brosch M, van Dongen S, Maziere P, Grocock RJ, Freilich S, Thornton J, Enright AJ: **Construction, visualisation, and clustering of transcription networks from microarray expression data.** *PLoS Comput Biol* 2007, **3:**2032-2042.

139. Theocharidis A, van Dongen S, Enright AJ, Freeman TC: **Network visualization and analysis of gene expression data using BioLayout Express(3D).** *Nat Protoc* 2009, **4:**1535-1550.

140. Cowley MJ, Cotsapas CJ, Williams RB, Chan EK, Pulvers JN, Liu MY, Luo OJ, Nott DJ, Little PF: **Intra- and inter-individual genetic differences in gene expression.** *Mamm Genome* 2009, **20:**281-295.

141. Edwards KD, Bombarely A, Story GW, Allen F, Mueller LA, Coates SA, Jones L: **TobEA: an atlas of tobacco gene expression from seed to senescence.** *BMC Genomics* 2010, **11:**142.

142. Inouye M, Silander K, Hamalainen E, Salomaa V, Harald K, Jousilahti P, Mannisto S, Eriksson JG, Saarela J, Ripatti S, et al: **An immune response network associated with blood lipid levels.** *PLoS Genet* 2010, **6.**

143. Mabbott NA, Kenneth Baillie J, Hume DA, Freeman TC: **Meta-analysis of lineage-specific gene expression signatures in mouse leukocyte populations.** *Immunobiology* 2010, **215:**724-736.

144. Natividad A, Freeman TC, Jeffries D, Burton MJ, Mabey DC, Bailey RL, Holland MJ: **Human conjunctival transcriptome analysis reveals the prominence of innate defense in Chlamydia trachomatis infection.** *Infect Immun* 2010, **78:**4895-4911.

145. Hume DA, Summers KM, Raza S, Baillie JK, Freeman TC: **Functional clustering and lineage markers: insights into cellular differentiation and gene function from large-scale microarray studies of purified primary cell populations.** *Genomics* 2010, **95:**328-338.

146. Lacaze P, Raza S, Sing G, Page D, Forster T, Storm P, Craigon M, Awad T, Ghazal P, Freeman TC: **Combined genome-wide expression profiling and targeted RNA interference in primary mouse macrophages reveals perturbation of transcriptional networks associated with interferon signalling.** *BMC Genomics* 2009, **10:**372.

147. Summers KM, Raza S, van Nimwegen E, Freeman TC, Hume DA: **Co-expression of FBN1 with mesenchyme-specific genes in mouse cell lines: implications for phenotypic variability in Marfan syndrome.** *Eur J Hum Genet* 2010, **18:**1209-1215.

148. Alfarano C, Andrade CE, Anthony K, Bahroos N, Bajec M, Bantoft K, Betel D, Bobechko B, Boutilier K, Burgess E, et al: **The Biomolecular Interaction Network Database and related tools 2005 update.** *Nucleic Acids Res* 2005, **33:**D418-424.

149. Hermjakob H, Montecchi-Palazzi L, Lewington C, Mudali S, Kerrien S, Orchard S, Vingron M, Roechert B, Roepstorff P, Valencia A, et al: **IntAct: an open source molecular interaction database.** *Nucleic Acids Res* 2004, **32:**D452-455.

150. **BioCarta: Charting Pathways of Life** [http://www.biocarta.com]

151. **Ingenuity Pathway Analysis** [IPA, http://www.ingenuity.com/]

152. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y: **KEGG for linking genomes to life and the environment.** *Nucleic Acids Res* 2008, **36:**D480-484.

153. Mishra GR, Suresh M, Kumaran K, Kannabiran N, Suresh S, Bala P, Shivakumar K, Anuradha N, Reddy R, Raghavan TM, et al: **Human protein reference database--2006 update.** *Nucleic Acids Res* 2006, **34:**D411-414.

154. Pico AR, Kelder T, van Iersel MP, Hanspers K, Conklin BR, Evelo C: **WikiPathways: pathway editing for the people.** *PLoS Biol* 2008, **6:**e184.

155. Schaefer CF, Anthony K, Krupa S, Buchoff J, Day M, Hannay T, Buetow KH: **PID: the Pathway Interaction Database.** *Nucleic Acids Res* 2009, **37:**D674-679.

156. Vastrik I, D'Eustachio P, Schmidt E, Gopinath G, Croft D, de Bono B, Gillespie M, Jassal B, Lewis S, Matthews L, et al: **Reactome: a knowledge base of biologic pathways and processes.** *Genome Biol* 2007, **8:**R39.

157. Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, Muller J, Doerks T, Julien P, Roth A, Simonovic M, et al: **STRING 8--a global view on proteins and their functional interactions in 630 organisms.** *Nucleic Acids Res* 2009, **37:**D412-416.

158. Le Novère N, Hucka M, Mi H, Moodie S, Shreiber F, Sorokin A, Demir E, Wegner K, Aladjem MI, Wimalaratne SM, Bergman FT, Gauges R, Ghazal P, Kawaji H, Li L, Matsuoka Y, Villéger A, Boyd SE, Calzone L, Courtot M, Dogrusoz U, Freeman TC, Funahashi A, Ghosh S, Jouraku A, Kim S, Kolpakov F, Luna A, Sahle S, Watterson S, Wu G, Goryanin I, Kell DB, Sander C, Sauro H, Snoep JL, Kohn K, Kitano H. : **The Systems Biology Graphical Notation.** *Nature Biotechnology* 2009, **27:**735-741.

159. Raza S, McDerment N, Lacaze PA, Robertson K, Watterson S, Chen Y, Chisholm M, Eleftheriadis G, Monk S, O'Sullivan M, et al: **Construction of a large scale integrated map of macrophage pathogen recognition and effector systems.** *BMC Syst Biol* 2010, **4:**63.

160. Zhong B, Tien P, Shu HB: **Innate immune responses: crosstalk of signaling and regulation of gene transcription.** *Virology* 2006, **352:**14-21.
161. Chaudhary PM, Eby M, Jasmin A, Bookwalter A, Murray J, Hood L: **Death receptor 5, a new member of the TNFR family, and DR4 induce FADD-dependent apoptosis and activate the NF-kappaB pathway.** *Immunity* 1997, **7:**821-830.
162. Freeman T.C. GL, Brosch M., van Dongen S., Mazière P., Grocock R.J, Freilich S., Thornton J. and Enright A.J. : **Construction, Visualisation and Clustering of Transcription Networks from Microarray Expression Data.** *PLoS Computational Biology* 2007, *in press*.
163. Grzegorczyk M, Husmeier D, Edwards KD, Ghazal P, Millar AJ: **Modelling non-stationary gene regulatory processes with a non-homogeneous Bayesian network and the allocation sampler.** *Bioinformatics* 2008, **24:**2071-2078.
164. Nishiya T, DeFranco AL: **Ligand-regulated chimeric receptor approach reveals distinctive subcellular localization and signaling properties of the Toll-like receptors.** *J Biol Chem* 2004, **279:**19008-19017.
165. Dragan AI, Carrillo R, Gerasimova TI, Privalov PL: **Assembling the human IFN-beta enhanceosome in solution.** *J Mol Biol* 2008, **384:**335-348.
166. Kirk P, Bazan JF: **Pathogen recognition: TLRs throw us a curve.** *Immunity* 2005, **23:**347-350.
167. Alexopoulou L, Holt AC, Medzhitov R, Flavell RA: **Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3.** *Nature* 2001, **413:**732-738.
168. Lund JM, Alexopoulou L, Sato A, Karow M, Adams NC, Gale NW, Iwasaki A, Flavell RA: **Recognition of single-stranded RNA viruses by Toll-like receptor 7.** *Proc Natl Acad Sci U S A* 2004, **101:**5598-5603.
169. Feuillet V, Medjane S, Mondor I, Demaria O, Pagni PP, Galan JE, Flavell RA, Alexopoulou L: **Involvement of Toll-like receptor 5 in the recognition of flagellated bacteria.** *Proc Natl Acad Sci U S A* 2006, **103:**12487-12492.
170. Hayashi F, Smith KD, Ozinsky A, Hawn TR, Yi EC, Goodlett DR, Eng JK, Akira S, Underhill DM, Aderem A: **The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5.** *Nature* 2001, **410:**1099-1103.
171. Nakao Y, Funami K, Kikkawa S, Taniguchi M, Nishiguchi M, Fukumori Y, Seya T, Matsumoto M: **Surface-expressed TLR6 participates in the recognition of diacylated lipopeptide and peptidoglycan in human cells.** *J Immunol* 2005, **174:**1566-1573.
172. Takeda K, Takeuchi O, Akira S: **Recognition of lipopeptides by Toll-like receptors.** *J Endotoxin Res* 2002, **8:**459-463.
173. Arbour NC, Lorenz E, Schutte BC, Zabner J, Kline JN, Jones M, Frees K, Watt JL, Schwartz DA: **TLR4 mutations are associated with endotoxin hyporesponsiveness in humans.** *Nat Genet* 2000, **25:**187-191.
174. Rhee SH, Hwang D: **Murine TOLL-like receptor 4 confers lipopolysaccharide responsiveness as determined by activation of NF kappa B and expression of the inducible cyclooxygenase.** *J Biol Chem* 2000, **275:**34035-34040.
175. Lamphier MS, Sirois CM, Verma A, Golenbock DT, Latz E: **TLR9 and the recognition of self and non-self nucleic acids.** *Ann N Y Acad Sci* 2006, **1082:**31-43.
176. Takeshita F, Leifer CA, Gursel I, Ishii KJ, Takeshita S, Gursel M, Klinman DM: **Cutting edge: Role of Toll-like receptor 9 in CpG DNA-induced activation of human cells.** *J Immunol* 2001, **167:**3555-3558.
177. Colonna M: **TLR pathways and IFN-regulatory factors: to each its own.** *Eur J Immunol* 2007, **37:**306-309.
178. Dong C, Davis RJ, Flavell RA: **MAP kinases in the immune response.** *Annu Rev Immunol* 2002, **20:**55-72.

179. Martinon F, Tschopp J: **Inflammatory caspases and inflammasomes: master switches of inflammation.** *Cell Death Differ* 2007, **14:**10-22.

180. Tschopp J, Martinon F, Burns K: **NALPs: a novel protein family involved in inflammation.** *Nat Rev Mol Cell Biol* 2003, **4:**95-104.

181. Martinon F, Gaide O, Petrilli V, Mayor A, Tschopp J: **NALP inflammasomes: a central role in innate immunity.** *Semin Immunopathol* 2007, **29:**213-229.

182. Takaoka A, Wang Z, Choi MK, Yanai H, Negishi H, Ban T, Lu Y, Miyagishi M, Kodama T, Honda K, et al: **DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response.** *Nature* 2007, **448:**501-505.

183. Burckstummer T, Baumann C, Bluml S, Dixit E, Durnberger G, Jahn H, Planyavsky M, Bilban M, Colinge J, Bennett KL, Superti-Furga G: **An orthogonal proteomic-genomic screen identifies AIM2 as a cytoplasmic DNA sensor for the inflammasome.** *Nat Immunol* 2009, **10:**266-272.

184. Fernandes-Alnemri T, Yu JW, Datta P, Wu J, Alnemri ES: **AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA.** *Nature* 2009, **458:**509-513.

185. Hornung V, Ablasser A, Charrel-Dennis M, Bauernfeind F, Horvath G, Caffrey DR, Latz E, Fitzgerald KA: **AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC.** *Nature* 2009, **458:**514-518.

186. Roberts TL, Idris A, Dunn JA, Kelly GM, Burnton CM, Hodgson S, Hardy LL, Garceau V, Sweet MJ, Ross IL, et al: **HIN-200 proteins regulate caspase activation in response to foreign cytoplasmic DNA.** *Science* 2009, **323:**1057-1060.

187. Yie J, Senger K, Thanos D: **Mechanism by which the IFN-beta enhanceosome activates transcription.** *Proc Natl Acad Sci U S A* 1999, **96:**13108-13113.

188. Panne D, Maniatis T, Harrison SC: **Crystal structure of ATF-2/c-Jun and IRF-3 bound to the interferon-beta enhancer.** *EMBO J* 2004, **23:**4384-4393.

189. Panne D, Maniatis T, Harrison SC: **An atomic model of the interferon-beta enhanceosome.** *Cell* 2007, **129:**1111-1123.

190. Karin M, Lin A: **NF-kappaB at the crossroads of life and death.** *Nat Immunol* 2002, **3:**221-227.

191. Barnes PJ, Karin M: **Nuclear factor-kappaB: a pivotal transcription factor in chronic inflammatory diseases.** *N Engl J Med* 1997, **336:**1066-1071.

192. Gilmore TD, Koedood M, Piffat KA, White DW: **Rel/NF-kappaB/IkappaB proteins and cancer.** *Oncogene* 1996, **13:**1367-1378.

193. Vallabhapurapu S, Karin M: **Regulation and function of NF-kappaB transcription factors in the immune system.** *Annu Rev Immunol* 2009, **27:**693-733.

194. Wietek C, O'Neill LA: **Diversity and regulation in the NF-kappaB system.** *Trends Biochem Sci* 2007, **32:**311-319.

195. Basak S, Shih VF, Hoffmann A: **Generation and activation of multiple dimeric transcription factors within the NF-kappaB signaling system.** *Mol Cell Biol* 2008, **28:**3139-3150.

196. Latimer M, Ernst MK, Dunn LL, Drutskaya M, Rice NR: **The N-terminal domain of IkappaB alpha masks the nuclear localization signal(s) of p50 and c-Rel homodimers.** *Mol Cell Biol* 1998, **18:**2640-2649.

197. Ganchi PA, Sun SC, Greene WC, Ballard DW: **I kappa B/MAD-3 masks the nuclear localization signal of NF-kappa B p65 and requires the transactivation domain to inhibit NF-kappa B p65 DNA binding.** *Mol Biol Cell* 1992, **3:**1339-1352.

198. **http://people.bu.edu/gilmore/nf-kb/**

199. Semple CA: **The comparative proteomics of ubiquitination in mouse.** *Genome Res* 2003, **13:**1389-1394.

200. Petroski MD: **The ubiquitin system, disease, and drug discovery.** *BMC Biochem* 2008, **9 Suppl 1:**S7.

201. Murata S, Yashiroda H, Tanaka K: **Molecular mechanisms of proteasome assembly.** *Nat Rev Mol Cell Biol* 2009, **10:**104-115.

202. Tanaka K: **The proteasome: overview of structure and functions.** *Proc Jpn Acad Ser B Phys Biol Sci* 2009, **85:**12-36.

203. **VisuaLyzer™. Medical Decision Logic, Inc.** [http://www.mdlogix.com]

204. **Network Workbench Tool.** [http://nwb.slis.indiana.edu]

205. Jünger M, Mutzel P,: **Pajek - Analysis and Visualization of Large Networks. .** In *Graph Drawing Software.* Berlin: Springer; 2003: 77-103

206. Theocharidis A, van Dongen S, Enright AJ, Freeman TC: **Network Visualisation and Analysis of Gene Expression Data using BioLayout *Express*^{3D}.** *Nature Protocols* 2009, **In press**.

207. Enright AJ, Ouzounis CA: **BioLayout--an automatic graph layout algorithm for similarity visualization.** *Bioinformatics* 2001, **17:**853-854.

208. Clement-Ziza M, Malabat C, Weber C, Moszer I, Aittokallio T, Letondal C, Rousseau S: **Genoscape: a Cytoscape plug-in to automate the retrieval and integration of gene expression data and molecular networks.** *Bioinformatics* 2009, **25:**2617-2618.

209. Kohn KW, Aladjem MI, Weinstein JN, Pommier Y: **Molecular interaction maps of bioregulatory networks: a general rubric for systems biology.** *Mol Biol Cell* 2006, **17:**1-13.

210. Oda K, Kimura T, Matsuoka Y, Funahashi A, Muramatsu M, Kitano H: **Molecular interaction map of a macrophage.** In *AfCS Research Reports*, vol. 2. pp. 12; 2004:12.

211. Raza S, McDerment N, Lacaze PA, Robertson K, Watterson S, Chen Y, Chisholm M, Eleftheriadis G, Monk S, O'Sullivan M, et al: **Construction of a large scale integrated map of macrophage pathogen recognition and effector systems.** *BMC Syst Biol* 2010, **4:**63.

212. Freeman TC, Raza S, Theocharidis A, Ghazal P: **The mEPN scheme: an intuitive and flexible graphical system for rendering biological pathways.** *BMC Syst Biol* 2010, **4:**65.

213. Marcucci F, Waller M, Kirchner H, Krammer P: **Production of immune interferon by murine T-cell clones from long-term cultures.** *Nature* 1981, **291:**79-81.

214. O'Malley JA, Nussbaum-Blumenson A, Sheedy D, Grossmayer BJ, Ozer H: **Identification of the T cell subset that produces human gamma interferon.** *J Immunol* 1982, **128:**2522-2526.

215. Munakata T, Semba U, Shibuya Y, Kuwano K, Akagi M, Arai S: **Induction of interferon-gamma production by human natural killer cells stimulated by hydrogen peroxide.** *J Immunol* 1985, **134:**2449-2455.

216. Walter MR, Windsor WT, Nagabhushan TL, Lundell DJ, Lunn CA, Zauodny PJ, Narula SK: **Crystal structure of a complex between interferon-gamma and its soluble high-affinity receptor.** *Nature* 1995, **376:**230-235.

217. Sakatsume M, Igarashi K, Winestock KD, Garotta G, Larner AC, Finbloom DS: **The Jak kinases differentially associate with the alpha and beta (accessory factor) chains of the interferon gamma receptor to form a functional receptor unit capable of activating STAT transcription factors.** *J Biol Chem* 1995, **270:**17528-17534.

218. Watling D, Guschin D, Muller M, Silvennoinen O, Witthuhn BA, Quelle FW, Rogers NC, Schindler C, Stark GR, Ihle JN, et al.: **Complementation by the protein tyrosine kinase JAK2 of a mutant cell line defective in the interferon-gamma signal transduction pathway.** *Nature* 1993, **366:**166-170.

219. Igarashi K, Garotta G, Ozmen L, Ziemiecki A, Wilks AF, Harpur AG, Larner AC, Finbloom DS: **Interferon-gamma induces tyrosine phosphorylation of interferon-gamma**

receptor and regulated association of protein tyrosine kinases, Jak1 and Jak2, with its receptor. *J Biol Chem* 1994, **269:**14333-14336.

220. Endo TA, Masuhara M, Yokouchi M, Suzuki R, Sakamoto H, Mitsui K, Matsumoto A, Tanimura S, Ohtsubo M, Misawa H, et al: **A new protein containing an SH2 domain that inhibits JAK kinases.** *Nature* 1997, **387:**921-924.

221. ten Hoeve J, de Jesus Ibarra-Sanchez M, Fu Y, Zhu W, Tremblay M, David M, Shuai K: **Identification of a nuclear Stat1 protein tyrosine phosphatase.** *Mol Cell Biol* 2002, **22:**5662-5668.

222. Simoncic PD, Lee-Loy A, Barber DL, Tremblay ML, McGlade CJ: **The T cell protein tyrosine phosphatase is a negative regulator of janus family kinases 1 and 3.** *Curr Biol* 2002, **12:**446-453.

223. Uddin S, Sassano A, Deb DK, Verma A, Majchrzak B, Rahman A, Malik AB, Fish EN, Platanias LC: **Protein kinase C-delta (PKC-delta) is activated by type I interferons and mediates phosphorylation of Stat1 on serine 727.** *J Biol Chem* 2002, **277:**14408-14416.

224. Naka T, Narazaki M, Hirata M, Matsumoto T, Minamoto S, Aono A, Nishimoto N, Kajita T, Taga T, Yoshizaki K, et al: **Structure and function of a new STAT-induced STAT inhibitor.** *Nature* 1997, **387:**924-929.

225. Alexander WS: **Suppressors of cytokine signalling (SOCS) in the immune system.** *Nat Rev Immunol* 2002, **2:**410-416.

226. Reith W, LeibundGut-Landmann S, Waldburger JM: **Regulation of MHC class II gene expression by the class II transactivator.** *Nat Rev Immunol* 2005, **5:**793-806.

227. Boss JM, Jensen PE: **Transcriptional regulation of the MHC class II antigen presentation pathway.** *Curr Opin Immunol* 2003, **15:**105-111.

228. **The Systems Biology Graphical Notation** [http://www.sbgn.org]

229. Yeung N, Cline MS, Kuchinsky A, Smoot ME, Bader GD: **Exploring biological networks with Cytoscape software.** *Curr Protoc Bioinformatics* 2008, **Chapter 8:**Unit 8 13.

230. **yEd Graph Editior - yWorks the diagramming company** [http://www.yworks.com]

231. **Interferome: The database of interferon regulated genes. Available at: http://www.interferome.org/**

232. Stark GR, Kerr IM, Williams BR, Silverman RH, Schreiber RD: **How cells respond to interferons.** *Annu Rev Biochem* 1998, **67:**227-264.

233. Meurs E, Chong K, Galabru J, Thomas NS, Kerr IM, Williams BR, Hovanessian AG: **Molecular cloning and characterization of the human double-stranded RNA-activated protein kinase induced by interferon.** *Cell* 1990, **62:**379-390.

234. Hovanessian AG: **Interferon-induced and double-stranded RNA-activated enzymes: a specific protein kinase and 2',5'-oligoadenylate synthetases.** *J Interferon Res* 1991, **11:**199-205.

235. Horisberger MA: **Interferons, Mx genes, and resistance to influenza virus.** *Am J Respir Crit Care Med* 1995, **152:**S67-71.

236. Pavlovic J, Schroder A, Blank A, Pitossi F, Staeheli P: **Mx proteins: GTPases involved in the interferon-induced antiviral state.** *Ciba Found Symp* 1993, **176:**233-243; discussion 243-237.

237. Zhou A, Paranjape JM, Der SD, Williams BR, Silverman RH: **Interferon action in triply deficient mice reveals the existence of alternative antiviral pathways.** *Virology* 1999, **258:**435-440.

238. Der SD, Zhou A, Williams BR, Silverman RH: **Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays.** *Proc Natl Acad Sci U S A* 1998, **95:**15623-15628.

239. Steimle V, Siegrist CA, Mottet A, Lisowska-Grospierre B, Mach B: **Regulation of MHC class II expression by interferon-gamma mediated by the transactivator gene CIITA.** *Science* 1994, **265:**106-109.

240. Chang CH, Guerder S, Hong SC, van Ewijk W, Flavell RA: **Mice lacking the MHC class II transactivator (CIITA) show tissue-specific impairment of MHC class II expression.** *Immunity* 1996, **4:**167-178.

241. Muller U, Steinhoff U, Reis LF, Hemmi S, Pavlovic J, Zinkernagel RM, Aguet M: **Functional role of type I and type II interferons in antiviral defense.** *Science* 1994, **264:**1918-1921.

242. Rubin BY, Gupta SL: **Differential efficacies of human type I and type II interferons as antiviral and antiproliferative agents.** *Proc Natl Acad Sci U S A* 1980, **77:**5928-5932.

243. Indraccolo S, Pfeffer U, Minuzzo S, Esposito G, Roni V, Mandruzzato S, Ferrari N, Anfosso L, Dell'Eva R, Noonan DM, et al: **Identification of genes selectively regulated by IFNs in endothelial cells.** *J Immunol* 2007, **178:**1122-1135.

244. Sanda C, Weitzel P, Tsukahara T, Schaley J, Edenberg HJ, Stephens MA, McClintick JN, Blatt LM, Li L, Brodsky L, Taylor MW: **Differential gene induction by type I and type II interferons and their combination.** *J Interferon Cytokine Res* 2006, **26:**462-472.

245. Tassiulas I, Hu X, Ho H, Kashyap Y, Paik P, Hu Y, Lowell CA, Ivashkiv LB: **Amplification of IFN-alpha-induced STAT1 activation and inflammatory function by Syk and ITAM-containing adaptors.** *Nat Immunol* 2004, **5:**1181-1189.

246. Waddell SJ, Popper SJ, Rubins KH, Griffiths MJ, Brown PO, Levin M, Relman DA: **Dissecting interferon-induced transcriptional programs in human peripheral blood cells.** *PLoS One* 2010, **5:**e9753.

247. Zhang S, Kim CC, Batra S, McKerrow JH, Loke P: **Delineation of diverse macrophage activation programs in response to intracellular parasites and cytokines.** *PLoS Negl Trop Dis* 2010, **4:**e648.

248. Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA: **DAVID: Database for Annotation, Visualization, and Integrated Discovery.** *Genome Biol* 2003, **4:**P3.

249. Huang da W, Sherman BT, Lempicki RA: **Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.** *Nat Protoc* 2009, **4:**44-57.

250. Pixley FJ, Stanley ER: **CSF-1 regulation of the wandering macrophage: complexity in action.** *Trends Cell Biol* 2004, **14:**628-638.

251. **YEd Graph Editor**

252. Krebs DL, Hilton DJ: **SOCS: physiological suppressors of cytokine signaling.** *J Cell Sci* 2000, **113 (Pt 16):**2813-2819.

253. Song MM, Shuai K: **The suppressor of cytokine signaling (SOCS) 1 and SOCS3 but not SOCS2 proteins inhibit interferon-mediated antiviral and antiproliferative activities.** *J Biol Chem* 1998, **273:**35056-35062.

254. Raices RM, Kannan Y, Bellamkonda-Athmaram V, Seshadri S, Wang H, Guttridge DC, Wewers MD: **A novel role for IkappaBzeta in the regulation of IFNgamma production.** *PLoS One* 2009, **4:**e6776.

255. Keller AD, Maniatis T: **Identification and characterization of a novel repressor of beta-interferon gene expression.** *Genes Dev* 1991, **5:**868-879.

256. Arlt A, Schafer H: **Role of the immediate early response 3 (IER3) gene in cellular stress response, inflammation and tumorigenesis.** *Eur J Cell Biol* 2010.

257. Takekawa M, Saito H: **A family of stress-inducible GADD45-like proteins mediate activation of the stress-responsive MTK1/MEKK4 MAPKKK.** *Cell* 1998, **95:**521-530.

258. Teng CH, Huang WN, Meng TC: **Several dual specificity phosphatases coordinate to control the magnitude and duration of JNK activation in signaling response to oxidative stress.** *J Biol Chem* 2007, **282:**28395-28407.

259. Zhang YW, Vande Woude GF: **Mig-6, signal transduction, stress response and cancer.** *Cell Cycle* 2007, **6:**507-513.

260. Marques JT, Williams BR: **Activation of the mammalian immune system by siRNAs.** *Nat Biotechnol* 2005, **23:**1399-1405.

261. Yoneyama M, Fujita T: **RNA recognition and signal transduction by RIG-I-like receptors.** *Immunol Rev* 2009, **227:**54-65.

262. Sioud M: **Innate sensing of self and non-self RNAs by Toll-like receptors.** *Trends Mol Med* 2006, **12:**167-176.

263. Sioud M: **RNA interference and innate immunity.** *Adv Drug Deliv Rev* 2007, **59:**153-163.

264. Colina R, Costa-Mattioli M, Dowling RJ, Jaramillo M, Tai LH, Breitbach CJ, Martineau Y, Larsson O, Rong L, Svitkin YV, et al: **Translational control of the innate immune response through IRF-7.** *Nature* 2008, **452:**323-328.

265. Grandvaux N, Servant MJ, tenOever B, Sen GC, Balachandran S, Barber GN, Lin R, Hiscott J: **Transcriptional profiling of interferon regulatory factor 3 target genes: direct involvement in the regulation of interferon-stimulated genes.** *J Virol* 2002, **76:**5532-5539.

266. Paun A, Reinert JT, Jiang Z, Medin C, Balkhi MY, Fitzgerald KA, Pitha PM: **Functional characterization of murine interferon regulatory factor 5 (IRF-5) and its role in the innate antiviral response.** *J Biol Chem* 2008, **283:**14295-14308.

267. Bonizzi G, Bebien M, Otero DC, Johnson-Vroom KE, Cao Y, Vu D, Jegga AG, Aronow BJ, Ghosh G, Rickert RC, Karin M: **Activation of IKKalpha target genes depends on recognition of specific kappaB binding sites by RelB:p52 dimers.** *EMBO J* 2004, **23:**4202-4210.

268. Du Z, Wei L, Murti A, Pfeffer SR, Fan M, Yang CH, Pfeffer LM: **Non-conventional signal transduction by type 1 interferons: the NF-kappaB pathway.** *J Cell Biochem* 2007, **102:**1087-1094.

269. Yang CH, Murti A, Pfeffer LM: **Interferon induces NF-kappa B-inducing kinase/tumor necrosis factor receptor-associated factor-dependent NF-kappa B activation to promote cell survival.** *J Biol Chem* 2005, **280:**31530-31536.

270. Mordmuller B, Krappmann D, Esen M, Wegener E, Scheidereit C: **Lymphotoxin and lipopolysaccharide induce NF-kappaB-p52 generation by a co-translational mechanism.** *EMBO Rep* 2003, **4:**82-87.

271. de Wit H, Dokter WH, Koopmans SB, Lummen C, van der Leij M, Smit JW, Vellenga E: **Regulation of p100 (NFKB2) expression in human monocytes in response to inflammatory mediators and lymphokines.** *Leukemia* 1998, **12:**363-370.

272. Krausgruber T, Blazek K, Smallie T, Alzabin S, Lockstone H, Sahgal N, Hussell T, Feldmann M, Udalova IA: **IRF5 promotes inflammatory macrophage polarization and T(H)1-T(H)17 responses.** *Nat Immunol* 2011, **12:**231-238.

273. Bogdan C: **Nitric oxide and the immune response.** *Nat Immunol* 2001, **2:**907-916.

274. Dobrovolskaia MA, Vogel SN: **Toll receptors, CD14, and macrophage activation and deactivation by LPS.** *Microbes Infect* 2002, **4:**903-914.

275. Gessani S, Belardelli F, Pecorelli A, Puddu P, Baglioni C: **Bacterial lipopolysaccharide and gamma interferon induce transcription of beta interferon mRNA and interferon secretion in murine macrophages.** *J Virol* 1989, **63:**2785-2789.

276. Karaghiosoff M, Steinborn R, Kovarik P, Kriegshauser G, Baccarini M, Donabauer B, Reichart U, Kolbe T, Bogdan C, Leanderson T, et al: **Central role for type I interferons and Tyk2 in lipopolysaccharide-induced endotoxin shock.** *Nat Immunol* 2003, **4:**471-477.

277. Thomas KE, Galligan CL, Newman RD, Fish EN, Vogel SN: **Contribution of interferon-beta to the murine macrophage response to the toll-like receptor 4 agonist, lipopolysaccharide.** *J Biol Chem* 2006, **281:**31119-31130.

278. Wiese M, Castiglione K, Hensel M, Schleicher U, Bogdan C, Jantsch J: **Small interfering RNA (siRNA) delivery into murine bone marrow-derived macrophages by electroporation.** *J Immunol Methods* 2010, **353:**102-110.

279. Kostarelos K, Miller AD: **Synthetic, self-assembly ABCD nanoparticles; a structural paradigm for viable synthetic non-viral vectors.** *Chem Soc Rev* 2005, **34:**970-994.

280. Reynolds A, Anderson EM, Vermeulen A, Fedorov Y, Robinson K, Leake D, Karpilow J, Marshall WS, Khvorova A: **Induction of the interferon response by siRNA is cell type- and duplex length-dependent.** *RNA* 2006, **12:**988-993.

281. Bantounas I, Phylactou LA, Uney JB: **RNA interference and the use of small interfering RNA to study gene function in mammalian systems.** *J Mol Endocrinol* 2004, **33:**545-557.

282. Panne D: **The enhanceosome.** *Curr Opin Struct Biol* 2008, **18:**236-242.

283. Wells CA, Ravasi T, Faulkner GJ, Carninci P, Okazaki Y, Hayashizaki Y, Sweet M, Wainwright BJ, Hume DA: **Genetic control of the innate immune response.** *BMC Immunol* 2003, **4:**5.

284. Hornung V, Guenthner-Biller M, Bourquin C, Ablasser A, Schlee M, Uematsu S, Noronha A, Manoharan M, Akira S, de Fougerolles A, et al: **Sequence-specific potent induction of IFN-alpha by short interfering RNA in plasmacytoid dendritic cells through TLR7.** *Nat Med* 2005, **11:**263-270.

285. Sioud M: **Single-stranded small interfering RNA are more immunostimulatory than their double-stranded counterparts: a central role for 2'-hydroxyl uridines in immune responses.** *Eur J Immunol* 2006, **36:**1222-1230.

286. Diebold SS, Massacrier C, Akira S, Paturel C, Morel Y, Reis e Sousa C: **Nucleic acid agonists for Toll-like receptor 7 are defined by the presence of uridine ribonucleotides.** *Eur J Immunol* 2006, **36:**3256-3267.

287. Goodchild A, Nopper N, King A, Doan T, Tanudji M, Arndt GM, Poidinger M, Rivory LP, Passioura T: **Sequence determinants of innate immune activation by short interfering RNAs.** *BMC Immunol* 2009, **10:**40.

288. Carralot JP, Kim TK, Lenseigne B, Boese AS, Sommer P, Genovesio A, Brodin P: **Automated high-throughput siRNA transfection in raw 264.7 macrophages: a case study for optimization procedure.** *J Biomol Screen* 2009, **14:**151-160.

289. Bartlett DW, Davis ME: **Insights into the kinetics of siRNA-mediated gene silencing from live-cell and live-animal bioluminescent imaging.** *Nucleic Acids Res* 2006, **34:**322-333.

290. Honma K, Udono H, Kohno T, Yamamoto K, Ogawa A, Takemori T, Kumatori A, Suzuki S, Matsuyama T, Yui K: **Interferon regulatory factor 4 negatively regulates the production of proinflammatory cytokines by macrophages in response to LPS.** *Proc Natl Acad Sci U S A* 2005, **102:**16001-16006.

291. Xu Z, Huang CX, Li Y, Wang PZ, Ren GL, Chen CS, Shang FJ, Zhang Y, Liu QQ, Jia ZS, et al: **Toll-like receptor 4 siRNA attenuates LPS-induced secretion of inflammatory cytokines and chemokines by macrophages.** *J Infect* 2007, **55:**e1-9.

292. Beutler B, Tkacenko V, Milsark I, Krochin N, Cerami A: **Effect of gamma interferon on cachectin expression by mononuclear phagocytes. Reversal of the lpsd (endotoxin resistance) phenotype.** *J Exp Med* 1986, **164:**1791-1796.

293. Williams JG, Jurkovich GJ, Hahnel GB, Maier RV: **Macrophage priming by interferon gamma: a selective process with potentially harmful effects.** *J Leukoc Biol* 1992, **52:**579-584.

294. Doughty L, Nguyen K, Durbin J, Biron C: **A role for IFN-alpha beta in virus infection-induced sensitization to endotoxin.** *J Immunol* 2001, **166:**2658-2664.

295. Erroi A, Fantuzzi G, Mengozzi M, Sironi M, Orencole SF, Clark BD, Dinarello CA, Isetta A, Gnocchi P, Giovarelli M, et al.: **Differential regulation of cytokine production in lipopolysaccharide tolerance in mice.** *Infect Immun* 1993, **61:**4356-4359.

296. Mackensen A, Galanos C, Wehr U, Engelhardt R: **Endotoxin tolerance: regulation of cytokine production and cellular changes in response to endotoxin application in cancer patients.** *Eur Cytokine Netw* 1992, **3:**571-579.

297. Zhang X, Morrison DC: **Lipopolysaccharide-induced selective priming effects on tumor necrosis factor alpha and nitric oxide production in mouse peritoneal macrophages.** *J Exp Med* 1993, **177:**511-516.

298. Ziegler-Heitbrock HW, Frankenberger M, Wedel A: **Tolerance to lipopolysaccharide in human blood monocytes.** *Immunobiology* 1995, **193:**217-223.

299. Granowitz EV, Porat R, Mier JW, Orencole SF, Kaplanski G, Lynch EA, Ye K, Vannier E, Wolff SM, Dinarello CA: **Intravenous endotoxin suppresses the cytokine response of peripheral blood mononuclear cells of healthy humans.** *J Immunol* 1993, **151:**1637-1645.

300. Takasuka N, Matsuura K, Yamamoto S, Akagawa KS: **Suppression of TNF-alpha mRNA expression in LPS-primed macrophages occurs at the level of nuclear factor-kappa B activation, but not at the level of protein kinase C or CD14 expression.** *J Immunol* 1995, **154:**4803-4812.

301. Biswas SK, Lopez-Collazo E: **Endotoxin tolerance: new mechanisms, molecules and clinical significance.** *Trends Immunol* 2009, **30:**475-487.

302. Hume DA, Underhill DM, Sweet MJ, Ozinsky AO, Liew FY, Aderem A: **Macrophages exposed continuously to lipopolysaccharide and other agonists that act via toll-like receptors exhibit a sustained and additive activation state.** *BMC Immunol* 2001, **2:**11.

303. Vadiveloo PK, Vairo G, Hertzog P, Kola I, Hamilton JA: **Role of type I interferons during macrophage activation by lipopolysaccharide.** *Cytokine* 2000, **12:**1639-1646.

304. Toshchakov V, Jones BW, Perera PY, Thomas K, Cody MJ, Zhang S, Williams BR, Major J, Hamilton TA, Fenton MJ, Vogel SN: **TLR4, but not TLR2, mediates IFN-beta-induced STAT1alpha/beta-dependent gene expression in macrophages.** *Nat Immunol* 2002, **3:**392-398.

305. Peters KL, Smith HL, Stark GR, Sen GC: **IRF-3-dependent, NFkappa B- and JNK-independent activation of the 561 and IFN-beta genes in response to double-stranded RNA.** *Proc Natl Acad Sci U S A* 2002, **99:**6322-6327.

306. Wang X, Hussain S, Wang EJ, Li MO, Garcia-Sastre A, Beg AA: **Lack of essential role of NF-kappa B p50, RelA, and cRel subunits in virus-induced type 1 IFN expression.** *J Immunol* 2007, **178:**6770-6776.

307. Basagoudanavar SH, Thapa RJ, Nogusa S, Wang J, Beg AA, Balachandran S: **Distinct Roles for the NF-{kappa}B RelA Subunit during Antiviral Innate Immune Responses.** *J Virol* 2011, **85:**2599-2610.

308. Kitano H: **Biological robustness.** *Nat Rev Genet* 2004, **5:**826-837.

309. Sweet MJ, Hume DA: **Endotoxin signal transduction in macrophages.** *J Leukoc Biol* 1996, **60:**8-26.

310. Reynolds A, Leake D, Boese Q, Scaringe S, Marshall WS, Khvorova A: **Rational siRNA design for RNA interference.** *Nat Biotechnol* 2004, **22:**326-330.

311. Chen PY, Weinmann L, Gaidatzis D, Pei Y, Zavolan M, Tuschl T, Meister G: **Strand-specific 5'-O-methylation of siRNA duplexes controls guide strand selection and targeting specificity.** *RNA* 2008, **14:**263-274.

312. Jackson AL, Burchard J, Leake D, Reynolds A, Schelter J, Guo J, Johnson JM, Lim L, Karpilow J, Nichols K, et al: **Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing.** *RNA* 2006, **12:**1197-1205.

313. Blanc M, Hsieh WY, Robertson KA, Watterson S, Shui G, Lacaze P, Khondoker M, Dickinson P, Sing G, Rodriguez-Martin S, et al: **Host defense against viral infection involves interferon mediated down-regulation of sterol biosynthesis.** *PLoS Biol* 2011, **9:**e1000598.

314. van de Wetering D, van Wengen A, Savage ND, van de Vosse E, van Dissel JT: **IFN-alpha cannot substitute lack of IFN-gamma responsiveness in cells of an IFN-gammaR1 deficient patient.** *Clin Immunol* 2011, **138:**282-290.

315. Goerdt S, Politz O, Schledzewski K, Birk R, Gratchev A, Guillot P, Hakiy N, Klemke CD, Dippel E, Kodelja V, Orfanos CE: **Alternative versus classical activation of macrophages.** *Pathobiology* 1999, **67:**222-226.

316. Ulevitch RJ, Tobias PS: **Recognition of gram-negative bacteria and endotoxin by the innate immune system.** *Curr Opin Immunol* 1999, **11:**19-22.

317. Schroder K, Sweet MJ, Hume DA: **Signal integration between IFNgamma and TLR signalling pathways in macrophages.** *Immunobiology* 2006, **211:**511-524.

318. Liew FY, Xu D, Brint EK, O'Neill LA: **Negative regulation of toll-like receptor-mediated immune responses.** *Nat Rev Immunol* 2005, **5:**446-458.

319. Held TK, Weihua X, Yuan L, Kalvakolanu DV, Cross AS: **Gamma interferon augments macrophage activation by lipopolysaccharide by two distinct mechanisms, at the signal transduction level and via an autocrine mechanism involving tumor necrosis factor alpha and interleukin-1.** *Infect Immun* 1999, **67:**206-212.

320. Lu YC, Yeh WC, Ohashi PS: **LPS/TLR4 signal transduction pathway.** *Cytokine* 2008, **42:**145-151.

321. Miyake K: **Innate immune sensing of pathogens and danger signals by cell surface Toll-like receptors.** *Semin Immunol* 2007, **19:**3-10.

322. de Veer MJ, Holko M, Frevel M, Walker E, Der S, Paranjape JM, Silverman RH, Williams BR: **Functional classification of interferon-stimulated genes identified using microarrays.** *J Leukoc Biol* 2001, **69:**912-920.

323. Hume DA, Wells CA, Ravasi T: **Transcriptional regulatory networks in macrophages.** *Novartis Found Symp* 2007, **281:**2-18; discussion 18-24, 50-13, 208-209.

324. Kota RS, Rutledge JC, Gohil K, Kumar A, Enelow RI, Ramana CV: **Regulation of gene expression in RAW 264.7 macrophage cell line by interferon-gamma.** *Biochem Biophys Res Commun* 2006, **342:**1137-1146.

325. Nilsson R, Bajic VB, Suzuki H, di Bernardo D, Bjorkegren J, Katayama S, Reid JF, Sweet MJ, Gariboldi M, Carninci P, et al: **Transcriptional network dynamics in macrophage activation.** *Genomics* 2006, **88:**133-142.

326. Roach JC, Smith KD, Strobe KL, Nissen SM, Haudenschild CD, Zhou D, Vasicek TJ, Held GA, Stolovitzky GA, Hood LE, Aderem A: **Transcription factor expression in lipopolysaccharide-activated peripheral-blood-derived mononuclear cells.** *Proc Natl Acad Sci U S A* 2007, **104:**16245-16250.

327. Negishi H, Ohba Y, Yanai H, Takaoka A, Honma K, Yui K, Matsuyama T, Taniguchi T, Honda K: **Negative regulation of Toll-like-receptor signaling by IRF-4.** *Proc Natl Acad Sci U S A* 2005, **102:**15989-15994.

328. Jeffrey KL, Camps M, Rommel C, Mackay CR: **Targeting dual-specificity phosphatases: manipulating MAP kinase signalling and immune responses.** *Nat Rev Drug Discov* 2007, **6:**391-403.

329. Chi H, Barry SP, Roth RJ, Wu JJ, Jones EA, Bennett AM, Flavell RA: **Dynamic regulation of pro- and anti-inflammatory cytokines by MAPK phosphatase 1 (MKP-1) in innate immune responses.** *Proc Natl Acad Sci U S A* 2006, **103:**2274-2279.

330. Hammer M, Mages J, Dietrich H, Servatius A, Howells N, Cato AC, Lang R: **Dual specificity phosphatase 1 (DUSP1) regulates a subset of LPS-induced genes and protects mice from lethal endotoxin shock.** *J Exp Med* 2006, **203:**15-20.

331. Salojin KV, Owusu IB, Millerchip KA, Potter M, Platt KA, Oravecz T: **Essential role of MAPK phosphatase-1 in the negative control of innate immune responses.** *J Immunol* 2006, **176:**1899-1907.

332. Zhao Q, Wang X, Nelin LD, Yao Y, Matta R, Manson ME, Baliga RS, Meng X, Smith CV, Bauer JA, et al: **MAP kinase phosphatase 1 controls innate immune responses and suppresses endotoxic shock.** *J Exp Med* 2006, **203:**131-140.

333. Kuwata H, Matsumoto M, Atarashi K, Morishita H, Hirotani T, Koga R, Takeda K: **IkappaBNS inhibits induction of a subset of Toll-like receptor-dependent genes and limits inflammation.** *Immunity* 2006, **24:**41-51.

334. Ding SW, Voinnet O: **Antiviral immunity directed by small RNAs.** *Cell* 2007, **130:**413-426.

335. Umbach JL, Cullen BR: **The role of RNAi and microRNAs in animal virus replication and antiviral immunity.** *Genes Dev* 2009, **23:**1151-1164.

336. Garmire LX, Shen Z, Briggs S, Yeo G, Subramaniam S, Glass C: **Regulatory network of microRNAs in RAW 264.7 macrophage cells.** *Conf Proc IEEE Eng Med Biol Soc* 2010, **2010:**6198-6201.

337. Schulte LN, Eulalio A, Mollenkopf HJ, Reinhardt R, Vogel J: **Analysis of the host microRNA response to Salmonella uncovers the control of major cytokines by the let-7 family.** *EMBO J* 2011.

338. Liu G, Friggeri A, Yang Y, Park YJ, Tsuruta Y, Abraham E: **miR-147, a microRNA that is induced upon Toll-like receptor stimulation, regulates murine macrophage inflammatory responses.** *Proc Natl Acad Sci U S A* 2009, **106:**15819-15824.

339. Barenco M, Brewer D, Papouli E, Tomescu D, Callard R, Stark J, Hubank M: **Dissection of a complex transcriptional response using genome-wide transcriptional modelling.** *Mol Syst Biol* 2009, **5:**327.

340. Cheadle C, Fan J, Cho-Chung YS, Werner T, Ray J, Do L, Gorospe M, Becker KG: **Control of gene expression during T cell activation: alternate regulation of mRNA transcription and mRNA stability.** *BMC Genomics* 2005, **6:**75.

341. Elkon R, Zlotorynski E, Zeller KI, Agami R: **Major role for mRNA stability in shaping the kinetics of gene induction.** *BMC Genomics* 2010, **11:**259.

342. Rabani M, Levin JZ, Fan L, Adiconis X, Raychowdhury R, Garber M, Gnirke A, Nusbaum C, Hacohen N, Friedman N, et al: **Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells.** *Nat Biotechnol* 2011, **29:**436-442.

343. Becker S, Devlin RB, Haskill JS: **Differential production of tumor necrosis factor, macrophage colony stimulating factor, and interleukin 1 by human alveolar macrophages.** *J Leukoc Biol* 1989, **45:**353-361.

344. Liang J, Wang J, Azfer A, Song W, Tromp G, Kolattukudy PE, Fu M: **A novel CCCH-zinc finger protein family regulates proinflammatory activation of macrophages.** *J Biol Chem* 2008, **283:**6337-6346.

345. Carballo E, Lai WS, Blackshear PJ: **Feedback inhibition of macrophage tumor necrosis factor-alpha production by tristetraprolin.** *Science* 1998, **281:**1001-1005.

346. Liang J, Song W, Tromp G, Kolattukudy PE, Fu M: **Genome-wide survey and expression profiling of CCCH-zinc finger family reveals a functional module in macrophage activation.** *PLoS One* 2008, **3:**e2880.

347. Lehner MD, Schwoebel F, Kotlyarov A, Leist M, Gaestel M, Hartung T: **Mitogen-activated protein kinase-activated protein kinase 2-deficient mice show increased susceptibility to Listeria monocytogenes infection.** *J Immunol* 2002, **168:**4667-4673.

348. Dalton DK, Pitts-Meek S, Keshav S, Figari IS, Bradley A, Stewart TA: **Multiple defects of immune cell function in mice with disrupted interferon-gamma genes.** *Science* 1993, **259:**1739-1742.

349. O'Brien BA, Archer NS, Simpson AM, Torpy FR, Nassif NT: **Association of SLC11A1 promoter polymorphisms with the incidence of autoimmune and inflammatory diseases: a meta-analysis.** *J Autoimmun* 2008, **31:**42-51.

350. Blackwell JM, Searle S, Mohamed H, White JK: **Divalent cation transport and susceptibility to infectious and autoimmune disease: continuation of the Ity/Lsh/Bcg/Nramp1/Slc11a1 gene story.** *Immunol Lett* 2003, **85:**197-203.

351. Ewen ME: **Where the cell cycle and histones meet.** *Genes Dev* 2000, **14:**2265-2270.

352. Sester DP, Stacey KJ, Sweet MJ, Beasley SJ, Cronau SL, Hume DA: **The actions of bacterial DNA on murine macrophages.** *J Leukoc Biol* 1999, **66:**542-548.

353. Krausgruber T, Blazek K, Smallie T, Alzabin S, Lockstone H, Sahgal N, Hussell T, Feldmann M, Udalova IA: **IRF5 promotes inflammatory macrophage polarization and TH1-TH17 responses.** *Nat Immunol* 2011, **12:**231-238.

354. Ohmori Y, Hamilton TA: **Requirement for STAT1 in LPS-induced gene expression in macrophages.** *J Leukoc Biol* 2001, **69:**598-604.

355. Masumoto J, Yang K, Varambally S, Hasegawa M, Tomlins SA, Qiu S, Fujimoto Y, Kawasaki A, Foster SJ, Horie Y, et al: **Nod1 acts as an intracellular receptor to stimulate chemokine production and neutrophil recruitment in vivo.** *J Exp Med* 2006, **203:**203-213.

356. Inohara N, Ogura Y, Chen FF, Muto A, Nunez G: **Human Nod1 confers responsiveness to bacterial lipopolysaccharides.** *J Biol Chem* 2001, **276:**2551-2554.

357. Viala J, Chaput C, Boneca IG, Cardona A, Girardin SE, Moran AP, Athman R, Memet S, Huerre MR, Coyle AJ, et al: **Nod1 responds to peptidoglycan delivered by the Helicobacter pylori cag pathogenicity island.** *Nat Immunol* 2004, **5:**1166-1174.

358. Schroder K, Muruve DA, Tschopp J: **Innate immunity: cytoplasmic DNA sensing by the AIM2 inflammasome.** *Curr Biol* 2009, **19:**R262-265.

359. Rescigno M, Citterio S, Thery C, Rittig M, Medaglini D, Pozzi G, Amigorena S, Ricciardi-Castagnoli P: **Bacteria-induced neo-biosynthesis, stabilization, and surface expression of functional class I molecules in mouse dendritic cells.** *Proc Natl Acad Sci U S A* 1998, **95:**5229-5234.

360. Lee KW, Lee Y, Kim DS, Kwon HJ: **Direct role of NF-kappaB activation in Toll-like receptor-triggered HLA-DRA expression.** *Eur J Immunol* 2006, **36:**1254-1266.

361. Casals C, Barrachina M, Serra M, Lloberas J, Celada A: **Lipopolysaccharide up-regulates MHC class II expression on dendritic cells through an AP-1 enhancer without affecting the levels of CIITA.** *J Immunol* 2007, **178:**6307-6315.

362. Sicher SC, Chung GW, Vazquez MA, Lu CY: **Augmentation or inhibition of IFN-gamma-induced MHC class II expression by lipopolysaccharides. The roles of TNF-alpha and nitric oxide, and the importance of the sequence of signaling.** *J Immunol* 1995, **155:**5826-5834.

363. Lombardo E, Alvarez-Barrientos A, Maroto B, Bosca L, Knaus UG: **TLR4-mediated survival of macrophages is MyD88 dependent and requires TNF-alpha autocrine signalling.** *J Immunol* 2007, **178:**3731-3739.

364. Saito K, Hirai MY, Yonekura-Sakakibara K: **Decoding genes with coexpression networks and metabolomics - 'majority report by precogs'.** *Trends Plant Sci* 2008, **13:**36-43.

365. Horan K, Jang C, Bailey-Serres J, Mittler R, Shelton C, Harper JF, Zhu JK, Cushman JC, Gollery M, Girke T: **Annotating genes of known and unknown function by large-scale coexpression analysis.** *Plant Physiol* 2008, **147:**41-57.

366. Lee HK, Hsu AK, Sajdak J, Qin J, Pavlidis P: **Coexpression analysis of human genes across many microarray data sets.** *Genome Res* 2004, **14:**1085-1094.

367. Sweet MJ, Campbell CC, Sester DP, Xu D, McDonald RC, Stacey KJ, Hume DA, Liew FY: **Colony-stimulating factor-1 suppresses responses to CpG DNA and expression of toll-like receptor 9 but enhances responses to lipopolysaccharide in murine macrophages.** *J Immunol* 2002, **168:**392-399.