RNA-binding as a new regulatory dimension in the antiviral innate immune response

Louisa Hunnewell von Clemm Iselin

Hilary 2024



St Catherine's College
University of Oxford

A thesis presented for the degree of Doctor of Philosophy

Acknowledgements

Firstly, I'd like to thank my supervisors, Alfredo Castello, Shabaz Mohammed, and Peter Simmonds.

Thank you Alfredo for your support, guidance and belief in me over the past four years. Thank you for always making time when I needed it and for taking time to help me learn and grow as a scientist. I have learnt so much from you both about the science we do and what it means to do science. I feel privileged to have had the chance to be a member of the Castello lab.

Thank you Shabaz for giving me the opportunity to learn proteomics from you and your team. Thank you for always making time to join for my weekly meetings, even when there was no proteomics in sight. You have been the yin to Alfredo's yang in project discussions and I so value the balance that you have brought.

Thank you Peter for your patience and understanding in what has been a somewhat unorthodox setup. Thank you for hosting me during the early stages of my project and for your valuable input and encouragement in our meetings since then.

I would also like to thank all the members of the Castello and Mohammed labs who have supported me both scientifically and personally over the course of my PhD.

Thank you to Wael and Chen for your help in the lab at the start of my PhD and for helping me to lay the foundations of my project. And thank you Aino for starting me on my R journey. I don't known where I would have been without your guidance and troubleshooting support.

Thank you Yana for all your amazing help with the proteomics part of my project. I have so valued the time I have been able to spend learning from you at the RFI. Also, thank you for all the time you have taken (and buses you have missed) to help get my samples prepped and run.

Thank you Enzo both for your friendship and mentorship. I am constantly in awe of your incredible work ethic, your deep love of science, and your remarkable capacity for 'chill'. These are all things I hope to emulate in my career and I feel lucky to have such a valuable example of them in you.

Thank you Azman for all the chats, both serious and silly, that have helped me through the last couple of years. Now that you're here, I cannot imagine the lab without your wit, wisdom, and, most importantly, your winning smile.

Lucia, I feel so lucky that you have been a part of my PhD journey. There is no one I would have rather spent three months failing at iCLIP with. Thank you for being such an amazing role model and for all the long discussions about science and so much more. I'm so happy that we can keep working together and can't wait to hopefully craft a lovely story together.

Thank you to Natasha. I'm genuinely not sure how I would have gotten to this point without you. Thank you for being the most incredible desk-mate, bench-mate, sometime roommate, and, most importantly, friend. Thank you for being simultaneously my biggest source of distraction and most valued support in the lab. Moving to Glasgow would have been worth it just to have met you.

Thank you to my friends, who have supported me from afar and through extended periods of radio silence. Especially to Georgia and Marina. Thank you for 15(!) years of the most incredible friendship. Your brilliance and creativity never cease to amaze me.

Thank you to my wonderful family. Thank you Mum and Dad for your love and support always. Ollie and Welles, thank you for being my biggest fans and for putting up with my craziness over the years. Also, thank you to Binti for being our UK home, wherever you are. I love you all so much.

Finally, thank you to Josh. There are no words that can adequately express what your support has meant to me over the past 4 years and, in particular, over the last three months. Thank you for coming along for every step of this wild ride. I love you.

Abstract

The type I interferon (IFN-I) response dramatically alters the intracellular environment to restrict virus infection. RNA-binding proteins (RBPs) are central to this phenomenon, sensing viral RNA (vRNA) and restricting infection at every stage of the vRNA's lifecycle. However, systematic profiling of the RNA-bound proteome (RBPome) in response to IFN-I is still lacking. To fill this gap, I combined RNA interactome capture (RIC) with IFN-I treatment and discovered over 250 IFN-regulated RBPs (IR-RBPs). Amongst upregulated IR-RBPs I identified known antiviral factors such as OAS3 and PKR, together with dozens of RBPs with unknown roles in immunity. Functional assays revealed that proteins upregulated by IFN-I play antiviral roles, while the opposite is observed for certain downregulated proteins. The majority of IR-RBPs do not show changes in protein abundance, indicating that their RNA-binding activity is being regulated post-translationally. Phosphoproteomic analysis of RIC eluates indicated that some of these post-translational events were a result of altered phosphorylation state. Using phosphomimetic mutants, combined with a dual-fluorescence RNA binding assay, I confirmed that phosphorylation was directly regulating RNA-binding activity for one antiviral IR-RBP, MATR3. To further determine what MATR3's antiviral role might be, I assessed how its RNA and protein interaction landscapes were altered by infection. I found that it associated with the 5' end of the Sindbis virus (SINV) genome and engaged with several ISGs with roles in translation control, including IFIT proteins. In summary, my work reveals a new level of IFN-I-mediated regulation, affecting the activity of many cellular RBPs. Studying these IR-RBPs can offer new insight into how IFN-I establishes an antiviral state.

Contents

List of Figures **9**

List of Tables **11**

1 Introduction **16**

1.1 The Type I IFN Response 16

1.1.1 Induction of IFN-I through pathogen sensing 16

1.1.2 IFN-I induced signalling 20

1.1.3 Antiviral effectors of the IFN-I response 24

1.2 RNA-binding proteins 28

1.2.1 Functional consequences of RNA-protein interactions 29

1.2.2 Properties of protein-RNA interfaces 33

1.2.3 An expanding universe of RNA-protein interaction dynamics 37

1.2.4 RBPs in viral infection 40

1.3 Aims 42

2 RNA-binding activity as a new regulatory layer in the type I IFN response **43**

2.1 Introduction 43

2.2 Results 45

2.2.1 The RNA-bound proteome is regulated by IFN-I treatment 45

2.2.2 A targeted screen to identify antiviral IR-RBPs 56

2.3 Discussion 64

2.3.1 Identifying new facets of known IFN-I regulatory mechanisms 64

2.3.2 Uncovering novel players in the antiviral immune response 67

3 A role for phosphorylation in regulating IFN-I-driven changes in RNA-binding activity **75**

3.1 Introduction 75

3.2 Results 78

3.2.1 Establishing the IFN-regulated RNA-bound phosphoproteome 78

3.2.2 MATR3's phosphorylation state alters its RNA-binding activity 88

3.2.3 What is responsible for MATR3's differential phosphorylation state? . . . 91

3.3 Discussion 94

3.3.1 Phosphoenrichment of RIC eluates is a new route to discovering functional phosphosites . . . 94
3.3.2 MAP4 S822: a new instance of phospho-driven riboregulation? . . . 97
3.3.3 GIGYF2 S30: modulating the 4EHP interface . . . 99
3.3.4 MATR3 S188: learning new tricks from an old phosphosite . . . 101
3.3.5 What is responsible for IFN-I-driven changes in phosphorylation? . . . 104

4 Establishing the interactome of MATR3 in virus infected cells 106
4.1 Introduction . . . 106
4.2 Results . . . 109
4.2.1 MATR3 alters its protein interaction network upon infection . . . 109
4.2.2 Using iCLIP to define MATR3's RNA interaction partners during infection 120
4.2.3 Does S188 phosphorylation state alter MATR3's binding specificity? . . . 128
4.2.4 Characterising the vRNA interactions of RBPs that regulate infection . . . 131
4.3 Discussion . . . 134
4.3.1 MATR3 PPI uncovers a wide network of potential RNA-dependent and independent interactions . . . 134
4.3.2 MATR3 alters its interaction network upon SINV infection . . . 136
4.3.3 MATR3's PRI motif is important for its antiviral activity . . . 141
4.3.4 MATR3 iCLIP data successfully recapitulates known binding properties . 142
4.3.5 vRNA interaction patterns offer insight into RBP functions during infection 143
4.3.6 Developing working models of MATR3's antiviral function . . . 145

5 General discussion and future directions 148
5.1 Shifting from an 'abundance' to an 'activity'-driven understanding of IFN-I regulation . . . 148
5.2 Expanding the repertoire of IFN-regulated RBPomes . . . 149
5.3 Beyond phosphorylation: probing PTMs as a general mechanism for IR-RBP control . . . 151
5.4 Mapping viral ribonucleoproteins: a new frontier in host-pathogen interactions . 153
5.5 Summary . . . 156

6 Materials 157
6.1 Reagents & Consumables . . . 157

6.2 In-House Buffers . . . 160
6.3 Antibodies & Dyes . . . 162
6.4 Plasmids . . . 163
6.5 Oligonucleotides . . . 164
6.6 Cell lines . . . 166
6.7 Instruments & Equipment . . . 167

7 Methods **168**

7.1 Cell Biology . . . 168
7.1.1 Maintenance of Cells . . . 168
7.1.2 Stable Transfection of HEK293 Flip/In T-REx . . . 168
7.1.3 siRNA Knockdowns . . . 168
7.1.4 Generation of viruses . . . 169
7.1.5 Titration of viruses . . . 170
7.1.6 Platereader assay . . . 171
7.1.7 Confocal microscopy . . . 171
7.2 Biochemical and molecular biology techniques . . . 172
7.2.1 Western blotting . . . 172
7.2.2 Silver staining . . . 172
7.2.3 Mutagenesis . . . 172
7.3 Dual fluorescence assay . . . 173
7.3.1 Sample harvesting . . . 173
7.3.2 IP and probe hybridisation . . . 174
7.3.3 Fluorescence measurement . . . 174
7.4 RNA Interactome Capture . . . 174
7.4.1 Sample harvesting . . . 174
7.4.2 IFN RIC . . . 175
7.4.3 Sample prep . . . 175
7.4.4 Phosphoenrichment . . . 176
7.4.5 Mass Spectrometry . . . 176
7.5 Protein-protein interaction analysis . . . 177
7.5.1 Sample harvesting . . . 177
7.5.2 co-IP . . . 177

7.5.3 Sample prep . . . 177
7.5.4 Mass Spectrometry . . . 178
7.6 iCLIP2 . . . 178
7.6.1 Sample harvesting . . . 178
7.6.2 iCLIP2 . . . 179
7.6.3 Sequencing . . . 180
7.7 Data analysis . . . 181
7.7.1 General . . . 181
7.7.2 IFN RIC . . . 181
7.7.3 IFN RIC phosphoproteomics . . . 182
7.7.4 Protein-protein interaction . . . 182
7.7.5 iCLIP . . . 183

8 Appendix 1 **185**

9 References **198**

List of Figures

1.1 Pathogen sensing pathways that induce IFN-I. 17
1.2 Signalling cascades induced by IFN-I. 20
1.3 ISG-induced antiviral processes. 24
1.4 RNA is regulated by and regulates RBPs. 29
1.5 RNA and protein-driven determinants of RBP binding specificity. 33
1.6 RBPs are involved in every stage of the viral lifecycle. 40
2.1 Schematic of IFN RIC protocol. 45
2.2 Quality control analysis of the RIC experiment applied to IFNα2 treated cells. . . 46
2.3 IFN RIC samples were analysed via LC-MS/MS. 48
2.4 Changes in the RNA-bound proteome become more pronounced with time. . . . 49
2.5 Changes in the RBPome do not correlate with changes in the transcriptome or proteome. 50
2.6 IR-RBPs are enriched in proteins linked to RNA processing and innate immunity. 51
2.7 Many IR-RBPs are members of the core vRNA interactome. 53
2.8 Groups of IR-RBPs show distinct kinetic profiles. 54
2.9 Features of candidates selected for a target screen. 56
2.10 Generating GFP-tagged overexpression lines. 59
2.11 Fluorescence-based tools for studying viral fitness. 60
2.12 Overexpression lines have varying effects on infection. 61
2.13 siRNA knockdowns provide complementary information on antiviral activity. . . . 62
3.1 Schematic of phosphoenrichment of RIC samples. 78
3.2 Zr-IMAC successfully enriched for phosphopeptides. 79
3.3 The RNA-bound phosphoproteome is altered upon IFN treatment. 80
3.4 Many changes in phosphopeptide abundance track changes in protein abundance. 81
3.5 Some proteins exhibit anti-correlation between RNA-binding activity and phosphorylation state. 82
3.6 A profile of MAP4's phosphorylation and RNA-binding properties. 83
3.7 A profile of GIGYF2's phosphorylation and RNA-binding properties. 84
3.8 Structural prediction of GIGYF2 S30's interaction with 4EHP 85
3.9 A profile of MATR3's phosphorylation and RNA-binding properties. 87
3.10 Schematic of dual-fluorescence RNA-binding assay. 88

3.11 Characterising MATR3 phosphomutant lines. 89
3.12 MATR3 phosphomutants show altered RNA-binding activity. 90
3.13 The overall pattern of PKA phosphorylation is not altered by IFN-I treatment. . . 91
3.14 AKAP8 and PPP1R10 do not appear to be responsible for MATR3's altered phosphorylation. 92
4.1 Schematic of comparisons in MATR3 PPI analysis. 109
4.2 Silver stain of MATR3 IP. 110
4.3 Quality control of MATR3 IP LC-MS/MS data. 111
4.4 Many proteins are enriched in the MATR3 IP over the parental control. 112
4.5 MATR3 has both RNA-dependent and independent interactions. 113
4.6 MATR3 alters its interaction partners following infection. 114
4.7 Comparing MATR3 IP with the whole cell proteome. 115
4.8 MATR3's SINV-specific interaction partners are linked to infection. 116
4.9 STRING networks of SINV-modified MATR3 interactions. 118
4.10 Disruption of MATR3/ PTBP interaction site alters SINV infection. 119
4.11 Schematic of the iCLIP protocol. 120
4.12 Principal Component Analysis of MATR3 WT iCLIP samples. 121
4.13 MATR3 has a core set of targets in mock and infected cells. 122
4.14 MATR3 primarily binds introns of protein-coding mRNA. 123
4.15 MATR3 binds to pyrimidine-rich motifs. 124
4.16 Secondary structure preferences of MATR3. 125
4.17 MATR3's binding profile on SINV RNA. 126
4.18 MATR3 binds to SINV's conserved 51nt 5' CSE. 127
4.19 There is overlap in binding targets of WT MATR3 and phosphomutants. 128
4.20 Phosphomutants share WT MATR3's transcript preferences. 129
4.21 Binding specificity is conserved in phosphomutants. 130
4.22 Phosphomutant viral RNA binding pattern is the same at WT MATR3. 130
4.23 Comparing iCLIP datasets for RBPs with roles in SINV infection. 131
4.24 RBPs show diverse binding profiles on SINV RNA. 133

List of Tables

2 Reagents & Consumables 157
3 In-House Buffers 160
4 Antibodies & Dyes 162
5 Plasmids 163
6 Oligonucleotides 164
7 Cell Lines 166
8 Instruments & Equipment 167
9 Linearisation reaction mix 169
10 T7 RNA polymerase mix 169
11 Lipofectamine 3000 transfection Tube 1 170
12 Lipofectamine 3000 transfection Tube 2 170
13 PCR mutagenesis mix composition 173
14 PCR mutagenesis thermocycling conditions 173
15 L3-IR-App adapter ligation mix 179
16 IFN RIC significance results 185

List of Abbreviations

Abbreviation	Definition
Å	Angstrom
AS	Alternative splicing
bp	Base pair
CBC	Cap-binding complex
cDNA	Complementary DNA
CDS	Coding sequence
cGAS	Cyclic GMP–AMP synthase
CHIKV	Chikungunya virus
CLIP	Crosslinking and immunoprecipitation
Co-IP	Co-immunoprecipitation
CTD	Carboxy terminal domain
Da	Dalton (unit)
DC	Dendritic cells
DENV	Dengue virus
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleoside triphosphate
dsRNA	Double-stranded RNA
EBOV	Ebola virus
eGFP	Enhanced green fluorescence protein
eIF	Eukaryotic initiation factor
EJC	Exon-junction complex
FBS	Foetal bovine serum
FDR	False discovery rate
FITR	Flp-In T-REx
FSK	Forskolin
GFP	Green fluorescent protein
GO	Gene ontology
gRNA	Genomic RNA

Abbreviation	Definition
HBV	Hepatitis C Virus
HCV	Hepatitis B Virus
HIV-1	Human immunodeficiency virus type 1
hnRNP	Heterogeneous nuclear ribonucleoprotein
hpi	Hours post infection
hpt	Hours post treatment
IAV	Influenza A virus
IDR	Intrinsically disordered region
IFN	Interferon
IFNAR	Interferon alpha receptor
IP	Immunoprecipitation
IR-RBP	IFN-regulated RNA-binding protein
IRF	IFN regulatory factor
ISG	IFN stimulated gene
ISRE	IFN stimulated response elements
JAK	Janus kinases
KD	Knockdown
KH	K-homology domain
KO	Knockout
KSHV	Kaposi's sarcoma-associated herpesvirus
LC-MS/MS	Liquid chromatography tandem mass spectrometry
lncRNA	Long non-coding RNA
MAPK	Mitogen-activated protein kinases
miRNA	microRNA
MLV	Moloney murine leukemia virus
MOI	Multiplicity of infection
mRNA	Messenger RNA
MS	Mass spectrometry
NCL	Non-crosslinked
nts	Nucleotides
ORF	Open reading frame

Abbreviation	Definition
PAMP	Pathogen associated molecular patterns
PCA	Principle component analysis
PCR	Polymerase chain reaction
pDC	Plasmacytoid dendritic cell
PKA	Protein kinase A
PKR	Protein kinase R
poly(A)	Polyadenylated
PPI	Protein-protein interaction
PPP	Triphosphate
PRR	Pattern recognition receptors
PTM	Post-translational modification
R	R programming language
RBD	RNA binding domain
RBP	RNA-binding protein
RBS	RNA binding site
RFP	Red fluorescent protein
RIC	RNA interactome capture
RIG-I	Retinoic acid-inducible gene I
RLR	RIG-I-like receptors
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
RRM	RNA recognition motif
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
sfRNA	Subgenomic flaviviral RNA
sgRNA	Subgenomic RNA
SINV	Sindbis virus
siRNA	Small interfering RNA
SMI	Size matched input
ssDNA	Single-stranded DNA
ssRNA	Single-stranded RNA

Abbreviation	Definition
STAT	Signal transducers and activators of transcription
STING	Stimulator of interferon genes
TLR	Toll-like receptors
UTR	Untranslated region
UV	Ultraviolet
vDNA	Viral DNA
vRNA/vRNP	Viral RNA/viral ribonucleoprotein
WCL	Whole cell lysate
WCP	Whole cell proteome
WNV	West Nile virus
WT	Wild-type
XL-MS	Crosslinking mass spectrometry
ZIKV	Zika virus
ZnF	Zinc finger

1 Introduction

1.1 The Type I IFN Response

The concept of an antiviral 'interferon' system was first introduced by Isaacs and Lindenmaan in 1957. In their seminal study, they reported that the fluid surrounding influenza A virus (IAV) infected chick embryos contained a non-viral protein that rendered noninfected cells resistant to IAV challenge [1]. Since this discovery, it has become clear that there is not a single 'interferon' (IFN) but rather a family of proteins that contribute to different aspects of immunity. IFNs are divided into three families: type I, type II, and type III [2]. The type I IFN (IFN-I) family consists of 17 proteins: 13 IFNα subtypes, IFNβ, IFNϵ, IFNκ, and IFNω. These proteins all bind to the heterodimeric IFNα receptor (IFNAR) to induce an intracellular antiviral program [3]. The type III IFNs (IFNλ1-4) induce a similarly antiviral program as a result of signalling through the heterodimeric IFNLR1/IL10R2 complex [4, 5]. Unlike IFNAR, which is expressed almost ubiquitously, the IFN-III receptor complex is expressed only in cells at mucosal barrier sites, meaning that the antiviral effects of IFN-IIIs are more spatially restricted [6, 7]. In contrast to type I and III IFNs, the IFN-II family consists of only one member, IFNγ, which signals through the IFNγ receptor (IFNGR). This cytokine is produced by cells of the innate and adaptive immune system and induces a pro-inflammatory response [8]. The rest of this thesis will be focused on IFN-Is, due to their central and ubiquitous role in the intracellular antiviral response.

1.1.1 Induction of IFN-I through pathogen sensing

Type I IFNs are produced following sensing of pathogen associated molecular patterns (PAMPs) by pattern recognition receptors (PRRs). Many of the PRRs that sense viral infection do so by recognising unusual features of viral nucleic acids. While this is not the only means of viral sensing, I will focus here on nucleic acid recognition receptors [9]. These include the RIG-I-like receptors (RLRs) that sense cytosolic double stranded (ds)RNA, the Toll-like receptors (TLRs) that sense endosomal RNA and DNA, and the cytosolic DNA sensors, cyclic GMP–AMP synthase (cGAS) and IFI16 [10]. Simplified signalling cascades for each of these proteins are outlined in Figure 1.1.

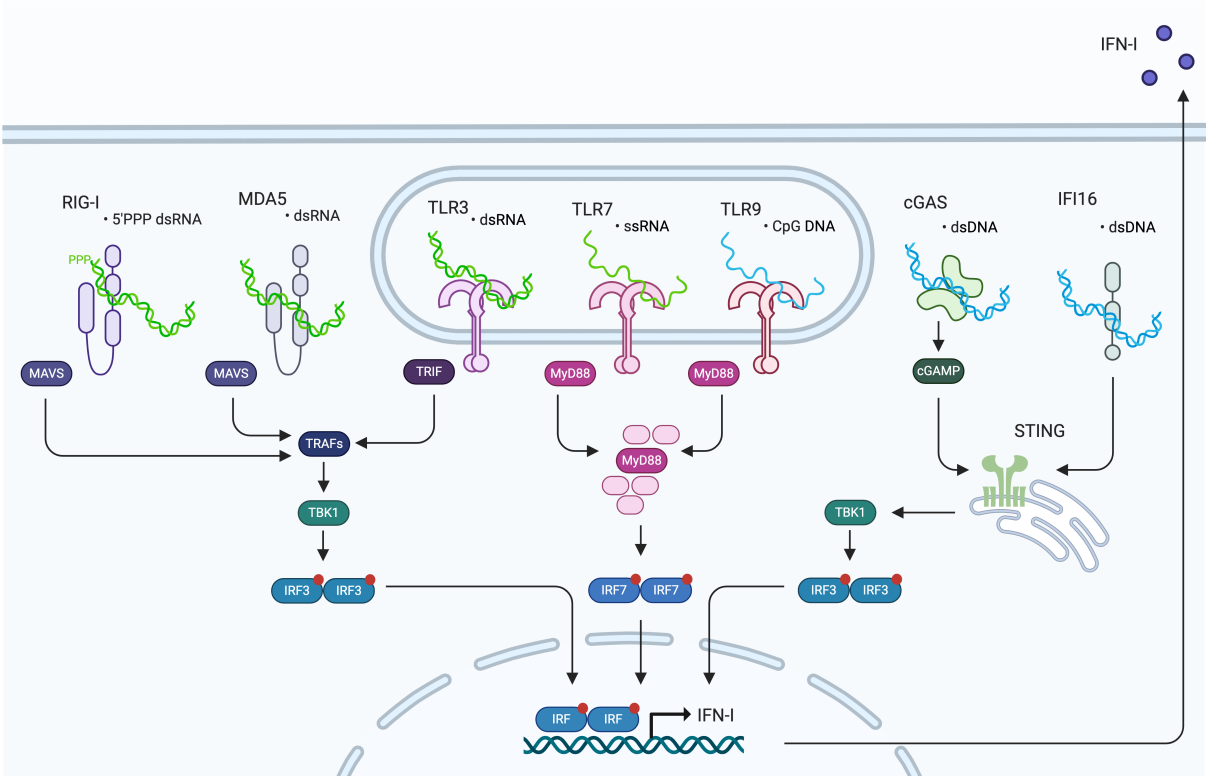


Figure 1.1: **Pathogen sensing pathways that induce IFN-I.** ***RLR pathway.*** RIG-I and MDA5 bind to 5' PPP ends in dsRNA and long stretches of dsRNA, respectively and both activate MAVS upon ligand recognition. MAVS recruits TRAF proteins upon activation, which recruits the kinase TBK1. TBK1 phosphorylates MAVS, which results in IRF3 recruitment of TBK1. TBK1 then directly phosphorylates IRF3, allowing for its dimerisation and nuclear translocation. ***TLR3 pathway.*** TLR3 is activated by dsRNA in the endosomal compartment. This results in activation of TRIF and recruitment of TRAF3. TRAF recruits TBK1, which promotes the phosphorylation and dimerisation of IRF3. ***TLR7/9 pathway.*** TLR7 and TLR9 sense endosomal RNA and unmethylated CpG in DNA, respectively and both signal through MyD88. When activated, MyD88 forms a complex with a group of proteins that includes IRAKs, TRAFs, IKKα, and IRF7. Phosphorylation of IRF7 in this context allows for dimerisation, nuclear translocation and IFN-I production. ***STING pathways.*** cGAMP is produced by cGAS in response to sensing cytosolic DNA. cGAMP binds to ER-associated STING, causing its relocalisation to the Golgi. In the Golgi, STING's C-terminal tail is phosphorylated by TBK1, creating a docking site for IRF3, which is brought into proximity with TBK1, resulting in its phosphorylation. IFI16 directly interacts with STING to promote its activation and boost cGAS signalling.

Nucleic acid binding PRRs can be divided into DNA and RNA sensors, both of which are involved in detecting DNA and RNA virus infections [11]. This reflects a level of redundancy in this essential system and speaks to the convergence of all these pathways on the central process of IFN-I production. This convergence begins at the level of interferon regulatory factor (IRF) phosphorylation. All IFN-I-inducing PRR signalling pathways described here culminate in the phosphorylation and homodimerisation of IRF3 or IRF7. This phosphorylation and

dimerisation results in IRF translocation into the nucleus, where they bind to the promoters of IFN-I genes and drive transcription [10]. Which IFN-I subtypes are produced is a function of cell type and IRF dimer composition. IRF3 is expressed across the majority of cell types and primarily drives IFNβ production [12]. IRF7, which is highly expressed in plasmacytoid dendritic cells (pDCs) favours IFNα production [13]. It is important to note that IFN-I production is often not the only consequence of viral sensing. Many of these PRRs induce a parallel NFκB driven pro-inflammatory signalling program which can also contribute to antiviral immunity [9].

RLRs. RIG-I and MDA5 are both members of the RLR family. They consist of two amino terminal caspase activation and recruitment domains (CARDs), a central helicase domain, and a carboxy terminal domain (CTD) [14]. Despite their shared architecture, RIG-I and MDA5 have different specificities, resulting distinct roles in pathogen sensing [15–18]. The primary target of RIG-I binding is dsRNA with a triphosphate or diphosphate end that lacks 2'-O-methylation on its first nucleotide [19–23]. This signature is found in both positive sense and negative sense single-stranded RNA (+ssRNA and -ssRNA) viruses [24, 25]. The ligand for MDA5 is less well defined but it is generally thought to recognise long, dsRNA, which is a common feature of the replication cycles of many viruses [18, 26]. Interestingly, cellular RNA targets that are released or exposed upon infection have been identified for both RIG-I and MDA5. For example, the cellular RNA5SP141 pseudogene is mis-localised and unmasked following herpesvirus infection, allowing for activation of RIG-I by a DNA virus [27]. MDA5 is activated by dsRNA released from damaged mitochondria, which frequently occurs during infection [28]. RIG-I and MDA5 both signal through mitrochondrial antiviral signalling protein (MAVS) in a signalling cascade the culminates in IRF3 phosphorylation and dimerisation (Figure 1.1) [29–31].

TLRs. TLRs are a group of ten type 1 membrane spanning receptors that sense a wide range of PAMPs [32]. They consist of leucine-rich repeats that determine ligand specificity, a cytoplasmic TIR domain, and a transmembrane domain, allowing for localisation either in plasma or endosomal membranes. TLRs 3, 7, 8, and 9 are all located in endosomal membranes and are involved in sensing viral nucleic acids. TLR3 was initially identified for its ability to induce NFκB and IFN signalling in response to the synthetic dsRNA mimic, poly(I:C) [33]. It is now known to induce signalling in response to regions of dsRNA greater than 90bp in length, which are frequently produced in the course of both DNA and RNA virus

infection [34]. TLR3 signals through a TRIF/TRAF cascade that leads to IRF3 phosphorylation (Figure 1.1). While TLR3 is expressed across a wide range of tissues [32], TLRs 7 and 9 are primarily localised to pDCs. These cells are 'professional' IFN-I producers, releasing large quantities of IFNα in response to sensing of viral infection. TLR7 binds to GU or AU-rich sequences in single-stranded vRNA located within the cell's endosomal compartments [35]. TLR9 binds to DNA containing unmethylated CpG motifs found in the DNA of bacteria and viruses [36, 37]. Both TLR7 and 9 induce IFN-I through a MyD88-driven signalling cascade (Figure 1.1) [38]. The endosomal localisation of these PRRs means that they primarily sense incoming viral nucleic acids. However, endosomal TLRs can also sample the cytosolic contents of their own or neighbouring cells through autophagy [39], or through cell-to-cell transfer of vRNA-containing exosomes [40]

Cytosolic DNA sensors. While TLR9 can sense vDNA within endosomes, there is also an important set of pathways that are responsible for cytosolic DNA sensing. The best characterised of these is the cGAS/STING pathway. cGAS is an enzyme that is activated by binding long regions of cytosolic DNA [41]. When activated, cGAS generates the second messenger 2'3'-cGAMP, which binds to and activates STING [42, 43]. The cGAS/ STING pathway has been implicated in sensing viral DNA from a range of DNA viruses and retroviruses. Additionally, it has been implicated in RNA virus sensing. In DENV infection, damage to mitochondria results in release of mtDNA, which is sensed by cGAS [44].

cGAS is not the only PRR that drives STING activation. The PYHIN protein, IFI16, can also activate STING in response to viral DNA binding in both the nucleus and the cytoplasm [45]. IFI16 appears to form puncta early in HSV infection, oligomerising on viral DNA to promote signalling [46]. STING activation by cGAS and IFI16 is cooperative. Both were found to be necessary for complete activation of the innate immune response to DNA virus infection in keratinocytes [47] and macrophages [48]. IFI16 can both directly interact with STING to promote translocation and phosphorylation, and can boost cGAS-driven signalling [47, 48]. Recent work has also hinted at a role for IFI16 in RNA sensing. It has been identified in association with the genomes of several RNA viruses, including CHIKV [49], SINV [50], and IAV [51], and appears to promote RIG-I-driven IFN-I production [51].

1.1.2 IFN-I induced signalling

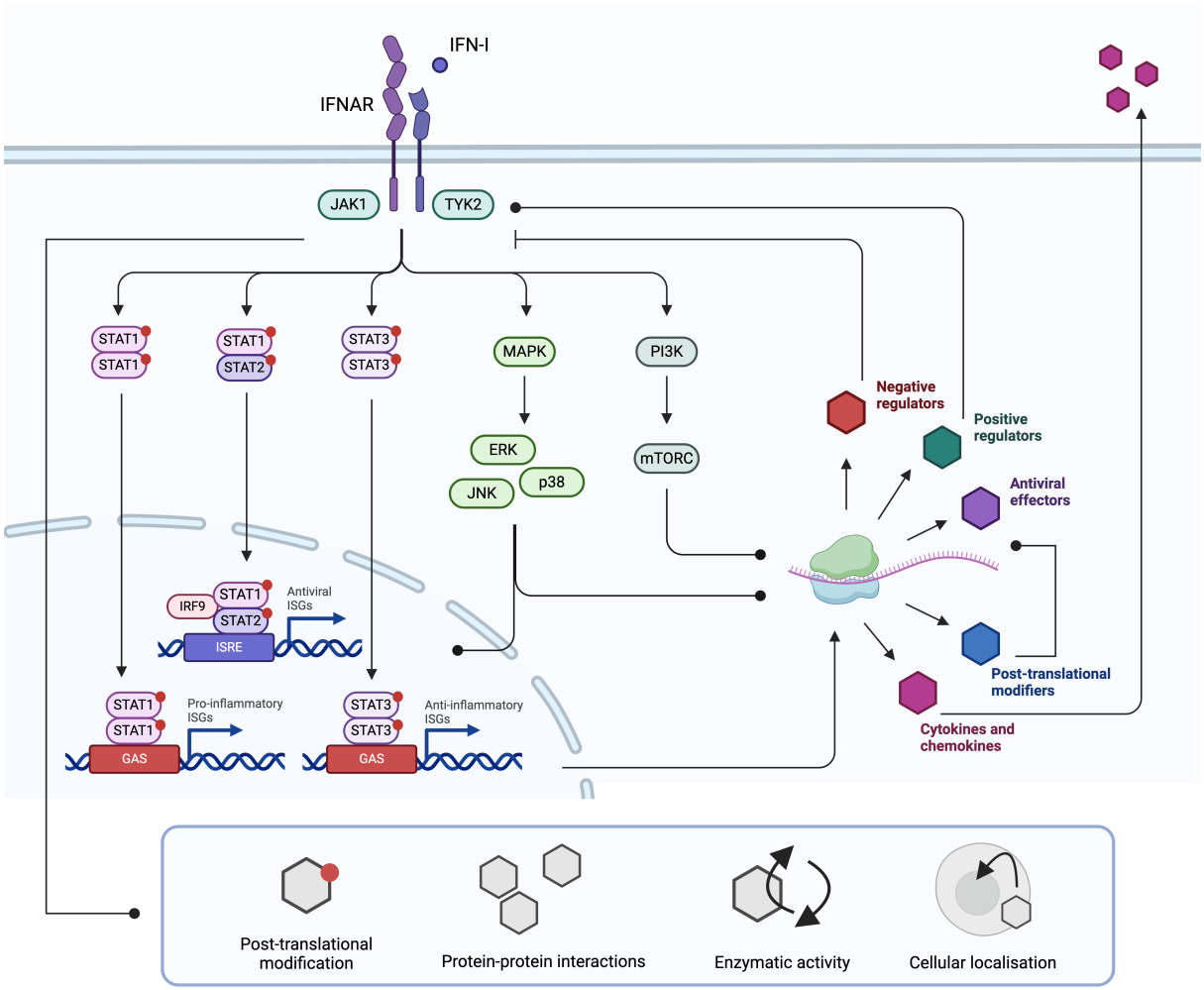


Figure 1.2: **Signalling cascades induced by IFN-I.** IFN-I binding to IFNAR gives rise to phosphorylation of JAK1 and TYK2. These can in turn phosphorylate STAT proteins to form either STAT1/STAT2, STAT1/STAT1 or STAT3/STAT3 dimers. STAT1/STAT2 binds to IRF9 in the nucleus to form ISGF3, which binds to ISREs and induces ISG expression. STAT1 and STAT3 homodimers bind to GAS elements to induce pro-/anti-inflammatory genes. STAT-induced mRNAs can then be translated into antiviral factors and regulators of the IFN-I response. IFNAR activation can also lead to activation of MAPK and PI3K pathways, which regulate transcription and translation. IFNAR signalling can also give rise to changes in activity of pre-existing proteins. These changes include: altered PTMs, altered PPIs, altered enzymatic activity, and altered cellular localisation.

IFN-I-induced transcriptional response. Following their synthesis and release, IFN-Is bind to the ubiquitously expressed IFNAR receptor. This receptor consists of two subunits (IFNAR1 and IFNAR2), which form a ternary complex upon binding of IFN-I [52]. Janus kinases (JAKs), bound to the cytosolic tail of IFNAR chains, are brought into proximity upon complex formation, allowing for their trans-phosphorylation and activation [53]. TYK2 and JAK1 are

bound to IFNAR1 and IFNAR2, respectively. Once activated, they phosphorylate tyrosine residues in the cytosolic region of the IFNAR subunit to which they are bound [54]. These phosphorylated tyrosine residues represent a platform for STATs (signal transducers and activators of transcription) to bind and become phosphorylated in turn. STATs are a group of seven proteins that are involved in downstream transcriptional regulation of a range of cytokine receptors [54]. In the canonical IFN-I signalling pathway, STAT1 and STAT2 are tyrosine phosphorylated by TYK2 and JAK1, causing them to dimerise.

Following dimerisation, STAT1/STAT2 translocate to the nucleus and bind to IRF9 to form the ISGF3 complex. This complex binds to IFN stimulated response elements (ISREs) within the promoters of IFN stimulated genes (ISGs), enhancing their transcription. While ISGF3 is central to ISG induction, it does not operate in isolation. In addition to recruiting factors involved in transcription initiation, it also recruits chromatin remodelling and modifying complexes [12]. Chromatin accessibility of ISG promoters is generally low under resting conditions but binding of ISGF3 has been shown through ChIP-seq to increase promoter accessibility [55]. This is mediated in part through recruitment of the nucleosome remodelling complex, BAF, which directly interacts with STAT2 and is essential for ISG induction [56].

ISGF3 can also act in concert with a range of other transcription factors, particularly IRFs, which fine-tune the IFN-I-induced transcriptional program. IRFs 1, 7, 8, and 9 are themselves transcriptionally induced by ISGF3 and have been proposed to participate in driving a delayed second wave of ISG induction [57]. Non-IRFs have also been implicated in transcription of specific ISG subsets. For example, there is a group of ISGs that is dependent on both ISGF3 and PLZF (another transcription factor) for induction. Interestingly, this effect is reliant on phosphorylation of PLZF by a member of the ERK family of mitogen-activated protein kinases (MAPKs) [58]. As discussed below, ERK activation is another pathway induced as part of the IFN-I response.

Other STAT dimers can also form as a result of IFNAR signalling, leading to parallel transcriptional programs that alter the nature of the antiviral response. Phosphorylated STAT1 and STAT3 form homodimers in response to IFN-I. The STAT1 homodimer is also known as gamma interferon activation factor (GAF) as it is also generated by IFNγ signalling. GAF translocates to the nucleus and binds to gamma interferon activation site (GAS) elements found in the promoters of genes that contribute to an inflammatory response [2]. Therefore, an IFN-I response in which STAT1 homodimerisation is prevalent will have a

more inflammatory signature [59]. On the other hand, STAT3 homodimers reduce IFN-I responsiveness by promoting transcription of negative regulators [60]. The balance of STAT1/STAT2, STAT1/STAT1 and STAT3/STAT3 dimers is driven by various factors, including relative STAT expression levels [61] and the phosphorylation state of STATs. For example, STAT1 phosphorylation by IKKϵ at S708 suppresses homodimer formation, thereby promoting ISGF3 formation and canonical ISG induction [62]. Of note, unphosphorylated STATs also play a role in transcription control. For example, a STAT2-IRF9 heterodimer has been found to sit on the promoters of ISGs in the absence of IFN-I signalling and result in low-level induction of these genes [63]. Together, these results indicate that the regulation of the IFN-I response by STATs stretches well beyond the canonical ISGF3 complex.

Post-transcriptional and post-translational regulation in the IFN-I response. Alongside the canonical JAK/STAT transcriptional response, there are STAT-independent IFN-driven signalling cascades. The best characterised of these are the MAPK and PI3K/AKT/mTOR pathways. These kinases can contribute to transcriptional regulation, as has been demonstrated for p38 MAPK [64]. However, another important part of their role is in IFN-I-driven translation regulation. Mammalian target of rapamycin (mTOR) positively regulates ISG translation directly through phosphorylating 4E-BP1, which drives formation of the translation initiation complex, and indirectly through phosphorylating 6SK. Activated 6SK phosphorylates eIF4B, enhancing ISG protein production [65]. 6SK can also be activated by MEK/ERK signalling, indicating that this mechanism is a convergence of mTOR and MAPK-driven cascades. ERK1/2 have also been found to contribute to Mnk1/2 activation, which leads to increased eIF4E phosphorylation. This in turn leads to increased translation of ISGs, including ISG15 and IFIT2 [66].

Alterations in the PTM landscape have also been identified as a means of IFN-I-driven regulation, independently of JAK/STAT signalling and ISG translation control. A study assessing phosphorylation in IFN-treated astrocytes and microglia found hundreds of differential phosphorylation events [67]. Differentially phosphorylated proteins were involved not only in canonical IFN-I signalling, but also in a wide range of other signalling pathways, reflecting other dimensions of the IFN-I response. Another post-translational modification that has been profiled in the context of the IFN-I response is ISGylation. This modification involves the addition of ISG15 to proteins in a process that is thought to occur co-translationally as part

of the IFN-I response [68]. Hundreds of proteins have been identified as ISGylation targets, many of which have not been linked to IFN signalling, indicating that regulation is occurring beyond the bounds of the canonical IFN-I response [69].

Regulation beyond ISGs has also been identified at the level of protein-protein interactions. One study that performed protein correlation profiling found thousands of altered protein-protein interactions in IFN-I treated cells [70]. The vast majority of these new interactions involved proteins whose abundance were not substantially altered by IFN-I (i.e., not ISGs). A similar result was observed in a study focused specifically on the interaction partners of exogenously expressed IFIT proteins. IFITs interacted with very few proteins under resting conditions but, upon IFN-I treatment, they obtained a network of interaction partners, the majority of which were not ISGs [71]. These altered protein-protein interactions might be driven by a variety of factors, including changes in the PTM landscape and changes in proteins' localisation patterns. How these events are regulated by the IFN-I response has yet to be elucidated.

Negative regulation of the IFN-I response. In addition to activating an antiviral program, the IFN-I response also induces mechanisms that will ultimately lead to its resolution. The negative regulatory mechanisms of the IFN-I response are as ancient as the response itself, speaking to their essential role in maintaining homeostasis and survival [72]. In the absence of appropriate negative regulation, the IFN-I response can result in excess tissue damage and drive autoimmune disease [73]. Indeed, people with in-born errors in these negative feedback pathways exhibit severe neurological pathology [74].

The first level of negative regulation occurs prior to initiation of the ISG transcriptional program [12]. IFNAR is rapidly internalised and degraded following activation through a p38 MAPK-driven process [75]. Additionally, co-stimulation of cells by TLR/ITAM agonists leads to activation of the phosphatase SHP2, which acts on STAT proteins to prevent their dimerisation and downstream signalling [76]. Protein inhibitor of activated STAT (PIAS) proteins also regulate the downstream signalling capacity of STATs. PIAS1 restricts STAT1's interaction with ISREs through directing SUMOylation dependent and independent pathways [77].

ISGs, including SOCS proteins and USP18, can also directly contribute to negative regulation [12]. SOCS proteins bind to phosphorylated tyrosine residues either on IFNAR or on JAKs, acting as a competitor for STAT binding and an inhibitor of JAK activity. Additionally,

by binding to phosphorylated IFNAR, they can promote its ubiquitination and degradation [72]. USP18 is the enzyme responsible for removing ISG15 conjugates [68]. Mice lacking USP18 exhibit much higher levels of ISGylation and are much more sensitive to IFN-I [78, 79]. However, this IFN-1 hypersensitivity was still observed for a catalytically inactive form of USP18, suggesting that at least part of its regulatory effect is ISG15-independent. Instead, it is thought to result from a direct interaction between IFNAR and USP18, which causes displacement of JAK1 and loss of downstream signalling [80].

1.1.3 Antiviral effectors of the IFN-I response

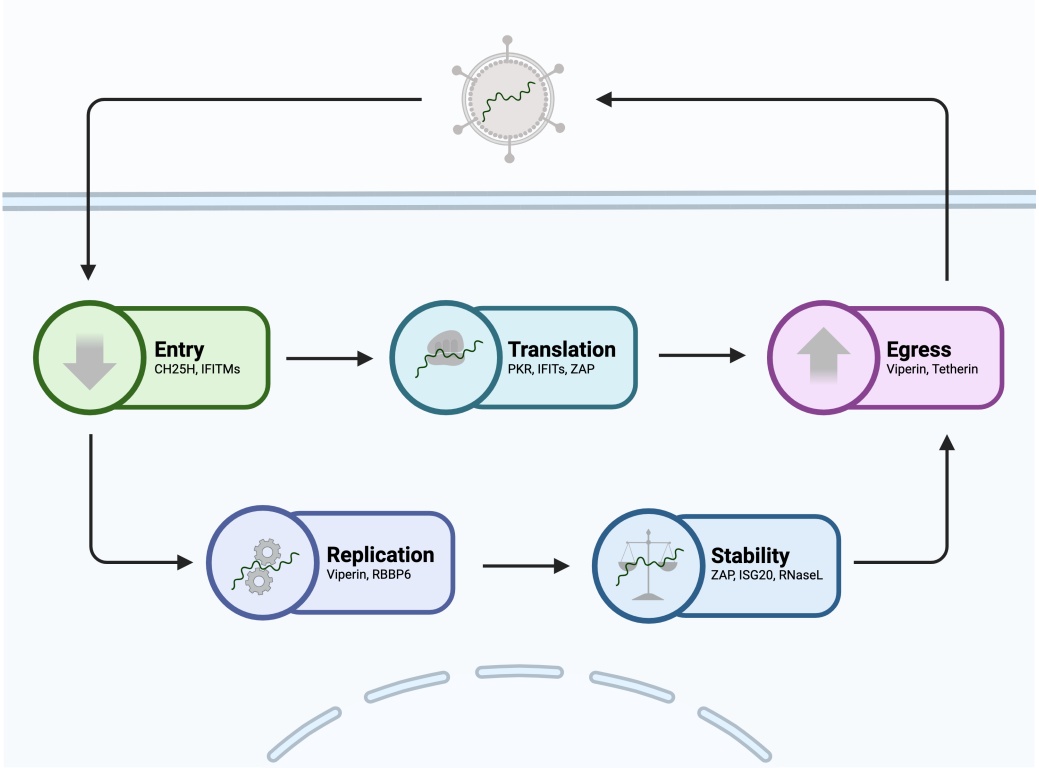


Figure 1.3: **ISG-induced antiviral processes.** Schematic showing the stages of the viral lifecycle targeted by ISGs, including examples of ISGs targeting each stage.

The canonical outcome of the IFN-I signalling cascade is an increase in transcription of hundreds of ISGs. These ISGs can have direct antiviral effects, boost intrinsic IFN-I signalling, contribute to paracrine signalling networks, or negatively regulate the response (as discussed above). ISGs have evolved to target every stage of the viral lifecycle. Indeed, often a single gene will have effects at multiple stages for the same or different viruses. A complete appraisal of antiviral ISGs is beyond the scope of this introduction so I will instead outline illustrative examples of the key steps that are inhibited (summarised in Figure 1.3).

Entry. ISGs that affect viral entry generally do so either through altering the ability of viruses to initiate membrane fusion, or through disrupting the endocytic pathway. The ISG cholesterol-25-hydroxylase (CH25H) is an enzyme responsible for converting cholesterol into 25HC. 25HC plays a central role in regulation of sterol metabolism and can also restrict infection of diverse enveloped viruses [81,82]. Treatment of cells with 25HC directly inhibits viral membrane fusion for a range of viruses, including SARS-CoV-2 [81, 83]. However, there is also evidence that CH25H can inhibit replication at a post-entry step for some viruses. For example, it has been found to restrict HCV through sequestration of its NS5A protein, which is essential for genome replication [84].

The IFITM proteins also restrict viral entry. These proteins were originally identified by Brass and colleagues in a screen of IAV modulating ISGs. Restriction was found to occur very early in infection not only for IAV, but also for WNV and DENV [85]. IFITMs are enriched in late endosomes and lysosomes, leading to the preferential restriction of viruses that require these structures for cell entry [86]. The exact mechanism of inhibition is still subject to debate but a prevailing theory is that IFITMs modulate the physical characteristics of membranes, including membrane curvature [87] and cholesterol content [88].

Translation. Translation of viral proteins is central to the viral lifecycle and relies heavily on host translational machinery. Perhaps because this step is so closely linked with cellular processes, it is also the stage most frequently targeted by antiviral factors [89]. These factors can either act globally to suppress translation, or can specifically modulate the translation of viral RNA.

PKR was one of the first ISGs to be assigned an antiviral function [90]. PKR is induced by IFN-I and activated through binding to regions of dsRNA and other structural features that are frequently found in viral genomes [91]. This results in homodimerisation, relieving an autoinhibitory conformation of PKR and activating its kinase domain [92]. Phosphorylation of eIF2α at serine 51 results in its interaction with eIF2B, an association that inhibits the guanine nucleotide exchange activity of eIF2B. This activity is required for GTP/GDP exchange on eIF2, which is an essential step in translation initiation [93]. Loss of this activity leads to translational arrest and accumulation of cellular and viral RNA in stress granules. However, there are viral RNAs that can evade this repression through employing unconventional translation initiation strategies, such as internal ribosome entry sites (IRESs) [94]. Other ISGs bind directly to viral

RNAs to inhibit their translation. The best characterised of these are the IFIT proteins. These proteins are induced to very high levels by IFN-Is [95, 96]. IFIT1 binds to unmethylated cap structures frequently found in viral RNA and engages with other IFIT proteins to compete with translation initiation factors for cap binding [97–99].

In addition to proteins that block translation of viral proteins, there are also those that alter these proteins co- or post-translationally. ISG15 (mentioned above) is highly induced by IFN-I and is added to proteins co-translationally by HERC5 [68]. HERC5 is also induced by IFN-I and both ISG15 and HERC5 have antiviral activity [100]. ISGylation of viral proteins can restrict infection by modifying viral proteins. For example, ISGylation of the CVB3 protein 2APro hinders its ability to cleave eIF4G and thus its ability to induce host translational shut-off [101].

Another important set of antiviral proteins are the E3 ubiquitin ligases, particularly those from the TRIM family. These enzymes mediate the addition of ubiquitin to proteins, which can either direct the protein for degradation or have downstream signalling effects. Ubiquitination of viral proteins to target them for degradation is a common antiviral mechanism. For example, TRIM22 is an ISG that targets diverse viral proteins for degradation, including EMCV's 3C protease [102] and IAV's nucleoprotein [103].

RNA stability. Regulation of RNA stability is closely linked with regulation of translation; many of the same antiviral factors have been implicated in modulation of both processes. One such ISG is ZAP, which binds to CpG dinucleotides in vRNA [104]. ZAP, with the help of TRIM25, recruits KHNYN which has endonucleolytic activity and thus degrades ZAP's RNA targets [105]. As well as degrading vRNA, ZAP can also inhibit translation through blocking the interaction between eIF4A and eIF4G [106]. In alphaviruses, this was found to be a necessary pre-requisite to ZAP-mediated degradation, indicating how closely linked these two processes are [107].

ISG20 is another antiviral factor that can affect both RNA stability and translation. ISG20 is a 3'-to-5' DEDD exonuclease that inhibits a wide range of viruses [89]. This inhibition is partially driven by ISG20's exonuclease activity. Overexpression of ISG20 reduces levels of vRNA but only if it is catalytically active [108]. How ISG20 specifically recognises vRNA for degradation is an open question. In the context of HBV infection, it is thought to be driven through N6-Methyladenosine (m6A) modifications on the vRNA [109]. However, this cannot fully explain its specificity as many cellular transcripts, including the majority of ISGs, also have

this modification [110]. ISG20 can also inhibit viral translation independently of its exonuclease activity. This can be mediated through ISG20-driven upregulation IFITs [111], or through recognition and recruitment of foreign RNAs to processing (P) bodies [112].

The OAS/RNase L system is another important contributor to vRNA degradation. OAS enzymes are activated through binding to double-stranded vRNA. Once activated, they synthesise 2'-5' oligoadenylates, which in turn activate the endoribonuclease RNase L [90]. While there are three catalytically active OASs, all of which are ISGs, OAS3 is thought to be the primary contributor to RNase L activation, due to having the highest level of RNA-binding activity [113, 114]. RNase L cleaves both cellular and viral RNA. Not only does this lead to a reduction in vRNA abundance but the cleavage products are also recognised by RIG-I, resulting in an amplification of pathogen sensing [115].

Genome replication. Genome replication is most frequently targeted through degradation or sequestration of viral proteins. The E3 ligase, RBBP6, is an ISG that represses EBOV RNA replication through direct interaction with VP30 [116]. RBBP6 binds to VP30 in a cleft that is usually occupied by the viral nucleoprotein (NP), acting as a competitive inhibitor of NP binding. The VP30-NP interaction is critical for EBOV RNA replication. Therefore, overexpression of full-length RBBP6, or the peptide that specifically occupies NP's binding cleft, inhibits replication. Overexpression of RBBP6 also reduces the protein abundance of VP30 in a manner that is dependent on the proteins' direct interaction. This suggests that RBBP6 exerts its antiviral effect through ligase-dependent and ligase-independent activity [116].

An interesting alternative replication inhibition strategy is employed by the pleiotropic antiviral protein, viperin. While initially identified as an inhibitor of viral particle budding [117], it is also a potent RNA replication inihibitor in HCV and DENV infection [118, 119]. Gizzi and colleagues found that viperin catalyses the formation of ddhCTP, a nucleoside analog that is incorporated into vRNA by the RNA-dependent RNA polymerase (RdRp) of flaviviruses. ddhCTP acts as a chain terminator, thus blocking RNA synthesis [120]. While potent against flaviviruses, ddhCTP only weakly inhibits the replication of viruses from other families, suggesting that the RdRps of other viruses might not be susceptible to its incorporation [121].

Egress. There are very few proteins that have been identified as inhibitors of viral egress. This might reflect its less frequent targeting by ISGs or technical challenges in isolating egress from other stages of the viral lifecycle. In addition to viperin (mentioned above), one protein

that has been definitively linked to blocking viral egress is tetherin, which physically 'tethers' viral particles to the cell membrane [122]. Many viruses from diverse families, including HIV-1, EBOV, KSHV, and IAV, have evolved mechanisms to counteract this tethering effect, indicating the importance of this restriction mechanism [2].

A consistent theme across every stage of the IFN-I response is the central role of RNA-binding proteins (RBPs). Firstly, many PRRs are RBPs that recognise unusual features of vRNA. Secondly, the IFN-I response induces changes in the transcriptional and post-transcriptional landscape, which are mediated by regulatory RBPs. Thirdly, a large proportion of ISG-driven antiviral mechanisms are mediated by RBPs that directly engage with vRNA. While there are extensive examples of RBPs' importance in these processes, there has yet to be a comprehensive assessment of their role and regulation in the IFN-I response. Methodological developments over the past 15 years have provided powerful new tools for understanding the scope, complexity and plasticity of the cellular RNA-bound proteome (RBPome). These tools, and the insights they have provided so far, along with key features of RBPs, are discussed in the next section.

1.2 RNA-binding proteins

The central role of RBPs in RNA metabolism has been known for decades. The first of these RBPs to be identified and characterised included heterogeneous ribonucleoproteins (hnRNPs) and poly(A) binding protein (PABP) [123]. These proteins provided a precedent for the importance of RBPs in regulating the mRNA lifecycle and offered insight into how RBPs recognise RNA. They were found to have a modular architecture, being composed of combinations of conserved RNA-binding domains (RBDs). This architecture was thought to represent the dominant means of RNA-protein interactions, and early efforts to identify new RBPs relied on searching for similar domains within the proteome [124]. However, even from the early days of RBP discovery, there were clear examples of proteins that did not follow this pattern of binding. For example, work in the 1990s found that the RBP iron regulatory protein 1 (IRP1) was the same protein as the metabolic enzyme, aconitase 1 (ACO1) [125]. IRP1 binds RNA and carries out its enzymatic function through the same globular domain, making these two processes mutually exclusive. The balance of iron in the cell determines whether IRP1 is able to coordinate an Fe4-S cluster within its enzymatic core. In its presence (when iron levels

are high), the protein functions as an enzyme but when it is lost (when iron levels are low), IRP1 switches to binding iron responsive elements (IREs) within the mRNAs of proteins linked to iron metabolism, regulating their translation and stability [126]. This is a beautiful example of the 'REM' theory, where a metabolite can regulate the transcriptional landscape through the dual functionality of an RNA-binding metabolic enzyme [127]. It also speaks more broadly to the complex regulatory processes that RNA-protein interactions can mediate, beyond RNA processing.

1.2.1 Functional consequences of RNA-protein interactions

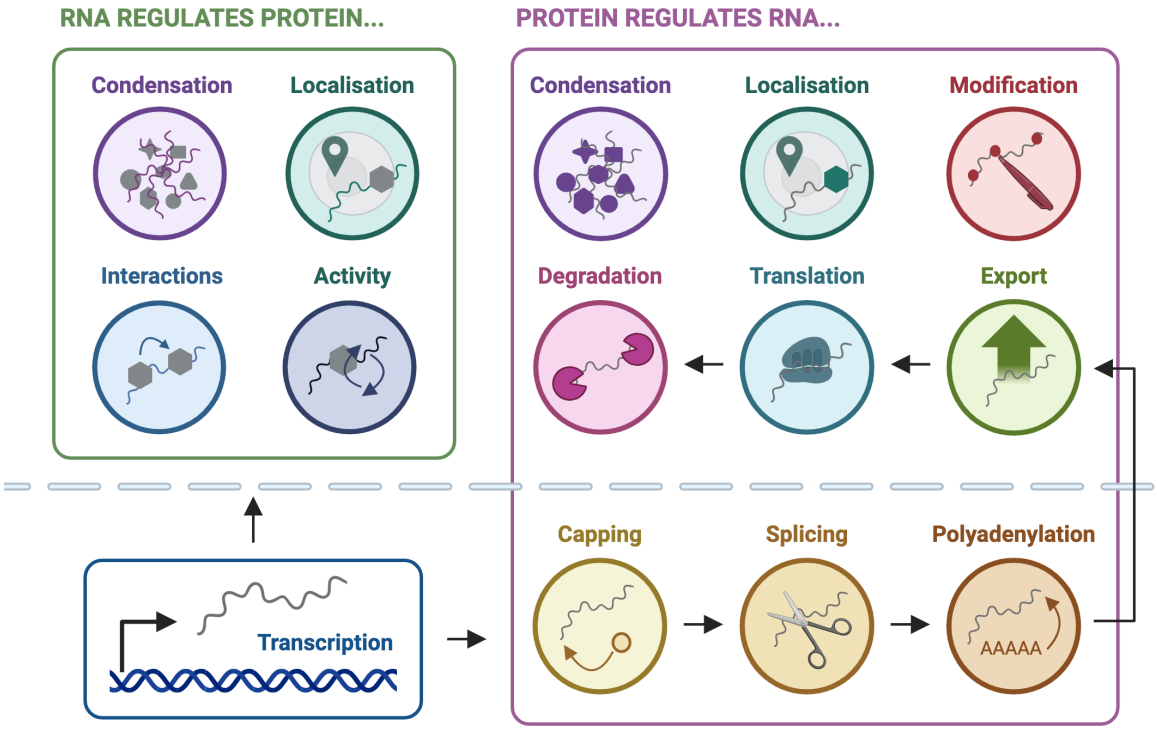


Figure 1.4: **RNA is regulated by and regulates RBPs.** The traditional path of an mRNA upon transcription is indicated by the direction of the arrows. The mRNA is capped, spliced and polyadenylated, before being exported from the nucleus. It can then be translated and will eventually be degraded. Alongside this simplified lifecycle, RNAs can be regulated by RBPs in other ways, including modulation of RNA localisation within the cell, modification, and liquid-liquid phase separation properties (condensation). RNA can also directly 'riboregulate' the activity of proteins, altering their condensation properties, enzymatic activity, localisation and interactions with other proteins.

Protein regulates RNA. RNAs exist in complex with proteins (as ribonucleoproteins, or RNPs) at every stage of their lifecycle. For mRNAs, the various sets of proteins that they

interact with from synthesis to degradation are key in shaping translational outcomes (Figure 1.4). These RBPs include 'core' proteins that invariably associate with mRNAs, and regulatory proteins, which selectively interact with specific transcripts or under certain conditions [128].

As soon as an mRNA emerges from the shadow of RNA polymerase II, it begins to interact with RBPs [128]. The first of these RBPs are those involved in mRNA capping. These mediate the attachment of a 7-methylguanosine 'cap' to the mRNA's 5' end, and drive the 2'-O methylation of the RNA's first and second nucleotides [129]. The cap structure formed in this process acts as an important attachment point for RBPs throughout the mRNA's lifecycle. It is initially bound by the nuclear cap binding complex (CBC), which contributes to RNA stability, splicing regulation, nuclear export, and recruitment of the ribosome for the pioneer round of translation [129]. In the cytoplasm, the CBC is replaced by eIF4E, the cap binding component of the translation initiation factor eIF4F, which promotes subsequent rounds of translation [130]. As well as mediating these important interactions, the cap structure, along with 2'-O methylation, also *prevent* binding of PRRs, such as IFIT1 and RIG-I, which sense unusual 5' ends, allowing the cell to distinguish self from non-self RNA [10].

The majority of genes are transcribed as pre-mRNAs, containing both the coding sequence (exons) and intervening non-coding sequence (introns) that must be removed. These introns are removed in a splicing reaction that is mediated by the spliceosome, a large molecular machine consisting of 5 RNA/protein subunits, known as snRNPs [131]. Additionally, there are tens of regulatory RBPs, including hnRNPs and SR proteins, that engage with the pre-mRNA and determine where the spliceosome acts [132]. Regulation of this process results in 'alternative splicing', in which different combinations of exons give rise to diverse mRNA and protein products from the same gene [131]. Sites of splicing in mRNAs are marked through deposition of the exon-junction complex (EJC) by the spliceosome. EJCs are another set of RBPs that provide important architectural information during post-transcriptional processing, and can promote mRNA export and translation [133].

Once the full mRNA has been transcribed, its 3' end is processed through cleavage and polyadenylation. This gives rise to an off-template poly(A) stretch, which is important in maintaining RNA stability and promoting circularisation during translation [128, 130]. A given mRNA can have multiple sites in its 3' UTR at which cleavage and polyadenylation can take place [134]. 3' UTRs are frequently bound by regulatory RBPs so alternative polyadenylation, which results in loss or gain of 3' UTR sequence, can profoundly alter RNA fates [134]. Various

RBPs have been implicated in the regulation of alternative polyadenylation, including Nova, which was found to modulate polyadenylation of neuronal transcripts to which it was directly bound [135]. Notably, alternative splicing and alternative polyadenylation have been reported to occur in many ISGs and proteins involved in IFN-I signalling [136]. This is a further indication of the importance of RBPs in shaping the IFN-I response.

After the mRNA is transcribed and processed, it must be exported to the cytoplasm to be translated. RBPs, including the EJC and TREX complex, are essential for mRNA export. Once in the cytoplasm, the mRNP is extensively remodelled to promote efficient translation. Important changes include replacement of the nuclear cap binding complex with cytoplasmic eIF4F [130]. Additionally, EJCs are actively removed from the RNA during the first round of translation to allow passage of the ribosome along the coding sequence [133]. While the mRNA coding sequence is 'naked' while it is being translated, factors that regulate translation bind extensively to mRNAs' UTRs. Above, I discussed some of the diverse mechanisms employed by cells to control translation of vRNA. Mechanisms involved in regulation of host mRNA translation are similarly diverse. One translation regulation mechanism involves a protein that we will encounter later in this thesis, GIGYF2. This protein frequently acts through interaction with 4EHP, which binds to the 5' cap in place of eIF4E without forming its essential interaction with eIF4G [137]. GIGYF2 is thought to mediate interactions with other repressive RBPs, restricting translation and/or marking mRNAs for decay [138]. It can also act independently of 4EHP, recruiting the CCR4-NOT complex, which deadenylates mRNAs, targeting them for degradation [139].

mRNA translation and decay are tightly linked processes which are often regulated by the same RBPs. For example, IRP1, which we encountered above, can promote RNA stability when it is bound to an mRNAs 3' UTR, but will block translation when associated with its 5' UTR [126]. Proteins that associate with the 3' UTR are frequently involved in regulating decay, which represents the final stage of an mRNA's lifecycle. RNA turnover is tightly regulated both for quality control purposes and to ensure precise control over protein expression levels. For example, TTP exerts control of the inflammatory response by binding to AU-rich elements in the 3' UTRs of proinflammatory cytokine mRNAs [140]. TTP recruits the CCR4-NOT to deadenylate these mRNAs and target them for degradation, ensuring that cytokines are only produced for as long as they are needed [141]. Regulation of translation and decay can also be established through mRNA sequestration in membraneless organelles, such as

P-bodies and stress granules, which act as 'storage facilities' for translationally repressed mRNAs. While their exact role in RNA metabolism is not yet understood, they appear to have important functions in cellular homeostasis and in response to cellular stress [142]. In summary, RBPs are absolutely central to regulation of RNAs from 'birth' to 'death' and RNA cannot be considered in isolation of the RNP complexes within which it sits.

RNA regulates protein. The canonical role of RNA-protein interactions is in regulating RNA metabolism. However, there is building evidence for the importance of RNA as a direct regulator of protein function through a process known as 'riboregulation' (Figure 1.4). For example, just as proteins can sequester RNAs within membraneless organelles, RNA can also cause recruitment of protein to these structures. NEAT1 is a long noncoding RNA that acts as a scaffold that drives formation of paraspeckles through recruitment of core paraspeckle proteins, including SFPQ, NONO, and PSPC1 [143]. This can indirectly alter the activity of these proteins through modulating their context and protein interaction partners but riboregulation can also directly alter protein activity. For example, activity of the autophagy receptor p62 is modulated through its interaction with the small non-coding, vtRNA1-1. Under steady state conditions, p62 is bound by vtRNA1-1, which prevents its oligomerisation. Upon cell starvation, levels of vtRNA1-1 are reduced, releasing p62, which can then homo-oligomerise and interact with autophagic effectors [144].

Another context that is rich in instances of riboregulation is viral infection. The RNA sensing pathways described above are all instances where RNA binding is directly regulating the activity of the bound PRR. For example, PKR dimerises and is released from its autoinhibitory state as a result of binding to dsRNA [91]. This activates its kinase activity, which leads to autophosphorylation, as well as phosphorylation of its key target, eIF2α. Interestingly, PKR can also be negatively riboregulated by cellular and viral decoy RNAs. The noncoding cellular RNA nc886 binds to PKR in the absence of infection and, similarly to p62, prevents its dimerisation and activation [145]. This is thought to be a mechanism of maintaining PKR expression in steady state cells without inducing translation arrest. High levels of HIV-1's TAR RNA element can also bind PKR and prevent its dimerisation and activation as a means of innate immune evasion [146].

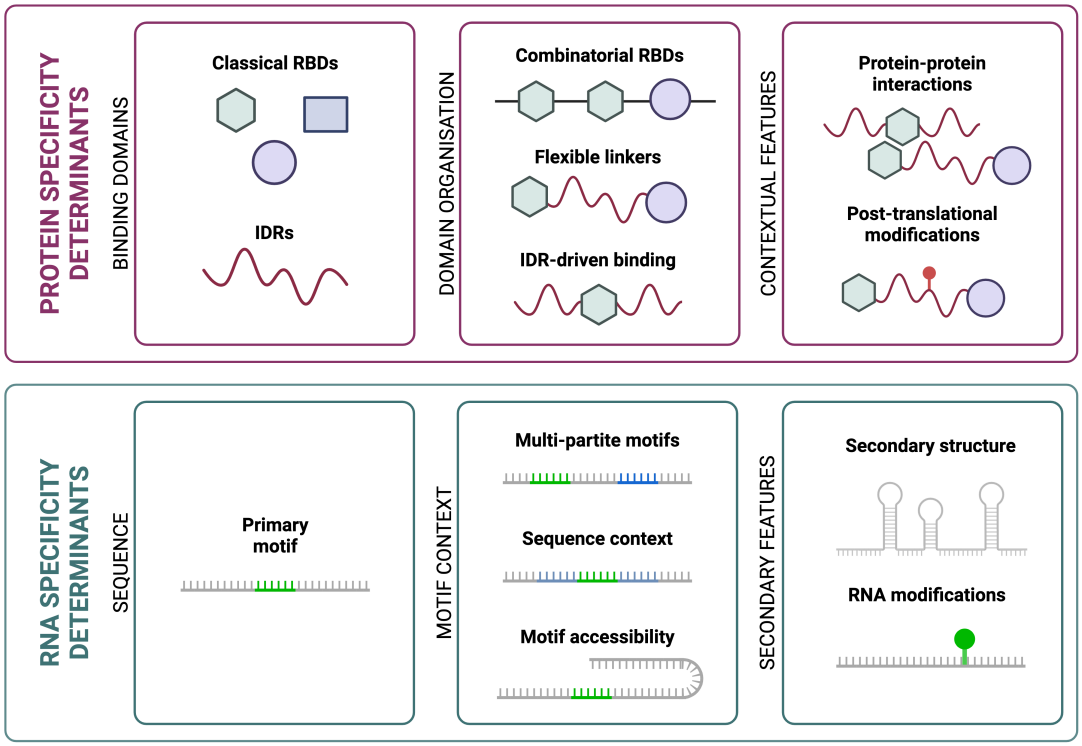


Figure 1.5: **RNA and protein-driven determinants of RBP binding specificity.** A schematic showing features of proteins (in purple) and RNA (in green) that affect their ability to interact with eachother.

1.2.2 Properties of protein-RNA interfaces

RNA binding domains. The initial paradigm of RNA binding was based on the interaction of RNA with a defined set of RNA-binding domains (RBDs) [147]. These domains include: RNA recognition motifs (RRMs), K-homology (KH) domains, Zing finger domains (ZnFs), and double-stranded (ds)RBDs, among others. Domains that bind single-stranded RNA (including RRMs, KH domains and ZnFs) typically contact the bases specifically, mediating binding through a combination of aromatic residue stacking, hydrogen bonding, and electrostatic interactions. By contrast, dsRBDs generally make contact with phosphate backbone through 2' OH groups, binding in a sequence-independent manner [148].

Usually, a single RBD will only recognise a very short stretch of RNA. The most extreme example of this is the Puf domain, which recognises a single nucleotide. However, Puf domains are found in arrays within pumillio (PUM) proteins, allowing for the precise recognition of eight nucleotide RNA sequences [149]. RBPs frequently contain arrays of either the same RBD, or combinations of RBDs to expand binding interfaces and increase their specificity [147]. An example of specificity derived from combinatorial RBPs is the STAR domain, which is

found in several proteins, including the splicing regulators, SF1 and Sam68. STAR consists of a KH domain whose specificity is extended by a QUA2 domain, allowing proteins to recognise extended motifs of up to seven nucleotides [150]. The almost infinite number of RBD combinations and arrangements make this a versatile code for recognising diverse RNA targets [147]. However, these canonical RBDs are still only part of the picture of how RBPs specifically bind RNA.

In addition to structured domains, intrinsically disordered regions (IDRs) are known to play important roles both in linking RBDs and as RNA binding interfaces in their own right. IDRs are characterised by their flexibility and high degree of surface exposure, making them well-suited for engaging with RNA and protein interaction partners [151]. IDR RNA-binding interfaces generally mirror the nucleic acid composition of canonical RBDs, albeit with an enrichment for disorder-promoting amino acids, such as serine, glycine, and proline [152]. There are also specific repeat patterns that occur within IDRs that are particularly well suited for RNA-binding. These include arginine-serine-rich (RS) repeats, arginine-glycine-rich (RG) repeats, basic patches, and [G/S]Y[G/S] repeats [153]. These different low complexity sequences can play specialised roles in RNA-binding regulation. For example, RS repeats are found in SR proteins, alongside RRMs. Within SR proteins, RS repeats directly interact with RNA at splice sites and modulate the binding activity of RRMs through their phosphorylation state [154]. [G/S]Y[G/S]-rich regions regulate, in a more general way, the ability of proteins to liquid-liquid phase separate (LLPS). Aromatic residues, which sit at the protein's surface due to their glycine/ serine-rich surroundings, will readily form interactions with amino acids or nucleotides, causing protein/RNA aggregates to form [153].

IDRs can bind RNA independently of other RBDs but also frequently act in concert with classical and non-classical RBDs, as outlined for SR proteins above. The contribution of IDRs to RBP specificity is generally thought to be minimal, although there are instances of specific RBP binding. For example, FMRP contains an RGG box element which specifically binds to G quadruplexes. However, the overall specificity of FMRP is still thought to be driven by its KH domain, with the RGG box's primary role being to increase overall affinity, a common feature of RNA-binding IDRs [155]. Another role that these regions can play is in defining linker properties between RBDs. For example, in PTBP1, a disordered linker between its third and fourth RRMs is essential for establishing an antiparallel conformation, which is in turn essential for PTBP1's function [156]. Therefore, IDRs are a powerful tool for fine-tuning RBPs' RNA-binding activity

and increasing the literal and figurative flexibility of the classical RBD code [151]. Interestingly, a comparison of RBP orthologues between yeast and humans identified expansion within IDR repeat regions in humans. It has been suggested that these regions might have evolved to exert more specific regulation as organism complexity increased, speaking to their central role in RNA-binding [157].

RNA determinants of RBP binding. Understanding how the almost infinite permutations of RBD conformations translate into RNA binding specificity has been a subject of intensive research. Classical RBDs often recognise short RNA sequence motifs defined by conserved amino acids within their structure [147]. Methods for identifying these motifs *in vitro* include SELEX [158], RNACompete [159], and RNA-Bind-and-Seq (RBNS) [160]. These methods rely on assessing the relative affinity of purified RBDs for a pool of RNA oligos. Such approaches have been used to identify the motifs favoured by many RBPs [161]. However, analyses of *in vivo* binding sites have indicated that these short motifs are rarely, if ever, sufficient to entirely account for an RBP's binding properties. Generally only a fraction (15-40%) of putative binding sites will actually be occupied *in vivo* [160]. Additionally, two proteins with the exact same linear motif can have unique sets of RNA targets within cells. This has been observed for HNRNPC and TIA-1, which both bind poly(U) tracts [162].

In vivo binding is commonly assessed using CLIP-based methods. These approaches rely on UV crosslinking of RNA and associated protein, followed by purification of protein-RNA complexes for the RBP of interest, and sequencing of associated RNA [163]. CLIP methods have been invaluable in establishing complex patterns of RBP binding and are highly complementary to *in vitro* approaches. The *in vitro* RBNS method has been effectively employed alongside eCLIP to fine-tune our understanding of RBP specificity. Taliaferro and colleagues used RBNS to assess the intron binding preferences of two splicing regulators: RBFOX2 and MBNL1. In addition to motif *occurrence*, they found that motif *accessibility* was a key determinant of binding; sequences that met both occurrence and accessibility criteria were much more likely to be identified as binding sites by eCLIP [160]. A further large-scale RBNS study of 78 RBPs found that individual motifs, bipartite motifs, motif sequence context, and secondary structure all contributed to defining the specificity of individual RBPs (Figure 1.5) [162]. Understanding how inherent specificity determinants relate to patterns of *in vivo* binding is key to making predictions about where RBPs are likely to bind in novel contexts (e.g.

different cell types) or in the presence of new potential targets (e.g. viral RNA). A recent focus of research has been on generating binding prediction algorithms based on available CLIP datasets, something that is starting to become more feasible with the current rapid advances in machine learning [164].

The proteins profiled using the methods outlined above are generally those whose specificity can be isolated to relatively contained linear sequences or defined secondary structure features. However, there are other RBPs for which this is not the case. For example, there are an important set of proteins that bind to structural features outside of the main body of the RNA. These include general cap binding proteins, such as EIF4E and NCBP2, as well as context-specific 5' end binding proteins like the PRRs, IFIT1 and RIG-I. As discussed above, IFIT1 binds to cap structures that lack 2' O methylation on the first nucleotide and RIG-I interacts with the 5' PPP or PP ends of dsRNA [14]. Another group of proteins recognise modifications made to the RNA. The best characterised of these are YTH proteins, which bind to - or 'read' - m6A modifications found extensively in cellular RNA [165].

In addition to containing small structural features, RNA can also form extensive 3D conformations that bind proteins through 'shape complementarity'. This is commonly observed in large protein-RNA enzymatic complexes, such as the spliceosome and ribosome, where proteins and RNA have co-evolved to perfectly fit one another's 3D structure in a manner that is reminiscent of protein-protein interactions [166]. Such co-evolution has also taken place between vRNA and host proteins. IRESs present in the genomes of many viruses, including hepatitis C virus (HCV) and poliovirus, have evolved RNA structures that mimic the binding interfaces of translation initiation factors to recruit the translation initiation machinery, even under conditions of translational arrest [94,167]. Still, other proteins have no intrinsic specificity but bind in a highly specific way as a result of PPIs. Such is the case for exon junction complex (EJC) proteins, which are deposited precisely 20nts upstream of splice sites, regardless of the RNA sequence, through a specific interaction with the spliceosome [131]. These examples reflect the diversity and complexity of RNA-protein interactions. Sequence and secondary structure features of RNA can be helpful for understanding the binding preferences of many RBPs, but there is also a large portion for which they are insufficient.

1.2.3 An expanding universe of RNA-protein interaction dynamics

Methods for identifying novel RBPs. The first high-throughput methods to discover RBPs comprehensively involved assessing the interaction of proteins with RNA probes *in vitro*, using either proteoarrays or mass spectrometry [168–170]. These approaches, applied to yeast, identified hundreds of novel RBPs, including some without clear links to RNA biology, such as metabolic enzymes [169, 170]. While an important first step in broadening the scope of RBPs, *in vitro* methods assess interactions outside of the protein's cellular context so do not account for protein-protein interactions or local factors that might contribute to binding. Additionally, these methods will detect indirect interactions so proteins identified cannot be definitively classed as RBPs without further validation.

An important step forward in RBP identification came with the advent of RNA interactome capture (RIC) [171, 172]. In RIC, proteins are covalently crosslinked to RNA in live cells by UV irradiation. Bases absorb at a specific wavelength, leading to the formation of free radicals that attack amino acids in close proximity (virtually zero distances; <2 Å) generating covalent bonds between RNA and proteins [152, 173]. The appropriate wavelength of UV light depends on the nature of the bases. Castello and colleagues performed RIC either on RNA containing all natural bases or RNA containing the nucleoside analog, 4-thiouridine (4SU), while Baltz and colleagues used exclusively 4SU-labelled RNA. Natural bases absorb optimally at 260nm, while 4SU is activated best at 365nm, allowing for selective crosslinking of 4SU-containing RNA. Oligo(dT) magnetic beads are then used to capture polyadenylated (poly(A)) RNA and associated proteins. Denaturing washes then remove any non-covalent interactions before proteins are released for analysis by LC-MS/MS [174].

RIC dramatically expanded the breadth of the known RBPome seemingly overnight. Castello and colleagues classified 860 proteins as RBPs in HeLa cells and Baltz and colleagues identified 791 putative RBPs in HEK293 cells. Strikingly, 543 of these proteins were shared between the two datasets. Since these initial studies, RIC has been applied to a wide range of species, including mice, fish [175], flies [176], worms [177], plants [178], yeast [179], and parasites [180]. Expanding the range of contexts studied has expanded the scope of the known RBPome. Additionally, comparative analysis of RBPomes across widely varying species have indicated important commonalities. Such comparisons have identified hundreds of RBPs that are conserved across eukaryotes [157].

While extremely powerful, RIC cannot identify interactions with RNA lacking a poly(A)

tail. This limits the study of proteins interacting with non-poly(A) non-coding RNAs and of species without poly(A) RNA, including bacteria and archaea. Several approaches have been developed to facilitate bulk RNP purification, including RNA labelling paired with click chemistry [181, 182], solid phase purification [183, 184], and organic phase separation-based methods [185–187]. While these methods allow for identification of interaction partners of the total RNA pool, coverage of poly(A) RNA interactors is much lower than in RIC as poly(A) RNA only represents around 10% of RNA in the cell. Therefore, these methods are highly complementary, with RIC offering in depth data of poly(A) RNA interactions, and bulk capture providing a shallower overview of interactions for all RNA types.

Uncovering new binding modalities in non-canonical RBPs. With the recent expansion in the number of proteins classified as RBPs, there has been an equivalent expansion in our understanding of the diverse ways that RBPs can interact RNA. In initial RIC studies, only around half of the RBPs identified contained classical RBDs, and one third contained no known RBDs, raising questions as to how these new RBPs engage with RNA [171, 172]. Methods, including RBDmap and RBS-ID, have been developed to experimentally define RNA-binding interfaces on a proteome-wide scale. RBDmap builds on RIC but employs two successive rounds of oligo(dT) capture, interspersed with protease digestion steps to specifically identify the tryptic peptides that are crosslinked to RNA [152]. RBS-ID uses a combination of crosslink enrichment, extensive digestion of crosslinked RNA fragments, and analysis of peptide mass shifts within MS spectra [188]. The high sensitivity of RBDmap, combined with the single nucleotide resolution of RBS-ID make them highly complementary methods and both have provided insight into the properties of non-canonical RBPs.

A large proportion of the crosslinked peptides identified by these methods were within known RBDs, speaking to dataset quality, but there were also many that overlapped with domains not previously known to bind RNA. Around half of RBSs identified by RBDmap mapped to IDRs, with 170 RBPs exhibiting binding exclusively through these regions [152]. This is in keeping with the known role of IDRs in RNA-binding and their known enrichment within RBPs [171]. Binding sites were also identified in globular domains with enzymatic cores, and domains that were thought to exclusively mediate protein-protein interactions, such as the WD40 domain, among others. Reassuringly, the chemical properties of these unconventional binding sites mirror the key features of classical RBDs; they were rich in positively charged

and aromatic residues, and generally depleted of large or negatively charged residues [152]. In summary, experimental approaches have been instrumental in expanding the scope of mechanisms that can govern RNA-protein interactions. Understanding where proteins bind RNA can also teach us about the roles that RNA binding is playing in regulating RNA and protein function.

The cellular RBPome is dynamic. An important paradigm that has been repeatedly strengthened by RNA-binding studies in a wide range of contexts is that the RBPome is highly dynamic. Sysoev and colleagues were the first to quantify changes in the RBPome in their study of the maternal-to-zygotic transition in Drosophila. They identified more than one hundred RBPs that exhibited quantifiable differences in binding between the two states. Roughly three quarters of these changes in RNA binding reflected equivalent changes in protein abundance. Interestingly, the other quarter exhibited changes in association with RNA that were independent of protein level, indicating that their RNA-binding activity is being regulated post-translationally. The authors of this study classified these proteins as 'dynamic binders' [176]. Many subsequent studies have identified similar patterns of RBP behaviour, albeit with a generally higher frequency of dynamic binders. Contexts studied include: stimulation of macrophages with LPS [189], treatment of cells with stress-inducing chemicals [185, 186], and virus infection [50, 190]. In each of these states the RBPome was found to be remodelled, often in ways that were specific to the stimulus being studied. For example, in infected cells, many of the proteins with altered RNA-binding activity were linked either to the innate antiviral response or to viral replication [50]. Garcia-Moreno and colleagues analysed the relative contribution of different factors to the changing RBPome upon SINV infection. They found that, while altered protein abundance contributed minimally, altered RNA abundance was a key driver of the changes observed. This is perhaps unsurprising, given that SINV infection induces massive degradation of cellular transcripts, along with extensive replication of its own genome. This results in a state in which 70% of polyadenylated RNA in the cell is viral. Therefore, the binding partners of this RNA will naturally come to dominate the RBPome [50]. It is likely that the balance of factors contributing to RBPome plasticity are determined by the cellular perturbation being studied and that other cellular states in which the transcriptome is less substantially altered would not be so heavily influenced by RNA abundance.

1.2.4 RBPs in viral infection

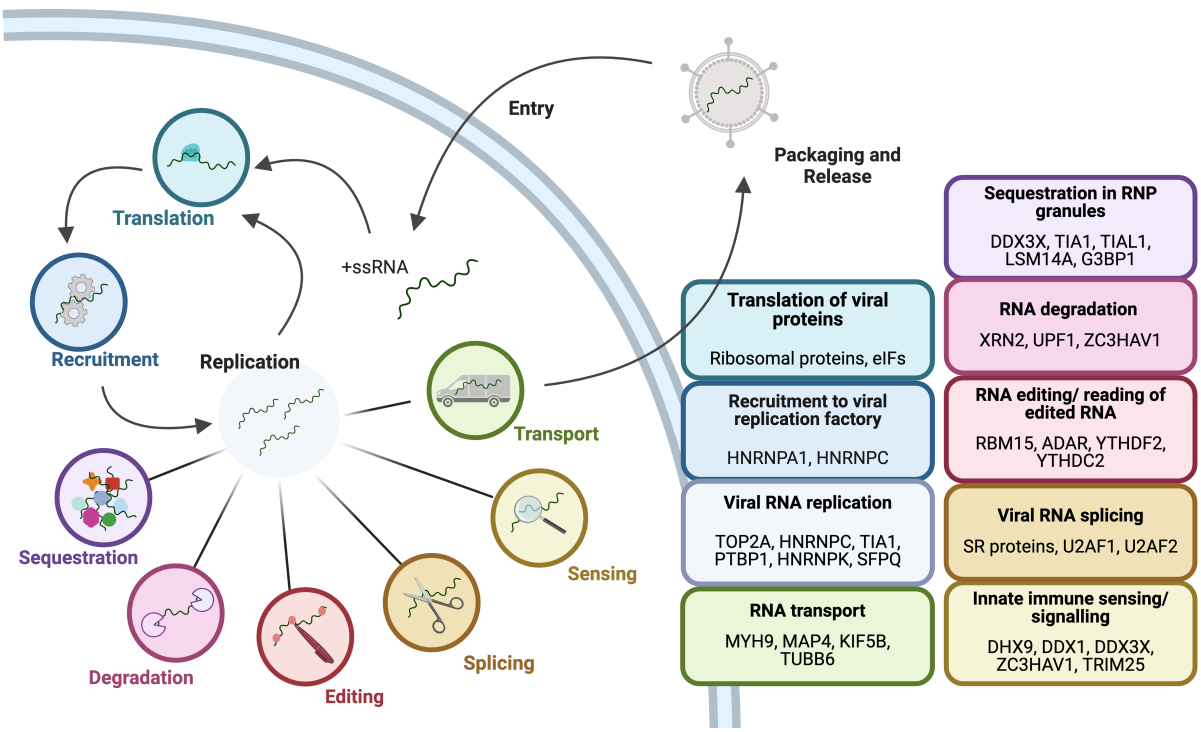


Figure 1.6: **RBPs are involved in every stage of the viral lifecycle.** A schematic showing the stages of the viral lifecycle altered by members of the vRNA interactome. These include core features of the viral lifecycle (Translation, Recruitment to viral replication factory, Replication, RNA transport, and RNA packaging and release) and secondary features that can promote or inhibit infection. Examples of core interactome members affecting each stage are outlined in the boxes on the right.

vRNA plays an active and central role in the infection cycles of all viruses but are particularly critical in RNA virus infection, where they must operate both as mRNA and viral genome. As with any cellular RNA, vRNA is continually packaged into ribonucleoprotein (vRNP) complexes. Components of these vRNPs can be the antiviral sensor and effector proteins discussed above, but can also be host factors that have been hijacked by the virus to promote its replication [191]. Due to the highly restricted coding capacity of their genomes, RNA viruses rely heavily on host RBPs for almost every stage of their replication cycles, including vRNA recruitment to replication factories, replication, translation, processing and packaging (outlined in Figure 1.6). While there are well-characterised examples of RBPs that are involved in each of these processes, the full complement of RBPs that contribute to viral replication, as well as the restriction of it, has yet to be elucidated.

Until recently, identifying cellular RBPs binding to vRNA relied on characterisation

of individual candidates, selected according to their hypothesised mechanisms of action. However, over the past decade, methods have been developed that allow unbiased characterisation of the 'vRNA interactome'; that is, the full complement of cellular proteins interacting with vRNA during infection. Given vRNA's central role in viral replication and the antiviral response, I hypothesised that vRNA interactome datasets would represent a rich resource in identifying new regulators of viral infection. To determine if this was the case, I compiled 21 vRNA interactome datasets, spanning 11 distinct viruses from 6 viral families. A summary of the vRNA interactome methods and of the results of my meta-analysis have been published as a review [192]. As part of this meta-analysis, I defined a 'core vRNA interactome', consisting of proteins present in at least one dataset from each of the three best-represented viral families: *Flaviviridae*, *Coronaviridae*, and *Togaviridae*. I predicted that conserved interaction partners might speak to pan-viral regulatory mechanisms and, indeed, found many proteins with known roles in infection. These included antiviral proteins like ZAP and TRIM25, as well as those that promote viral replication like components of the translation machinery and some hnRNPs. Core interactome members that are linked to various pro- and antiviral processes are highlighted in Figure 1.6. Around two thirds of these proteins have been linked to viral infection in the literature. However, this still leaves close to seventy conserved vRNA interactors without any known links to infection. This speaks to how much we still don't know about how RBPs shape viral infection outcomes and how much there is to learn from RNA interactome studies.

1.3 Aims

The IFN-I response is a critical component of antiviral innate immunity. While it has been primarily studied as a transcriptionally driven response, there is evidence to suggest that IFN-I can also drive changes in the post-transcriptional and post-translational landscapes of the cell. Given the central role of RBPs both in the viral lifecycle, and in the antiviral innate immune response, I hypothesised that RNA-binding activity might represent a hitherto unexplored layer of IFN-I driven regulation. The overarching aim of this thesis was to determine whether RNA-binding activity is regulated by the IFN-I response and, if so, to establish the consequences of this regulation for infection. I sought to address this through addressing three key aims:

1. Use RIC to identify IFN-regulated RBPs (IR-RBPs) and determine whether these IR-RBPs are playing a role in infection.
2. Assess whether phosphorylation is playing a role in differential RNA-binding activity of IR-RBPs.
3. Establish the infection-specific RNA and protein interactomes of a prototypical antiviral IR-RBP

2 RNA-binding activity as a new regulatory layer in the type I IFN response

2.1 Introduction

Since the discovery of the IFN-I response in the 1950s, identifying the antiviral factors that it regulates has been a subject of intensive study. The primary focus of such studies has been on defining the transcriptional changes induced by IFN-I. This was initially made possible through microarrays, which led to the classification of hundreds of genes as ISGs [193–196]. RNA sequencing (RNAseq) has since taken this number into the thousands, with datasets spanning a wide range of contexts [197]. These have provided insight into aspects of the IFN-I response that are conserved across evolution, cell types, IFN subtypes, and disease states, as well as aspects that are context-dependent [55, 198, 199]. RNAseq has also allowed for the identification of regulation beyond transcriptional abundance. It is now clear that alternative splicing (AS), alternative start site usage, and alternative polydenylation are all mechanisms employed to shape the transcriptional landscape in response to IFN-I [72, 200]. Additionally, regulation of non-coding RNA transcription has emerged as an important aspect of the IFN-I response. IFN-induced long non-coding (lnc)RNAs and micro (mi)RNAs can have direct antiviral effects and/or indirect regulatory roles in the IFN-I response [72].

These results have shed light on the complexity of IFN-I-driven regulation and hint at additional layers of regulation beyond ISG transcription. Several studies have sought to explore these additional layers by combining transcriptomics with other 'omic' approaches. These can probe IFN-I's control of the stages during (ATAC- seq, Chip-seq) or after (proteomics) transcription [199–206]. Although there are several IFN-I proteomic datasets available, many of these suffer from a lack of depth. The absolute abundance of antiviral proteins, even after IFN-I stimulation, is generally low. This means that proteomic studies that only reach depths in the low thousands miss qualitative and quantitative data for many ISGs, limiting comparison with transcriptomic datasets. A recent proteomic analysis of IFN-I treated cells was performed by Castello lab member Dr. Honglin Chen. This dataset reached a depth of more than ten thousand proteins, allowing the quantification of many low-abundant ISGs. Surprisingly, despite more than 400 hundred mRNAs exhibiting increased expression 20 hours post-treatment, fewer than one hundred proteins exhibited an equivalent change in protein abundance (Chen et al, In Preparation). As described in Chapter 1, IFN-I is known to

directly regulate protein translation and post-translational activity through PTMs and PPIs [70]. Together, these results indicate that we should move away from an exclusive focus on ISGs and focus in parallel on IFN-regulated proteins, which represent an overlapping but non-identical set of factors that contribute to antiviral immunity.

Despite this expansion in our understanding of IFN-I-driven regulatory mechanisms in recent years, the most prevalent approach for identifying IFN-regulated antiviral proteins is still to overexpress [207–212] or deplete (siRNA/CRISPR-cas9) panels of ISGs [213–218]. These studies have been key in identifying new antiviral factors and laying the foundations for unveiling their molecular mechanisms [219–221]. Their expansion to new cellular systems and viral species will likely continue to yield meaningful results. However, they are limited by their exclusive focus on genes regulated by IFN-I at the transcriptional level. Interestingly, in two genome-wide knockdown screens assessing antiviral activity of proteins after IFN-I treatment, a large proportion of the genes modulating infection were not ISGs [217, 222]. This implies that critical IFN-regulated antiviral factors are being missed by focusing exclusively on ISGs. Notably, in both of these studies, the most enriched group of proteins, besides classical ISGs, are RBPs, speaking to the potential of RBPs to operate as non-ISG antiviral proteins in the IFN-I response [217, 222].

In recent years, RNA-bound proteomes have become a fruitful place to identify antiviral factors. These studies have focused on either affinity purification using PAMPs within viral RNA as baits [223] or assessing RNA-binding to viral or cellular RNA in the context of infection [50,190,192]. For example, when changes to the RBPome were assessed upon SINV infection, GEMIN5 was identified among the RBPs whose RNA-binding activity was upregulated. The authors found that GEMIN5 binds to the 5' cap of viral RNAs and represses infection [50]. However, these studies have centered around changes induced by virus infection or viral RNA. I hypothesised that the IFN-I response might be regulating RNA-binding activity independently of infection, switching the intracellular environment from permissive to hostile. In this chapter, I explore the use of RIC to identify IFN-regulated RBPs in a systematic and comprehensive manner.

2.2 Results

2.2.1 The RNA-bound proteome is regulated by IFN-I treatment

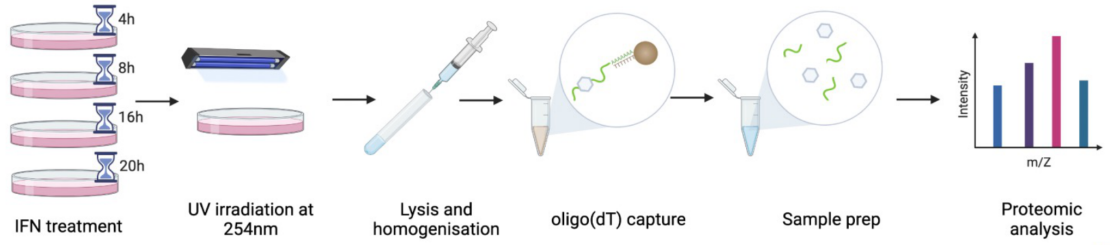


Figure 2.1: **Schematic of IFN RIC protocol.**

RNA-binding proteins are central to the antiviral innate immune response. Therefore, I hypothesized that the IFN-I response might alter the RNA-binding capacity of antiviral RBPs to promote the antiviral state of the cell. To test this hypothesis, I used RIC, which captures proteins in close contact with poly(A) RNA [174, 224]. In brief, cells are irradiated with 150 mJ/cm^2 UV light at a wavelength of 254nm. Cells are then lysed under denaturing conditions and homogenised by passing the lysate through a 27G needle. This is important to shear DNA and reduce sample viscosity. Poly(A) RNA is then isolated using oligo(dT)$_{25}$ magnetic beads. The beads are washed under denaturing conditions in both high and low salt buffers to ensure removal of non-covalently associated protein and RNA, respectively. RNA and crosslinked proteins are eluted from the beads by heating to 55°C and eluates are digested using a combination of RNase A and RNase T1 to release proteins.

In designing my IFN RIC experiment, there were several parameters that needed to be considered, including cell line, IFN-I subtype, IFN-I concentration, and treatment times. For all of these parameters, I decided to match those used by Dr Chen for his whole cell proteome and transcriptome datasets so that I would have a direct point of comparison for my RBPome dataset. In his experiment, HEK293 cells were treated with 200U IFNα2a for either 4 or 20 hours before harvesting. HEK293 cells were selected for their ability to induce a robust and comprehensive antiviral response upon IFNα treatment (Chen et al, In preparation) and SINV infection [50]. IFNα2a is often considered to by the 'archetypal' IFNα subtype, due to its early discovery and high specific activity [225–227]. While IFNβ induces a more robust response than any of the IFNα subtypes, it also gives rise to antiproliferative effects, which were not of interest for this study and might have increased cell toxicity [52]. In experiments performed by Dr Chen, 200U was found to be the concentration at which levels of STAT phosphorylation

started to plateau (Chen et al, In preparation). Therefore, it was selected to maximise the response, while minimising side effects. Early (4h) and late (20h) timepoints were selected to assess the rapid and delayed effects of the IFN-I response. I chose to use both of these timepoints and add two intermediate timepoints – 8h and 16h – to improve kinetic resolution over the course of the IFN-I response. Mock-treated and non-UV irradiated samples were used as controls, harvested alongside the first and last time points [174, 224]. I harvested samples using the above conditions and performed RIC for four replicates. Subsequent to successful quality control checks by Western blotting and silver stain, these samples were processed for analysis by liquid chromatography tandem mass spectrometry (LC-MS/MS).

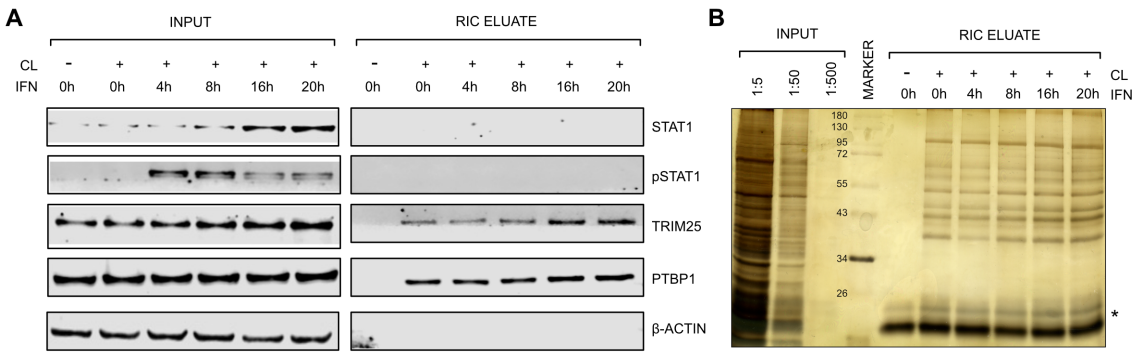


Figure 2.2: **Quality control analysis of the RIC experiment applied to IFNα2 treated cells.** HEK293 cells were treated with IFNα2 for either 0h, 4h, 8h, 16h, or 20h, before being UV crosslinked and harvested. Polyadenylated RNA and associated proteins were captured using oligo(dT) magnetic beads, washed, and eluted at 55°C. (**A**) Western blotting was performed on input and RIC eluate samples against STAT1, pSTAT1, TRIM25, PTBP1, and β-Actin. (**B**) Silver staining was performed against RIC eluate samples. The mock input sample was also loaded in serial dilution to allow for comparison of band patterns. RNase bands are marked with a *.

To assess if the IFN-I treatment was successful for IFN RIC samples, I checked total STAT1 and phosphorylated STAT1 (pSTAT1) levels in RIC inputs via Western blotting (Figure 2.2A). STAT1 is accumulated and phosphorylated upon IFNα treatment [54]. Accordingly, I observed a gradual increase in total STAT1 throughout the infection, while pSTAT1 levels were highest at 4 hpt and 8 hpt. Levels of the ISG TRIM25 in the whole cell lysate also increased throughout the treatment, while levels of β-Actin and the non-ISG, PTBP1, were unaltered. These results confirmed that IFN-I treatment successfully recapitulated the pattern of induction previously observed by Dr Chen.

RIC should selectively capture RBPs in a UV irradiation-dependent manner. To assess whether my RIC experiments were successful, I performed SDS-PAGE gel electrophoresis,

combined with silver staining on RIC eluates and mock input sample (Figure 2.2B). The banding pattern of UV-irradiated RIC eluates was clearly different from the input, indicating enrichment for a specific set of proteins. In contrast, the protein pattern of UV-irradiated RIC eluates across different conditions were consistent. This is what we would expect as the most abundant housekeeping RBPs remain unaltered by most biological conditions tested, including virus infection [50]. Additionally, the non-irradiated control lane was devoid of proteins, except for the RNase band at the bottom of the gel, implying that isolation of proteins is indeed UV-dependent. Similar results were obtained for all four biological replicates, supporting the robustness of the method.

To further confirm that RIC efficiently and specifically isolated known RBPs, I performed Western blotting analysis of RIC eluates (Figure 2.2A). RBPs such as TRIM25 and PTBP1 were strongly enriched in eluates in a UV-dependent manner, while proteins lacking RNA-binding activity, such as STAT1 and β-Actin were efficiently depleted. This is consistent with the specific enrichment of cellular RBPs that has previously been described [171]. Of note, changes in levels of RBPs in RIC eluates over the course of the IFN-I response can be observed by Western blotting. TRIM25 abundance increased over time in eluates to an equivalent or even greater extent than in inputs. Surprisingly, PTBP1 showed a subtle but progressive increase in eluates, despite not displaying a detectable changes in the input. This provided a first indication that IFN-I-induced changes in RNA-binding activity can occur independently of protein abundance in some instances.

Eluates were analysed by LC-MS/MS to identify RBPs associated with poly(A) RNA under each experimental condition. Peptides and proteins were identified and quantified in MaxQuant [228], with 1065 proteins identified in at least one sample. The distribution of protein intensities across crosslinked samples was consistent, while non-crosslinked sample intensities were low in comparison (Figure 2.3A), in line the results of silver staining (Figure 2.2B). Principal component analysis (PCA) also indicates that non-crosslinked samples are substantially different from all crosslinked samples (Figure 2.3B), further demonstrating the UV-dependence of protein isolation. When non-crosslinked eluates are excluded from the PCA, more subtle differences between UV-irradiated samples were observable (Figure 2.3C). Two distinct groups could be distinguished, one consisting of mock (0h), 4h and 8h, and the other consisting of 16h and 20h. This suggests that most observable changes in RNA-binding activity, relative to mock, are occurring after 8 hours.

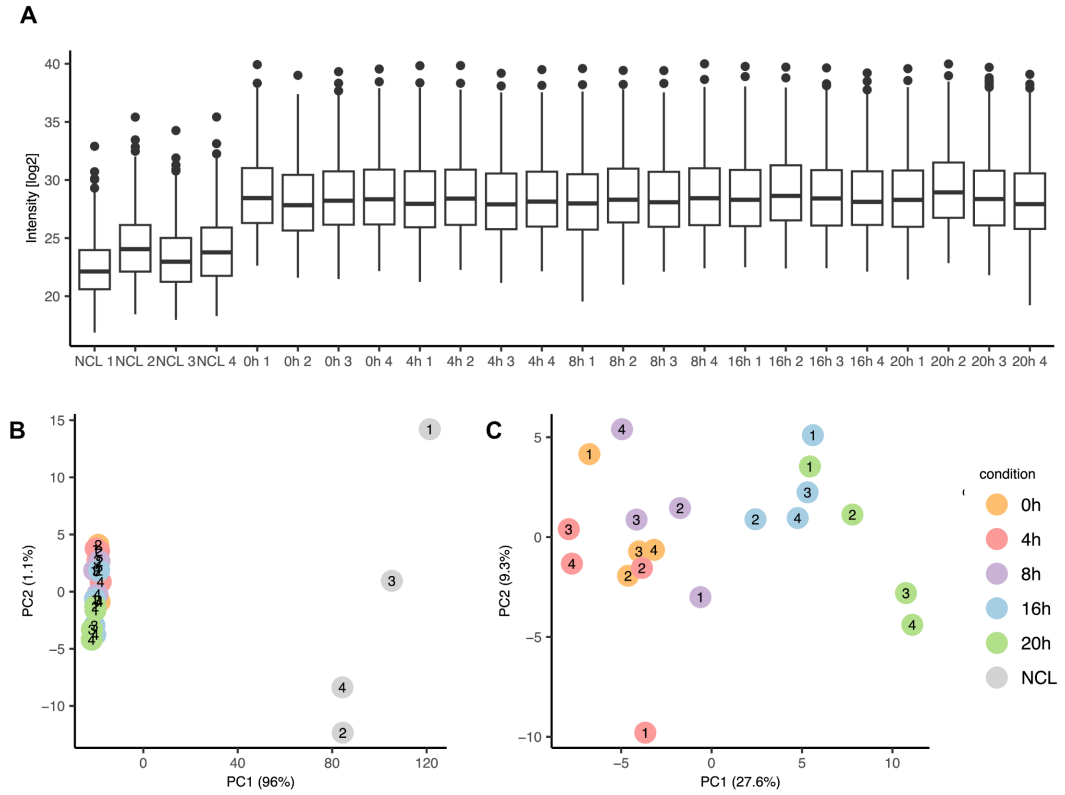


Figure 2.3: **IFN RIC samples were analysed via LC-MS/MS.** (**A**) Boxplots showing distributions of log2 intensities across all samples. (**B+C**) Principal component analysis was performed for all samples (**B**) or for only crosslinked samples (**C**). The first two principal components were plotted as a 2D scatter plot, with percentage of variance explained by each component given in brackets in the axes titles. Condition is represented by the colour of each point and replicate is indicated by the number inside each point (0h = mock, 4h-20h = hours post IFN-I treatment, NCL = non-crosslinked).

To identify which proteins had significantly altered RNA-binding activity after IFN-I treatment, I compared each IFN-I treated condition to mock using the limma package in R [229]. This analysis revealed a small group of proteins with altered RNA-binding activity at 4 and 8 hpt. The number of significantly altered proteins was increased at 16hpt (Figure 2.4C) and was most pronounced at 20hpt, with 272 proteins exhibiting differential RNA-binding activity (10% FDR) (Figure 2.4A-D). A complete list of IFN-I-regulated RBPs (IR-RBPs) can be found in the Appendix. Among the proteins with the strongest increase in RNA binding at 20h were well-known antiviral proteins, such as OAS3, HELZ2, IFIT1, and PKR [72]. These results indicate that some antiviral RBPs that are conventionally thought to recognise patterns specific to foreign RNA can engage with cellular RNA upon IFN-I stimulation. Other known IFN-regulated RBPs were enriched at 10% FDR, including TRIM25 [230], TRIM56 [231], and

IFI16 [45]. The presence of *bona fide* antiviral RBPs within the 10% FDR group underscores the importance of considering this set of proteins.

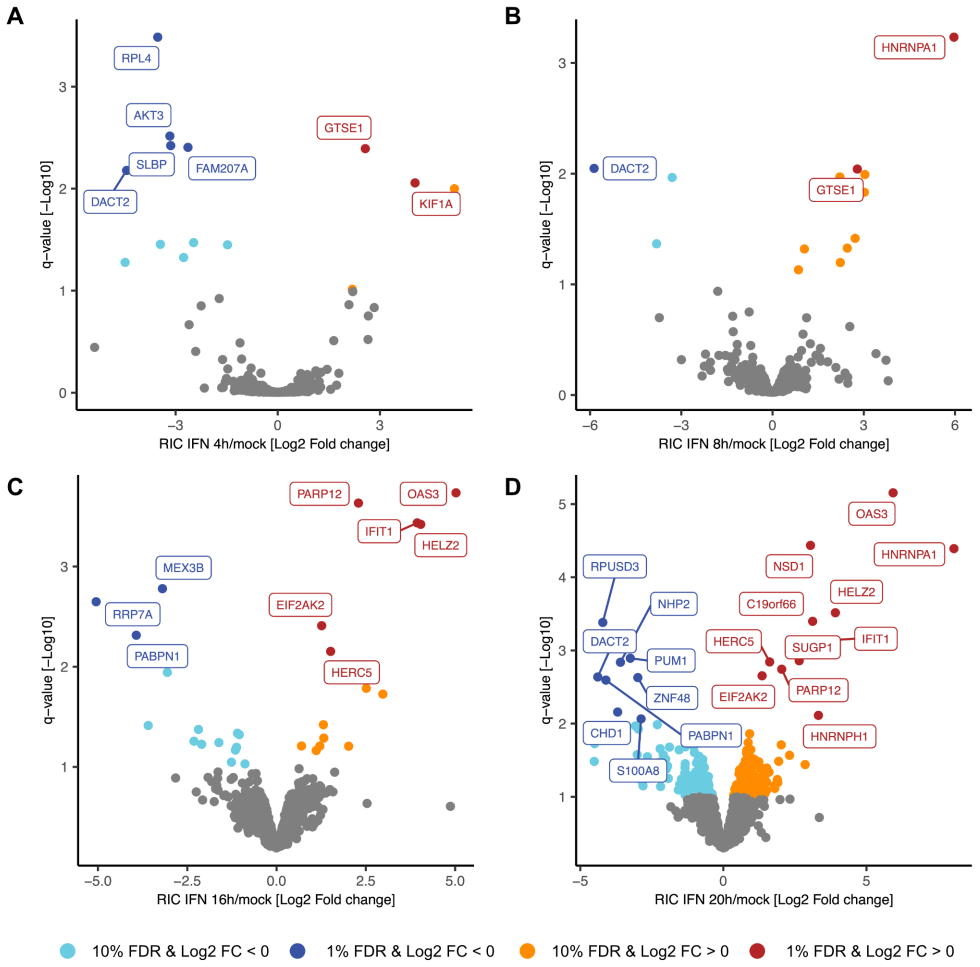


Figure 2.4: **Changes in the RNA-bound proteome become more pronounced with time.** Volcano plots summarising the results of limma testing for four sets of comparisons: mock v 4 hours (**A**), mock v 8 hours (**B**), mock v 16 hours (**C**), and mock v 20 hours (**D**). Log2 fold-change between the conditions is plotted against -log10 of the q-value. Proteins with increased binding after IFNα2 treatment are represented by red (1% FDR) and orange (10% FDR) points. Proteins with decreased binding after IFNα2 treatment are represented by dark blue (1% FDR) and light blue (10% FDR) points. Proteins with no significant change in binding are shown in grey. Proteins with <1% FDR are labeled.

RNA-binding activity can be dependent on protein abundance (more protein, more binding) or independent (more or less binding with the same amount of protein). To determine which type of regulation was prevalent in the IFN-I timecourse, I compared the fold changes of RBPs with significantly altered RNA-binding at 20 hpt (10% FDR) with those significantly changing in the transcriptome (5% FDR) and whole cell proteome (10% FDR). Surprisingly, only 9 RBPs

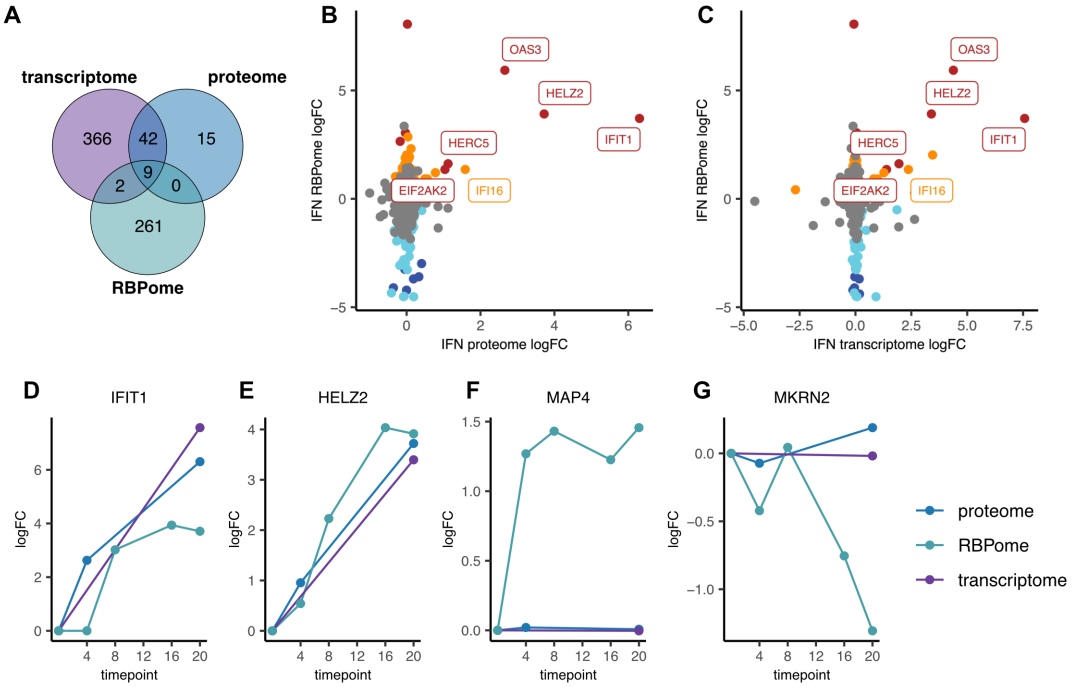


Figure 2.5: **Changes in the RBPome do not correlate with changes in the transcriptome or proteome.** (**A**) A Venn diagram comparing proteins that change significantly in transcriptome, proteome, or RBPome at 20 hours post IFNα2 treatment. Significance is defined as 10% FDR for proteome and RBPome, and 5% FDR for transcriptome. Both up- and down-regulated proteins were included. (**B+C**) Log2 fold-change 20 hours post-IFN treatment at the proteome (B) or transcriptome (C) level was plotted against RBPome Log2 fold-change. Proteins were coloured according to significance in the RBPome dataset (defined in Figure 2.4). (**D-G**) Log2 fold-change at proteome (blue), transcriptome (purple), and RBPome (teal) level were plotted across all four timepoints for four illustrative IR-RBPs. Transcriptomic and proteomic data from Chen et al (in preparation).

were changing significantly in all three datasets (Figure 2.5A). Plotting fold change in RBPome versus fold change in either the proteome (Figure 2.5B) or transcriptome (Figure 2.5C), I could distinguish two sets of RBPs. There was a small group of well-characterised antiviral ISGs for which the correlation between RNA/protein abundance and RNA-binding activity was high, and a large group for which there was discordance between datasets. These two distinct behaviours could also be observed in the kinetic profiles of individual proteins. ISGs IFIT1 (Figure 2.5D) and HELZ2 (Figure 2.5E) showed similar or smaller changes at the level of RNA binding relative to protein and RNA abundance. This is consistent with a model in which changes in RNA and, subsequently, protein levels drive the observed change in RNA-binding. IR-RBPs MAP4 (Figure 2.5F) and MKRN2 (Figure 2.5G), on the other hand, showed robust changes in RNA-binding activity with no changes at the RNA or protein levels. Most IR-RBPs

exhibited behaviour similar to MAP4 and MKRN2, suggesting that modulation at the level RNA-binding activity, rather than at the level of abundance, is the dominant force in shaping the IFN-I-regulated RBPome.

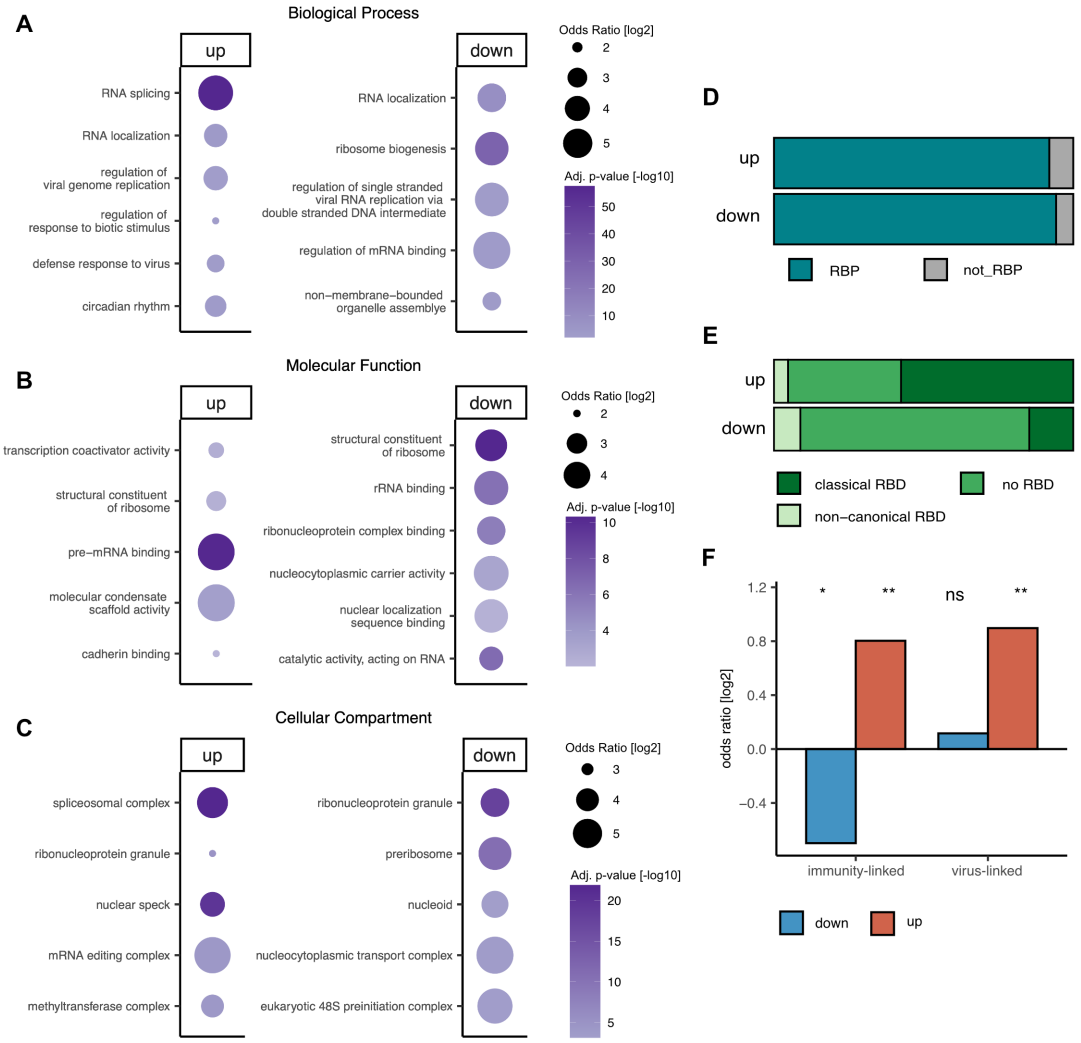


Figure 2.6: **IR-RBPs are enriched in proteins linked to RNA processing and innate immunity.** (**A-C**) GO enrichment analysis of proteins with significantly increased (up) or decreased (down) RNA-binding activity 20 hours post IFNα2 treatment. The most enriched GO terms after grouping GO terms based on similarity are shown. Shade of purple represents -log10 p-value, and point size represents enrichment odds ratio. (**D**) Percentage of proteins categorised as RBPs for 'up'/ 'down' proteins, as defined for A-C. Proteins were classed as RBPs if they were present in at least three datasets in RBPbase. (**E**) Percentage of proteins in 'up'/ 'down' groups with either classical RBDs, non-canonical RBDs, or no RBDs, based on Pfam annotation. Proteins with both classical and non-canonical RBDs were counted in 'classical RBD' group. (**F**) A bar plot showing enrichment of 'up'/ 'down' groups in proteins linked to either immunity or viruses in the literature. A protein is considered 'linked' if there are 6 or more papers with the protein's name and either 'virus'/'viral' or 'immunity'/'immune' mentioned in the title or abstract in PubMed. Statistical enrichment was assessed using Fisher's Exact Test (P<0.01 **, P<0.05 *).

IR-RBPs represent a new layer of IFN-regulated proteins that require characterisation. To determine the overall composition of the IR-RBPome, I performed GO enrichment analysis on significantly up and downregulated IR-RBPs at 20hpt. The predominance of RNA-related annotations in all three ontology classes is a reassuring indication that RIC successfully enriched for RBPs (Figure 2.6A-C). This is further supported by the fact that almost all IR-RBPs in the dataset have been previously classified as RBPs by three or more experiments in RBPbase (Figure 2.6D). These RBPs include both those with classical RBDs and those without, indicating that both canonical and non-canonical RBPs are regulated by the IFN-I response (Figure 2.6E). While both upregulated and downregulated IR-RBPs are enriched in GO terms related to RNA metabolism, there are also differences between the functions that are enriched. Upregulated IR-RBPs are strongly enriched in proteins linked to splicing, which might reflect a role for IR-RBPs in regulating IFN-I-driven alternative splicing. There is also enrichment for GO terms related to the antiviral response ('Regulation of viral genome replication', 'regulation of response to biotic stimulus', and 'defense response to virus'). This indicates that there are strong links with innate immunity, despite the limited correlation with IFN-driven proteome and transcriptome-level changes observed in Figure 2.5. A link with the antiviral response is supported by the results of text mining analysis (Figure 2.6F). I used the RISmed package in R to search for abstracts containing the name of each IR-RBP alongside 'immunity' or 'virus'-related terms. This analysis revealed significant enrichment for immunity and virus-linked proteins among upregulated IR-RBPs, whereas downregulated RBPs were significantly under-enriched in immunity-linked proteins. Taken together, these analyses revealed i) that the identified proteins are compatible with *bona fide* RBPs, and ii) that they are linked to viral gene expression and immunity. This speaks to the quality of the dataset and its potential value in understanding new aspects of the antiviral response.

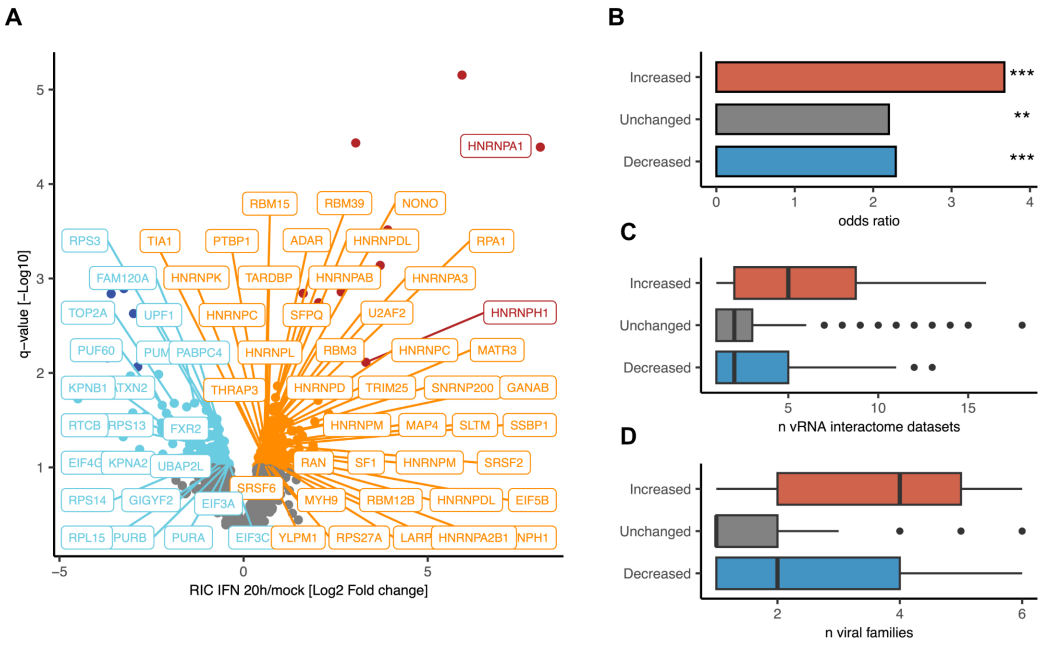


Figure 2.7: **Many IR-RBPs are members of the core vRNA interactome.** (**A**) The volcano plot in Figure 2.4D is shown again here with all significantly changing proteins that are members of the core vRNA interactome labelled. (**B**) A bar plot showing enrichment of proteins in the IFN RIC datasets in members of the core vRNA interactome. Enrichment was tested using Fisher's Exact Tests for proteins with increase, decreased, or unchanged RNA-binding activity 20 hpt (P<0.0001 ***, P<0.01 **). (**C-D**) Boxplots showing the distribution of number of datasets (**C**) or viral families (**D**) among proteins with increase, decreased, or unchanged RNA-binding activity 20 hpt.

In Chapter 1 and in Iselin et al (2022) [192], I outlined a meta-analysis that I performed to identify conserved interaction partners of vRNA. The rationale behind this analysis was that proteins that were consistently found to bind to vRNA were likely to be playing important roles in the viral lifecycle or in the host defence response. To determine whether any core interactome members were also IR-RBPs, I labelled all significantly changing vRNA interactome members on the IFN RIC volcano plot at 20 hpt (Figure 2.7A). Excitingly, in addition to proteins with well-characterised roles in infection like TRIM25, I found proteins that I had highlighted in my review as reflecting interesting new themes in vRNA-protein interactions. These include MAP4 and MYH9, which are proteins associated with the cytoskeleton, and which are both upregulated IR-RBPs. In addition to individual examples, I also wanted to determine whether there was overall enrichment of vRNA binding proteins among IR-RBPs. Based on enrichment analysis using Fisher's exact test, I found that all groups of RBPs in the IFN RIC experiment (downregulated and upregulated IR-RBPs as well as unaltered RBPs) were significantly enriched in core vRNA interactome members, relative to the whole RBPome.

However, the enrichment is most pronounced among upregulated IR-RBPs, possibly due to an enrichment of antiviral RBPs that engage with vRNA (Figure 2.7B). Downregulated IR-RBPs do not show pronounced enrichment in core vRNA interactome members, relative to non-IR-RBPs. However, there is a visible shift towards downregulated IR-RBPs being present in more vRNA interactome datasets and occurring in the vRNA interactomes of viruses from a larger number viral families relative to non-IR-RBPs (Figure 2.7C&D). In all, many RBPs whose binding activity is regulated by IFN-I interact directly with vRNA. Whether these two properties are linked and whether IFN-I can directly regulate vRNA binding will need to be assessed in future.

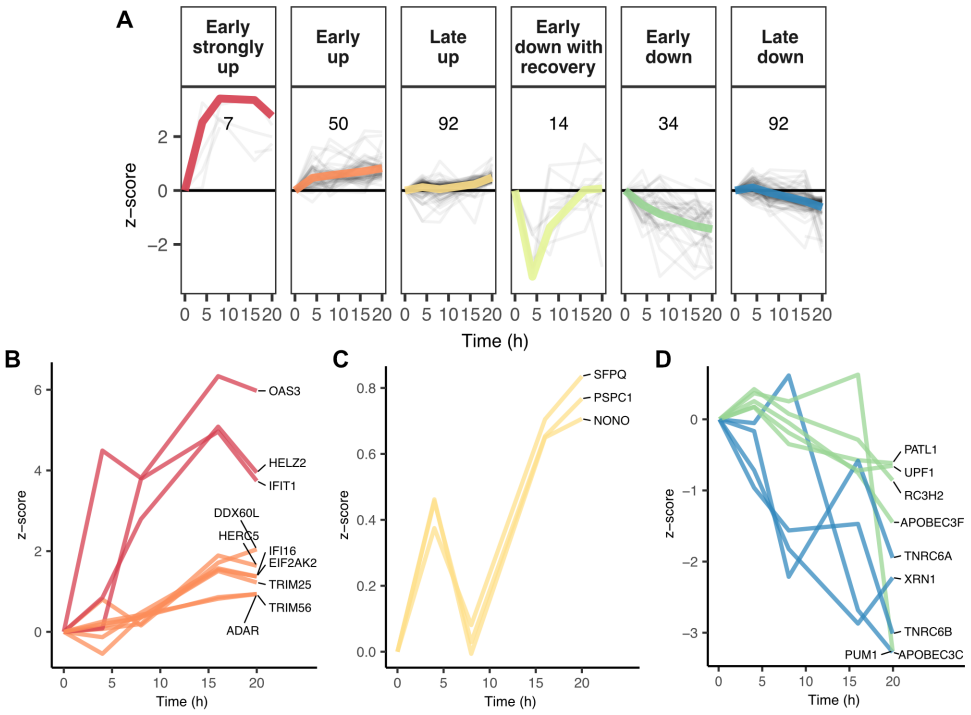


Figure 2.8: **Groups of IR-RBPs show distinct kinetic profiles.** (**A**) Results of k-means clustering are shown as line plots, with time plotted against z-score. Z-score represents the number of standard deviations away from the mean log2 intensity that a given intensity is at a given timepoint. Each grey line represents the profile of z-scores across all four timepoints for a given protein. Proteins are separated based on k-means clustering results. The thick, coloured line represents the mean z-score profile across each cluster. Cluster size is indicated within each plot. Kinetic profiles for specific groups of proteins are plotted, as in panel A, in panels B-D. Each line is labelled with the name of the protein it represents. Lines are colours according to the cluster that they belong to (as specified in panel A). (**B**) shows ISGs, defined as proteins that are significantly upregulated in the IFN-I transcriptome dataset (adjusted p-value ≤ 0.05). (**C**) Shows core paraspeckle proteins, based on how they are defined in the literature. (**D**) shows proteins annotated with the GO term 'P-body'.

A strength of my IFN RIC dataset is the kinetic resolution provided by having four IFN-I treatment timepoints. Proteins that co-vary over time at the level of protein abundance [232] and RNA-binding activity [190] often share functional properties, offering insights into proteins with unknown functions. To identify proteins that co-varied in this dataset, I used k-means clustering to group proteins based on the direction and magnitude of change observed across the four timepoints. Clustering highlighted six distinct patterns of behaviour: Early strong enhancement; early enhancement; late enhancement; early inhibition with recovery; early inhibition; and late inhibition (Figure 2.8A). I found that proteins sharing biological properties were often grouped together in a given cluster. For example, all ISGs cluster in the 'early strongly up' or 'early up' groups (Figure 2.8B). Among those that show early strong enhancement is IFIT1, which is known to be one of the earliest responders to IFN-I treatment [233]. The separation of ISGs into two clusters may be a reflection of different roles played in the IFN-I response. Interestingly, the core components of paraspeckles not only cluster together in the 'late up' group, but show strikingly similar kinetic profiles (Figure 2.8C). This is true not only for the core paraspeckle proteins, but for a wider group of proteins that have been annotated as 'paraspeckle associated' (data not shown). Such consistent behaviour strongly suggests that they are regulated in concert via a similar or highly parallel mechanism. Interestingly, all P-body associated proteins identified in the dataset exhibit decreased RNA-binding activity in response to IFN-I (Figure 2.8D). While the patterns of change are not as consistent as for paraspeckles, there are groups of proteins within this category that show similar trends. For example, PATL1, UPF1, and RC3H2 (also known as roquin-2), which all play roles in RNA degradation, all exhibit a consistent pattern of change.

2.2.2 A targeted screen to identify antiviral IR-RBPs

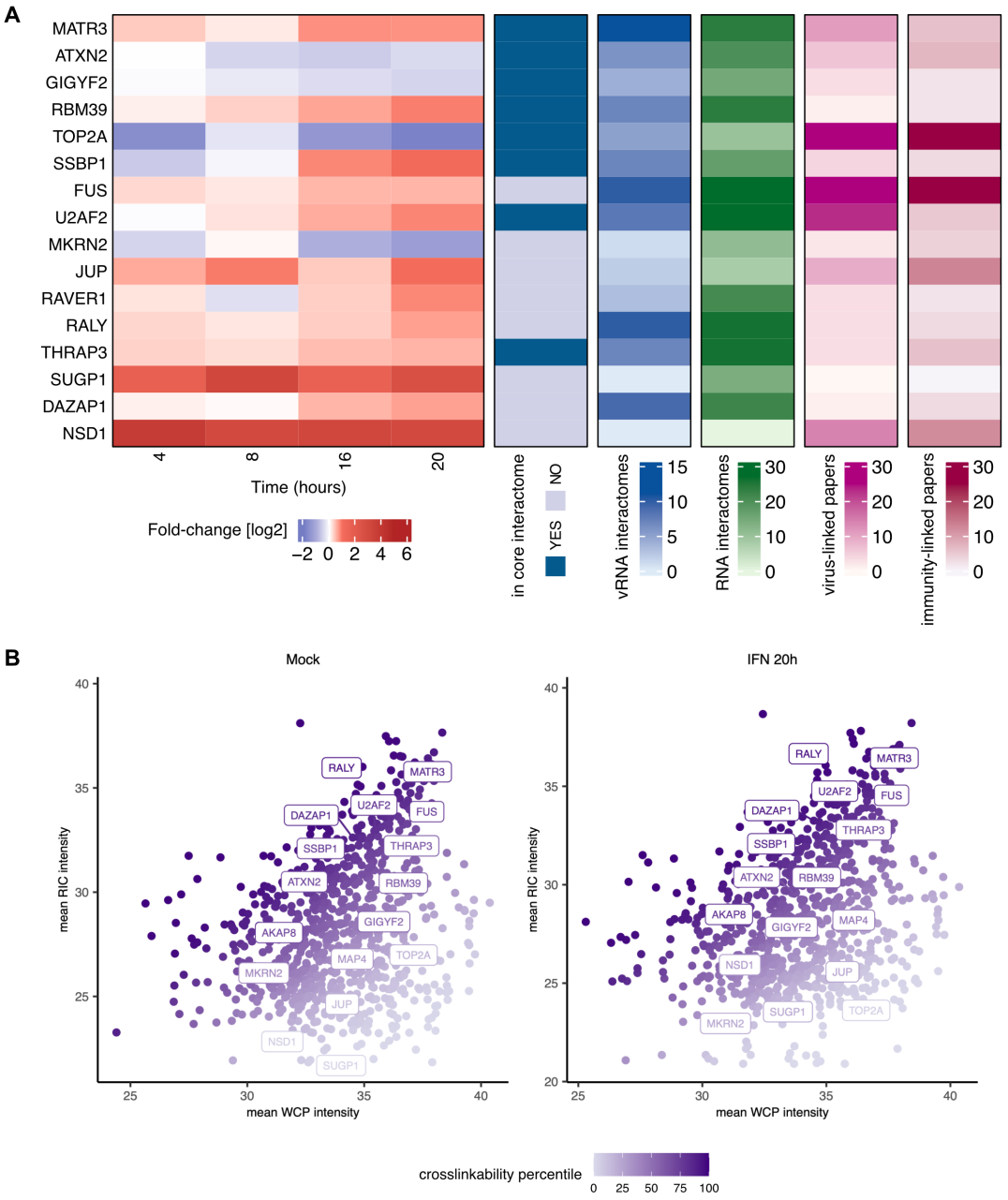


Figure 2.9: **Features of candidates selected for a target screen.** (**A**) A heatmap summarising some of the metrics used to identify candidates for follow-up as part of a targeted screen. (**B**) Scatter plots showing the mean log2 intensity values for proteins in the whole cell proteome dataset generated by Dr Chen, versus the log2 intensity values in the IFN RIC datasets. Results are plotted both for mock and IFN 20-hour conditions. Proteins were ranked based on the ratio of RIC to WCP intensity. The shade of purple of each point reflects its percentile position in this 'crosslinkability' ranking. Proteins included in the screen are labelled.

Results of the IFN RIC experiment indicate that RNA-binding activity is a new layer of IFN-I-driven regulation. Given the role of the IFN-I response in establishing an antiviral state, I predicted that among the upregulated IR-RBPs identified there would be novel antiviral proteins. Additionally, I predicted that among downregulated IR-RBPs, there might be proteins that promote viral infection that are being modulated by IFN-I to reduce this 'pro-viral' activity. To test these hypotheses, I selected 16 candidate IR-RBPs for a small-scale, targeted screen.

I selected these proteins based on several criteria:

1. *Induction pattern.* IR-RBPs with a similar pattern of induction to well-characterised antiviral proteins represented an obvious first choice for selection. Additionally, I immediately excluded proteins with erratic kinetic profiles, reasoning that these were more likely to be stochastic, non-regulatory proteins. As mentioned above, I hypothesised that proteins with decreased binding activity might be dependency factors in viral infection so I decided to include both upregulated and downregulated IR-RBPs in the screen.

2. *vRNA binding.* Direct interaction with of IR-RBPs with vRNA is a strong link to virus infection. Given the pan-viral activity of the IFN-I response, I prioritised proteins that were members of the core vRNA interactome or proteins that were present in a large number of vRNA interactome datasets [192]. However, I did not exclude non-vRNA-interacting proteins that met other criteria as direct vRNA interaction is not the only way in which a protein might alter infection outcomes.

3. *Strong evidence for cellular RNA-binding activity*. Since this study is focused on IFN-regulated *RBPs*, I prioritised proteins for which there is robust evidence of RNA-binding activity to decrease the risk of false positives. I considered both the number of prior studies supporting the RNA-binding activity of a given IR-RBP and its 'crosslinkability'. I determined 'crosslinkability' in both mock and IFN-treated conditions by determining the mean log2 intensity of each protein in the IFN RIC experiment relative to the whole-cell proteome. Proteins with low and high 'crosslinkability' can be biologically important RBPs, so for this screen I chose proteins with a wide range of crosslinking profiles. However, I focused on proteins with high 'crosslinkability' for follow-up, as proteins that crosslink poorly are more challenging to probe mechanistically.

4. *Prior literature.* The goal of this screen was to assess the value of the IR-RBPome as a source of *new* players in the cellular antiviral programme, rather than provide

confirmation of those that have been extensively characterised. Therefore, I chose to focus on proteins with relatively little or no evidence linking them to viruses or immunity. Additionally, based on a literature search, I prioritised proteins for which I could envision potential mechanistic links with infection.

An overview of the selected candidates, along with metrics that contributed to their selection, is provided in Figure 2.9A. Additionally, the 'crosslinkability' of selected candidates are highlighted in Figure 2.9B.

To assess the effect of these selected candidates on viral infection, I first established overexpression lines using the HEK293 Flp-In T-REx system [234]. In this system, a parental cell line is first generated and clonally selected. This line contains a single integration of an FRT/lacZ-Zeocin construct, and an independent integration of a Tet repressor. To overexpress the protein of interest, two further plasmids are subsequently transfected: the pcDNA5/FRT plasmid, containing the gene of interest under the control of a cytomegalovirus (CMV) promoter, and the pOG44 plasmid, which expresses the Flp recombinase enzyme. The Flp recombinase is an enzyme derived from S. cerevisiae that causes specific insertion at the FRT site, defined by the initial integration event in the parental cell line. This integration also results in a switch from zeocin to hygromycin resistance, allowing for the selection of only cells in which integration has occurred successfully. Once cells have undergone selection, the protein can be inducibly expressed through treatment with tetracycline or doxycycline.

I generated stable lines with inducible expression for 11 of the 16 candidates selected for my screen. The lines expressed the protein of interest fused to the enhanced green fluorescence protein (eGFP) at their C-terminus via a flexible Gly/Ser-rich linker that allows for independent folding of protein and tag. eGFP allows us to assess overexpression based on green fluorescence signal under the EVOS microscope (Figure 2.10). For all 11 lines, I was able to see fluorescent signal. Additionally, for most proteins, I could observe a localisation pattern that appeared to be consistent with the known subcellular distribution of the endogenous protein. MATR3, RBM39, U2AF2, RAVER1, RALY, THRAP3, and SUGP1 exhibited distribution consistent with nuclear localisation, while ATXN2 and MKRN2 displayed a pattern indicative of cytoplasmic localisation. However, it is important to note that, without DAPI staining, I cannot confirm this nuclear/ cytoplasmic distribution pattern. Similarly, for JUP and SSBP1, which localise to cell-cell junctions and mitochondria, respectively, I cannot confirm their appropriate localisation without protein markers of these sites. Western blotting

against eGFP resulted in bands at heights consistent with the molecular weight of each protein with the addition of the eGFP tag (Figure 2.12). Although green fluorescence signal was visible for all lines, it was clearly stronger in some lines than others. This is an important consideration when interpreting results from these lines, as lines with low expression levels may display lower phenotype penetrance regardless of the importance of the protein.

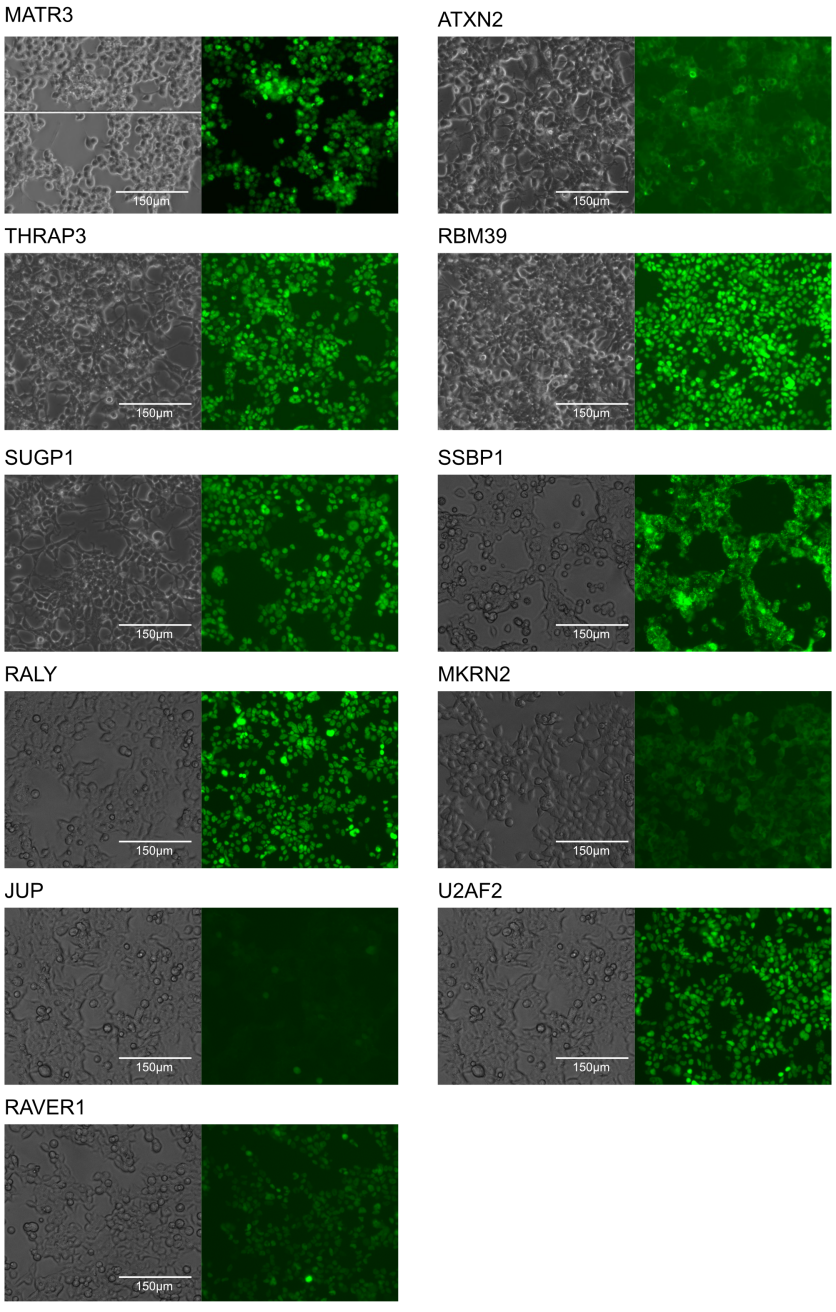


Figure 2.10: **Generating GFP-tagged overexpression lines.** Fluorescence and brightfield images of live cells taken at 20x magnification, showing the localisation of GFP-tagged proteins after 24 hours of induction with doxycycline.

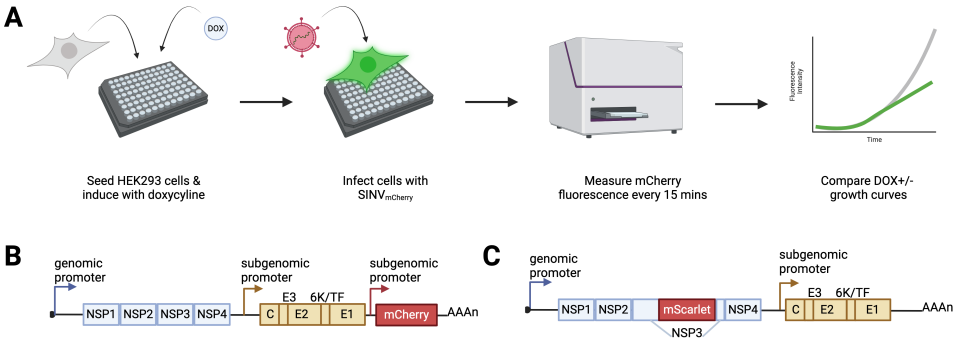


Figure 2.11: **Fluorescence-based tools for studying viral fitness.** **A** Schematic of a near real time fluorescence platereader assay to assess the effect of protein overexpression. **B** SINV$_{mCherry}$ construct schematic. **C** SINV$_{nsP3-Scarlet}$ construct schematic.

I determined the effect of overexpression of each of these IR-RBPs on SINV infection using a fluorescence-based near real-time viral fitness assay (outlined in Figure 2.11A), and endpoint Western blot analysis. I selected SINV for this screen because it is a prototypical alphavirus for which the Castello lab has several fluorescent chimeric versions. For the viral fitness assay, I used SINV$_{mCherry}$, which expresses mCherry from a second subgenomic promoter (Figure 2.11B) [50]. HEK293 FITR cells were induced with doxycycline for 24 hours, before being infected with 0.1 MOI of SINV$_{mCherry}$. They were then incubated in a CO_2 and temperature-controlled platereader. Fluorescence readings were taken every 15 minutes to allow detailed assessment of viral gene expression kinetics in cells overexpressing the protein of interest (induced) relative to the uninduced line (Figure 2.11A). A protein can be considered 'antiviral' if viral replication is significantly reduced in induced cells relative to uninduced cells. I observed a substantial reduction in viral replication following induction for MATR3, ATXN2, and THRAP3, and a subtler reduction for RBM39, SUGP1, and RALY. No difference between induced and uninduced conditions were observed with the other lines (Figure 2.12 Main panels). These trends were corroborated in all cases by Western blot, with the exception of SSBP1 and U2AF2, which showed no effect in the fluorescence-based assay but caused reduced capsid expression by Western blot (Figure 2.12 Inset panels). This disparity might be caused by the fact that capsid is expressed from a genuine sgRNA that contains translation regulatory sequences at the 5' end of the capsid sequence. Conversely, mCherry mRNA lacks these sequences, making it suitable to assess effects on replication but not to recapitulate all regulatory processes involved in translation of the sgRNA [235].

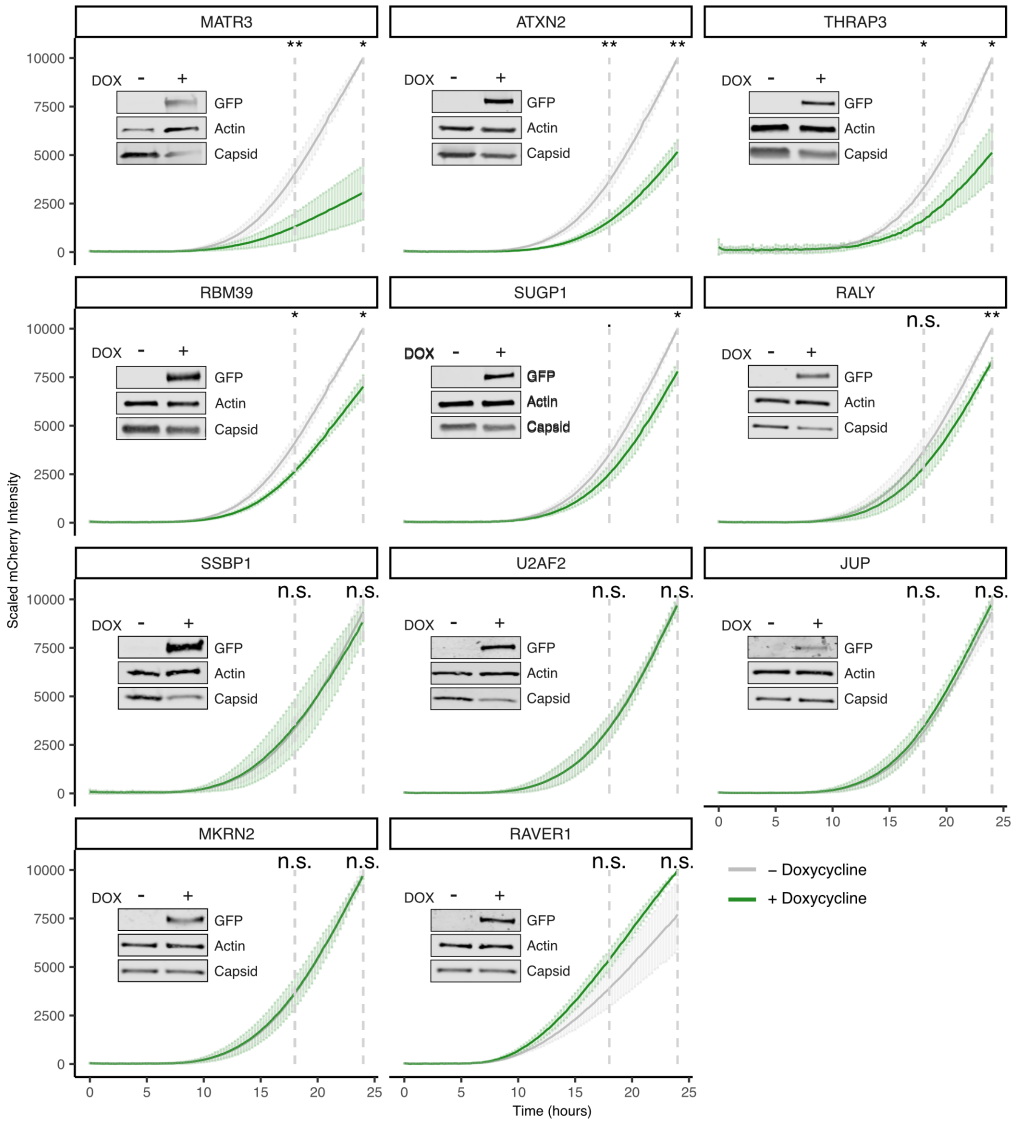


Figure 2.12: **Overexpression lines have varying effects on infection.** (**Main panels**) HEK293 overexpression lines were grown in a 96-well plate and either induced with doxycycline (+doxycycline) for 24 hours or left uninduced (-doxycycline). Cells were infected with SINV-mCherry at an MOI of 0.1 and fluorescence measurements were taken every 15 minutes for 24 hours. Fluorescence intensity, with minimum signal level set to 0, and maximum signal set to 10000, is shown as line plots. Error bars represent standard deviation across three biological replicates. Value for each biological replicate is the mean of signal intensity across three wells per condition. '+ doxycycline' condition for each protein is shown in green, and '- doxycycline' condition is shown in grey. Statistical significance was tested using two-tailed T-test (P<0.001 ***, P<0.01 **, P<0.05 *, P<0.1 ·). (**Inset panels**) HEK293 overexpression lines were seeded in 24-well plates and either induced with doxycycline for 24 hours or left uninduced. After 24 hours, cells were infected with WT SINV at an MOI of 0.5. Lysates were harvested after 18 hours and blotted against GFP, β-Actin, and SINV Capsid. 'GFP' bands reflect the band found at the molecular weight consistent with the protein of interest plus the eGFP tag.

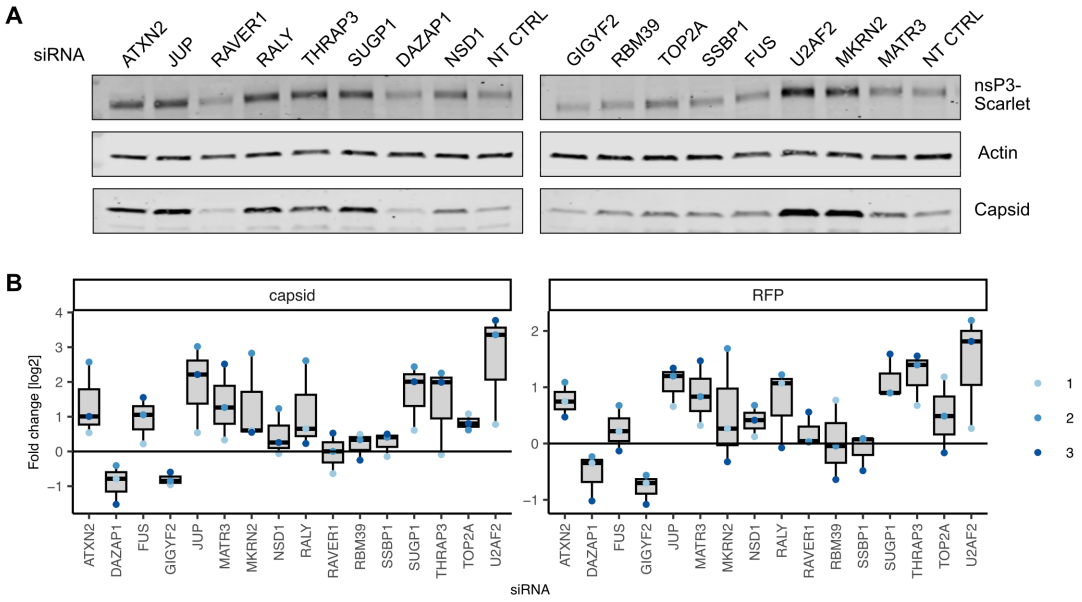


Figure 2.13: **siRNA knockdowns provide complementary information on antiviral activity.** HEK293 FITR parental cells were seeded in 24-well plates and transfected with siRNA against the protein of interest after 24 hours. After a further 24 hours of knockdown, cells were infected with SINV nsP3-Scarlet at an MOI of 0.1. After 18 hours, lysates were harvested, normalised by protein concentration, and blotted against RFP (nsP3-Scarlet), β-Actin, and SINV Capsid. (**A**) A representative Western blot, showing all siRNAs screened, split across two membranes. The non-targeting control (NT CTRL) was loaded on each membrane. (**B**) Intensity of RFP and capsid bands from three independent replicates were quantified and normalised to β-Actin intensity for each siRNA knockdown. Here, capsid/ RFP signal is represented as a fold-change relative to intensity in the non-targeting control. Replicate number of each point is indicated by shade of blue.

Since IR-RBPs are predominantly regulated at the activity level, rather than at the level of abundance, I reasoned that the native levels of expression for some of these proteins may already suffice to produce maximal antiviral effect. In this scenario, overexpression would fail to produce a detectable phenotype. To assess this possibility, I decided to deplete the IR-RBPs using siRNAs. HEK293 FITR (parental) cells were transfected with siRNAs against each of the 16 candidate genes and were infected at 24 hours post transfection with SINV$_{nsP3-scarlet}$ at an MOI of 0.1 and analysed at 18hpi . Figure 5A shows a representative Western blot for one of the replicates and Figure 4B summarises capsid and nsP3-Scarlet quantification (normalised to β-Actin) as a log2 fold-change relative to the non-targeting control across all replicates. ATXN2, THRAP3, and MATR3 depletion caused a reproducible enhancement of viral gene expression. This is consistent with results of overexpression, which resulted in the opposite effect. For

RALY, JUP, SUGP1, U2AF2, and MKRN2, I observed an upregulation of viral gene expression, despite having mild effects in the overexpression assay. For RBM39 there was no discernible effect with knockdown, despite there being a reproducible reduction in viral replication upon overexpression. Additionally, I assessed the effect of knockdown of NSD1, GIGYF2, DAZAP1, TOP2A, and FUS, for which I was not able to generate overexpression lines. Both DAZAP1 and GIGYF2 showed a reduction in viral gene expression upon knockdown, while FUS and TOP2A did not show consistent effects. However, since I did not assess knockdown efficiency, we cannot rule out that proteins for which we did not observed an effect might still be important for viral infection. In conclusion, using siRNA knockdown and an inducible overexpression system, I validated the importance of IR-RBPs in regulating virus infection. These two screens together highlight the value of using activity-based proteome-wide approaches such as RNA binding surveys to identify new players in innate immunity and viral infection.

2.3 Discussion

To date, the majority of studies categorising IFN-I-driven regulation have focused on changes in RNA or, to a lesser extent, protein abundance. These studies cannot account for other types of modulation, such as potential post-translational regulation of protein activity. In this chapter, I have shown that modulation of RNA-binding activity represents an additional layer of IFN-I-driven regulation. I applied RIC to IFN-α2-treated cells at four different times post IFN-I addition to determine not only if RBPs are stimulated or inhibited by the treatment, but also what their patterns of change are. My data showed that 272 RBPs displayed altered RNA-binding activity after 20 hours of IFN-a2 treatment, with 150 showing increased and 122 showing decreased association with RNA. Together, these data reveal that IFN-I profoundly alters the landscape of active RBPs in the cell, potentially promoting a hostile environment for infecting viruses.

2.3.1 Identifying new facets of known IFN-I regulatory mechanisms

A small group of IR-RBPs are also transcriptionally upregulated by IFN-I and so are classified as ISGs. The presence of these ISGs is not only a reassuring indication of dataset quality, but also offers an opportunity to better understand the regulatory mechanisms and RNA-binding behaviours of these proteins. For example, one of the ISGs that exhibited increased RNA-binding activity following IFN-I treatment was HERC5, which is responsible for co-translational ISGylation of proteins produced following IFN-I treatment [68, 236]. HERC5 has been shown to exert this co-translational effect through direct interaction with the ribosome [68]. However, the direct interaction of HERC5 with cellular mRNA the I observe could reflect a means of HERC5 recruitment to the ribosome or a way of stabilising it at the appropriate position for ISGylation. HERC5's mechanism of RNA binding has not been established, probably due to its low abundance and low RNA-binding levels prior to IFN-I treatment. Its RNA binding interface could be determined through use of RBDmap or RBS-ID applied to IFN-I-treated cells. Knowledge of its RNA binding interface would allow us to develop RNA binding-deficient mutants. By comparing the ISGylation profile of cells expressing WT or RNA-binding deficient HERC5, we could determine whether RNA-binding activity was required for its ISGylation activity. It would also be interesting to compare proteomic ISGylation data with HERC5 iCLIP data to assess whether binding of HERC5 to specific transcripts affected the subsequent modification profiles of their protein products.

Among the ISGs in my dataset, I was surprised to find several viral PRRs, including OAS3,

PKR, IFIT1, and IFI16. A core part of these proteins' canonical functions is their capacity for self/non-self differentiation, which allows them to act exclusively in response to or upon vRNA. However, there is emerging evidence that some of these proteins can also bind to mis-localised or mis-processed cellular RNA, such as mitochondrial dsRNA that is released during viral infection [237]. While this has been documented as an amplification strategy during infection, there is no evidence that it can directly activate PRRs in response to IFN-I alone. This is likely because sterile activation of PRRs has the potential to be extremely damaging. Indeed, PRR activation by endogenous RNA is a hallmark of many highly pathogenic type I interferonopathies [74]. Therefore, my prediction is that the binding I observe is not a reflection of canonical PRR activation by endogenous ligands.

Another possible explanation for cellular RNA-binding activity is that it relates to non-canonical functions of these PRRs, which operate directly on cellular RNA targets. For example, PKR binds to a secondary structure found in the 5' UTR of the IFNγ mRNA and represses its translation [238]. This is thought to be a means by which IFNγ signalling is repressed in the presence of an active IFN-I response and might not represent the only instance of such a role for PKR. An informative next step would be to perform iCLIP in IFN-I treated cells to determine whether there is a specific set of binding targets, and whether any of these targets are regulated in the same way as the IFNγ mRNA. A potential challenge of applying iCLIP in this context is that the binding targets of PKR are double-stranded and so would not crosslink optimally [171]. A strategy that has been employed for PKR, outside of the context of IFN-I treatment, is fCLIP-seq. This relies on formaldehyde rather than UV crosslinking so does not pose the same efficiency challenge. This approach identified many nuclear and mitochondrial dsRNA interaction partners and could be employed in combination with IFN-I treatment to identify which interactions are enhanced in the IFN-I response [239]. An alternative approach would be to use a method like hyperTRIBE or STAMP [240, 241], in which the catalytic subunit of an RNA editing enzyme (either ADAR or APOBEC) is fused to the protein of interest. Specific patterns of mutation can then be observed upon RNA sequencing, giving a sense of where the protein is bound.

There is convincing evidence of binding to cellular mRNA for PKR, but why OAS3, IFI16, and IFIT1 are being captured by RIC is less clear. OAS3 has been documented as binding to endo-RNAs, which are dsRNA fragments that arise from jammed intronic decay [242]. However, these would not contribute to RIC signal as they are not polyadenylated. IFI16

has only been identified as a vRNA sensor relatively recently and the pattern in vRNA that it recognises is still not known [45]. There is evidence for IFI16 positively regulating the IFN-I response through binding to the IFNα gene promoter [243]. One possibility is that RNA binding, as well as DNA binding, contributes to this regulatory role. IFIT1 binds to cap0 structures, which should not be present in cellular mRNA transcripts. To my knowledge, there is no evidence of IFIT1 binding to cellular RNA, except for in unphysiological contexts, such as knockdown of the 2' O methyltransferase, CMTR1 [244]. Determining the cellular targets of all of these PRRs could offer valuable insights into their activity and regulation.

While the vast majority of ISGs in the IR-RBPome exhibited increased RNA-binding activity, commensurate with their increased RNA/protein abundance, two ISGs exhibited significantly decreased RNA-binding activity. These were the cytidine deaminases, APOBEC3C (A3C) and APOBEC3F (A3F). While these proteins did not exhibit altered abundance in the transcriptomic and proteomic datasets that I was comparing my RIC data to, they have been previously identified as ISGs in several cellular contexts [245, 246]. The APOBEC3 protein family are responsible for C to U editing events, which can occur in both cellular and viral nucleic acids. For example, hypermutation of proviral DNA in retroviruses, including HIV-1, is thought to be both an important antiviral mechanism and a key driver of viral evolution. While some A3s, such as A3A and A3G, have dual RNA and DNA editing activity, A3C and A3F are only classified as ssDNA editors [247]. However, even A3s that do not edit RNA are RBPs, with RNA binding playing a central role in their regulation [248]. A3s are packaged into large RNP complexes as soon as they are translated and are sequestered within P-bodies and stress granules under resting conditions [249]. Their interaction with RNA is thought to directly inhibit their editing ability, preventing damaging mutations of cellular DNA [250]. Upon infection, APOBEC3s appear to be released from these complexes, allowing them to engage with viral nucleic acids. Interestingly, A3G and A3F are packaged into HIV-1 viral particles through their interaction with viral or cellular RNAs. These interactions still appear to be inhibitory, with editing activity only being activated upon degradation of vRNA by HIV-1's RNase H [251]. These results indicate that the decreased RNA-binding activity that I observed for A3C and A3F could reflect their release from repressive RNPs. It is interesting to note that A3C and A3F cluster with other P-body-associated proteins in IFN RIC data. All of these proteins exhibit decreased RNA-binding activity in response to IFN-I, raising the possibility that altered P-body dynamics might be a general mechanism or result of IFN-I-driven regulation. How P-bodies are regulated in

response to IFN-I treatment warrants investigation in future.

Another surprising member of the downregulated IR-RBPome is AKT3. As with the APOBECs, this kinase was not identified as an ISG in the transcriptomic data generated by Dr Chen but has been previously identified as an ISG [252] and is an important regulator of IFN-I induction. It interacts with and promotes phosphorylation of IRF3 and IRF7, inducing IFN-I signalling in response to non-canonical pathogen sensing pathways. For example, AKT3 induces IRF3 phosphorylation in macrophages in response to perturbations in cholesterol metabolism [253] and activates IRF3/7 in DCs, following viral sensing by PARP9 [254]. AKT3 has not been characterised as an RBP but it has been found to frequently target RBPs for phosphorylation, thereby regulating AS so it is possible that RNA-binding activity is involved in its engagement with these proteins [255]. In my data, I observe a significant decrease in RNA-binding activity at both 4hpt and 8hpt, followed by a return to baseline at 16hpt. One possibility is that induction of the IFN-I response causes AKT3 to disengage from homeostatic, RNA-linked regulatory functions to allow for engagement with IFN signalling complexes. Alternatively, RNA might be a direct regulator of AKT3 activity. In this case, reduced RNA-binding could represent a means of promoting phosphorylation of IFN-related targets, representing an exciting instance of 'riboregulation'. AKT3's RNA-binding properties will need to be characterised before the role of this RNA binding can be assessed. If its RNA-binding interface can be established, a next step would be to establish what the consequences of disrupting its RNA-binding activity were on its phosphorylation targets.

2.3.2 Uncovering novel players in the antiviral immune response

In addition to known IFN-regulated proteins, I identified a large set of proteins that are not ISGs. This set is enriched in proteins that have been linked to infection in the literature and/or that are direct interactors of vRNA. To determine whether the IR-RBPome could not only rediscover proteins with known roles in host defence, but identify new ones, I performed a targeted knockdown and overexpression screen. This screen identified several antiviral candidates that each warrant further investigation. It also highlighted some key themes that contribute to our understanding of the IR-RBPome as a whole.

The most enriched set of proteins that exhibited upregulation in the IR-RBPome were those involved in splicing. These included 8 SR proteins, 16 HNRNPs, and many other proteins that have been annotated as splicing regulators. By contrast, no annotated splicing regulators

were identified among downregulated IR-RBPs. This striking trend speaks to the importance of splicing regulation in shaping the IFN-I response. A role for AS has been documented for every stage of the IFN-I response, with sensors, signalling components, and antiviral ISGs all exhibiting AS. RIG-I, TBK1, MAVS, IRF3, TLR3, and IFNAR2 genes can all give rise to both the active protein and truncated proteins that play inhibitory roles, contributing to IFN-response resolution [256]. AS variants of antiviral ISGs can also shape their antiviral effects. For example, ZAP can give rise to both long and short splice variants (ZAP-L and ZAP-S). While both of these variants have antiviral activity, differences in localisation result in targeting of different viruses. Interestingly, one study found that both proteins are constitutively expressed in hepatocytes but only ZAP-S is induced by IFN-I treatment [257]. Since ZAP-S has been identified as a regulator of the IFN-I response, this might also reflect a negative feedback mechanism to prevent over-activation. Instances of IFN-induced AS that give rise to antiviral activity have also been recorded. For example, IFN-I treatment leads to increased inclusion of exon 4 in OASL, giving rise to the antiviral isoform of this protein [200, 258].

A recent transcriptomic study of IFN-treated pancreatic islet cells identified thousands of changes in AS, reflecting global remodelling of the splicing landscape. In search of a cause of this remodelling, the authors assessed differential expression of splicing regulators. They found no change in abundance for the majority of these proteins, indicating that splicing regulation is instead being driven by alterations in protein activity [200]. My IFN RIC data indicates that changes in RNA-binding activity of splicing regulators could be one route to IFN-driven AS. For example, IRF3's AS is regulated by HNRNPA1 and HNRNPA2B1. Binding of these two HNRNPs promotes inclusion of exons 2 and 3 of IRF3, giving rise to full-length protein, rather than the protein's truncated isoform, which has a negative regulatory role [259]. Both HNRNPA1 and HNRNPA2B1 have substantially increased RNA-binding activity following IFN-I treatment, which could increase binding to IRF3 pre-mRNA, promoting production of its active form. Interestingly, another splicing regulator that shows increased RNA-binding activity following IFN-I treatment is DAZAP1. This protein directly interacts with and negatively regulates the activity of HNRNPA1 at weak splice sites [260]. DAZAP1 was one of the upregulated IR-RBPs that I included in my screen and the only one that led to a consistent decrease in viral replication. One possibility is that DAZAP1 inhibits HNRNPA1-driven antiviral AS of proteins like IRF3, thus negatively regulating the IFN-I response. While HNRNPA1's RNA-binding activity is induced early in the IFN-I response, DAZAP1 exhibits more delayed

induction. Therefore, this upregulation of DAZAP1 might represent a negative feedback mechanism that contributes to response resolution. However, this is highly conjectural and will need to be assessed experimentally through identifying changes in the AS landscape in HNRNPA1/ DAZAP1 KD cells, and/or assessment of potential overlap in HNRNPA1/ DAZAP1 binding sites on ISG RNA.

In addition to DAZAP1, I assessed the antiviral activity of several other known splicing regulators, including SUGP1, U2AF2, THRAP3, RBM39, and MATR3, which all exhibited varying degrees of antiviral effect. SUGP1 was among the most strongly induced IR-RBPs in my dataset but has never been previously linked to the IFN-I response, making it a very interesting candidate for study. Overexpression of SUGP1 results in a modest decrease in SINV gene expression, while knockdown leads to a robust and consistent increase in levels of nsP3 and capsid. This modest effect upon overexpression could either be a result of the protein being in excess at steady-state or could reflect a requirement of co-factors for SUGP1 to exert its antiviral effect. While its exact role in splicing has not been defined, SUGP1 is known to associate with components of the spliceosome, including SF3B1, U2AF2 and SF1. Disruption of the interaction between SUGP1 and SF3B1 has been identified as a frequent cause of aberrant splicing in cancer [261,262], and SUGP1 has also been found to have a role in cholesterol homeostasis through AS regulation [263]. An interesting next step would be to assess whether knockdown of SUGP1 resulted in changes in patterns of AS following IFN-I response induction. This could be determined either through assessment of a panel of known IFN-induced splicing events by RT-qPCR or through global analysis of AS by RNA-seq.

Unlike SUGP1, which does not appear in any vRNA interactome datasets, U2AF2, THRAP3, RBM39, and MATR3 are all members of the core vRNA interactome. This interaction with vRNA could either reflect a mechanism of these proteins' antiviral activity that is unrelated to splicing, or could speak to an immune evasion strategy, whereby proteins important for antiviral AS are sequestered by the virus. This has been observed in reovirus infection, where the viral μ2 protein sequesters SRSF2 within nuclear speckles, leading to changes in AS. Knockdown of SRSF2 was found to promote infection, indicating that reovirus might sequester this protein to prevent its restrictive effects [264]. In Chapter 4, I will present iCLIP and protein-protein interaction data indicating that MATR3 engages specifically with the 5' end of vRNA and interacts with a set of known regulators of translation during the antiviral response. Assessment of cellular binding sites also did not provide any indication of splicing regulation occurring in

ISGs. This suggests that MATR3 is more likely to be acting directly on vRNA than indirectly through changes in splicing. Each of the antiviral IR-RBPs identified in this study will need to be assessed to determine whether they are having direct antiviral effects or indirect effects through shaping the IFN regulatory landscape. An initial indication of this could be provided through assessing IFN-I signalling induction in the presence/absence of the protein of interest. An experiment that I would have done had time allowed was to treat KD cells with poly(I:C) and assess their ability to produce IFN (through ELISA) and respond to IFN (through Western blot of STAT/pSTAT and well-known ISGs like IFIT1).

As well as proteins with increased binding activity, I was also interested in proteins whose activity was reduced upon IFN-I treatment. I hypothesised that these proteins might be negative regulators of IFN-I signalling or positive regulators of viral infection. Their downregulation could contribute to the IFN-I response either through de-repressing IFN-I signalling or sequestering factors that would otherwise engage with vRNA to promote its processing. IFN-driven downregulation of suppressive factors at the transcriptional level has been previously demonstrated. KDM5B is an H3K4 histone methyltransferase that negatively regulates IFNβ production [265]. This gene was found to be one of a small group of conserved IFN repressed genes in a study comparing transcriptional signatures across a range of species [198]. Suppression of transcription of this gene could represent a positive feedback mechanism to boost IFNβ production. Interestingly, the downregulated IR-RBP, GIGYF2, is also a negative regulator of IFNβ production that is co-opted by SARS-CoV2 to inhibit the antiviral response [266]. In my screen, I found that GIGYF2 knockdown inhibited SINV infection, suggesting that GIGYF2 could promote infection in a range of contexts. Determining whether IFNβ production is enhanced by GIGYF2 KD, both in response to poly(I:C) and SINV infection would be an important next step in validating this mechanism in my infection system. GIGYF2 will be discussed further in Chapter 3, where I will present data on how its phosphorylation state is altered in the IFN-I response.

In addition to de-repressing IFN-I signaling, reducing RNA-binding activity might be a way of reducing availability of RBPs required by viruses for their replication. Among downregulated IR-RBPs are several proteins that members the Castello group and others have identified as promoting viral infection. For example, XRN1 is a downregulated IR-RBP that promotes infection of a range of viruses. Although this pro-viral effect can occur through several distinct mechanisms, all of these rely on its ability to interact with RNA. XRN1 promotes flavivirus

infection through partial degradation of the viral genome, which gives rise to a subgenomic flaviviral (sf)RNA with pro-viral properties. sfRNA is formed as a result of secondary structure features that physically block the processivity of XRN1 [267]. These secondary structural features, along with a reliance on XRN1, are found across a range of flaviviruses [268]. Although the exact mechanism of action is unknown, it has been reported to sequester RBPs that might otherwise contribute to antiviral immunity, like TRIM25 [269]. XRN1 has also been found to promote IAV infection through increasing degradation of ISGs, leading to a reduction in immune activation [270]. Additionally, in work led by Dr Vincenzo Ruscica, we have begun to uncover a new XRN1-driven mechanism of enhancement in alphavirus infection. We found that XRN1 causes global degradation of cellular transcripts as a means of releasing nucleotides for use by the virus in replicating its genome (Ruscica et al, In Preparation). In addition to XRN1, we found that other members of the decay machinery were also required for this pro-viral activity. PATL1 is one of these accessory proteins and is also a downregulated IR-RBP. In addition to its role in supporting XRN1's activity, PATL1 has been found to play an important role in HCV replication, independently of XRN1 [271]. Both of these proteins play key roles in supporting diverse viral infections both independently and in combination. Whether the cell has evolved a mechanism to reduce availability of these proteins as a means of restricting viral infection will need to be assessed in future.

XRN1 and PATL1 are just two of many IR-RBPs that have been identified as playing pro-viral roles in infection. In work by members of the Castello lab alone, PURB, FAM120A, RTCB, and CPEB4 have all been found to promote viral infection in either SINV, HIV-I, or both ([272]; Palmalux et al, In Preparation; Noerenberg et al, In Preparation). To determine whether my IR-RBPome data could represent a source of novel pan-viral host dependency factors, I selected several downregulated IR-RBPs to include in my screen, namely GIGYF2, ATXN2, MKRN2 and TOP2A. As mentioned above, knockdown of GIGYF2 led to reduced infection, consistent with what has been seen previously for SARS-CoV2 [266]. For ATXN2 and MKRN2, I observed the opposite effect, and for TOP2A I didn't observe any effect of knockdown on infection.

ATXN2 is a stress granule (SG)-associated protein that has been previously identified as playing a role in ZIKV [273] and HCV infection [274]. It was also identified in the vRNA interactomes not just of flaviviruses, but of SINV and SARS-CoV2 as well, leading me to hypothesise that it might be playing a panviral role in promoting infection. Surprisingly, I found that overexpression of ATXN2 resulted in reduced infection, while knockdown led to increased

infection, a pattern which is more consistent with an antiviral effect. An obvious next step will be to test the effect of knockdown and overexpression in my experimental system on ZIKV infection. This would allow me to determine whether the differences from previous studies are driven by technical or biological factors. If the effects are biological, it will be interesting to determine how differences between the lifecycles of these two families of cytoplasmic RNA viruses result in these different outcomes. Since ATXN2 is an SG-associated protein, one possibility is that it reflects differences in how the viruses interact with SGs.

Another protein that I anticipated would have a pro-viral effect was MKRN2. When I selected this protein, it hadn't been previously probed as a modulator of infection. However, it had been identified as a negative regulator of inflammation [275] and was a member of the core vRNA interactome [192]. A paper has also recently been deposited in bioRxiv indicating that MKRN2 positively contributes to IAV infection [276]. Subsequent work by another member of the Castello group, Dr Wael Kamel, found that knockdown of MKRN2 could result in either an increase or decrease in viral replication, depending on which siRNA was used. Together, these data indicate that MKRN2 could still be a pro-viral factor but that different strategies will be required to definitively assess this. It would be of value to generate an MKRN2 knockout line, which would allow us to assess the significance of MKRN2 without any potential off-target effects introduced by siRNAs.

The examples of MKRN2 and ATXN2 speak to important considerations in interpreting the results of my screen and point to clear next steps. The different phenotypes observed with different MKRN2 siRNAs highlights the importance of validating results derived from individual siRNAs. Firstly, in this study due to limitations on resources, I could not check the knockdown efficiency for every protein. Therefore, it is possible that mild, non-existent, or contrasting phenotypes reflect variation in knockdown efficiency. Differences in the amplitude of effects seen across replicates might also be explained by this. One commonly used way of assessing knockdown efficiency is to perform RT-qPCR but, given the well-known disconnect between RNA and protein levels, this approach also presents technical limitations [277]. To minimise off target effects and increase the likelihood of a successful knockdown, I used Silencer Select siRNAs which are validated by the manufacturer. However, the variation seen between different MKRN2 siRNAs speaks to the fact that individual siRNAs should not be trusted in isolation. Therefore, confirmation of phenotypes with other siRNAs or, ideally, by generating knockout lines would be an important next step for all candidates, not just MKRN2.

Beyond ensuring efficient and specific knockdown, another important layer of characterisation is to ensure that successful knockdown/overexpression is not toxic to the cell. Due to the number of proteins included in this screen, there were many aspects of both knockdown and overexpression cells that I was not able to characterise in depth. I did not extensively assess the effect on uninfected cells of protein knockdown or overexpression, beyond visual inspection of cell morphology and cell growth rates. One explanation for phenotypes is that very high or low levels of a protein could have cytotoxic effects, which indirectly reduce levels of infection. For example, MATR3 overexpression led to a change in cell morphology, which may or may not be related to the antiviral phenotype. Including both overexpression and knockdown has the benefit of reducing the risk that phenotypes are only a result of cytotoxic or off-target effects; MATR3 also shows an antiviral effect based on siRNA screening, without any effect on cell morphology. I observed this myself and it was also observed by Dr. Wael Kamel, using different MATR3-targeting siRNAs.

The example of MATR3 highlights the value of having both knockdown and overexpression lines to confirm that a phenotype is biologically driven. This is also the case for ATXN2, which exhibits effects consistent with an antiviral role through both knockdown and overexpression, without having any observable effects on cell morphology. While it is still possible that the effects are technically driven in both cases, it is also distinctly possible that ATXN2 promotes infection in flaviviruses, while restricting it in alphaviruses. The antiviral immune response has evolved under pressure from a wide range of pathogens so it is highly likely that there are factors that restrict certain viruses but not others [220]. Additionally, there are examples of viruses co-opting antiviral factors to promote their replication. For example, while viperin is antiviral in a wide range of viruses, it is co-opted by HMCV to promote cytoskeletal remodelling and enhance viral infection [278]. For these reasons, performing this screen in viruses from other families with alternative replication strategies could yield highly complementary results.

In summary, this targeted multi-angle screen validates the importance of several IR-RBPs in modulating infection and pinpoints exciting candidates for follow up. Not only will it be important to assess how these proteins are modulating infection, but also how their RNA-binding activity is being regulated by IFN-I in the first place. RBPome changes can be driven by protein abundance changes, by changes in the pool of target RNAs, or by post-translational modifications of the RBPs themselves. In the next chapter, I will present data indicating that alterations in the phosphorylation landscape are one driving factor behind observed changes

in the IR-RBPome, specifically focusing on the example of MATR3. In chapter 4, I will present further characterisation of MATR3's RNA and protein interactomes during infection. Together, these results highlight one route towards characterising a given IR-RBP. I envision a similar strategy being employed for other IR-RBPs to further enhance our understanding of this new layer of IFN-driven regulation.

3 A role for phosphorylation in regulating IFN-I-driven changes in RNA-binding activity

3.1 Introduction

In Chapter 2, I presented evidence that the IFN-I response regulates the RNA-binding activity of hundreds of RBPs, independently of their protein abundance. One route by which this regulation is likely to be occurring is through post-translation modifications (PTMs). PTMs are a key mechanism by which proteins' activity can be regulated and there is extensive evidence for the importance of PTMs in regulation of the IFN-I response [279]. PTMs are a diverse group of modifications, ranging from small chemical groups, to proteins, to complex glycan moieties. There are predicted to be as many as four hundred distinct PTMs, each of which can alter protein function in a variety of ways, greatly expanding the flexibility and tunability of the proteome [280].

Phosphorylation is the most common and most commonly studied PTM. It is estimated that between one third [281] and three-quarters of the proteome can be phosphorylated [282]. There are over five hundred kinases and forty phosphatases, in addition to hundreds of regulatory factors, responsible for shaping the phosphoproteome [283]. Phosphorylation involves the replacement of a neutral hydroxyl group with a tetrahedral phosphoryl group with two negative charges, profoundly changing an amino acid's chemical and electrostatic properties. This can, in turn, alter almost any aspect of a protein's function, including structure, localisation, stability, protein interaction partners, enzymatic activity, and, importantly, RNA binding. RNA-binding sites are typically positively charged, allowing them to interact with the negatively charged phosphate backbone of RNA. Introducing a negatively charged phosphate group into this interface alters its electrostatic properties, resulting in disruption or modulation of RNA-binding activity [284].

RBPs are rich in sites of post-translational modification. A study by England and colleagues compiling available PTM databases indicated that RBPs show a significantly greater frequency of modification than proteins overall, with 95% of proteins found to have at least one modification. The vast majority of these modifications are phosphorylation events [285]. This high rate of modification is likely due, in part, to the fact that RBPs are also more disordered than the proteome at large [171]. IDRs are focal points for PTMs due to their large surface area, which provides extensive access for protein binding and modification [286].

However, phosphorylation within structured domains of RBPs is also prevalent. Indeed, one meta-analysis of phosphosites found that the DNA- and RNA-binding C2H2 ZnF domain and RRMs were among the most frequently modified protein domains [287]. An analysis of the properties of experimentally determined RNA-binding sites further indicated that they are significantly enriched in PTM sites, and particularly phosphorylation sites [152]. This speaks to the importance of RBPs as phosphorylation targets and, by extension, the potential for phosphorylation to regulate RNA-binding activity.

There are diverse instances of phosphorylation-driven regulation of RBPs, many of which occur in the context of the immune response. The GAIT complex, which regulates the translation of pro-inflammatory transcripts following IFNγ signalling, is a beautiful example of how phosphorylation can alter RNA-binding behaviour of a group of proteins in a highly regulated manner [288]. Assembly of the GAIT complex involves two temporally distinct sets of phosphorylation events. The first is phosphorylation of EPRS at residues 886 and 999, causing it to leave the multiaminoacyl tRNA synthetase complex and associate with NSAP1, forming the pre-GAIT complex. The second is phosphorylation of L13a, which causes it to leave the ribosome and engage with the pre-GAIT complex, along with GAPDH, forming the mature GAIT complex. This complex specifically recognises an RNA structure consisting of two stem-loops, which is found in the 3' UTR of target RNAs [289]. In this example, phosphorylation alters the protein interaction partners, RNA-binding specificity, and downstream functionality of these four proteins, demonstrating its immense power in modulating key cellular processes.

Phosphorylation is also central to every stage of the IFN-I response. Viral sensing pathways often involve phosphorylation. PKR's phosphorylation of eIF2α to induce translational shut-off is probably the best characterised example of this. However, phosphorylation has also been linked to signalling downstream of RLR, TLR, and cGAS-STING pathways, indicating its universal importance in pathogen sensing [30, 279]. Phosphorylation is also central to signalling downstream of IFN-I/IFNAR interaction (as described in Chapter 1). In addition to key roles in signalling, phosphorylation is also important in modulating the nature, magnitude and resolution of the IFN-I response [12, 279]. For example, serine phosphorylation of STAT1 by IKKϵ can qualitatively alter the nature of the IFN-I response by reducing the frequency of STAT1 homodimerisation and thus increasing ISGF3-driven signalling [290]. Phosphorylation has also been linked to IFN-I response resolution, as outlined in Chapter 1. For example, P38 MAPK-mediated phosphorylation of IFNAR on a degron sequence promotes its internalisation,

ubiquitination, and degradation [75,291]. Given the central role of phosphorylation in regulating of both RBPs and the IFN-I response, I decided to assess how the phosphorylation state of the RNA-bound proteome might be altered by IFN-I treatment. The results of this line of investigation are presented next.

3.2 Results

3.2.1 Establishing the IFN-regulated RNA-bound phosphoproteome

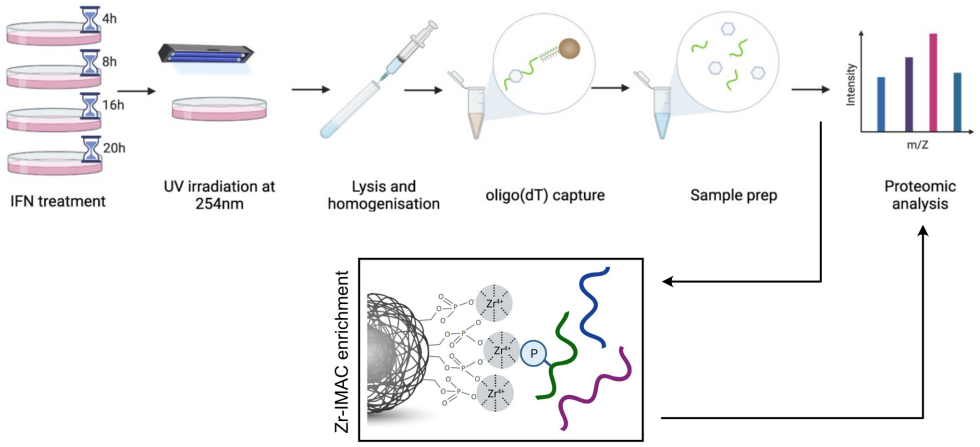


Figure 3.1: **Schematic of phosphoenrichment of RIC samples.**

Phosphorylation is an important regulator of RNA-binding activity and a key aspect of IFN-induced signaling. Given the challenges in achieving depth in whole-cell phosphoproteomic studies and the substoichiometric nature of phosphorylation [292], I decided to specifically assess how the phosphorylation state of poly(A) RNA-associated proteins changed over the course of the IFN-I response. To this end, I performed phosphoenrichment on half of the IFN RIC samples described in Chapter 2 using Zr4+ IMAC bead enrichment (Figure 3.1). This enrichment strategy relies on the chelation of immobilised Zr4+ cations with the negatively charged phosphate groups within phosphopeptides [293]. I chose to use Zr-IMAC beads for this study due to their efficiency and selectivity, particularly when starting material is scarce [294].

MaxQuant was used to identify and quantify peptides detected in phosphoenriched samples by LC-MS/MS with phosphate included in the search as a variable modification. An average of 961 phosphosites were identified across all crosslinked samples, while 173 phosphosites were identified in their non-UV-irradiated counterparts (Figure 3.2A). While phosphorylated peptides represent roughly one third of all peptides identified by LC-MS/MS, an average of 68% of overall signal was represented by phosphopeptides (Figure 3.2B). This reflects their higher abundance within the sample as a result of enrichment, relative to non-phosphorylated peptides. Intensity coming from phosphopeptides in the non-crosslinked

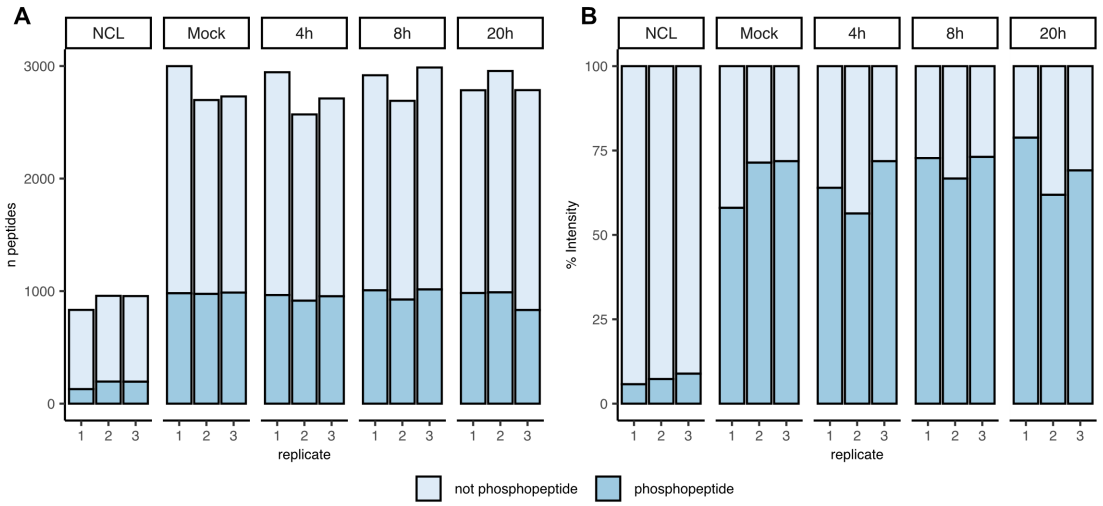


Figure 3.2: **Zr-IMAC successfully enriched for phosphopeptides.** (**A**) A bar plot showing the number of peptides identified by LC-MS/MS for each sample. Number of phosphorylated and non-phosphorylated peptides are indicated by darker and lighter shades of blue, respectively. (**B**) A bar plot indicating the proportion of overall signal coming from phosphorylated or non-phosphorylated peptides.

condition, on the other hand, was only 7%, reflecting lower enrichment efficiency due to the 'emptiness' of the input samples (See chapter 2 Figure 2.3A).

To determine whether RNA-associated phosphosites exhibited altered abundance during the IFN-I response, I performed limma testing at each timepoint, relative to the mock condition. While there were four timepoints in the original RIC experiment, I used the 16hpt samples to optimise the phosphoenrichment protocol with RIC samples. 4hpt, 8hpt and 20hpt were processed using this optimised protocol for comparison with IFN RIC inputs. Because these two sets of data are generated from the same samples, we can confidently draw comparisons between them. Overall, a similar trend to the original IFN RIC data is visible, with relatively few significantly changing phosphosites at 4hpt (Figure 3.3A) and 8hpt (Figure 3.3B), and more significant changes at 20hpt. In total, there are 65 significantly altered RBP phosphosites 20 hours after IFNα2 treatment.

Changes in RBP phosphosite intensity might be a result of the altered phosphorylation state of the protein or altered protein abundance in the original RIC eluates. To determine which of these two factors was driving phosphopeptide abundance, I compared the fold change of each phosphosite with the whole protein fold change in RIC eluates (Figure 3.4A). For proteins/phosphosites that are significantly changing in both datasets, the majority exhibit changes in the same direction. Several examples of this can be seen in Figure 3.4B. While the

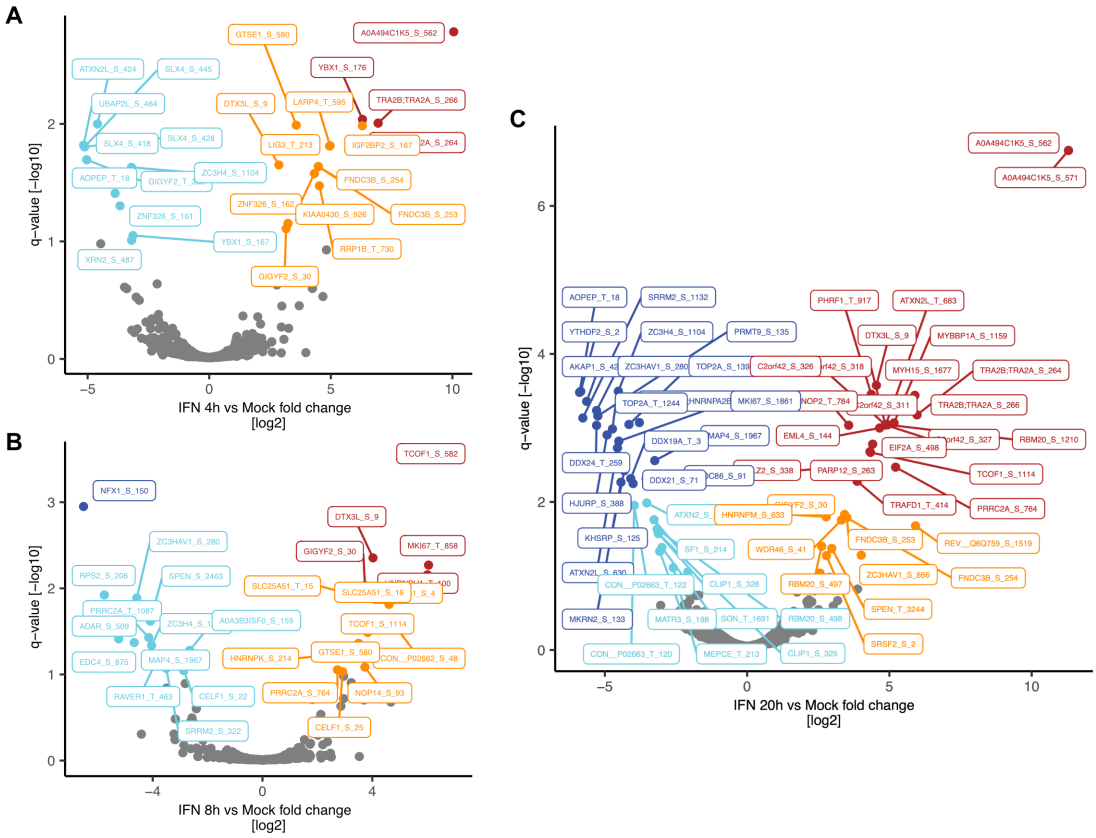


Figure 3.3: **The RNA-bound phosphoproteome is altered upon IFN treatment.** Volcano plots summarising the results of limma testing for three sets of comparisons: mock v 4 hours (**A**), mock v 8 hours (**B**) and mock v 20 hours (**C**). Log2 fold-change between the conditions is plotted against -log10 of the q-value. Phosphopeptides with increased abundance after IFNα2 treatment are represented by red (1% FDR) and orange (10% FDR) points and those with decreased abundance are represented by dark blue (1% FDR) and light blue (10% FDR) points. Phosphopeptides with <10% FDR are labeled with gene name, phosphorylated residue symbol (S or T) and phosphorylated residue position.

amplitude of change is different between the RBP and its phosphosite, the overall pattern of change is preserved. In the IFN RIC dataset, protein intensities and, therefore, fold-changes are an average over the three most abundant peptides. However, the phospho-RIC intensities are derived from the single phosphopeptide, meaning they are subject to more variation, which might explain differences in the magnitude of change. Alternatively, these differences might be biologically driven; a protein can be phosphorylated alongside having increased RNA-binding activity, or dephosphorylated at the same time that RNA-binding activity is decreased. However, deconvoluting these two possibilities is very challenging in this context.

While most phosphosites exhibit changes that correlate with changes in overall protein levels, there are some IR-RBPs for which the opposite is observed. Several proteins,

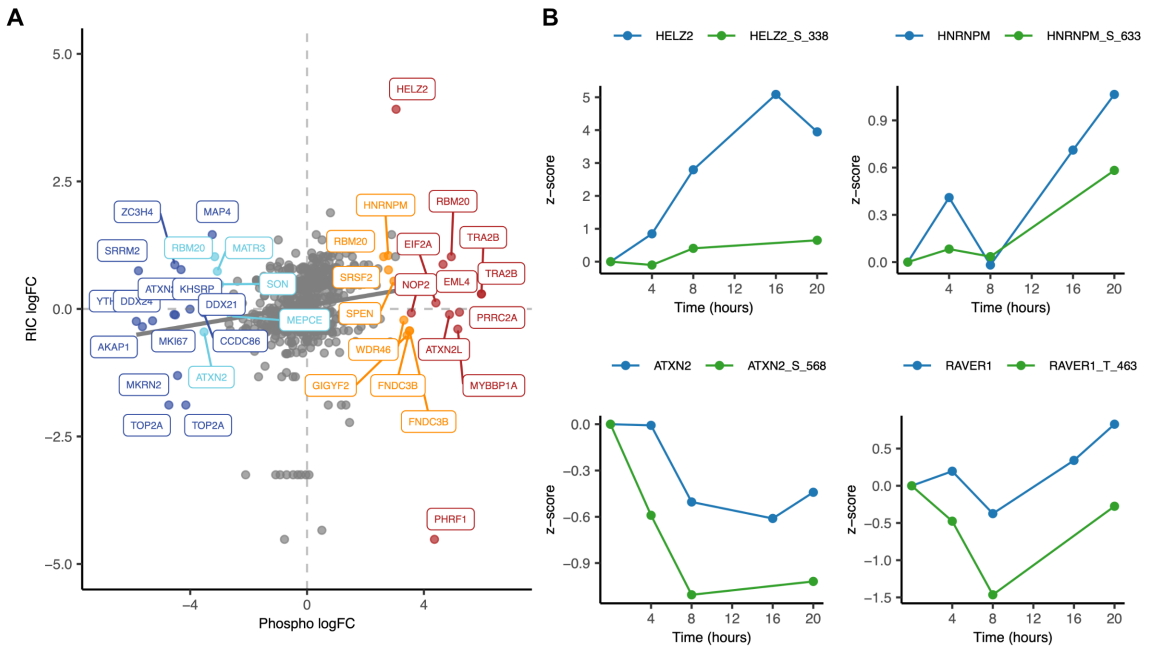


Figure 3.4: **Many changes in phosphopeptide abundance track changes in protein abundance.** (**A**) A scatter plot comparing the log2 fold-change of phosphopeptides following phosphoenrichment to the log2 fold-change of the proteins from which they are derived, prior to enrichment. Points are coloured based on significance of phosphopeptides. Phosphopeptides that are significantly changing (<10% FDR) are labelled with the gene names of the proteins from which they are derived. (**B**) Illustrative kinetic profiles of IR-RBPs that have significantly changing phosphopeptides. Changes in protein (blue) and phosphopeptide (green) abundance at each timepoint are represented as z-scores.

including RBM20, KHSRP, ZC3H4, MAP4, GIGYF2, and MATR3, exhibit a significant change in phosphosite abundance that is in direct opposition to the change observed at the protein level (Figure 3.5). This might reflect a regulatory event where phosphorylation or dephosphorylation can directly modulate RNA-binding activity. In the cases of RBM20, KHSRP, ZC3H4, MAP4 and MATR3, there is a significant increase in RNA-binding activity in response to IFN-I, while phosphorylation is significantly reduced (Figure 3.5A-D & F). GIGYF2, on the other hand, has reduced overall binding activity, while showing an increase in phosphosite intensity (Figure 3.5E). One of the benefits of having temporal information is that we can compare *trends* of change, in addition to overall differences in fold-change. Excitingly, ZC3H4 and MATR3 protein/phosphosite data show patterns of change that are an apparent mirror of each other, suggesting that phosphorylation and RNA-binding are linked. To further assess whether phosphorylation is likely to be regulating RBP's binding activity, I assessed how identified phosphosites relate to other protein features for three biologically interesting examples: MAP4,

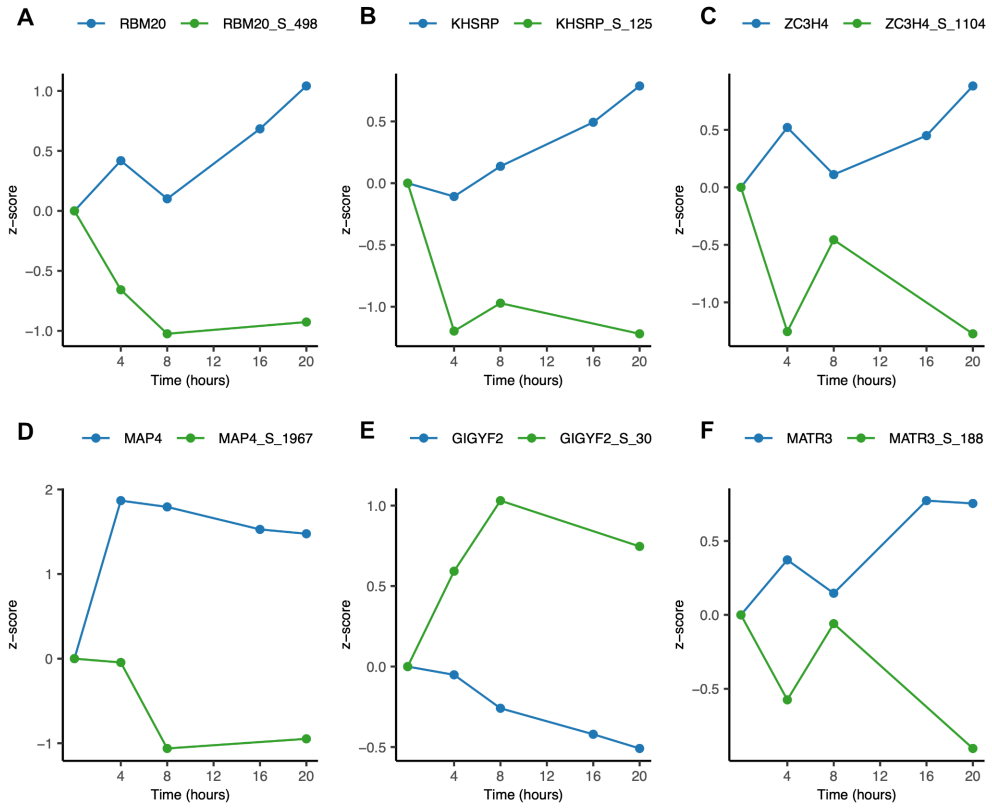


Figure 3.5: **Some proteins exhibit anti-correlation between RNA-binding activity and phosphorylation state.** Kinetic profile of changes in RNA-binding and phosphopeptide abundance over the IFN-I response for: (**A**) RBM20, (**B**) KHSRP, (**C**) ZC3H4, (**D**) MAP4, (**E**) GIGYF2, and (**F**) MATR3. Changes in protein (blue) and phosphopeptide (green) abundance at each timepoint are represented as z-scores.

GIGYF2, and MATR3.

MAP4 is a microtubule-associated protein which shows a robust increase in RNA-binding activity upon IFN-I treatment. This is in contrast to S822 site intensity, which is reduced from 8hpt onwards (Figure 3.5D). The S822 phosphosite sits within a disordered region, near to MAP4's microtubule-associated domains (Figure 3.6A). It is also located between two sites identified by RBDmap as directly interacting with RNA (Figure 3.6A). To assess the overall phosphorylation status of MAP4, I used a list of phosphosites curated by Ochoa and colleagues [295]. Plotting the frequency of phosphosites across the length of MAP4 highlights that it is a heavily phosphorylated protein (Figure 3.6B). Indeed, 10% of residues in MAP4 have a recorded phosphosite. The frequency of phosphosites is highest in the disordered region within which the S822 site sits; there are eight phosphosites within the ten nucleotides either side of S822 (Figure 3.6C). Based on meta-data collected by Ochoa and colleagues, all of these

phosphosites have been observed in at least 10 biological samples, indicating that they are unlikely to be stochastic events (Figure 3.6C). This phosphorylation cluster is highly conserved, with four of the phosphosites having evolved over 100 million years ago, which might reflect an important role for phosphorylation within this region (Figure 3.6D). In all, MAP4 S822 sits within a highly phosphorylated region close to RNA and microtubule binding domains, indicating that there is a high likelihood of it serving a biologically meaningful role, potentially in modulating RNA-binding activity.

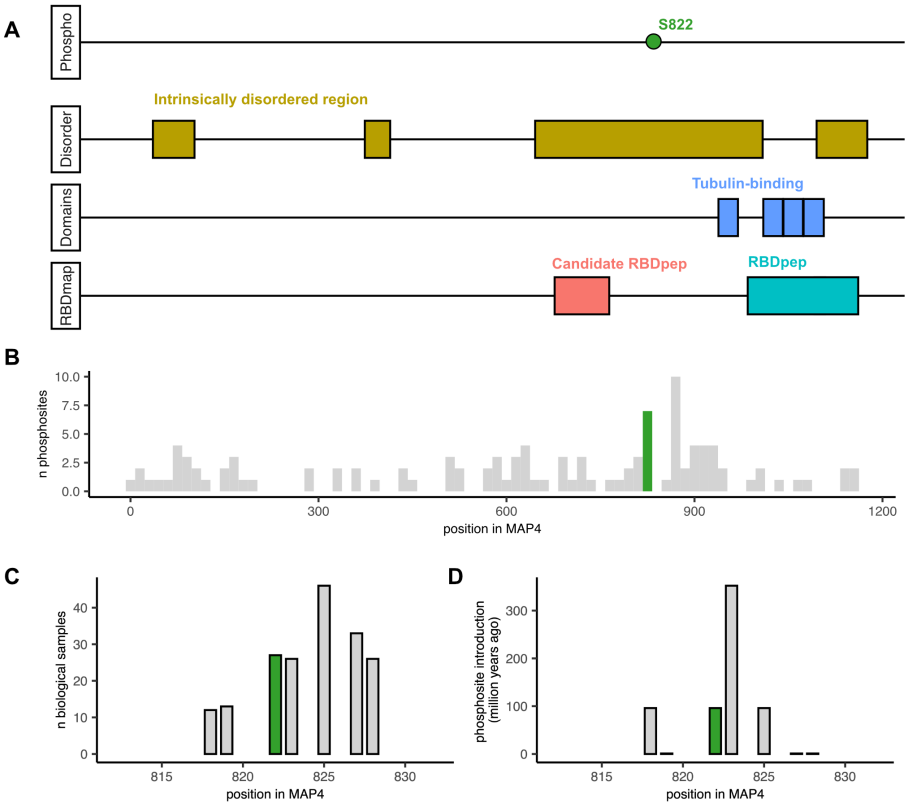


Figure 3.6: **A profile of MAP4's phosphorylation and RNA-binding properties.** (**A**) A schematic of MAP4's pertinent features. (*Phospho*) Position of phosphosite is marked by the green point. (*Disorder*) Regions annotated in Uniprot as disordered are highlighted in gold. (*Domain*) Pfam annotated tubulin-binding domains are indicated in blue. (*RBDmap*) The position of RNA-bound peptides identified at 1% FDR (RBDpep) and 10% FDR (Candidate RBDpep) are highlighted in turquoise and pink. (**B**) A histogram of the frequency of phosphosites across the length of MAP4, based on [295]. The bin containing the S822 phosphosite is highlighted in green. (**C** & **D**) Bar charts showing the number of biological samples, and the evolutionary age of phosphosites within 10 amino acids on either side of S822. Evolutionary age is represented as millions of years since a common ancestor in which the phosphosite is found. Where the value is 0, the phosphosite is *Homo sapiens*-specific. Data for both plots was derived from [295].

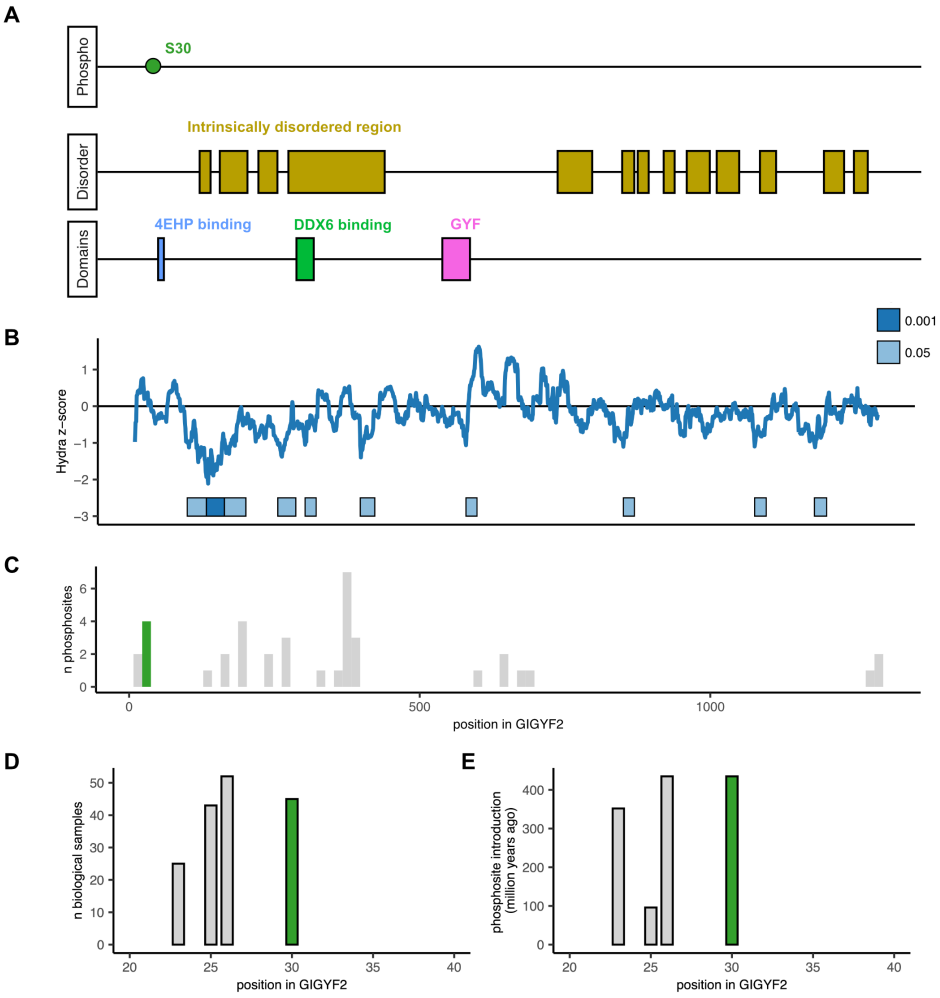


Figure 3.7: **A profile of GIGYF2's phosphorylation and RNA-binding properties.** (**A**) A schematic of GIGYF2's pertinent features. (*Phospho*) Position of phosphosite is marked by the green point. (*Disorder*) Regions annotated in Uniprot as disordered are highlighted in gold. (*Domain*) The position of the Pfam annotated GYF domain is indicated in pink. 4EHP and DDX6 binding motifs are highlighted in blue and green. (**B**) A plot showing the summary results of the HydRA RNA-binding site prediction algorithm. The line plot represents the aggregate score of all tools used to predict RNA-binding probability across the sequence of GIGYF2. The blocks below represent regions predicted to bind RNA at a p-value of 0.01 (dark blue) or 0.05 (light blue). (**C**) A histogram of the frequency of phosphosites across the length of GIGYF2, based on [295]. The bin containing the S30 phosphosite is highlighted in green. (**D** & **E**) Bar charts showing the number of biological samples, and the evolutionary age of phosphosites within 10 amino acids on either side of S30. Evolutionary age is represented as millions of years since a common ancestor in which the phosphosite is found. Where the value is 0, the phosphosite is *Homo sapiens*-specific. Data for both plots was derived from [295].

GIGYF2 has decreased RNA-binding activity in response to IFN-I treatment and is pro-viral in SINV infection (Figure 2.13). In opposition to GIGYF2's RNA-binding activity, S30 phosphorylation is enhanced following IFN-I treatment (Figure 3.5E). The S30 site is positioned

next to GIGYF2's 4EHP interaction domain (Figure 3.7A) but experimental evidence of where GIGYF2 binds RNA is lacking. To determine whether S30 might overlap with an RNA-binding region, I predicted GIGYF2's RNA-binding interfaces using HydRA. This is a tool that aggregates a range of machine-learning methods that predict RNA-binding properties based on protein sequence and interaction information (Figure 3.7B) [296]. HydRA prediction indicates that the region of GIGYF2 most likely to interact with RNA is after the S30 site and 4EHP interaction domain. This would suggest that the phosphosite is more likely to be having an indirect effect on RNA-binding activity.

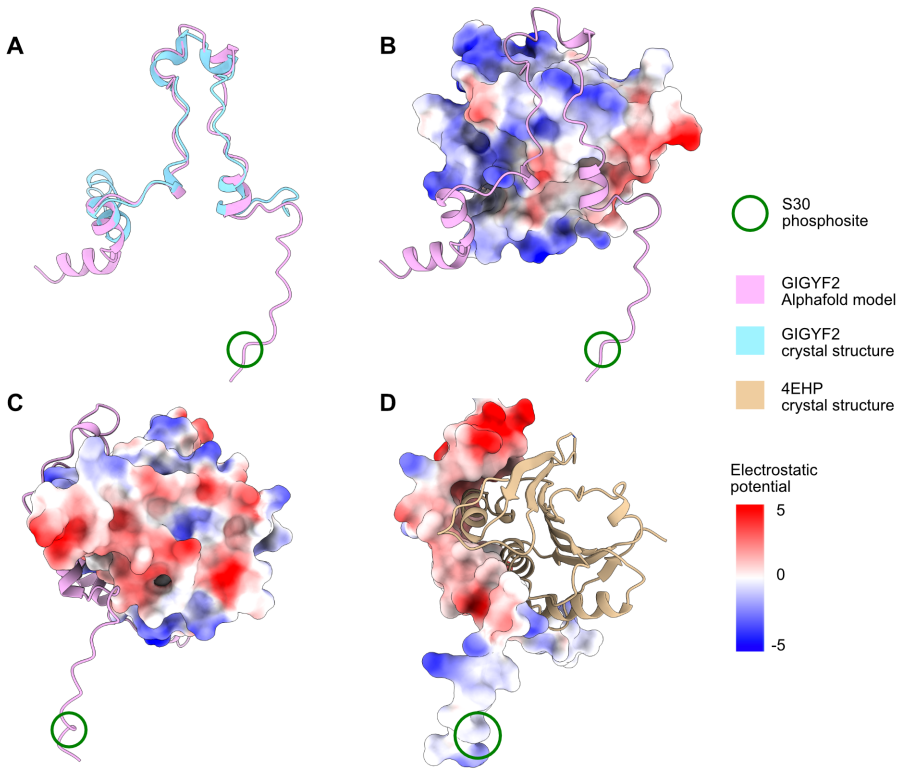


Figure 3.8: **Structural prediction of GIGYF2 S30's interaction with 4EHP**. (**A**) Alignment of crystal structure and AlphaFold predicted structures for GIGYF2's 4EHP interaction domain. The AlphaFold structure is extended to include the S30 residue. (**B**-**D**) Co-folded structures of GIGYF2 and 4EHP, with either 4EHP (B&C) or GIGYF2 (D) coloured according to electrostatic potential. In all cases, the S30 site is marked by a green circle. This figure was made in collaboration with Dr Lucia Alvarez.

One route through which S30 might be regulating GIGYF2's RNA-binding activity is through inhibiting its interaction with 4EHP. To test this hypothesis, I wanted to assess how the S30 site physically relates to 4EHP. While there is a crystal structure of the 4EHP/GIGYF2 interaction,

the GIGYF2 region used begins at residue 35 so does not include S30 [297]. However, the AlphaFold prediction closely mimics the crystal structure of this region so I used an extended region which includes this site (Figure 3.8A) and determined where it would be positioned relative to the known interaction of GIGYF2 and 4EHP (Figure 3.8B). Electrostatic mapping indicates that the N-terminal region of the interaction domain of GIGYF2, extending to S30, is positively charged (Figure 3.8D), while the region of 4EHP that it interacts with is highly negatively charged (Figure 3.8C). Whether the S30 region would interact with this positively charged interface cannot be determined from this structure prediction as it is in an unstructured region. Further experimental characterisation will be required to determine whether phosphorylation at this residue would reduce the 4EHP/GIGYF2 interaction affinity. Although GIGYF2 exhibits a lower overall profile of phosphorylation than MAP4, the regulated phosphosite is similarly placed within a phosphorylation 'hotspot' (Figure 3.8C). S30 is one of four reproducibly identified and evolutionarily conserved phosphorylated residues between residues 20 and 30 (Figure 3.8D&E). This provides a further indication that this phosphosite might be involved in regulation of an important aspect of GIGYF2's function.

MATR3 is a nuclear matrix protein with a range of functions in RNA metabolism [298]. MATR3 has two RRMs, which are generally thought to be its main sites of RNA binding [298]. However, RBDmap data suggests that there are several other regions of the protein that contact RNA (Figure 3.9A). One of these regions is an N-terminal disordered region, within which the S188 phosphosite is located (Figure 3.9A). S188 phosphorylation mirrors MATR3's RNA-binding activity following IFN-I treatment, suggesting that it might be directly modulating the protein's ability to bind RNA (Figure 3.5F). While MATR3 is a highly phosphorylated protein (Figure 3.9B), unlike GIGYF2 and MAP4, the S188 site does not sit within a cluster of phosphorylation sites. However, this site is highly reproducible and conserved (Figure 3.9C&D). Due to the striking mirroring of MATR3's RNA-binding and phosphorylation patterns, in addition to the fact that S188 site is within a high-confidence RNA-binding site, I decided to test experimentally whether S188 phosphorylation regulates MATR3's RNA-binding activity.

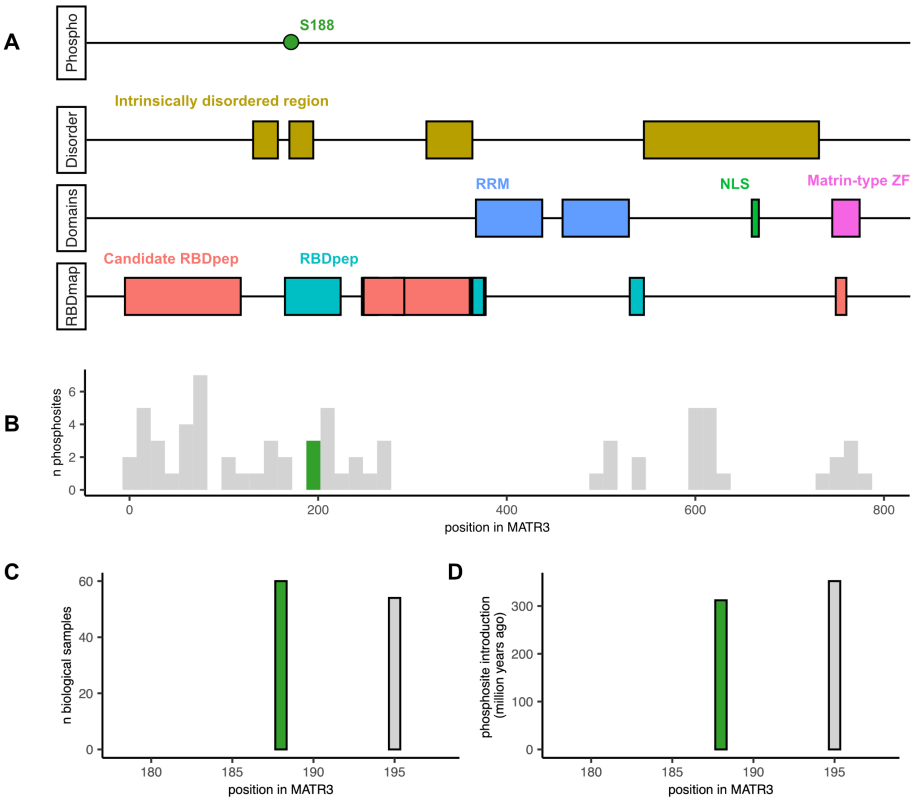


Figure 3.9: **A profile of MATR3's phosphorylation and RNA-binding properties.** (**A**) A schematic of MATR3's pertinent features. (*Phospho*) Position of the phosphosite is marked by the green point. (*Disorder*) Regions annotated in Uniprot as disordered are highlighted in gold. (*Domain*) Pfam annotated RRMs are indicated in blue. The nuclear localisation signal (NLS) is marked in green and its DNA-binding ZnF is marked in pink. (*RBDmap*) The position of RNA-bound peptides identified at 1% FDR (RBDpep) and 10% FDR (Candidate RBDpep) are highlighted in turquoise and pink. (**B**) A histogram of the frequency of phosphosites across the length of MATR3, based on [295]. The bin containing the S188 phosphosite is highlighted in green. (**C** & **D**) Bar charts showing the number of biological samples, and the evolutionary age of phosphosites within 10 amino acids on either side of S188. Evolutionary age is represented as millions of years since a common ancestor in which the phosphosite is found. Where the value is 0, the phosphosite is *Homo sapiens*-specific. Data for both plots was derived from [295].

3.2.2 MATR3's phosphorylation state alters its RNA-binding activity

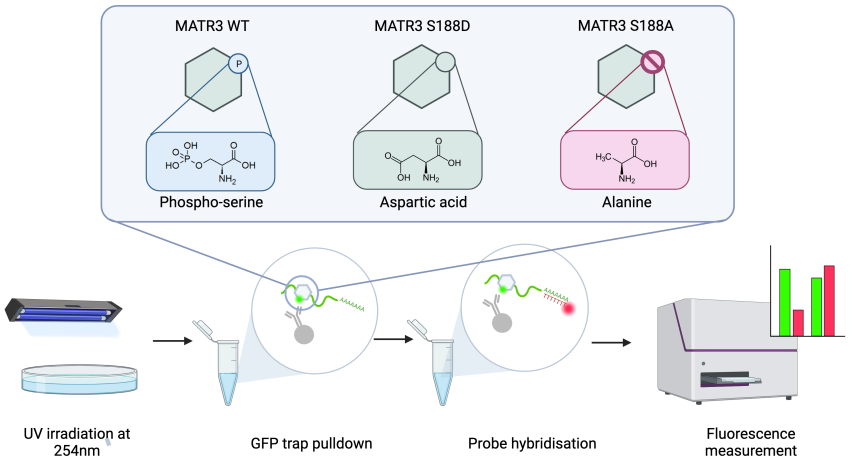


Figure 3.10: **Schematic of dual-fluorescence RNA-binding assay.** The protein constructs overexpressed in each of the three lines whose RNA-binding activity is being assessed are depicted in the blue box.

MATR3's phosphorylation at S188 exactly mirrors its overall RNA-binding activity, leading me to hypothesise that phosphorylation might directly regulate RNA-binding activity. To experimentally assess the effect of S188's phosphorylation state on MATR3, I generated HEK293 FITR lines overexpressing MATR3-GFP with mutations at residue 188 (mutagenesis was performed with the help of Dr Azman Embarc Buh). The serine was mutated to either an aspartic acid (S188D) or an alanine (S188A) (Figure 3.10A). Aspartic acid is a 'phosphomimetic' so will simulate a state in which all MATR3 is phosphorylated. Alanine, on the other hand, cannot be phosphorylated but has similar dimensions to serine, so it replicates a scenario in which there is no S188 phosphorylation.

To establish whether mutant MATR3-GFP lines exhibit the same patterns of behaviour as the line overexpressing WT MATR3-GFP, I imaged cells after 24 hours of induction by confocal microscopy (in collaboration with Dr Azman Embarc Buh). Although all three MATR3-GFP lines show similar protein localisation, the signal intensity differs markedly between lines, with the S188D line exhibiting a slight reduction in GFP intensity, and the S188A mutant showing substantially reduced signal (Figure 3.11A). This finding was corroborated by a direct read-out for GFP fluorescence intensity over the first 24 hours of doxycycline induction (Figure 3.11B). This confirms that the lines induce to different levels and with different dynamics; something that it is important to bear in mind when interpreting results.

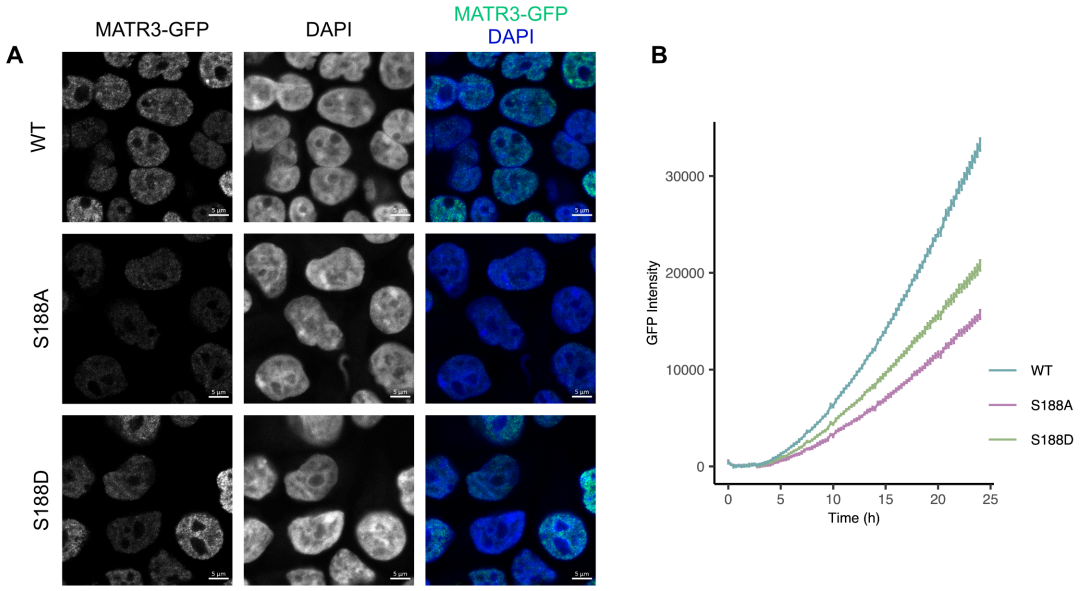


Figure 3.11: **Characterising MATR3 phosphomutant lines.** (**A**) HEK293 FITR MATR3-GFP lines (WT, S188A and S188D) were induced with doxycyline for 24h, before being fixed with paraformaldehyde and stained with DAPI. Cells were imaged by fluorescence microscopy. Scale bar = 5 µm. (**B**) HEK293 FITR MATR3-GFP lines were seeded on a 96 well-plate. 24h after seeding, doxycycline was added to media and cells were incubated for a further 24h in a temperature and CO2 controlled platereader. GFP fluorescence readings were taken every 15mins.

To establish how the phosphomimetic and phospho-null mutants affect MATR3's RNA-binding activity, I employed a dual-fluorescence RNA-binding assay (Figure 3.10B) [299]. In this assay, HEK293 FITR cells overexpressing the protein of interest tagged with GFP are crosslinked at 254nm, as in RIC. The protein of interest, along with any RNA directly interacting with it, is then captured via GFP-trap pulldown and washed to remove non-covalent interactions. Fluorescenct probes targeting the RNA of interest are then hybridised to the protein-bound RNA. Green fluorescent signal from the protein and red fluorescent signal from the RNA probes can then be read by a fluorescence platereader. The ratio of these two signals indicates the level of RNA binding. Since this assay relies on a ratio of signals, rather than absolute intensity, it can be effectively employed in contexts where input levels of GFP-tagged protein are not constant.

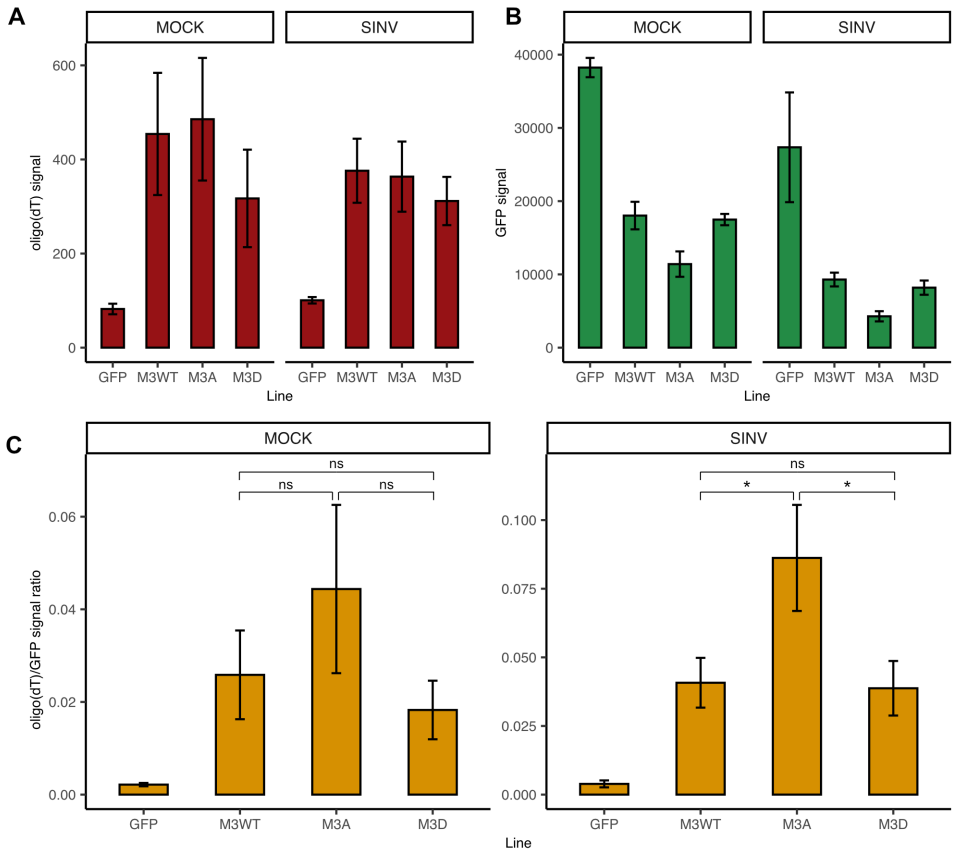


Figure 3.12: **MATR3 phosphomutants show altered RNA-binding activity.** Bar plots showing the results of the dual-fluorescence RNA-binding assay. (**A**) Alexa Fluor 594 fluorescence intensity was measured as a read-out for oligo(dT) signal. (**B**) GFP fluorescence intensity represents signal from the IPed protein. (**C**) Ratio of oligo(dT) to GFP signal is a reflection of RNA-binding level. All intensity values were blank corrected. The blank was a well in which all steps were equivalent to sample wells but no sample was added. n = 3. * $<$ 0.05 by Student's two sample t-test.

I performed the dual fluorescence assay on mock and infected samples for the WT and mutant MATR3 lines, using oligo(dT)$_{25}$ probes to assess overall binding to poly(A) RNA. A HEK293 FITR line overexpressing GFP alone was included as a control. I found that red fluorescence levels, which reflect the amount of poly(A) RNA present, were substantially lower in the GFP sample than in any of the MATR3 lines, despite the GFP sample having the highest levels of green fluorescence (Figures 3.12 A & B). This validates that the protocol is exclusively detecting RNA pulled down as a result of direct protein-RNA interaction. Due to the differences in levels of MATR3-GFP protein present in the three samples, I normalised red fluorescence intensity to GFP intensity. The normalised intensities indicate that the S188A mutant, which

cannot be phosphorylated, exhibits higher RNA-binding activity than either WT MATR3 or the S188D mutant in both mock and infected conditions, although this effect is only significant in infected conditions (Figure 3.12C). This agrees with my hypothesis that S188 phosphorylation reduces RNA-binding activity. While the S188D mutant exhibits a trend towards reduced RNA-binding activity in the mock condition, there is no difference in infected cells (Figure 3.12C). In all, while I cannot draw definitive conclusions from the phosphomimetic mutant, results from the phospho-null mutant provide a strong indication that MATR3's RNA-binding activity is modulated by its phosphorylation state.

3.2.3 What is responsible for MATR3's differential phosphorylation state?

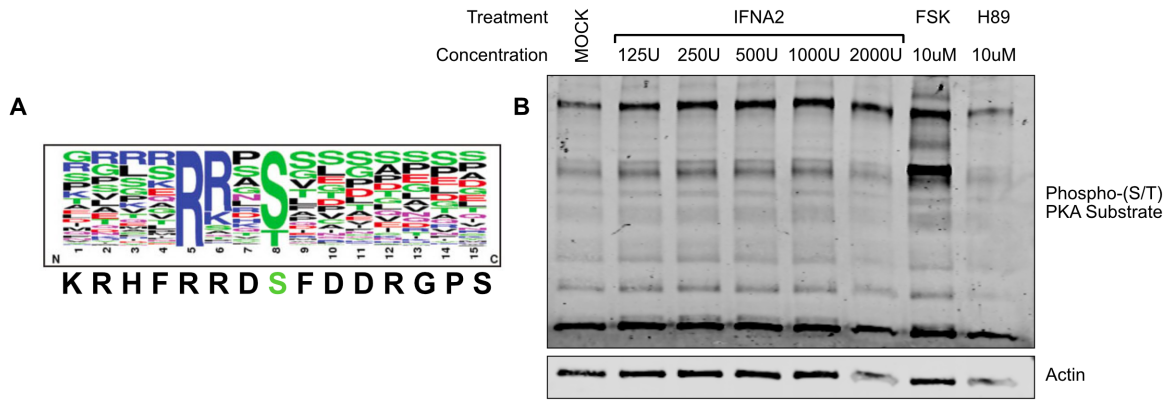


Figure 3.13: **The overall pattern of PKA phosphorylation is not altered by IFN-I treatment.** (**A**) PKA phosphosite target motif taken from PTMscan with MATR3 S188 site sequence beneath. (**B**) Western blot analysis of PKA phosphosubstrate intensity at different IFNα2 concentrations or in the presence of PKA activator (FSK, 10uM) or inhibitor (H89, 10uM). All samples were collected at 20hpt. n = 1.

Results from phosphoproteomics and the dual-fluorescence assay provide an exciting indication that the IFN-I response is modulating at least one RBP's binding activity through altering its phosphorylation state. My next question was: can we identify what is driving MATR3's differential phosphorylation? Prior literature and motif analysis, have identified PKA as one kinase that can phosphorylate MATR3 at S188 (Figure 3.13A) [300]. The simplest explanation for reduced MATR3 phosphorylation is that PKA activity is reduced. To test whether there was an overall reduction in PKA phosphorylation in the IFN-I response, I treated cells with a range of IFN-I doses for 20 hours and Western blotted using an antibody against phosphorylated sites that match PKA's substrate specificity (PKA phosphosubstrates). Each

band represents one protein target of PKA and the overall pattern gives a sense of the level of PKA-driven phosphorylation in the cell. As controls, I included Forskolin (FSK), which is a PKA activator, and H89, which is an inhibitor. There was no visible difference in the overall level of PKA phosphorylation relative to mock, even at high IFN-I doses (Figure 3.13B). This is in contrast to FSK or H89, which show increased and decrease phosphosubstrate abundance, respectively (Figure 3.13B). However, it is important to note that, based on the loading control, there was an overall reduction in protein in the H89 sample, which might account for the reduced signal. While phosphosubstrate abundance is only a proxy for PKA activity, it would appear from this experiment that there are not global changes in the PKA phosphorylation landscape following IFN-I treatment. This was performed only once, however, so replicates would be required to further confirm this result.

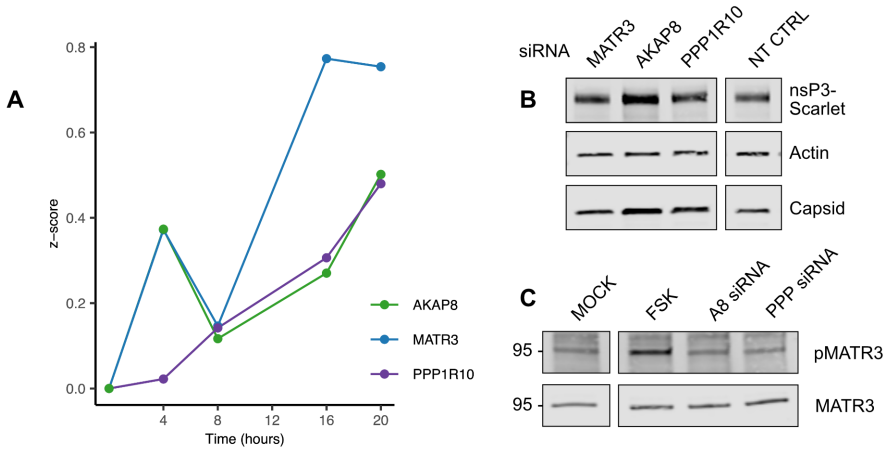


Figure 3.14: **AKAP8 and PPP1R10 do not appear to be responsible for MATR3's altered phosphorylation.** (**A**) IFN RIC kinetic profiles of MATR3 (blue), AKAP8 (green) and PPP1R10 (purple). (**B**) Western blot analysis of infection after siRNA knockdown of AKAP8, PPP1R10, or a non-targeting control (NT CTRL). Cells were transfected with 10nM siRNA and incubated for 24 hours, before being infected with SINVnsP3-Scarlet at an MOI of 0.1. Samples were analysed after 18 hours. Representative image of results from 3 independent replicates. (**C**) Western blot analysis of eluates of a MATR3 endogenous IP. Cells treated with FSK, AKAP8 siRNA or PPP1R10 siRNA, or left untreated, were harvested after 24 hours and MATR3 was IPed. Samples were blotted using MATR3 and PKA phosphosubstrate antibodies (phosphorylated/ pMATR3). n = 1.

The overall pattern of PKA phosphorylation does not appear to change following IFN-I treatment. This indicates that MATR3's phosphorylation/dephosphorylation is being regulated in response to IFN-I either by a different kinase or through a more complex regulatory network of which PKA is one part. Two candidates for proteins that might be part of such a regulatory

network are the PKA adaptor AKAP8, and the Protein Phosphatase 1 (PP1) adaptor PPP1R10, or PNUTS. Both of these proteins are RBPs whose RNA-binding activity is significantly enhanced following IFN-I treatment (Figure 3.14A). MATR3 protein-protein interaction data (shown in Chapter 4) also indicates that both of these proteins interact with MATR3 and that AKAP8 increases its association with MATR3 following SINV infection.

siRNA knockdowns of PPP1R10 and AKAP8 indicate that both of these proteins individually have antiviral effects, with knockdown of AKAP8 resulting in a robust and reproducible increase in SINV infection (Figure 3.14B). Since both of these proteins are IR-RBPs, this provides two additional interesting candidates for follow-up to add to those identified in the screen in Chapter 2. However, it does not provide an indication of whether this effect is related to an interaction with MATR3. To determine whether either of these proteins are able to alter the phosphorylation state of MATR3, I performed IPs against endogenous MATR3 in AKAP8 and PPP1R10 knockdown cells, and in cells treated with FSK. I then performed Western blotting, using antibodies against PKA phosphosubstrates and MATR3 to assess the relative proportion of PKA phosphorylated (p)MATR3 in each condition (Figure 3.14C). While FSK led to an increase in MATR3 phosphorylation, there was no change in pMATR3 level with either AKAP8 or PPP1R10 knockdown. This suggests that these are not the proteins responsible for MATR3's differential phosphorylation, although it is also possible that the lack of effect is a reflection of poor knockdown efficiency. This experiment would need to be repeated and knockdown validated with protein-specific antibodies or RT-qPCR to confirm this conclusion.

3.3 Discussion

3.3.1 Phosphoenrichment of RIC eluates is a new route to discovering functional phosphosites

My IFN RIC data suggests that many IR-RBPs are regulated post-translationally. What is responsible for the regulation of these proteins remains an open and important question. In this chapter, I explored phosphorylation as one potential regulatory avenue. To establish the phosphorylation landscape of IR-RBPs, I performed phospho-enrichment on IFN RIC samples. This allowed me to identify 65 phosphosites that significantly changed in abundance within the RBPome following IFN-I treatment. Among these were several instances of phosphosite abundance changing in the opposite direction to RBP abundance, hinting at potential regulatory events. I was able to confirm one of these instances of putative regulation for MATR3, where its RNA-binding activity is increased by reduced S188 phosphorylation.

Enriching for phosphopeptides within the RBPome, rather than in the whole cell proteome offers several advantages. One of the main benefits is that it allows us to identify phosphosites that might not otherwise be detected. Achieving consistent detection of phosphosites within the whole cell proteome is extremely challenging, due to the substoichiometric, labile, and diverse nature of phosphorylation events [292]. Since only a very small subpopulation of a given protein might be modified at a given residue, phosphopeptide signal can often be drowned out by signal from non-phosphorylated or differentially phosphorylated peptide. The additional layer of stochasticity derived from peptide selection in data-dependent acquisition (DDA) strategies means that consistent detection and quantification of phosphopeptides is hard to achieve. By enriching for phosphopeptides among RBPs only, we are significantly reducing the complexity of the sample and increasing the chances of consistently detecting phosphopeptides. We are also increasing the likelihood of any given phosphosite being relevant to the context of interest, which is, in this case, RBPs.

Another benefit of combining RIC and phosphoenrichment is that it allows us to directly generate hypotheses about how phosphorylation events modulate protein behaviour. One of the most significant challenges facing the field of phosphoproteomics is how to assign biological functions to the hundreds of thousands of phosphosites that have been identified to date. It is estimated that as few as 3% of phosphosites have an annotated function, leading some to suggest that the majority of phosphosites observed are ‘silent’ events

without biological significance [287]. Given the highly sophisticated system of regulation surrounding phosphorylation, and the dramatic effects that phosphorylation can have on a protein's structure and function, it seems more likely that strategies for probing phosphosite function have simply not kept pace with phosphoproteomics, a common issue for 'omics'-based discovery approaches. Indeed, functionally characterising individual phosphosites is extremely labour-intensive, meaning that it is necessary to focus on individual sites for which a function can be predicted.

There have been several efforts to narrow down the phosphoproteome to sites that are most likely to be functionally relevant, based on a range of criteria [287, 295]. These include: evolutionary conservation, intersection with an interaction surface, co-occurrence with other PTMs, and stoichiometry (a measure of the proportion of the protein is phosphorylated at a given site). However, probably the most frequently used selection approach is to identify sites that are altered in response to a given stimulus. In this study, we are not only determining how phosphosites change in response to a stimulus, but also how this change relates to a change in activity of the protein to which the phosphosite belongs. By combining this activity-level information with what is already known about the protein from the literature, we can rapidly generate testable hypotheses as to how the phosphosite might be operating. Such testable hypotheses for three proteins (MAP4, GIGYF2, and MATR3) will be discussed below, although these are not the only proteins in the dataset that would be interesting to characterise further.

The ability to combine stimulus-driven changes in a phosphosite with potentially phosphosite-driven changes in activity has the potential to be extremely powerful. However, the results require careful interpretation. In this experiment, we only see protein and, therefore, phosphosite that is already bound to RNA. Therefore, when we classify a phosphosite as inhibiting RNA-binding, we must keep in mind that any level of phosphorylation that we can detect is present on RNA-bound RBP. In the case of MATR3, for example, S188 phosphorylation is detected at all four timepoints on the RNA-bound protein, despite evidence that S188 phosphorylation reduces RNA-binding activity. This means that this approach is a valuable way of assessing phosphosites that *modulate* RNA-binding activity but not those that abolish it. If phosphorylation completely ablates RNA binding, no phosphorylated protein will be enriched by RIC so this phosphosite will not be detected. However, since many phosphorylation events are likely to alter but not completely abolish RNA-binding activity, there is still much that we can learn from this method.

This experiment represents an important starting point in uncovering regulation within the RNA-bound phosphoproteome. However, paradoxically, one of the most exciting aspects of my IFN RIC phosphoproteomic dataset is how much room there is for technical improvement. The decision to perform phosphoenrichment on these samples was taken after they had already been processed for RIC, meaning that the scale of the experiment and the sample collection protocol were not optimised for phosphoproteomics. In this experiment, around 70% of signal came from phosphopeptides. This was somewhat lower than pilot enrichments of whole cell lysates, where over 90% of signal was phosphopeptide-derived. The explanation for this likely relates to the small amount of input material. Despite using the minimum bead quantity advised, it is likely that the beads were far from being saturated. Under these conditions, there are not sufficient phosphopeptides to out-compete weaker interactions formed between the beads and non-phosphorylated peptides. This is clearly illustrated by the very poor enrichment in the non-irradiated control samples, which contained by far the smallest amount of starting material. Due to low amounts of starting material, I also made the decision to perform phosphoenrichment washes in NaCl, rather than glycolic acid, which is more commonly used and is much more stringent. Despite increasing non-specific interactions, this was an important step to keep sample loss to a minimum during the enrichment process. If this experiment were to be repeated with larger amounts of starting material, it would be possible to perform a more stringent and efficient phosphoenrichment. This would likely result in improved coverage of the RNA-bound phosphoproteome, and more instances of IFN-regulated phosphosites.

More than 95% of known phosphorylation events in humans occur on serine or threonine residues. However, these are not the only residues that are phosphorylated. Between 0.4 and 4% of phosphorylation events in the cell are predicted to occur on tyrosine residues (pTyr) [282]. pTyr is a critically important modification in driving signalling cascades in many cellular processes, including the IFN-I response [301]. Due to this central role in signalling, pTyr is highly regulated. Basal levels of the modification are low and highly active pTyr-specific phosphatases mean that it is very labile, making it difficult to detect without the inclusion of phosphatase inhibitors during lysis [302]. Despite challenges in studying pTyr, there is evidence that it can modulate RNA-binding activity [303]. Additionally, analysis of amino acid composition of RNA-protein interfaces indicates that Tyr is highly abundant, making it a promising candidate for regulation [152]. Given the central role of tyrosine phosphorylation in the IFN-I response, it will be exciting to know whether it is responsible for any instances of

IR-RBP regulation. I was not able to detect many instances of pTyr phosporylation due to its substantially lower frequency than pThr and pSer and because I did not include a phosphatase inhibitor in my lysis buffer. If I were repeating this experiment, I would not only include this inhibitor but would also perform pTyr enrichment of a portion of each sample using a pTyr-specific antibody to specifically explore this key regulatory modification [304].

Regardless of the type of phosphorylation being studied, this protocol might benefit from using a different MS data acquisition strategy. In this experiment, we used DDA, which, as described above, is not best suited to phosphoproteomics due to the stochastic nature of peptide selection for quantification. Data-independent acquisition (DIA) strategies, that measure all fragments within a given mass/charge (m/z) window simultaneously, have the potential to increase both dynamic range and reproducibility of phosphopeptide detection [305]. However, DIA relies on sophisticated computational pipelines and pre-defined spectral libraries to deconvolute highly complex spectra. Traditionally, DIA relies on experiment-specific spectral libraries, which have to be manually generated. More recent approaches have moved towards spectral libraries generated *in silico* using prior knowledge of how peptides fragment. However, the additional layer of complexity added by phosphopeptides makes this particularly computationally challenging, especially if phosphosite localisation is desired [306]. DIA and its application to phosphoproteomics is a rapidly evolving field so it is likely that this will become a more viable option in the coming years.

Despite the subtlety of interpretation required and the potential for technical improvements, there are several clear-cut incidents of what looks like phosphosite-driven regulation in my dataset. These are cases where RNA-associated phosphosite abundance shows the opposite trend to RNA-binding activity. Below, I will discuss three particularly interesting instances of this that speak to the exciting potential of this approach in hypothesis generation.

3.3.2 MAP4 S822: a new instance of phospho-driven riboregulation?

MAP4 is a microtubule-associated protein that acts to stabilise microtubules by directly binding to them, and by preventing the binding of 'catastrophe factors' that promote microtubule disassembly. Despite the highly disordered nature of MAP4, cryogenic electron microscopy (cryo-EM) has been used to establish the nature of its interaction with microtubules [307]. This has suggested a model of binding in which there is a central stable interaction with tubulin through the four tubulin-binding domains, as well as peripheral, flexible interactions through

surrounding disordered regions. The proline-rich region directly preceding the tubulin-binding domains is thought to be particularly important for mediating this interaction [308]. Interestingly, RBDmap data indicates that two regions of MAP4 interact with RNA: one high-confidence site, overlapping with the tubulin-binding domains, and one candidate site within the proline-rich region [152].

Phosphorylation at several residues of this proline-rich region is a key mechanism for regulating MAP4's association with microtubules [309–311]. Phosphorylation at residues S696, S768 and S787 causes MAP4 to dissociate from microtubules, leading to their destabilisation. Interestingly, hypoxia-induced phosphorylation at these three sites in neonatal cardiomyocytes results in the translocation of MAP4 into mitochondria [312]. In mitochondria, MAP4 promotes apoptosis through inducing permeability transition pore opening in a microtubule-binding-independent manner. This would indicate that, when released from microtubules, MAP4 can take on moonlighting functions.

While research has primarily focused on these three phosphosites, MAP4 is a heavily phosphorylated protein. Indeed, there is evidence for phosphorylation of ~10% of all of MAP4's amino acids. This is ten times higher than the average rate of phosphorylation across annotated phosphoproteins [295]. The majority of these phosphosites are in the proline-rich region, with the highest density occurring around the S822 phosphosite regulated in this study. This hotspot of phosphorylation is a strong indication of functionality; PTM co-occurrence is a key determinant of whether a site is biologically meaningful [313]. Protein-protein interaction sites within disordered regions often exhibit a high degree of conservation in phosphosite occurrence but not phosphosite position [314]. This is likely a reflection of the fact that phosphorylation is changing the overall electrostatic properties at the interface, rather than having a more precise effect. While this analysis was performed for protein interaction interfaces, it is likely to also apply in the case of protein-RNA interactions. Of note, the MaxQuant search results for MAP4 S822 phosphopeptide indicated that, in addition to the relatively high probability localisation at S822, there was a lower probability prediction that the phosphosite was located at S823. This might reflect the presence of two interchangeable sub-populations, in which one or other of these sites is phosphorylated. Interestingly, the S823 is a more evolutionarily ancient site than surrounding sites in the region. S823 might represent a 'founder' site, with additional phosphosites accumulating over evolutionary time to introduce redundancy or flexibility at an important site of regulation. This potential redundancy

is something it would be important to keep in mind when designing mutants to test the effect of phosphorylation.

Whether S822/823 phosphorylation is negatively regulating microtubule-binding in a way similar to S696, S768 and S787 would be an important question to address experimentally. However, in this study, I see that a reduction in phosphorylation results in an increase in RNA-binding activity. Given that MAP4's primary RNA binding site is within its tubulin-binding domains, I would expect that microtubule binding and RNA binding are mutually exclusive processes. In this case, dephosphorylation would result in a decrease in microtubule-binding and an increase in RNA-binding activity. This could be a mechanism by which MAP4 is sequestered away from microtubules through the riboregulatory action of RNA [166]. Alternatively, it might be that MAP4 serves an entirely distinct, moonlighting function as an RBP.

In addition to addressing how S822/823 affects the balance of RNA-binding/ microtubule-binding activity, another important question to address is why MAP4 is being regulated by the IFN-I response. Many viruses are known to exploit microtubule-based transport systems to move around the cell [315]. Indeed, MAP4 is a dependency factor for the early stages of infection for both HIV-1 and HCV [316, 317]. Conversely, its microtubule-stabilising function is antiviral in IAV infection, which relies on flexibility of the microtubule network for productive infection [318]. Despite these opposing roles in different viruses, we might hypothesise that the IFN-I response drives sequestration of MAP4 away from microtubules to cause their destabilisation and prevent viral transport. Complicating this picture further, however, is the finding that MAP4 associates directly with the vRNA of eight viruses from five viral families [192]. This highly conserved interaction with vRNA suggests that MAP4's RNA-binding activity might be serving its own function in infection independently of reducing its microtubule association. MAP4 represents an exciting example of an uncharacterised RBP, which appears to be being regulated by IFN-I at the level of RNA-binding / microtubule-binding activity, through its phosphorylation state. Further research will be required to characterise its RNA-binding properties and its role in infection.

3.3.3 GIGYF2 S30: modulating the 4EHP interface

GIGYF2 is a translational repressor that has both 4EHP-dependent and independent activity [137, 139]. 4EHP is a homolog of eIF4E, which binds to the 5' cap of mRNA but lacks

eIF4E's capacity for interaction with eIF4G. Instead, it interacts with GIGYF2, which forms subsequent interactions with other RBPs, including the miRNA repression machinery, to mediate sequence-specific repression [137, 319, 320]. Repression can constitute translation inhibition but can also involve mRNA degradation, allowing 4EHP/GIGYF2 to act as regulators of mRNA quality control [138, 321].

One important instance of specific 4EHP/GIGYF2-mediated inhibition is in the translation of IFNβ. Zhang and colleagues found that 4EHP was able to suppress IFNβ translation in a miR-34a-dependent manner [322]. While they did not assess the role of GIGYF2 in this paper, it has since been found that SARS-CoV2's nsP2 promotes IFNβ repression through its direct interaction with GIGYF2 [266]. This repression was 4EHP-dependent and nsP2 was found to directly promote the GIGYF2/4EHP interaction. Subsequent work has shown that nsP2 co-opts GIGYF2/4EHP to promote repression driven by a range of miRNAs [323]. It will be interesting to know how GIGYF2 would alter proteome composition following SARS-CoV2 infection.

Given this role for GIGYF2 in complex with 4EHP in innate immune suppression, it is interesting to note that the S30 phosphosite is positioned immediately before the GIGYF2/4EHP interaction domain. The crystal structure of how GIGYF2 and 4EHP interact does not extend to S30 [297]. However, the electrostatic map of the likely interaction interface between S30 and 4EHP indicates that the negative charge of the phosphate group could disrupt the interaction of a positively charged region on GIGYF2 with a negatively charged region of 4EHP. As with MAP4, S30 is one of several phosphosites concentrated within this region. Again, this indicates that this region represents an important regulatory interface. Combined with RNA-binding predictions suggesting that S30 does not sit within an RNA-binding region, this suggests that S30 has an indirect effect on RNA-binding through modulating 4EHP interaction. This would need to be tested experimentally. It would be important to establish whether IFN-I treatment or an S30 phosphomimetic did indeed reduce 4EHP interaction. By comparing the effect on RNA-binding activity in WT versus 4EHP knockout cells, I could also assess whether a loss of 4EHP interaction would be likely to result in the observed reduction in RNA binding.

If S30 does indeed reduce interaction with 4EHP, this could be a mechanism of selectively modulating one aspect of GIGYF2's function, while leaving other translation inhibition functions intact. Not only would this amplify the IFN-I response by de-repressing IFNβ translation but it might also reduce availability of the 4EHP/GIGYF2 system to viral hijacking mechanisms. The

nsP2-GIGYF2 interaction is conserved across a range of coronaviruses and other viruses are known to upregulate miRNAs, including MiR34a [324]. Therefore, co-opting this translation repression machinery could represent a conserved innate immune evasion strategy that is partially offset through reduced GIGYF2/4EHP interaction.

While it might benefit the cell to reduce activity of this system in the early stages of the IFN-I response, it is important that it is active during the resolution phase of the response to prevent chronic inflammation [322]. Therefore, it would also be interesting to know how S30 phosphorylation and RNA-binding activity behaved at timepoints later than 20 hours post IFN-I treatment.

3.3.4 MATR3 S188: learning new tricks from an old phosphosite

MATR3 is a highly disordered and extensively phosphorylated protein. Phosphorylation has been identified as a key regulator of MATR3's nuclear/cytoplasmic localisation. The nuclear form of MATR3 is more heavily phosphorylated than the cytoplasmic form [325] and treating cells with a broad-spectrum kinase inhibitor results in its relocalisation to the cytoplasm [326]. In addition to overall phosphorylation, individual phosphosites have been identified as important promoters of nuclear retention. For example, S208 is phosphorylated by ATM following DNA damage, resulting in retention of MATR3 in the nucleus, where it plays a role in the DNA damage response [327]. Additionally, S150 phosphorylation by alpha herpesviruses results in nuclear retention, although the consequences of this for viral infection have not been determined [328].

There have also been some reports that phosphorylation can alter MATR3's nucleic acid binding, although evidence of a direct relationship is lacking. One study reported that phosphorylation was important for DNA-binding activity but this is more likely to be a reflection of increased nuclear retention [329]. In another study, PKA phosphorylation was proposed to increase MATR3 binding to pY RNA1-s2 in the retina [330]. In this study, they assessed the amount of MATR3 interacting with pY RNA1-s2-coupled streptavidin beads in retina cell lysates. While they found that adding a phosphatase inhibitor increased MATR3 pull-down, they were not able to detect PKA-specific MATR3 phosphorylation in untreated or treated lysates. When exogenous PKA was added to lysates, on the other hand, they were able to detect phosphorylation but did not see an increase in MATR3 pull-down. This suggests that the effect of the phosphatase inhibitor is indirect and that PKA-driven phosphorylation is not

responsible for increased MATR3 binding to pY RNA1-s2.

In this study, I found that MATR3 phosphorylation at S188 is anti-correlated with RNA binding, suggesting that this phosphorylation event is a negative regulator of RNA-binding activity. S188 phosphorylation was previously suggested to promote degradation of MATR3 [300]. This is surprising because neither in the current study nor in a previous study expressing a MATR3 S188A mutant is there any evidence that phosphorylation reduces the protein's stability [328]. The observation that S188 phosphorylation resulted in degradation was based on an analysis of NMDA-induced PKA activation in cerebellar neurons. While, inhibition of PKA was found to prevent both S188 phosphorylation and MATR3 degradation, a link between these two processes was not demonstrated. PKA has thousands of targets, making it possible that phosphorylation of another target could indirectly result in degradation [331]. It is also possible that, in this cellular context, other factors downstream of the activated NMDA receptor contribute to a destabilising effect on MATR3.

To experimentally determine whether this phosphosite was regulating RNA-binding activity, I used a dual-fluorescence RNA-binding assay [299]. This approach indicated that the phospho-null mutant had increased RNA-binding activity relative to WT MATR3. This is in keeping with my hypothesis that S188 dephosphorylation following IFN-I treatment increases MATR3's RNA-binding activity. The phosphomimetic mutant, on the other hand, did not result in the decreased RNA-binding activity that we might have expected. This could be because phosphomimetic mutants do not exactly recapitulate the biological properties of a phosphorylation event. Aspartic acid only has a single negative charge, rather than a 2- charge, and has a much smaller hydration shell than a phosphate group. This means that its effect will be weaker than phosphorylation.

There are two possibilities of how phosphorylation might be altering MATR3's RNA-binding activity. It might be regulating its affinity for existing RNA targets or it might be resulting in a shift in target affinity, which has the secondary effect of changing the level of RNA-binding. Given what we know about MATR3's RNA-binding properties, the first of these possibilities is more likely to be the case. MATR3 possesses two RRMs, which are known to be the primary drivers of its target specificity; the motif recognised by MATR3's RRMs *in vitro* matches the whole protein's *in vivo* specificity, based on iCLIP [161, 332], although it is possible that MATR3 could have secondary specificities, akin to what was observed with different disease-associated mutations in TDP-43 [333]. Another possibility is that the disordered region within which S188

sits is involved in anchoring the protein to RNA following RRM binding. A model for how this might be operating is presented by Malki and colleagues in their study of Sam68's RNA-binding activity [334]. Like MATR3, Sam68 is a splicing regulator composed of a central canonical RBD flanked by lower affinity, flexible regions that contribute to RNA binding. Malki et al found that two threonine phosphorylation sites in these flanking regions were regulated over the course of the cell cycle. Reduced RNA-binding as a result of increased phosphorylation at these sites was found to alter the splicing profile of the protein, possibly as a result of Sam68's reduced ability to out-compete other splicing factors. The parallels between the cases of Sam68 and MATR3 make this an intriguing hypothesis to follow up on. iCLIP data presented in the next chapter will provide further evidence that the effect is driven by changes in MATR3 affinity, as opposed to specificity.

The dual-fluorescence assay has the important benefit of allowing us to determine the functional consequences of differential phosphorylation in a biologically relevant context. However, studying the effect of phosphorylation within this biological context makes it hard to isolate the exact function of S188 phosphorylation. We cannot rule out the possibility that the effect on RNA-binding activity is an indirect one without assessing its effect on purified MATR3, using *in vitro* binding assays. However, this presents substantial technical barriers. Full-length MATR3 is extremely difficult to purify [335]. While it is possible to specifically purify its RRMs, the S188 site occurs in a distal disordered region. Since S188 phosphorylation is likely modulating RRM-driven RNA-binding activity, it would be important to test the effect of S188 phosphorylation in a 384 amino acid construct that also contained the RRMs, which would likely be similarly challenging to purify. Establishing whether it would be possible to generate a hybrid construct, consisting of the RRMs and the disordered region containing S188 joined by a flexible linker would be an important next step.

Another step in understanding the consequences of MATR3 S188 phosphorylation in the IFN-I response will be to determine how it affects MATR3's antiviral activity. This is not something that I was able to assess using the GFP-tagged overexpression lines. Since the lines overexpress to different levels, we would only be able to compare their antiviral activity if, as with the dual fluorescence RNA-binding assay, we were able to normalise to GFP levels. However, such normalisation would only be valid if antiviral activity was directly proportional to protein quantity, which is highly unlikely. An alternative to using GFP overexpression lines would be to perform site-directed mutagenesis to directly modify the endogenous protein. This

would reduce the risk of toxicity associated with overexpression and would also ensure that the only MATR3-driven effect we were seeing would be driven by the mutant.

3.3.5 What is responsible for IFN-I-driven changes in phosphorylation?

Since MATR3 S188 is a known target of PKA, I was interested to see whether the change we observe in phosphorylation state is a direct result of altered PKA activity. While there is no direct evidence of this in the literature, there is extensive evidence that PKA promotes viral infection. Viruses including HCV, ZIKV, AD5, EBOV, and HIV-1 all show increased infection upon PKA inhibition [336–339]. Additionally, PKA is known to negatively regulate the innate immune response by phosphorylating key signaling molecules, including IRFs [340–342]. Levels of cAMP, which is the second messenger responsible for PKA activation, are also decreased following innate immune activation in mouse bone marrow-derived macrophages [343]. Together, these results would point to PKA as a negative regulator of innate immune signalling whose suppression might play a role in promoting the IFN-I response.

Despite these indications, I did not observe any change in PKA phosphosubstrate abundance following IFN-I treatment, even at high concentrations. However, this does not rule out that PKA is regulated. Western blot of PKA substrates is only a rough proxy for PKA activity; less than 20 out of hundreds of PKA substrates are visible by blotting, meaning this is only a fraction of the picture. Had resources allowed, I would have preferred to use antibodies directly against PKA and PKA's activated form to determine whether activation is directly regulated.

PKA is also highly regulated, not only in a given cell, but in given compartments of a cell. Therefore, it is possible that phosphorylation of a given set of substrates is modulated without an overall change in activity. These local effects are driven by AKAPs, which scaffold PKA, its targets, and components of the PKA regulatory machinery that keep activation and phosphorylation levels under tight control [344]. AKAP8 (also known as AKAP95) is a nuclear protein thought to act as a scaffold for a PKA nuclear microdomain [345]. Since AKAP8 is regulated by the IFN-I response and interacts with MATR3, I hypothesised that it might be mediating the altered phosphorylation of MATR3. I did not see any effect on MATR3 phosphorylation following knockdown of AKAP8 but I did see a strong increase in SINV infection. This might suggest that AKAP8 is having another effect in infection. Possibly, through its interaction with PKA, or through one of its other scaffolding functions in splicing or transcription, for example [346, 347].

Another possibility is that MATR3's dephosphorylation is a result of increased phosphatase activity, rather than reduced kinase activity. As with PKA, phosphatases are heavily regulated through interactions with phosphatase-associated proteins. Indeed, Protein Phosphatase 1 (PP1), which is responsible for the bulk of dephosphorylation events in the cell, is thought to form as many as 650 distinct complexes, allowing it to precisely regulate levels of phosphorylation, despite its inherent promiscuity [348]. One of its interaction partners is PPP1R10, or PNUTS. This protein has increased RNA-binding activity following IFN-I treatment and was also identified by Dr Kamel, another member of the Castello group, as a protein that shuttles from the nucleus to the cytoplasm and associates with vRNA upon SINV infection (Kamel et al, under review). PNUTS is a predominantly nuclear protein that targets PP1 to specific substrates, often in association with RNA due to its inherent RNA-binding activity [349]. The PNUTS-PP1 complex has been found to inhibit KSHV gene expression and replication [350]. I also see that PNUTS has an antiviral effect in SINV, although this is much milder than AKAP8. As with AKAP8, I did not see any effect on MATR3 phosphorylation levels. As discussed for the siRNA screen described in Chapter 2, lack of an effect could be a reflection of poor knockdown efficiency so knockdown will need to be confirmed for both AKAP8 and PPP1R10 before entirely ruling out these proteins' involvement in MATR3 regulation.

Although I was not able to identify the mechanism behind MATR3's differential phosphorylation, it is still an exciting instance of IFN-driven activity-level regulation. A more in depth phosphoproteomic analysis, as well as follow-up on other differentially phosphorylated proteins will undoubtedly yield more instances of such regulation.

4 Establishing the interactome of MATR3 in virus infected cells

4.1 Introduction

MATR3 is a highly conserved 125kDa protein that was initially identified as a component of the nuclear matrix [351]. As illustrated in the previous chapter, MATR3 is a highly disordered protein, containing two RRMs, which specifically bind RNA, and two C2H2 zinc finger domains, which are thought to primarily mediate DNA binding but can also bind to RNA [152, 325]. This protein architecture allows MATR3 to interact with RNA, DNA, and, through its IDRs, a wide range of proteins. These wide-ranging interactions, in turn, allow MATR3 to engage in diverse cellular processes, many of which rely on its capacity to act as a bridge between proteins and nucleic acids [352].

In the nucleus, MATR3 contributes to the establishment of 'functional neighbourhoods' through its association with chromatin [353]. Additionally, its interaction with non-coding, regulatory regions of DNA, such as enhancers, have suggested a role for MATR3 in transcription regulation [354]. For example, in the DNA damage response, MATR3 regulates the induction of several P53 targets through its association with a lncRNA, PINCR, as well as P53 [355]. As well as regulating transcription, MATR3 has been linked to practically every subsequent stage of an RNA's lifecycle. It regulates RNA alternative splicing, nuclear export, translation, and stability.

Splicing regulation by MATR3 is widespread. MATR3 knockdown leads to hundreds of changes in AS events, many of which occur at sites directly bound by MATR3, based on CLIP data [332]. RNA immunoprecipitation with sequencing (RIP-seq) and iCLIP have both indicated that MATR3 predominantly binds to pyrimidine-rich sequences within introns [332, 335, 356]. While many splicing regulators exhibit sharp peaks in binding at a defined distance from the splice site, MATR3 instead displays distributed binding surrounding a pyrimidine-rich motif. It has been suggested that this reflects a 'bind and spread' model, in which an initial specific binding event is followed by subsequent, lower affinity interactions that coat the RNA and obscure potential splice sites [332]. This occurs at multivalent binding sites within LINE elements, where it binds alongside PTBP1 to prevent recognition of cryptic splicing or polyadenylation sites [356]. While PTBP1 often binds concurrently with MATR3, only 18% of regulated splicing events are PTBP1-dependent, indicating that MATR3 is able to act both in conjunction with and independently of this interaction partner [332, 335]. MATR3 can also

regulate splicing as part of other complexes. For example, MATR3 is a component of the LASR complex, which consists of several RBPs and is present at almost every site of RBFOX-driven splicing regulation [357].

In addition to observing altered patterns of splicing following MATR3 knockdown, several studies have identified changes in RNA stability [358]. These changes are generally more subtle than changes in splicing, occurring in a relatively small number of transcripts [332, 335]. A portion of these instances of stability regulation can also be attributed to AS as MATR3 can prevent inclusion of premature stop codons that lead to nonsense-mediated decay [332]. However, there are validated instances of stabilisation independently of this process. For example, MATR3 was found to bind to and stabilise the mRNAs of puripotency factor LIN28A and NANOG [359]. Additionally, large-scale 3' UTR tethering assays have indicated that it has a stabilising effect on a reporter gene [104].

Interestingly, one set of RNAs that are stabilised by MATR3 are those derived from HIV-1. In 2011, two parallel studies demonstrated that MATR3 binds to HIV-1 RNA, promoting Rev-dependent nuclear export and stabilisation of unspliced RNA [360, 361]. It has since been shown that knockdown of MATR3 substantially increases ZAP's capacity to restrict HIV-1 and, to a lesser extent, MLV (Moloney murine leukemia virus). This increased restriction was found to be primarily driven by a ZAP-dependent reduction in RNA stability [362].

In contrast to this apparently pro-viral function in retrovirus infection, MATR3 has been linked to antiviral activity in RNA and DNA virus infection. Morchikh et al identified MATR3 in association with HEXIM-1, which forms a complex required for DNA sensing through interaction with NEAT1 [363]. However, this study did not validate that MATR3 is present in the complex involved in DNA sensing. As presented in Chapter 4, siRNA knockdown data indicates that MATR3 is antiviral in SINV infection. This has been further corroborated by Dr Wael Kamel, a member of the Castello lab, who has shown, in unpublished data, the same effect using different siRNAs in SINV infection by Western blot, RT-qPCR and plaque assay. Additionally, he observed antiviral effects for Semliki Forest Virus (SFV), which is another alphavirus, and the picornavirus, Coxsackievirus B3 (CVB3). Dr Kamel identified MATR3 as one of a group of nuclear proteins that relocalise to the cytoplasm following infection [364]. Interestingly, MATR3 is in the vRNA interactomes of a range of cytoplasmic RNA viruses, including SINV, suggesting that cytoplasmic MATR3 might be playing a broader role in infection [192].

While MATR3 is generally considered a nuclear protein, a fraction of the protein localises

to the cytoplasm, even under resting conditions [325]. This was first reported by Hibino and colleagues in 2006, who not only found that MATR3 localised in the cytoplasm, but associated with polysomes [325]. This hinted at a role in translation that has since been confirmed by other studies. Rajgor et al identified a truncated form of MATR3 in association with P-bodies and found that MATR3 was important for miRISC-mediated silencing [365]. Conversely, Pollini et al found that a cytoplasmic pool of MATR3 in iPSCs promoted translation of certain pluripotency factors [359]. The extent and significance of MATR3's roles in the cytoplasm and specifically in translation control is likely context dependent and requires further investigation.

To date, the study of MATR3's protein and nucleic acid interaction landscapes have been critical in defining its cellular functions. However, this diversity of functions and interaction partners makes it challenging to predict how MATR3 might be restricting infection of cytoplasmic RNA viruses. In this chapter, I define the SINV-specific RNA and protein interactomes of MATR3. This is an important step in establishing which of MATR3's many functions is responsible for its antiviral activity.

4.2 Results

4.2.1 MATR3 alters its protein interaction network upon infection

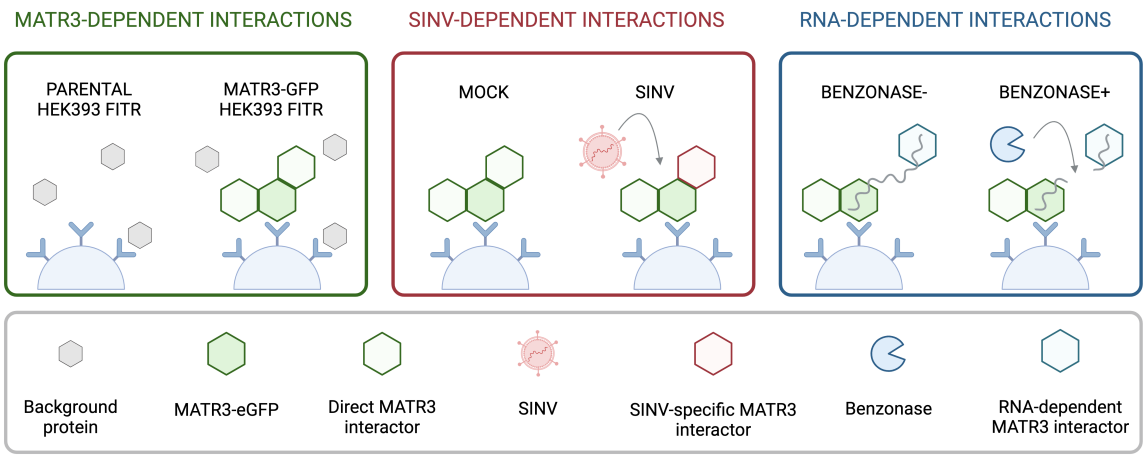


Figure 4.1: **Schematic of comparisons in MATR3 PPI analysis.**

MATR3 is a highly disordered protein which establishes many contacts to carry out diverse functions. Which of these functions and, by extension, which interaction partners might be important for MATR3's antiviral activity is an open question. To address this, I performed a GFP-trap IP of MATR3-GFP and its protein interaction partners. In this experiment, I wanted to assess protein enrichment at three different levels (Outlined in Figure 4.1). Firstly, I wanted to focus on proteins specifically interacting with MATR3-GFP, excluding non-specific interactions with agarose beads or nanobodies. For this, I included a HEK293 FITR parental control, in which binding to GFP-trap agarose will consist entirely of non-specific interactions. Secondly, I wanted to identify changes in MATR3's interaction partners following infection. To this end, I included both mock samples and samples infected with SINV for 9 hours. MATR3 is predominantly nuclear but SINV is a cytoplasmic RNA virus and MATR3 is known to migrate to the cytoplasm upon infection. Therefore, I chose to reduce the contribution of nuclear MATR3 to the IP signal by using a mild lysis buffer sedimenting the nuclei using a low-speed spin prior to IP. Thirdly, I wanted differentiate between RNA-dependent and RNA-independent interactions. After IP and an initial wash step, I split the samples in two and treated half with the nuclease, benzonase. Further wash steps result in RNA-dependent interactors being lost in the benzonase treated sample.

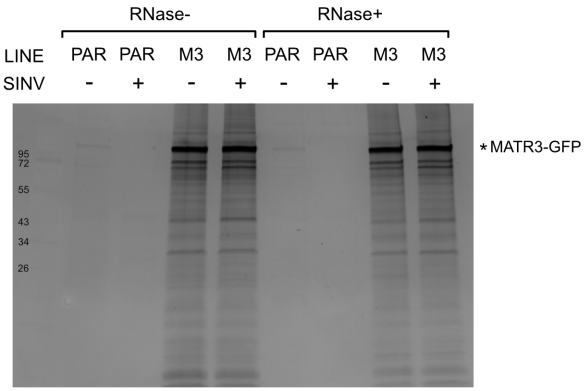


Figure 4.2: **Silver stain of MATR3 IP.** Mock and SINV infected HEK293 FITR parental cells or cells overexpressing MATR3-GFP were harvested 9hpi. Immunoprecipitation was performed using GFP-Trap agarose. After capture and an initial wash, samples were divided in two and either treated with benzonase (RNase+) or left untreated (RNase-). Following further washes, proteins were eluted in 1% SDS. Silver staining was performed on 10% of eluates. Band corresponding to the size of MATR3-GFP is marked with a *.

To assess the quality of the IP, I performed silver staining. This provides several indications of IP quality. Firstly, MATR3-GFP is the most enriched protein across all conditions. Secondly, the pattern of bands across all MATR3 IP samples look highly similar. This indicates that the strongest interaction partners of MATR3 are RNA-independent and are preserved upon infection. Additionally, the lanes with parental samples show considerably lower signal than IP samples, indicating that the IP is clean of non-MATR3-specific interactors. Together, this indicates that the IP resulted in selective capture of MATR3-GFP and its interaction partners.

Eluates were analysed by LC-MS/MS to determine MATR3's interaction network. Peptides and proteins were identified and quantified using MaxQuant [228]. Protein intensity distributions concur with results of silver staining; overall intensities were consistent across IP samples and reduced in parental control samples (Figure 4.3A). PCA further demonstrates that the majority of variation in the dataset is explained by the differences between MATR3-GFP and parental lines (Figure 4.3B). When the parental samples were removed, RNase+/- samples and mock/infected samples can be clearly differentiated (Figure 4.3C). This reflects the high degree of reproducibility between replicates and the presence of biological differences as a result of treatment condition.

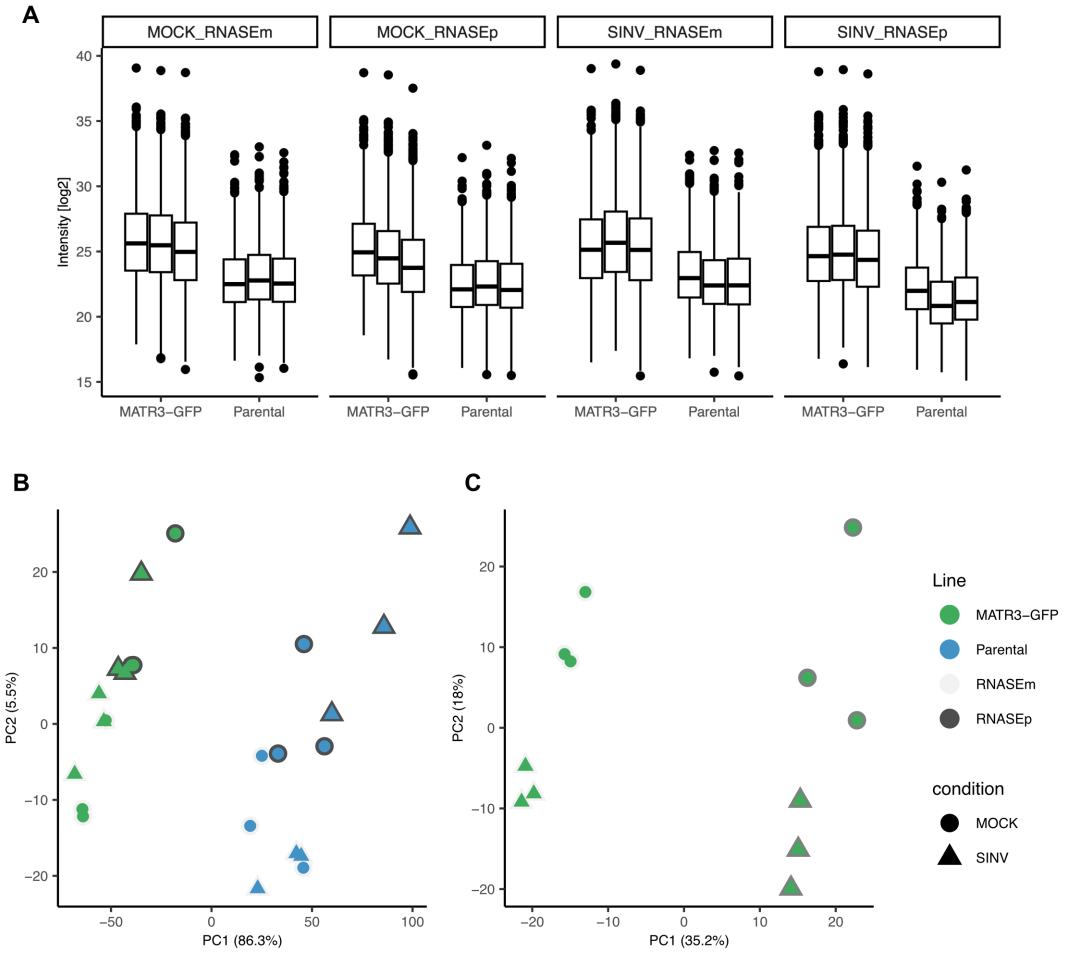


Figure 4.3: **Quality control of MATR3 IP LC-MS/MS data.** MATR3 IP samples were analysed by LC-MS/MS and peptides/proteins were identified and quantified in MaxQuant. (**A**) A boxplot showing the log2 intensity distribution of triplicate samples for each condition without normalisation. (**B**) & (**C**) PCA of MATR3 IP samples with and without the parental control included, respectively. In both cases, the first two principal components were plotted as a 2D scatter plot, with percentage of variance explained by each component given in brackets in the axes titles. MATR3-GFP samples are shown in green and parental control samples in blue. Triangles represent infected samples and uninfected samples are shown as circles. RNase treatment is indicated by the points' borders; RNase-treated samples have a grey border, while untreated samples do not.

An average of 2045 proteins were identified in each MATR3-GFP sample. Many of these might be true interactors but it is also possible that some are contaminants. To identify proteins specifically interacting with MATR3, I assessed enrichment over the parental control using the limma package in R (Figure 4.4). Using a stringent p-value cut-off (0.01% FDR), I identified an average of 898 proteins as significantly enriched above the parental control in each dataset. I defined proteins significantly enriched in at least one sample as 'MATR3 interactors' (a total of

1464 proteins). I filtered the total list of identified proteins to include only 'MATR3 interactors' for subsequent analyses.

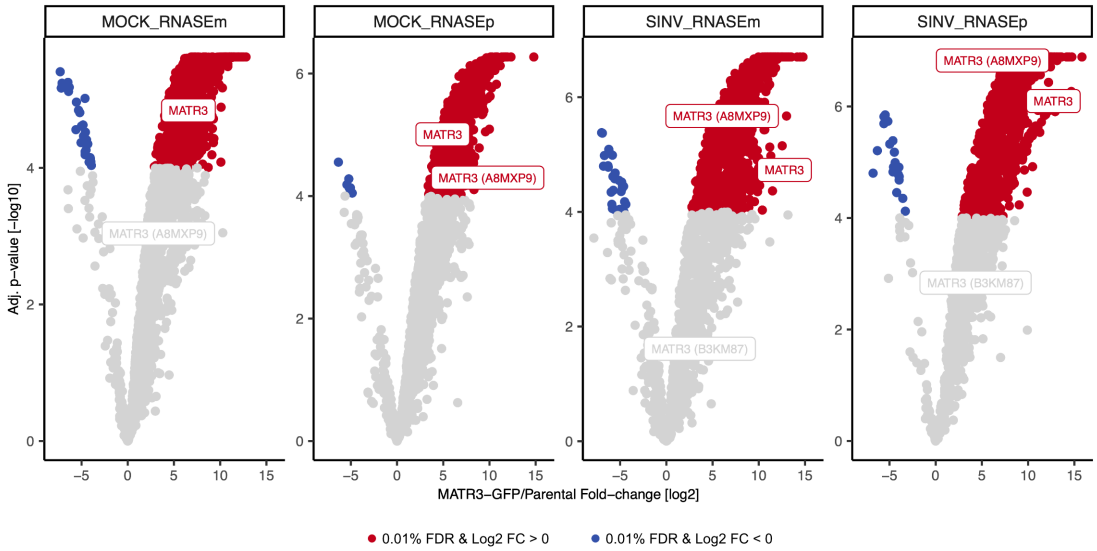


Figure 4.4: **Many proteins are enriched in the MATR3 IP over the parental control.** Volcano plots summarising the results of limma testing of each MATR3-GFP condition against its parental control. Log2 fold-change between the conditions is plotted against -log10 of the adjusted p-value. Proteins enriched in the MATR3-GFP sample over the parental control at an FDR of 0.01% are coloured red. Proteins enriched in parental control at an FDR of 0.01% are coloured dark blue. Proteins that do not meet this significance threshold are shown in grey. All proteoforms of MATR3 detected in the dataset are labelled.

MATR3 interacts with both RNA and protein. Therefore, I expected that a proportion of MATR3's interactions are bridged by RNA. To determine whether this was the case and identify RNA-dependent versus RNA-independent interaction partners, I performed limma testing comparing samples processed with or without benzonase treatment (Figure 4.5A). I classified proteins as 'RNA-dependent' if they had a significantly greater interaction with MATR3 in the presence of RNA. Any protein that didn't show a significant change or was enriched in the absence of RNA was classified as 'RNA-independent'. 67% and 50% of interactions were RNA-dependent in mock and infected conditions, respectively. To determine whether there were specific features of RNA-dependent/ independent interaction partners, I performed GO enrichment of the two groups of proteins (Figure 4.5B). Interestingly, some of the same GO terms appear in both sets of proteins, including 'non-membrane-bounded organelle assembly', 'ribonucleoprotein complex biogenesis', and 'RNA localisation', and both groups are enriched in GO terms related to RNA processing. This suggests that MATR3's RNA-dependent and -independent interaction

partners might represent parts of the same functional complexes, which are partially stabilised as a result of RNA-binding.

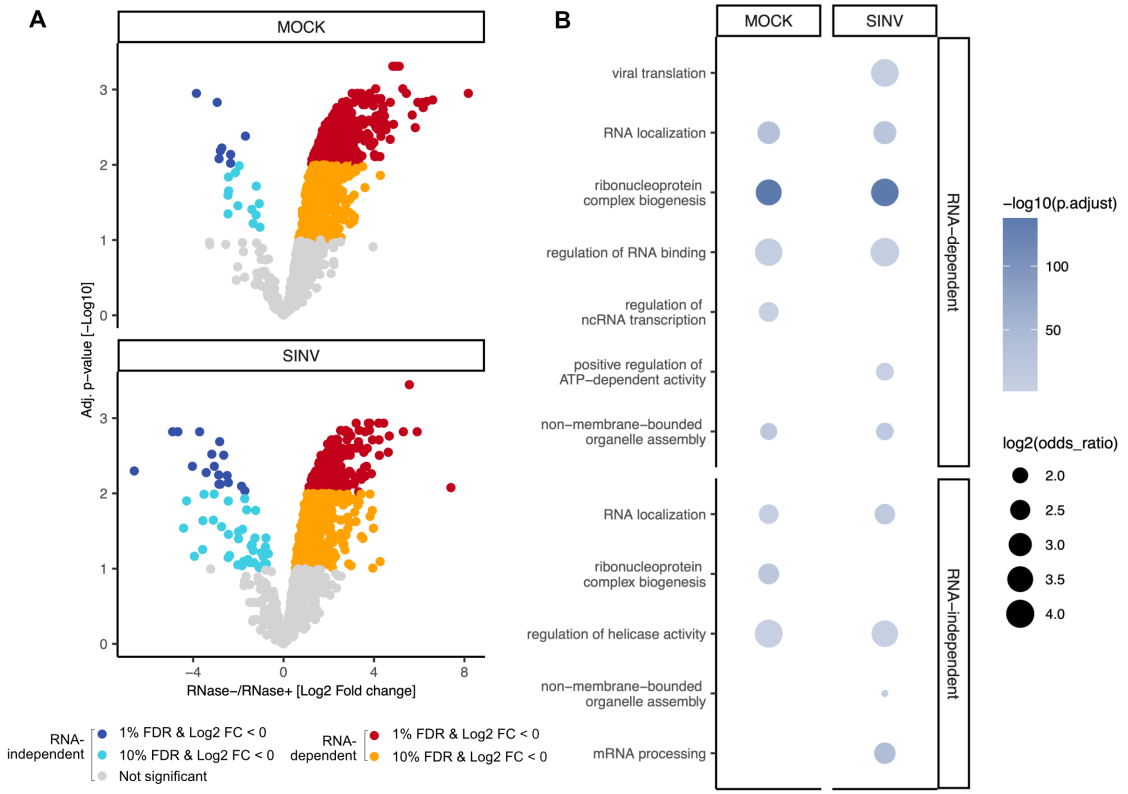


Figure 4.5: **MATR3 has both RNA-dependent and independent interactions.** (**A**) Volcano plots summarising the results of limma testing comparing MATR3-GFP mock and SINV samples in the presence and absence of benzonase. Testing was performed on a filtered list of proteins identified as significantly enriched (0.01% FDR) over the parental control in at least one condition. Log2 fold-change between the conditions is plotted against -log10 of the adjusted p-value. Proteins enriched in benzonase-untreated samples are coloured red (1% FDR) and orange (10% FDR). These are classified as RNA-dependent interactions. Proteins enriched in benzonase-treated samples are coloured dark blue (1% FDR) and light blue (10% FDR). Along with proteins coloured in grey, these are classified as RNA-independent interactions. (**B**) GO enrichment analysis of RNA-dependent and independent interactions in mock and infected samples. Enrichment analysis was performed with clusterProfiler and overlapping GO terms were removed, using a similarity threshold of 0.3. The size of each point reflects the log2 odds ratio of enrichment and the shade of purple reflects the significance of the enrichment, with darker purple indicating more significant enrichment.

The goal of this study was to establish whether MATR3's interaction network is altered by SINV infection. Quantitative analysis comparing MATR3's interaction partners in uninfected and infected cells indicates that there are changes. These changes are more pronounced in

benzonase-untreated cells, with 52 proteins significantly increasing and 68 proteins significantly decreasing at an FDR threshold of 10% (Figure 4.6A). In benzonase-treated samples, there were 27 proteins with increased and 34 proteins with decreased binding (Figure 4.6B). In both cases, MATR3 did not show a significant change, as would be expected, and the most significantly changing proteins were the viral proteins nsP1, nsP2, nsP3 and capsid. This is what we would expect, since these proteins are absent from uninfected cells and are highly abundant following infection.

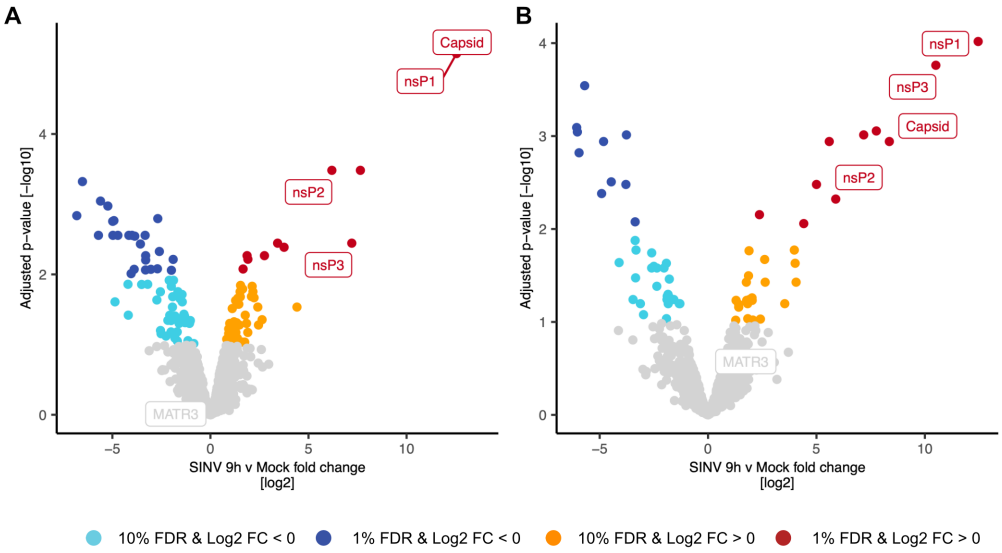


Figure 4.6: **MATR3 alters its interaction partners following infection.** Volcano plots summarising the results of limma testing for enrichment of MATR3-GFP interaction partners in mock and SINV samples in either benzonase treated (**A**) or untreated (**B**) samples. Testing was performed on a filtered list of proteins identified as significantly enriched (0.01% FDR) over the parental control in at least one condition. Log2 fold-change between the conditions is plotted against -log10 of the adjusted p-value. Proteins enriched in infected samples are coloured red (1% FDR) and orange (10% FDR). Proteins enriched in uninfected samples are coloured dark blue (1% FDR) and light blue (10% FDR). Proteins that do not show a significant change are colour grey. MATR3 and the viral proteins are labelled.

A concern with protein-protein interaction analysis, particularly with a dataset as large as the MATR3 interactome, is that it simply represents a diluted version of the whole cell proteome. To determine whether proteins with increased binding following infection are true interactors of MATR3 or are detected as a result of their high abundance in the cell, I compared protein intensities from the MATR3 IP analysis to whole cell proteome (WCP) intensities from a previous analysis of SINV infected HEK293 cells [50]. This indicates that the majority of altered

proteins are enriched in the MATR3 IP, relative to the WCP. Among the most highly enriched proteins is ZC3HAV1 (ZAP, PARP13), which is a known interaction partner of MATR3 and is important in the innate immune response against alphaviruses. AKAP8, which we encountered in the previous chapter, is also highly enriched in the MATR3 IP. In addition to proteins where we can calculate a fold-change between the two conditions, there are also many that are only present in the MATR3 IP and are below the detection threshold for the WCP. This is a strong indication that these are specifically pulled down with MATR3.

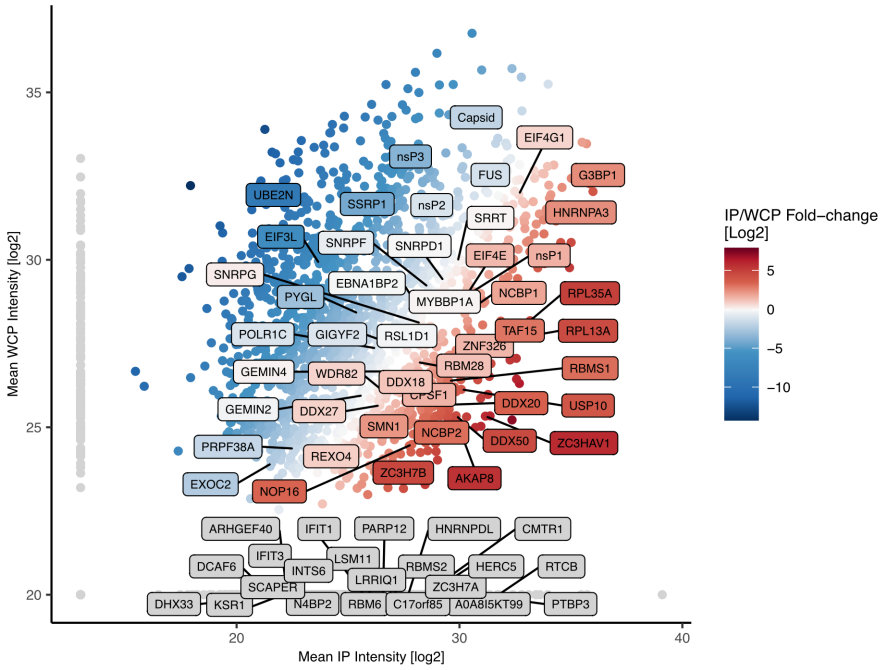


Figure 4.7: **Comparing MATR3 IP with the whole cell proteome.** Scatter plots comparing MATR3-GFP IP log2 intensities at SINV 9hpi without benzonase treatment to whole cell proteome log2 intensities 18hpi. Colour reflects log2 fold-change in MATR3 IP v WCP. Blue points are proteins with higher intensity in the whole cell proteome and red proteins are those with higher intensity in the MATR3 IP. Points in grey are proteins only present in one or other dataset. Proteins significantly enriched in MATR3 binding after SINV infection are labelled.

In addition to assessing enrichment of cellular proteins, we can also determine whether the level of binding to viral proteins is a reflection of true interaction. WCP samples were harvested 18hpi, whereas the MATR3 IP samples were harvested 9hpi. While differences in cellular proteins are minimal between the two timepoints, the exponential growth of the virus means that viral proteins will be more abundant in the WCP of this analysis than in the MATR3 IP inputs [50]. Despite this difference, nsP1 is still enriched in the MATR3 IP, relative to the WCP,

while the other viral proteins are not. This indicates that MATR3 forms a genuine interaction with nsP1, whereas the other viral proteins are more likely being picked up due to their high abundance or through indirect interaction via nsP1.

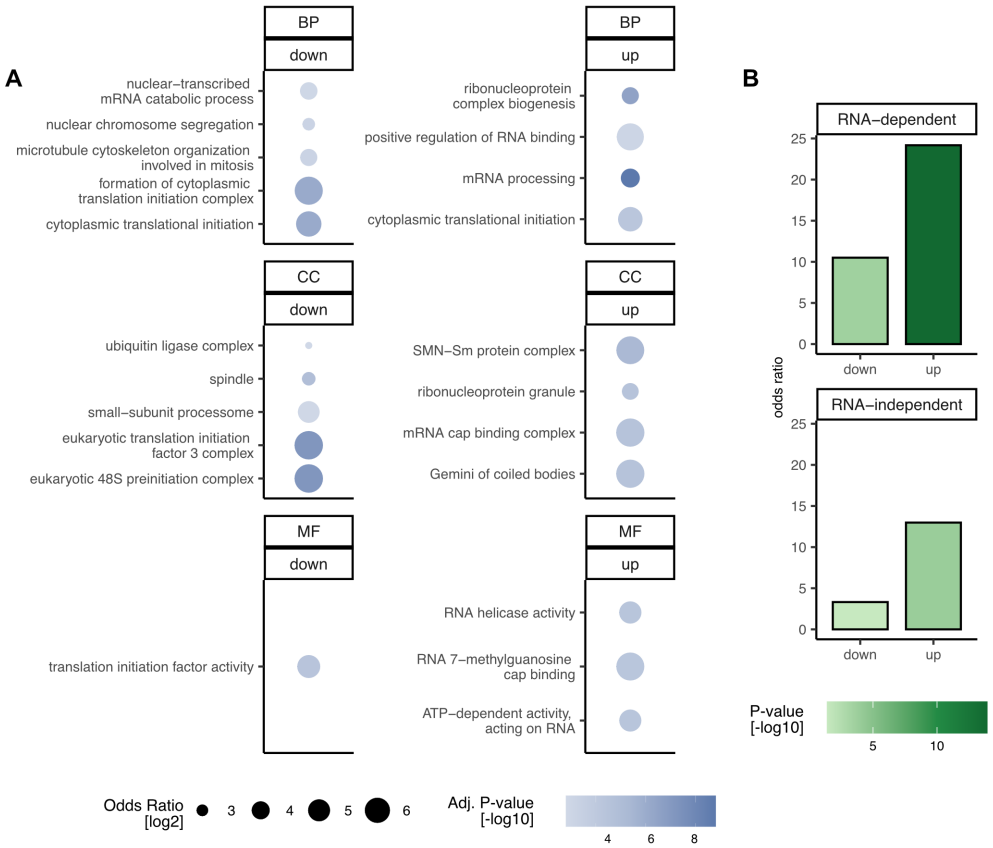


Figure 4.8: **MATR3's SINV-specific interaction partners are linked to infection.** (**A**) GO enrichment analysis of proteins enriched in the mock sample (down) and the SINV infected sample (up). Enriched proteins in benzonase+/- conditions were combined for this analysis. Enrichment analysis was performed with clusterProfiler and overlapping GO terms were removed, using a similarity threshold of 0.5. The size of each point reflects the log2 odds ratio of enrichment and the shade of blue reflects the significance of the enrichment, with darker blue indicating more significant enrichment. (**B**) Bar plots showing enrichment of MATR3 interactors in the vRNA interactome. Proteins were separated into RNA-dependent and RNA-independent groups, based on their classification defined in Figure 4.5 in infected cells. Bar height reflects the odds ratio of enrichment, relative to the whole proteome and shade of green reflects the significance of enrichment. Significance was tested using Fisher's Exact Test. Darker green reflects more significant enrichment.

Proteins that show altered MATR3 interaction following infection might speak to a functional shift in MATR3's activity that relates to its antiviral effect. To determine the nature of these proteins, I performed GO enrichment analysis on both upregulated and downregulated interaction

partners (Figure 4.8A). For this analysis, I combined unique significant hits from benzonase+ and benzonase- analyses. Upregulated interaction partners were highly enriched in members of the SMN-Sm protein complex, and proteins involved in mRNA capping and cap binding (Figure 4.9). Proteins involved in cap binding include both members of the canonical nuclear cap binding complex (NCBP1 and NCBP2), the non-canonical nuclear cap binding protein, NCBP3, and the cap methyltransferase, CMTR1. Additionally, two members of the eIF4F complex are present: the cap binding protein, eIF4E, and eIF4G1, which is important for RNA translation initiation. There is also a notable group of antiviral proteins, including IFIT1, IFIT3, HERC5, PARP12, and ZC3HAV1 (ZAP) (Figure 4.9A).

Among the most enriched GO terms in downregulated interaction partners were those related to the microtubule cytoskeleton and formation of the mitotic spindle. Indeed, STRING network analysis indicates that a large proportion of downregulated interactors are linked to the microtubule cytoskeleton (Figure 4.9B). Another prominent group of proteins enriched among downregulated interactors are those involved in translation initiation and regulation. This includes five members of eIF3 complex, and the antiviral kinase EIF2AK2 (or PKR). While the majority of the microtubule cytoskeleton-associated interactions are RNA-independent, association with the eIF3 complex is bridged by RNA.

In addition to assessing cellular functions of MATR3's altered interaction partners, I wanted to assess whether any of these proteins were linked with SINV infection. To this end, I determined whether up or down-regulated interaction partners were enriched in members of the SINV vRNA interactome [364] (Figure 4.8B). For this analysis, I separated proteins into RNA-dependent/-independent groups based on their classification in SINV infected cells (Figure 4.5A). All groups of proteins were significantly enriched in members of the vRNA interactome, relative to the proteome as a whole. This is likely a reflection of the fact that many of MATR3's interaction partners are RBPs. Notably, the greatest enrichment (25-fold) was in RNA-dependent interactors with increased interaction following SINV infection. This suggests that SINV RNA is responsible for bridging many of the infection-specific RNA-dependent interactions of MATR3.

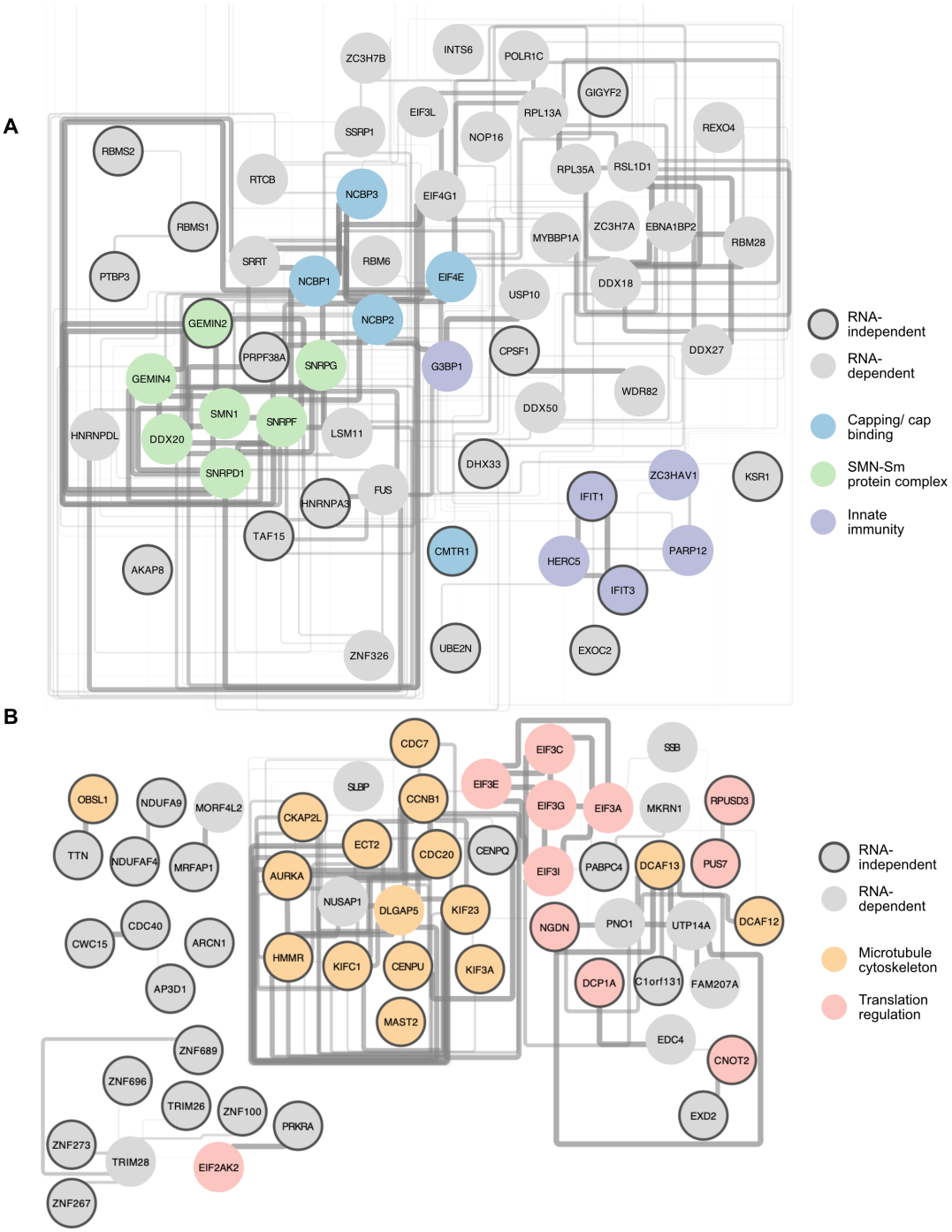


Figure 4.9: **STRING networks of SINV-modified MATR3 interactions.** Networks of upregulated (**A**) and downregulated (**B**) MATR3 interactions following infection. The width of the connecting line reflects the interaction score, with wider lines reflecting stronger interactions. Proteins outlined in black constitute RNA-independent interactions, based on their classification defined in Figure 4.5 in infected cells. Proteins were coloured based on GO annotation. Network analysis was performed using STRING and results were plotted with Cytoscape.

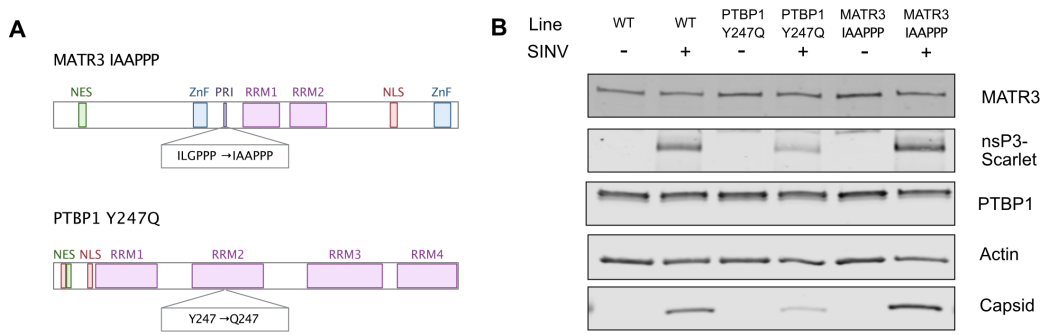


Figure 4.10: **Disruption of MATR3/ PTBP1's interaction site alters SINV infection.** (**A**) Schematics of MATR3 and PTBP1 uniprot domain structure. The position and nature of the amino acid substitutions used to abolish PTBP1-MATR3 interaction are highlighted. (**B**) Western blotting showing the effect of the PTBP1 Y247Q mutation and the MATR3 IAAPPP mutation in HEK293T cells. WT and mutant HEK293T lines were infected with $SINV_{nsp3-Scarlet}$ at an MOI of 0.1 and harvested 18hpi. Blotting was performed against the viral proteins nsP3-Scarlet and capsid, as well as MATR3, PTBP1 and β-Actin as a loading control. Representative image of results from 3 independent replicates.

A next step in establishing the relative importance of MATR3's interaction partners for its antiviral function is to directly disrupt specific protein interfaces. MATR3 is rich in short linear motifs (SLiMs), which are sequences that occur in disordered regions and mediate protein-protein interactions [366]. However, the interaction interfaces of most of MATR3's binding partners are not known. One SLiM that has been well characterised in MATR3 is the PRI motif, which is responsible for MATR3's interaction with PTBPs 1-3 [367]. I found that MATR3's interaction with PTBP1 and PTBP2 is unchanged upon infection but, interestingly, its interaction with PTBP3 is increased. Given the importance of the MATR3-PTBP1 interaction and the altered binding of PTBP3, I wanted to determine whether disruption of the PRI domain had any effect on infection. To this end, I made use of HEK293T lines in which the interaction interface in either MATR3 or PTBP1 has been mutated (a kind gift from the lab of Chris Smith). This involves mutation of the PRI motif in MATR3 (IAAPPP) and a residue in RRM2 for PTBP1 (Y247Q) (Figure 4.10A). Western blot results from three independent replicates indicate that infection is reduced in PTBP1 Y247Q cells, while it is enhanced in the MATR3 IAAPPP line (Figure 4.10B). Protein abundance of MATR3 and PTBP1 was not altered by mutations (Figure 4.10B). This suggests that MATR3's PRI domain is contributing to its antiviral activity but not through interaction with PTBP1. An equivalent mutant line has not been generated for PTBP3 but a next step will be to establish if this protein is contributing to MATR3's antiviral role.

4.2.2 Using iCLIP to define MATR3's RNA interaction partners during infection

The majority of MATR3's functions rely on its RNA-binding activity and I found that its RNA-binding activity is regulated as part of the IFN-I response. This led me to hypothesise that MATR3's antiviral activity might involve either its interaction with vRNA, or altered interaction with cellular RNA targets. To determine whether MATR3 exhibits a specific pattern of binding to vRNA or interacts differently with cellular RNA following infection, I employed iCLIP2. This method allows for the identification of RNA-binding sites with single nucleotide resolution [163] (Figure 4.11). It relies on stabilisation of protein-RNA interactions by UV crosslinking, followed by partial RNase digestion. The protein of interest is then immunoprecipitated under stringent wash conditions. Protein-RNA complexes of the correct size are separated and extracted by SDS-PAGE gel electrophoresis and protein is digested, leaving RNA fragments with a small protein adduct at the site of crosslinking. During subsequent reverse transcription, this adduct causes termination in ~80-90% of cases [368]. This termination means that the 5' end of the cDNA produced will, in a majority of cases, represent the crosslinking site. These cDNAs are then processed and sequenced to identify MATR3's binding sites .

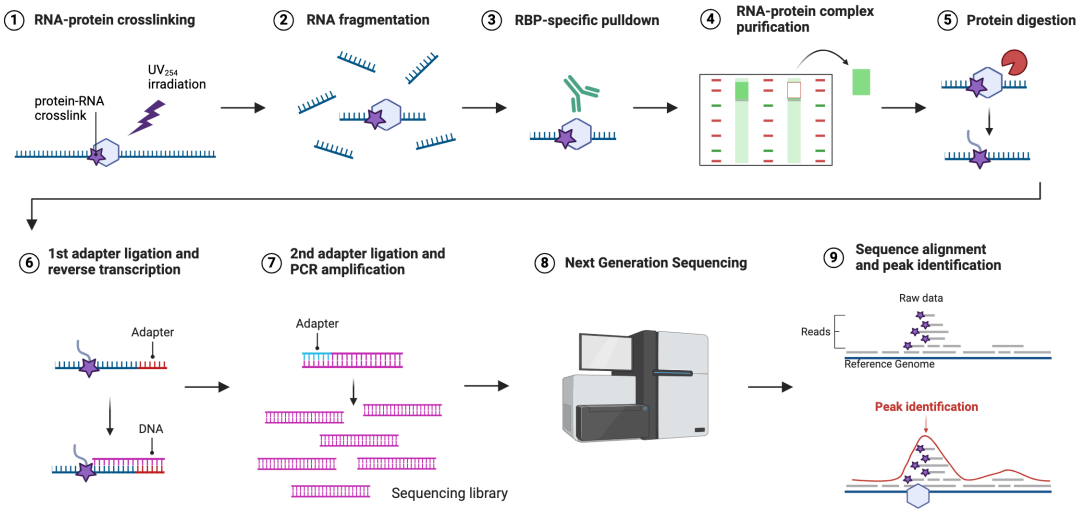


Figure 4.11: **Schematic of the iCLIP protocol.**

To identify how MATR3's RNA targets change following infection, I performed iCLIP2 on mock and SINV-infected samples in MATR3-GFP HEK293 FITR cells. SINV-infected cells were harvested after 9 hours of infection, when growth is in the exponential phase but before the cells are completely overwhelmed with viral RNA. As a control, I harvested HEK293 FITR

parental cells under the same conditions for use as a size-matched input (SMI). SMI samples are processed through a parallel pipeline, in which no protein is immunoprecipitated but the same region of membrane is extracted, providing a background signal in the specific region of interest. Usually, the SMI for a given sample is taken directly from the input used for IP. However, when an overexpression line is used, the overexpressed protein can represent a large fraction of the protein purified, even in the SMI, resulting in high levels of 'signal' in the control. Therefore, in this experiment, I used HEK293 FITR parental cells harvested at the same time under the same conditions to circumvent this issue.

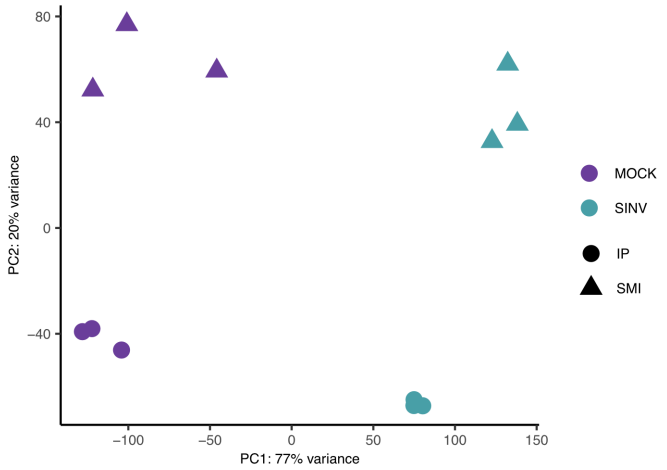


Figure 4.12: **Principal Component Analysis of MATR3 WT iCLIP samples.** The first two principal components were plotted as a 2D scatter plot, with percentage of variance explained by each component given in brackets in the axes titles. Mock samples are shown in purple and SINV infected samples in teal. Triangles represent SMI samples and IP samples are shown as circles.

Using an in-house analysis pipeline, I aligned sequencing data to a combined human and SINV genome annotation and defined 'crosslink sites' as the first nucleotide before the start of each read. I then identified regions with significantly enriched crosslink frequency over the SMI using the ht-seqCLIP/ DEWseq analysis pipeline [369]. Using the 1000 most variable sites across all samples, I performed PCA to assess how the samples compared to each other (Figure 4.12). PCA demonstrates that there are pronounced differences between the IP and SMI samples, as well as between mock and infected conditions. Replicates of the same condition cluster together closely, reflecting a low degree of technical variation.

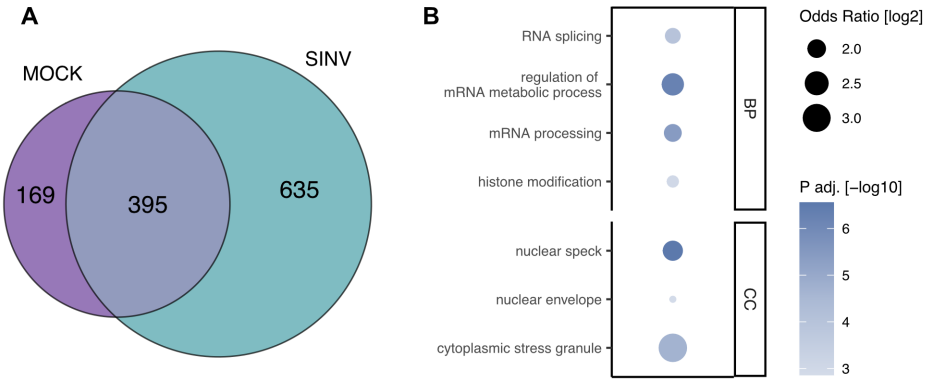


Figure 4.13: **MATR3 has a core set of targets in mock and infected cells.** (**A**) A Venn diagram showing the overlap of MATR3 target genes in mock and infected cells. Target genes are defined as genes with at least one significant binding site, based on DEW-seq enrichment analysis (adj. p-value < 0.01). (**B**) GO enrichment analysis of overlapping MATR3 interactors. Enrichment analysis was performed with clusterProfiler and overlapping GO terms were removed, using a similarity threshold of 0.5. The size of each point reflects the log2 odds ratio of enrichment and the shade of blue reflects the significance of the enrichment, with darker blue indicating more significant enrichment.

By comparing the list of genes with MATR3 binding sites (adjusted p-val <0.01 enrichment over SMI) in mock and infected cells, I found that there was a substantial overlap between target genes under the two conditions (Figure 4.13A). These genes likely reflect the core interaction partners of MATR3. Interestingly, GO enrichment analysis of genes interacting under both conditions indicated binding to genes involved in splicing regulation, as well as with nuclear and cytoplasmic membraneless organelles, indicating that MATR3 might bind to and regulate the mRNAs encoding the proteins it functionally interacts with. MATR3-bound RNAs include MATR3 itself, as well as known interaction partners of MATR3, such as SFPQ, which is a core member of paraspeckles. Additionally, the lncRNA NEAT1, which is central to paraspeckle assembly, interacts under both conditions (Figure 4.13B).

To determine if there were general differences in MATR3's binding behaviour following infection, I assessed the properties of MATR3 binding sites in cellular RNA. The majority of binding sites unique to each condition and shared between the two conditions are in protein-coding genes, with a small fraction in lncRNAs (Figure 4.14A). The fraction of lncRNA binding sites is slightly higher in targets identified in both mock and infected conditions. Among protein-coding genes, the majority of binding sites were localised to introns. This proportion of intron

binding is substantially higher than the average proportion found across RBPs globally (eCLIP superset in ENCODE; Figure 4.14B). By plotting the signal peak distribution within intron binding sites, and normalising these to intron lengths, I defined MATR3's pattern of binding in intronic sequences (Figure 4.14C). In contrast to many other splicing regulators, MATR3 shows distributed binding across the length of the intron with a slight peak at its start and end in both mock and infected cells. This is consistent with MATR3's known binding modality [356]. In addition to binding introns, there is also a fraction of MATR3's binding sites in the 3' UTR (Figure 4.14D). Interestingly, this proportion is increased upon infection. This is consistent with the observed shift in MATR3's localisation from nucleus to cytoplasm, which would lead to a reduction in intron binding (Kamel et al, In preparation). It might also reflect an additional function for MATR3 in binding 3' UTRs that is related to its known role in translation regulation.

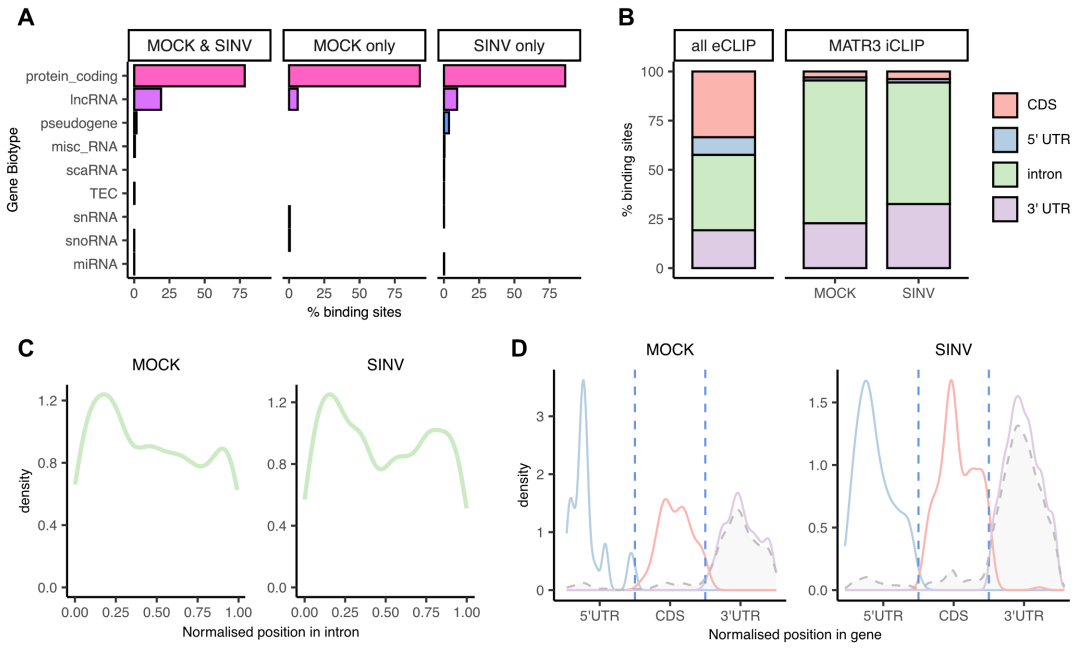


Figure 4.14: **MATR3 primarily binds introns of protein-coding mRNA.** (**A**) Bar plots showing the proportion of target RNAs corresponding to different RNA biotypes for shared interactors in mock and infected conditions, and those specific to each condition. (**B**) A stacked bar plot showing the distribution of enriched binding sites in mRNAs across gene regions. The average distribution of binding sites across all available eCLIP datasets from ENCODE is included for comparison. (**C**) A line plot showing the normalised position of each intron binding peak, relative to intron length. (**D**) Meta-gene profiles of binding site distribution across 5' UTR, CDS, and 3' UTR. Coloured lines reflect binding within each region. The grey shaded area reflects the overall pattern of binding across all regions.

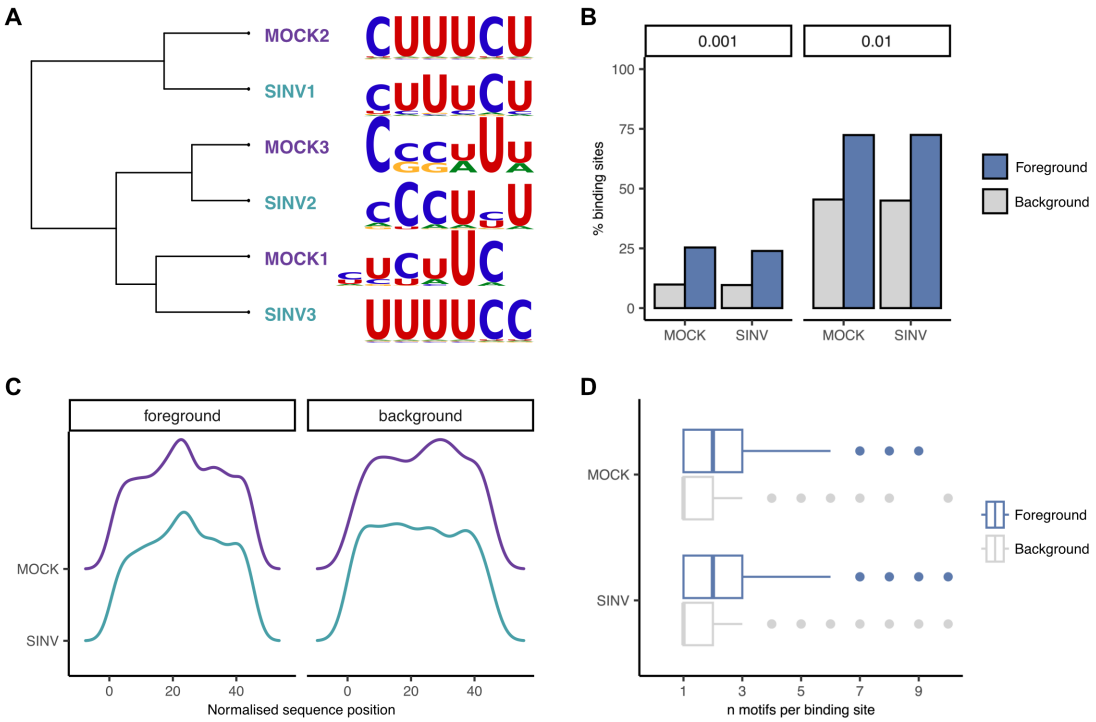


Figure 4.15: **MATR3 binds to pyrimidine-rich motifs.** (**A**) Hierarchical clustering of the top three most enriched motifs in mock and infected binding sites. Motif enrichment was performed on 50 nucleotide windows, centred around the signal peak of each binding site. Enrichment was relative to gene and gene region-matched sequences generated for each binding site, and was performed using STREME from the MEME suite. (**B**) Percentage of input sequences containing at least one motif at two motif similarity thresholds (0.001 and 0.01). Percentage of binding sites is represented by the blue bars and percentage of background sequences by the grey bars. (**C**) Profiles of motif start site distribution within 50 nucleotide binding and background sequences at a motif similarity threshold of 0.01. (**D**) A boxplot depicting the distribution of the number of motifs within each binding site or background sequence at a motif similarity threshold of 0.01.

MATR3 has altered binding targets following infection. To assess whether the binding specificity itself is changing, I performed a motif enrichment analysis in each condition. To this end, I took a 50nt window for each binding site in cellular RNA, taking the signal peak as the window's centre, and searched for enrichment over a set of gene and gene region-matched background sequences using STREME [370]. To determine whether there were multiple different motifs enriched, I performed hierarchical clustering of the top 3 motifs. These motifs all reflect MATR3's known preference for pyrimidine-rich sequences. Motifs enriched among cellular RNA targets in mock and SINV infected cells cluster together, suggesting that there is not a shift in binding specificity. Under both conditions, the proportion of sequences

containing a motif was high (Figure 4.15B). At a stringent similarity threshold, around 25% of binding sites contained a motif and at a less stringent threshold, it was close to 75%. In both cases, it was substantially above the occurrence rate in the background. Additionally, while in background sequences, there is even distribution of motif occurrences across the length of the sequence, in MATR3 binding sites, there is a clear peak at the centre of the sequence (Figure 4.15C). In these sequences, the midpoint reflects the peak in signal, indicating that crosslinking is occurring at the site of the motif. While there is a peak in binding at the midpoint, there is also relatively high levels of motif occurrence along the length of the sequence. Indeed, many binding sites contain more than one motif (Figure 4.15D). In conclusion, MATR3's binding specificity in both mock and infected conditions recapitulates what has been observed previously, increasing confidence in the results. This would also suggest that the difference in binding targets between MATR3 mock and infected samples might reflect differences in protein localisation or different interaction partners, rather than a change in MATR3's specificity.

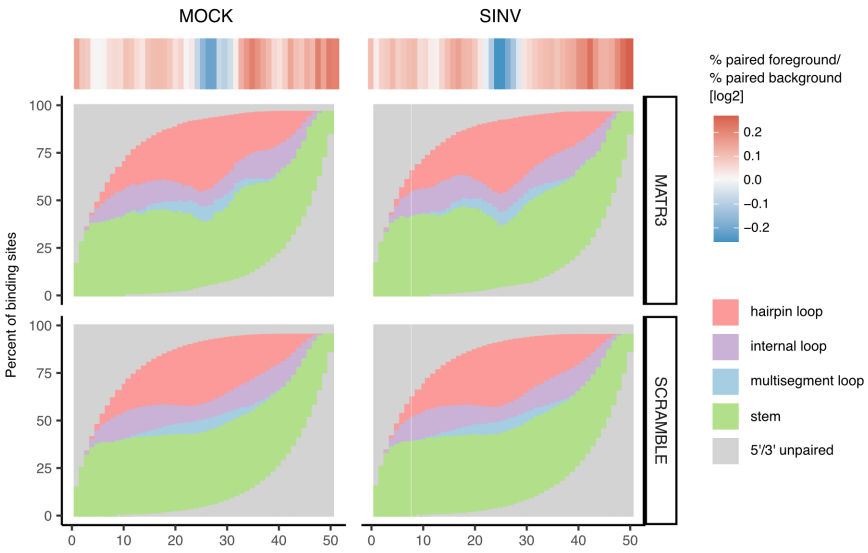


Figure 4.16: **Secondary structure preferences of MATR3.** (**top**) Heatmap plotting log2 of the % of paired nucleotides in MATR3 binding sites versus motif prediction background, according to RNAfold. Binding sites were defined 50 nucleotide windows, centred around the signal peak of each binding site. (**bottom**) Stacked bar chart plotting percentage of each structural feature at each position in foreground and background sequences, determined by Forgi.

Another aspect of an RBP's specificity is its secondary structure preference. This has not previously been determined for MATR3. To establish MATR3's structural preference, I used RNAfold to fold each input sequence used for motif prediction (25nts either side of the peak

in signal for each binding site). I also folded 10 scrambles of each sequence and assessed the frequency of paired sequence at each position, relative to this scrambled background. This indicates that in both mock and infected conditions, MATR3 preferentially binds within unstructured regions as the central regions of binding sites are under-enriched in paired nucleotides (Figure 4.16A). To assess the nature of these unstructured regions, I used the forgi package to determine what features they were likely to belong to. There is a slight peak in frequency of hairpin loop features at the centre of the sequence but, in general, there is no clear preference for specific structural features (Figure 4.16B).

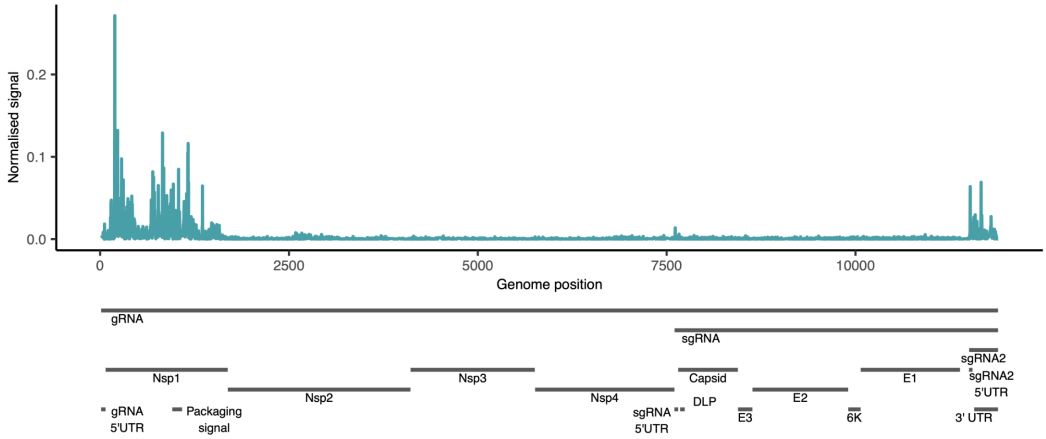


Figure 4.17: **MATR3's binding profile on SINV RNA.** A line plot depicting MATR3-GFP signal on SINV RNA. The density of crosslinked sites at each position is calculated was a percentage of total signal across the genome. Percent signal at each position in the SMI was subtracted from IP to give the MATR3 binding profile.

The above analyses have focused on differences in binding to cellular RNA. However, the most pronounced difference between mock and infected samples is the binding of MATR3 to SINV RNA. By subtracting the signal pattern observed in the size-matched input, we can observe a MATR3-specific pattern of binding (Figure 4.17). There is a strong peak in binding at the 5' end of the genome, with binding distributed along the length of nsP1. Additionally, there is a smaller binding region at the 3' end of the genome. Interestingly, the highest binding peak occurs within the second stem loop of the 51nt 5' CSE (Figure 4.18)A. This feature is highly conserved at the level of structure and sequence (Figure 4.18B). SHAPE-MaP data from Kutchko and colleagues demonstrates that there is a high degree of structure conservation between alphaviruses [371]. Multiple sequence alignment further shows that the sequence of this region is also highly conserved (Figure 4.18C). Plotting the signal intensity in iCLIP onto

the secondary structure, the peak in MATR3 can be seen to occur within the loop of the stem loop. This fits with MATR3's preference for binding to ssRNA (Figure 4.18D). The sequence in this loop is pyrimidine-rich, further reflecting MATR3's known specificity. It is important to note that analysis of MATR3's viral genome binding is still ongoing. Therefore, while exciting, these results should be considered preliminary.

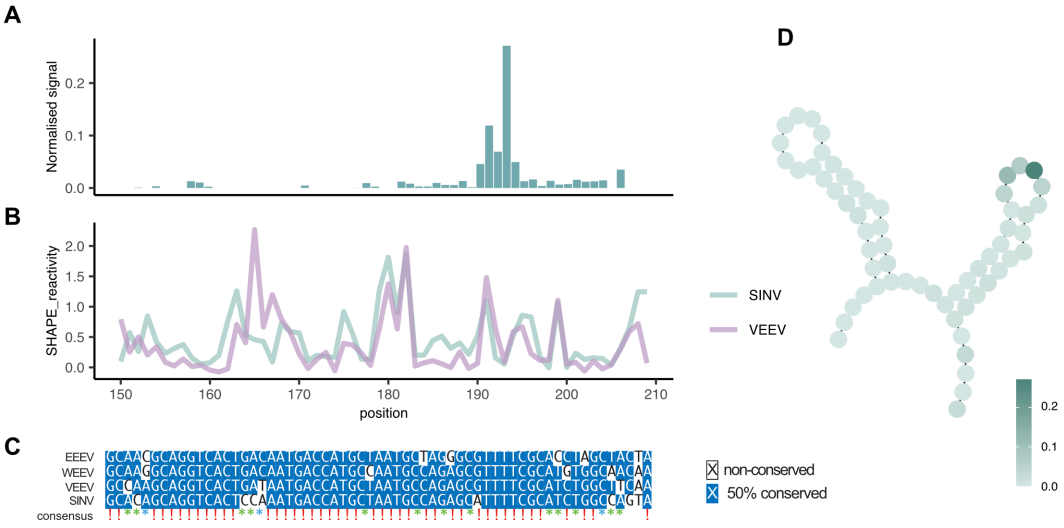


Figure 4.18: **MATR3 binds to SINV's conserved 51nt 5' CSE.** (**A**) A zoomed in bar plot of the normalised signal shown in Figure 4.17, focusing on the 51nt 5' CSE. (**B**) SHAPE-MaP reactivity for SINV (teal) and VEEV (mauve) taken from [371]. (**C**) Multiple sequence alignment of SINV to three New World alphaviruses. Nucleotides in blue are conserved across all four species. (**D**) Secondary structure plot showing the results of RNAfold for the 51nt 5' CSE. SHAPE-MaP data was included to inform folding. Crosslinking intensity at each position is reflected by colour.

4.2.3 Does S188 phosphorylation state alter MATR3's binding specificity?

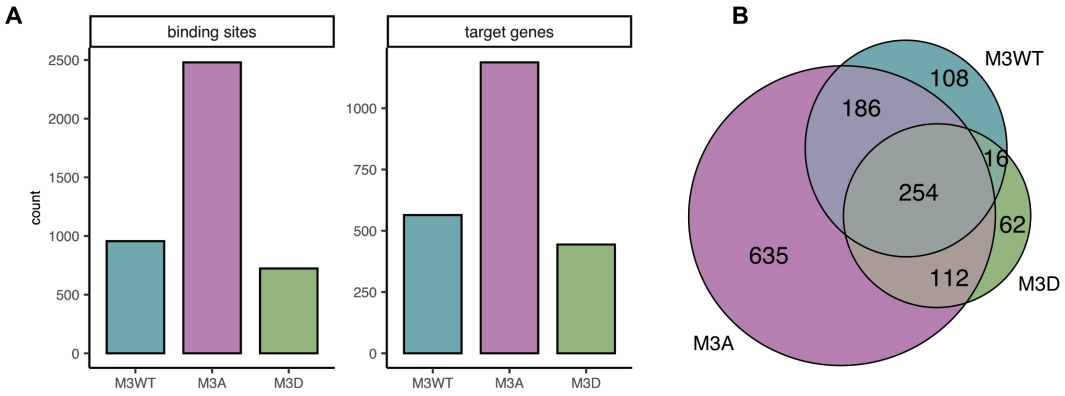


Figure 4.19: **There is overlap in binding targets of WT MATR3 and phosphomutants.** (**A**) Bar plots showing the number of binding sites and target genes in mock samples from WT, S188A, and S188D MATR3-GFP. (**B**) A Venn diagram showing overlap between WT and mutant MATR3 target genes. Target genes are defined as genes with at least one significant binding site, based on DEW-seq enrichment analysis (adj. p-value < 0.01).

Evidence presented in Chapter 3 indicates that MATR3's RNA-binding activity can be modulated by its phosphorylation state at S188. To determine whether the differences in binding activity were driven by a change in specificity or affinity, I performed iCLIP on MATR3-GFP S188A (phospho-null) and S188D (phosphomimetic) mutants. Strikingly, in mock conditions the pattern in number of binding sites and target genes identified across the three cell lines mirrors that observed by dual-fluorescence assay (Figure 4.19A & Figure 3.12C). Although the extensive sample processing required for iCLIP2 limits its quantitative potential, the consistency between these two methods offers further validation of the results observed by dual-fluorescence assay. There is a high degree of overlap between the target RNAs identified in the WT and mutant MATR3 samples in mock conditions (Figure 4.19B). This indicates that there is a core set of interactions shared regardless of the protein's phosphorylation state.

In addition to an overlap in targets, I also wanted to assess whether the general preferences of MATR3 are also preserved by these mutations. As with WT MATR3, both S188A and S188D mutants predominantly bind to protein coding genes (Figure 4.20A). They also show the same pattern of binding within protein coding genes, with the vast majority of binding sites occurring in introns. Notably, in both mutants, there is the same increase in the proportion of 3' UTR binding sites following SINV infection, indicating that this behaviour is not related to

phosphorylation state.

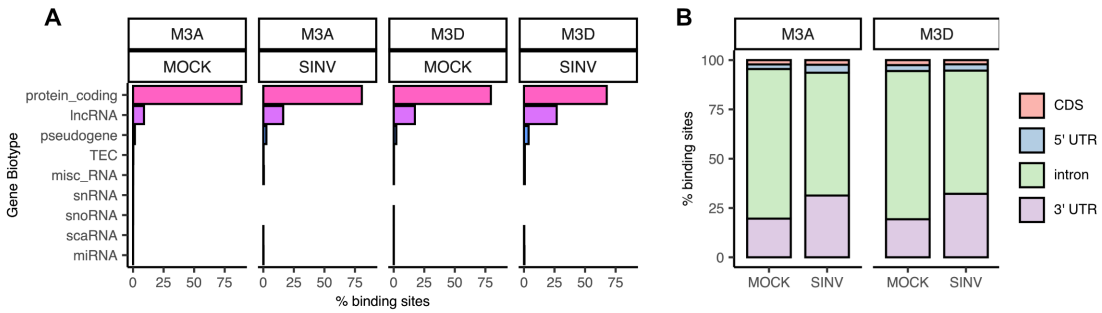


Figure 4.20: **Phosphomutants share WT MATR3's transcript preferences.** (**A**) Bar plots showing the proportion of target RNAs corresponding to different RNA biotypes. This percentage was calculated for mock and infected conditions for both mutant lines. (**B**) A stacked bar plot showing the distribution of significantly enriched binding sites in protein coding mRNAs within each region of the RNA, based on genome annotation.

Since specificity is thought to be driven by MATR3's RRMs, I hypothesised that S188 phosphorylation, which occurs in a disordered region upstream of the RRMs would be unlikely to alter motif binding preference. To determine whether specificity was altered, I performed motif enrichment for the mutants in mock and infected conditions. As anticipated, when comparing the most enriched motif across all samples, they all reflected MATR3's pyrimidine-rich sequence preference (Figure 4.21A). The proportion of sequences containing the motif was also found to be consistent across all samples, indicating that motif-driven binding is not affected by the protein's phosphorylation state (Figure 4.21B).

While I was not able to establish how S188 phosphorylation altered MATR3's antiviral activity, I was interested to know whether there were changes in the pattern of binding to vRNA. I found that the pattern of binding is very similar across WT and mutants (Figure 4.22). This indicates that, even in this case of unconventional binding, MATR3's preference is not altered by its phosphorylation state. Together, these results indicate that MATR3 S188 phosphorylation is able to alter the affinity of MATR3's RNA interaction without altering its specificity in binding to both cellular and viral RNA.

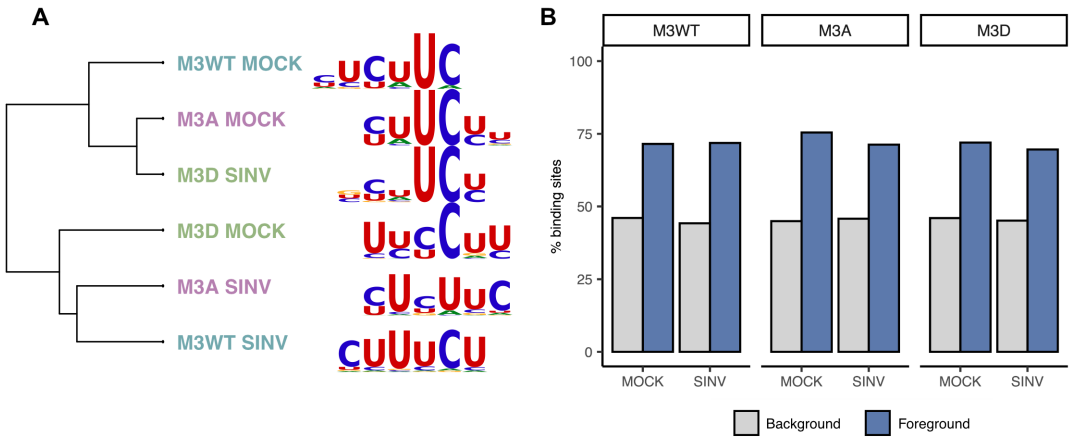


Figure 4.21: **Binding specificity is conserved in phosphomutants.** (**A**) Hierarchical clustering of the most enriched motif in mock and infected binding sites for WT and mutant MATR3. Motif enrichment was performed on 50 nucleotide windows, centred around the signal peak of each binding site. Enrichment was relative to gene and gene region-matched sequences generated for each binding site, and was performed using STREME from the MEME suite. (**B**) Percentage of input sequences containing at least one motif at a similarity threshold of, e.g., 0.01. Percentage of binding sites is represented by the blue bars and percentage of background sequences by the grey bars.

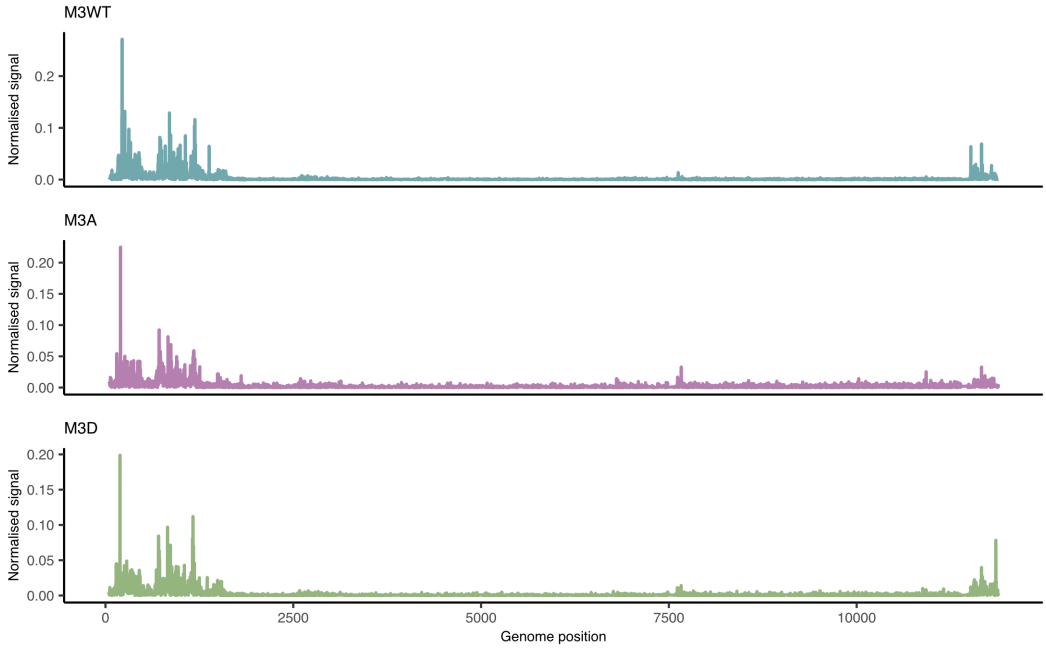


Figure 4.22: **Phosphomutant viral RNA binding pattern is the same at WT MATR3.** Line plots depicting MATR3-GFP WT and mutant signal on SINV RNA. The density of crosslinked sites at each position is calculated was a percentage of total signal across the genome. Percent signal at each position in the SMI was subtracted from IP to give each protein's binding profile.

4.2.4 Characterising the vRNA interactions of RBPs that regulate infection

Studying binding of cellular RBPs to vRNA with single nucleotide resolution has the potential to greatly enhance our understanding of host-virus interactions. During my PhD, I have had the opportunity not only to analyse MATR3 iCLIP data in SINV infected cells, but also datasets for several other proteins. These have included: TRIM56 (iCLIP performed by me), TRIM25 (iCLIP performed by Dr Lucia Alvarez [372]), XRN1 (iCLIP performed by Dr Vincenzo Ruscica, in preparation), DDX1, and RTCB (iCLIP performed by Natasha Palmalux, in preparation). While the exact conditions and goals of each experiment varied, by looking at these datasets side-by-side, we can begin to observe trends and understand potential biases in CLIP-based analyses of virus infected cells.

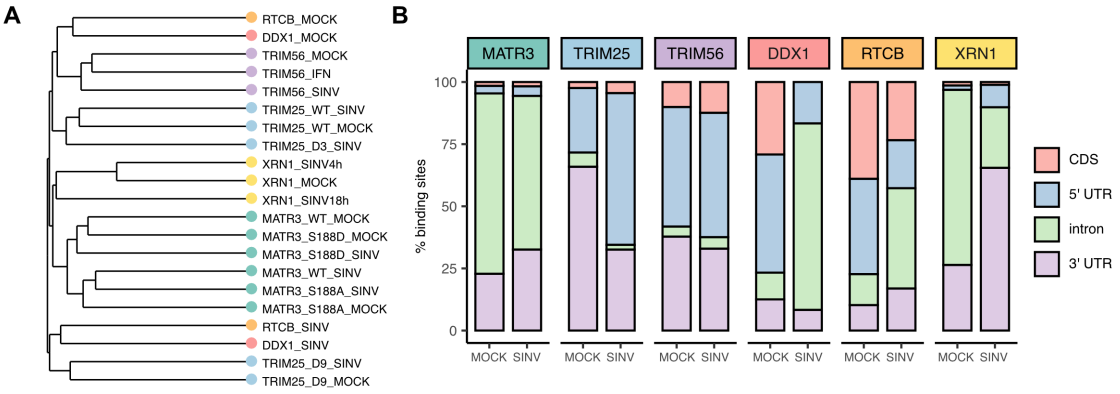


Figure 4.23: **Comparing iCLIP datasets for RBPs with roles in SINV infection.** (**A**) A tree depicting results of hierarchical clustering analysis of RBP target genes. RBPs were clustered based on a one-hot encoding matrix of genes bound by each RBP at an FDR threshold of 1%. Point colour represents the RBP from which each sample was derived. (**B**) A stacked bar plot showing the distribution of significantly enriched binding sites in protein coding mRNAs within each region of the RNA, based on genome annotation. Bars are separated by condition and RBP.

One benefit of having several datasets generated in the same cell line and analysed by the same pipeline is that we can explore similarities and differences across different RBPs and conditions. For example, comparing MATR3 mutants, their pattern of binding to both cellular and viral RNA looks strikingly similar. However, I wanted to rule out the possibility that similarities between these datasets is a reflection of inherent biases in the iCLIP method, rather than biologically meaningful effects. To this end, I assessed similarity of all the iCLIP datasets based on their cellular RNA targets using hierarchical clustering (Figure 4.23A). In

this analysis, all MATR3 samples clustered together on their own branch of the tree, indicating that their similarity is biologically, rather than technically driven.

While MATR3 mutant samples cluster together, the TRIM25 D9 (RNA-binding-null mutant) samples are in an entirely separate branch of the tree to the rest of the iCLIP samples, along with RTCB and DDX1 infected samples (Figure 4.23A). Unlike MATR3 S188 mutations, which have subtle effects on binding affinity, the TRIM25 D9 mutant has lost all RNA-binding activity, based on nuclear magnetic resonance (NMR) and isothermal titration calorimetry (ITC) analysis [372]. Therefore, the target genes identified in these datasets are likely to predominantly reflect background or contaminant signal. For example, there are many mitochondrial ribosomal genes among the binding targets, which are known to be common contaminants in iCLIP data. During infection, DDX1 and RTCB show a dramatic shift from cellular RNA to viral RNA. This would explain the fact that it essentially resembles a 'no binding' profile when focusing on cellular RNA.

In addition to looking at specific binding targets, we can also compare binding site distribution within genes. For MATR3, I observed that it exhibited a predominantly intronic localisation but that there was a shift towards more 3' UTR binding following infection (Figure 4.14B). Interestingly, a shift in binding site distribution seems to be a common feature of infection (Figure 4.23B). XRN1 shifts from intron to 3' UTR binding, while TRIM25 shifts from binding predominantly 5' UTR to CDS. DDX1 and RTCB, which bind to the 5' UTR and the start of coding sequence in mock conditions, exhibit a loss of clearly defined binding preference due to the loss of cellular binding sites. These results together suggest that, for proteins with diverse binding specificities, SINV infection can often result in changes in cellular binding preferences.

In addition to binding cellular RNA, MATR3, TRIM25, TRIM56, DDX1, RTCB, and XRN1 also bind viral RNA. Comparing the normalised signal distribution of each protein across the genome, it becomes very clear that there are distinct patterns of binding (Figure 4.24). Unlike GFP, which does not show specific binding to vRNA and exhibits a uniform distribution of signal, each protein has distinct peaks in signal across the genome. MATR3 (WT and mutants) bind to the 5' and 3' ends of the viral genome, as described above. In contrast, XRN1 predominantly binds to the 3' end of the genome, consistent with its binding to 3' UTRs of cellular RNA following infection. DDX1 and RTCB exhibit pronounced binding at the 5' end of the subgenomic RNA (sgRNA), which is also the start of the gRNA 3' UTR. This 5' end binding is also consistent with DDX1/RTCB's cellular RNA binding preference, TRIM25

and TRIM56 have sites of binding within the sgRNA region which reflect their structure and sequence preferences. This would indicate that cellular RBPs have specific binding sites within viral RNA, which could provide clues as to their roles in infection.

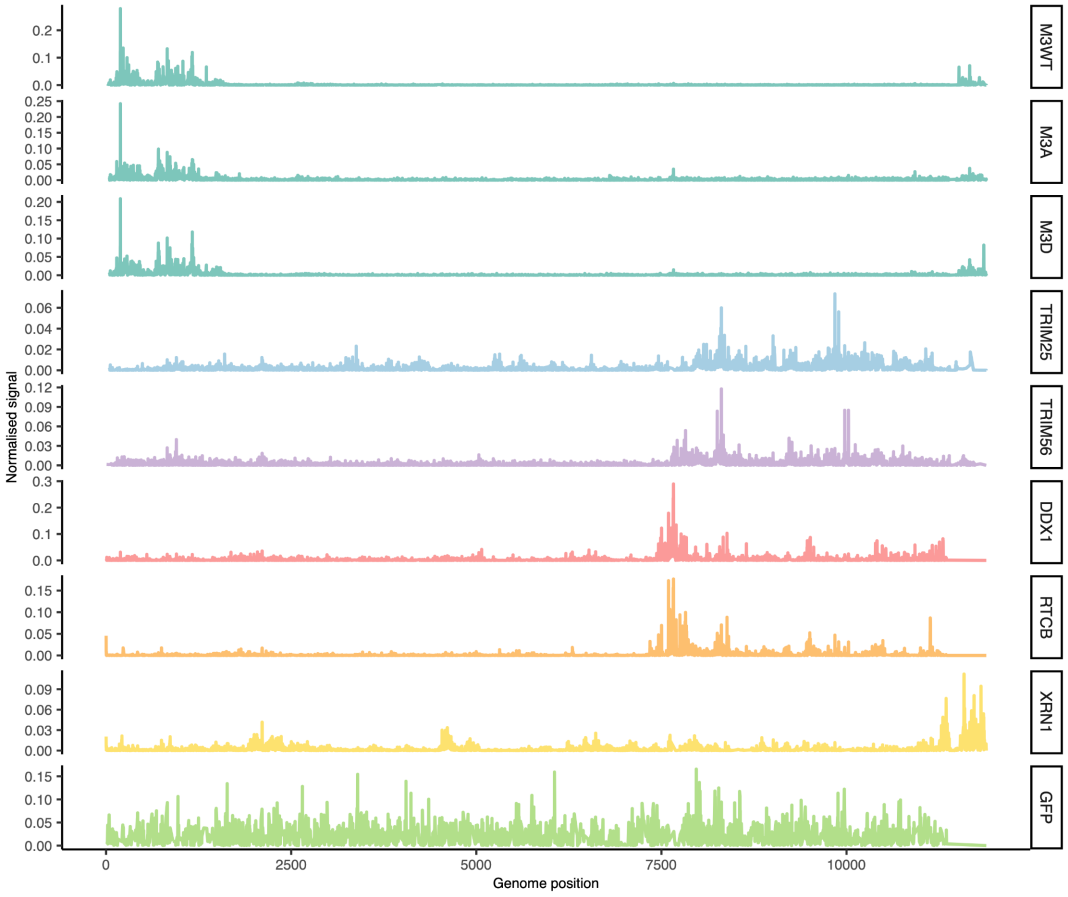


Figure 4.24: **RBPs show diverse binding profiles on SINV RNA.** Line plots depicting RBP signal on SINV RNA. The density of crosslinked sites at each position is calculated as a percentage of total signal across the genome. Percent signal at each position in the SMI was subtracted from IP to give each protein's binding profile.

4.3 Discussion

4.3.1 MATR3 PPI uncovers a wide network of potential RNA-dependent and independent interactions

MATR3 is a nuclear matrix protein with diverse roles in RNA metabolism. I have identified MATR3 as an IR-RBP, whose RNA-binding activity is modulated through its phosphorylation state and which exhibits antiviral activity against SINV. To assess which of the processes MATR3 has been linked to might be contributing to this antiviral activity, I performed a PPI analysis in mock and infected cells. In addition to comparing mock and infected conditions, I also compared the interaction network in the presence and absence of RNA. Together, these data allow me to assess the role of infection and of RNA in shaping MATR3's interactions.

In each condition, I found that hundreds of proteins were enriched over the parental control at a stringent FDR threshold. While MATR3 is undoubtedly interacting with a wide range of proteins, the number of proteins enriched above the parental control in this data is not likely a reflection of only direct MATR3 interactions. Rather, it represents the wider network of binding within which MATR3 sits. I purposefully used native wash conditions (150mM NaCl) to preserve weak interactions. These wash conditions, combined with a very clean control and the extremely high sensitivity of the instrument used for LC-MS/MS analysis means that even peripheral and indirect interactors of MATR3 will be detected as enriched.

Diverse studies, using affinity purification- mass spectrometry (AP-MS) or the yeast two-hybrid system, have sought to establish the interactome of MATR3, or have identified MATR3 as an important interaction partner of other proteins of interest. In total, there are close to 700 annotated interaction partners of MATR3, based on the curated BioGRID database [373]. Therefore, the goal of this study was not to establish a list of MATR3's direct interactions. Rather, I wanted to assess how the interaction network of MATR3 is altered following infection. However, it is reassuring to note that many validated interaction partners of MATR3 are enriched in this IP. These include: DDX5, RALY, ZAP, RBM45, AGO2, PTBPs 1-3, DHX9, HNRNPK, P53, NONO, SFPQ, and TARDBP [114, 327, 358, 362, 365]. The presence of previously validated interaction partners confirms dataset quality. However, an important next step in this study will be to validate interactions of interest in the context of infection.

Being able to detect both dominant and peripheral interaction partners has benefits. However, it is also important to differentiate between them to identify those that are likely to be contributing

directly to MATR3's function. An initial indication of whether proteins are likely to be true interactors is to assess their enrichment in the IP relative to the cell's proteome. Most of the proteins that are enriched following infection are also enriched in the IP over the WCP, suggesting that they are not simply highly abundant proteins in the cell. Interactions of interest that exhibit enrichment over the WCP will next need to be validated by IP under more stringent wash conditions and using an endogenous antibody (rather than GFP-Trap), followed by Western blot. It would also be valuable to perform an IP of the protein of interest and determine whether MATR3 can be detected among its interaction partners.

In addition to experimental validation of specific interactions, there are also *in silico* tools that can provide an indication of whether and how two proteins might be interacting. Computational docking methods have been employed for decades to identify likely binding interfaces between proteins [374]. The recent advent of deep learning tools like Alphafold and RoseTTa have made it possible to perform co-folding 'experiments' at scale *in silico*. This has been applied in several contexts to provide structural insight into known interactions and predict new ones [375]. One of the challenges of identifying potential MATR3 interaction interfaces is that the protein is highly disordered. However, there is evidence that Alphafold can effectively predict conditional folding within disordered regions that occurs as a result of PPIs [376].

In silico co-folding can be supported and complemented by experimental approaches. One such approach is crosslinking mass spectrometry (XL-MS). In XL-MS, a chemical crosslinker is employed either prior to cell lysis or after IP. Crosslinks act to tether interacting proteins during IP and, importantly, interacting peptides during subsequent MS/MS analysis [377]. By detecting linked peptides, it becomes possible to directly map intra- and inter-protein interfaces [374]. This is a powerful way of validating the presence of direct interactions. It can also provide constraints for co-folding analysis, improving our understanding of how proteins interact [374]. However, it is important to note that, given the low abundance of any individual crosslinking event and consequent dynamic range issue in MS analysis, the absence of crosslinking cannot rule out direct interaction.

Many proteins that are genuine interaction partners of MATR3 might rely on RNA to bridge or stabilise their interaction. Indeed, low-throughput validation of several proteins identified in MATR3 PPIs, including HNRNPK, DHX9, ZAP, and the HIV-1 protein Rev, have indicated that interaction is reduced or lost following RNase treatment [327,360,362]. The importance of RNA in bridging PPIs, particularly for RNA-binding proteins is becoming increasingly apparent. A

study performed by Caudron-Herger and colleagues in 2019 assessed proteins' RNA-dependence by comparing their sucrose gradient migration profiles in the presence and absence of RNases [378]. 37% of proteins identified (1784 in total) exhibited a significant shift in migration pattern following RNase treatment, indicating that their interactions are, to some degree, RNA-dependent. For some proteins, RNase treatment resulted in a dramatic downward shift into a single peak, reflective of monomeric protein. For others, MATR3 included, there is a peak shared between RNase treated and untreated conditions in addition to a treatment-specific peak. This likely reflects the presence of MATR3 in both RNA-dependent and independent complexes. This is consistent with what I observe in the present study. I found that, following benzonase treatment, 70% and 50% of proteins had significantly altered interaction with MATR3 in mock and infected cells, respectively. Both RNA-dependent and independent interaction sets are enriched in GO terms related to RNA-binding. While this might seem surprising for RNA-independent interactions, evidence suggests that RBPs' RNA-dependent and independent interaction partners are highly likely to also be RBPs [379].

In this study, I have defined two groups: RNA-dependent and RNA-independent interactors. This classification is a helpful starting point for understanding the interaction landscape but does not capture the spectrum of interaction modalities encompassed by the concept of 'RNA-dependency'. Some proteins identified in this study might simply be sitting nearby to MATR3 on the same RNA. We would expect these proteins to have the greatest fold-change or have 'on-off' changes, where the protein is only detected in the absence of benzonase. Alternatively, some proteins might directly interact with MATR3 in a manner that is stabilised by the presence of RNA. Therefore, for proteins of interest, it would be important to individually assess their level of RNA-dependency. The interaction validation strategies described above would also be important in determining whether proteins are likely to be directly interacting and what the role of RNA might be in each instance.

4.3.2 MATR3 alters its interaction network upon SINV infection

The goal of this PPI analysis was to determine whether the interaction network of MATR3 altered upon SINV infection. In total, I found 65 proteins with increased binding and 83 proteins with decreased binding in the benzonase+/- samples following SINV infection. Among the increased interaction partners are many cap binders and translation regulators. Decreased binders were enriched in proteins involved in microtubule cytoskeleton and the mitotic spindle,

and a different subset of proteins involved in translation.

This altered interaction network might, in part, reflect the altered localisation of the protein following infection. MATR3 exhibits a shift towards the cytoplasm following SINV infection, which might account for the reduced interaction with proteins associated with chromatin organisation and transcription control [364]. In this study, I used gentle NP40 lysis buffer and low-speed spin to sediment nuclei. Since MATR3 is predominantly nuclear, even after infection, the purpose of this was to allow me to better observe interactions occurring within the cytoplasmic fraction of the protein. However, while I intended to enrich for the cytoplasm, I did not perform a complete nuclear/cytoplasmic fractionation and did not validate that nuclear proteins were excluded. Given the proportionally very high levels of nuclear MATR3, it is highly likely that a meaningful proportion of signal reflects nuclear interaction partners. Therefore, it is possible that some of the changes observed reflect a nuclear/cytoplasmic shift.

Another potential explanation for MATR3's altered interactions is that they reflect a functional shift following infection. iCLIP data indicates that MATR3 shows a subtle shift in cellular RNA-binding towards 3' UTRs, as well as exhibiting a clear peak in binding at the 5' end of the viral genome. Many of the proteins with increased binding following infection are RNA-dependent and are members of the vRNA interactome. The presence of MATR3 at the 5' end of the viral RNA could cause it to come into contact with a new set of cap-binding or 5' end binding proteins. Excitingly, among these proteins are several that have been linked to antiviral innate immunity, including PARP12, ZC3HAV1 (ZAP), HERC5, IFIT1, and IFIT3.

ZAP exhibits increased MATR3 binding following SINV infection and interacts with MATR3 in an RNA-dependent manner. This is not the first study to demonstrate an RNA-dependent ZAP-MATR3 interaction following infection. A study identifying MATR3's interaction partners after HIV-1 infection found that it associates with ZAP, as well as DDX17 and EXOSC3, which are both involved in RNA degradation [362]. Knockdown of MATR3 was found to result in more pronounced restriction of HIV-1 following ZAP overexpression. This effect was also observed to a lesser extent in Moloney murine leukemia virus (MLV) infection. Based on this, the authors concluded that MATR3 is directly inhibiting the ability of ZAP to degrade viral RNA. Two previous studies have already identified MATR3 as an interaction partner of Rev that promotes infection by aiding in Rev-dependent RNA export and stabilisation [360,361]. MATR3 is known to stabilise a small subset of cellular transcripts so this function for MATR3 is not unprecedented [327]. However, it is interesting to note that two other nuclear matrix proteins,

SFPQ and HNRNPU were found to have the same effect upon ZAP overexpression [362]. This suggests that the effect is not specific to MATR3 but instead relates to an RNA metabolic process taking place in the nucleus. While MATR3 and other nuclear matrix proteins appear to be co-opted to stabilise HIV-1 RNA, retroviruses are processed through a pathway much closer in nature to cellular mRNA processing. Therefore, it would not be surprising if MATR3's role of HIV-1 infection differs from that in cytoplasmic RNA viruses.

ZAP is known to be an important antiviral factor in alphavirus infection. There is evidence of ZAP inhibiting replication of SFV, RRV, VEEV, and SINV [380]. Unlike most RNA viruses, alphaviruses do not suppress CpG dinucleotides in their genome [104,381]. CpGs are recognised by ZAP through direct binding, resulting in translation inhibition and/or RNA degradation [382]. Using a temperature sensitive SINV mutant, which can undergo translation but not replication, a 2003 study found that ZAP-driven alphavirus inhibition occurs during the initial translation stage [380]. There is evidence that ZAP might interfere with the interaction between eIF4A and eIF4G, so it is possible that it could bind at the 5' end of the genome to mediate this, thus sitting in proximity to MATR3 [106]. However, whether MATR3 is important for its antiviral function is not clear.

Unlike ZAP, IFIT1 and IFIT3 interact with MATR3 in an RNA-independent manner. IFIT proteins are strongly and rapidly induced ISGs [383]. They can also be upregulated directly by IRF3 signalling so are classified as 'viral stress-inducible genes' [95, 96]. IFIT1 binds to both 5' PPP structures [71] and capped RNAs that are not 2'-O-methylated on the first nucleotide (cap0 RNA), features that are often found in vRNA [97, 98]. IFIT1 is the only IFIT protein that has a binding pocket large enough to accommodate the cap structure [384] but IFIT2 and IFIT3 can form complexes with it to promote its antiviral activity [99]. The primary mechanism by which IFIT1 is thought to inhibit infection is through out-competing eIF4E in binding to the cap and thus preventing translation initiation, although it can also inhibit translation through eIF3 binding [385]. Since RNA polymerase II transcribed sequences have either cap1 or cap2 structures, IFIT1's specificity represents a powerful way of differentiating self from non-self RNA [386]. This allows for the continued translation of ISGs, while viral RNA translation is blocked.

Unlike other viral families that have evolved 2'-O-methylation strategies [387], alphavirus RNAs are not 2'-O-methylated, making them targets for restriction by IFIT proteins. Interestingly, a secondary structure element has evolved within the 5' UTR of alphaviruses which reduces the

binding of IFIT1 [388]. Differences in 5' UTR structures have been found to substantially alter alphavirus sensitivity to IFNβ and this altered sensitivity is predominantly driven by differential IFIT1 binding [389]. While this 5' UTR structure can partially overcome IFIT1 restriction, it does not completely abolish it. SINV infection is reduced by IFIT1 overexpression and is enhanced by its knockdown [207], suggesting that there is still interaction with the cap0 structure.

One intriguing possibility is that MATR3 could stabilise the binding of IFIT1/IFIT3 to the viral genome. Interestingly, MATR3 is one of a relatively small group of proteins that has been found to bind to IFIT2 following IFN-I treatment [71]. This indicates that MATR3's protein interaction partners, as well as its RNA-binding activity, might be regulated by the IFN-I response to promote its antiviral effect and that one aspect of this effect might involve engagement with IFIT proteins. It will be important to establish whether IFIT1/3's antiviral activity is modulated by altered MATR3 levels and to assess how the proteins are interacting to determine whether MATR3 is indeed contributing to IFIT-mediated restriction.

In addition to proteins involved in antiviral immune response, I identified several nuclear 5' end binding/ processing proteins. Among these is CMTR1, the 2'-O-methyltransferase responsible for formation of the cap1 structure [390]. This protein is an ISG that is induced to ensure that ISG transcripts possess the cap1 structure and can thus evade restriction by ISGs like IFIT1 [23, 386]. Additionally, there are the three nuclear cap binding proteins: NCBP1, NCBP2 and NCBP3. NCBP1 and 2 form the canonical cap binding complex which is essential for various aspects of RNA metabolism, including regulation of splicing, export and stability [129]. NCBP3 is a non-canonical member of the complex that can either substitute for NCBP2 or engage with both proteins as an accessory factor [391]. These proteins are exported with the RNA and are thought to contribute to the pioneer round of translation, before being replaced by eIF4F [392]. NCBP3 KO results in increased infection of a range of viruses, including the alphavirus SFV. This increase was paired with a decrease in innate immune activation at the protein but not the RNA level, indicating that NCBP3 is involved in translational control [393]. NCBP1 binds to SINV RNA so it is possible that association of MATR3 with this complex occurs on the vRNA [364]. However, what the role of this binding might be is unclear. It seems unlikely that CMTR1 is binding to MATR3 in the context of vRNA since SINV is not 2'-O-methylated. Whether these nuclear 5' end processing protein interactions relate directly to vRNA binding or to another cellular function of MATR3 will need to be assessed.

In addition to these 5' end regulatory factors, there are also several proteins involved in

canonical translation that exhibit altered binding following infection. eIF4G and eIF4E exhibit increased RNA-dependent binding. These are members of the eIF4F complex that binds to the 5' end of mRNA transcripts to promote initiation [130]. Interestingly, several members of the eIF3 complex show decreased binding. Unlike the eIF4F complex, which binds directly to the mRNA prior to ribosome recruitment, the eIF3 complex associates with the 40S ribosomal subunit in advance of its engaging with RNA [130]. The presence of eIF4 but not eIF3 factors could indicate incomplete initiation without subsequent translation. MATR3 associates with the polysome fraction of uninfected cells [325]. It is possible that MATR3 is associating with translationally repressed cellular mRNAs or that it is playing a role in repressing translation of SINV RNA. It would be interesting to determine by polysome profiling whether MATR3's association is altered following infection. It would also be interesting to know whether this altered association reflects behaviour on viral or cellular RNA. Viral sgRNA translation occurs through an eIF4F-independent mechanism [235]. However, this might not be the case for gRNA translation, which is what we would expect MATR3 to be involved in, given its RNA binding profile.

An important question uniting all of the proteins discussed above is which of them are binding in concert, which are mutually exclusive, and what the RNA contexts that they are binding in are. For example, we would not expect binding of the nuclear cap binding complex and eIF4F to the same RNA, nor would we expect either of these complexes to be binding in the presence of IFITs, which physically obscure the cap structure. An approach to determining which proteins are binding together would be to carry out size exclusion chromatography of native complexes from MATR3 IP eluates. This approach can be combined either with Western blotting of specific targets or LC-MS/MS to assess co-elution patterns globally [394]. Additionally, RNA present in the sample can be analysed by RT-qPCR to determine which complexes are most strongly associated with vRNA. Such an approach was used to assess the composition of a HEXIM1-associated protein complex that is involved in cGAS/STING signalling [363]. Using glycerol density gradient centrifugation, distinct sets of protein interactions that eluted separately could be identified. Additionally, RT-qPCR analysis of the elution profile of the HEXIM1's canonical RNA target, 7SK, indicated that these interactions were 7SK-independent.

By combining fractionation with the XL-MS strategies described above, it might also be possible to achieve sufficient depth to assess the interaction dynamics of specific MATR3

complexes. The potential of XL-MS for differentiating between sets of interactions of a given protein has already been demonstrated. Herzog and colleagues perfomed XL-MS following IP of several proteins known to associate with PP2A as part of different holoenzymes [395]. Combining XL-MS with fractionation would reduce sample complexity within each fraction, increasing the likelihood of detecting crosslinked peptides. Additionally, co-elution provides another layer of constraint in establishing complex composition.

4.3.3 MATR3's PRI motif is important for its antiviral activity

XL-MS has the potential to enhance our understanding of which proteins MATR3 is directly interacting with and where their interaction interfaces might be. One feature that we already know plays a central role in MATR3's protein interaction dynamics is the PRI motif. This is a 7 amino acid region that sits within one of MATR3's disordered regions and mediates MATR3's interaction with PTBP1 [332]. The PRI motif has been identified in several PTBP1 interaction partners and is characterised by the consensus [S/G][I/L]LGxφP, where φ is a small hydrophobic side-chain [396].

The PRI domain is sufficient to mediate MATR3's interaction with PTBP1 and mutations within this domain almost completely abolish binding [332]. Surprisingly, mutation of the PRI domain globally disrupts MATR3's regulation of alternative splicing, despite only 18% of these regulatory events requiring PTBP1. This suggests that other proteins can form important interactions through the PRI domain. In this study, I found that cells expressing PRI mutant MATR3 exhibit increased levels of infection relative to cells expressing WT MATR3, indicating that the PRI motif is important for MATR3's antiviral activity. Surprisingly, mutation of the PRI binding site in PTBP1 had the opposite effect, suggesting that loss of PTBP1 interaction is not responsible for MATR3-dependent antiviral activity.

A comparative affinity purification of WT and PRI mutant MATR3 indicated that the most significantly altered interaction partners of MATR3 were PTBP proteins [332]. Other proteins that were found to have reduced binding to PRI mutant MATR3 are known interactors of PTBP1 and do not contain clear RRM2-like structures that would be indicative of PRI binding. Therefore, it is unclear how PRI is contributing to MATR3's activity. One possibility is that other PTBP paralogs are responsible for the effect. Interestingly, PTBP3 has increased RNA-independent binding to MATR3 following infection. PTBP3 is a poorly-understood PTBP1 paralog that is present at much lower levels than PTBP1 in the cell [367]. PTBP1 RRM2

mutation might free up MATR3 for PTBP3 interaction, explaining why infection is reduced in these cells. However, given how little research there is into PTBP3, how it would contribute to antiviral innate immunity is unclear. Another possibility is that the PRI mediates an interaction that occurs specifically in response to infection, such as the interaction with IFIT proteins. It would be interesting to determine whether there are any differential interactions between WT and mutant MATR3 specifically in the context of infection.

4.3.4 MATR3 iCLIP data successfully recapitulates known binding properties

MATR3's interactions with RNA are central to its various functions in the cell. To determine MATR3's RNA interaction properties following infection, I performed iCLIP2 in mock and SINV infected cells. I found that MATR3's binding specificity remains broadly the same following infection, with a slight shift in preference from intron to 3' UTR binding. However, I found that MATR3 shows a clear peak of binding to the SINV genome, consistent with vRNA interactome data.

In this study I found that MATR3 binds predominantly to introns in protein coding genes, consistent with its role in AS regulation. As discussed in this chapter's introduction, previous studies have found a very similar pattern of binding, speaking to the quality of this dataset [332, 335]. Additionally, I found enrichment for binding to pyrimidine-rich motifs, which also reflects the findings of SELEX/RNAcompete-style approaches and CLIP datasets. The most enriched motif by RNAcompete is AUCUU [161]. In contrast, CLIP methods are generally enriched for motifs without the A nucleotide. This is likely a reflection of the differences between binding properties *in vivo* and *in vitro*.

In vitro RNA-binding strategies like RNAcompete rely on assessing interaction affinity of purified RNA-binding domains with short oligonucleotide sequences. These approaches have been extremely powerful in identifying specificity within canonical RNA-binding domains but they have also led to a focus on short linear motifs as the best way of defining RBP specificity. While some proteins, such as Rbfox, exhibit extremely specific binding to a single motif, for most proteins, this fails to capture the true complexity of RNA binding preference [162, 397]. Indeed, given the frequency at which short motifs appear in the genome, this would not offer sufficient selectivity to allow for specific interaction and function.

In an attempt to move away from this overly simplified definition of specificity, I began to look at clusters of enriched motifs, rather than focusing only on the top hit given by STREME.

This has yielded some interesting results. For example, in an analysis of PURA and PURB iCLIP, I found that the specificity of these proteins is made up of two distinct features: a poly(U) tract and a purine-rich sequence [272]. Positional analysis indicated that the poly(U) motif was consistently found before the purine motif, giving us a better understanding of how these proteins are able to achieve binding specificity. For MATR3, all of the enriched motifs reflect variations on a theme; they are all pyrimidine-rich, with the majority containing a strong 'UC'. For MATR3, an exact motif does not seem to be as important as the general sequence properties. This is clear when comparing the proportion of binding sites at low and high motif similarity thresholds. Even when a less stringent threshold is used, there is still pronounced enrichment of motifs in MATR3 binding sites, relative to background. Together, these results indicate that MATR3 can bind to regions in the same sequence space comparable to, if not identical to, the most enriched motifs. Another notable feature of MATR3-bound sequences is the higher incidence of multiple motifs within a given sequence. This is consistent with MATR3's known interaction with multivalent binding sites. This can facilitate homo-oligomerisation of MATR3 on target RNAs, or the interaction of MATR3 with other RBPs. For example, PTBP1 and MATR3 share very similar specificity, allowing them to bind simultaneously to LINE elements [356].

RBPs can also display specific structural preferences, either in combination with or independently of sequence specificity. Most RBPs exhibit a preference for binding within unstructured regions, as is the case with MATR3 [162]. However, it is important to note that protein binding to ssRNA is more likely to be crosslinked and detected via methods such as iCLIP2 than dsRNA regions, resulting in technical bias [368]. While some proteins simply show preference for 'low structure' at the site of crosslinking, there are others that have more specific structural preference. For example, *in vitro* binding studies have indicated that TRIM25 binds preferentially to purine-rich sequences within the loop of stem loop structures. I observed this same binding preference in iCLIP data, with a strong peak in frequency of 'hairpin loop' structural features at the centre of binding sites [372].

4.3.5 vRNA interaction patterns offer insight into RBP functions during infection

Determining RBPs' cellular binding preferences can inform our understanding of how and why RBPs might be engaging with vRNA. For each of the proteins whose iCLIP data I have analysed, features of their cellular binding preferences are preserved in their binding to vRNA,

providing evidence that binding is specific and functional, as opposed to being random. MATR3's main binding site within the viral genome reflects its cellular specificity. The highest intensity of signal occurs in a pyrimidine-rich tract within the unpaired end of a hairpin loop. The strongest peaks of TRIM25 binding similarly reflect both TRIM25's sequence and structure preference, occurring within AG-rich hairpin loops [372]. I also found that there was some overlap in both the binding specificity and genome binding sites of TRIM25 and TRIM56. These are two antiviral TRIM proteins which are thought to bind RNA through very different modalities. TRIM25 binds through its PRY/SPRY and coiled-coiled domains [372]. The exact binding modality of TRIM56 is not known but, based on homology to related TRIM proteins, it is likely that it is binding through its C-terminal NHL domain [398]. The fact that these two proteins are binding similarly to vRNA raises the interesting question as to whether this binding might represent a convergent evolutionary mechanism. This is something that will need to be explored further in future.

For proteins without clear sequence specificity, vRNA binding reflects other aspects of their binding preference. For example, RTCB and DDX1, which both form part of the tRNA ligase complex, bind at the 5' end of the SINV sgRNA. This is consistent with the fact that they preferentially bind to the 5' UTR and the very start of the coding sequence on cellular RNAs. What drives this binding preference is not clear as motif enrichment analysis did not indicate sequence specificity, nor did structural analysis indicate any specific feature enrichment. It might be that specificity is driven instead by PPIs. XRN1 is found preferentially binding to the 3' UTR of cellular RNAs during infection. This is a reflection of its role as a 5' to 3' exonuclease [50]. We hypothesise that XRN1 is 'captured' by iCLIP at the 3' end of RNA as this region is more highly structured than the coding sequence and will stall XRN1's degradation. The presence of XRN1 adjacent to viral replication factories, in addition to the abundance of viral RNA in the cell, make it highly likely that XRN1 will degrade some viral RNA (Ruscica et al, in preparation). XRN1's 3' end binding in the viral genome could be a reflection of this degradation. In all of these examples, RBPs' vRNA binding profiles reflect aspects of their biology that have been determined by orthogonal methods. This offers valuable confirmation that the peaks we see are instances of specific RBP binding, rather than artifacts of vRNA abundance. This is further confirmed by the variation in profiles seen between RBPs.

Such specific patterns of binding also give us confidence in developing hypotheses as to what the role of binding might be. The fact that MATR3 binds to the 51nt 5' CSE is an

important consideration in assessing its function. The 51nt 5' CSE consists of a set of two hairpins that enhance synthesis of vRNA in both mosquito and vertebrate cells, based on mutational analysis [399, 400]. It is a highly conserved structure found across all alphaviruses. In a study comparing the SHAPE-MaP profiles of SINV and VEEV, the 5' CSE was one of only three secondary structure features that was conserved between the two viruses, despite both genomes having high levels of overall structure [371]. Since the structure and sequence of this region is so highly conserved, it would be interesting to establish whether MATR3 binds in a similar way to the genomes of other alphaviruses. iCLIP in SFV has been optimised and successfully carried out by Dr Alvarez in the Castello lab, making this a viable next step.

A more challenging but equally interesting future direction would be to confirm whether MATR3 binding is being driven by sequence preference. To this end, it would be necessary to establish how MATR3's interaction with vRNA is altered by 5' CSE mutations. One approach to this would be to perform ITC on a sequence of just the hairpin (WT or mutant) using purified MATR3 RBDs. However, since the whole protein cannot be purified and since this occurs outside of the cellular context, it would also be valuable to determine how binding to the viral genome is altered by mutation during the course of a natural infection cycle. Also, in light of the 'bind and spread' model, whereby a single high affinity MATR3 binding site initiates non-specific coating of a region of RNA, it would be interesting to see whether loss of this peak in binding led to more global reduction of binding across the region [332]. One of the challenges of generating mutants within this feature is that almost any change to the coding sequence results in non-viable virus [401]. It would be necessary to determine whether it was possible to create a mutant that preserved coding sequence and structure, while removing the pyrimidine tract that MATR3 is binding to. To circumvent this challenge in studying 5' structural elements, one group introduced silent mutations into the 51nt 5' CSE of the VEEV genome to abolish its secondary structure and added in the 5' promoter sequences upstream of nsP1, thereby removing the sequence constraints associated with nsP1 coding [402]. A construct similar to this could be employed to assess the effect of mutations on MATR3 binding.

4.3.6 Developing working models of MATR3's antiviral function

MATR3 is thought to regulate AS by physically coating splice sites to prevent recognition by splicing factors [332]. One possibility is that MATR3's binding to the 51nt 5' CSE has a similar effect, blocking the structure from carrying out its role in promoting infection. In this model,

the presence of so many 5' end binding proteins would be a reflection of proximity, rather than functionality. Alternatively, the 5' CSE might simply represent a highly conserved binding platform to allow for MATR3 to associate with the vRNA. This association might be important in promoting its interaction with other 5' end cellular proteins, like IFITs 1 and 3, for which it could act as an accessory protein.

There are several follow-up experiments that could help us to differentiate between these possible mechanisms. Firstly, if we were to perform the XL-MS experiment proposed above, we could try to establish the nature of MATR3's interaction with nsP1. Recent cryoEM analyses of CHIKV's nsP1 have demonstrated that it forms dodecomeric pores in the membranes of viral replication factories through which newly synthesised vRNA is extruded [403]. It tethers the rest of the replicase complex to the inner membrane, while mediating reactions with cellular proteins at its outer interface. The fact that MATR3 interacts most strongly with nsP1, of all the non-structural proteins, is an initial indication that it might be predominantly found outside the replication factories. Additionally, if we were able to identify where on nsP1 MATR3 binds, we could determine if it was a region found on the inner or outer membrane surface.

Establishing the nature of MATR3's interaction with nsP1 could also offer direct indications as to its antiviral role. My data suggests that nsP1 is a strongly enriched RNA-independent interactor of MATR3 following infection. It is worth noting that, as with many of MATR3's cellular interactors, nsP1 is involved in 5' end processing, having both capping and decapping activity [404, 405]. Therefore, we cannot rule out that MATR3's interaction with nsP1 might represent an aspect of its function, rather than be a function of proximity.

Another important avenue of research is to directly establish whether replication or translation is being inhibited. These two processes are intimately linked from the earliest stages of the SINV lifecycle. When the viral genome enters the cell, it undergoes an initial round of translation to form the replicase complex, without which replication cannot take place [406]. Similarly, replication defects will inevitably result in translation defects by reducing the available pool of viral mRNA. Replicon systems are a valuable way to assess the relative contributions of these processes to a given phenotype. Replicons probing translation and replication independently, using luciferase as a read-out of activity, have been developed for alphaviruses [407, 408]. Determining which of these phases is more strongly blocked by MATR3 KD would be key to establishing its antiviral role.

In conclusion, in this chapter, I have presented the results of protein and RNA interaction

analyses for MATR3. These results indicate that MATR3 alters its RNA and protein interaction partners following infection. By probing these altered interactions, we can begin to generate testable hypotheses as to what its antiviral function might be.

5 General discussion and future directions

5.1 Shifting from an 'abundance' to an 'activity'-driven understanding of IFN-I regulation

Understanding how the IFN-I response establishes an antiviral state has been a focus of research since IFN-I's initial discovery in the 1950s. The primary goal of this research has been to identify and characterise genes that are transcriptionally induced by IFN-I (ISGs). However, recent studies have demonstrated that IFN-I-driven regulation stretches well beyond transcription. Changes in transcript processing [136, 256], translation control [409], PPIs [70], and post-translational modification (beyond the JAK/STAT signalling cascade) [279], have all been identified as products of the IFN-I response. All of these factors contribute to shaping ISG outcomes, but can also alter the activity of non-ISGs to endow them with antiviral properties. Altering protein activity directly, without the need for transcription and translation, is a powerful mechanism to save on energy-costly bio-synthetic processes, while increasing the speed at which antiviral defences can be triggered.

In the present study, I add a new layer to this growing repertoire of activity-driven effects by finding that the RBPome is extensively remodelled upon IFN-I treatment. IR-RBPs included a handful of known ISGs, along with many that had not previously been classified as 'IFN regulated'. Among these were many proteins that had been previously linked to viral infection in literature, and/or that were conserved interactors of vRNA. This indicated that IR-RBPs are enriched in known regulators of infection and could offer a valuable resource in identifying novel antiviral factors. I confirmed this through a small-scale screen, in which I identified several antiviral candidates, including ATXN2, JUP, RBM39, SUGP1, THRAP3, U2AF2, and MATR3.

No layer of IFN-I-driven regulation acts in isolation. Alterations in proteins' RNA-binding activity will almost certainly have consequences for the other activity-level changes described above. For example, many IR-RBPs have been linked to splicing regulation, a process which is central in shaping the nature of IFN-I-induced transcriptome and proteome [256]. Additionally, several IR-RBPs are protein modifying enzymes that could be contributing to changes in the PTM landscape of IFN-I-treated cells [279]. These enzymes include the E3 ubiquitin ligases TRIM25 and TRIM56, the E3 ISG15 protein ligase HERC5, the mono-ADP-ribosyltransferase PARP12, and the kinase PKR. An exciting possibility is that modulation of RNA-binding by IFN-I could be a means of 'riboregulating' these proteins' enzymatic activity or altering the targets

of this activity. Indeed, both PKR and TRIM25 are known targets of 'riboregulation' during infection [91, 372]. I am planning to assess whether IFN-driven change in TRIM56's RNA binding have any effect on its ubiquitination targets when I return to the lab. In collaboration with Dr Lucia Alvarez, I have designed and generated a putative RNA-binding-null mutant of TRIM56. I plan to compare the IFN-I-specific ubiquitination targets of TRIM56 with and without RNA-binding activity using a recently developed ubiquitin-specific proximity-based labelling approach [410]. This will allow me to differentiate between activity that is dependent on TRIM56's RNA binding and that which isn't.

In Chapter 3, I provided evidence that changes in RNA-binding activity might not only lead to changes in PTMs, but might be also be a direct result of such changes. I identified several putative instances of RBPs being regulated through individual phosphosites and confirmed this regulation for MATR3. These examples speak to complex and integrated networks of IFN-driven regulation, stretching well beyond ISG induction. Viewing these layers of regulation in isolation risks losing out on valuable insight, such as the reciprocal regulatory dynamic between PTMs and RNA-binding activity. My expectation is that, as our understanding of each individual layer improves, so will our ability to view these layers as parts of an integrated whole. An exciting future direction would be to combine RIC not only with PTM profiling but with profiling of other layers of activity regulation, such as enzymatic activity probes, to start to assess how the pieces of this puzzle fit together.

5.2 Expanding the repertoire of IFN-regulated RBPomes

Analyses of IFN-I-driven transcriptomic changes performed using different cell types, species and IFN subtypes have demonstrated that there are both 'core' and context-dependent ISGs [193, 198]. It is likely that these classifications can also be applied to IR-RBPs. In this study, I assessed changes in RNA-binding activity in a single cell line (HEK293) using a single IFN subtype (IFNα2). Establishing how IR-RBPomes vary across different contexts will be a highly informative way of establishing core IFN-driven regulatory events.

There are 17 IFN-I subtypes in total, all of which bind to and form structurally similar ternary complexes with IFNAR [411]. However, differences in receptor affinity between subtypes can result in quantitative and qualitative differences in the IFN-I response [412]. Schreiber and Piehler have proposed a model in which there are both 'robust' and 'tunable' aspects of the IFN-I response [52]. They suggest that the 'antiviral' aspects of the response are 'robust'

(consistent across subtypes) and are driven by the canonical JAK/STAT cascade. In contrast, antiproliferative and immunomodulatory functions are considered 'tunable' and are driven by non-canonical signalling pathways, such as MAPK and PI3K [413]. However, there is also evidence for 'tunability' of antiviral activity; a recent study of inhibition of SARS-CoV2 indicated a spectrum of antiviral effects across different subtypes, driven by substantial differences in core ISG induction [205]. Given the inherently plastic nature of RNA-binding activity, we might expect the IR-RBPome to be a particularly 'tunable' aspect of the response. To establish this, a clear next step would be to assess whether the IR-RBPome differs between low, mid and high binding affinity IFN-Is, and whether there are core IR-RBPs that are consistent across affinity levels. Comparing the IR-RBPomes induced by type I and type III IFNs would also be informative. Although IFN-IIIs bind to the IFNLR1 and IL10Rβ heterodimer, rather than IFNAR, both subtypes induce a JAK/STAT signalling cascade that leads to ISGF3 formation and upregulation of similar sets of ISGs [414]. Despite these similarities, there is evidence for differences in non-canonical signalling cascades induced by IFNs I and III, which might result in differences in their IR-RBPomes [415].

While IFN-I and IFN-III induce relatively similar responses, their target cell types differ substantially as a result of differences in receptor expression. While, IFNAR is expressed ubiquitously, IFNLR1/ IL10Rβ are almost exclusively expressed by cells at mucosal barrier sites [414]. Differences in gene expression patterns can also result in more subtle differences in IFN responsiveness, even to the same subtype. Levels of components of the JAK/STAT signalling cascade can quantitatively and qualitatively alter the nature of the IFN-I response [12]. For example, cells expressing high levels of STAT1 will have a higher frequency of STAT1 homodimerisation, which in turn results in a more pro-inflammatory response profile [59]. Activity-level regulation, such as regulation of RNA-binding, is reliant on the pool of pre-existing proteins in the cell so is likely to be highly cell type dependent. A clear example of this was seen in a study of proteins binding to SARS-CoV2 RNA in Huh7 and VeroE6 cells. 83 interaction partners were common to both cell types, while 146 and 83 were specific to Huh7 and VeroE6 cells, respectively, something that might contribute to different infection dynamics in the two cell types [416]. By studying IR-RBPomes across different cell lines, we could broaden the scope of context-specific IR-RBPs, and better understand heterogeneity in viral susceptibility across cell types and tissues.

Many IR-RBPs identified in this study directly interact with vRNA. Modulation of the binding

activity of these proteins by IFN-I could be a way of increasing (antiviral proteins) or decreasing (dependency factors) engagement with vRNA. If time had allowed, my next step in this project would have been to establish how the vRNA interactome is altered by IFN-I pre-treatment as a way of determining whether IFN-I is able to alter the engagement of IR-RBPs with vRNA. This has been attempted before in the context of early CHIKV infection. However, this study did not perform a quantitative comparison of IFN-dependent interactions, so insights that could be drawn from it were limited [49]. A reason that this quantitative comparison was lacking might have been that the levels of vRNA captured in each condition were substantially different due to IFN-driven restriction. Therefore, careful optimisation of any future approach will be necessary to ensure an appropriate balance of IFN-I response and viral infection. Viral infection will be reduced even by low doses of IFN-I, so it might be necessary to measure vRNA levels by RT-qPCR and adjust sample amounts prior to downstream processing. Having IFN-dependent vRNA interactome data across viruses from different families would be a powerful way of identifying conserved and family-specific IFN-driven antiviral mechanisms. It would also offer insight into the role of IFN-driven regulation of RNA-binding activity in shaping vRNA interactomes.

5.3 Beyond phosphorylation: probing PTMs as a general mechanism for IR-RBP control

Phosphorylation is a key regulator of both the IFN-I response [279] and RNA-binding activity [284]. However, to my knowledge, there are no recorded instances of IFN-I directly regulating a protein's RNA-binding activity through altering its phosphorylation state. In this study, I identified several putative instances of such regulation. By comparing my original IFN RIC data with data generated from the same samples following phosphoenrichment, I was able to visualise the relationship between RNA-binding activity and phosphosite abundance for individual RBPs. For several of these proteins, including MATR3, MAP4, and GIGYF2, I found that these two features were anti-correlated, strongly hinting that phosphorylation is directly regulating RNA-binding activity. I was able to validate that dephosphorylation of S188 in MATR3 increased its RNA-binding activity using a phospho-null mutant, demonstrating the value of combining RIC with phosphoenrichment in identifying regulatory phosphosites. As discussed in Chapter 3, a next step in this line of study will be to repeat this experiment at scale to increase the number of phosphosites that can be detected. However, phosphorylation

is not the only PTM with dual regulatory roles in the IFN-I response and RNA-binding. Another exciting future direction will be to combine IFN RIC with enrichment strategies for other PTMs to further enhance our understanding of the post-translational regulatory landscape of the IR-RBPome.

A group of modifications with particularly strong links to both RNA-binding regulation and the IFN-I response are the ubiquitin-like (UBL) modifiers, including ubiquitination, ISGylation and SUMOylation. These are small proteins that vary in sequence but that share a beta-grasp structure, as well as a GG motif at their C termini, which allows for their protein conjugation [417]. While UBLs have important commonalities, they serve very different roles in innate immunity. Ubiquitination plays precise and specialised roles in innate immune signalling [418]. ISGylation, on the other hand, is believed to globally modify newly translated proteins during the IFN-I response, as discussed in Chapter 1 [68]. While its exact role in the antiviral response is still not clear, loss of ISG15 in mice results in severe defects in anti-viral immunity [419]. In contrast to ISG15, when SUMOylation is lost in mice, there is an upregulation in inflammatory and IFN-dependent signalling [420]. This is consistent with the observed function of SUMOylation in modulating innate immunity, although it has also been found to stabilise some antiviral proteins [421].

All three of these UBLs are also known to modulate RNA-binding activity. Ubiquitination and SUMOylation are frequently found to modify RBPs [285]. There are diverse instances of SUMO-driven regulation of RNA metabolism, even in the context of antiviral immune response [422]. For example, SUMOylation of ADAR1 leads to changes in its RNA editing activity [423]. Many ISGylated proteins are also involved in cellular RNA metabolic processes and ISGylation has been found to directly regulate the RNA-binding activity of viral proteins, such as IAV's NS1 [424].

UBLs represent exciting candidates for regulators of the IR-RBPome. However, isolating and distinguishing between UBLs is non-trivial. Ubiquitin, ISG15, and another UBL, NEDD8, all yield the same diglycine remnant upon trypsin digestion. This means that enrichment of diglycine modified peptides, which is the most efficient method for identifying UBL modifications, cannot differentiate between them [425]. SUMO, on the other hand, is digested into a very large remnant, which prevents SUMOylated peptides from being accurately identified by MS [426]. Both of these challenges can be addressed through the use of modified cell lines or transient transfection of modified proteins [417]. For example, to differentiate between UBLs

that produce diglycine remnants, tagged proteins can be introduced to allow for their specific enrichment. To circumvent the challenge of SUMO's large digestion fragment, constructs have been designed with mutations that result in shorter digestion remnants that can be identified easily by MS [417]. However, introducing these additional complications into the already complex IFN RIC protocol might prove unfeasible. Alternative approaches rely on creative use of proteases to create differentially digested peptides [427–429]. These can be individually enriched using remnant-specific antibodies. Such an approach might prove more achievable and would allow for sequential enrichment of multiple modifications from the same RIC samples.

UBLs are just some of the many PTMs that have the potential to modulate RNA-binding activity in the IFN-I response. These include: acetylation, methylation, PARylation, oxidation, and nitrosylation [430]. Since any one IR-RBP might be regulated by any combination of these modifications, building up a complete picture of post-translational control in the IFN-I response poses challenges [431]. However, with each new PTM that is analysed, we will move closer to understanding the factors that shape the IR-RBPome.

5.4 Mapping viral ribonucleoproteins: a new frontier in host-pathogen interactions

vRNA interactome studies have provided insight into the wide range of cellular RBPs contributing vRNP composition [192]. While these methods can tell us *whether* a protein is binding to vRNA, they do not give us an indication as to *where* it is binding. Therefore, complementary methods are required to define where RBPs are binding with the ultimate goals of establishing how and why they are interacting with vRNA.

In Chapter 4, I demonstrated the potential of iCLIP to enhance our understanding of vRNA interacting proteins. I identified a specific peak of MATR3 binding within the 51nt 5' CSE structure of SINV gRNA, which could relate to its antiviral function. Additionally, I was able to compare the vRNA binding profiles of five other RBPs. Each of these profiles has a story to tell. From the profiles of RTCB and DDX1, we see a pattern of binding consistent with sgRNA translation control. From TRIM25 and TRIM56, we see indications of convergent evolution in the binding patterns of two antiviral TRIM proteins. And from XRN1, we observe a pattern of binding consistent with its exoribonuclease activity, indicating that at least some of the vast pool of vRNA in the cell is being targeted for degradation.

The value of CLIP-based methods in understanding vRNA interaction dynamics has also

been demonstrated for viruses beyond SINV, including HIV-1 [272, 432], IAV [433, 434], and SARS-CoV2 [435]. Each of these datasets enriches our understanding of an individual protein in the context of a specific virus. For example, iCLIP allowed authors of one study to track the changing specificity of Gag over the course of the HIV-1 infection cycle, providing valuable insights into genome packaging mechanisms [436]. However, to begin to understand vRNP dynamics more holistically, we will need many more datasets across different vRNA interacting proteins and different viral genomes. iCLIP is not a method suited to such an undertaking. Each iCLIP experiment is extremely resource- and time-intensive and it is an inherently low throughput approach. In addition, much of the data generated in each experiment, while potentially interesting, is superfluous for studying vRNA binding. Therefore, there is a need to develop new strategies for high-throughput, targeted characterisation of vRNP organisation.

There are already methods available that would allow us to explore certain aspects of vRNP architecture. One recently published method, TREX, sets out to assess the interactome of specific regions of RNAs [437]. In this method, organic phase separation is combined with specific probe binding and RNase H digestion to release and isolate RBPs bound to the sequence of interest. This was used to define the interactome of specific regions of the 45S ribosomal RNA, including the 361 base 3' ETS region [437]. While this would not provide the specificity of CLIP-based methods, it could be used to identify which region of the viral genome a given protein is predominantly binding to. For example, we could assess whether MATR3 interaction partners are similarly localised to MATR3 on the SINV genome or determine whether the pattern of localisation of MATR3 is conserved across different alphaviruses and even across other viral families.

A drawback of this approach is that assessing binding windows across the entire length of a viral genome would require a huge amount of sample processing and MS run time. One way of reducing this would be to focus on regions that are known 'hotspots' of protein binding or regulation. RNP-MaP is another recently developed approach that would allow us to identify these hotspots [438]. RNP-MaP is designed to identify the overall distribution of binding sites, regardless of the proteins from which they are derived, across an RNA of interest. It relies on a combination of RNA-protein chemical crosslinking and a low fidelity reverse transcriptase that incorporates off-template nucleotides when sites of crosslinking are encountered. This reverse transcriptase will read through these crosslinked sites, meaning that multiple protein binding sites can be registered in a given read. RNP-MaP was used to establish the distribution of

binding sites across the length of the 20kb non-coding RNA, XIST. It identified binding hotspots in regions that are known to be important in regulation, as well as previously uncharacterised regions [438]. We could use this approach to understand the overall distribution of protein binding to vRNA. It would allow us to determine whether proteins distribute evenly across the length of the vRNA or are enriched in specific regions that might have regulatory significance. If there are specific regions that are highly enriched in binding sites, these could be targeted in a TREX-style experiment to assess protein composition.

While this application of RNP-MaP could offer valuable insight into the viral RNA binding landscape, the method is most powerful when combined with CLIP data so that binding sites can be assigned to specific proteins. Weidmann and colleagues aligned the global binding profile of XIST RNA with the binding sites of specific RBPs derived from publicly available eCLIP data [438]. Because each read can contain multiple instances of crosslinking, they were able to correlate binding of specific RBPs. They found that there were distinct 'communities' of functionally-related proteins whose binding sites clustered together and which frequently bound to the same RNA molecule. While a method like TREX could provide a sense of protein groupings, understanding combinatorial binding patterns on vRNA would still require CLIP-derived binding profiles on viral genomes.

An alternative to replacing CLIP would be to improve its efficiency, cost-effectiveness and multiplexing capability. One such effort has given rise to antibody barcode eCLIP [439]. In this method, unique DNA barcodes are tethered to antibodies against the proteins of interest. During the eCLIP protocol, the DNA barcodes are transferred from the antibody to the RNA associated with the immunoprecipitated protein. In this way, eCLIP can be performed for multiple proteins in the same sample, reducing requirements for starting material and increasing the rate at which datasets can be generated. Alternatively, machine learning-based approaches could be employed to predict binding sites in new contexts based on data that is already available [164]. However, applying these approaches to viral genomes relies on the assumption that the mode in which RBPs engage with vRNA is the same as with cellular RNA. Without a better understanding of vRNA binding properties of cellular RBPs (which must necessarily come from experimental data), this assumption feels premature. Hopefully, with continued development of both experimental and computational strategies, the global study of vRNP architecture and dynamics will begin to yield exciting results that will inform our understanding of the regulatory roles played by IR-RBPs.

5.5 Summary

In this thesis, I have outlined my contribution towards defining and understanding a new layer of IFN-I driven regulation. Using RIC, I found that the RNA-binding activity of hundreds of proteins are altered by the IFN-I response, independently of protein abundance. I identified new players in viral infection among these IR-RBPs. Each of these candidates warrant follow-up to assess their roles in infection and regulatory mechanisms. To globally probe one layer of potential regulation, I also established a new approach for assessing the role of PTMs in modulating the RBPome. By performing phosphoenrichment of RIC samples, I was able to establish how changes in phosphorylation directly related to changes in RNA-binding activity. This strategy allowed me to identify several exciting candidates for further study, including the novel antiviral protein, MATR3. I found that MATR3's RNA-binding activity was directly regulated through its phosphorylation state in response to IFN-I. This example highlights the power of combinatorial profiling in defining regulatory networks. An important next step will be to combine IFN RIC with enrichment methods for other PTMs to build a better understanding of IFN-I induced regulation. To further characterise MATR3's role in infection, I assessed how its RNA and protein interactomes were shaped by SINV infection. I found a pattern of MATR3 behaviour that hinted at a potential role in viral translation control, something that will need to be assessed in future. Together, these data indicate that the IR-RBPome is a valuable resource in identifying new players in viral infection. They speak to the importance of looking beyond ISGs in thinking about the factors driving infection and highlight complex regulatory networks shaping the antiviral state that have yet to be uncovered.

6 Materials

6.1 Reagents & Consumables

Reagent	Supplier	Identifier
10x Tris/Glycine/SDS Running buffer	BioRad	1610732
2X Phusion HF PCR Master mix	NEB	M0531L
4–15% Mini-PROTEAN TGX Precast Protein Gel	Bio-Rad	4561084
5' Deadenylase	NEB	M0331S
96-well, Cell Culture-Treated, Black Flat-Bottom Microplate	Greiner	10369081
Acetic acid glacial	Fisher Chemical	A/0360/PB17
Acetonitrile (Mass Spec Grade)	Fisher Scientific	15492857
AEBSF	BioChemica	A14210100
Ammonium persulfate	Thermo Fisher	17874
Benzonase Nuclease	Millipore	E1014
Colour prestained protein standard, broad range	NEB	P7719S
DAPI	Invitrogen	D1306
DH5α E. coli	NEB	C2987H
DMEM	Gibco	11965092
DMEM	Thermo Fisher	11995065
DMEM (no phenol red)	Thermo Fisher	21063029
DMSO	Sigma-Aldrich	D2260

Table 2: Reagents & Consumables

Reagent	Supplier	Identifier
DpnI	NEB	R0176S
DTT	Sigma Aldrich	D1532
Dynabeads MyOne Silane	Thermo Fisher	37002D
EDTA	Millipore	324503
Ethanol absolute		
EvaGreen	Biotium	31000
FastAP alkaline phosphatase	Thermo Fisher	EF0654
FBS	Gibco	16000044
Formic acid (Mass Spec Grade)	Sigma	00940
GFP-Trap agarose beads	Chromotek	gta-20
GFP-Trap Multiwell Plate	ChromoTek	gtp
Glass Circle Coverslips	Fisherbrand	12333138
HiScribe T7 Arca	NEB	E2065
Hygromycin B	Thermo Fisher	J60681.MD
iBLOT2 nitrocellulose membrane	Thermo Fisher	IB23001
IGEPAL CA-630	Sigma-Aldrich	I8896
Lithium chloride	Sigma-Aldrich	L9650
Lithium-dodecyl sulfate	Sigma-Aldrich	L4632
Luna Universal One-step RT-qPCR kit	NEB	E3005E
Lys-C	Wako	129-02541
MagReSyn Zr-IMAC beads	Resyn Biosciences	MR-ZRM002
NuPage LDS Sample Buffer	Invitrogen	NP0007
NuPAGETM4-12% Bis-Tris 1mm gel	Invitrogen	NP0322BOX
Oasis MAX 96-well µElution Plate	Waters	186001829
Oligo(dT)25 beads	NEB	S1419S
Opti-MEM reduced serum Medium	Gibco	31985062
Paraformaldehyde, 16% w/v aq. soln.	Thermo Fisher	043368.9M
PBS	Gibco	10010023
Penicillin/streptomycin	Gibco	15140122

Reagent	Supplier	Identifier
PFU Turbo DNA polymerase	Agilent	600252
Phenol:Chloroform:Isoamyl Alcohol	Sigma-Aldrich	P3803
Pierce control agarose resin	Thermo Fisher	26150
PNK	New England Biolabs	M0201L
poly-L-Lysine	Sigma-Aldrich	P4832
ProNex	Promega	NG2001
Protein Lobind tubes 1.5 ml	Eppendorf	0030108116
Proteinase K	Roche	3115828001
Qubit 1X dsDNA HS Assay Kit	Invitrogen	Q33230
RecJf endonuclease	New England Biolabs	M0264S
Recombinant Human IFNα2	Novus Biologicals	NBP2-34971
RiboLock RNase Inhibitor	Thermo Fisher	EO0381
Ribonucleoside Vanadyl Complex	NEB	S1402S
RNase I	Thermo Fisher	AM2294
SilverQuest staining kit	Invitrogen	LC6070
SlowFade Diamond antifade mountant	Invitrogen	S36967
Sodium Pyruvate	Thermo Fisher	11360070
SpeedBeads Magnetic Carboxylate Modified Particles	Sigma-Aldrich	GE651521 05050250
Superscript IV	Thermo Fisher	18090010
T4 DNA ligase	NEB	M0202S
T4 RNA ligase	Thermo Fisher	EL0021
Triton X-100	Promega	H5141
Trypan Blue Stain (0.4%)	Invitrogen	T10282
TrypLE Express Enzyme	Gibco	12604013
Trypsin Gold (Mass Spec Grade)	Promega	V5280
TurboDNase	Thermo Fisher	AM2238
Tween-20	Sigma Aldrich	P1379
XhoI	NEB	R0146S

6.2 In-House Buffers

Buffer	Composition
DF 5x dilution buffer	1.25M NaCl, 100mM Tris pH 7.5
DF blocking buffer	200mM LiCl, 1mM EDTA, 20mM Tris HCl pH 7.5, 0.01% IGEPAL, 100ug/ml yeast tRNA, 100ug/ml BSA
DF hybridisation buffer	500mM LiCl, 1mM EDTA, 20mM Tris HCl pH 7.5, 0.01% IGEPAL, 0.05% LiDS, 5mM DTT
DF lysis buffer	500mM KCl, 5mM MgCl2, 10mM Tris pH 7.5, 1% IGEPAL, 1mM DTT
DF medium salt buffer	250mM NaCl, 1mM MgCl2, 20mM Tris HCl pH 7.5, 0.05% IGEPAL, 0.025% SDS
DF wash buffer 1	500mM LiCl, 1mM EDTA, 20mM Tris HCl pH 7.5, 0.01% IGEPAL, 0.01% LiDS, 5mM DTT
DF wash buffer 2	200mM LiCl, 1mM EDTA, 20mM Tris HCl pH 7.5, 0.01% IGEPAL, 0.01% LiDS, 5mM DTT
iCLIP 5x PNK buffer	350 mM Tris HCl pH 6.5, 50 mM MgCl2, 25 mM DTT
iCLIP high salt buffer	500 mM NaCl, 1 mM MgCl2, 20 mM Tris HCl pH 7.5, 0.05% NP40, 0.1% SDS, 1 mM DTT, 1x AEBSF
iCLIP medium salt buffer	250 mM NaCl, 20 mM Tris HCl pH 7.5, 1 mM MgCl2, 0.05% NP40, 1x AEBSF
iCLIP PK-SDS solution	10 mM Tris HCl pH 7.4, 100 mM NaCl, 1 mM EDTA, and 0.2% SDS
iCLIP PNK buffer	20 mM Tris HCl pH 7.5, 10 mM MgCl2. 0.2% Tween-20

Table 3: In-House Buffers

Buffer	Composition
IP dilution buffer	10 mM Tris HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA
IP lysis buffer	10 mM Tris HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5 % IGEPAL
IP wash buffer	10 mM Tris HCl pH 7.5, 150 mM NaCl, 0.05 % IGEPAL, 0.5 mM EDTA
PE loading buffer	80% acetonitrile, 6% TFA
PE wash buffer 1	50% acetonitrile, 0.5% TFA, 200mM NaCl
PE wash buffer 2	50% acetonitrile, 0.1% TFA
RIC elution buffer	20 mM Tris HCl pH 7.5, 1 mM EDTA
RIC lysis buffer	500 mM LiCl, 20 mM Tris HCl pH 7.5, 0.5% LiDS, 1 mM EDTA, 5 mM DTT
RIC wash buffer 1	500 mM LiCl, 20 mM Tris HCl pH 7.5, 0.1% LiDS, 1 mM EDTA, 0.1% NP40, 5 mM DTT
RIC wash buffer 2	500 mM LiCl, 20 mM Tris HCl pH 7.5, 1 mM EDTA, 0.01% NP40, 5 mM DTT
RIC wash buffer 3	200 mM LiCl, 20 mM Tris HCl pH 7.5, 1 mM EDTA, 5 mM DTT
RIPA lysis buffer	10 mM Tris HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.1% SDS, 1% Triton X-100, 1% sodium deoxycholate

6.3 Antibodies & Dyes

Reagent	Supplier	Identifier
beta-actin primary antibody	Sigma	A1978
GFP primary antibody	chromotek	3h9-100
IRDye 680RD Donkey anti-Rabbit IgG	LI-COR Biosciences	926-68073
IRDye 680RD Goat anti-Mouse IgG	LI-COR Biosciences	926-68070
IRDye 800CW Donkey anti-Rabbit IgG	LI-COR Biosciences	926-32213
IRDye 800CW Goat anti-Human IgG	LI-COR Biosciences	926-32232
IRDye 800CW Goat anti-Rat IgG	LI-COR Biosciences	926-32219
MATR3 primary antibody	Proteintek	12202-2-AP
oligo(dT) stellaris probe	Biosearch tech	N/A
Phospho-(Ser/Thr) PKA primary antibody	Cell signalling	9621S
Phospho-STAT1 primary antibody	Cell signalling	9167S
PTBP1 primary antibody	Sigma	WH000 5725M1
RFP primary antibody	chromotek	5F8-100
SINV Capsid primary antibody	Lab of L. Carrasco	N/A
STAT1 primary antibody	Cell signalling	9176S
TRIM25 primary antibody	Abcam	ab167154

Table 4: Antibodies & Dyes

6.4 Plasmids

Plasmid	Backbone	Source
pcDNA5-FRT-TO-ATXN2-GFP	pcDNA5-FRT-TO	Biobasic
pcDNA5-FRT-TO-DAZAP1-GFP	pcDNA5-FRT-TO	Biobasic
pcDNA5-FRT-TO-FUS-GFP	pcDNA5-FRT-TO	Biobasic
pcDNA5-FRT-TO-JUP-GFP	pcDNA5-FRT-TO	Biobasic
pcDNA5-FRT-TO-MAP4-GFP	pcDNA5-FRT-TO	Biobasic
pcDNA5-FRT-TO-MATR3-GFP	pcDNA5-FRT-TO	Biobasic
pcDNA5-FRT-TO-MATR3-S188A-GFP	pcDNA5-FRT-TO	Made in this project
pcDNA5-FRT-TO-MATR3-S188D-GFP	pcDNA5-FRT-TO	Made in this project
pcDNA5-FRT-TO-MKRN2-GFP	pcDNA5-FRT-TO	Biobasic
pcDNA5-FRT-TO-RALY-GFP	pcDNA5-FRT-TO	Biobasic
pcDNA5-FRT-TO-RAVER1-GFP	pcDNA5-FRT-TO	Biobasic
pcDNA5-FRT-TO-RBM39-GFP	pcDNA5-FRT-TO	Biobasic
pcDNA5-FRT-TO-SSBP1-GFP	pcDNA5-FRT-TO	Biobasic
pcDNA5-FRT-TO-SUGP1-GFP	pcDNA5-FRT-TO	Biobasic
pcDNA5-FRT-TO-THRAP3-GFP	pcDNA5-FRT-TO	Biobasic
pcDNA5-FRT-TO-TOP2A-GFP	pcDNA5-FRT-TO	Biobasic
pcDNA5-FRT-TO-TRIM56-GFP	pcDNA5-FRT-TO	Castello lab
pcDNA5-FRT-TO-U2AF2-GFP	pcDNA5-FRT-TO	Biobasic
pOG44 Flp-recombinase expr vector	pOG44	Thermo Fisher
pT7-SVmCherry	pTE3'2J1	Castello lab
pT7-SVnsP3Scarlet	pTE3'2J1	Castello lab
pT7-SVwt	pTE3'2J1	Lab of L. Carrasco

Table 5: Plasmids

6.5 Oligonucleotides

Oligonucleotide	Sequence
L01clip2.0	NNNNATCACGNNNNNAGATCGGAAGAGCGTCGTG
L02clip2.0	NNNNCGATGTNNNNNAGATCGGAAGAGCGTCGTG
L03clip2.0	NNNNTTAGGCNNNNNAGATCGGAAGAGCGTCGTG
L04clip2.0	NNNNTGACCANNNNNAGATCGGAAGAGCGTCGTG
L05clip2.0	NNNNACAGTGNNNNNAGATCGGAAGAGCGTCGTG
L06clip2.0	NNNNGCCAATNNNNNAGATCGGAAGAGCGTCGTG
L07clip2.0	NNNNCAGATCNNNNNAGATCGGAAGAGCGTCGTG
L08clip2.0	NNNNACTTGANNNNNAGATCGGAAGAGCGTCGTG
L09clip2.0	NNNNGATCAGNNNNNAGATCGGAAGAGCGTCGTG
L10clip2.0	NNNNTAGCTTNNNNNAGATCGGAAGAGCGTCGTG
L11clip2.0	NNNNATGAGCNNNNNAGATCGGAAGAGCGTCGTG
L12clip2.0	NNNNCTTGTANNNNNAGATCGGAAGAGCGTCGTG
L13clip2.0	NNNNAGTCAANNNNNAGATCGGAAGAGCGTCGTG
L14clip2.0	NNNNAGTTCCNNNNNAGATCGGAAGAGCGTCGTG
L15clip2.0	NNNNATGTCANNNNNAGATCGGAAGAGCGTCGTG
L16clip2.0	NNNNCCGTCCNNNNNAGATCGGAAGAGCGTCGTG
L17clip2.0	NNNNCAACTANNNNNAGATCGGAAGAGCGTCGTG
L18clip2.0	NNNNGTCCGCNNNNNAGATCGGAAGAGCGTCGTG
L19clip2.0	NNNNGTGAAANNNNNAGATCGGAAGAGCGTCGTG
L20clip2.0	NNNNCACCGGNNNNNAGATCGGAAGAGCGTCGTG

Table 6: Oligonucleotides

Oligonucleotide	**Sequence**
L3-App	AGATCGGAAGAGCGGTTCAG
M3-S188A-Fw	TCCCCGATCGTCAAAGGCATCTCTCCGGAAGTGC
M3-S188A-Rv	GCACTTCCGGAGAGATGCCTTTGACGATCGGGGA
M3-S188D-Fw	GTCCCCGATCGTCAAAGTCATCTCTCCGGAAGTGCC
M3-S188D-Rv	GGCACTTCCGGAGAGATGACTTTGACGATCGGGGAC
P3Solexa	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTC CTGCTGAACCGCTCTTCCGATCT
P3Solexa_s	CTGAACCGCTCTTCCGATCT
P5Solexa	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTA CACGACGCTCTTCCGATCT
P5Solexa_s	ACACGACGCTCTTCCGATCT
RToligo	GGATCCTGAACCGCT

6.6 Cell lines

Cell line	Source	Modified?
HEK293	ECACC #85120602	Parental
HEK293T	Castello lab	Parental
HEK293 Flp/In T-Rex	Thermo Fisher #R78007	Parental
VeroE6	Castello lab	Parental
BHK	Castello lab	Parental
BSR-T7	Castello lab	Parental
HEK293 Flp/In T-rex TRIM56-GFP	Castello lab	Modified
HEK293 Flp/In T-rex THRAP3-GFP	Made in this project	Modified
HEK293 Flp/In T-rex MATR3-GFP	Made in this project	Modified
HEK293 Flp/In T-rex MATR3-S188A-GFP	Made in this project	Modified
HEK293 Flp/In T-rex MATR3-S188D-GFP	Made in this project	Modified
HEK293 Flp/In T-rex RBM39-GFP	Made in this project	Modified
HEK293 Flp/In T-rex ATXN2-GFP	Made in this project	Modified
HEK293 Flp/In T-rex SUGP1-GFP	Made in this project	Modified
HEK293 Flp/In T-rex RAVER1-GFP	Made in this project	Modified
HEK293 Flp/In T-rex RALY-GFP	Made in this project	Modified
HEK293 Flp/In T-rex SSBP1-GFP	Made in this project	Modified
HEK293 Flp/In T-rex JUP-GFP	Made in this project	Modified
HEK293 Flp/In T-rex U2AF2-GFP	Made in this project	Modified
HEK293 Flp/In T-rex MKRN2-GFP	Made in this project	Modified
HEK293 Flp/In T-rex FUS-GFP	Made in this project	Modified
HEK293 Flp/In T-rex DAZAP1-GFP	Made in this project	Modified
HEK293 Flp/In T-rex MAP4-GFP	Made in this project	Modified
HEK293 Flp/In T-rex TOP2A-GFP	Made in this project	Modified
HEK293 Flp/In T-rex GIGYF2-GFP	Made in this project	Modified

Table 7: Cell Lines

6.7 Instruments & Equipment

Instrument	Supplier
Clariostar platereader	BMG LABTECH
Countess II FL Automated Cell Counter	Thermo Fisher
EVOS M5000	Thermo Fisher
Extraction Plate Manifold for Oasis 96-Well Plates	Waters
GelDoc Imaging System	BioRad
NanoDrop Microvolume Spectrophotometer	Thermo Fisher
Odyssey CLx Imaging System	LiCor
Orbitrap Exploris	Thermo Fisher
Orbitrap Fusion Ascend Tribrid Mass Spectrometer	Thermo Fisher
Orbitrap Fusion Lumos Tribrid Mass Spectrometer	Thermo Fisher
PCR machine	ABI
PowerPac Basic Power Supply	BioRad
QuantStudio real-time PCR system	Applied Biosystems
Qubit Fluorometer	Thermo Fisher
Transblot Turbo Transfer System	BioRad
254nm UV Crosslinker	Roth Selection

Table 8: Instruments & Equipment

7 Methods

7.1 Cell Biology

7.1.1 Maintenance of Cells

All cells were kept at 37°C with 5% CO_2. Cells were maintained in Dulbecco's modified eagle medium (DMEM) with 10% fetal bovine serum (FBS) and 1x penicillin-streptomycin. In addition to standard culture conditions, HEK293 Flp/In T-REx stable cell lines were kept in 150 µg mL^{-1} hygromycin B and 5 µg mL^{-1} blasticidin. Cells were passaged regularly to maintain a confluency of $<$90%. To split cells, media was removed and cells were washed once with 1x PBS. 1x TrypLE Express enzyme was then added and cells were incubated for 2 min at 37°C with 5% CO_2. Cells were resuspended in 10% FBS DMEM and seeded at the desired density. When required, cells were counted Trypan Blue Stain (0.4%) and the Countess II FL automated cell counter.

7.1.2 Stable Transfection of HEK293 Flip/In T-REx

2×10^6 HEK293 Flip/In T-REx parental cells were seeded in a T25 flask in DMEM (10% FBS, 1x P/S). 24 hours after seeding and 1 hour before transfection, tissue culture media was changed to fresh DMEM (10% FBS, 1x P/S). 5.33 µg pOG44 Flp-recombinase expression vector and 0.67 µg (9:1 ratio) were added to 600 µL optiMEM and mixed gently. 18 µL X-tremeGENE 360 (equilibrated to room temperature) was then added (1:3 ratio) and incubated at room temperature for 15 min. Transfection complex was then added dropwise to the cells. After 24 hours, media was changed to fresh DMEM (10% FBS, 1x P/S). After a further 24 hours, cells were split into a T75 flask and 150 µg mL^{-1} hygromycin B was added to cell culture media.

7.1.3 siRNA Knockdowns

1.5×10^5 HEK293 Flip/In T-REx parental cells per well were seeded in a 24 well plate in 1ml DMEM (10% FBS, 1x P/S). 24 hours after seeding, 500ul media was removed from each well and replaced with 500 µL serum-free DMEM. Cells were incubated for 1 hour more prior to transfection. Transfections were performed using Lipofectamine RNAiMAX transfection reagent. Solution 1 (3 µL RNAiMAX & 50 µL OptiMEM) and solution 2 (1 µL 10 µM siRNA & 50 µL OptiMEM) were prepared separately then combined. Combined solutions were incubated

for 5 minutes at RT before being pipetted drop-wise onto cells. Cells were incubated for 24 hours before being infected with SINV at an MOI of 0.1. For infection, 750 µL media was removed from each well and 250 µL of serum-free media (with virus) was added. Cells were incubated for 1 hour. Media was then replaced with fresh DMEM (5% FBS). After a further 18 hours, samples were harvested in RIPA buffer and processed for western blot. Protein concentrations were measured using Qubit Protein BR assay and loading volumes were normalised to the lowest protein concentration.

7.1.4 Generation of viruses

SINV was either generated directly from the plasmid or through expansion of a pre-existing virus stock. For production from plasmid, the plasmid was first linearised by restriction digestion with XhoI.

Component	Amount
DNA	1 µg
10X rCutSmart Buffer	5 µL (1X)
XhoI	1 µL
Nuclease Free Water	Up to 50 µL

Table 9: Linearisation reaction mix

Plasmid was incubated in the above mix for 15 minutes at 37°C, then at 65°C for 20 minutes. mRNA was then synthesised by incubation of linearised plasmid at 37°C for 30 minutes with T7 RNA polymerase mix.

Component	Amount
DNA	1 µg
2X ARCA/NTP Mix	10 µL
T7 RNA Polymerase Mix	2 µL
Nuclease Free Water	Up to 20 µL

Table 10: T7 RNA polymerase mix

DNase I was then added and the mixture was incubated for 15 minutes at 37°C. RNA was purified using RNeasy Mini Kit and quantified using nanodrop. 2.5 µg mRNA was transfected into BHK21 cells at 70-90% confluency in a 10cm dish. Lipofectamine 3000 kit was used for

transfection. Tube 1 and Tube 2 were prepared, as per Table 11 and Table 12, mixed, and incubated at RT for 10-15 minutes. RNA-lipid complex was then pipetted directly onto cells.

Component	**Amount**
Opti-MEM Medium	750 µL
Lipofectamine 3000 Reagent	45 µL DNA

Table 11: Lipofectamine 3000 transfection Tube 1

Component	**Amount**
Opti-MEM Medium	750 µL
RNA	2.5 µg
P3000 Reagent	2 µL μg^{-1} DNA

Table 12: Lipofectamine 3000 transfection Tube 2

Cells were incubated for 48-72 hours (37°C, 5% CO_2). Cells were monitored regularly and pH was buffered with HEPES if media became too acidic. The supernatant was harvested when nearly all cells were showing cytopathic effects. To harvest, supernatant was transferred to a centrifuge tube and HEPES solution was added to a final concentration of 500 mM. Supernatant was centrifuged at 1000rpm for 5 minutes to remove cell debris and was then passed through a 0.45 µm filter. Single-use aliquots were stored at -80°C. For expansion stocks, a T175 flask of BHK21 cells were infected with virus at an MOI of 0.1 in 15ml DMEM. Cells were incubated for 48 hours (37°C, 5% CO_2) and harvested as described above. No more than one round of expansion was performed for tagged virus stocks.

7.1.5 Titration of viruses

SINV was titrated by plaque assay. 1x10^5 cells per well were seeded in a 24-well plate in DMEM (2% FBS, 1x P/S) and incubated overnight. Serial dilution of virus stock or viral supernatant was prepared in DMEM (2% FBS, 1x P/S). Media was carefully removed from 24-well plate and 100 µL of virus stock serial dilution was added to each well. Cells were incubated for 1 hour at 37°C, before 0.5 µL of prepared 0.6% Avicell overlay was added. Cells were gently shaken to evenly distribute the overlay. After 3 days of incubation (37°C, 5% CO_2), Avicell was removed and 500 µL 10% Formaldehyde was added to each well. Plates were incubated in the fume hood for 1 hour, then rinsed twice with PBS. Coomassie blue was added and plates were

incubated for 1 hour, before plates were washed in water and plaques were counted. Titer was calculated as follows:

$$\text{PFU/ml} = \frac{\text{Average number of plaques}}{\text{Dilution of stock} \times \text{Volume of inoculum}}$$

7.1.6 Platereader assay

2.5x10^4 cells per well were seeded on a black, clear-bottomed 96 well plate in 100 µL colourless DMEM (5% FBS, 1x P/S). When using a HEK293 Flip/In T-REx inducible cell line, media for the 'induced' condition was supplemented with 1 µg mL^{-1} Doxycycline. 24 hours after seeding/ induction, 100 µL colourless DMEM (0% FBS, 1x P/S) containing virus at the desired MOI was added. Cells were then incubated at 37°C with 5% CO_2 in a CLARIOstar fluorescence plate reader. mCherry/ mScarlet signal was measured every 15 min over a 24-hour period to give a read-out of virus replication.

7.1.7 Confocal microscopy

Coverslips were placed into the wells of a 24 well plate, washed in 100% ethanol for 30 minutes and pre-coated with poly-L-Lysine solution for 5 minutes at 37°C. 2x10^5 HEK293 FITR MATR3-eGFP cells were plated onto coverslips incubated in 10% FBS, 1% Pen/Strep DMEM supplemented with 1 µg mL^{-1} doxycycline. After 24 hours of induction, cells adherent to coverslips were washed twice with 1X PBS and fixed in 4% PFA for 30 minutes at room temperature. The cells were washed again with 1X PBS two times and permeabilized using PBS with 0.1% Triton-X100 for 10 minutes at room temperature. After permebeabilization, the coverslips were washed with 1X PBS two times and incubated with DAPI at a ratio of 1:1000 for 10 minutes at 37°C. The cells were washed twice in 1X PBS and then in nuclease-free water and mounted on glass slides using SlowFade Diamond antifade mountant. Images were acquired using a 63x/1.40 oil DIC M27 Plano Apochromat objective lens with a total magnification of 630x on a Zeiss LSM 880 Axio-Observer confocal microscope. The exciting wavelength used was 405nm for DAPI and 488nm for GFP and emission wavelengths were 498nm and 536nm, respectively.

7.2 Biochemical and molecular biology techniques

7.2.1 Western blotting

Samples were prepared by mixing with NuPAGE 4x LDS sample buffer (Invitrogen) and heated to 95°C for 5 minutes. Proteins were then loaded into 1.5mm SDS- polyacrylamide gels (10%), prepared using the TGX FastCast acrylamide kit (Bio-Rad). Proteins were separated by SDS-PAGE for 55 minutes at 180V and transferred to nitrocellulose membranes using the Trans-Blot Turbo Transfer System (Bio-Rad). Membranes were blocked for 1 hour at RT or overnight at 4°C. 5% skim milk prepared in 0.1% PBS-T was used for blocking and for preparing antibody dilutions. Membranes were incubated with primary antibody either for 1 hour at RT or overnight at 4°C. After three washes in 0.1% PBS-T, membranes were incubated with the relevant secondary antibodies for 1 hour at RT. Membranes were then washed three more times in 0.1% PBS-T before imaging. Imaging was performed on the LI-COR Odyssey Fc or LI-COR Odyssey CLx.

7.2.2 Silver staining

Samples were prepared by mixing with NuPAGE 4x LDS sample buffer (Invitrogen) and heated to 95°C for 5 minutes. Proteins were then loaded into 1.5mm SDS- polyacrylamide gels (10%), prepared using the TGX FastCast acrylamide kit (Bio-Rad). Proteins were separated by SDS-PAGE for 55 minutes at 180V and then washed briefly in ultrapure water. Staining was performed using the SilverQuest silver staining kit, according to their basic staining protocol.

7.2.3 Mutagenesis

For PCR mutagenesis, the plasmid to be mutated was replicated using primers designed with Agilent's QuikChange Primer Design tool. The PCR mix and PCR cycling conditions were as outlined in Table 13 and Table 14. After the PCR reaction was complete, 0.5 µL of DpnI was added and the mixture was incubated at 37°C for 1 hour. The end result of this reaction was directly transformed into DH5α E. coli and resulting colonies were sequenced to confirm the mutation had been introduced correctly.

Component	Concentration/Amount
Buffer	1X
dNTPs	0.2 mM
FwPrimer	2.5 pmols
RvPrimer	2.5 pmols
Plasmid	20 ng
PFU turbo DNA polymerase	5 U
H_2O	-

Table 13: PCR mutagenesis mix composition

Step	Number of Cycles	Temperature (℃)	Time
1	1	92	2 min
2	10	92	10 s
		Primer Tm - 5℃	30 s
		68	2 min/kb
3	20	92	10 s
		Primer Tm - 5℃	30 s
		68	2 min/kb plus 10 s/cycle

Table 14: PCR mutagenesis thermocycling conditions

7.3 Dual fluorescence assay

7.3.1 Sample harvesting

HEK293 FITR MATR3-eGFP (WT, S188A, and S188D) and HEK293 FITR eGFP lines were seeded in 6 well plates at a density of 1x10^6. Cells were induced with doxycyline for 16 hours. Cells were then either infected with SINV at an MOI of 1 or left uninfected. Media was changed in all wells and fresh doxycycline was added. After 18 hours of infection, media was aspirated, and cells were washed twice with PBS. 6 well plates were placed on ice blocks and UV irradiated at 254nm at an energy of 150mJ/cm^2. Cells were resuspended in PBS and spun down at 1000xg to pellet. PBS was removed and 160 µL DF lysis buffer was added. Cells were incubated on ice for 30 mins and cleared via max speed centrifugation (10 mins, 4°C). Samples were diluted in 5x DF Dilution buffer and SDS was added to a final concentration of 0.05%.

7.3.2 IP and probe hybridisation

Samples were added to the ChromoTek GFP-Trap Multiwell Plate and incubated overnight at 4°C with gentle rocking. Flowthrough was then removed and wells were washed once with 180 µL DF medium salt buffer. 180 µL DF blocking solution was then added and samples were incubated for 15 minutes at 4°C with gentle rocking. Samples were then incubated in 50 µL DF hybridisation buffer, supplemented with 160nM oligo(dT)$_{25}$ stellaris probe (Alexa Fluor 594), for 1 hour at room temperature. Samples were washed once with 180 µL DF wash buffer 1 and twice with 180 µL DF wash buffer 2. 100 µL of DF wash buffer 2 was added for fluorescence measurement.

7.3.3 Fluorescence measurement

Fluorescence measurements were taken using the BMG Clariostar platereader. Measurements were taken for GFP (470-15/515-20) and Alexa Fluor 594 (575-20/630-40). Each well was analysed as a 10x10 matrix in 'well scan' mode and the median fluorescence was calculated. Enhanced dynamic range was used and focal height was set using auto-focus, with measurements taken from above. All values were normalised to a 'blank' well, in which no IP had been performed.

7.4 RNA Interactome Capture

7.4.1 Sample harvesting

Two 15cm dishes per condition were seeded with 20x10^6 HEK293 cells per dish in DMEM (10% FBS, 1x P/S). After 24 hours, media was aspirated from all plates. IFN-containing DMEM was added to the 4hr, 8hr, 16hr, and 20hr plates, and fresh media was added to the mock crosslinked and non-crosslinked control plates. Recombinant human IFNα 2a at a concentration of 200 units/ml was used. At 4, 8, 16, and 20 hours after IFN treatment, the plates for each timepoint were crosslinked and harvested. One plate from mock crosslinked and non-crosslinked controls was harvested at 4 hours and the other was harvested at 20 hours. Media was aspirated, and plates were washed twice in 5ml PBS. Plates for crosslinked samples were placed on ice blocks and UV irradiated at 254nm at an energy of 150mJ/cm^2 . 1ml RIC lysis buffer was added to each plate, and cells were harvested using a cell scraper and lysates from the two plates per condition were combined. Lysates were stored at -80°C.

7.4.2 IFN RIC

Lysates were thawed and lysis buffer volume was increased to 5ml/sample. Homogenisation was achieved by passing the lysate through a 5ml syringe with 27G needle gauge until clear. 600 µL of pre-calibrated oligo(dT)$_{25}$ magnetic beads were added to each sample and samples were incubated for 2 hours at 4°C with gentle rotation. At the end of this incubation, beads were separated from lysate on a magnetic rack and lysates were transferred to tubes containing a fresh set of beads and a second round of capture was performed. The beads were washed in fresh lysis buffer for 5 minutes on ice, with 10 inversions each minute. This wash process was repeated twice with 1ml/plate ice-cold buffer 1 and twice with 1ml/plate ice-cold buffer 2, replacing the samples on the magnet after each wash and removing the supernatant. The beads were then washed twice with room temperature buffer 3, with 10 inversions but without incubation. Following removal of buffer 3, 360 µL elution buffer (20mM Tris-HCl pH 7.5 and 1mM EDTA) was added and the samples were incubated for 5 minutes at 55°C with agitation (300rpm). Beads were separated on the magnetic rack and the eluates were removed to fresh tubes. Samples were then RNase treated with RNase T1 and RNase A (1:1 in 10-fold dilution with elution buffer, 1 µL per sample) and incubated at 37°C for 1 hour. The wash procedure was repeated following the second round of capture and the eluates from each round were combined.

7.4.3 Sample prep

Samples were processed for Mass Spectrometry via the bead-based single-pot, solid-phase-enhanced sample-preparation (SP3) method [440]. First, samples were incubated with 10mM TCEP and 50mM C-IAA at RT in the dark for 30 minutes to reduce and alkylate proteins. Next, samples were split to ensure maximum volume per tube of 300 µL and 100 µg Carboxylated Beads were added to each tube. 100% ACN was added to achieve a final concentration of 80% and samples were incubated for 30 minutes at RT. Samples were then placed on a magnetic rack and left to clear for a further 10 minutes. Samples were then washed twice with 70% EtOH and once with 100% ACN in a loop until samples were free from detergent.

Samples were incubated overnight at 37°C with 250ng trypsin. The next day, tubes were placed on the magnetic rack and the supernatant was transferred to fresh tubes. The remaining peptides were eluted in 2% DMSO for 2 minutes off the rack. Eluates were combined and acidified with neat FA prior to loading on the Mass Spectrometer.

7.4.4 Phosphoenrichment

Following digestion, the peptide mixture was treated with benzonase nuclease (1:1000) and incubated at 37°C for 30 min. Samples were then cleaned up using the Oasis HLB 96-well µElution Plate. Columns were prepared by adding 200 µL ACN and setting the vacuum manifold to 3" Hg. Once all liquid had flowed through, 200 µL 0.5% FA was added and wells were emptied as above. The samples were then added and the wells were washed twice with 0.5% FA. Peptides were eluted twice in 25 µL SPE elution buffer.

For phosphoenrichment, 0.4mg of equilibrated Zr-IMAC beads were added to each sample. Samples were incubated for 20 minutes (RT, 1350 RPM), then placed on a magnetic rack to separate beads. Samples were then washed once in 400 µL Loading Buffer, twice in 400 µL Wash Buffer 1, and twice in 400 µL Wash Buffer 2. For each wash, samples were incubated for 2 min (RT, 1350 RPM), before being placed on the magnetic rack. Samples were incubated for 10 min (RT, 1350 RPM) in 1% ammonium hydroxide to elute peptides. Elution was repeated twice more and eluates were acidified with formic acid to a final concentration of 4%. Clean-up was then repeated, as detailed above. Samples were eluted once in 25 µL SPE elution buffer and then diluted in 100 µL 5% DMSO 5% FA before loading onto the Mass Spectrometer.

7.4.5 Mass Spectrometry

Analysis of IFN RIC peptides was carried out using an Ultimate 3000 nano-LC 1000 system coupled to an Orbitrap Exploris (Thermo Fisher Scientific). Phosphoenriched peptides were analysed using an Ultimate 3000 nano-LC 1000 system coupled to an Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Fisher Scientific). Peptides were initially trapped on a C18 PepMap100 pre-column (300 µm inner diameter x 5 mm, 100A) and then separated on an in-house built C18 column (Reprosil-Gold, Dr. Maisch, 1.9 µm particle size) column (ID: 75 µm, length: 50 cm) at a flow rate of 200 nL/min. IFN RIC peptides were separated over 60 min (12-35%B) and phosphoenriched peptides were separated over 30 min (12-38% B). In both cases, mobile phase A (water and 0.1% formic acid) and mobile phase B (acetonitrile and 0.1% formic acid) were used. Separated peptides were directly electrosprayed into the mass spectrometer. Mass spectra were acquired in the orbitrap (350-1400 m/z, resolution 60000, AGC target 3 x 10^6, maximum injection time 50 ms) in a data-dependent mode. The top 40 most abundant peaks in the survey scan were fragmented using HCD (resolution 7500, AGC target 4 x10^4, maximum injection time 64 ms).

7.5 Protein-protein interaction analysis

7.5.1 Sample harvesting

$10x10^6$ HEK293 Flp-In T-REx MATR3-eGFP cells were seeded per condition in a 10cm dish in DMEM supplemented with 2% FBS. Cells were induced with 1 µg/ml of doxycycline overnight. For the infected condition, cells were infected with 3 MOI of SINV 9 hours before harvesting. When harvesting, cells were washed twice in 5ml PBS, then resuspended in 1ml PBS and transferred to a 1.5ml Eppendorf. Samples were spun down for 5 minutes (1500rpm, 4°C) and PBS was removed. 450 µL IP lysis buffer was added and samples were incubated on ice for 30 minutes. Samples were centrifuged for 3 minutes (3000rpm, 4°C) and supernatant was transferred to fresh tubes. Supernatant was spun for 15 minutes (max speed, 4°C) and the supernatant was transferred to another fresh tube, snap frozen on dry ice and stored at -80°C.

7.5.2 co-IP

Samples were first incubated 25 µL agarose control beads for 30 minutes with rotation at 4°C to pre-clear them. They were then incubated with 50 µL GFP-Trap beads for 2 hours. Samples were spun down (2500xg, 4°C, 5 minutes) and the supernatant was removed. Beads were resuspended in 1ml IP Wash Buffer and divided between two tubes, one for Benzonase treatment and one to be left untreated. Tubes were then spun down again, as above, and beads were resuspended in either 500 µL IP Wash Buffer supplemented with 1 µL mL^{-1} Benzonase or in wash buffer without Benzonase. Benzonase-treated samples were incubated for 15 minutes at 37°C and untreated samples were stored on ice. All samples were then washed twice more in 500 µL IP Wash Buffer. To elute, beads were resuspended in 50 µL 1% SDS and incubated for 5 minutes at 55°C with rotation (1100rpm). Samples were then spun down for 2 minutes (RT, 2500xg) and the eluate was transferred to a fresh tube. Elution was repeated once more and eluates were combined.

7.5.3 Sample prep

Samples were prepared using SP3, as described in the RNA Interactome Capture Sample Preparation section.

7.5.4 Mass Spectrometry

Analysis of peptides was carried out using an Ultimate 3000 nano-LC 1000 system coupled to an Orbitrap Fusion Ascend Tribrid Mass Spectrometer (Thermo Fisher Scientific). Peptides were initially trapped on a C18 PepMap100 pre-column (300 µm inner diameter x 5 mm, 100A) and then separated on an in-house built C18 column (Reprosil-Gold, Dr. Maisch, 1.9 µm particle size) column (ID: 50 µm, length: 50 cm) at a flow rate of 100 nL/min. Peptides were separated over 60 min (12-35%B) using mobile phase A (water and 5% DMSO, 0.1% formic acid) and mobile phase B (acetonitrile and 5% DMSO, 0.1% formic acid). Separated peptides were directly electrosprayed into an Orbitrap Ascend mass spectrometer and analysed in a data-dependent mode. MS1 spectra were acquired in the orbitrap (350-1400 m/z, resolution 60000, AGC target 3 x 10^6, maximum injection time 50 ms). The top 40 most abundant peaks in the survey scan were fragmented using HCD and analysed in the ion trap (scan speed Turbo, AGC target 2 x10^4, maximum injection time 32 ms, normalised collision energy 26%).

7.6 iCLIP2

7.6.1 Sample harvesting

For IP samples, 10x10^6 HEK293 Flp-In T-REx RBP-eGFP cells were seeded per condition in a 10cm dish in DMEM supplemented with 2% FBS. Cells were induced with 1 µg/ml of doxycycline overnight. For the infected condition, cells were infected with 3 MOI of SINV 9 hours before harvesting. For the IFN-treated condition, cells 500U/ml IFNA2 was added to media 20 hours before harvesting. For Size Matched Input (SMI) samples, parental HEK293 Flp-In T-REx cells were seeded and treated in the same way. Next, cells were washed twice in 1X PBS and crosslinked at 0.3 J/cm2 UV light irradiation at 254 nm. Cells were then lysed in 1 ml RIPA buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1% Triton X-100, 0.1% SDS, 0.5% wt/vol Na deoxycholate and 0.2 mM AEBSF). Lysates were incubated for 30 min on ice, then homogenized by passing through a 27G needle five times. Finally, lysates were cleared by centrifugation (18000xg for 10 min at 4℃), snap-frozen in dry ice, and stored at -80℃ until use.

7.6.2 iCLIP2

4 U TurboDNase and 5 U RNase I were added to lysates. These were mixed by vortexing and incubated for 3 min at 37℃ (1100 rpm). 200 U RiboLock RNase Inhibitor was then added, and lysates were incubated for a further 3 min on ice. Lysates were pre-cleared with 25 µL of pre-equilibrated control agarose beads for 30 min at 4℃ with gentle rotation followed by centrifugation at 2500xg for 2 min. Lysates were transferred to fresh tubes and incubated with 25 µL of pre-equilibrated GFP_Trap agarose bead slurry for 2 h at 4℃ with gentle rotation. Beads were washed twice with 900 µL of cold high-salt wash buffer, twice with 900 µL of cold medium-salt wash buffer, and twice with 900 µL of cold PNK wash buffer. RNA 3'-end dephosphorylation was performed at 37℃ for 40 min (1100 rpm) in 5x PNK buffer with 5 U PNK, 0.25 U FastAP alkaline phosphatase, 0.5 U TurboDNase and 20 U RNasin. Beads were washed once with 500 µL cold PNK wash buffer, twice with 900 µL of cold high-salt wash buffer, and twice with 900 µL cold PNK wash buffer. L3-IR-App adapter [441] ligation was performed overnight in the dark (16℃, 1100 rpm) in a mixture composed as outlined in Table 15. Beads were then washed once with 500 µL PNK wash buffer, twice with 900 µL of high-salt wash buffer, and twice with 900 µL PNK wash buffer.

Component	**Concentration/Amount**
L3-IR-App adapter	125 nM
T4 RNA ligase	30 U
Ribolock RNasin	20 U
T4 PNK	4 U
PEG8000	22.5%
DMSO	5%
10x Ligation buffer	1x

Table 15: L3-IR-App adapter ligation mix

IP and SMI samples were denatured in 1X NuPage LDS Sample Buffer with 100 mM DTT for 5 min at 70℃. IP samples were spun down for 2 min at 2500xg and eluate was transferred to a fresh tube. Elution was repeated once more and eluates were combined. Samples were separated on a 4–15% Mini-PROTEAN TGX Precast Protein Gel for 65 min at 180 V. Protein-RNA complexes were transferred onto an iBLOT2 nitrocellulose membrane using the Bio-Rad Trans-blot Turbo (1.5mm gel setting) and visualized on a LI-COR Odyssey Fc imaging system.

The region corresponding to the RBP-eGFP band and above was cut (for both IP and SMI samples) and digested using 350 µg Proteinase K in 180 µL PK-SDS solution for 60 min at 50℃ (1100 rpm). RNA was purified by adding 1X volume of Phenol:Chloroform:Isoamyl Alcohol pH 6.6-6.9, incubating for 10 min at 37℃ (1100 rpm) and centrifugation at 16000xg for 5 min in MaxTract tubes. RNA was cleaned using Zymo RNA Clean & Concentrator-5.

For SMI library preparation, SMI control samples were then treated with 5 U PNK, 0.5 U FastAP, and 20 U RNAsin in PNK buffer pH 6.5 for 40 min at 37℃ (1100 rpm). RNA was cleaned up with Dynabeads MyOne Silane. L3-IR-App adapter ligation was performed with 45 U T4 RNA ligase in 1X T4 RNA Ligase Reaction Buffer with 2% DMSO, 27% PEG8000, and 133 nM L3-IR-adapter for 75 min at room temperature, followed by MyOne bead purification. SMI samples were then treated with 25 U 5' Deadenylase and 15 U RecJf endonuclease in 1X New England Biolabs buffer 2 with 20 U RNAsin and 20% PEG8000 for 1 h at 30℃ and then 30 min at 37℃ (1100 rpm), followed by a MyONE clean-up. RNA from IP and SMI samples were reverse transcribed using Superscript IV reverse transcriptase and hydrolyzed by adding 1.25 µL of 1 M NaOH for 15 min at 85℃, before neutralization with 1.25 µL of 1 M HCl.

cDNA was purified using MyOne silane beads. L#clip2.0 adapters with barcodes for multiplexing [368] were ligated to cDNA by mixing 2 µL of 10 µM adapter with 5 µL of cDNA, adding 1 µL of DMSO, and incubating at 75℃ for 2 min before placing on ice. Then, ligation mix (45 U T4 RNA ligase in 1X RNA ligase buffer with 22.5% PEG8000) was added to the cDNA-bead solution and incubated overnight at 20℃ (1100 rpm). cDNA was cleaned up with MyONE beads before PCR amplification.

Pre-amplification was performed using 2X Phusion HF PCR Master mix with P5Solexa_s and P3Solexa_s primers for six cycles, followed by ProNex size-selective purification. Optimal qPCR cycles were determined by Real-Time qPCR, using EvaGreen, 2X Phusion HF PCR Master mix, and P5/P3 Solexa primers. Final PCR products were purified using two consecutive rounds of ProNex Size selection.

7.6.3 Sequencing

Samples were quantified using Qubit DNA HS and library size was measured using High Sensitivity D1000 TapeStation. Each group of samples was pooled equimolarly and then mixed at the following proportions: 75% IP library pool, 25% SMI library pool. Sequencing was performed on a NextSeq 550 sequencer with a 75 cycle High-output kit v2.5.

7.7 Data analysis

7.7.1 General

GO enrichment analyses were performed using clusterProfiler in R [442]. A p-value cut-off of 0.01 and a q-value cut-off of 0.05 were used and the Benjamini-Hochberg method was used to correct for multiple testing. Non-unique GO terms were collapsed using clusterProfiler's simplify function.

vRNA interactome datasets were compiled as described in Iselin et al [192]. Briefly, all vRNA interactome datasets generated in the context of whole-virus infection that were available at the time were used. Hits were defined based on the statistical threshold set within each study. The core vRNA interactome was defined as any protein occurring in at least one dataset generated for a virus from the *Coronaviridae*, *Togaviridae*, and *Flaviviridae* families.

Text mining was performed in R using the RISmed package. Separate searches of PubMed titles and abstracts were done for the terms: virus-linked ([gene name] AND [virus OR viral]), immune-linked ([gene name] AND [immune OR immunity]), and IFN-linked ([gene name] AND [IFN OR interferon]. No minimum date was set and the maximum date was set to the date of analysis. A gene was considered 'linked' to the above terms if there were 6 or more papers returned.

Principal component analyses were performed using the base R package on log2 transformed values. Data was filtered to exclude NAs and normalised and batch corrected to match limma analysis processing where appropriate.

7.7.2 IFN RIC

Protein identification and quantification were performed using Andromeda search engine implemented in MaxQuant [228]. Peptides were searched v Human Uniprot database. False discovery rate (FDR) was set at 1% using the both peptide and protein identification and 'match between runs' was turned on. Otherwise, default parameters were used.

Data analysis of ProteinGroups file from MaxQuant was performed in R. Data was normalised using the 'vsn' package. Rows were filtered to remove any with $>$2 NA values in each condition under study. Minimum value imputation was performed for on/off changes (all NA values in one condition and $<$2 NA values in the other). Only values for replicates corresponding to non-NA values in the other condition were imputed. Fold-changes and p values were calculated using

the limma package [229], with replicate included as a block in the regression analysis, and FDR was calculated from p values using the fdrtool package.

For kinetic clustering and plotting, values were first imputed for missing timepoints using minimum value imputation and fold-changes relative to mock were calculated. Where limma analysis had been performed, fold-changes calculated during this analysis were used. Values were then transformed from fold-changes into z-scores. K-means clustering was performed using the FKM function from the fclust package with k set to 6.

7.7.3 IFN RIC phosphoproteomics

Protein identification and quantification were performed using Andromeda search engine implemented in MaxQuant [228]. Peptides were searched using the Human Uniprot database. False discovery rate (FDR) was set at 1% for both peptides and 'match between runs' was turned on. Phospho(STY) was set as a variable modification. Otherwise, default parameters were used.

Data analysis of Phospho(STY) file was performed in R. Data was normalised using the 'vsn' package. Rows were filtered to remove any with $>$2 NA values in each condition under study. Minimum value imputation was performed for on/off changes (all NA values in one condition and $<$2 NA values in the other). Only values for replicates corresponding to non-NA values in the other condition were imputed. Fold-changes and p values were calculated using the limma package [229], with replicate included as a block in the regression analysis, and FDR was calculated from p values using the fdrtool package.

7.7.4 Protein-protein interaction

Protein identification and quantification were performed using Andromeda search engine implemented in MaxQuant [228]. Peptides were searched using the Human Uniprot database with viral proteins added. False discovery rate (FDR) was set at 1% for both peptide and protein identification and 'match between runs' was turned on. Phospho(STY) was set as a variable modification. Otherwise, default parameters were used.

Data analysis of ProteinGroups file from MaxQuant was performed in R. For MATR3-GFP v parental comparison and benzonase + v benzonase -, no normalisation was performed. For mock v SINV comparison, data was normalised using the 'vsn' package. For benzonase + v benzonase - and mock v SINV comparisons, data was filtered to include only proteins that reached the 0.01% FDR threshold in at least one MATR3-GFP v parental comparison. Rows

were filtered to remove any with >2 NA values in each condition under study. Minimum value imputation was performed for on/off changes (all NA values in one condition and <2 NA values in the other). Only values for replicates corresponding to non-NA values in the other condition were imputed. Fold-changes and p values were calculated using the limma package [229] and FDR was calculated from p values using the fdrtool package.

To generate interaction networks, protein IDs were imported into string-db [443] and interactions were filtered to include only those with a confidence score >0.4. Networks were plotted using Cytoscape.

7.7.5 iCLIP

Raw FASTQ files were demultiplexed using the Je Suite [444] and adapters were trimmed using Cutadapt [445]. STAR was used to align reads to a concatenated human (GRCh38, ENSEMBL Release 106) and SINV (pT7-SVwt) genome in end-to-end alignment mode [446]. Only uniquely aligned reads were retained for downstream analysis. PCR duplicates were collapsed using unique molecular identifiers (UMIs) with the Je Suite. The crosslink truncation site for each read (-1 from the 5' start site of the read) was extracted using BEDTools [447].

Peak calling was performed with HTSeq-clip and the R/Bioconductor package, DEW-seq [369]. HTSeq-clip was used to generate a sliding window annotation of the human and SINV genome (50nt window, 20nt step size) and calculate the frequency of crosslink truncation sites within each window. DEW-Seq was then used to calculate the differential enrichment of each window relative to size-matched input control samples, with a cut-off of >2 log2 fold change and >0.01 adjusted p-value. Multiple hypothesis correction was performed using the Independent Hypothesis Weighting (IHW) method [448]. Overlapping windows were merged to form binding regions.

PCA was performed using DESeq2 [449] . Following size correction and variance stabilisation, the 1000 most variable sliding windows were selected and used for PCA plotting.

Binding site properties, including gene name, biotype, and gene feature, were extracted from the ENSEMBL genome annotation using the GenomicRanges package. Metagene analyses were performed using functions from the cliProfiler package.

Sequences for motif prediction and secondary structure prediction were defined for each binding site as a 50-nucleotide region, centered on the peak in BigWig signal. For motif prediction, a gene and gene region-matched background sequence was extracted for each

binding site to allow for differential enrichment analysis. Enrichment analysis was performed using STREME from the MEME suite [370]. Universalmotif was used for motif processing and motifStack was used to cluster and plot motifs.

Secondary structures were predicted using RNAfold from ViennaRNA [450] and specific structure features were extracted by Forgi, also from ViennaRNA. Enrichment for paired sequence was calculated relative to 10 scrambled controls of each input sequence. For plotting of SINV viral genome structures, SHAPE data from Kutchko et al [371] was included as a constraint in the structure prediction. Coordinates for plotting specific structures were generated with the RRNA package.

SINV genome coverage in reads per million was calculated using BEDTools [447]. Percent of total signal was then calculated at each position in the IP and SMI samples. SMI signal was subtracted from IP signal for plotting.

8 Appendix 1

Table displaying all significantly changing genes in IFN RIC experiment. Significance at each timepoint is indicated by +/- symbols. **++** log2FC $>$ 0 & FDR $\leq$ 1%; **+** log2FC $>$ 0 & FDR $\leq$ 10%; **–** log2FC $<$ 0 & FDR $\leq$ 1%; **-** log2FC $<$ 0 & FDR $\leq$ 10%.

Gene	**Sig 4h**	**Sig 8h**	**Sig 16h**	**Sig 20h**
OAS3		+	++	++
NSD1		+	+	++
HNRNPA1	+	++		++
HELZ2		+	++	++
C19orf66				++
RPUSD3				–
IFIT1		+	++	++
SUGP1				++
PUM1				–
HERC5			++	++
NHP2				–
DACT2	–	–		–
PARP12 (Q9H0J9)			++	++
PABPN1			–	–
EIF2AK2			++	++
ZNF48				–
HNRNPH1 (H0YBK1)				++
S100A8				–
CHD1				–
KRI1			-	-
C1orf52				-
NOC2L			-	-
SCAMP4				-

Table 16: IFN RIC significance results

Gene	Sig 4h	Sig 8h	Sig 16h	Sig 20h
PHRF1			-	-
PCMT1				-
ADAR			+	+
DDX60L				+
PUF60 (A0A0J9YXJ8)				-
TOP2A			-	-
RBM39				+
TNRC6B				-
NHP2L1				-
TEX10				-
ERAL1				-
PUM2				-
ATXN2 (F8W0B5)				-
KIF1B				-
HNRNPDL (A0A087WUK2)				+
SFPQ				+
SPCS2				+
UFL1				-
CCDC124				+
CPEB4				-
HNRNPAB				+
KPNB1				-
GAPVD1				+

Gene	Sig 4h	Sig 8h	Sig 16h	Sig 20h
RPF2				-
JUP				+
FAM170A			+	+
MBNL1 (H7C4T5)				-
DDX42				+
PABPC4				-
SUGP2				+
KHSRP				+
KIF1C				-
CSTF2T				+
SNRNP200				+
ZC3H7A				-
HNRNPC (G3V2Q1)			+	+
RAVER1				+
PHF5A			+	+
TNRC6A				-
U2AF2				+
HNRNPA3				+
RBM3				+
TRIM25			+	+
CTSD				+
RBM20				+
XPO5				-

Gene	Sig 4h	Sig 8h	Sig 16h	Sig 20h
RPS26				-
DIEXF				-
APOBEC3C			-	-
RPA1				+
FCF1				-
NONO				+
STRBP				-
A0A7I2V323				+
SNRPG				+
RBM15B				+
RBM6				+
UPF1				-
HNRNPD (H0Y8G5)				+
RPL35			-	-
RBMX				+
XRN1				-
NOB1				+
IFI16				+
GANAB				+
RPS3				-
GRWD1				-
RBM15				+
MATR3				+

Gene	Sig 4h	Sig 8h	Sig 16h	Sig 20h
RALY				+
CIRH1A				-
SENP3				+
DDX55				-
HNRNPM (P52272)				+
RC3H2				-
CHERP				+
NAT10				-
NOP58				-
RPLP1				-
MRPL17			-	-
RBM4				+
DDX17				+
SART3				+
SNRPD1				+
ISG20L2				-
CSTF3				-
TARDBP				+
MAP4				+
RBMXL1				+
RBMX2				+
SRSF3				+
CEBPZ				-

Gene	Sig 4h	Sig 8h	Sig 16h	Sig 20h
FAM120A				-
MRPS7				-
HNRNPUL1				+
HIST1H1E				+
TSN				+
GOLGA3				+
SLC25A3			+	+
PLEC				+
MKRN2				-
NAA10				-
RPS13				-
SEC61B				-
SLTM				+
CELF2				+
HNRNPL				+
FXR2				-
SSBP1				+
SRSF2				+
RBM5				+
HIST1H1C				+
NOM1				-
ZC3H18				+
POLR2A				+

Gene	Sig 4h	Sig 8h	Sig 16h	Sig 20h
PSPC1				+
ESF1				-
U2SURP				+
UTP11L				-
TBRG4				-
MYEF2 (A0A0A0MR39)				+
PTBP1				+
MRPS9				-
YARS2				-
EFTUD2				+
ESPL1				+
RTCB				-
EIF4G1				-
ALYREF				-
NSUN2				-
MEX3C				-
TNPO1			-	-
PEG10				-
MCAT				-
HNRNPC				+
SUPT5H				+
NOL8				-
FBLL1				-

Gene	Sig 4h	Sig 8h	Sig 16h	Sig 20h
SF1 (Q15637)				+
KPNA2				-
MRPL57				+
MRPS23				-
WDR3				-
GLTSCR2				+
EWSR1				+
HNRNPM				+
A0A3B3ITJ4				+
APOBEC3F				-
NOL6				-
NSA2				-
ZC3H4				+
PPAN-P2RY11				-
NOL10				-
TIA1 (F8W8I6)				+
RPS14				-
THRAP3				+
SRSF10				+
HNRNPK				+
PATL1				-
EML4				+
HNRNPH1				+

Gene	Sig 4h	Sig 8h	Sig 16h	Sig 20h
EIF3L				-
UBAP2L				-
RPS20				+
SRSF12				+
PWP2				-
HNRNPDL				+
RPL12				+
RAN				+
UTP15				-
TRUB2				-
RPF1				-
SRSF4				+
RPL23A				+
RBM14				+
MRPS31				-
TEFM				-
MRPL47				+
LARP4B				+
MRPS26				+
HNRNPUL2				+
PPP1R10				+
ARHGEF1				+
DDX10				-

Gene	Sig 4h	Sig 8h	Sig 16h	Sig 20h
SBF1				-
SNRPD2				+
KHDRBS1				+
BTF3L4				+
RRP7A			–	
EIF5B				+
EIF4A3				-
FUS				+
DAZAP1				+
RBM4B				+
HNRNPF				+
TRIM56				+
TRMT10C			-	-
GIGYF2				-
NCOA5				+
RBM7				+
DDX47				-
CSTF2				+
POLRMT				-
RPL30				-
FAM98B				-
MYH10				+
LDHB				+

Gene	Sig 4h	Sig 8h	Sig 16h	Sig 20h
RBM34				-
SRSF6				+
ZCCHC11				-
MRPL4				-
FARSB				+
HNRNPR				+
RBM10				+
METTL16				-
MIF				+
ELAVL1				+
GNL3L				-
HNRNPH2				+
HNRNPA2B1				+
SRSF9				+
SRSF7				+
KDM5A				+
NARS				+
DDX31				-
SPEN				+
AKAP8				+
DDX28				-
WDR75				-
EIF3A				-

Gene	Sig 4h	Sig 8h	Sig 16h	Sig 20h
RPL21				-
ZC3H13				+
PURB				-
RBM45				+
RPS27A				+
PURA				-
RBM12B				+
EIF3C				-
YLPM1				+
SECISBP2				-
MEX3B			–	
MYH9				+
SARNP				-
HNRNPLL				+
API5				+
CASC3				-
MRPL42				+
CKAP5				-
RPL15				-
A0A5H1ZRQ2				+
NOL11				-
MRPS21			-	
ARL6IP4	-			

Gene	Sig 4h	Sig 8h	Sig 16h	Sig 20h
SLBP	–			
BOLA2	-			
SLC25A6	-			
AKT3	–	-		
NAP1L1			-	
NACA		-		
HSP90B1			-	
SPCS1	-			
SNRPF	-			
KRR1		+		
RPL4	–			
FAM207A	–			
HMGB2		+		

9 References

[1] A. Isaacs and J. Lindenmann. Virus interference. I. The interferon. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 147(927):258–267, September 1957.

[2] John W. Schoggins. Interferon-Stimulated Genes: What Do They All Do? *Annual Review of Virology*, 6(1):567–584, 2019. _eprint: https://doi.org/10.1146/annurev-virology-092818-015756.

[3] Sidney Pestka, Christopher D. Krause, and Mark R. Walter. Interferons, interferon-like cytokines, and their receptors. *Immunological Reviews*, 202:8–32, December 2004.

[4] Paul Sheppard, Wayne Kindsvogel, Wenfeng Xu, Katherine Henderson, Stacy Schlutsmeyer, Theodore E. Whitmore, Rolf Kuestner, Ursula Garrigues, Carl Birks, Jenny Roraback, Craig Ostrander, Dennis Dong, Jinu Shin, Scott Presnell, Brian Fox, Betty Haldeman, Emily Cooper, David Taft, Teresa Gilbert, Francis J. Grant, Monica Tackett, William Krivan, Gary McKnight, Chris Clegg, Don Foster, and Kevin M. Klucher. IL-28, IL-29 and their class II cytokine receptor IL-28R. *Nature Immunology*, 4(1):63–68, January 2003. Publisher: Nature Publishing Group.

[5] Sergei V. Kotenko, Grant Gallagher, Vitaliy V. Baurin, Anita Lewis-Antes, Meiling Shen, Nital K. Shah, Jerome A. Langer, Faruk Sheikh, Harold Dickensheets, and Raymond P. Donnelly. IFN-lambdas mediate antiviral protection through a distinct class II cytokine receptor complex. *Nature Immunology*, 4(1):69–77, January 2003.

[6] Caroline Sommereyns, Sophie Paul, Peter Staeheli, and Thomas Michiels. IFN-Lambda (IFN-λ) Is Expressed in a Tissue-Dependent Fashion and Primarily Acts on Epithelial Cells In Vivo. *PLOS Pathogens*, 4(3):e1000017, March 2008. Publisher: Public Library of Science.

[7] Alexandra I. Wells and Carolyn B. Coyne. Type III Interferons in Antiviral Defenses at Barrier Surfaces. *Trends in Immunology*, 39(10):848–858, October 2018.

[8] Kate Schroder, Paul J Hertzog, Timothy Ravasi, and David A Hume. Interferon-γ: an overview of signals, mechanisms and functions. *Journal of Leukocyte Biology*, 75(2):163–189, February 2004.

[9] Michael Carty, Coralie Guy, and Andrew G. Bowie. Detection of Viral Infections by Innate Immunity. *Biochemical Pharmacology*, 183:114316, January 2021.

[10] Martin Schlee and Gunther Hartmann. Discriminating self from non-self in nucleic acid sensing. *Nature Reviews Immunology*, 16(9):566–580, September 2016. Number: 9 Publisher: Nature Publishing Group.

[11] Chunmei Cai, Yan-Dong Tang, Guocai Xu, and Chunfu Zheng. The crosstalk between viral RNA- and DNA-sensing mechanisms. *Cellular and Molecular Life Sciences*, 78(23):7427–7434, 2021.

[12] Lionel B. Ivashkiv and Laura T. Donlin. Regulation of type I interferon responses. *Nature Reviews Immunology*, 14(1):36–49, January 2014. Number: 1 Publisher: Nature Publishing Group.

[13] Kenya Honda, Hideyuki Yanai, Hideo Negishi, Masataka Asagiri, Mitsuharu Sato, Tatsuaki Mizutani, Naoya Shimada, Yusuke Ohba, Akinori Takaoka, Nobuaki Yoshida, and Tadatsugu Taniguchi. IRF-7 is the master regulator of type-I interferon-dependent immune responses. *Nature*, 434(7034):772–777, April 2005.

[14] Jan Rehwinkel and Michaela U. Gack. RIG-I-like receptors: their regulation and roles in RNA sensing. *Nature Reviews Immunology*, 20(9):537–551, September 2020. Publisher: Nature Publishing Group.

[15] Mitsutoshi Yoneyama, Mika Kikuchi, Takashi Natsukawa, Noriaki Shinobu, Tadaatsu Imaizumi, Makoto Miyagishi, Kazunari Taira, Shizuo Akira, and Takashi Fujita. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nature Immunology*, 5(7):730–737, July 2004.

[16] J. Andrejeva, K. S. Childs, D. F. Young, T. S. Carlos, N. Stock, S. Goodbourn, and R. E. Randall. The V proteins of paramyxoviruses bind the IFN-inducible RNA helicase, mda-5, and inhibit its activation of the IFN-beta promoter. *Proceedings of the National Academy of Sciences of the United States of America*, 101(49):17264–17269, December 2004.

[17] Hiroki Kato, Osamu Takeuchi, Shintaro Sato, Mitsutoshi Yoneyama, Masahiro Yamamoto, Kosuke Matsui, Satoshi Uematsu, Andreas Jung, Taro Kawai, Ken J. Ishii,

Osamu Yamaguchi, Kinya Otsu, Tohru Tsujimura, Chang-Sung Koh, Caetano Reis e Sousa, Yoshiharu Matsuura, Takashi Fujita, and Shizuo Akira. Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature*, 441(7089):101–105, May 2006.

[18] Hiroki Kato, Osamu Takeuchi, Eriko Mikamo-Satoh, Reiko Hirai, Tomoji Kawai, Kazufumi Matsushita, Akane Hiiragi, Terence S. Dermody, Takashi Fujita, and Shizuo Akira. Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5. *The Journal of Experimental Medicine*, 205(7):1601–1610, July 2008.

[19] Veit Hornung, Jana Ellegast, Sarah Kim, Krzysztof Brzózka, Andreas Jung, Hiroki Kato, Hendrik Poeck, Shizuo Akira, Karl-Klaus Conzelmann, Martin Schlee, Stefan Endres, and Gunther Hartmann. 5'-Triphosphate RNA is the ligand for RIG-I. *Science (New York, N.Y.)*, 314(5801):994–997, November 2006.

[20] Andreas Pichlmair, Oliver Schulz, Choon Ping Tan, Tanja I. Näslund, Peter Liljeström, Friedemann Weber, and Caetano Reis e Sousa. RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates. *Science (New York, N.Y.)*, 314(5801):997–1001, November 2006.

[21] Martin Schlee, Andreas Roth, Veit Hornung, Cristina Amparo Hagmann, Vera Wimmenauer, Winfried Barchet, Christoph Coch, Markus Janke, Aleksandra Mihailovic, Greg Wardle, Stefan Juranek, Hiroki Kato, Taro Kawai, Hendrik Poeck, Katherine A. Fitzgerald, Osamu Takeuchi, Shizuo Akira, Thomas Tuschl, Eicke Latz, Janos Ludwig, and Gunther Hartmann. Recognition of 5' triphosphate by RIG-I helicase requires short blunt double-stranded RNA as contained in panhandle of negative-strand virus. *Immunity*, 31(1):25–34, July 2009.

[22] Delphine Goubau, Martin Schlee, Safia Deddouche, Andrea J. Pruijssers, Thomas Zillinger, Marion Goldeck, Christine Schuberth, Annemarthe G. Van der Veen, Tsutomu Fujimura, Jan Rehwinkel, Jason A. Iskarpatyoti, Winfried Barchet, Janos Ludwig, Terence S. Dermody, Gunther Hartmann, and Caetano Reis e Sousa. Antiviral immunity via RIG-I-mediated recognition of RNA bearing 5'-diphosphates. *Nature*, 514(7522):372–375, October 2014.

[23] Christine Schuberth-Wagner, Janos Ludwig, Ann Kristin Bruder, Anna-Maria Herzner, Thomas Zillinger, Marion Goldeck, Tobias Schmidt, Jonathan L. Schmid-Burgk, Romy Kerber, Steven Wolter, Jan-Philip Stümpel, Andreas Roth, Eva Bartok, Christian Drosten, Christoph Coch, Veit Hornung, Winfried Barchet, Beate M. Kümmerer, Gunther Hartmann, and Martin Schlee. A Conserved Histidine in the RNA Sensor RIG-I Controls Immune Tolerance to N1-2'O-Methylated Self RNA. *Immunity*, 43(1):41–51, July 2015.

[24] Maxime Chazal, Guillaume Beauclair, Ségolène Gracias, Valérie Najburg, Etienne Simon-Lorière, Frédéric Tangy, Anastassia V. Komarova, and Nolwenn Jouvenet. RIG-I Recognizes the 5' Region of Dengue and Zika Virus Genomes. *Cell Reports*, 24(2):320–328, July 2018.

[25] Aartjan J. W. Te Velthuis, Joshua C. Long, David L. V. Bauer, Rebecca L. Y. Fan, Hui-Ling Yen, Jane Sharps, Jurre Y. Siegers, Marian J. Killip, Hollie French, Maria José Oliva-Martín, Richard E. Randall, Emmie de Wit, Debby van Riel, Leo L. M. Poon, and Ervin Fodor. Mini viral RNAs act as innate immune agonists during influenza virus infection. *Nature Microbiology*, 3(11):1234–1242, November 2018.

[26] Antonio Gregorio Dias Junior, Natalia G. Sampaio, and Jan Rehwinkel. A Balancing Act: MDA5 in Antiviral Immunity and Autoinflammation. *Trends in Microbiology*, 27(1):75–85, January 2019.

[27] Jessica J. Chiang, Konstantin M. J. Sparrer, Michiel van Gent, Charlotte Lässig, Teng Huang, Nikolaus Osterrieder, Karl-Peter Hopfner, and Michaela U. Gack. Viral unmasking of cellular 5S rRNA pseudogene transcripts induces RIG-I-mediated immunity. *Nature Immunology*, 19(1):53–62, January 2018.

[28] Ashish Dhir, Somdutta Dhir, Lukasz S. Borowski, Laura Jimenez, Michael Teitell, Agnès Rötig, Yanick J. Crow, Gillian I. Rice, Darragh Duffy, Christelle Tamby, Takayuki Nojima, Arnold Munnich, Manuel Schiff, Claudia Ribeiro de Almeida, Jan Rehwinkel, Andrzej Dziembowski, Roman J. Szczesny, and Nicholas J. Proudfoot. Mitochondrial double-stranded RNA triggers antiviral signalling in humans. *Nature*, 560(7717):238–242, August 2018.

[29] Eric D. Tang and Cun-Yu Wang. MAVS self-association mediates antiviral innate immune signaling. *Journal of Virology*, 83(8):3420–3428, April 2009.

[30] Siqi Liu, Xin Cai, Jiaxi Wu, Qian Cong, Xiang Chen, Tuo Li, Fenghe Du, Junyao Ren, You-Tong Wu, Nick V. Grishin, and Zhijian J. Chen. Phosphorylation of innate immune adaptor proteins MAVS, STING, and TRIF induces IRF3 activation. *Science*, 347(6227):aaa2630, March 2015. Publisher: American Association for the Advancement of Science.

[31] Katherine A. Fitzgerald, Sarah M. McWhirter, Kerrie L. Faia, Daniel C. Rowe, Eicke Latz, Douglas T. Golenbock, Anthony J. Coyle, Sha-Mei Liao, and Tom Maniatis. IKKepsilon and TBK1 are essential components of the IRF3 signaling pathway. *Nature Immunology*, 4(5):491–496, May 2003.

[32] Sandra N. Lester and Kui Li. Toll-Like Receptors in Antiviral Innate Immunity. *Journal of Molecular Biology*, 426(6):1246–1264, March 2014.

[33] Lena Alexopoulou, Agnieszka Czopik Holt, Ruslan Medzhitov, and Richard A. Flavell. Recognition of double-stranded RNA and activation of NF-κB by Toll-like receptor 3. *Nature*, 413(6857):732–738, October 2001. Publisher: Nature Publishing Group.

[34] Joshua N. Leonard, Rodolfo Ghirlando, Janine Askins, Jessica K. Bell, David H. Margulies, David R. Davies, and David M. Segal. The TLR3 signaling complex forms by cooperative receptor dimerization. *Proceedings of the National Academy of Sciences*, 105(1):258–263, January 2008. Publisher: Proceedings of the National Academy of Sciences.

[35] Florian Heil, Hiroaki Hemmi, Hubertus Hochrein, Franziska Ampenberger, Carsten Kirschning, Shizuo Akira, Grayson Lipford, Hermann Wagner, and Stefan Bauer. Species-Specific Recognition of Single-Stranded RNA via Toll-like Receptor 7 and 8. *Science*, 303(5663):1526–1529, March 2004. Publisher: American Association for the Advancement of Science.

[36] Hiroaki Hemmi, Osamu Takeuchi, Taro Kawai, Tsuneyasu Kaisho, Shintaro Sato, Hideki Sanjo, Makoto Matsumoto, Katsuaki Hoshino, Hermann Wagner, Kiyoshi Takeda, and Shizuo Akira. A Toll-like receptor recognizes bacterial DNA. *Nature*, 408(6813):740–745, December 2000. Publisher: Nature Publishing Group.

[37] Jennifer Lund, Ayuko Sato, Shizuo Akira, Ruslan Medzhitov, and Akiko Iwasaki. Toll-like Receptor 9–mediated Recognition of Herpes Simplex Virus-2 by Plasmacytoid Dendritic Cells. *Journal of Experimental Medicine*, 198(3):513–520, August 2003.

[38] Takumi Kawasaki and Taro Kawai. Toll-Like Receptor Signaling Pathways. *Frontiers in Immunology*, 5:461, September 2014.

[39] Mary J. Manuse, Caitlin M. Briggs, and Griffith D. Parks. Replication-independent activation of human plasmacytoid dendritic cells by the paramyxovirus SV5 Requires TLR7 and autophagy pathways. *Virology*, 405(2):383–389, September 2010.

[40] Ken Takahashi, Shinichi Asabe, Stefan Wieland, Urtzi Garaigorta, Pablo Gastaminza, Masanori Isogawa, and Francis V. Chisari. Plasmacytoid dendritic cells sense hepatitis C virus-infected cells, produce interferon, and inhibit infection. *Proceedings of the National Academy of Sciences of the United States of America*, 107(16):7431–7436, April 2010.

[41] Xu Zhang, Jiaxi Wu, Fenghe Du, Hui Xu, Lijun Sun, Zhe Chen, Chad A. Brautigam, Xuewu Zhang, and Zhijian J. Chen. The Cytosolic DNA Sensor cGAS Forms an Oligomeric Complex with DNA and Undergoes Switch-like Conformational Changes in the Activation Loop. *Cell Reports*, 6(3):421–430, February 2014.

[42] Jiaxi Wu, Lijun Sun, Xiang Chen, Fenghe Du, Heping Shi, Chuo Chen, and Zhijian J. Chen. Cyclic GMP-AMP Is an Endogenous Second Messenger in Innate Immune Signaling by Cytosolic DNA. *Science*, 339(6121):826–830, February 2013.

[43] Pu Gao, Manuel Ascano, Yang Wu, Winfried Barchet, Barbara L. Gaffney, Thomas Zillinger, Artem A. Serganov, Yizhou Liu, Roger A. Jones, Gunther Hartmann, Thomas Tuschl, and Dinshaw J. Patel. Cyclic [G(2',5')pA(3',5')p] Is the Metazoan Second Messenger Produced by DNA-Activated Cyclic GMP-AMP Synthase. *Cell*, 153(5):1094–1107, May 2013.

[44] Bo Sun, Karin B. Sundström, Jun Jie Chew, Pradeep Bist, Esther S. Gan, Hwee Cheng Tan, Kenneth C. Goh, Tanu Chawla, Choon Kit Tang, and Eng Eong Ooi. Dengue virus activates cGAS through the release of mitochondrial DNA. *Scientific Reports*, 7(1):3594, June 2017.

[45] Leonie Unterholzner, Sinead E. Keating, Marcin Baran, Kristy A. Horan, Søren B. Jensen, Shruti Sharma, Cherilyn M. Sirois, Tengchuan Jin, Eicke Latz, T. Sam Xiao,

Katherine A. Fitzgerald, Søren R. Paludan, and Andrew G. Bowie. IFI16 is an innate immune sensor for intracellular DNA. *Nature Immunology*, 11(11):997–1004, November 2010. Number: 11 Publisher: Nature Publishing Group.

[46] Krystal K. Lum, Timothy R. Howard, Catherina Pan, and Ileana M. Cristea. Charge-Mediated Pyrin Oligomerization Nucleates Antiviral IFI16 Sensing of Herpesvirus DNA. *mBio*, 10(4):10.1128/mbio.01428–19, July 2019.

[47] Jessica F. Almine, Craig A. J. O'Hare, Gillian Dunphy, Ismar R. Haga, Rangeetha J. Naik, Abdelmadjid Atrih, Dympna J. Connolly, Jordan Taylor, Ian R. Kelsall, Andrew G. Bowie, Philippa M. Beard, and Leonie Unterholzner. IFI16 and cGAS cooperate in the activation of STING during DNA sensing in human keratinocytes. *Nature Communications*, 8:14392, February 2017.

[48] K. L. Jønsson, A. Laustsen, C. Krapp, K. A. Skipper, K. Thavachelvam, D. Hotter, J. H. Egedal, M. Kjolby, P. Mohammadi, T. Prabakaran, L. K. Sørensen, C. Sun, S. B. Jensen, C. K. Holm, R. J. Lebbink, M. Johannsen, M. Nyegaard, J. G. Mikkelsen, F. Kirchhoff, S. R. Paludan, and M. R. Jakobsen. IFI16 is required for DNA sensing in human macrophages by promoting production and function of cGAMP. *Nature Communications*, 8(1):14391, February 2017.

[49] Byungil Kim, Sarah Arcos, Katherine Rothamel, Jeffrey Jian, Kristie L. Rose, W. Hayes McDonald, Yuqi Bian, Seth Reasoner, Nicholas J. Barrows, Shelton Bradrick, Mariano A. Garcia-Blanco, and Manuel Ascano. Discovery of Widespread Host Protein Interactions with the Pre-replicated Genome of CHIKV Using VIR-CLASP. *Molecular Cell*, 78(4):624–640.e7, May 2020. Publisher: Elsevier.

[50] Manuel Garcia-Moreno, Marko Noerenberg, Shuai Ni, Aino I. Järvelin, Esther González-Almela, Caroline E. Lenz, Marcel Bach-Pages, Victoria Cox, Rosario Avolio, Thomas Davis, Svenja Hester, Thibault J. M. Sohier, Bingnan Li, Gregory Heikel, Gracjan Michlewski, Miguel A. Sanz, Luis Carrasco, Emiliano P. Ricci, Vicent Pelechano, Ilan Davis, Bernd Fischer, Shabaz Mohammed, and Alfredo Castello. System-wide Profiling of RNA-Binding Proteins Uncovers Key Regulators of Virus Infection. *Molecular Cell*, 74(1):196–211.e11, April 2019. Publisher: Elsevier.

[51] Zhimin Jiang, Fanhua Wei, Yuying Zhang, Tong Wang, Weihua Gao, Shufang Yu, Honglei Sun, Juan Pu, Yipeng Sun, Mingyang Wang, Qi Tong, Chengjiang Gao, Kin-Chow Chang, and Jinhua Liu. IFI16 directly senses viral RNA and enhances RIG-I transcription and activation to restrict influenza virus infection. *Nature Microbiology*, 6(7):932–945, July 2021.

[52] Gideon Schreiber and Jacob Piehler. The molecular basis for functional plasticity in type I interferon signaling. *Trends in Immunology*, 36(3):139–149, March 2015.

[53] L. Velazquez, M. Fellous, G. R. Stark, and S. Pellegrini. A protein tyrosine kinase in the interferon alpha/beta signaling pathway. *Cell*, 70(2):313–322, July 1992.

[54] George R. Stark and James E. Darnell. The JAK-STAT Pathway at Twenty. *Immunity*, 36(4):503–514, April 2012.

[55] Sara Mostafavi, Hideyuki Yoshida, Devapregasan Moodley, Hugo LeBoité, Katherine Rothamel, Towfique Raj, Chun Jimmie Ye, Nicolas Chevrier, Shen-Ying Zhang, Ting Feng, Mark Lee, Jean-Laurent Casanova, James D. Clark, Martin Hegen, Jean-Baptiste Telliez, Nir Hacohen, Philip L. De Jager, Aviv Regev, Diane Mathis, and Christophe Benoist. Parsing the Interferon Transcriptional Network and Its Disease Associations. *Cell*, 164(3):564–578, January 2016. Publisher: Elsevier.

[56] Zhijiang Yan, Kairong Cui, Darryl M. Murray, Chen Ling, Yutong Xue, Amy Gerstein, Ramon Parsons, Keji Zhao, and Weidong Wang. PBAF chromatin-remodeling complex requires a novel specificity subunit, BAF200, to regulate expression of selective interferon-responsive genes. *Genes & Development*, 19(14):1662–1667, July 2005.

[57] David E. Levy, Isabelle Marié, Eric Smith, and Arun Prakash. Enhancement and diversification of IFN induction by IRF-7-mediated positive feedback. *Journal of Interferon & Cytokine Research: The Official Journal of the International Society for Interferon and Cytokine Research*, 22(1):87–93, January 2002.

[58] Dakang Xu, Michelle Holko, Anthony J. Sadler, Bernadette Scott, Shigeki Higashiyama, Windy Berkofsky-Fessler, Melanie J. McConnell, Pier Paolo Pandolfi, Jonathan D. Licht, and Bryan R. G. Williams. Promyelocytic leukemia zinc finger protein regulates interferon-mediated innate immunity. *Immunity*, 30(6):802–816, June 2009.

[59] Ioannis Tassiulas, Xiaoyu Hu, Hao Ho, Yogita Kashyap, Paul Paik, Yongmei Hu, Clifford A. Lowell, and Lionel B. Ivashkiv. Amplification of IFN-α-induced STAT1 activation and inflammatory function by Syk and ITAM-containing adaptors. *Nature Immunology*, 5(11):1181–1189, November 2004.

[60] Wei-Bei Wang, David E. Levy, and Chien-Kuo Lee. STAT3 Negatively Regulates Type I IFN-Mediated Antiviral Response. *The Journal of Immunology*, 187(5):2578–2585, September 2011.

[61] Hao H. Ho and Lionel B. Ivashkiv. Role of STAT3 in type I interferon responses. Negative regulation of STAT1-dependent inflammatory gene activation. *The Journal of Biological Chemistry*, 281(20):14111–14118, May 2006.

[62] Sze-Ling Ng, Brad A. Friedman, Sonja Schmid, Jason Gertz, Richard M. Myers, Benjamin R. Tenoever, and Tom Maniatis. IκB kinase epsilon (IKK(epsilon)) regulates the balance between type I and type II interferon responses. *Proceedings of the National Academy of Sciences of the United States of America*, 108(52):21170–21175, December 2011.

[63] Ekaterini Platanitis, Duygu Demiroz, Anja Schneller, Katrin Fischer, Christophe Capelle, Markus Hartl, Thomas Gossenreiter, Mathias Müller, Maria Novatchkova, and Thomas Decker. A molecular switch from STAT2-IRF9 to ISGF3 underlies interferon-induced gene transcription. *Nature Communications*, 10(1):2921, July 2019. Publisher: Nature Publishing Group.

[64] S. Uddin, B. Majchrzak, J. Woodson, P. Arunkumar, Y. Alsayed, R. Pine, P. R. Young, E. N. Fish, and L. C. Platanias. Activation of the p38 mitogen-activated protein kinase by type I interferons. *The Journal of Biological Chemistry*, 274(42):30127–30131, October 1999.

[65] Barbara Kroczynska, Surinder Kaur, Efstratios Katsoulidis, Beata Majchrzak-Kita, Antonella Sassano, Sara C. Kozma, Eleanor N. Fish, and Leonidas C. Platanias. Interferon-dependent engagement of eukaryotic initiation factor 4B via S6 kinase (S6K)- and ribosomal protein S6K-mediated signals. *Molecular and Cellular Biology*, 29(10):2865–2875, May 2009.

[66] Sonali Joshi, Surinder Kaur, Amanda J. Redig, Katy Goldsborough, Kevin David, Takeshi Ueda, Rie Watanabe-Fukunaga, Darren P. Baker, Eleanor N. Fish, Rikiro Fukunaga, and Leonidas C. Platanias. Type I interferon (IFN)-dependent activation of Mnk1 and its role in the generation of growth inhibitory responses. *Proceedings of the National Academy of Sciences of the United States of America*, 106(29):12097–12102, July 2009.

[67] Barney Viengkhou, Melanie Y. White, Stuart J. Cordwell, Iain L. Campbell, and Markus J. Hofer. A novel phosphoproteomic landscape evoked in response to type I interferon in the brain and in glial cells. *Journal of Neuroinflammation*, 18(1):237, October 2021.

[68] Larissa A. Durfee, Nancy Lyon, Kyungwoon Seo, and Jon M. Huibregtse. The ISG15 Conjugation System Broadly Targets Newly Synthesized Proteins: Implications for the Antiviral Function of ISG15. *Molecular Cell*, 38(5):722–732, June 2010.

[69] Fabien Thery, Denzel Eggermont, and Francis Impens. Proteomics Mapping of the ISGylation Landscape in Innate Immunity. *Frontiers in Immunology*, 12, 2021.

[70] Craig H. Kerr, Michael A. Skinnider, Daniel D. T. Andrews, Angel M. Madero, Queenie W. T. Chan, R. Greg Stacey, Nikolay Stoynov, Eric Jan, and Leonard J. Foster. Dynamic rewiring of the human interactome by interferon signaling. *Genome Biology*, 21(1):140, June 2020.

[71] Andreas Pichlmair, Caroline Lassnig, Carol-Ann Eberle, Maria W. Górna, Christoph L. Baumann, Thomas R. Burkard, Tilmann Bürckstümmer, Adrijana Stefanovic, Sigurd Krieger, Keiryn L. Bennett, Thomas Rülicke, Friedemann Weber, Jacques Colinge, Mathias Müller, and Giulio Superti-Furga. IFIT1 is an antiviral protein that recognizes 5'-triphosphate RNA. *Nature Immunology*, 12(7):624–630, July 2011. Publisher: Nature Publishing Group.

[72] William M. Schneider, Meike Dittmann Chevillotte, and Charles M. Rice. Interferon-Stimulated Genes: A Complex Web of Host Defenses. *Annual Review of Immunology*, 32(1):513–545, 2014. _eprint: https://doi.org/10.1146/annurev-immunol-032713-120231.

[73] John C. Hall and Antony Rosen. Type I interferons: crucial participants in disease amplification in autoimmunity. *Nature Reviews Rheumatology*, 6(1):40–49, January 2010. Publisher: Nature Publishing Group.

[74] Yanick J. Crow and Daniel B. Stetson. The type I interferonopathies: 10 years on. *Nature Reviews Immunology*, 22(8):471–483, August 2022. Publisher: Nature Publishing Group.

[75] Sabyasachi Bhattacharya, Juan Qian, Christos Tzimas, Darren P. Baker, Constantinos Koumenis, J. Alan Diehl, and Serge Y. Fuchs. Role of p38 Protein Kinase in the Ligand-independent Ubiquitination and Down-regulation of the IFNAR1 Chain of Type I Interferon Receptor. *The Journal of Biological Chemistry*, 286(25):22069–22076, June 2011.

[76] M. You, D. H. Yu, and G. S. Feng. Shp-2 tyrosine phosphatase functions as a negative regulator of the interferon-stimulated Jak/STAT pathway. *Molecular and Cellular Biology*, 19(3):2416–2424, March 1999.

[77] Richard S. Rogers, Curt M. Horvath, and Michael J. Matunis. SUMO modification of STAT1 and its role in PIAS-mediated inhibition of gene activation. *The Journal of Biological Chemistry*, 278(32):30091–30097, August 2003.

[78] Michael P. Malakhov, Oxana A. Malakhova, Keun Il Kim, Kenneth J. Ritchie, and Dong-Er Zhang. UBP43 (USP18) specifically removes ISG15 from conjugated proteins. *The Journal of Biological Chemistry*, 277(12):9976–9981, March 2002.

[79] Kenneth J. Ritchie, Chang S. Hahn, Keun Il Kim, Ming Yan, Dabralee Rosario, Li Li, Juan Carlos de la Torre, and Dong-Er Zhang. Role of ISG15 protease UBP43 (USP18) in innate immunity to viral infection. *Nature Medicine*, 10(12):1374–1378, December 2004.

[80] Oxana A. Malakhova, Keun Il Kim, Jiann-Kae Luo, Weiguo Zou, K. G. Suresh Kumar, Serge Y. Fuchs, Ke Shuai, and Dong-Er Zhang. UBP43 is a novel regulator of interferon signaling independent of its ISG15 isopeptidase activity. *The EMBO journal*, 25(11):2358–2367, June 2006.

[81] Su-Yang Liu, Roghiyh Aliyari, Kelechi Chikere, Guangming Li, Matthew D. Marsden, Jennifer K. Smith, Olivier Pernet, Haitao Guo, Rebecca Nusbaum, Jerome A. Zack, Alexander N. Freiberg, Lishan Su, Benhur Lee, and Genhong Cheng. Interferon-Inducible Cholesterol-25-Hydroxylase Broadly Inhibits Viral Entry by Production of 25-Hydroxycholesterol. *Immunity*, 38(1):92–105, January 2013.

[82] Mathieu Blanc, Wei Yuan Hsieh, Kevin A. Robertson, Kai A. Kropp, Thorsten Forster, Guanghou Shui, Paul Lacaze, Steven Watterson, Samantha J. Griffiths, Nathanael J. Spann, Anna Meljon, Simon Talbot, Kathiresan Krishnan, Douglas F. Covey, Markus R. Wenk, Marie Craigon, Zsolts Ruzsics, Jürgen Haas, Ana Angulo, William J. Griffiths, Christopher K. Glass, Yuqin Wang, and Peter Ghazal. The transcription factor STAT-1 couples macrophage synthesis of 25-hydroxycholesterol to the interferon antiviral response. *Immunity*, 38(1):106–118, January 2013.

[83] Ruochen Zang, James Brett Case, Eylan Yutuc, Xiucui Ma, Sheng Shen, Maria Florencia Gomez Castro, Zhuoming Liu, Qiru Zeng, Haiyan Zhao, Juhee Son, Paul W. Rothlauf, Alex J. B. Kreutzberger, Gaopeng Hou, Hu Zhang, Sayantan Bose, Xin Wang, Michael D. Vahey, Kartik Mani, William J. Griffiths, Tom Kirchhausen, Daved H. Fremont, Haitao Guo, Abhinav Diwan, Yuqin Wang, Michael S. Diamond, Sean P. J. Whelan, and Siyuan Ding. Cholesterol 25-hydroxylase suppresses SARS-CoV-2 replication by blocking membrane fusion. *Proceedings of the National Academy of Sciences of the United States of America*, 117(50):32105–32113, December 2020.

[84] Yongzhi Chen, Shanshan Wang, Zhaohong Yi, Huabin Tian, Roghiyh Aliyari, Yanhua Li, Gang Chen, Ping Liu, Jin Zhong, Xinwen Chen, Peishuang Du, Lishan Su, F. Xiao-Feng Qin, Hongyu Deng, and Genhong Cheng. Interferon-inducible cholesterol-25-hydroxylase inhibits hepatitis C virus replication via distinct mechanisms. *Scientific Reports*, 4:7242, December 2014.

[85] Abraham L. Brass, I.-Chueh Huang, Yair Benita, Sinu P. John, Manoj N. Krishnan, Eric M. Feeley, Bethany J. Ryan, Jessica L. Weyer, Louise van der Weyden, Erol Fikrig, David J. Adams, Ramnik J. Xavier, Michael Farzan, and Stephen J. Elledge. The IFITM proteins mediate cellular resistance to influenza A H1N1 virus, West Nile virus, and dengue virus. *Cell*, 139(7):1243–1254, December 2009.

[86] I.-Chueh Huang, Charles C. Bailey, Jessica L. Weyer, Sheli R. Radoshitzky, Michelle M. Becker, Jessica J. Chiang, Abraham L. Brass, Asim A. Ahmed, Xiaoli Chi, Lian Dong, Lindsay E. Longobardi, Dutch Boltz, Jens H. Kuhn, Stephen J. Elledge, Sina Bavari, Mark R. Denison, Hyeryun Choe, and Michael Farzan. Distinct Patterns of IFITM-Mediated Restriction of Filoviruses, SARS Coronavirus, and Influenza A Virus. *PLOS Pathogens*, 7(1):e1001258, January 2011. Publisher: Public Library of Science.

[87] Xiangyang Guo, Jan Steinkühler, Mariana Marin, Xiang Li, Wuyuan Lu, Rumiana Dimova, and Gregory B. Melikyan. Interferon-Induced Transmembrane Protein 3 Blocks Fusion of Diverse Enveloped Viruses by Altering Mechanical Properties of Cell Membranes. *ACS nano*, 15(5):8155–8170, May 2021.

[88] Samad Amini-Bavil-Olyaee, Youn Jung Choi, Jun Han Lee, Mude Shi, I.-Chueh Huang, Michael Farzan, and Jae U. Jung. The antiviral effector IFITM3 disrupts intracellular cholesterol homeostasis to block viral entry. *Cell Host & Microbe*, 13(4):452–464, April 2013.

[89] Emily Yang and Melody M. H. Li. All About the RNA: Interferon-Stimulated Genes That Interfere With Viral RNA Processes. *Frontiers in Immunology*, 11, 2020.

[90] P. Lengyel. Biochemistry of interferons and their actions. *Annual Review of Biochemistry*, 51:251–282, 1982.

[91] James L. Cole. Activation of PKR: an open and shut case? *Trends in Biochemical Sciences*, 32(2):57–62, February 2007. Publisher: Elsevier.

[92] Madhusudan Dey, Chune Cao, Arvin C. Dar, Tomohiko Tamura, Keiko Ozato, Frank Sicheri, and Thomas E. Dever. Mechanistic link between PKR dimerization, autophosphorylation, and eIF2alpha substrate recognition. *Cell*, 122(6):901–913, September 2005.

[93] Andrew M. Bogorad, Kai Ying Lin, and Assen Marintchev. Novel mechanisms of eIF2B action and regulation by eIF2α phosphorylation. *Nucleic Acids Research*, 45(20):11962–11979, November 2017.

[94] Sylvain de Breyne, Yingpu Yu, Anett Unbehaun, Tatyana V. Pestova, and Christopher U. T. Hellen. Direct functional interaction of initiation factor eIF4G with type 1 internal ribosomal entry sites. *Proceedings of the National Academy of Sciences*, 106(23):9197–9202, June 2009. Publisher: Proceedings of the National Academy of Sciences.

[95] Nathalie Grandvaux, Marc J. Servant, Benjamin tenOever, Ganes C. Sen, Siddarth Balachandran, Glen N. Barber, Rongtuan Lin, and John Hiscott. Transcriptional profiling of interferon regulatory factor 3 target genes: direct involvement in the regulation of interferon-stimulated genes. *Journal of Virology*, 76(11):5532–5539, June 2002.

[96] Saumendra N. Sarkar and Ganes C. Sen. Novel functions of proteins encoded by viral stress-inducible genes. *Pharmacology & Therapeutics*, 103(3):245–259, September 2004.

[97] Taishi Kimura, Hiroshi Katoh, Hisako Kayama, Hiroyuki Saiga, Megumi Okuyama, Toru Okamoto, Eiji Umemoto, Yoshiharu Matsuura, Masahiro Yamamoto, and Kiyoshi Takeda. Ifit1 Inhibits Japanese Encephalitis Virus Replication through Binding to 5' Capped 2'-O Unmethylated RNA. *Journal of Virology*, 87(18):9997–10003, September 2013.

[98] Matthias Habjan, Philipp Hubel, Livia Lacerda, Christian Benda, Cathleen Holze, Christian H. Eberl, Angelika Mann, Eveline Kindler, Cristina Gil-Cruz, John Ziebuhr, Volker Thiel, and Andreas Pichlmair. Sequestration by IFIT1 Impairs Translation of 2'O-unmethylated Capped RNA. *PLOS Pathogens*, 9(10):e1003663, October 2013. Publisher: Public Library of Science.

[99] Renata C Fleith, Harriet V Mears, Xin Yun Leong, Thomas J Sanford, Edward Emmott, Stephen C Graham, Daniel S Mansur, and Trevor R Sweeney. IFIT3 and IFIT2/3 promote IFIT1-mediated translation inhibition by enhancing binding to non-self RNA. *Nucleic Acids Research*, 46(10):5269–5285, June 2018.

[100] Nicholas A. Mathieu, Ermela Paparisto, Stephen D. Barr, and Donald E. Spratt. HERC5 and the ISGylation Pathway: Critical Modulators of the Antiviral Immune Response. *Viruses*, 13(6), June 2021. Publisher: Multidisciplinary Digital Publishing Institute (MDPI).

[101] Anna Rahnefeld, Karin Klingel, Anett Schuermann, Nicola L. Diny, Nadine Althof, Anika Lindner, Philipp Bleienheuft, Konstantinos Savvatis, Dorota Respondek, Elisa Opitz, Lars Ketscher, Martina Sauter, Ulrike Seifert, Carsten Tschöpe, Wolfgang Poller, Klaus-Peter Knobeloch, and Antje Voigt. Ubiquitin-like protein ISG15 (interferon-stimulated gene of 15 kDa) in host defense against heart failure in a mouse model of virus-induced cardiomyopathy. *Circulation*, 130(18):1589–1600, October 2014.

[102] Patrick Eldin, Laura Papon, Alexandra Oteiza, Emiliana Brocchi, T. Glen Lawson, and Nadir Mechti. TRIM22 E3 ubiquitin ligase activity is required to mediate antiviral activity against encephalomyocarditis virus. *The Journal of General Virology*, 90(Pt 3):536–545, March 2009.

[103] Andrea Di Pietro, Anna Kajaste-Rudnitski, Alexandra Oteiza, Lucia Nicora, Greg J. Towers, Nadir Mechti, and Elisa Vicenzi. TRIM22 inhibits influenza A virus infection by targeting the viral nucleoprotein for degradation. *Journal of Virology*, 87(8):4523–4533, April 2013.

[104] Xiu Luo, Xinlu Wang, Yina Gao, Jingpeng Zhu, Songqing Liu, Guangxia Gao, and Pu Gao. Molecular Mechanism of RNA Recognition by Zinc-Finger Antiviral Protein. *Cell Reports*, 30(1):46–52.e4, January 2020. Publisher: Elsevier.

[105] Mattia Ficarelli, Harry Wilson, Rui Pedro Galão, Michela Mazzon, Irati Antzin-Anduetza, Mark Marsh, Stuart JD Neil, and Chad M Swanson. KHNYN is essential for the zinc finger antiviral protein (ZAP) to restrict HIV-1 containing clustered CpG dinucleotides. *eLife*, 8:e46767, 2019.

[106] Yiping Zhu, Xinlu Wang, Stephen P Goff, and Guangxia Gao. Translational repression precedes and is required for ZAP-mediated mRNA decay. *The EMBO Journal*, 31(21):4236–4246, November 2012. Num Pages: 4246 Publisher: John Wiley & Sons, Ltd.

[107] Tatsuya Kozaki, Michihiro Takahama, Takuma Misawa, Yoshiharu Matsuura, Shizuo Akira, and Tatsuya Saitoh. Role of zinc-finger anti-viral protein in host defense against Sindbis virus. *International Immunology*, 27(7):357–364, July 2015.

[108] Chean Ring Leong, Kenji Funami, Hiroyuki Oshiumi, Deng Mengao, Hiromi Takaki, Misako Matsumoto, Hussein H. Aly, Koichi Watashi, Kazuaki Chayama, and Tsukasa Seya. Interferon-stimulated gene of 20 kDa protein (ISG20) degrades RNA of hepatitis B virus to impede the replication of HBV in vitro and in vivo. *Oncotarget*, 7(42):68179–68193, October 2016.

[109] Hasan Imam, Geon-Woo Kim, Saiful Anam Mir, Mohsin Khan, and Aleem Siddiqui. Interferon-stimulated gene 20 (ISG20) selectively degrades N6-methyladenosine modified Hepatitis B Virus transcripts. *PLoS pathogens*, 16(2):e1008338, February 2020.

[110] Roni Winkler, Ella Gillis, Lior Lasman, Modi Safra, Shay Geula, Clara Soyris, Aharon Nachshon, Julie Tai-Schmiedel, Nehemya Friedman, Vu Thuy Khanh Le-Trilling, Mirko

Trilling, Michal Mandelboim, Jacob H. Hanna, Schraga Schwartz, and Noam Stern-Ginossar. m6A modification controls the innate immune response to infection by targeting type I interferons. *Nature Immunology*, 20(2):173–182, February 2019. Publisher: Nature Publishing Group.

[111] Christopher M. Weiss, Derek W. Trobaugh, Chengqun Sun, Tiffany M. Lucas, Michael S. Diamond, Kate D. Ryman, and William B. Klimstra. The Interferon-Induced Exonuclease ISG20 Exerts Antiviral Activity through Upregulation of Type I Interferon Response Proteins. *mSphere*, 3(5):e00209–18, September 2018.

[112] Nannan Wu, Xuan-Nhi Nguyen, Li Wang, Romain Appourchaux, Chengfei Zhang, Baptiste Panthu, Henri Gruffat, Chloé Journo, Sandrine Alais, Juliang Qin, Na Zhang, Kevin Tartour, Frédéric Catez, Renaud Mahieux, Theophile Ohlmann, Mingyao Liu, Bing Du, and Andrea Cimarelli. The interferon stimulated gene 20 protein (ISG20) is an innate defense antiviral factor that discriminates self versus non-self translation. *PLOS Pathogens*, 15(10):e1008093, October 2019. Publisher: Public Library of Science.

[113] Mikkel Søes Ibsen, Hans Henrik Gad, Karthiga Thavachelvam, Thomas Boesen, Philippe Desprès, and Rune Hartmann. The 2'-5'-oligoadenylate synthetase 3 enzyme potently synthesizes the 2'-5'-oligoadenylates required for RNase L activation. *Journal of Virology*, 88(24):14222–14231, December 2014.

[114] Yize Li, Shuvojit Banerjee, Yuyan Wang, Stephen A. Goldstein, Beihua Dong, Christina Gaughan, Robert H. Silverman, and Susan R. Weiss. Activation of RNase L is dependent on OAS3 expression during infection with diverse human viruses. *Proceedings of the National Academy of Sciences*, 113(8):2241–2246, February 2016. Publisher: Proceedings of the National Academy of Sciences.

[115] Krishnamurthy Malathi, Takeshi Saito, Nannette Crochet, David J. Barton, Michael Gale, and Robert H. Silverman. RNase L releases a small RNA from HCV RNA that refolds into a potent PAMP. *RNA (New York, N.Y.)*, 16(11):2108–2119, November 2010.

[116] Jyoti Batra, Judd F. Hultquist, Dandan Liu, Olena Shtanko, John Von Dollen, Laura Satkamp, Gwendolyn M. Jang, Priya Luthra, Toni M. Schwarz, Gabriel I. Small, Eusondia Arnett, Manu Anantpadma, Ann Reyes, Daisy W. Leung, Robyn Kaake, Paige Haas, Carson B. Schmidt, Larry S. Schlesinger, Douglas J. LaCount, Robert A. Davey, Gaya K.

Amarasinghe, Christopher F. Basler, and Nevan J. Krogan. Protein Interaction Mapping Identifies RBBP6 as a Negative Regulator of Ebola Virus Replication. *Cell*, 175(7):1917–1930.e13, December 2018.

[117] Xiuyan Wang, Ella R. Hinson, and Peter Cresswell. The interferon-inducible protein viperin inhibits influenza virus release by perturbing lipid rafts. *Cell Host & Microbe*, 2(2):96–105, August 2007.

[118] Karla J. Helbig, Nicholas S. Eyre, Evelyn Yip, Sumudu Narayana, Kui Li, Guillaume Fiches, Erin M. McCartney, Rohit K. Jangra, Stanley M. Lemon, and Michael R. Beard. The antiviral protein viperin inhibits hepatitis C virus replication via interaction with nonstructural protein 5A. *Hepatology (Baltimore, Md.)*, 54(5):1506–1517, November 2011.

[119] Karla J. Helbig, Jillian M. Carr, Julie K. Calvert, Satiya Wati, Jennifer N. Clarke, Nicholas S. Eyre, Sumudu K. Narayana, Guillaume N. Fiches, Erin M. McCartney, and Michael R. Beard. Viperin Is Induced following Dengue Virus Type-2 (DENV-2) Infection and Has Anti-viral Actions Requiring the C-terminal End of Viperin. *PLoS Neglected Tropical Diseases*, 7(4):e2178, April 2013.

[120] Anthony S. Gizzi, Tyler L. Grove, Jamie J. Arnold, Joyce Jose, Rohit K. Jangra, Scott J. Garforth, Quan Du, Sean M. Cahill, Natalya G. Dulyaninova, James D. Love, Kartik Chandran, Anne R. Bresnick, Craig E. Cameron, and Steven C. Almo. A naturally occurring antiviral ribonucleotide encoded by the human genome. *Nature*, 558(7711):610–614, June 2018. Number: 7711 Publisher: Nature Publishing Group.

[121] Chunyu Wei, Caishang Zheng, Jianhong Sun, Dan Luo, Yan Tang, Yuan Zhang, Xianliang Ke, Yan Liu, Zhenhua Zheng, and Hanzhong Wang. Viperin Inhibits Enterovirus A71 Replication by Interacting with Viral 2C Protein. *Viruses*, 11(1):13, December 2018.

[122] David Perez-Caballero, Trinity Zang, Alaleh Ebrahimi, Matthew W. McNatt, Devon A. Gregory, Marc C. Johnson, and Paul D. Bieniasz. Tetherin inhibits HIV-1 release by directly tethering virions to cells. *Cell*, 139(3):499–511, October 2009.

[123] Guramrit Singh, Gabriel Pratt, Gene W. Yeo, and Melissa J. Moore. The Clothes Make the mRNA: Past and Present Trends in mRNP Fashion. *Annual Review of Biochemistry*, 84(Volume 84, 2015):325–354, June 2015. Publisher: Annual Reviews.

[124] Vivek Anantharaman, Eugene V. Koonin, and L. Aravind. Comparative genomics and evolution of proteins involved in RNA metabolism. *Nucleic Acids Research*, 30(7):1427–1464, April 2002.

[125] M W Hentze and P Argos. Homology between IRE-BP, a regulatory RNA-binding protein, aconitase, and isopropylmalate isomerase. *Nucleic acids research*, 19(8):1739–1740, April 1991.

[126] Alfredo Castello, Matthias W. Hentze, and Thomas Preiss. Metabolic Enzymes Enjoying New Partnerships as RNA-Binding Proteins. *Trends in Endocrinology & Metabolism*, 26(12):746–757, December 2015. Publisher: Elsevier.

[127] Matthias W. Hentze and Thomas Preiss. The REM phase of gene regulation. *Trends in Biochemical Sciences*, 35(8):423–426, August 2010.

[128] Niels H. Gehring, Elmar Wahle, and Utz Fischer. Deciphering the mRNP Code: RNA-Bound Determinants of Post-Transcriptional Gene Regulation. *Trends in Biochemical Sciences*, 42(5):369–382, May 2017.

[129] Thomas Gonatopoulos-Pournatzis and Victoria H. Cowling. Cap-binding complex (CBC). *Biochemical Journal*, 457(Pt 2):231–242, January 2014.

[130] Jailson Brito Querido, Irene Díaz-López, and V. Ramakrishnan. The molecular basis of translation initiation and its regulation in eukaryotes. *Nature Reviews Molecular Cell Biology*, 25(3):168–186, March 2024. Number: 3 Publisher: Nature Publishing Group.

[131] A. Gregory Matera and Zefeng Wang. A day in the life of the spliceosome. *Nature Reviews Molecular Cell Biology*, 15(2):108–121, February 2014. Publisher: Nature Publishing Group.

[132] Xiang-Dong Fu and Manuel Ares. Context-dependent control of alternative splicing by RNA-binding proteins. *Nature Reviews Genetics*, 15(10):689–701, October 2014. Publisher: Nature Publishing Group.

[133] Niels H. Gehring, Styliani Lamprinaki, Andreas E. Kulozik, and Matthias W. Hentze. Disassembly of Exon Junction Complexes by PYM. *Cell*, 137(3):536–548, May 2009.

[134] Bin Tian and James L. Manley. Alternative cleavage and polyadenylation: the long and short of it. *Trends in Biochemical Sciences*, 38(6):312–320, June 2013.

[135] Donny D. Licatalosi, Aldo Mele, John J. Fak, Jernej Ule, Melis Kayikci, Sung Wook Chi, Tyson A. Clark, Anthony C. Schweitzer, John E. Blume, Xuning Wang, Jennifer C. Darnell, and Robert B. Darnell. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature*, 456(7221):464–469, November 2008. Publisher: Nature Publishing Group.

[136] Xin Jia, Shaochun Yuan, Yao Wang, Yonggui Fu, Yong Ge, Yutong Ge, Xihong Lan, Yuchao Feng, Feifei Qiu, Peiyi Li, Shangwu Chen, and Anlong Xu. The role of alternative polyadenylation in the antiviral innate immune response. *Nature Communications*, 8:14605, February 2017.

[137] Masahiro Morita, Lian Wee Ler, Marc R. Fabian, Nadeem Siddiqui, Michael Mullin, Valerie C. Henderson, Tommy Alain, Bruno D. Fonseca, Galina Karashchuk, Christopher F. Bennett, Tomohiro Kabuta, Shinji Higashi, Ola Larsson, Ivan Topisirovic, Robert J. Smith, Anne-Claude Gingras, and Nahum Sonenberg. A Novel 4EHP-GIGYF2 Translational Repressor Complex Is Essential for Mammalian Development. *Molecular and Cellular Biology*, 32(17):3585–3593, September 2012.

[138] Ramona Weber, Min-Yi Chung, Csilla Keskeny, Ulrike Zinnall, Markus Landthaler, Eugene Valkov, Elisa Izaurralde, and Cátia Igreja. 4EHP and GIGYF1/2 Mediate Translation-Coupled Messenger RNA Decay. *Cell Reports*, 33(2), October 2020. Publisher: Elsevier.

[139] Cinthia C Amaya Ramirez, Petra Hubbe, Nicolas Mandel, and Julien Béthune. 4EHP-independent repression of endogenous mRNAs by the RNA-binding protein GIGYF2. *Nucleic Acids Research*, 46(11):5792–5808, June 2018.

[140] Ester Carballo, Wi S. Lai, and Perry J. Blackshear. Feedback Inhibition of Macrophage Tumor Necrosis Factor-α Production by Tristetraprolin. *Science*, 281(5379):1001–1005, August 1998. Publisher: American Association for the Advancement of Science.

[141] Marc R. Fabian, Filipp Frank, Christopher Rouya, Nadeem Siddiqui, Wi S. Lai, Alexey Karetnikov, Perry J. Blackshear, Bhushan Nagar, and Nahum Sonenberg. Structural basis for the recruitment of the human CCR4–NOT deadenylase complex by tristetraprolin. *Nature Structural & Molecular Biology*, 20(6):735–739, June 2013. Publisher: Nature Publishing Group.

[142] Mauro Costa-Mattioli and Peter Walter. The integrated stress response: From mechanism to disease. *Science (New York, N.Y.)*, 368(6489):eaat5314, April 2020.

[143] Finn McCluggage and Archa H. Fox. Paraspeckle nuclear condensates: Global sensors of cell stress? *BioEssays*, 43(5):2000245, 2021. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/bies.202000245.

[144] Rastislav Horos, Magdalena Büscher, Rozemarijn Kleinendorst, Anne-Marie Alleaume, Abul K. Tarafder, Thomas Schwarzl, Dmytro Dziuba, Christian Tischer, Elisabeth M. Zielonka, Asli Adak, Alfredo Castello, Wolfgang Huber, Carsten Sachse, and Matthias W. Hentze. The Small Non-coding Vault RNA1-1 Acts as a Riboregulator of Autophagy. *Cell*, 176(5):1054–1067.e12, February 2019.

[145] Ewelina Golec, Liza Lind, Munazza Qayyum, Anna M. Blom, and Ben C. King. The Noncoding RNA nc886 Regulates PKR Signaling and Cytokine Production in Human Cells. *Journal of Immunology (Baltimore, Md.: 1950)*, 202(1):131–141, January 2019.

[146] Guerline Clerzius, Jean-François Gélinas, and Anne Gatignol. Multiple levels of PKR inhibition during HIV-1 replication. *Reviews in Medical Virology*, 21(1):42–53, 2011. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/rmv.674.

[147] Bradley M. Lunde, Claire Moore, and Gabriele Varani. RNA-binding proteins: modular design for efficient function. *Nature Reviews Molecular Cell Biology*, 8(6):479–490, June 2007. Number: 6 Publisher: Nature Publishing Group.

[148] Meredith Corley, Margaret C. Burns, and Gene W. Yeo. How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. *Molecular Cell*, 78(1):9–29, April 2020. Publisher: Elsevier.

[149] Xiaoqiang Wang, Juanita McLachlan, Phillip D. Zamore, and Traci M. Tanaka Hall. Modular recognition of RNA by a human pumilio-homology domain. *Cell*, 110(4):501–512, August 2002.

[150] Corine Vernet and Karen Artzt. STAR, a gene family involved in signal transduction and activation of RNA. *Trends in Genetics*, 13(12):479–484, December 1997. Publisher: Elsevier.

[151] Anna Balcerak, Alicja Trebinska-Stryjewska, Ryszard Konopinski, Maciej Wakula, and Ewa Anna Grzybowska. RNA–protein interactions: disorder, moonlighting and junk contribute to eukaryotic complexity. *Open Biology*, 9(6):190096, June 2019. Publisher: Royal Society.

[152] Alfredo Castello, Bernd Fischer, Christian K. Frese, Rastislav Horos, Anne-Marie Alleaume, Sophia Foehr, Tomaz Curk, Jeroen Krijgsveld, and Matthias W. Hentze. Comprehensive Identification of RNA-Binding Domains in Human Cells. *Molecular Cell*, 63(4):696–710, August 2016. Publisher: Elsevier.

[153] Aino I. Järvelin, Marko Noerenberg, Ilan Davis, and Alfredo Castello. The new (dis)order in RNA regulation. *Cell Communication and Signaling*, 14(1):9, April 2016.

[154] Michael A. Jamros, Brandon E. Aubol, Malik M. Keshwani, Zhaiyi Zhang, Stefan Stamm, and Joseph A. Adams. Intra-domain Cross-talk Regulates Serine-arginine Protein Kinase 1-dependent Phosphorylation and Splicing Function of Transformer 2β1. *The Journal of Biological Chemistry*, 290(28):17269–17281, July 2015.

[155] Anh Tuân Phan, Vitaly Kuryavyi, Jennifer C. Darnell, Alexander Serganov, Ananya Majumdar, Serge Ilin, Tanya Raslin, Anna Polonskaia, Cynthia Chen, David Clain, Robert B. Darnell, and Dinshaw J. Patel. Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. *Nature structural & molecular biology*, 18(7):796–804, June 2011.

[156] Florian C. Oberstrass, Sigrid D. Auweter, Michèle Erat, Yann Hargous, Anke Henning, Philipp Wenter, Luc Reymond, Batoul Amir-Ahmady, Stefan Pitsch, Douglas L. Black, and Frédéric H.-T. Allain. Structure of PTB bound to RNA: specific binding and implications for splicing regulation. *Science (New York, N.Y.)*, 309(5743):2054–2057, September 2005.

[157] Benedikt M. Beckmann, Rastislav Horos, Bernd Fischer, Alfredo Castello, Katrin Eichelbaum, Anne-Marie Alleaume, Thomas Schwarzl, Tomaž Curk, Sophia Foehr, Wolfgang Huber, Jeroen Krijgsveld, and Matthias W. Hentze. The RNA-binding

proteomes from yeast to man harbour conserved enigmRBPs. *Nature Communications*, 6(1):10127, December 2015. Publisher: Nature Publishing Group.

[158] C. Tuerk and L. Gold. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science (New York, N.Y.)*, 249(4968):505–510, August 1990.

[159] Debashish Ray, Hilal Kazan, Esther T. Chan, Lourdes Peña Castillo, Sidharth Chaudhry, Shaheynoor Talukder, Benjamin J. Blencowe, Quaid Morris, and Timothy R. Hughes. Rapid and systematic analysis of the RNA recognition specificities of RNA-binding proteins. *Nature Biotechnology*, 27(7):667–670, July 2009.

[160] J. Matthew Taliaferro, Nicole J. Lambert, Peter H. Sudmant, Daniel Dominguez, Jason J. Merkin, Maria S. Alexis, Cassandra A. Bazile, and Christopher B. Burge. RNA Sequence Context Effects Measured In Vitro Predict In Vivo Protein Binding and Regulation. *Molecular Cell*, 64(2):294–306, October 2016.

[161] Debashish Ray, Hilal Kazan, Kate B. Cook, Matthew T. Weirauch, Hamed S. Najafabadi, Xiao Li, Serge Gueroussov, Mihai Albu, Hong Zheng, Ally Yang, Hong Na, Manuel Irimia, Leah H. Matzat, Ryan K. Dale, Sarah A. Smith, Christopher A. Yarosh, Seth M. Kelly, Behnam Nabet, Desirea Mecenas, Weimin Li, Rakesh S. Laishram, Mei Qiao, Howard D. Lipshitz, Fabio Piano, Anita H. Corbett, Russ P. Carstens, Brendan J. Frey, Richard A. Anderson, Kristen W. Lynch, Luiz O. F. Penalva, Elissa P. Lei, Andrew G. Fraser, Benjamin J. Blencowe, Quaid D. Morris, and Timothy R. Hughes. A compendium of RNA-binding motifs for decoding gene regulation. *Nature*, 499(7457):172–177, July 2013. Number: 7457 Publisher: Nature Publishing Group.

[162] Daniel Dominguez, Peter Freese, Maria S. Alexis, Amanda Su, Myles Hochman, Tsultrim Palden, Cassandra Bazile, Nicole J. Lambert, Eric L. Van Nostrand, Gabriel A. Pratt, Gene W. Yeo, Brenton R. Graveley, and Christopher B. Burge. Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. *Molecular Cell*, 70(5):854–867.e9, June 2018. Publisher: Elsevier.

[163] Markus Hafner, Maria Katsantoni, Tino Köster, James Marks, Joyita Mukherjee, Dorothee Staiger, Jernej Ule, and Mihaela Zavolan. CLIP and complementary methods. *Nature Reviews Methods Primers*, 1(1):20, December 2021.

[164] Lei Deng, Youzhi Liu, Yechuan Shi, Wenhao Zhang, Chun Yang, and Hui Liu. Deep neural networks for inferring binding sites of RNA-binding proteins by using distributed representations of RNA primary sequence and secondary structure. *BMC Genomics*, 21(13):866, December 2020.

[165] Lian Cui, Rui Ma, Jiangluyi Cai, Chunyuan Guo, Zeyu Chen, Lingling Yao, Yuanyuan Wang, Rui Fan, Xin Wang, and Yuling Shi. RNA modifications: importance in immune cell biology and related diseases. *Signal Transduction and Targeted Therapy*, 7(1):1–26, September 2022. Publisher: Nature Publishing Group.

[166] Matthias W. Hentze, Alfredo Castello, Thomas Schwarzl, and Thomas Preiss. A brave new world of RNA-binding proteins. *Nature Reviews Molecular Cell Biology*, 19(5):327–341, May 2018. Number: 5 Publisher: Nature Publishing Group.

[167] D. V. Sizova, V. G. Kolupaeva, T. V. Pestova, I. N. Shatsky, and C. U. Hellen. Specific interaction of eukaryotic translation initiation factor 3 with the 5' nontranslated regions of hepatitis C virus and classical swine fever virus RNAs. *Journal of Virology*, 72(6):4775–4782, June 1998.

[168] Falk Butter, Marion Scheibe, Mario Mörl, and Matthias Mann. Unbiased RNA–protein interaction screen by quantitative proteomics. *Proceedings of the National Academy of Sciences*, 106(26):10626–10631, June 2009.

[169] Tanja Scherrer, Nitish Mittal, Sarath Chandra Janga, and André P. Gerber. A Screen for RNA-Binding Proteins in Yeast Indicates Dual Functions for Many Enzymes. *PLOS ONE*, 5(11):e15499, November 2010. Publisher: Public Library of Science.

[170] Nikoleta G. Tsvetanova, Daniel M. Klass, Julia Salzman, and Patrick O. Brown. Proteome-Wide Search Reveals Unexpected RNA-Binding Proteins in Saccharomyces cerevisiae. *PLOS ONE*, 5(9):e12671, September 2010.

[171] Alfredo Castello, Bernd Fischer, Katrin Eichelbaum, Rastislav Horos, Benedikt M. Beckmann, Claudia Strein, Norman E. Davey, David T. Humphreys, Thomas Preiss, Lars M. Steinmetz, Jeroen Krijgsveld, and Matthias W. Hentze. Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell*, 149(6):1393–1406, June 2012. Publisher: Elsevier.

[172] Alexander G. Baltz, Mathias Munschauer, Björn Schwanhäusser, Alexandra Vasile, Yasuhiro Murakawa, Markus Schueler, Noah Youngs, Duncan Penfold-Brown, Kevin Drew, Miha Milek, Emanuel Wyler, Richard Bonneau, Matthias Selbach, Christoph Dieterich, and Markus Landthaler. The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts. *Molecular Cell*, 46(5):674–690, June 2012.

[173] I. G. Pashev, S. I. Dimitrov, and D. Angelov. Crosslinking proteins to nucleic acids by ultraviolet laser irradiation. *Trends in Biochemical Sciences*, 16(9):323–326, September 1991.

[174] Alfredo Castello, Rastislav Horos, Claudia Strein, Bernd Fischer, Katrin Eichelbaum, Lars M. Steinmetz, Jeroen Krijgsveld, and Matthias W. Hentze. System-wide identification of RNA-binding proteins by interactome capture. *Nature Protocols*, 8(3):491–500, March 2013. Number: 3 Publisher: Nature Publishing Group.

[175] Vladimir Despic, Mario Dejung, Mengting Gu, Jayanth Krishnan, Jing Zhang, Lydia Herzel, Korinna Straube, Mark B. Gerstein, Falk Butter, and Karla M. Neugebauer. Dynamic RNA-protein interactions underlie the zebrafish maternal-to-zygotic transition. *Genome Research*, 27(7):1184–1194, July 2017.

[176] Vasiliy O. Sysoev, Bernd Fischer, Christian K. Frese, Ishaan Gupta, Jeroen Krijgsveld, Matthias W. Hentze, Alfredo Castello, and Anne Ephrussi. Global changes of the RNA-bound proteome during the maternal-to-zygotic transition in Drosophila. *Nature Communications*, 7(1):12128, July 2016. Number: 1 Publisher: Nature Publishing Group.

[177] Ana M. Matia-González, Emma E. Laing, and André P. Gerber. Conserved mRNA-binding proteomes in eukaryotic organisms. *Nature Structural & Molecular Biology*, 22(12):1027–1033, December 2015. Number: 12 Publisher: Nature Publishing Group.

[178] Marlene Reichel, Yalin Liao, Mandy Rettel, Chikako Ragan, Maurits Evers, Anne-Marie Alleaume, Rastislav Horos, Matthias W. Hentze, Thomas Preiss, and Anthony A. Millar. In Planta Determination of the mRNA-Binding Proteome of Arabidopsis Etiolated Seedlings. *The Plant Cell*, 28(10):2435–2452, October 2016.

[179] Sarah F. Mitchell, Saumya Jain, Meipei She, and Roy Parker. Global Analysis of Yeast mRNPs. *Nature structural & molecular biology*, 20(1):127–133, January 2013.

[180] Devki Nandan, Sneha A. Thomas, Anne Nguyen, Kyung-Mee Moon, Leonard J. Foster, and Neil E. Reiner. Comprehensive Identification of mRNA-Binding Proteins of Leishmania donovani by Interactome Capture. *PLOS ONE*, 12(1):e0170068, January 2017. Publisher: Public Library of Science.

[181] Xichen Bao, Xiangpeng Guo, Menghui Yin, Muqddas Tariq, Yiwei Lai, Shahzina Kanwal, Jiajian Zhou, Na Li, Yuan Lv, Carlos Pulido-Quetglas, Xiwei Wang, Lu Ji, Muhammad J. Khan, Xihua Zhu, Zhiwei Luo, Changwei Shao, Do-Hwan Lim, Xiao Liu, Nan Li, Wei Wang, Minghui He, Yu-Lin Liu, Carl Ward, Tong Wang, Gong Zhang, Dongye Wang, Jianhua Yang, Yiwen Chen, Chaolin Zhang, Ralf Jauch, Yun-Gui Yang, Yangming Wang, Baoming Qin, Minna-Liisa Anko, Andrew P. Hutchins, Hao Sun, Huating Wang, Xiang-Dong Fu, Biliang Zhang, and Miguel A. Esteban. Capturing the interactome of newly transcribed RNA. *Nature Methods*, 15(3):213–220, March 2018. Publisher: Nature Publishing Group.

[182] Rongbing Huang, Mengting Han, Liying Meng, and Xing Chen. Transcriptome-wide discovery of coding and noncoding RNA-binding proteins. *Proceedings of the National Academy of Sciences*, 115(17):E3879–E3887, April 2018. Publisher: Proceedings of the National Academy of Sciences.

[183] Claudio Asencio, Aindrila Chatterjee, and Matthias W. Hentze. Silica-based solid-phase extraction of cross-linked nucleic acid-bound proteins. *Life Science Alliance*, 1(3):e201800088, June 2018.

[184] Vadim Shchepachev, Stefan Bresson, Christos Spanos, Elisabeth Petfalski, Lutz Fischer, Juri Rappsilber, and David Tollervey. Defining the RNA interactome by total RNA-associated protein purification. *Molecular Systems Biology*, 15(4):e8689, April 2019.

[185] Rayner M. L. Queiroz, Tom Smith, Eneko Villanueva, Maria Marti-Solano, Mie Monti, Mariavittoria Pizzinga, Dan-Mircea Mirea, Manasa Ramakrishna, Robert F. Harvey, Veronica Dezi, Gavin H. Thomas, Anne E. Willis, and Kathryn S. Lilley. Comprehensive identification of RNA-protein interactions in any organism using orthogonal organic phase separation (OOPS). *Nature Biotechnology*, 37(2):169–178, February 2019.

[186] Jakob Trendel, Thomas Schwarzl, Rastislav Horos, Ananth Prakash, Alex Bateman, Matthias W. Hentze, and Jeroen Krijgsveld. The Human RNA-Binding Proteome and Its Dynamics during Translational Arrest. *Cell*, 176(1):391–403.e19, January 2019.

[187] Erika C. Urdaneta, Carlos H. Vieira-Vieira, Timon Hick, Hans-Herrmann Wessels, Davide Figini, Rebecca Moschall, Jan Medenbach, Uwe Ohler, Sander Granneman, Matthias Selbach, and Benedikt M. Beckmann. Purification of cross-linked RNA-protein complexes by phenol-toluol extraction. *Nature Communications*, 10(1):990, March 2019. Publisher: Nature Publishing Group.

[188] Jong Woo Bae, S. Chul Kwon, Yongwoo Na, V. Narry Kim, and Jong-Seo Kim. Chemical RNA digestion enables robust RNA-binding site mapping at single amino acid resolution. *Nature Structural & Molecular Biology*, 27(7):678–682, July 2020. Publisher: Nature Publishing Group.

[189] Anke Liepelt, Isabel S. Naarmann-de Vries, Nadine Simons, Katrin Eichelbaum, Sophia Föhr, Stuart K. Archer, Alfredo Castello, Björn Usadel, Jeroen Krijgsveld, Thomas Preiss, Gernot Marx, Matthias W. Hentze, Dirk H. Ostareck, and Antje Ostareck-Lederer. Identification of RNA-binding Proteins in Macrophages by Interactome Capture. *Molecular & Cellular Proteomics : MCP*, 15(8):2699–2714, August 2016.

[190] Wael Kamel, Marko Noerenberg, Berati Cerikan, Honglin Chen, Aino I. Järvelin, Mohamed Kammoun, Jeffrey Y. Lee, Ni Shuai, Manuel Garcia-Moreno, Anna Andrejeva, Michael J. Deery, Natasha Johnson, Christopher J. Neufeldt, Mirko Cortese, Michael L. Knight, Kathryn S. Lilley, Javier Martinez, Ilan Davis, Ralf Bartenschlager, Shabaz Mohammed, and Alfredo Castello. Global analysis of protein-RNA interactions in SARS-CoV-2-infected cells reveals key regulators of infection. *Molecular Cell*, 81(13):2851–2867.e7, July 2021. Publisher: Elsevier.

[191] Erika Girardi, Sebastien Pfeffer, Thomas F. Baumert, and Karim Majzoub. Roadblocks and fast tracks: How RNA binding proteins affect the viral RNA journey in the cell. *Seminars in Cell & Developmental Biology*, 111:86–100, March 2021.

[192] Louisa Iselin, Natasha Palmalux, Wael Kamel, Peter Simmonds, Shabaz Mohammed, and Alfredo Castello. Uncovering viral RNA–host cell interactions on a proteome-wide scale. *Trends in Biochemical Sciences*, 47(1):23–38, January 2022.

[193] Sandy D. Der, Aimin Zhou, Bryan R. G. Williams, and Robert H. Silverman. Identification of genes differentially regulated by interferon α, β, or γ using oligonucleotide arrays. *Proceedings of the National Academy of Sciences*, 95(26):15623–15628, December 1998. Publisher: Proceedings of the National Academy of Sciences.

[194] Michael J. de Veer, Michelle Holko, Mathias Frevel, Eldon Walker, Sandy Der, Jayashree M. Paranjape, Robert H. Silverman, and Bryan R. G. Williams. Functional classification of interferon-stimulated genes identified using microarrays. *Journal of Leukocyte Biology*, 69(6):912–920, 2001. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1189/jlb.69.6.912.

[195] Stefan F. Wieland, Raquel G. Vega, Rolf Müller, Claire F. Evans, Brian Hilbush, Luca G. Guidotti, J. Gregor Sutcliffe, Peter G. Schultz, and Francis V. Chisari. Searching for Interferon-Induced Genes That Inhibit Hepatitis B Virus Replication in Transgenic Mouse Hepatocytes. *Journal of Virology*, 77(2):1227–1236, January 2003. Publisher: American Society for Microbiology.

[196] Robert E. Lanford, Bernadette Guerra, Helen Lee, Deborah Chavez, Kathleen M. Brasky, and Catherine B. Bigger. Genomic response to interferon-α in chimpanzees: Implications of rapid downregulation for hepatitis C kinetics. *Hepatology*, 43(5):961–972, 2006. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hep.21167.

[197] Irina Rusinova, Sam Forster, Simon Yu, Anitha Kannan, Marion Masse, Helen Cumming, Ross Chapman, and Paul J. Hertzog. Interferome v2.0: an updated database of annotated interferon-regulated genes. *Nucleic Acids Research*, 41(Database issue):D1040–1046, January 2013.

[198] Andrew E. Shaw, Joseph Hughes, Quan Gu, Abdelkader Behdenna, Joshua B. Singer, Tristan Dennis, Richard J. Orton, Mariana Varela, Robert J. Gifford, Sam J. Wilson, and Massimo Palmarini. Fundamental properties of the mammalian innate immune system revealed by multispecies comparison of type I interferon responses. *PLOS Biology*, 15(12):e2004086, December 2017. Publisher: Public Library of Science.

[199] Kenneth Hodge, Jiradej Makjaroen, Jonathan Robinson, Sakda Khoomrung, and Trairak Pisitkun. Deep Proteomic Deconvolution of Interferons and HBV Transfection Effects on a Hepatoblastoma Cell Line. *ACS Omega*, 5(27):16796–16810, June 2020.

[200] Maikel L. Colli, Mireia Ramos-Rodríguez, Ernesto S. Nakayasu, Maria I. Alvelos, Miguel Lopes, Jessica L. E. Hill, Jean-Valery Turatsinze, Alexandra Coomans de Brachène, Mark A. Russell, Helena Raurell-Vila, Angela Castela, Jonàs Juan-Mateu, Bobbie-Jo M. Webb-Robertson, Lars Krogvold, Knut Dahl-Jorgensen, Lorella Marselli, Piero Marchetti, Sarah J. Richardson, Noel G. Morgan, Thomas O. Metz, Lorenzo Pasquali, and Décio L. Eizirik. An integrated multi-omics approach identifies the landscape of interferon-α-mediated responses of human pancreatic beta cells. *Nature Communications*, 11(1):2584, May 2020. Number: 1 Publisher: Nature Publishing Group.

[201] Wei Yan, Hookeun Lee, Eugene C. Yi, David Reiss, Paul Shannon, Bartlomiej K. Kwieciszewski, Carlos Coito, Xiao-jun Li, Andrew Keller, Jimmy Eng, Timothy Galitski, David R. Goodlett, Ruedi Aebersold, and Michael G. Katze. System-based proteomic analysis of the interferon response in human liver cells. *Genome Biology*, 5(8):R54, July 2004.

[202] Nicola Ternette, Cynthia Wright, Holger B. Kramer, Mikael Altun, and Benedikt M. Kessler. Label-free quantitative proteomics reveals regulation of interferon-induced protein with tetratricopeptide repeats 3 (IFIT3) and 5'-3'-exoribonuclease 2 (XRN2) during respiratory syncytial virus infection. *Virology Journal*, 8(1):442, September 2011.

[203] Dominik A. Megger, Jos Philipp, Vu Thuy Khanh Le-Trilling, Barbara Sitek, and Mirko Trilling. Deciphering of the Human Interferon-Regulated Proteome by Mass Spectrometry-Based Quantitative Analysis Reveals Extent and Dynamics of Protein Induction and Repression. *Frontiers in Immunology*, 8, 2017.

[204] Jiradej Makjaroen, Poorichaya Somparn, Kenneth Hodge, Witthaya Poomipak, Nattiya Hirankarn, and Trairak Pisitkun. Comprehensive Proteomics Identification of IFN-λ3-regulated Antiviral Proteins in HBV-transfected Cells. *Molecular & cellular proteomics: MCP*, 17(11):2197–2215, November 2018.

[205] Jonas Schuhenn, Toni Luise Meister, Daniel Todt, Thilo Bracht, Karin Schork, Jean-Noel Billaud, Carina Elsner, Natalie Heinen, Zehra Karakoese, Sibylle Haid, Sriram Kumar, Linda Brunotte, Martin Eisenacher, Yunyun Di, Jocelyne Lew, Darryl Falzarano, Jieliang Chen, Zhenghong Yuan, Thomas Pietschmann, Bettina Wiegmann, Hendrik Uebner, Christian Taube, Vu Thuy Khanh Le-Trilling, Mirko Trilling, Adalbert Krawczyk, Stephan

Ludwig, Barbara Sitek, Eike Steinmann, Ulf Dittmer, Kerry J. Lavender, Kathrin Sutter, and Stephanie Pfaender. Differential interferon-α subtype induced immune signatures are associated with suppression of SARS-CoV-2 infection. *Proceedings of the National Academy of Sciences of the United States of America*, 119(8):e2111600119, February 2022.

[206] Mingzhi Zhao, Liang Li, Linhui Zhai, Qi Yue, Hongyu Liu, Suping Ren, Xingwei Jiang, Fenghua Gao, Shanshan Bai, Honghao Li, Ying Zhang, Hongwei Xu, Liying Zhang, Pinghuang Liu, Minjia Tan, and Qun Yu. Comparative Transcriptomic and Proteomic Analyses Prove that IFN-λ1 is a More Potent Inducer of ISGs than IFN-α against Porcine Epidemic Diarrhea Virus in Porcine Intestinal Epithelial Cells. *Journal of Proteome Research*, 19(9):3697–3707, September 2020. Publisher: American Chemical Society.

[207] Yugen Zhang, Crystal W. Burke, Kate D. Ryman, and William B. Klimstra. Identification and Characterization of Interferon-Induced Proteins That Inhibit Alphavirus Replication. *Journal of Virology*, 81(20):11246–11255, October 2007. Publisher: American Society for Microbiology.

[208] Y. Itsui, N. Sakamoto, M. Kurosaki, N. Kanazawa, Y. Tanabe, T. Koyama, Y. Takeda, M. Nakagawa, S. Kakinuma, Y. Sekine, S. Maekawa, N. Enomoto, and M. Watanabe. Expressional screening of interferon-stimulated genes for antiviral activity against hepatitis C virus replication. *Journal of Viral Hepatitis*, 13(10):690–700, 2006. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2893.2006.00732.x.

[209] Hangfei Qi, Virginia Chu, Nicholas C. Wu, Zugen Chen, Shawna Truong, Gurpreet Brar, Sheng-Yao Su, Yushen Du, Vaithilingaraja Arumugaswami, C. Anders Olson, Shu-Hua Chen, Chung-Yen Lin, Ting-Ting Wu, and Ren Sun. Systematic identification of anti-interferon function on hepatitis C virus genome reveals p7 as an immune evasion protein. *Proceedings of the National Academy of Sciences*, 114(8):2018–2023, February 2017. Publisher: Proceedings of the National Academy of Sciences.

[210] Sam J. Wilson, John W. Schoggins, Trinity Zang, Sebla B. Kutluay, Nolwenn Jouvenet, Mudathir A. Alim, Julia Bitzegeio, Charles M. Rice, and Paul D. Bieniasz. Inhibition of HIV-1 Particle Assembly by 2',3'-Cyclic-Nucleotide 3'-Phosphodiesterase. *Cell Host & Microbe*, 12(4):585–597, October 2012.

[211] Melissa Kane, Trinity M. Zang, Suzannah J. Rihn, Fengwen Zhang, Tonya Kueck, Mudathir Alim, John Schoggins, Charles M. Rice, Sam J. Wilson, and Paul D. Bieniasz. Identification of Interferon-Stimulated Genes with Antiretroviral Activity. *Cell Host & Microbe*, 20(3):392–405, September 2016. Publisher: Elsevier.

[212] Junjie Feng, Arthur Wickenhagen, Matthew L. Turnbull, Veronica V. Rezelj, Felix Kreher, Natasha L. Tilston-Lunel, Gillian S. Slack, Benjamin Brennan, Elina Koudriakova, Andrew E. Shaw, Suzannah J. Rihn, Charles M. Rice, Paul D. Bieniasz, Richard M. Elliott, Xiaohong Shi, and Sam J. Wilson. Interferon-Stimulated Gene (ISG)-Expression Screening Reveals the Specific Antibunyaviral Activity of ISG20. *Journal of Virology*, 92(13):e02140–17, June 2018. Publisher: American Society for Microbiology.

[213] Philippe Metz, Eva Dazert, Alessia Ruggieri, Johanna Mazur, Lars Kaderali, Artur Kaul, Ulf Zeuge, Marc P. Windisch, Martin Trippler, Volker Lohmann, Marco Binder, Michael Frese, and Ralf Bartenschlager. Identification of type I and type II interferon-induced effectors controlling hepatitis C virus replication. *Hepatology*, 56(6):2082–2093, 2012. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hep.25908.

[214] Jianqing Li, Steve C. Ding, Hyelim Cho, Brian C. Chung, Michael Gale, Sumit K. Chanda, and Michael S. Diamond. A Short Hairpin RNA Screen of Interferon-Stimulated Genes Identifies a Novel Negative Regulator of the Cellular Antiviral Response. *mBio*, 4(3):e00385–13, June 2013. Publisher: American Society for Microbiology.

[215] Gayatri Subramanian, Teodora Kuzmanovic, Ying Zhang, Cara Beate Peter, Manoj Veleeparambil, Ritu Chakravarti, Ganes C. Sen, and Saurabh Chattopadhyay. A new mechanism of interferon's antiviral action: Induction of autophagy, essential for paramyxovirus replication, is inhibited by the interferon stimulated gene, TDRD7. *PLoS pathogens*, 14(1):e1006877, January 2018.

[216] Dahlene N. Fusco, Cynthia Brisac, Sinu P. John, Yi-Wen Huang, Christopher R. Chin, Tiao Xie, Hong Zhao, Nikolaus Jilg, Leiliang Zhang, Stephane Chevaliez, Daniel Wambua, Wenyu Lin, Lee Peng, Raymond T. Chung, and Abraham L. Brass. A Genetic Screen Identifies Interferon-α Effector Genes Required to Suppress Hepatitis C Virus Replication. *Gastroenterology*, 144(7):1438–1449.e9, June 2013. Publisher: Elsevier.

[217] Hong Zhao, Wenyu Lin, Kattareeya Kumthip, Du Cheng, Dahlene N. Fusco, Oliver Hofmann, Nikolaus Jilg, Andrew W. Tai, Kaku Goto, Leiliang Zhang, Winston Hide, Jae Young Jang, Lee F. Peng, and Raymond T. Chung. A functional genomic screen reveals novel host genes that mediate interferon-alpha's effects against hepatitis C virus. *Journal of Hepatology*, 56(2):326–333, February 2012. Publisher: Elsevier.

[218] Molly OhAinle, Louisa Helms, Jolien Vermeire, Ferdinand Roesch, Daryl Humes, Ryan Basom, Jeffrey J Delrow, Julie Overbaugh, and Michael Emerman. A virus-packageable CRISPR screen identifies host factors mediating interferon inhibition of HIV. *eLife*, 7:e39823, December 2018. Publisher: eLife Sciences Publications, Ltd.

[219] Su-Yang Liu, David Jesse Sanchez, Roghiyh Aliyari, Sun Lu, and Genhong Cheng. Systematic identification of type I and type II interferon-induced antiviral factors. *Proceedings of the National Academy of Sciences*, 109(11):4239–4244, March 2012. Publisher: Proceedings of the National Academy of Sciences.

[220] John W. Schoggins, Sam J. Wilson, Maryline Panis, Mary Y. Murphy, Christopher T. Jones, Paul Bieniasz, and Charles M. Rice. A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature*, 472(7344):481–485, April 2011. Number: 7344 Publisher: Nature Publishing Group.

[221] John W. Schoggins, Donna A. MacDuff, Naoko Imanaka, Maria D. Gainey, Bimmi Shrestha, Jennifer L. Eitson, Katrina B. Mar, R. Blake Richardson, Alexander V. Ratushny, Vladimir Litvak, Rea Dabelic, Balaji Manicassamy, John D. Aitchison, Alan Aderem, Richard M. Elliott, Adolfo García-Sastre, Vincent Racaniello, Eric J. Snijder, Wayne M. Yokoyama, Michael S. Diamond, Herbert W. Virgin, and Charles M. Rice. Pan-viral specificity of IFN-induced genes reveals new roles for cGAS in innate immunity. *Nature*, 505(7485):691–695, January 2014. Number: 7485 Publisher: Nature Publishing Group.

[222] R. Blake Richardson, Maikke B. Ohlson, Jennifer L. Eitson, Ashwani Kumar, Matthew B. McDougal, Ian N. Boys, Katrina B. Mar, Pamela C. De La Cruz-Rivera, Connor Douglas, Genevieve Konopka, Chao Xing, and John W. Schoggins. A CRISPR screen identifies IFI6 as an ER-resident interferon effector that blocks flavivirus replication. *Nature Microbiology*, 3(11):1214–1223, November 2018. Number: 11 Publisher: Nature Publishing Group.

[223] Jenny J. Seo, Soo-Jin Jung, Jihye Yang, Da-Eun Choi, and V. Narry Kim. Functional viromic screens uncover regulatory RNA elements. *Cell*, 186(15):3291–3306.e21, July 2023.

[224] Joel I. Perez-Perri, Marko Noerenberg, Wael Kamel, Caroline E. Lenz, Shabaz Mohammed, Matthias W. Hentze, and Alfredo Castello. Global analysis of RNA-binding protein dynamics by comparative and enhanced RNA interactome capture. *Nature Protocols*, 16(1):27–60, January 2021. Number: 1 Publisher: Nature Publishing Group.

[225] D V Goeddel, H M Shepard, E Yelverton, D Leung, R Crea, A Sloma, and S Pestka. Synthesis of human fibroblast interferon by E. coli. *Nucleic Acids Research*, 8(18):4057–4074, September 1980.

[226] Michel Streuli, Shigekazu Nagata, and Charles Weissmann. At Least Three Human Type α Interferons: Structure of α2. *Science*, 209(4463):1343–1347, September 1980. Publisher: American Association for the Advancement of Science.

[227] Franciane Paul, Sandra Pellegrini, and Gilles Uzé. IFNA2: The prototypic human alpha interferon. *Gene*, 567(2):132–137, August 2015.

[228] Jürgen Cox and Matthias Mann. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, 26(12):1367–1372, December 2008.

[229] G. K. Smyth. limma: Linear Models for Microarray Data. In Robert Gentleman, Vincent J. Carey, Wolfgang Huber, Rafael A. Irizarry, and Sandrine Dudoit, editors, *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, Statistics for Biology and Health, pages 397–420. Springer, New York, NY, 2005.

[230] Michaela U. Gack, Young C. Shin, Chul-Hyun Joo, Tomohiko Urano, Chengyu Liang, Lijun Sun, Osamu Takeuchi, Shizuo Akira, Zhijian Chen, Satoshi Inoue, and Jae U. Jung. TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. *Nature*, 446(7138):916–920, April 2007. Number: 7138 Publisher: Nature Publishing Group.

[231] Jie Wang, Baoming Liu, Nan Wang, Young-Min Lee, Chunming Liu, and Kui Li. TRIM56 Is a Virus- and Interferon-Inducible E3 Ubiquitin Ligase That Restricts Pestivirus

Infection. *Journal of Virology*, 85(8):3733–3745, April 2011. Publisher: American Society for Microbiology.

[232] Georg Kustatscher, Piotr Grabowski, Tina A. Schrader, Josiah B. Passmore, Michael Schrader, and Juri Rappsilber. Co-regulation map of the human proteome enables identification of protein functions. *Nature Biotechnology*, 37(11):1361–1371, November 2019. Number: 11 Publisher: Nature Publishing Group.

[233] Lucas J. M. Bruurs, Micha Müller, Jelle G. Schipper, Huib H. Rabouw, Sanne Boersma, Frank J. M. van Kuppeveld, and Marvin E. Tanenbaum. Antiviral responses are shaped by heterogeneity in viral replication dynamics. *Nature Microbiology*, 8(11):2115–2129, November 2023. Publisher: Nature Publishing Group.

[234] Stephen O'Gorman, Daniel T. Fox, and Geoffrey M. Wahl. Recombinase-Mediated Gene Activation and Site-Specific Integration in Mammalian Cells. *Science*, 251(4999):1351–1355, March 1991. Publisher: American Association for the Advancement of Science.

[235] Luis Carrasco, Miguel Angel Sanz, and Esther González-Almela. The Regulation of Translation in Alphavirus-Infected Cells. *Viruses*, 10(2):70, February 2018. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute.

[236] Joyce Jing Yi Wong, Yuh Fen Pung, Newman Siu-Kwan Sze, and Keh-Chuang Chin. HERC5 is an IFN-induced HECT-type E3 protein ligase that mediates type I IFN-induced ISGylation of protein targets. *Proceedings of the National Academy of Sciences*, 103(28):10735–10740, July 2006. Publisher: Proceedings of the National Academy of Sciences.

[237] Coralie Guy and Andrew G Bowie. Recent insights into innate immune nucleic acid sensing during viral infection. *Current Opinion in Immunology*, 78:102250, October 2022.

[238] Smadar Cohen-Chalamish, Anat Hasson, Dahlia Weinberg, Lise Sarah Namer, Yona Banai, Farhat Osman, and Raymond Kaempfer. Dynamic refolding of IFN-gamma mRNA enables it to function as PKR activator and translation template. *Nature Chemical Biology*, 5(12):896–903, December 2009.

[239] Yoosik Kim, Joha Park, Sujin Kim, MinA Kim, Myeong-Gyun Kang, Chulhwan Kwak, Minjeong Kang, Baekgyu Kim, Hyun-Woo Rhee, and V. Narry Kim. PKR Senses

Nuclear and Mitochondrial Signals by Interacting with Endogenous Double-Stranded RNAs. *Molecular Cell*, 71(6):1051–1063.e6, September 2018. Publisher: Elsevier.

[240] Reazur Rahman, Weijin Xu, Hua Jin, and Michael Rosbash. Identification of RNA-binding protein targets with HyperTRIBE. *Nature Protocols*, 13(8):1829–1849, August 2018. Publisher: Nature Publishing Group.

[241] Kristopher W. Brannan, Isaac A. Chaim, Ryan J. Marina, Brian A. Yee, Eric R. Kofman, Daniel A. Lorenz, Pratibha Jagannatha, Kevin D. Dong, Assael A. Madrigal, Jason G. Underwood, and Gene W. Yeo. Robust single-cell discovery of RNA targets of RNA-binding proteins and ribosomes. *Nature Methods*, 18(5):507–519, May 2021. Publisher: Nature Publishing Group.

[242] Alisha Chitrakar, Kristina Solorio-Kirpichyan, Eliza Prangley, Sneha Rath, Jin Du, and Alexei Korennykh. Introns encode dsRNAs undetected by RIG-I/MDA5/interferons and sensed via RNase L. *Proceedings of the National Academy of Sciences*, 118(46):e2102134118, November 2021. Publisher: Proceedings of the National Academy of Sciences.

[243] Mikayla R. Thompson, Shruti Sharma, Maninjay Atianand, Søren B. Jensen, Susan Carpenter, David M. Knipe, Katherine A. Fitzgerald, and Evelyn A. Kurt-Jones. Interferon γ-inducible Protein (IFI) 16 Transcriptionally Regulates Type I Interferons and Other Interferon-stimulated Genes and Controls the Interferon Response to both DNA and RNA Viruses*. *Journal of Biological Chemistry*, 289(34):23568–23581, August 2014.

[244] Graham D. Williams, Nandan S. Gokhale, Daltry L. Snider, and Stacy M. Horner. The mRNA Cap 2'-O-Methyltransferase CMTR1 Regulates the Expression of Certain Interferon-Stimulated Genes. *mSphere*, 5(3):e00202–20, May 2020.

[245] Marianne Bonvin, François Achermann, Isabell Greeve, Deborah Stroka, Adrian Keogh, Daniel Inderbitzin, Daniel Candinas, Peter Sommer, Simon Wain-Hobson, Jean-Pierre Vartanian, and Jobst Greeve. Interferon-inducible expression of APOBEC3 editing enzymes in human hepatocytes and inhibition of hepatitis B virus replication. *Hepatology (Baltimore, Md.)*, 43(6):1364–1374, June 2006.

[246] Gang Peng, Ke Jian Lei, Wenwen Jin, Teresa Greenwell-Wild, and Sharon M. Wahl. Induction of APOBEC3 family proteins, a defensive maneuver underlying interferon-

induced anti–HIV-1 activity. *Journal of Experimental Medicine*, 203(1):41–46, January 2006.

[247] Riccardo Pecori, Salvatore Di Giorgio, J. Paulo Lorenzo, and F. Nina Papavasiliou. Functions and consequences of AID/APOBEC-mediated DNA and RNA deamination. *Nature Reviews Genetics*, 23(8):505–518, August 2022. Publisher: Nature Publishing Group.

[248] Harold C. Smith. RNA binding to APOBEC deaminases; Not simply a substrate for C to U editing. *RNA Biology*, 14(9):1153–1165, November 2016.

[249] Sarah Gallois-Montbrun, Beatrice Kramer, Chad M. Swanson, Helen Byers, Steven Lynham, Malcolm Ward, and Michael H. Malim. Antiviral Protein APOBEC3G Localizes to Ribonucleoprotein Complexes Found in P Bodies and Stress Granules. *Journal of Virology*, 81(5):2165–2178, March 2007. Publisher: American Society for Microbiology.

[250] Bogdan Polevoda, William M. McDougall, Bradley N. Tun, Michael Cheung, Jason D. Salter, Alan E. Friedman, and Harold C. Smith. RNA binding to APOBEC3G induces the disassembly of functional deaminase complexes by displacing single-stranded DNA substrates. *Nucleic Acids Research*, 43(19):9434–9445, October 2015.

[251] Vanessa B. Soros, Wes Yonemoto, and Warner C. Greene. Newly Synthesized APOBEC3G Is Incorporated into HIV Virions, Inhibited by HIV RNA, and Subsequently Activated by RNase H. *PLOS Pathogens*, 3(2):e15, February 2007. Publisher: Public Library of Science.

[252] Philipp Hubel, Christian Urban, Valter Bergant, William M. Schneider, Barbara Knauer, Alexey Stukalov, Pietro Scaturro, Angelika Mann, Linda Brunotte, Heinrich H. Hoffmann, John W. Schoggins, Martin Schwemmle, Matthias Mann, Charles M. Rice, and Andreas Pichlmair. A protein-interaction network of interferon-stimulated genes extends the innate immune system landscape. *Nature Immunology*, 20(4):493–502, April 2019. Number: 4 Publisher: Nature Publishing Group.

[253] Jun Xiao, Weiyun Li, Xin Zheng, Linlin Qi, Hui Wang, Chi Zhang, Xiaopeng Wan, Yuxiao Zheng, Ruiyue Zhong, Xin Zhou, Yao Lu, Zhiqi Li, Ying Qiu, Chang Liu, Fang Zhang, Yanbo Zhang, Xiaoyan Xu, Zhongzhou Yang, Hualan Chen, Qiwei Zhai, Bin Wei, and Hongyan Wang. Targeting 7-Dehydrocholesterol Reductase Integrates Cholesterol

Metabolism and IRF3 Activation to Eliminate Infection. *Immunity*, 52(1):109–122.e6, January 2020.

[254] Junji Xing, Ao Zhang, Yong Du, Mingli Fang, Laurie J. Minze, Yong-Jun Liu, Xian Chang Li, and Zhiqiang Zhang. Identification of poly(ADP-ribose) polymerase 9 (PARP9) as a noncanonical sensor for RNA virus in dendritic cells. *Nature Communications*, 12(1):2681, May 2021.

[255] Ioannis Sanidas, Christos Polytarchou, Maria Hatziapostolou, Scott A. Ezell, Filippos Kottakis, Lan Hu, Ailan Guo, Jianxin Xie, Michael J. Comb, Dimitrios Iliopoulos, and Philip N. Tsichlis. Phosphoproteomics Screen Reveals Akt Isoform-Specific Signals Linking RNA Processing to Lung Cancer. *Molecular Cell*, 53(4):577–590, February 2014. Publisher: Elsevier.

[256] Kuo-Chieh Liao and Mariano A. Garcia-Blanco. Role of Alternative Splicing in Regulating Host Response to Viral Infection. *Cells*, 10(7):1720, July 2021.

[257] Richeng Mao, Hui Nie, Dawei Cai, Jiming Zhang, Hongyan Liu, Ran Yan, Andrea Cuconati, Timothy M. Block, Ju-Tao Guo, and Haitao Guo. Inhibition of Hepatitis B Virus Replication by the Host Zinc Finger Antiviral Protein. *PLOS Pathogens*, 9(7):e1003494, July 2013. Publisher: Public Library of Science.

[258] Joao Marques, Jangawar Anwar, Signe Eskildsen-Larsen, Dominique Rebouillat, Soren R. Paludan, Ganes Sen, Bryan R. G. Williams, and Rune Hartmann. The p59 oligoadenylate synthetase-like protein possesses antiviral activity that requires the C-terminal ubiquitin-like domain. *The Journal of General Virology*, 89(Pt 11):2767–2772, November 2008.

[259] Rong Guo, Yong Li, Jinying Ning, Dan Sun, Lianjun Lin, and Xinmin Liu. HnRNP A1/A2 and SF2/ASF regulate alternative splicing of interferon regulatory factor-3 and affect immunomodulatory functions in human non-small cell lung cancer cells. *PloS One*, 8(4):e62729, 2013.

[260] Chunaram Choudhary, Brian T. Weinert, Yuya Nishida, Eric Verdin, and Matthias Mann. The growing landscape of lysine acetylation links metabolism and cell signalling. *Nature Reviews Molecular Cell Biology*, 15(8):536–550, August 2014. Publisher: Nature Publishing Group.

[261] Jian Zhang, Abdullah M. Ali, Yen K. Lieu, Zhaoqi Liu, Jianchao Gao, Raul Rabadan, Azra Raza, Siddhartha Mukherjee, and James L. Manley. Disease-Causing Mutations in SF3B1 Alter Splicing by Disrupting Interaction with SUGP1. *Molecular Cell*, 76(1):82–95.e7, October 2019. Publisher: Elsevier.

[262] Zhaoqi Liu, Jian Zhang, Yiwei Sun, Tomin E. Perea-Chamblee, James L. Manley, and Raul Rabadan. Pan-cancer analysis identifies mutations in SUGP1 that recapitulate mutant SF3B1 splicing dysregulation. *Proceedings of the National Academy of Sciences*, 117(19):10305–10312, May 2020. Publisher: Proceedings of the National Academy of Sciences.

[263] Mee J. Kim, Chi-Yi Yu, Elizabeth Theusch, Devesh Naidoo, Kristen Stevens, Yu-Lin Kuang, Erin Schuetz, Amarjit S. Chaudhry, and Marisa W. Medina. SUGP1 is a novel regulator of cholesterol metabolism. *Human Molecular Genetics*, 25(14):3106–3116, July 2016.

[264] Efraín E. Rivera-Serrano, Ethan J. Fritch, Elizabeth H. Scholl, and Barbara Sherry. A Cytoplasmic RNA Virus Alters the Function of the Cell Splicing Protein SRSF2. *Journal of Virology*, 91(7):10.1128/jvi.02488–16, March 2017. Publisher: American Society for Microbiology.

[265] Catherine Ptaschinski, Sumanta Mukherjee, Martin L. Moore, Mareike Albert, Kristian Helin, Steven L. Kunkel, and Nicholas W. Lukacs. RSV-Induced H3K4 Demethylase KDM5B Leads to Regulation of Dendritic Cell-Derived Innate Cytokines and Exacerbates Pathogenesis In Vivo. *PLOS Pathogens*, 11(6):e1004978, June 2015. Publisher: Public Library of Science.

[266] Zhang Xu, Jung-Hyun Choi, David L. Dai, Jun Luo, Reese Jalal Ladak, Qian Li, Yimeng Wang, Christine Zhang, Shane Wiebe, Alex C. H. Liu, Xiaozhuo Ran, Jiaqi Yang, Parisa Naeli, Aitor Garzia, Lele Zhou, Niaz Mahmood, Qiyun Deng, Mohamed Elaish, Rongtuan Lin, Lara K. Mahal, Tom C. Hobman, Jerry Pelletier, Tommy Alain, Silvia M. Vidal, Thomas Duchaine, Mohammad T. Mazhab-Jafari, Xiaojuan Mao, Seyed Mehdi Jafarnejad, and Nahum Sonenberg. SARS-CoV-2 impairs interferon production via NSP2-induced repression of mRNA translation. *Proceedings of the National Academy of Sciences*, 119(32):e2204539119, August 2022. Publisher: Proceedings of the National Academy of Sciences.

[267] Gorben P. Pijlman, Anneke Funk, Natasha Kondratieva, Jason Leung, Shessy Torres, Lieke van der Aa, Wen Jun Liu, Ann C. Palmenberg, Pei-Yong Shi, Roy A. Hall, and Alexander A. Khromykh. A Highly Structured, Nuclease-Resistant, Noncoding RNA Produced by Flaviviruses Is Required for Pathogenicity. *Cell Host & Microbe*, 4(6):579–591, December 2008. Publisher: Elsevier.

[268] Andrii Slonchak and Alexander A. Khromykh. Subgenomic flaviviral RNAs: What do we know after the first decade of research. *Antiviral Research*, 159:13–25, November 2018.

[269] Gayathri Manokaran, Esteban Finol, Chunling Wang, Jayantha Gunaratne, Justin Bahl, Eugenia Z. Ong, Hwee Cheng Tan, October M. Sessions, Alex M. Ward, Duane J. Gubler, Eva Harris, Mariano A. Garcia-Blanco, and Eng Eong Ooi. Dengue subgenomic RNA binds TRIM25 to inhibit interferon expression for epidemiological fitness. *Science (New York, N.Y.)*, 350(6257):217–221, October 2015.

[270] Yen-Chin Liu, Bobo Wing-Yee Mok, Pui Wang, Rei-Lin Kuo, Honglin Chen, and Shin-Ru Shih. Cellular 5'-3' mRNA Exoribonuclease XRN1 Inhibits Interferon Beta Activation and Facilitates Influenza A Virus Replication. *mBio*, 12(4):10.1128/mbio.00945–21, July 2021. Publisher: American Society for Microbiology.

[271] Nicoletta Scheller, Leonardo Bruno Mina, Rui Pedro Galão, Ashwin Chari, Mireia Giménez-Barcons, Amine Noueiry, Utz Fischer, Andreas Meyerhans, and Juana Díez. Translation and replication of hepatitis C virus genomic RNA depends on ancient cellular proteins that control mRNA fates. *Proceedings of the National Academy of Sciences*, 106(32):13517–13522, August 2009. Publisher: Proceedings of the National Academy of Sciences.

[272] Manuel Garcia-Moreno, Robin Truman, Honglin Chen, Louisa Iselin, Caroline E. Lenz, Jeff Lee, Kate Dicker, Marko Noerenberg, Thibault J. M. Sohier, Natasha Palmalux, Aino I. Järvelin, Wael Kamel, Vincenzo Ruscica, Emiliano P. Ricci, Ilan Davis, Shabaz Mohammed, and Alfredo Castello. Incorporation of genome-bound cellular proteins into HIV-1 particles regulates viral infection, June 2023.

[273] Gaston Bonenfant, Nina Williams, Rachel Netzband, Megan C. Schwarz, Matthew J. Evans, and Cara T. Pager. Zika Virus Subverts Stress Granules To Promote and Restrict

Viral Gene Expression. *Journal of Virology*, 93(12):10.1128/jvi.00520–19, May 2019. Publisher: American Society for Microbiology.

[274] Yasuo Ariumi, Misao Kuroki, Yukihiro Kushima, Kanae Osugi, Makoto Hijikata, Masatoshi Maki, Masanori Ikeda, and Nobuyuki Kato. Hepatitis C Virus Hijacks P-Body and Stress Granule Components around Lipid Droplets. *Journal of Virology*, 85(14):6882–6892, July 2011. Publisher: American Society for Microbiology.

[275] Chanyoung Shin, Yuma Ito, Shota Ichikawa, Makio Tokunaga, Kumiko Sakata-Sogawa, and Takashi Tanaka. MKRN2 is a novel ubiquitin E3 ligase for the p65 subunit of NF-κB and negatively regulates inflammatory responses. *Scientific Reports*, 7(1):46097, April 2017. Publisher: Nature Publishing Group.

[276] Stefano Bonazza, Hannah Coutts, Swathi Sukumar, Hannah Turkington, and David G. Courtney. Identifying cellular RNA-binding proteins during infection uncovers a role for MKRN2 in influenza mRNA trafficking, October 2023. Pages: 2023.10.30.564690 Section: New Results.

[277] Yansheng Liu, Andreas Beyer, and Ruedi Aebersold. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell*, 165(3):535–550, April 2016.

[278] Jun-Young Seo, Rakina Yaneva, Ella R. Hinson, and Peter Cresswell. Human cytomegalovirus directly induces the antiviral protein viperin to enhance infectivity. *Science (New York, N.Y.)*, 332(6033):1093–1097, May 2011.

[279] Yilong Zhou, Chenxi He, Lin Wang, and Baoxue Ge. Post-translational regulation of antiviral innate signaling. *European Journal of Immunology*, 47(9):1414–1426, 2017. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/eji.201746959.

[280] George A. Khoury, Richard C. Baliban, and Christodoulos A. Floudas. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Scientific Reports*, 1(1):90, September 2011. Publisher: Nature Publishing Group.

[281] Friedrich Marks. *Protein Phosphorylation*. John Wiley & Sons, 1996. Google-Books-ID: jbJajU1WesAC.

[282] Kirti Sharma, Rochelle C. J. D'Souza, Stefka Tyanova, Christoph Schaab, Jacek R. Wiśniewski, Jürgen Cox, and Matthias Mann. Ultradeep Human Phosphoproteome Reveals a Distinct Regulatory Nature of Tyr and Ser/Thr-Based Signaling. *Cell Reports*, 8(5):1583–1594, September 2014.

[283] Li-Rong Yu and Timothy D. Veenstra. Characterization of Phosphorylated Proteins Using Mass Spectrometry. *Current protein & peptide science*, 22(2):148–157, 2021.

[284] Alejandro Velázquez-Cruz, Blanca Baños-Jaime, Antonio Díaz-Quintana, Miguel A. De la Rosa, and Irene Díaz-Moreno. Post-translational Control of RNA-Binding Proteins and Disease-Related Dysregulation. *Frontiers in Molecular Biosciences*, 8, 2021.

[285] Whitney E England, Jingtian Wang, Siwei Chen, Pierre Baldi, Ryan A Flynn, and Robert C Spitale. An atlas of posttranslational modifications on RNA binding proteins. *Nucleic Acids Research*, 50(8):4329–4339, May 2022.

[286] Alaji Bah and Julie D. Forman-Kay. Modulation of Intrinsically Disordered Protein Function by Post-translational Modifications. *The Journal of Biological Chemistry*, 291(13):6696–6705, March 2016.

[287] Elise J. Needham, Benjamin L. Parker, Timur Burykin, David E. James, and Sean J. Humphrey. Illuminating the dark phosphoproteome. *Science Signaling*, 12(565):eaau8645, January 2019. Publisher: American Association for the Advancement of Science.

[288] Abul Arif, Peng Yao, Fulvia Terenzi, Jie Jia, Partho Sarothi Ray, and Paul L. Fox. The GAIT translational control system. *WIREs RNA*, 9(2):e1441, 2018. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wrna.1441.

[289] Prabha Sampath, Barsanjit Mazumder, Vasudevan Seshadri, and Paul L. Fox. Transcript-selective translational silencing by gamma interferon is directed by a novel structural element in the ceruloplasmin mRNA 3' untranslated region. *Molecular and Cellular Biology*, 23(5):1509–1519, March 2003.

[290] Benjamin R. tenOever, Sze-Ling Ng, Mark A. Chua, Sarah M. McWhirter, Adolfo García-Sastre, and Tom Maniatis. Multiple Functions of the IKK-Related Kinase IKKε in Interferon-Mediated Antiviral Immunity. *Science*, 315(5816):1274–1278, March 2007. Publisher: American Association for the Advancement of Science.

[291] Serge Y. Fuchs. Hope and Fear for Interferon: The Receptor-Centric Outlook on the Future of Interferon Therapy. *Journal of Interferon & Cytokine Research*, 33(4):211–225, April 2013.

[292] Fiorella A. Solari, Margherita Dell'Aica, Albert Sickmann, and René P. Zahedi. Why phosphoproteomics is still a challenge. *Molecular BioSystems*, 11(6):1487–1493, May 2015.

[293] Shun Feng, Mingliang Ye, Houjiang Zhou, Xiaogang Jiang, Xingning Jiang, Hanfa Zou, and Bolin Gong. Immobilized Zirconium Ion Affinity Chromatography for Specific Enrichment of Phosphopeptides in Phosphoproteome Analysis *. *Molecular & Cellular Proteomics*, 6(9):1656–1665, September 2007. Publisher: Elsevier.

[294] Ignacio Arribas Diez, Ireshyn Govender, Previn Naicker, Stoyan Stoychev, Justin Jordaan, and Ole N. Jensen. Zirconium(IV)-IMAC Revisited: Improved Performance and Phosphoproteome Coverage by Magnetic Microparticles for Phosphopeptide Affinity Enrichment. *Journal of Proteome Research*, 20(1):453–462, January 2021.

[295] David Ochoa, Andrew F. Jarnuczak, Cristina Viéitez, Maja Gehre, Margaret Soucheray, André Mateus, Askar A. Kleefeldt, Anthony Hill, Luz Garcia-Alonso, Frank Stein, Nevan J. Krogan, Mikhail M. Savitski, Danielle L. Swaney, Juan A. Vizcaíno, Kyung-Min Noh, and Pedro Beltrao. The functional landscape of the human phosphoproteome. *Nature Biotechnology*, 38(3):365–373, March 2020. Number: 3 Publisher: Nature Publishing Group.

[296] Wenhao Jin, Kristopher W. Brannan, Katannya Kapeli, Samuel S. Park, Hui Qing Tan, Maya L. Gosztyla, Mayuresh Mujumdar, Joshua Ahdout, Bryce Henroid, Katherine Rothamel, Joy S. Xiang, Limsoon Wong, and Gene W. Yeo. HydRA: Deep-learning models for predicting RNA-binding capacity from protein interaction association context and protein sequence. *Molecular Cell*, 83(14):2595–2611.e11, July 2023.

[297] Daniel Peter, Ramona Weber, Felix Sandmeir, Lara Wohlbold, Sigrun Helms, Praveen Bawankar, Eugene Valkov, Cátia Igreja, and Elisa Izaurralde. GIGYF1/2 proteins use auxiliary sequences to selectively bind to 4EHP and repress target mRNA expression. *Genes & Development*, 31(11):1147–1161, June 2017.

[298] Miguel B. Coelho, Jan Attig, Jernej Ule, and Christopher W.J. Smith. Matrin3: connecting gene expression with the nuclear matrix. *WIREs RNA*, 7(3):303–315, 2016. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wrna.1336.

[299] Claudia Strein, Anne-Marie Alleaume, Ulrich Rothbauer, Matthias W. Hentze, and Alfredo Castello. A versatile assay for RNA-binding proteins in living cells. *RNA*, 20(5):721–731, May 2014. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

[300] Gennaro Giordano, Ana María Sánchez-Pérez, Carmina Montoliu, Ronald Berezney, Kishore Malyavantham, Lucio G. Costa, Juan José Calvete, and Vicente Felipo. Activation of NMDA receptors induces protein kinase A-mediated phosphorylation and degradation of matrin 3. Blocking these effects prevents NMDA-induced neuronal death. *Journal of Neurochemistry*, 94(3):808–818, 2005. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1471-4159.2005.03235.x.

[301] Juan Liu, Cheng Qian, and Xuetao Cao. Post-Translational Modification Control of Innate Immunity. *Immunity*, 45(1):15–30, July 2016. Publisher: Elsevier.

[302] Tony Hunter. The Genesis of Tyrosine Phosphorylation. *Cold Spring Harbor Perspectives in Biology*, 6(5):a020644, May 2014.

[303] Lawrence L. Wang, Stéphane Richard, and Andrey S. Shaw. p62 Association with RNA Is Regulated by Tyrosine Phosphorylation (*). *Journal of Biological Chemistry*, 270(5):2010–2013, February 1995. Publisher: Elsevier.

[304] Teck Yew Low, M. Aiman Mohtar, Pey Yee Lee, Nursyazwani Omar, Houjiang Zhou, and Mingliang Ye. Widening the Bottleneck of Phosphoproteomics: Evolving Strategies for Phosphopeptide Enrichment. *Mass Spectrometry Reviews*, 40(4):309–333, 2021. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mas.21636.

[305] Dorte B. Bekker-Jensen, Oliver M. Bernhardt, Alexander Hogrebe, Ana Martinez-Val, Lynn Verbeke, Tejas Gandhi, Christian D. Kelstrup, Lukas Reiter, and Jesper V. Olsen. Rapid and site-specific deep phosphoproteome profiling by data-independent acquisition without the need for spectral libraries. *Nature Communications*, 11(1):787, February 2020.

[306] Ana Martínez-Val, Kyle Fort, Claire Koenig, Leander Van der Hoeven, Giulia Franciosa, Thomas Moehring, Yasushi Ishihama, Yu-ju Chen, Alexander Makarov, Yue Xuan, and Jesper V. Olsen. Hybrid-DIA: intelligent data acquisition integrates targeted and discovery proteomics to analyze phospho-signaling in single spheroids. *Nature Communications*, 14(1):3599, June 2023.

[307] Hideki Shigematsu, Tsuyoshi Imasaki, Chihiro Doki, Takuya Sumi, Mari Aoki, Tomomi Uchikubo-Kamo, Ayako Sakamoto, Kiyotaka Tokuraku, Mikako Shirouzu, and Ryo Nitta. Structural insight into microtubule stabilization and kinesin inhibition by Tau family MAPs. *Journal of Cell Biology*, 217(12):4155–4163, October 2018.

[308] M. Katsuki, K. Tokuraku, H. Murofushi, and S. Kotani. Functional analysis of microtubule-binding domain of bovine MAP4. *Cell Structure and Function*, 24(5):337–344, October 1999.

[309] A Mori, Aizawa H, Saido Tc, Kawasaki H, Mizuno K, Murofushi H, Suzuki K, and Sakai H. Site-specific phosphorylation by protein kinase C inhibits assembly-promoting activity of microtubule-associated protein 4. *Biochemistry*, 30(38), September 1991. Publisher: Biochemistry.

[310] K Ookata, Hisanaga S, Sugita M, Okuyama A, Murofushi H, Kitazawa H, Chari S, Bulinski Jc, and Kishimoto T. MAP4 is the in vivo substrate for CDC2 kinase in HeLa cells: identification of an M-phase specific and a cell cycle-independent phosphorylation site in MAP4. *Biochemistry*, 36(50), December 1997. Publisher: Biochemistry.

[311] V Srsen, Kitazawa H, Sugita M, Murofushi H, Bulinski Jc, Kishimoto T, and Hisanaga S. Serum-dependent phosphorylation of human MAP4 at Ser696 in cultured mammalian cells. *Cell structure and function*, 24(5), October 1999. Publisher: Cell Struct Funct.

[312] J. Hu, Z. Chu, J. Han, Q. Zhang, D. Zhang, Y. Dang, J. Ren, H. C. Chan, J. Zhang, and Y. Huang. Phosphorylation-dependent mitochondrial translocation of MAP4 is an early step in hypoxia-induced apoptosis in cardiomyocytes. *Cell Death & Disease*, 5(9):e1424–e1424, September 2014. Number: 9 Publisher: Nature Publishing Group.

[313] Romain A. Studer, Ricard A. Rodriguez-Mias, Kelsey M. Haas, Joanne I. Hsu, Cristina Viéitez, Carme Solé, Danielle L. Swaney, Lindsay B. Stanford, Ivan Liachko, René Böttcher, Maitreya J. Dunham, Eulàlia de Nadal, Francesc Posas, Pedro Beltrao,

and Judit Villén. Evolution of protein phosphorylation across 18 fungal species. *Science*, 354(6309):229–232, October 2016. Publisher: American Association for the Advancement of Science.

[314] Liam J. Holt, Brian B. Tuch, Judit Villén, Alexander D. Johnson, Steven P. Gygi, and David O. Morgan. Global Analysis of Cdk1 Substrate Phosphorylation Sites Provides Insights into Evolution. *Science*, 325(5948):1682–1686, September 2009. Publisher: American Association for the Advancement of Science.

[315] Débora M. Portilho, Roger Persson, and Nathalie Arhel. Role of non-motile microtubule-associated proteins in virus trafficking. *Biomolecular Concepts*, 7(5-6):283–292, December 2016.

[316] Daniel E. Gallo and Thomas J. Hope. Knockdown of MAP4 and DNAL1 produces a post-fusion and pre-nuclear translocation impairment in HIV-1 replication. *Virology*, 422(1):13–21, January 2012.

[317] Qisheng Li, Yong-Yuan Zhang, Stephan Chiu, Zongyi Hu, Keng-Hsin Lan, Helen Cha, Catherine Sodroski, Fang Zhang, Ching-Sheng Hsu, Emmanuel Thomas, and T. Jake Liang. Integrative Functional Genomics of Hepatitis C Virus Infection Identifies Host Dependencies in Complete Viral Replication Cycle. *PLOS Pathogens*, 10(5):e1004163, May 2014.

[318] Flora De Conto, Enrica Di Lonardo, Maria Cristina Arcangeletti, Carlo Chezzi, Maria Cristina Medici, and Adriana Calderaro. Highly Dynamic Microtubules Improve the Effectiveness of Early Stages of Human Influenza A/NWS/33 Virus Infection in LLC-MK2 Cells. *PLOS ONE*, 7(7):e41207, July 2012. Publisher: Public Library of Science.

[319] Rui Fu, Myanna T. Olsen, Kristofor Webb, Eric J. Bennett, and Jens Lykke-Andersen. Recruitment of the 4EHP-GYF2 cap-binding complex to tetraproline motifs of tristetraprolin promotes repression and degradation of mRNAs with AU-rich elements. *RNA*, 22(3):373–382, March 2016. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

[320] Isabel Myriam Schopp, Cinthia Claudia Amaya Ramirez, Jerneja Debeljak, Elisa Kreibich, Merle Skribbe, Klemens Wild, and Julien Béthune. Split-BioID a conditional proteomics approach to monitor the composition of spatiotemporally defined protein complexes. *Nature Communications*, 8(1):15690, June 2017. Publisher: Nature Publishing Group.

[321] Kelsey L. Hickey, Kimberley Dickson, J. Zachery Cogan, Joseph M. Replogle, Michael Schoof, Karole N. D'Orazio, Niladri K. Sinha, Jeffrey A. Hussmann, Marco Jost, Adam Frost, Rachel Green, Jonathan S. Weissman, and Kamena K. Kostova. GIGYF2 and 4EHP Inhibit Translation Initiation of Defective Messenger RNAs to Assist Ribosome-Associated Quality Control. *Molecular Cell*, 79(6):950–962.e6, September 2020. Publisher: Elsevier.

[322] Xu Zhang, Clément Chapat, Peng Wang, Jung-Hyun Choi, Qian Li, Jun Luo, Shane Wiebe, Sung-Hoon Kim, Nathaniel Robichaud, Isabela Fabri Karam, David Dai, Angela P. Hackett, Rongtuan Lin, Tommy Alain, Long Yang, Seyed Mehdi Jafarnejad, and Nahum Sonenberg. microRNA-induced translational control of antiviral immunity by the cap-binding protein 4EHP. *Molecular Cell*, 81(6):1187–1199.e5, March 2021.

[323] Parisa Naeli, Xu Zhang, Patric Harris Snell, Susanta Chatterjee, Muhammad Kamran, Reese Jalal Ladak, Nick Orr, Thomas Duchaine, Nahum Sonenberg, and Seyed Mehdi Jafarnejad. The SARS-CoV-2 protein NSP2 enhances microRNA-mediated translational repression. *Journal of Cell Science*, 136(19):jcs261286, October 2023.

[324] Joana Ribeiro, Joana Marinho-Dias, Paula Monteiro, Joana Loureiro, Inês Baldaque, Rui Medeiros, and Hugo Sousa. miR-34a and miR-125b Expression in HPV Infection and Cervical Cancer Development. *BioMed Research International*, 2015:e304584, June 2015. Publisher: Hindawi.

[325] Yasuhide Hibino, Tatsuhiro Usui, Yasuhiro Morita, Noriko Hirose, Mari Okazaki, Nobuhiko Sugano, and Koichi Hiraga. Molecular properties and intracellular localization of rat liver nuclear scaffold protein P130. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression*, 1759(5):195–207, May 2006.

[326] Shoji Hisada-Ishii, Mizuki Ebihara, Nao Kobayashi, and Yasuo Kitagawa. Bipartite nuclear localization signal of matrin 3 is essential for vertebrate cells. *Biochemical and Biophysical Research Communications*, 354(1):72–76, March 2007.

[327] Maayan Salton, Ran Elkon, Tatiana Borodina, Aleksey Davydov, Marie-Laure Yaspo, Eran Halperin, and Yosef Shiloh. Matrin 3 Binds and Stabilizes mRNA. *PLOS ONE*, 6(8):e23882, August 2011. Publisher: Public Library of Science.

[328] Angela Erazo, Michael B. Yee, Bruce W. Banfield, and Paul R. Kinchington. The Alphaherpesvirus US3/ORF66 Protein Kinases Direct Phosphorylation of the Nuclear Matrix Protein Matrin 3. *Journal of Virology*, 85(1):568–581, January 2011. Publisher: American Society for Microbiology.

[329] Yasuhide Hibino, Hiromitsu Ohzeki, Noriko Hirose, and Nobuhiko Sugano. Involvement of phosphorylation in binding of nuclear scaffold proteins from rat liver to a highly repetitive DNA component. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression*, 1396(1):88–96, March 1998.

[330] Fumiyoshi Yamazaki, Hyun Hee Kim, Pierre Lau, Christopher K. Hwang, P. Michael Iuvone, David Klein, and Samuel J. H. Clokie. pY RNA1-s2: A Highly Retina-Enriched Small RNA That Selectively Binds to Matrin 3 (Matr3). *PLOS ONE*, 9(2):e88217, February 2014. Publisher: Public Library of Science.

[331] Kiyoshi Isobe, Hyun Jun Jung, Chin-Rang Yang, J'Neka Claxton, Pablo Sandoval, Maurice B. Burg, Viswanathan Raghuram, and Mark A. Knepper. Systems-level identification of PKA-dependent signaling in epithelial cells. *Proceedings of the National Academy of Sciences*, 114(42):E8875–E8884, October 2017. Publisher: Proceedings of the National Academy of Sciences.

[332] Miguel B Coelho, Jan Attig, Nicolás Bellora, Julian König, Martina Hallegger, Melis Kayikci, Eduardo Eyras, Jernej Ule, and Christopher WJ Smith. Nuclear matrix protein Matrin3 regulates alternative splicing and forms overlapping regulatory networks with PTB. *The EMBO Journal*, 34(5):653–668, March 2015. Publisher: John Wiley & Sons, Ltd.

[333] Martina Hallegger, Anob M. Chakrabarti, Flora C. Y. Lee, Bo Lim Lee, Aram G. Amalietti, Hana M. Odeh, Katie E. Copley, Jack D. Rubien, Bede Portz, Klara Kuret, Ina Huppertz,

Frédérique Rau, Rickie Patani, Nicolas L. Fawzi, James Shorter, Nicholas M. Luscombe, and Jernej Ule. TDP-43 condensation properties specify its RNA-binding and regulatory repertoire. *Cell*, 184(18):4680–4696.e22, September 2021.

[334] Idir Malki, Inara Liepina, Nora Kogelnik, Hollie Watmuff, Sue Robinson, Adam Lightfoot, Oksana Gonchar, Andrew Bottrill, Andrew M Fry, and Cyril Dominguez. Cdk1-mediated threonine phosphorylation of Sam68 modulates its RNA binding, alternative splicing activity and cellular functions. *Nucleic Acids Research*, 50(22):13045–13062, December 2022.

[335] Yuri Uemura, Takuya Oshima, Munetaka Yamamoto, Charles Jourdan Reyes, Pedro Henrique Costa Cruz, Toshiharu Shibuya, and Yukio Kawahara. Matrin3 binds directly to intronic pyrimidine-rich sequences and controls alternative splicing. *Genes to Cells*, 22(9):785–798, 2017. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/gtc.12512.

[336] Michelle J. Farquhar, Helen J. Harris, Mandy Diskar, Sarah Jones, Christopher J. Mee, Søren U. Nielsen, Claire L. Brimacombe, Sonia Molina, Geoffrey L. Toms, Patrick Maurel, John Howl, Friedrich W. Herberg, Sven C. D. van IJzendoorn, Peter Balfe, and Jane A. McKeating. Protein Kinase A-Dependent Step(s) in Hepatitis C Virus Entry and Infectivity. *Journal of Virology*, 82(17):8797–8811, September 2008. Publisher: American Society for Microbiology.

[337] Julian Scherer, Julie Yi, and Richard B. Vallee. PKA-dependent dynein switching from lysosomes to adenovirus: A novel form of host–virus competition. *Journal of Cell Biology*, 205(2):163–177, April 2014.

[338] Fan Cheng, Suzane Ramos da Silva, I-Chueh Huang, Jae U. Jung, and Shou-Jiang Gao. Suppression of Zika Virus Infection and Replication in Endothelial Cells and Astrocytes by PKA Inhibitor PKI 14-22. *Journal of Virology*, 92(4):e02019–17, January 2018. Publisher: American Society for Microbiology.

[339] Lin Zhu, Ting Gao, Yi Huang, Jing Jin, Di Wang, Leike Zhang, Yanwen Jin, Ping Li, Yong Hu, Yan Wu, Hainan Liu, Qincai Dong, Guangfei Wang, Tong Zheng, Caiwei Song, Yu Bai, Xun Zhang, Yaoning Liu, Weihong Yang, Ke Xu, Gang Zou, Lei Zhao, Ruiyuan Cao, Wu Zhong, Xianzhu Xia, Gengfu Xiao, Xuan Liu, and Cheng Cao. Ebola

virus VP35 hijacks the PKA-CREB1 pathway for replication and pathogenesis by AKIP1 association. *Nature Communications*, 13(1):1–13, April 2022. Number: 1 Publisher: Nature Publishing Group.

[340] Alicia Collado-Hidalgo, Caroline Sung, and Steve Cole. Adrenergic inhibition of innate anti-viral response: PKA blockade of Type I interferon gene transcription mediates catecholamine support for HIV-1 replication. *Brain, Behavior, and Immunity*, 20(6):552–563, November 2006.

[341] Bing-Ru Yan, Lu Zhou, Ming-Ming Hu, Mi Li, Heng Lin, Yan Yang, Yan-Yi Wang, and Hong-Bing Shu. PKACs attenuate innate antiviral response by phosphorylating VISA and priming it for MARCH5-mediated degradation. *PLOS Pathogens*, 13(9):e1006648, September 2017. Publisher: Public Library of Science.

[342] Yi Guo, Xia-Nan Zhang, Shan Su, Zi-Lun Ruan, Ming-Ming Hu, and Hong-Bing Shu. β-adrenoreceptor-triggered PKA activation negatively regulates the innate antiviral response. *Cellular & Molecular Immunology*, pages 1–14, January 2023. Publisher: Nature Publishing Group.

[343] Chengfei Zhang, Yan Yan, Hongwang He, Li Wang, Na Zhang, Jie Zhang, Hongjun Huang, Nannan Wu, Hua Ren, Min Qian, Mingyao Liu, and Bing Du. IFN-stimulated P2Y13 protects mice from viral infection by suppressing the cAMP/EPAC1 signaling pathway. *Journal of Molecular Cell Biology*, 11(5):395–407, May 2019.

[344] Philipp Skroblin, Solveig Grossmann, Gesa Schäfer, Walter Rosenthal, and Enno Klussmann. Chapter Five - Mechanisms of Protein Kinase A Anchoring. In Kwang Jeon, editor, *International Review of Cell and Molecular Biology*, volume 283, pages 235–330. Academic Press, January 2010.

[345] Terri Clister, Eric C. Greenwald, George S. Baillie, and Jin Zhang. AKAP95 Organizes a Nuclear Microdomain to Control Local cAMP for Regulating Nuclear PKA. *Cell Chemical Biology*, 26(6):885–891.e4, June 2019. Publisher: Elsevier.

[346] Jing Hu, Alireza Khodadadi-Jamayran, Miaowei Mao, Kushani Shah, Zhenhua Yang, Md Talat Nasim, Zefeng Wang, and Hao Jiang. AKAP95 regulates splicing through scaffolding RNAs and RNA processing factors. *Nature Communications*, 7(1):13347, November 2016. Number: 1 Publisher: Nature Publishing Group.

[347] Xiaohui Hu, Samuel E. Harvey, Rong Zheng, Jingyi Lyu, Caitlin L. Grzeskowiak, Emily Powell, Helen Piwnica-Worms, Kenneth L. Scott, and Chonghui Cheng. The RNA-binding protein AKAP8 suppresses tumor metastasis by antagonizing EMT-associated alternative splicing. *Nature Communications*, 11(1):486, January 2020. Number: 1 Publisher: Nature Publishing Group.

[348] Mathieu Bollen, Wolfgang Peti, Michael J. Ragusa, and Monique Beullens. The extended PP1 toolkit: designed to create specificity. *Trends in Biochemical Sciences*, 35(8):450–458, August 2010.

[349] Young-Mi Kim, Takuo Watanabe, Patrick B. Allen, Young-Myoung Kim, Shin-Jeong Lee, Paul Greengard, Angus C. Nairn, and Young-Guen Kwon. PNUTS, a Protein Phosphatase 1 (PP1) Nuclear Targeting Subunit: CHARACTERIZATION OF ITS PP1- AND RNA-BINDING DOMAINS AND REGULATION BY PHOSPHORYLATION *. *Journal of Biological Chemistry*, 278(16):13819–13828, April 2003. Publisher: Elsevier.

[350] Anne M. Devlin, Ashutosh Shukla, Julio C. Ruiz, Spencer D. Barnes, Ashwin Govindan, Olga V. Hunter, Anna M. Scarborough, Iván D'Orso, and Nicholas K. Conrad. The PNUTS-PP1 complex acts as an intrinsic barrier to herpesvirus KSHV gene expression and replication. *Nature Communications*, 13(1):7447, December 2022. Number: 1 Publisher: Nature Publishing Group.

[351] R. Berezney and D. S. Coffey. Nuclear matrix. Isolation and characterization of a framework structure from rat liver nuclei. *The Journal of Cell Biology*, 73(3):616–637, June 1977.

[352] Ahmed M. Malik and Sami J. Barmada. Matrin 3 in neuromuscular disease: physiology and pathophysiology. *JCI Insight*, 6(1), January 2021. Publisher: American Society for Clinical Investigation.

[353] Kishore S. Malyavantham, Sambit Bhattacharya, Marcos Barbeitos, Lopamudra Mukherjee, Jinhui Xu, Frank O. Fackelmayer, and Ronald Berezney. Identifying functional neighborhoods within the cell nucleus: Proximity analysis of early S-phase replicating chromatin domains to sites of transcription, RNA polymerase II, HP1γ, matrin 3 and SAF-A. *Journal of Cellular Biochemistry*, 105(2):391–403, 2008. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcb.21834.

[354] Dorota Skowronska-Krawczyk, Qi Ma, Michal Schwartz, Kathleen Scully, Wenbo Li, Zhijie Liu, Havilah Taylor, Jessica Tollkuhn, Kenneth A. Ohgi, Dimple Notani, Yoshinori Kohwi, Terumi Kohwi-Shigematsu, and Michael G. Rosenfeld. Required enhancer–matrin-3 network interactions for a homeodomain transcription program. *Nature*, 514(7521):257–261, October 2014. Number: 7521 Publisher: Nature Publishing Group.

[355] Ritu Chaudhary, Berkley Gryder, Wendy S Woods, Murugan Subramanian, Matthew F Jones, Xiao Ling Li, Lisa M Jenkins, Svetlana A Shabalina, Min Mo, Mary Dasso, Yuan Yang, Lalage M Wakefield, Yuelin Zhu, Susan M Frier, Branden S Moriarity, Kannanganattu V Prasanth, Pablo Perez-Pinera, and Ashish Lal. Prosurvival long noncoding RNA PINCR regulates a subset of p53 targets in human colorectal cancer cells by binding to Matrin 3. *eLife*, 6:e23244, June 2017. Publisher: eLife Sciences Publications, Ltd.

[356] Jan Attig, Federico Agostini, Clare Gooding, Anob M. Chakrabarti, Aarti Singh, Nejc Haberman, Julian A. Zagalak, Warren Emmett, Christopher W. J. Smith, Nicholas M. Luscombe, and Jernej Ule. Heteromeric RNP Assembly at LINEs Controls Lineage-Specific RNA Processing. *Cell*, 174(5):1067–1081.e17, August 2018.

[357] Andrey Damianov, Yi Ying, Chia-Ho Lin, Ji-Ann Lee, Diana Tran, Ajay A. Vashisht, Emad Bahrami-Samani, Yi Xing, Kelsey C. Martin, James A. Wohlschlegel, and Douglas L. Black. Rbfox Proteins Regulate Splicing as Part of a Large Multiprotein Complex LASR. *Cell*, 165(3):606–619, April 2016. Publisher: Elsevier.

[358] Maayan Salton, Yaniv Lerenthal, Shih-Ya Wang, David J. Chen, and Yosef Shiloh. Involvement of Matrin 3 and SFPQ/NONO in the DNA damage response. *Cell Cycle*, 9(8):1568–1576, April 2010. Publisher: Taylor & Francis _eprint: https://doi.org/10.4161/cc.9.8.11298.

[359] Daniele Pollini, Rosa Loffredo, Federica Maniscalco, Marina Cardano, Mariachiara Micaelli, Isabelle Bonomo, Nausicaa Valentina Licata, Daniele Peroni, Weronika Tomaszewska, Annalisa Rossi, Valeria Crippa, Erik Dassi, Gabriella Viero, Alessandro Quattrone, Angelo Poletti, Luciano Conti, and Alessandro Provenzani. Multilayer and MATR3-dependent regulation of mRNAs maintains pluripotency in human induced pluripotent stem cells. *iScience*, 24(3):102197, February 2021.

[360] Anna Kula, Jessica Guerra, Anna Knezevich, Danijela Kleva, Michael P. Myers, and Alessandro Marcello. Characterization of the HIV-1 RNA associated proteome identifies Matrin 3 as a nuclear cofactor of Rev function. *Retrovirology*, 8(1):60, July 2011.

[361] Venkat SRK Yedavalli and Kuan-Teh Jeang. Matrin 3 is a co-factor for HIV-1 Rev in regulating post-transcriptional viral gene expression. *Retrovirology*, 8(1):61, July 2011.

[362] Angela Erazo and Stephen P. Goff. Nuclear matrix protein Matrin 3 is a regulator of ZAP-mediated retroviral restriction. *Retrovirology*, 12(1):57, July 2015.

[363] Mehdi Morchikh, Alexandra Cribier, Raoul Raffel, Sonia Amraoui, Julien Cau, Dany Severac, Emeric Dubois, Olivier Schwartz, Yamina Bennasser, and Monsef Benkirane. HEXIM1 and NEAT1 Long Non-coding RNA Form a Multi-subunit Complex that Regulates DNA-Mediated Innate Immune Response. *Molecular Cell*, 67(3):387–399.e5, August 2017. Publisher: Elsevier.

[364] Wael Kamel, Vincenzo Ruscica, Manuel Garcia-Moreno, Natasha Palmalux, Louisa Iselin, Maximilian Hannan, Aino Järvelin, Marko Noerenberg, Samantha Moore, Andres Merits, Ilan Davis, Javier Martinez, Shabaz Mohammed, and Alfredo Castello. Compositional analysis of Sindbis virus ribonucleoproteins reveals an extensive co-opting of key nuclear RNA-binding proteins, October 2021. Pages: 2021.10.06.463336 Section: New Results.

[365] Dipen Rajgor, Jonathan G. Hanley, and Catherine M. Shanahan. Identification of novel nesprin-1 binding partners and cytoplasmic matrin-3 in processing bodies. *Molecular Biology of the Cell*, 27(24):3894–3902, December 2016. Publisher: American Society for Cell Biology (mboc).

[366] Holger Dinkel, Sushama Michael, Robert J. Weatheritt, Norman E. Davey, Kim Van Roey, Brigitte Altenberg, Grischa Toedt, Bora Uyar, Markus Seiler, Aidan Budd, Lisa Jödicke, Marcel A. Dammert, Christian Schroeter, Maria Hammer, Tobias Schmidt, Peter Jehl, Caroline McGuigan, Magdalena Dymecka, Claudia Chica, Katja Luck, Allegra Via, Andrew Chatr-aryamontri, Niall Haslam, Gleb Grebnev, Richard J. Edwards, Michel O. Steinmetz, Heike Meiselbach, Francesca Diella, and Toby J. Gibson. ELM—the database of eukaryotic linear motifs. *Nucleic Acids Research*, 40(Database issue):D242–D251, January 2012.

[367] Miguel B. Coelho, David B. Ascher, Clare Gooding, Emma Lang, Hannah Maude, David Turner, Miriam Llorian, Douglas E.V. Pires, Jan Attig, and Christopher W.J. Smith. Functional interactions between polypyrimidine tract binding protein and PRI peptide ligand containing proteins. *Biochemical Society Transactions*, 44(4):1058–1065, August 2016.

[368] Andreas Buchbender, Holger Mutter, F. X. Reymond Sutandy, Nadine Körtel, Heike Hänel, Anke Busch, Stefanie Ebersberger, and Julian König. Improved library preparation with the new iCLIP2 protocol. *Methods*, 178:33–48, June 2020.

[369] Sudeep Sahadevan, Thileepan Sekaran, Nadia Ashaf, Marko Fritz, Matthias W Hentze, Wolfgang Huber, and Thomas Schwarzl. htseq-clip: a toolset for the preprocessing of eCLIP/iCLIP datasets. *Bioinformatics*, page btac747, November 2022.

[370] Timothy L. Bailey, James Johnson, Charles E. Grant, and William S. Noble. The MEME Suite. *Nucleic Acids Research*, 43(Web Server issue):W39–W49, July 2015.

[371] Katrina M Kutchko, Emily A Madden, Clayton Morrison, Kenneth S Plante, Wes Sanders, Heather A Vincent, Marta C Cruz Cisneros, Kristin M Long, Nathaniel J Moorman, Mark T Heise, and Alain Laederach. Structural divergence creates new functional features in alphavirus genomes. *Nucleic Acids Research*, 46(7):3657–3670, April 2018.

[372] Lucia Alvarez, Kevin Haubrich, Louisa Iselin, Laurent Gillioz, Vincenzo Ruscica, Karine Lapouge, Sandra Augsten, Ina Huppertz, Nila Roy Choudhury, Bernd Simon, Pawel Masiewicz, and Mathilde Lethier. The molecular dissection of TRIM25's RNA-binding mechanism provides key insights into its antiviral activity, December 2023.

[373] Rose Oughtred, Jennifer Rust, Christie Chang, Bobby-Joe Breitkreutz, Chris Stark, Andrew Willems, Lorrie Boucher, Genie Leung, Nadine Kolas, Frederick Zhang, Sonam Dolma, Jasmin Coulombe-Huntington, Andrew Chatr-aryamontri, Kara Dolinski, and Mike Tyers. The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Science : A Publication of the Protein Society*, 30(1):187–200, January 2021.

[374] Lolita Piersimoni, Panagiotis L. Kastritis, Christian Arlt, and Andrea Sinz. Cross-Linking Mass Spectrometry for Investigating Protein Conformations and Protein–Protein

Interactions-A Method for All Seasons. *Chemical Reviews*, 122(8):7500–7531, April 2022. Publisher: American Chemical Society.

[375] Ian R. Humphreys, Jimin Pei, Minkyung Baek, Aditya Krishnakumar, Ivan Anishchenko, Sergey Ovchinnikov, Jing Zhang, Travis J. Ness, Sudeep Banjade, Saket R. Bagde, Viktoriya G. Stancheva, Xiao-Han Li, Kaixian Liu, Zhi Zheng, Daniel J. Barrero, Upasana Roy, Jochen Kuper, Israel S. Fernández, Barnabas Szakal, Dana Branzei, Josep Rizo, Caroline Kisker, Eric C. Greene, Sue Biggins, Scott Keeney, Elizabeth A. Miller, J. Christopher Fromme, Tamara L. Hendrickson, Qian Cong, and David Baker. Computed structures of core eukaryotic protein complexes. *Science*, 374(6573):eabm4805, November 2021. Publisher: American Association for the Advancement of Science.

[376] T. Reid Alderson, Iva Pritišanac, Đesika Kolarić, Alan M. Moses, and Julie D. Forman-Kay. Systematic identification of conditionally folded intrinsically disordered regions by AlphaFold2. *Proceedings of the National Academy of Sciences*, 120(44):e2304302120, October 2023. Publisher: Proceedings of the National Academy of Sciences.

[377] Francis J. O'Reilly and Juri Rappsilber. Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. *Nature Structural & Molecular Biology*, 25(11):1000–1008, November 2018. Number: 11 Publisher: Nature Publishing Group.

[378] Maiwen Caudron-Herger, Scott F. Rusin, Mark E. Adamo, Jeanette Seiler, Vera K. Schmid, Elsa Barreau, Arminja N. Kettenbach, and Sven Diederichs. R-DeeP: Proteome-wide and Quantitative Identification of RNA-Dependent Proteins by Density Gradient Ultracentrifugation. *Molecular Cell*, 75(1):184–199.e10, July 2019.

[379] Kristopher W. Brannan, Wenhao Jin, Stephanie C. Huelga, Charles A. S. Banks, Joshua M. Gilmore, Laurence Florens, Michael P. Washburn, Eric L. Van Nostrand, Gabriel A. Pratt, Marie K. Schwinn, Danette L. Daniels, and Gene W. Yeo. SONAR Discovers RNA-Binding Proteins from Analysis of Large-Scale Protein-Protein Interactomes. *Molecular Cell*, 64(2):282–293, October 2016. Publisher: Elsevier.

[380] Matthew J. Bick, John-William N. Carroll, Guangxia Gao, Stephen P. Goff, Charles M. Rice, and Margaret R. MacDonald. Expression of the Zinc-Finger Antiviral Protein

Inhibits Alphavirus Replication. *Journal of Virology*, 77(21):11555–11562, November 2003. Publisher: American Society for Microbiology.

[381] Jennifer L. Meagher, Matthew Takata, Daniel Gonçalves-Carneiro, Sarah C. Keane, Antoine Rebendenne, Heley Ong, Victoria K. Orr, Margaret R. MacDonald, Jeanne A. Stuckey, Paul D. Bieniasz, and Janet L. Smith. Structure of the zinc-finger antiviral protein in complex with RNA reveals a mechanism for selective targeting of CG-rich viral sequences. *Proceedings of the National Academy of Sciences*, 116(48):24303–24309, November 2019. Publisher: Proceedings of the National Academy of Sciences.

[382] Mattia Ficarelli, Stuart J.D. Neil, and Chad M. Swanson. Targeted Restriction of Viral Gene Expression and Replication by the ZAP Antiviral System. *Annual Review of Virology*, 8(1):265–283, 2021. _eprint: https://doi.org/10.1146/annurev-virology-091919-104213.

[383] H. A. Bluyssen, R. J. Vlietstra, P. W. Faber, E. M. Smit, A. Hagemeijer, and J. Trapman. Structure, chromosome localization, and regulation of expression of the interferon-regulated mouse Ifi54/Ifi56 gene family. *Genomics*, 24(1):137–148, November 1994.

[384] Yazan M. Abbas, Beatrice Theres Laudenbach, Saúl Martínez-Montero, Regina Cencic, Matthias Habjan, Andreas Pichlmair, Masad J. Damha, Jerry Pelletier, and Bhushan Nagar. Structure of human IFIT1 with capped RNA reveals adaptable mRNA binding and mechanisms for sensing N1 and N2 ribose 2'-O methylations. *Proceedings of the National Academy of Sciences*, 114(11):E2106–E2115, March 2017. Publisher: Proceedings of the National Academy of Sciences.

[385] Michael S. Diamond and Michael Farzan. The broad-spectrum antiviral functions of IFIT and IFITM proteins. *Nature Reviews Immunology*, 13(1):46–57, January 2013. Number: 1 Publisher: Nature Publishing Group.

[386] Francisco Inesta-Vaquera and Victoria H Cowling. Regulation and function of CMTR1-dependent mRNA cap methylation. *WIREs RNA*, 8(6):e1450, 2017. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wrna.1450.

[387] Stephane Daffis, Kristy J. Szretter, Jill Schriewer, Jianqing Li, Soonjeon Youn, John Errett, Tsai-Yu Lin, Stewart Schneller, Roland Zust, Hongping Dong, Volker Thiel, Ganes C. Sen, Volker Fensterl, William B. Klimstra, Theodore C. Pierson, R. Mark Buller,

Michael Gale Jr, Pei-Yong Shi, and Michael S. Diamond. 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature*, 468(7322):452–456, November 2010. Publisher: Nature Publishing Group.

[388] Jennifer L. Hyde, Christina L. Gardner, Taishi Kimura, James P. White, Gai Liu, Derek W. Trobaugh, Cheng Huang, Marco Tonelli, Slobodan Paessler, Kiyoshi Takeda, William B. Klimstra, Gaya K. Amarasinghe, and Michael S. Diamond. A Viral RNA Structural Element Alters Host Recognition of Nonself RNA. *Science*, 343(6172):783–787, February 2014. Publisher: American Association for the Advancement of Science.

[389] Josephine M. Reynaud, Dal Young Kim, Svetlana Atasheva, Aliaksandra Rasalouskaya, James P. White, Michael S. Diamond, Scott C. Weaver, Elena I. Frolova, and Ilya Frolov. IFIT1 Differentially Interferes with Translation and Replication of Alphavirus Genomes and Promotes Induction of Type I Interferon. *PLoS Pathogens*, 11(4):e1004863, April 2015.

[390] François Bélanger, Janusz Stepinski, Edward Darzynkiewicz, and Jerry Pelletier. Characterization of hMTr1, a Human Cap1 2'-O-Ribose Methyltransferase *. *Journal of Biological Chemistry*, 285(43):33037–33044, October 2010. Publisher: Elsevier.

[391] Xavier Rambout and Lynne E. Maquat. NCBP3: A Multifaceted Adaptive Regulator of Gene Expression. *Trends in Biochemical Sciences*, 46(2):87–96, February 2021.

[392] P. Fortes, T. Inada, T. Preiss, M. W. Hentze, I. W. Mattaj, and A. B. Sachs. The yeast nuclear cap binding complex can interact with translation factor eIF4G and mediate translation initiation. *Molecular Cell*, 6(1):191–196, July 2000.

[393] Anna Gebhardt, Valter Bergant, Daniel Schnepf, Markus Moser, Arno Meiler, Dieudonnée Togbe, Claire Mackowiak, Line S. Reinert, Søren R. Paludan, Bernhard Ryffel, Alexey Stukalov, Peter Staeheli, and Andreas Pichlmair. The alternative cap-binding complex is required for antiviral defense in vivo. *PLOS Pathogens*, 15(12):e1008155, December 2019. Publisher: Public Library of Science.

[394] Pierre C. Havugimana, G. Traver Hart, Tamás Nepusz, Haixuan Yang, Andrei L. Turinsky, Zhihua Li, Peggy I. Wang, Daniel R. Boutz, Vincent Fong, Sadhna Phanse, Mohan Babu, Stephanie A. Craig, Pingzhao Hu, Cuihong Wan, James Vlasblom, Vaqaar-un-Nisa Dar, Alexandr Bezginov, Gregory W. Clark, Gabriel C. Wu, Shoshana J. Wodak, Elisabeth

R. M. Tillier, Alberto Paccanaro, Edward M. Marcotte, and Andrew Emili. A census of human soluble protein complexes. *Cell*, 150(5):1068–1081, August 2012.

[395] Franz Herzog, Abdullah Kahraman, Daniel Boehringer, Raymond Mak, Andreas Bracher, Thomas Walzthoeni, Alexander Leitner, Martin Beck, Franz-Ulrich Hartl, Nenad Ban, Lars Malmström, and Ruedi Aebersold. Structural Probing of a Protein Phosphatase 2A Network by Chemical Cross-Linking and Mass Spectrometry. *Science*, 337(6100):1348–1352, September 2012. Publisher: American Association for the Advancement of Science.

[396] Sonali Joshi, Surinder Kaur, Barbara Kroczynska, and Leonidas C. Platanias. Mechanisms of mRNA translation of interferon stimulated genes. *Cytokine*, 52(1-2):123–127, 2010.

[397] Klara Kuret, Aram Gustav Amalietti, D. Marc Jones, Charlotte Capitanchik, and Jernej Ule. Positional motif analysis reveals the extent of specificity of protein-RNA interactions observed by CLIP. *Genome Biology*, 23(1):191, September 2022.

[398] Inga Loedige, Mathias Stotz, Saadia Qamar, Katharina Kramer, Janosch Hennig, Thomas Schubert, Patrick Löffler, Gernot Längst, Rainer Merkl, Henning Urlaub, and Gunter Meister. The NHL domain of BRAT is an RNA-binding domain that directly contacts the hunchback mRNA for regulation. *Genes & Development*, 28(7):749–764, April 2014. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

[399] H G Niesters and J H Strauss. Mutagenesis of the conserved 51-nucleotide region of Sindbis virus. *Journal of Virology*, 64(4):1639–1647, April 1990.

[400] I. Frolov, R. Hardy, and C. M. Rice. Cis-acting RNA elements at the 5' end of Sindbis virus genome RNA regulate minus- and plus-strand RNA synthesis. *RNA*, 7(11):1638–1651, November 2001. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

[401] Rafik Fayzulin and Ilya Frolov. Changes of the Secondary Structure of the 5' End of the Sindbis Virus Genome Inhibit Virus Growth in Mosquito Cells and Lead to Accumulation

of Adaptive Mutations. *Journal of Virology*, 78(10):4953–4964, May 2004. Publisher: American Society for Microbiology.

[402] Gilles Michel, Olga Petrakova, Svetlana Atasheva, and Ilya Frolov. Adaptation of Venezuelan equine encephalitis virus lacking 51-nt conserved sequence element to replication in mammalian and mosquito cells. *Virology*, 362(2):475–487, June 2007.

[403] Rhian Jones, Gabriel Bragagnolo, Rocío Arranz, and Juan Reguera. Capping pores of alphavirus nsP1 gate membranous viral replication factories. *Nature*, 589(7843):615–619, January 2021. Publisher: Nature Publishing Group.

[404] Rhian Jones, Michael Hons, Nadia Rabah, Noelia Zamarreño, Rocío Arranz, and Juan Reguera. Structural basis and dynamics of Chikungunya alphavirus RNA capping by nsP1 capping pores. *Proceedings of the National Academy of Sciences*, 120(12):e2213934120, March 2023. Publisher: Proceedings of the National Academy of Sciences.

[405] Michelle Cheok Yien Law, Kuo Zhang, Yaw Bia Tan, Trinh Mai Nguyen, and Dahai Luo. Chikungunya virus nonstructural protein 1 is a versatile RNA capping and decapping enzyme. *Journal of Biological Chemistry*, 299(12), December 2023. Publisher: Elsevier.

[406] Jonathan C. Rupp, Kevin J. Sokoloski, Natasha N. Gebhart, and Richard W. Hardy. Alphavirus RNA synthesis and non-structural protein functions. *The Journal of General Virology*, 96(Pt 9):2483–2500, September 2015.

[407] Miguel A. Sanz, Alfredo Castelló, and Luis Carrasco. Viral Translation Is Coupled to Transcription in Sindbis Virus-Infected Cells. *Journal of Virology*, 81(13):7061–7068, July 2007. Publisher: American Society for Microbiology.

[408] Age Utt, Tania Quirin, Sirle Saul, Kirsi Hellström, Tero Ahola, and Andres Merits. Versatile Trans-Replication Systems for Chikungunya Virus Allow Functional Analysis and Tagging of Every Replicase Protein. *PLOS ONE*, 11(3):e0151616, March 2016.

[409] Derek Walsh, Michael B. Mathews, and Ian Mohr. Tinkering with translation: protein synthesis in virus-infected cells. *Cold Spring Harbor Perspectives in Biology*, 5(1):a012351, January 2013.

[410] Urbi Mukhopadyay, Sophie Levantovsky, Sarah Gharbi, Frank Stein, Christian Behrends, and Sagar Bhogaraju. A ubiquitin-specific, proximity-based labeling approach for the identification of ubiquitin ligase substrates, February 2024. Pages: 2023.09.04.556194 Section: New Results.

[411] Christoph Thomas, Ignacio Moraga, Doron Levin, Peter O. Krutzik, Yulia Podoplelova, Angelica Trejo, Choongho Lee, Ganit Yarden, Susan E. Vleck, Jeffrey S. Glenn, Garry P. Nolan, Jacob Piehler, Gideon Schreiber, and K. Christopher Garcia. Structural Linkage between Ligand Discrimination and Receptor Activation by Type I Interferons. *Cell*, 146(4):621–632, August 2011.

[412] Diego A. Jaitin, Laila C. Roisman, Eva Jaks, Martynas Gavutis, Jacob Piehler, Jose Van der Heyden, Gilles Uze, and Gideon Schreiber. Inquiring into the Differential Action of Interferons (IFNs): an IFN-α2 Mutant with Enhanced Affinity to IFNAR1 Is Functionally Similar to IFN-β. *Molecular and Cellular Biology*, 26(5):1888–1897, March 2006. Publisher: Taylor & Francis _eprint: https://doi.org/10.1128/MCB.26.5.1888-1897.2006.

[413] Candice Mazewski, Ricardo E. Perez, Eleanor N. Fish, and Leonidas C. Platanias. Type I Interferon (IFN)-Regulated Activation of Canonical and Non-Canonical Signaling Pathways. *Frontiers in Immunology*, 11:606456, November 2020.

[414] Helen M. Lazear, John W. Schoggins, and Michael S. Diamond. Shared and Distinct Functions of Type I and Type III Interferons. *Immunity*, 50(4):907–923, April 2019.

[415] Kalliopi Pervolaraki, Megan L. Stanifer, Stephanie Münchau, Lynnsey A. Renn, Dorothee Albrecht, Stefan Kurzhals, Elena Senís, Dirk Grimm, Jutta Schröder-Braunstein, Ronald L. Rabin, and Steeve Boulant. Type I and Type III Interferons Display Different Dependency on Mitogen-Activated Protein Kinases to Mount an Antiviral State in the Human Gut. *Frontiers in Immunology*, 8:459, 2017.

[416] Ryan A. Flynn, Julia A. Belk, Yanyan Qi, Yuki Yasumoto, Jin Wei, Mia Madel Alfajaro, Quanming Shi, Maxwell R. Mumbach, Aditi Limaye, Peter C. DeWeirdt, Cameron O. Schmitz, Kevin R. Parker, Elizabeth Woo, Howard Y. Chang, Tamas L. Horvath, Jan E. Carette, Carolyn R. Bertozzi, Craig B. Wilen, and Ansuman T. Satpathy. Discovery

and functional interrogation of SARS-CoV-2 RNA-host protein interactions. *Cell*, 184(9):2394–2411.e16, April 2021. Publisher: Elsevier.

[417] Chongyang Li, Trent G. Nelson, Alfred C. O. Vertegaal, and Pierre Thibault. Proteomic strategies for characterizing ubiquitin-like modifications. *Nature Reviews Methods Primers*, 1(1):1–21, July 2021. Publisher: Nature Publishing Group.

[418] Meredith E. Davis and Michaela U. Gack. Ubiquitination in the antiviral immune response. *Virology*, 479-480:52–65, May 2015.

[419] Deborah J. Lenschow, Caroline Lai, Natalia Frias-Staheli, Nadia V. Giannakopoulos, Andrew Lutz, Thorsten Wolff, Anna Osiak, Beth Levine, Robert E. Schmidt, Adolfo García-Sastre, David A. Leib, Andrew Pekosz, Klaus-Peter Knobeloch, Ivan Horak, and Herbert Whiting Virgin. IFN-stimulated gene 15 functions as a critical antiviral molecule against influenza, herpes, and Sindbis viruses. *Proceedings of the National Academy of Sciences*, 104(4):1371–1376, January 2007. Publisher: Proceedings of the National Academy of Sciences.

[420] Adrien Decque, Olivier Joffre, Joao G. Magalhaes, Jack-Christophe Cossec, Ronnie Blecher-Gonen, Pierre Lapaquette, Aymeric Silvin, Nicolas Manel, Pierre-Emmanuel Joubert, Jacob-Sebastian Seeler, Matthew L. Albert, Ido Amit, Sebastian Amigorena, and Anne Dejean. Sumoylation coordinates the repression of inflammatory and anti-viral gene-expression programs during innate sensing. *Nature Immunology*, 17(2):140–149, February 2016. Publisher: Nature Publishing Group.

[421] Faten El-Asmi, Francis P. McManus, Carlos Eduardo Brantis-de Carvalho, Jose Carlos Valle-Casuso, Pierre Thibault, and Mounira K. Chelbi-Alix. Cross-talk between SUMOylation and ISGylation in response to interferon. *Cytokine*, 129:155025, May 2020.

[422] Patricia Richard, Vasupradha Vethantham, and James L. Manley. Roles of Sumoylation in mRNA Processing and Metabolism. *Advances in experimental medicine and biology*, 963:15–33, 2017.

[423] Joana M.P. Desterro, Liam P. Keegan, Ellis Jaffray, Ron T. Hay, Mary A. O'Connell, and Maria Carmo-Fonseca. SUMO-1 Modification Alters ADAR1 Editing Activity. *Molecular Biology of the Cell*, 16(11):5115–5126, November 2005.

[424] YuJie Tang, Gongxun Zhong, Lianhui Zhu, Xing Liu, Yufei Shan, Huapeng Feng, ZhiGao Bu, Hualan Chen, and Chen Wang. Herc5 Attenuates Influenza A Virus by Catalyzing ISGylation of Viral NS1 Protein. *The Journal of Immunology*, 184(10):5777–5790, May 2010.

[425] Guoqiang Xu, Jeremy S. Paige, and Samie R. Jaffrey. Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nature Biotechnology*, 28(8):868–873, August 2010. Publisher: Nature Publishing Group.

[426] Marie Lork, Gauthier Lieber, and Benjamin G. Hale. Proteomic Approaches to Dissect Host SUMOylation during Innate Antiviral Immune Responses. *Viruses*, 13(3):528, March 2021.

[427] Vyacheslav Akimov, Inigo Barrio-Hernandez, Sten V. F. Hansen, Philip Hallenborg, Anna-Kathrine Pedersen, Dorte B. Bekker-Jensen, Michele Puglia, Stine D. K. Christensen, Jens T. Vanselow, Mogens M. Nielsen, Irina Kratchmarova, Christian D. Kelstrup, Jesper V. Olsen, and Blagoy Blagoev. UbiSite approach for comprehensive mapping of lysine and N-terminal ubiquitination sites. *Nature Structural & Molecular Biology*, 25(7):631–640, July 2018. Publisher: Nature Publishing Group.

[428] Ivo A. Hendriks, David Lyon, Dan Su, Niels H. Skotte, Jeremy A. Daniel, Lars J. Jensen, and Michael L. Nielsen. Site-specific characterization of endogenous SUMOylation across species and organs. *Nature Communications*, 9(1):2456, June 2018. Publisher: Nature Publishing Group.

[429] Frédéric Lamoliatte, Francis P. McManus, Ghizlane Maarifi, Mounira K. Chelbi-Alix, and Pierre Thibault. Uncovering the SUMOylation and ubiquitylation crosstalk in human cells using sequential peptide immunopurification. *Nature Communications*, 8(1):14109, January 2017. Publisher: Nature Publishing Group.

[430] Erin L. Sternburg, Lara A. Gruijs Da Silva, and Dorothee Dormann. Post-translational modifications on RNA-binding proteins: accelerators, brakes, or passengers in neurodegeneration? *Trends in Biochemical Sciences*, 47(1):6–22, January 2022.

[431] Adam P. Lothrop, Matthew P. Torres, and Stephen M. Fuchs. Deciphering post-translational modification codes. *FEBS Letters*, 587(8):1247–1257, 2013. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1016/j.febslet.2013.01.047.

[432] Paul D. Bieniasz and Sebla B. Kutluay. CLIP-related methodologies and their application to retrovirology. *Retrovirology*, 15:35, May 2018.

[433] Nara Lee, Valerie Le Sage, Adalena V. Nanni, Dan J. Snyder, Vaughn S. Cooper, and Seema S. Lakdawala. Genome-wide analysis of influenza viral RNA and nucleoprotein association. *Nucleic Acids Research*, 45(15):8968–8977, September 2017.

[434] Graham D. Williams, Dana Townsend, Kristine M. Wylie, Preston J. Kim, Gaya K. Amarasinghe, Sebla B. Kutluay, and Adrianus C. M. Boon. Nucleotide resolution mapping of influenza A virus nucleoprotein-RNA interactions reveals RNA features required for replication. *Nature Communications*, 9(1):465, January 2018. Publisher: Nature Publishing Group.

[435] Arthur Wickenhagen, Elena Sugrue, Spyros Lytras, Srikeerthana Kuchi, Marko Noerenberg, Matthew L. Turnbull, Colin Loney, Vanessa Herder, Jay Allan, Innes Jarmson, Natalia Cameron-Ruiz, Margus Varjak, Rute M. Pinto, Jeffrey Y. Lee, Louisa Iselin, Natasha Palmalux, Douglas G. Stewart, Simon Swingler, Edward J. D. Greenwood, Thomas W. M. Crozier, Quan Gu, Emma L. Davies, Sara Clohisey, Bo Wang, Fabio Trindade Maranhão Costa, Monique Freire Santana, Luiz Carlos de Lima Ferreira, Lee Murphy, Angie Fawkes, Alison Meynert, Graeme Grimes, ISARIC4C Investigators, Joao Luiz Da Silva Filho, Matthias Marti, Joseph Hughes, Richard J. Stanton, Eddie C. Y. Wang, Antonia Ho, Ilan Davis, Ruth F. Jarrett, Alfredo Castello, David L. Robertson, Malcolm G. Semple, Peter J. M. Openshaw, Massimo Palmarini, Paul J. Lehner, J. Kenneth Baillie, Suzannah J. Rihn, and Sam J. Wilson. A prenylated dsRNA sensor protects against severe COVID-19. *Science*, 374(6567):eabj3624, September 2021. Publisher: American Association for the Advancement of Science.

[436] Sebla B. Kutluay, Trinity Zang, Daniel Blanco-Melo, Chelsea Powell, David Jannain, Manel Errando, and Paul D. Bieniasz. Global Changes in the RNA Binding Specificity of HIV-1 Gag Regulate Virion Genesis. *Cell*, 159(5):1096–1109, November 2014. Publisher: Elsevier.

[437] Martin Dodel, Giulia Guiducci, Maria Dermit, Sneha Krishnamurthy, Emilie L. Alard, Federica Capraro, Zeinab Rekad, Lovorka Stojic, and Faraz K. Mardakheh. TREX

reveals proteins that bind to specific RNA regions in living cells. *Nature Methods*, 21(3):423–434, March 2024. Publisher: Nature Publishing Group.

[438] Chase A. Weidmann, Anthony M. Mustoe, Parth B. Jariwala, J. Mauro Calabrese, and Kevin M. Weeks. Analysis of RNA–protein networks with RNP-MaP defines functional hubs on RNA. *Nature Biotechnology*, 39(3):347–356, March 2021. Publisher: Nature Publishing Group.

[439] Daniel A. Lorenz, Hsuan-Lin Her, Kylie A. Shen, Katie Rothamel, Kasey R. Hutt, Allan C. Nojadera, Stephanie C. Bruns, Sergei A. Manakov, Brian A. Yee, Karen B. Chapman, and Gene W. Yeo. Multiplexed transcriptome discovery of RNA-binding protein binding sites by antibody-barcode eCLIP. *Nature Methods*, pages 1–5, December 2022. Publisher: Nature Publishing Group.

[440] Christopher S. Hughes, Sophie Moggridge, Torsten Müller, Poul H. Sorensen, Gregg B. Morin, and Jeroen Krijgsveld. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nature Protocols*, 14(1):68–85, January 2019. Number: 1 Publisher: Nature Publishing Group.

[441] Brian J. Zarnegar, Ryan A. Flynn, Ying Shen, Brian T. Do, Howard Y. Chang, and Paul A. Khavari. irCLIP platform for efficient characterization of protein–RNA interactions. *Nature Methods*, 13(6):489–492, June 2016. Number: 6 Publisher: Nature Publishing Group.

[442] Tianzhi Wu, Erqiang Hu, Shuangbin Xu, Meijun Chen, Pingfan Guo, Zehan Dai, Tingze Feng, Lang Zhou, Wenli Tang, Li Zhan, Xiaocong Fu, Shanshan Liu, Xiaochen Bo, and Guangchuang Yu. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*, 2(3), August 2021. Publisher: Elsevier.

[443] Christian von Mering, Lars J. Jensen, Berend Snel, Sean D. Hooper, Markus Krupp, Mathilde Foglierini, Nelly Jouffre, Martijn A. Huynen, and Peer Bork. STRING: known and predicted protein–protein associations, integrated and transferred across organisms. *Nucleic Acids Research*, 33(suppl_1):D433–D437, January 2005.

[444] Charles Girardot, Jelle Scholtalbers, Sajoscha Sauer, Shu-Yi Su, and Eileen E.M. Furlong. Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. *BMC Bioinformatics*, 17(1):419, October 2016.

[445] Marcel Martin. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1):10–12, May 2011. Number: 1.

[446] Alexander Dobin, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1):15–21, January 2013.

[447] Aaron R. Quinlan and Ira M. Hall. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6):841–842, March 2010.

[448] Nikolaos Ignatiadis, Bernd Klaus, Judith B. Zaugg, and Wolfgang Huber. Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. *Nature Methods*, 13(7):577–580, July 2016. Number: 7 Publisher: Nature Publishing Group.

[449] Michael I. Love, Wolfgang Huber, and Simon Anders. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12):550, December 2014.

[450] Ronny Lorenz, Stephan H. Bernhart, Christian Höner zu Siederdissen, Hakim Tafer, Christoph Flamm, Peter F. Stadler, and Ivo L. Hofacker. ViennaRNA Package 2.0. *Algorithms for Molecular Biology*, 6(1):26, November 2011.