

THE STRUCTURE OF IFIT PROTEINS AND THEIR RECOGNITION OF VIRAL RNA

by

Yazan Mahmoud Abbas

Department of Biochemistry

McGill University

Montreal, Quebec, Canada

April, 2017

A thesis submitted to McGill University in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

© Yazan M. Abbas, April 2017

The first principle is that you must not fool yourself - and you are the easiest person to fool.

- Richard P. Feynman

For Mom and Dad

ABSTRACT

Interferons are antiviral cytokines which orchestrate a powerful host response to viral infection by upregulating more than 300 interferon stimulated genes (ISGs). Many of these are effector molecules that altogether mediate diverse antiviral activities and impart cells with resistance to viral infection. Amongst the most potently induced ISGs are the interferon induced proteins with tetratricopeptide repeats (IFITs), a family of innate-immune effectors that in humans comprises 4 well characterized members: IFIT1, IFIT2, IFIT3, and IFIT5. IFITs were initially thought to protect cells by modulating cellular processes through disruptive protein-protein interactions. However, they were recently discovered to have a more direct antiviral role which relies on the recognition of viral 5' RNA structures.

IFIT1 and IFIT5 were recently shown to recognize viral RNA displaying 5'-triphosphates (PPP), a molecular signature of 'non-self' that distinguishes it from host RNA. To gain insight into IFIT antiviral mechanisms, I determined the crystal structure of human IFIT5, its complex with PPP-RNAs, and an N-terminal fragment of human IFIT1. The structures revealed a novel helical fold which forms a narrow, positively-charged tunnel designed to recognize single-stranded PPP-RNAs in a sequence non-specific manner. Solution X-ray analysis and limited protease digestion suggested a role for conformational changes in RNA binding. Gel shift assays showed that IFIT1 and IFIT5 could engage only single-stranded RNA, or base-paired RNA with a 5'-overhang of approximately 3-5 nucleotides. Mutational analysis confirmed that IFIT1 and IFIT5 have a conserved RNA recognition mechanism, and demonstrated that RNA binding is required for optimal antiviral activity in cell culture. Taken together, PPP-RNA recognition potentially allows IFIT1 or IFIT5 to sequester viral genomic or intermediate RNAs to limit virus replication.

I then focused on IFIT1's ability to prevent virus propagation by selectively inhibiting viral protein synthesis. It does so by competing with the translation initiation factor eIF4F for binding to the 5'-end of viral mRNA. Host mRNAs are protected from IFIT1 as they are ribose 2'-O methylated at the first cap proximal nucleotide (N1); a mechanism which viruses often exploit to escape IFIT1 detection. I determined several crystal structures of RNA-bound human IFIT1, which showed that its positively-charged RNA-binding tunnel is distinct from that found in IFIT5, and further extended to allow binding of both capped and uncapped RNAs. IFIT1 uses a relatively plastic and non-specific mode of binding allowing it to recognize multiple forms and

conformations of the cap. The IFIT1 tunnel encircles four cap proximal nucleotides in a manner that excludes host mRNAs bearing N1 methylation. Gel-shift and *in vitro* translation assays confirm that N1 methylation interferes with IFIT1 recognition, but in an RNA-dependent manner. Structural and functional analysis show that 2'-O methylation at N2, another abundant mRNA modification, is also occluded and could potentially synergize with N1 methylation to further protect host mRNAs from aberrant IFIT1 recognition. This work defines the molecular basis for human IFIT1 activity and uncovers additional mechanisms which govern self vs non-self mRNA discernment by IFIT1.

RESUMÉ

Les interférons sont des cytokines antivirales qui provoquent chez l'hôte une réponse immunitaire aux infections virales en régulant à la hausse plus de 300 gènes stimulés par l'interféron (ISGs). Plusieurs de ces gènes sont des effecteurs antiviraux qui confèrent aux cellules une résistance à l'infection virale. Parmi les ISGs les plus hautement induits, on trouve la famille de protéines induites par l'interféron contenant des répétitions tétratricopeptides (IFITs) – un groupe d'effecteurs du système immunitaire inné dont les membres sont IFIT1, IFIT2, IFIT3 et IFIT5. Les effets protecteurs des IFITs étaient initialement attribués à la modulation de processus cellulaires par des interactions protéine-protéine. Récemment, un rôle antiviral direct de cette famille a été découvert, impliquant la reconnaissance de l'extrémité 5' de l'ARN viral.

IFIT1 et IFIT5 ont récemment été démontrés à reconnaître l'ARN viral contenant une extrémité 5'-triphosphate (PPP), un motif moléculaire qui le différencie de l'ARN de l'hôte. Afin de mieux comprendre les mécanismes antiviraux des IFITs, j'ai déterminé la structure cristalline de la protéine IFIT5 humaine, incluant des complexes avec des ARNs PPP, ainsi que celle d'un fragment N-terminal de la protéine IFIT1 humaine. Ces structures révèlent un nouveau repliement hélicoïdal formant un tunnel étroit et chargé positivement dont le rôle est la reconnaissance de l'ARN PPP monocaténaire, sans spécificité pour la séquence. L'analyse aux rayons X en solution et la protéolyse limitée suggèrent des changements de conformation de la protéine suite à la liaison à l'ARN. Des tests de retard sur gel démontrent qu'IFIT1 et IFIT5 ne se lient qu'à l'ARN monocaténaire ou à l'ARN à deux brins pourvu d'un surplomb 5' de trois à cinq nucléotides. Une analyse mutationnelle confirme que le mécanisme de liaison à l'ARN de ces deux protéines est conservé, et que cette liaison est requise pour une activité antivirale optimale en culture cellulaire. Ainsi, la reconnaissance de l'ARN PPP permettrait à IFIT1 ou IFIT5 de séquestrer l'ARN génomique viral, ou d'autres ARNs intermédiaires, afin de limiter la réplication du virus.

J'ai ensuite étudié l'inhibition de la traduction de l'ARN messenger viral par IFIT1, dont l'effet est d'empêcher la propagation du virus. IFIT1 accomplit ceci par la compétition avec le facteur d'initiation de la traduction eIF4F pour la liaison à l'extrémité 5' de l'ARN messenger viral. Les ARNs messagers de l'hôte sont protégés contre IFIT1 par la méthylation 2'-O sur le ribose du premier nucléotide suivant la coiffe (N1). Ce mécanisme est aussi souvent utilisé par des virus pour éviter la détection par IFIT1. J'ai déterminé plusieurs structures de la protéine IFIT1 humaine

liée à l'ARN, révélant que son tunnel de liaison à l'ARN, chargé positivement, est différent de celui d'IFIT5. Dans IFIT1, celui-ci est plus profond et permet la liaison à l'ARN avec ou sans coiffe; la flexibilité du mode de liaison permet la reconnaissance de plusieurs formes et conformations de la coiffe. Le tunnel d'IFIT1 entoure les quatre nucléotides suivant la coiffe d'une façon qui exclut l'entrée des ARNs messagers de l'hôte, porteurs d'une méthylation N1. Des analyses de retard sur gel et de traduction *in vitro* confirment que la méthylation N1 interfère avec la reconnaissance par IFIT1; cet effet dépend cependant de l'ARN en question. Une analyse structurelle et fonctionnelle démontre que la méthylation 2'-O du deuxième nucléotide (N2), une autre modification fréquente des ARNs messagers, empêche aussi l'entrée de ceux-ci, agissant ainsi en synergie avec la méthylation N1 pour protéger les ARNs messagers de l'hôte. En résumé, ces études ont établi les principes de l'action d'IFIT1 chez l'humain et ont découvert des mécanismes additionnels de distinction entre l'ARN de l'hôte et étranger par cette protéine.

TABLE OF CONTENTS

ABSTRACT	vii
RESUMÉ	ix
TABLE OF CONTENTS	xi
LIST OF FIGURES	xvii
LIST OF TABLES	xix
LIST OF ABBREVIATIONS	xx
PREFACE	xxiii
CONTRIBUTION OF AUTHORS	xxiv
ORIGINAL CONTRIBUTIONS TO KNOWLEDGE	xxv
ACKNOWLEDGEMENTS	xxvii
CHAPTER 1: GENERAL INTRODUCTION	29
1.1 Overview of antiviral innate immunity	29
1.1.1 Classification of viruses	29
1.1.1.1 Group I dsDNA viruses and group II ssDNA viruses.	30
1.1.1.2 Group III dsRNA viruses	31
1.1.1.3 Group IV positive-strand ssRNA viruses	31
1.1.1.4 Group V negative-strand ssRNA viruses (NSV).....	31
1.1.1.5 The reverse-transcribing group VI (+) ssRNA and group VII DNA viruses.	32
1.1.2 Detection of viral nucleic acids.....	33
1.1.2.1 Endosomal recognition of viral nucleic acids	33
1.1.2.2 Cytosolic recognition of viral nucleic acids	34

1.1.2.3 PPP-RNA recognition by RIG-I.....	35
1.1.3 Interferons and Interferon Stimulated Genes (ISGs)	37
1.1.4 Protein Kinase R is an ISG that recognizes long dsRNA & single stranded PPP-RNA.	38
1.2 The interferon induced proteins with tetratricopeptide repeats (IFITs).....	40
1.2.1 The IFIT gene family	40
1.2.2 IFIT expression patterns <i>in vitro</i> and <i>in vivo</i>	42
1.2.3 Primary structure of IFIT proteins: The Tetratricopeptide Repeat motif.....	43
1.2.4 IFIT biochemical activity	44
1.2.5 IFIT1 and IFIT5 are RNA binding proteins that recognize viral 5' PPP RNA.....	45
1.3 Eukaryotic and viral mRNA cap synthesis.....	47
1.3.1 A brief note on nomenclature.....	47
1.3.2 mRNA cap recognition in cap-dependent translation initiation.....	48
1.3.3 Eukaryotic mRNA capping mechanisms	49
1.3.4 Eukaryotic mRNA ribose 2'-O methylation at N1 (Cap1) and N2 (Cap2).....	51
1.3.4.1 The Cap1 methyltransferase CMTr1	51
1.3.4.2 The Cap2 methyltransferase CMTr2.....	52
1.3.5 Viral mRNA capping and 2'-O methylation.....	53
1.3.5.1 Retrovirus and DNA virus capping in the nucleus.....	54
1.3.5.2 Vaccinia Virus and Reovirus 'conventional' capping in the cytoplasm	54
1.3.5.3 Flavivirus, coronavirus, and alphavirus capping in the cytoplasm	56
1.3.5.4 <i>Rhabdoviridae</i> -like capping in the cytoplasm.....	57
1.3.5.5 Cap snatching by segmented nsRNA viruses in the nucleus or cytoplasm.....	59
1.3.6 Uncapped viral mRNA 5' ends.....	60
1.3.7 Potential biological roles of mRNA ribose 2'-O methylation	61

1.4 mRNA Cap1 methylation is a molecular signature of self	62
1.4.1 Ribose 2'-O methylation of viral mRNA enhances virulence and evades IFIT activity	62
1.4.2 IFIT1 & IFIT1B selectively inhibit translation of viral Cap0-mRNA	63
1.4.3 Cap1 methylation also interferes with RIG-I Like Receptor signalling	64
1.5 Overview of thesis and objectives	65
CHAPTER 2: STRUCTURAL BASIS FOR VIRAL 5'-PPP-RNA RECOGNITION BY HUMAN IFIT PROTEINS	67
2.1 Abstract.....	67
2.2 Introduction.....	67
2.3 Results	69
2.3.1 Crystal structures of IFIT5 and nIFIT1	69
2.3.2 IFIT5 specifically binds PPP-RNAs	76
2.3.3 PPP-RNA recognition is non-sequence specific	83
2.3.4 PPP-RNA binding involves a conformational change	84
2.3.5 IFIT5 and IFIT1 bind only PPP-ssRNAs	89
2.3.6 Functional validation of PPP-RNA binding to IFITs.....	89
2.4 Discussion	92
2.5 Materials and Methods.....	93
2.5.1 Protein expression and purification.....	93
2.5.2 RNA and <i>in vitro</i> transcription.....	94
2.5.3 Crystallization and structure determination	95
2.5.4 Small Angle X-ray Scattering	96
2.5.5 Gel shift assays.....	96

2.5.6 Limited Proteolysis	97
2.5.7 Mutational analysis and pulldowns	97
2.5.8 Flu polymerase activity (IFIT1 antiviral activity assay)	97
2.5.9 HEK-Flip-In (IFIT5 antiviral activity assay)	98
2.6 Tables	98
2.6 Acknowledgements	100
CONNECTING TEXT	101
CHAPTER 3: STRUCTURE OF HUMAN IFIT1 WITH CAPPED RNA REVEALS ADAPTABLE MRNA BINDING AND MECHANISMS FOR SENSING N1 AND N2 RIBOSE 2'-O METHYLATIONS.....	102
3.1 Abstract.....	102
3.2 Introduction.....	103
3.3 Results	105
3.3.1 RNA binding and inhibition of <i>in vitro</i> translation by human IFIT1	105
3.3.2 Overall structure of full-length human IFIT1.	108
3.3.3 The IFIT1 RNA-binding tunnel houses a functionally distinct cap binding pocket.	110
3.3.4 IFIT1 can non-specifically accommodate multiple forms & conformations of cap.	115
3.3.5 Mutational analysis of cap recognition.	118
3.3.6 The cap binding mechanism is conserved in IFIT1 and IFIT1B proteins.....	120
3.3.7 IFIT1 cap binding is distinct from canonical cap binding proteins.	122
3.3.8 Binding of cap-proximal nucleotides.....	123
3.3.9 IFIT1 senses ribose 2'-O methylation at N1 and N2.....	126
3.3.10 Functional validation of IFIT1 activity against 2'-O mtase deficient human CoV.....	130

3.4 Discussion	131
3.5 Materials and Methods.....	134
3.5.1 Cloning, protein expression, and purification of IFIT proteins.	134
3.5.2 Cloning, protein expression, and purification of RNA modifying enzymes.....	135
3.5.3 Crystallization and data collection.	136
3.5.4 Structure determination, model building, and refinement.	137
3.5.5 Sequence, structure and RNA analysis.	137
3.5.6 Enzymatic preparation and purification of RNA for EMSAs.	138
3.5.7 RNA electromobility gel shift assay (EMSA).	139
3.5.8 Preparation of reporter mRNA.....	139
3.5.9 <i>In vitro</i> translation assay with Krebs extracts.	140
3.5.10 Mammalian constructs, cells, reagents and viruses.	140
3.5.11 Virus infections and determination of virus titers.....	141
3.5.12 Real-time RT-PCR.....	141
3.5.13 Chemical synthesis of PPP-AAAA.....	141
3.5.14 Chemical synthesis of Gppp-AAAA.....	143
3.5.15 Purification of chemically synthesized PPP-AAAA and Gppp-AAAA.	143
3.5.16 Preparation of m7Gppp-AAAA.....	143
3.5.17 Characterization of RNA by LC-MS.	144
3.6 Tables	145
3.7 Acknowledgements	147
CHAPTER 4: GENERAL DISCUSSION	149
4.1 Self vs non-self mRNA discernment by IFIT1	149

4.1.1 mRNA cap recognition by human IFIT1	150
4.1.2 CMTr1, CMTr2, and other enzymes protecting endogenous mRNA from IFIT1	150
4.1.4 Viral evasion of IFIT1.....	152
4.1.4.1 Viral N1 methylation.....	152
4.1.4.2 Viral 5' RNA structure and other modifications	153
4.1.4.3 New perspectives on IFIT1 activity	154
4.1.4.4 The differing roles of IFIT1 and IFIT1B.....	155
4.2 Viral PPP-RNA recognition	156
4.2.1 PPP-RNA recognition by IFIT5	156
4.2.2 PPP-RNA recognition by IFIT1	157
4.2.3 Diverse <i>in vitro</i> nucleic acid binding by IFIT5	158
4.2.3.1 Does IFIT5 specifically bind capped-RNA?	159
4.3 Additional differences between human and mouse IFIT genes.....	160
4.4 Therapeutic potential for targeting IFIT1.....	161
4.4.1 mRNA therapeutics.....	161
4.4.2 Implications for vaccine design	162
REFERENCES.....	164

LIST OF FIGURES

Figure 1.1	Baltimore classification of viruses	30
Figure 1.2	Virally activated signalling pathways and effector activity of ISGs.....	32
Figure 1.3	Relationship between IFIT proteins and primary structure.....	41
Figure 1.4	Working model for IFIT antiviral function on PPP-RNA.....	46
Figure 1.5	Schematic of differentially methylated mRNA cap structures.....	48
Figure 1.6	Eukaryotic and viral capping mechanisms	53
Figure 1.7	Model of mRNA self vs non-self discernment by IFIT1 and IFIT1B.....	64
Figure 2.1	Structural overview of human IFIT5.....	70
Figure 2.2	Sequence alignment coloured by BLOSUM62 conservation score	72
Figure 2.3	Comparison of IFIT5 Subdomain I to other TPR proteins.....	73
Figure 2.4	Crystal structure of nIFIT1 (IFIT1 residues 7-279)	74
Figure 2.5	Comparison of the IFIT5 superhelix to O-linked GlcNac transferase (OGT).....	75
Figure 2.6	Structure of IFIT5 bound to PPP-RNA	77
Figure 2.7	F _o -F _c maps of RNA in the three IFIT5-PPP-RNA co-crystal structures	78
Figure 2.8	Structural alignment of nIFIT1 and PPP-RNA bound IFIT5	79
Figure 2.9	Close-up view of the metal ion-binding site in IFIT5	79
Figure 2.10	Sugar pucker of nucleotides 1 and 2 (N1 and N2)	80
Figure 2.11	Interaction between IFIT5 and PPP-RNA at N3 and N4	82
Figure 2.12	The interaction between IFIT5 and PPP-RNA is non-sequence specific.....	83
Figure 2.13	IFIT5 undergoes a conformational change upon binding PPP-RNA	84
Figure 2.14	Comparison between RNA-free IFIT5 and RNA-bound IFIT5	85
Figure 2.15	Small-angle X-ray scattering data NRA-free IFIT5 and RNA-bound IFIT5	88
Figure 2.16	Functional analyses of IFIT binding to PPP-RNA	90
Figure 2.17	IFIT1 & IFIT5 preferentially bind ssRNA or base-paired RNA with 5' overhangs. 91	
Figure 2.18	Pull-downs between myc-tagged IFITs and RNA-coated beads.....	92
Figure 3.1	RNA binding and inhibition of <i>in vitro</i> translation by human IFIT1	106
Figure 3.2	RNA binding and inhibition of <i>in vitro</i> translation cont'd.....	107
Figure 3.3	Overall structure of monomeric, RNA-bound human IFIT1.....	108
Figure 3.4	Overall structure of monomeric, RNA-bound human IFIT1 cont'd	109

Figure 3.5 IFIT1 mRNA cap binding mechanism	111
Figure 3.6 IFIT1 PPP- and m7Gppp- binding mechanism	114
Figure 3.7 IFIT1 can accommodate multiple forms and conformations of the cap.....	116
Figure 3.8 IFIT1 can accommodate multiple forms and conformations of the cap and mutational analysis of cap recognition.....	118
Figure 3.9 Functional validation of cap recognition.....	119
Figure 3.10 ClustalO sequence alignment of select IFIT genes from humans, mice, & rabbits	121
Figure 3.11 Comparison to the canonical cap binding proteins.....	122
Figure 3.12 RNA binding mechanism at N1-N4	125
Figure 3.13 IFIT1 forms a positively-charged, solvent-exposed RNA binding groove	125
Figure 3.14 Sensing of N1 and N2 ribose 2'-O methylation by IFIT1	127
Figure 3.15 Sensing of N1 and N2 ribose 2'-O methylation by IFIT1 cont'd	130

LIST OF TABLES

Table 2.1 Data collection and refinement statistics IFIT5 and nIFIT1	98
Table 2.2 Data collection and refinement statistics IFIT5 with PPP-RNAs.....	99
Table 2.3 Sequences of DNA templates used to generate the PPP-RNA for crystallization.....	99
Table 2.4 Sequences of RNAs (transcribed and synthetic), used in Gel shifts and pull downs .	100
Table 3.1 Data collection and refinement statistics IFIT1 with capped and uncapped RNAs ...	145
Table 3.2 Sequences of <i>in vitro</i> transcribed RNA, DNA oligos templates for <i>in vitro</i> transcription, and primers for quantitative real-time PCR.....	146
Table 3.3 Summary of Mass Spec data.....	147

LIST OF ABBREVIATIONS

AMP – Adenosine monophosphate
CBC – Cap binding complex
CBP20/CBP80 – Cap binding protein
CDN – Cyclic dinucleotide
CE – Capping enzyme
cGAMP – cyclic GMP-AMP
cGAS – cGAMP synthase
CoV – Coronavirus
CTD – C-terminal Domain
DENV – Dengue virus
Dxo/Dom3Z – Decapping and exoribonuclease protein
eIF – Eukaryotic initiation factor
EMCV – Encephalomyocarditis virus
G-patch – Glycine rich domain
GMP – Guanosine monophosphate
GTase – Guanylyltransferase
hCMTR – Human Cap-specific mRNA (nucleoside-2'-O-)-methyltransferase
HCoV – Human coronavirus
HCV – Hepatitis C virus
HSV-1 – Herpes simplex virus type 1
IFN – Interferon
IFNAR – Interferon α/β receptor
IFNGR – Interferon γ receptor
IFNLR – Interferon λ receptor
IRF – Interferon regulatory factor
ISG – Interferon stimulated gene
ISGF3 – Interferon stimulated gene factor 3
IVTmRNA – *In vitro* transcribed messenger RNA drugs
JAK –Janus kinase

JEV – Japanese encephalitis virus
K_d – Equilibrium dissociation constant
LC-MS/MS – Liquid chromatography–mass spectrometry
LGP2 – Laboratory of genetics and physiology 2
m7G – *N*7-methylguanosine
m7Gppp – *N*7-methylguanosine triphosphate
Mce – Mammalian capping enzyme
MDA5 – Melanoma differentiation-associated gene 5
MEF – Mouse embryonic fibroblasts
MERS – Middle East respiratory syndrome
MHV – Mouse hepatitis virus
MTase – Methyltransferase
NF-κB – Nuclear factor kappa-light-chain-enhancer of activated B cells
NS/nsP/nsp – nonstructural/nonstructural protein
nsNSV – non-segmented negative strand RNA virus
OGT – UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase
PABP – Poly(A) binding protein
PAMP – Pathogen associated molecular pattern
PIC – Preinitiation complex
PKR – Protein kinase, RNA-activated
PP5 – Protein phosphatase 5
PPP – Triphosphate
PRNTase – Polyribonucleotidyl transferase
PRR – Pattern recognition receptor
RdRp – RNA-dependent RNA polymerase
RFM – Rossmann-fold methyltransferase
RIG-I – Retinoic acid-inducible gene I
RLR – RIG-I-like receptor
RNA pol II – RNA polymerase II
RNAi – RNA interference
RNGTT – RNA guanylyl transferase and 5' triphosphatase

RNMT – RNA (guanine-N7-) methyltransferase
rRNA – Ribosomal RNA
RTPase – RNA Triphosphatase
SAM – S-Adenosyl methionine
SARS – Severe acute respiratory syndrome
siRNA – Small interfering RNA
snRNA – Small nuclear RNA
sNSV – Segmented negative strand RNA virus
STAT – Signal transducer and activator of transcription
TLR – Toll-like receptor
TPR – Tetratricopeptide repeat
tRNA – Transfer RNA
UTR – Untranslated region
VACV – Vaccinia virus
VP – Viral protein
VPg – Viral protein, genome linked
VSV – Vesicular stomatitis virus
WNV – West Nile virus
WW – Tryptophan containing domain
XRN2 – 5'-3' exoribonuclease 2
YFV – Yellow fever virus

PREFACE

This is a manuscript-based thesis consisting of two published articles.

Chapter 2

Abbas, Y.M., Pichlmair, A., Górna, M.W., Superti-Furga, G., and Nagar, B. (2013). Structural basis for viral 5'-PPP-RNA recognition by human IFIT proteins. *Nature* **494**, 60-64.

Chapter 3

Abbas, Y.M., Laudenbach, B.T., Martínez-Montero, S., Cencic, R., Habjan, M., Pichlmair, A., Damha, M.J., Pelletier, J., and Nagar, B. (2017). Structure of human IFIT1 with capped RNA reveals adaptable mRNA binding and mechanisms for sensing N1 and N2 ribose 2'-O methylations. *Proc Natl Acad Sci USA* **114**, E2106-E2115.

CONTRIBUTION OF AUTHORS

Chapter 2

The project was conceived by Dr. Giulio Superti-Furga and Dr. Bhushan Nagar. I designed experiments and performed all the work in the manuscript under the supervision of Dr. Nagar except: CLS staff assisted with mail-in X-ray data collection; gel-shifts with 5' modified 7SKas-RNA as well as mutational analysis of IFIT1 and IFIT5 in HEK293 cells were performed by Dr. Maria W. Górna in the lab of Dr. Giulio Superti-Furga (Figures 2.16 B and C, and 2.18); Flu virus and VSV-GFP infectivity assays were performed by Dr. Andreas Pichlmair in the lab of Dr. Giulio Superti-Furga (Figures 2.16 D and E). I prepared a draft of the manuscript which was modified by Dr. Nagar, and all authors contributed to the final version.

Chapter 3

The project was conceived by both myself and Dr. Bhushan Nagar. I designed experiments and performed all the work in the manuscript under the supervision of Dr. Nagar except: CLS staff assisted with mail-in X-ray data collection; Dr. Regina Cencic performed translation assays with assistance from Patrick Senechal in the lab of Dr. Jerry Pelletier (Figures 3.1 C, 3.2 D, 3.8 E, 3.9 B, 3.14 C and E); Dr. Saul Martínez-Montero from the lab of Dr. Masad J. Damha prepared chemically synthesized oligos used for crystallization in the latter stages of the project; Beatrice Laudénbach and Dr. Matthias Habjan from the lab of Dr. Andreas Pichlmair performed virus infectivity assays (Figures 3.14 F and G); and Alexei Gorelik and Katalin Illes from my lab kindly assisted with insect cell expression of human Cap2 methyltransferase. I prepared a draft of the manuscript which was edited in close collaboration with Dr. Nagar.

ORIGINAL CONTRIBUTIONS TO KNOWLEDGE

Chapter 2: Structural basis for viral 5'-PPP-RNA recognition by human IFIT proteins

- Determined the crystal structures of human IFIT5 and an N-terminal fragment of human IFIT1, the first structures of an IFIT protein, which revealed an unexpected arrangement of TPR motifs and a novel protein fold.
- Determined the crystal structure of IFIT5 bound to different 5' triphosphate RNA, uncovering the structural basis for IFIT-RNA interactions and the structural basis for specific triphosphate recognition by IFIT5, thereby validating the recently discovered paradigm for IFIT protein function as effectors which sense viral 5' structure.
- The structures of IFIT5 with different sequences also suggested a mechanism for sequence non-specific RNA recognition, thus implicating IFIT5 in potentially broad virus recognition.
- TPR containing proteins are classical protein-protein interaction motifs, but the structures are the first and (along with IFIT1) the only examples of TPR-RNA co-crystal structures to date.
- Combined structural and biochemical analysis by X-ray crystallography, small-angle X-ray scattering, and limited proteolysis uncovered a role for conformational changes in RNA binding that is potentially important for activity.
- Structure guided binding assays showed that IFIT1 and IFIT5 target RNA with single stranded 5'-ends, distinguishing them from other known cytosolic sensors of viral RNA, which recognize double stranded RNA.
- Structure-guided mutational analysis performed with the help of Dr. Giulio Superti-Furga's group showed that RNA binding is required for the full antiviral activity of IFIT1 and IFIT5 in cell culture.

Chapter 3: Structure of human IFIT1 with capped RNA reveals adaptable mRNA binding and mechanisms for sensing ribose N1 and N2 methylation

- Determined the crystal structure of full-length human IFIT1 bound to capped RNA, showing how IFIT1 uses a potentially unique and relatively plastic cap binding mechanism that can recognize multiple forms and conformations of cap (e.g. unmethylated guanosine cap and adenosine cap).
- Crystal structures with unmethylated guanosine-capped RNA and uncapped triphosphate RNA provided additional insight into IFIT1's multiple 5' specificities.
- The structures revealed the molecular basis for IFIT1 inhibition of capped mRNA translation, showing how IFIT1 forms a tight molecular surface near the ribose 2'-hydroxyls at N1 and N2, thereby restricting IFIT1 activity to foreign mRNA not methylated at these two positions.
- Discovered a role for ribose methylation at N2 in protecting endogenous mRNA from IFIT1 recognition, and demonstrated potential synergy between ribose N1 and N2 methylations in preventing IFIT1 recognition of self mRNAs.

ACKNOWLEDGEMENTS

First and foremost, I wish to thank Bhushan Nagar, for his guidance and support, for always keeping an open door for questions and an open mind for all scientific discourse, and for the opportunity to work on such an outstanding project. This opportunity would not have been possible without Dr. Giulio Superti-Furga's collaboration with our group, for which I am indebted. I would like to thank my research advisory committee members, Dr. Albert Berghuis, Dr. Masad Damha, and Dr. Jerry Pelletier, for their support, advice and direct contributions to my research.

This thesis would also not have been possible if not for the discoveries made by Dr. Andreas Pichlmair during his time in Giulio's lab - I can only ever hope to continue standing on the shoulders of such giants. I am also grateful for the contributions made by Andreas and Dr. Maria Góna to my research, and for their continued collaboration. From Dr. Pelletier's lab, Dr. Regina Cencic and Patrick Senechal assisted with crucial experiments; from Dr. Damha's lab, Dr. Saul Martínez-Montero generated essential reagents; and from Andreas's lab, Beatrice Laudenschlag and Dr. Matthias Habjan performed experiments for Chapter 3. At the Canadian Macromolecular Crystallography Facility, Shaun Labiuk and the rest of the CMCF staff have been incredibly helpful over the years with data collection at the CLS. I am grateful for financial support from the CIHR Strategic Training Initiative in Chemical Biology, the NSERC CREATE Training Program in Bionanomachines, Groupe de Recherche Axé sur la Structure des Protéines (GRASP), and McGill University. I also thank members of GRASP for infrastructure support that made much of this work possible.

I am fortunate to have worked alongside wonderful colleagues during my extended tenure in the Nagar lab, and I wish to thank all lab members, past and present, for making the lab a fun and pleasant work place. A few people deserve special mention. First off, Lama Talje has been a friend, classmate, colleague, and sister throughout my entire time in Montreal. I am grateful to her for encouraging me to apply to Bhushan's lab, and for showing me the ropes early on. I am also eternally grateful to Rose Szittner, for being such a warm and kind-hearted colleague, for helping me with experiments early on, for helping keep the Nagar lab running, and for being generous with sharing home-cooked meals throughout the years. Ahmad Kanaan, for his constant energy and sense of wonder, and for being a 'brother-in-arms' in science, music, photography, football, movies, and much more. Filipp Frank, for his camaraderie, and for setting such a high standard of

scientific accomplishment, giving us all something to shoot for. Katalin Illes, for her infectious enthusiasm and constant support over the years. Alexei Gorelik, for being such a model colleague and for his wonderful sense of humor over the years; Alexei also translated the abstract of this thesis. Geneviève Virgili, for being such a wonderful friend and scientist - you started out as the little sister of the lab, but grew to be our big sister. Jonathan Labriola, for being an incredible friend and brother inside and outside the lab. Ahmed Gebai, for taking sarcasm to new levels of fun. I was also lucky enough to mentor many wonderful undergrads over the years. I am especially thankful for the contributions of Irene Xie, Jad Belle, Poom Pratumsuwan, Mirandi Shi, Ying Chou, and Steffanie Neun to this thesis and other side projects.

To my dear friends inside and outside Bellini – Ghaleb, Janice, Juliana, Shane, Ali, Anna, Kim and Jesse, – thank you for all the great science and non-science memories, for making the grad school experience that much better, and for being such an excellent source of support that helped me get through some tough times at grad school. To my roommate for much of grad school, David Wang, thanks for all the fantastic science, nerd, and whiskey times, and for proofreading parts of this thesis.

Most importantly, I thank my family – Zahra Ghosn, Mahmoud Abbas, and Nabil Abbas, for their continued love and support, and for the sacrifices they made which helped me pursue a career in research and science.

CHAPTER 1: GENERAL INTRODUCTION

1.1 Overview of antiviral innate immunity

Our immune system can be broadly categorized into innate and adaptive immunity. The adaptive immune system is tailored towards specific pathogens, and results in antigen-specific and long-lasting protective immunity. However, this high specificity comes at the cost of a delayed response to infection, which can take up to one or two weeks to mount, during which the host relies on the innate immune system to fight off pathogens. Not only does the innate immune system constitute our primary form of defense during the early phase of infection, but its proper activation is also required to mount an appropriate adaptive response (Fensterl et al., 2015). In this way, the innate immune system shapes the overall response to infection.

1.1.1 Classification of viruses

A recent study estimated that there exists at least 320,000 different virus species which infect mammals awaiting to be discovered (Anthony et al., 2013). Despite this apparent staggering diversity, viruses can still be classified in an hierarchical manner (i.e. by order, family, genus, and species) based only on several simple criteria, starting with the nature of their nucleic acid genome, followed by capsid symmetry, then presence or absence of an envelope, and finally particle dimensions. An elegant way to group these viruses and further simplify the viral universe is to classify them according to the Baltimore scheme (Baltimore, 1971), in which viruses are grouped based on the type of nucleic acid genome they carry, and the manner in which information flows between genome and messenger RNA (mRNA) (**Fig. 1.1**). In this way, the Baltimore scheme mirrors the central dogma of molecular biology.

Under the Baltimore system, virus families are grouped into 7 distinct classes, which consider the genome building block (DNA or RNA), genome structure (single stranded (ss) or double stranded (ds)), and genome polarity with respect to the mRNA coding sequence (positive sense (+) or negative sense (-)). The 7 classes of viruses are described below, along with some details on replication for specific virus families relevant to this study. Although this information is by no means exhaustive, it is intended as a reference for future readers and to better appreciate the relationship between virus families. Most of the information below is adapted from (Flint et al.).

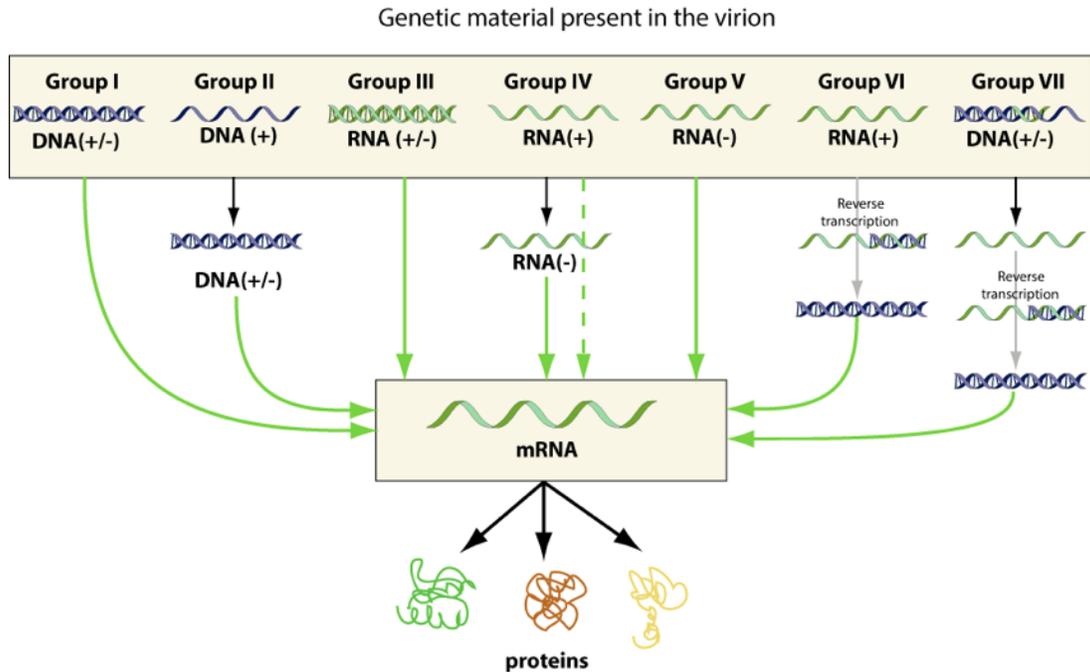


Figure 1.1 Baltimore classification of viruses

Relationship between virus genetic material in virions and viral mRNA. Adapted from (Hulo et al., 2010).

A few commonalities to bear in mind concerning viruses and their replication. All viruses generally proceed through the same basic steps to establish productive infection: attachment, entry, uncoating, transcription, mRNA translation, genome replication, assembly, and release. As is evident from the Baltimore scheme, mRNA is central to viruses where they are completely reliant on the host for translation (Li et al., 2015). Throughout the virus lifecycle, viral nucleic acids are rarely naked (Schlee and Hartmann, 2016). For ssRNA viruses, the (+) and (-) strands are coated with viral nucleocapsid proteins forming ribonucleoprotein structures. Finally, to increase efficiency of replication and assembly, and to shield viruses from host antiviral defenses, many induce the formation of ‘viral factories’ by reorganizing cellular membranes and the host cytoskeleton (Netherton and Wileman, 2011).

1.1.1.1 Group I dsDNA viruses and group II ssDNA viruses.

Prominent examples of Group I viruses include *Herpesviridae*, *Papillomaviridae*, and *Poxviridae*. With few exceptions (e.g. *Poxviridae*), group I and II DNA viruses replicate in the nucleus and are heavily reliant on the host for DNA and RNA synthesis. *Poxviridae* such as vaccinia virus and variola virus (causative agent of smallpox) replicate in the cytoplasm and encode their own

polymerases for DNA and mRNA synthesis. Vaccinia virus encodes nearly 185 genes, and is therefore minimally reliant on the host.

1.1.1.2 Group III dsRNA viruses

This group includes *Reoviridae* (such as mammalian orthoreovirus, bluetongue virus, and rotavirus). Genomes of *Reoviridae* viruses contain 10-12 segments that encode nearly as many proteins. The dsRNA segments are encapsidated within a viral core and each is bound to a viral polymerase. The (-) strand is used for mRNA synthesis, which in turn is used for more (-) strand synthesis.

1.1.1.3 Group IV positive-strand ssRNA viruses

Notable members include *Coronaviridae*, *Picornaviridae*, *Togaviridae*, and *Flaviviridae*. The genomes of these viruses are functional mRNA molecules that are translated directly upon entry, producing polyproteins that require proteolytic cleavage. For *Togaviridae*, an additional subgenomic mRNA is produced which encodes another polyprotein, while *Coronaviridae* viruses produce multiple subgenomic mRNAs. The term subgenomic mRNA is used to distinguish these molecules from full-length, genomic RNA/mRNA. The (+) ssRNA viruses need to synthesize (-) strand templates before more mRNA or genomic (+) RNA can be produced.

1.1.1.4 Group V negative-strand ssRNA viruses (NSV)

This group includes the segmented NSVs (sNSV) *Arenaviridae*, *Bunyaviridae*, and *Orthomyxoviridae*, as well as the non-segmented NSVs (nsNSV) *Bornaviridae*, *Rhabdoviridae*, *Filoviridae*, and *Paramyxoviridae*. Replication of *Arenaviridae* and *Bunyaviridae* is cytosolic, while replication of *Orthomyxoviridae* is nuclear. *Orthomyxoviridae* such as influenza contain 8 (-) strand segments, each associated with viral RNA polymerase. These are copied into mRNA in the nucleus; for replication, (-) strands are also copied into (+) strand in the nucleus, which in turn are used as templates for synthesis of more genomic (-) strands.

The nsNSVs are all part of a larger order known as *Mononegavirales*, for which vesicular stomatitis virus (VSV) is the prototype. The VSV genome is made up of a single (-) strand RNA that is associated with viral RNA polymerase; this is used as a template for synthesizing 5 subgenomic mRNAs. During replication, the (-) strand is used to synthesize a (+) strand copy, which is then used as a template for more genomic (-) strand synthesis.

1.1.1.5 The reverse-transcribing group VI (+) ssRNA and group VII DNA viruses.

These viruses are nuclear or nucleocytoplasmic. They rely on the host for RNA synthesis in the nucleus, but encode their own DNA polymerase (reverse-transcriptase). For ssRNA retroviruses, their genomes are copied into linear dsDNA in the cytosol by the reverse transcriptase before entering the nucleus.

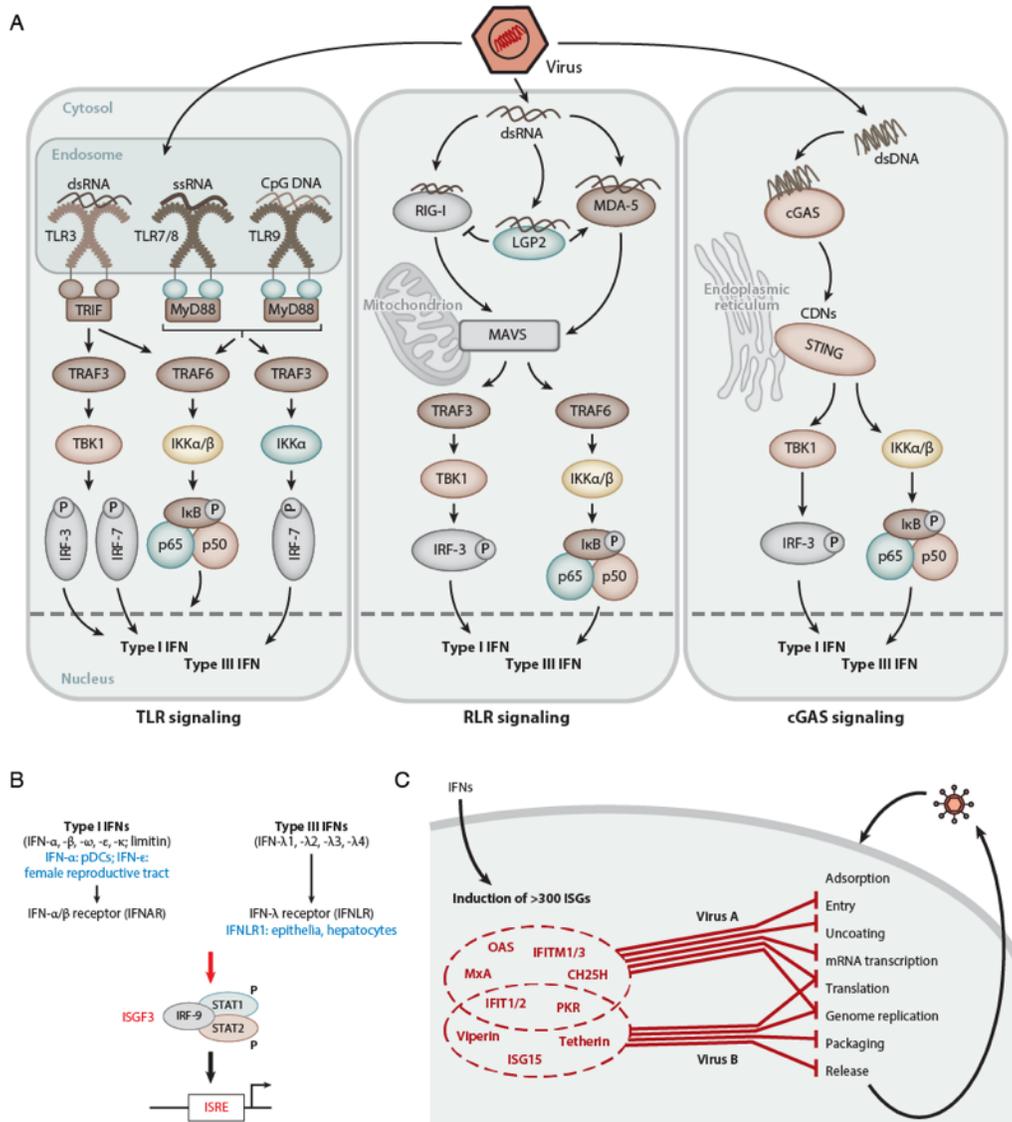


Figure 1.2 Virally activated signalling pathways and effector activity of IFN stimulated genes (ISGs)

(A) TLR, RLR, and cGAS signaling in response to viral PAMP detection. (B) Type I and type III IFNs activate IFNAR and IFNLR triggering ISGF3 activation and synthesis of more than 300 ISGs. (C) Many ISGs have effector activity that directly block viral replication. Individual ISGs usually target one step in the viral lifecycle, and groups of ISGs work in concert to prevent viral replication. Specific ISGs usually target specific virus families, while some ISGs (e.g. IFIT1/2 and PKR) can target multiple virus families. Shown are some ISG examples and the viral replication steps they interfere with. Figure adapted from (Fensterl et al., 2015).

1.1.2 Detection of viral nucleic acids

The proper activation of innate antiviral defenses is dependent on intricate mechanisms of self vs non-self discernment which target nucleic acids, since RNA and DNA are a broadly conserved feature of viral infection. To distinguish between host and viral nucleic acids, the host relies on a limited set of germline-encoded pattern-recognition receptors (PRRs) that recognize motifs not commonly found in the host (Schlee and Hartmann, 2016); these are known as pathogen-associated molecular patterns (PAMPs). To assist in PAMP recognition, the innate immune system also integrates specific localization of viral nucleic acids into its program (Schlee and Hartmann, 2016). PAMP detection by PRRs triggers interferon (IFN) secretion which in turn activates pathways that culminate with the induction of more than 300 host effector molecules known as IFN stimulated genes (ISGs). Altogether, ISGs mediate diverse effects to promote an antiviral state in cells of the body (Fensterl et al., 2015).

1.1.2.1 Endosomal recognition of viral nucleic acids

Toll-like receptors (TLRs) are transmembrane glycoproteins that recognize a wide array of PAMPs indicative of microbial infection (Barbalat et al., 2011). Of the 10 TLRs in humans, four are involved in sensing viral nucleic acids, TLR3, TLR7, TLR8, and TLR9; these are usually expressed in cells of the immune system (Barbalat et al., 2011). Although TLR3 can be found on the surface of fibroblasts (Matsumoto et al., 2003), TLRs are usually localized to endosomes and lysosomes where they monitor the lumen of these compartments for virus derived nucleic acids (**Fig. 1.2 A**) (Barbalat et al., 2011). The strategic placement of TLRs in these vesicles allows them to detect viruses which gain entry through endocytosis. Importantly, TLRs exist as monomers that are activated through PAMP binding to their leucine-rich-repeat (LRR) containing extracellular domains, which causes dimerization and signal transduction via their cytosolic domains (Botos et al., 2011). The critical transcription factors activated by TLRs are NF- κ B, IRF3 and IRF7 (interferon regulatory factors) leading to IFN induction. IRF3, IRF7, and other IRFs recognize interferon stimulated response elements (ISRE) in gene promoters (Fensterl et al., 2015).

Double stranded RNA is a hallmark of viral infection, as it is not commonly found in the host but is associated with dsRNA viruses, and is a by-product of replication of many ssRNA and DNA viruses (Barbalat et al., 2011). dsRNA PAMPs are detected by TLR3 in a sequence non-specific manner, but a minimum length of ~ 45 base pairs is required for receptor dimerization and

activation (Liu et al., 2008); TLR3 also recognizes stem loop structures within viral ssRNA (Tatematsu et al., 2013). Its expression in macrophages and some dendritic cell subsets suggest that, in addition to incoming virus, TLR3 also detects viral infection through phagocytosis of infected, apoptotic cells (Schulz et al., 2005).

TLR7 and TLR8 mediate potential sequence specific PAMP binding by recognizing uridine and guanine rich ssRNA (Barbalat et al., 2011). Crystal structures of TLR8 revealed that it harbors two ligand binding sites, one for uridine and one for UG/UUG oligonucleotides, that bind and activate TLR8 in a synergistic manner, suggesting that TLR8 recognizes degradation products of viral ssRNA (Tanji et al., 2015). Similarly, structural analysis of TLR7 revealed two distinct binding sites, one for guanosine and one U-rich ssRNA, that cooperate to activate TLR7 (Zhang et al., 2016). In contrast to TLR7 and TLR8, TLR9 detects viral DNA through recognition of unmethylated CpG dideoxynucleotide motifs (Ohto et al., 2015); TLR9 can also recognize RNA:DNA hybrids which are indicative of retrovirus replication (Rigby et al., 2014).

1.1.2.2 Cytosolic recognition of viral nucleic acids

In contrast to TLRs, cytosolic nucleic acid sensors are more widely expressed in different cell types (Barbalat et al., 2011). Cytosolic DNA is a potent activator of IFN signalling, and its presence not only suggests DNA virus infection but is also a ‘danger’ signal indicative of nuclear damage. Multiple studies have recently uncovered a wide range of cytosolic DNA sensors, chief amongst them is cGAS (cyclic-GMP-AMP synthase) (**Fig. 1.2 A**) (Schlee and Hartmann, 2016). Binding of dsDNA by cGAS results in the synthesis of a cyclic dinucleotide (CDN) containing a unique, 2′-5′ and 3′-5′ linked cyclic-GMP-AMP (cGAMP), which functions as a second messenger activating downstream signalling leading to IFN induction via IRF3 and NF-κB (Chen et al., 2016). Similar to TLR9, cGAS is also activated by RNA:DNA hybrids (Mankan et al., 2014).

The primary RNA sensors in the cytosol belong to the RIG-I-like receptor (RLR) family, composed of RIG-I, MDA5, and LGP2 (**Fig. 1.2 A**). RIG-I and MDA5 were both initially discovered as dsRNA sensors that promote IFN production via activation of IRF3 and NF-κB (Kato et al., 2006; Yoneyama et al., 2004). Together they mediate an antiviral response to a wide array of ssRNA and dsRNA virus families (Kato et al., 2006). RLRs belong to the DExD/H-box RNA helicase family, and have similar domain architecture, made up of two N-terminal CARD domains, a helicase domain, and a C-terminal regulatory domain (CTD). The helicase and CTD

are involved in RNA binding, and RLRs undergo conformational changes upon PAMP recognition to promote downstream signalling via their CARDs domains (Bin Wu and Hur, 2015). LGP2 lacks the CARD domains and instead modulates RLR responses: it negatively regulates RIG-I signaling but promotes MDA5 activation by enhancing MDA5-RNA interactions (Bruns et al., 2014; Schlee and Hartmann, 2016; Uchikawa et al., 2016). Whereas MDA5 recognizes long dsRNA with potentially higher order structure (Kato et al., 2008; Pichlmair et al., 2009), RIG-I prefers short, blunt-end dsRNA displaying 5' triphosphates (PPP) (Schlee et al., 2009).

1.1.2.3 PPP-RNA recognition by RIG-I

The recognition of viral 5' ends is a common theme in cytosolic RNA sensing, as it provides a powerful means to distinguish between host and virus RNA. Nascent RNAs synthesized in the host nucleus all bear a 5' triphosphate since the transcription process generally initiates with an NTP. With few exceptions, all of these nuclear transcripts undergo maturation or processing steps that result in 5' monophosphate RNA (on rRNA and tRNA), or *N*⁷-methylguanosine caps (on mRNA and snRNA) (Hornung et al., 2006). In contrast, most negative-strand RNA viruses retain a 5' triphosphate on their genomic RNA, and all RNA viruses in general produce transient transcriptional or replicative intermediates that also harbor a 5' triphosphate (Hornung et al., 2006). Thus, cytosolic PPP-RNA is a marker of viral infection.

The discovery of PPP-RNA as a virus-associated PAMP, and its recognition by RIG-I, was rather serendipitous. Soon after small interfering RNAs (siRNAs) became a wide-spread tool in molecular biology, for their ability to mediate highly sequence-specific, post-transcriptional gene silencing through the RNA interference (RNAi) pathway (Elbashir et al., 2001), researchers were attempting to harness their therapeutic potential. An emerging idea at the time was the application of RNAi to specifically knock-down viral genes, thereby opening the door for its use in treating viral diseases (Kapadia et al., 2003; McCaffrey et al., 2003). However, it was becoming apparent that these siRNAs could also mediate sequence-independent anti-viral effects in the form of a potent IFN response (Karikó et al., 2004; Sledz et al., 2003). One landmark study observed that siRNAs prepared *in vitro* using bacteriophage T7 RNA polymerase could block viral infection in a sequence-independent manner; in contrast, the chemically synthesized siRNAs were less effective (Kim et al., 2004). A key feature of their siRNAs prepared using T7 and other bacteriophage RNA polymerases is the 5' triphosphate, and its removal resulted in a complete loss of IFN induction. This was the first study to demonstrate a role for PPP containing RNA in eliciting

an IFN response, although the IFN-stimulating PPP-sensor was not identified at the time (Kim et al., 2004).

Soon afterwards, several groups independently showed that RIG-I was the PPP-RNA receptor (Hornung et al., 2006; Pichlmair et al., 2006; Plumet et al., 2007). Even though RIG-I sensing of dsRNA was well-characterized at the time (Marques et al., 2006), PPP-RNA sensing by RIG-I was initially thought to be an independent mechanism and ssRNA specific, as *in vitro* T7-transcribed ssRNAs were potent activators of the RIG-I pathway (Cui et al., 2008; Hornung et al., 2006; Pichlmair et al., 2006; Takahasi et al., 2008). However, it turned out that the *in vitro* transcription process generated unintended hairpin-containing, single-stranded by-products with complementary 5' and 3' ends; these PPP-RNA containing contaminants were responsible for RIG-I activation (Schmidt et al., 2009). Thus, RIG-I integrates multiple signatures into its sensing mechanism: a 5' PPP and base-pairing at the 5' end, to distinguish between host and viral RNA. The optimal RIG-I ligand is defined as short dsRNA with at least 10-20 bps, a 5' triphosphate, and a blunt-end at the same side of the PPP; single-stranded hairpin RNA displaying similar features are also potent activators (Luo et al., 2012; Schlee et al., 2009). The PPP-end engages the C-terminal regulatory domain of RIG-I, which also discriminates between blunt-ended or overhang containing dsRNA (Cui et al., 2008; Lu et al., 2010; Takahasi et al., 2008; Wang et al., 2010), while the helicase domain binds the dsRNA region (Civril et al., 2011; Jiang et al., 2011; Kowalinski et al., 2011; Luo et al., 2011). The helicase domain couples its ATPase activity to self vs non-self discrimination and downstream signaling (Lässig et al., 2015; Peisley et al., 2013), which is stimulated by PPP binding to the C-terminal domain (Myong et al., 2009; Takahasi et al., 2008). Synergy between the two domains is therefore required to relieve RIG-I from its auto-inhibited conformation and expose the CARDS for downstream signalling.

This optimal RIG-I activating motif can be displayed by viruses *in vivo* in several ways (Luecke and Paludan, 2016; Schlee and Hartmann, 2016). First, segmented negative-strand ssRNA viruses like Influenza and arenaviruses contain single stranded genomes with partially complementary 5' and 3' ends; when bound to nucleocapsid protein, these genomes are circularized to form a so-called 'panhandle' structure (Moeller et al., 2012), which contains a base-paired blunt-end and a 5' triphosphate that has been shown to activate RIG-I (Rehwinkel et al., 2010; Weber et al., 2013). Second, many non-segmented negative strand ssRNA viruses generate defective genome particles that are by-products of replication, which contain a PPP and perfectly

complementary ends; these also form panhandle structures with long stretches of dsRNA regions and are therefore highly immunostimulatory via RIG-I (Strahle et al., 2006; Sun et al., 2015). Third, the genomes of dsRNA viruses have two distinct 5' ends on each strand, one that is capped and unlikely to trigger RIG-I, while the complementary strand contains a 5' diphosphate (Banerjee, 1980); this double stranded PP-RNA end can also trigger RIG-I signalling (Goubau et al., 2014). Finally, perfectly complementary double-stranded RNA with 5' triphosphates are transiently formed by all RNA viruses during replication or transcription (Hornung et al., 2006).

The importance of the blunt-ended 5' triphosphate for RIG-I detection is underscored by viral evasion strategies that attempt to mask it. *Picornaviridae* (positive-strand ssRNA viruses) use a protein linked primer to initiate genome replication, thereby masking transiently formed PPP ends (Goodfellow, 2011); pathogenic viruses such as Borna disease virus, Hantaan virus and Crimean-Congo hemorrhagic fever virus process their genomic 5' triphosphate into a 5' monophosphate to escape RIG-I detection (Habjan et al., 2008); and arenaviruses, which form circular panhandle genome segments as described above, produce a single nucleotide 5' overhang that also escapes RIG-I detection (Marq et al., 2010).

1.1.3 Interferons and Interferon Stimulated Genes (ISGs)

Interferons (IFN) were initially discovered in 1957 by Isaacs and Lindenmann as factors which are produced and secreted by cells in response to viral challenge, and when added to uninfected cells, have the property to interfere with subsequent viral infection (Isaacs and Lindenmann, 2015). IFNs therefore act upon cells to promote an antiviral state (Fensterl et al., 2015), and are essential in this process as the ablation of major IFN receptors in mice renders them susceptible to viral infection (Muller et al., 1994).

Vertebrates encode a large repertoire of IFN genes which are classified into three main types based on the cell-surface receptors and signalling pathways they trigger (reviewed in (Fensterl et al., 2015)). The Type I IFNs are the largest group and include, among others, IFN- β and more than 10 subtypes of IFN- α , all of which engage the ubiquitously expressed IFN- α/β receptor (IFNAR) (**Fig. 1.2 B**). Through IFNAR, type I IFNs activate the JAK-STAT pathway resulting in the formation of an activated ISGF3 complex (IFN-stimulated gene factor 3) which is composed of phosphorylated STAT1, phosphorylated STAT2, and the transcription factor IRF9. Type III IFNs comprise 4 subtypes of IFN- λ which bind to the IFN- λ receptor (IFNLR); this also

activates ISGF3 by signalling through the same JAK-STAT pathway as IFNAR (**Fig. 1.2 B**). However, unlike IFNAR, IFNLR expression is restricted to epithelial cells and hepatocytes. IFN- γ is the only type II IFN, and is produced by cells of the immune system (e.g. T and B cells); it signals via the IFN- γ receptor (IFNGR) resulting in a phosphorylated STAT1 homodimer known as the gamma-activated factor (GAF).

Activated ISGF3 and GAF translocate to the nucleus where they act on targets by recognizing ISREs and gamma-activated sites (GAS), respectively, in the promoter region of many target ISGs (Fensterl et al., 2015). This results in the transcriptional induction of > 300 ISGs which are responsible for mediating the pleiotropic effects of IFNs (Der et al., 1998). Functional assortment of ISGs reveals their involvement in a multitude of cellular processes such as apoptosis, antigen presentation and processing, and pro-inflammatory signalling, thereby explaining IFNs ability to exert anti-proliferative effects as well as to stimulate innate immune cells and help activate the adaptive immune system (de Veer et al., 2001; Fensterl et al., 2015). Interestingly, many of the constitutively expressed molecules involved in sensing viral infection (e.g. RLRs, TLRs, and cGAS) as well as those which promote antiviral signalling (e.g. IRFs, STAT1 and STAT2) are interferon inducible as well, and are therefore implicated in a positive feedback loop (Fensterl et al., 2015). One of the largest group of ISGs encodes protein products that exert cell-intrinsic antiviral activity through a direct effect on viral replication (de Veer et al., 2001). These ISGs can block viral replication by targeting one or more key steps in the virus lifecycle, for e.g. translation, genome replication, or viral egress (Fensterl et al., 2015). However, most ISGs encode weak antiviral activity when expressed on their own, and thus a proper antiviral response requires the concerted effect of multiple ISGs to simultaneously target multiple steps in a virus' lifecycle (**Fig. 1.2 C**) (Fensterl et al., 2015; Schoggins et al., 2011). Indeed, co-expression of multiple ISGs results in additive antiviral effects (Karki et al., 2012; Schoggins et al., 2011).

1.1.4 Protein Kinase R is an ISG that recognizes long dsRNA and single stranded PPP-RNA

One of the more well characterized ISG mechanisms is that of protein kinase R (PKR), which is composed of two N-terminal dsRNA-binding-motifs (dsRBMs) and a C-terminal kinase domain (Hull and Bevilacqua, 2016). When uninfected cells are primed with IFN, PKR is upregulated but remains in a latent, inactive state. Following infection, the presence of viral PAMPS such as long dsRNA is recognized by the dsRBM domain, resulting in PKR dimerization and subsequent

activation via autophosphorylation. Activated PKR phosphorylates the eukaryotic initiation factor 2 α (eIF2 α), resulting in general inhibition of translation initiation, thereby halting the spread of virus. Importantly, PKR activation requires a marker of viral infection, such as dsRNA with a minimum length of \sim 33 bp, for dimerization and activation (Hull and Bevilacqua, 2016). In this way, uninfected cells are protected from any detrimental effects of PKR, as its activity is limited in the absence of virus. Many other ISGs also target viral PAMPs as part of their antiviral program (Schlee and Hartmann, 2016).

Interestingly, while searching for novel activators of PKR, Bevilacqua and colleagues serendipitously discovered that PKR could also specifically recognize 5' triphosphate RNA (Nallagatla et al., 2007; Zheng and Bevilacqua, 2004). Through *in vitro* selection of aptamers from a large random pool of RNA, a novel motif was identified which was composed of an imperfect hairpin \sim 16 bp in length and flanked by 10-15 nt single stranded tails at either the 5' end, 3' end, or both (Zheng and Bevilacqua, 2004). However, as these RNAs were prepared using T7 RNA polymerase, PKR activation by the ssRNA turned out to be 5' triphosphate dependent (Nallagatla et al., 2007). Indeed, activation of PKR by the hairpin-containing ssRNA was abolished when the 5' PPP was replaced by a hydroxyl, monophosphate, diphosphate, or *N*7-methylguanosine cap. In contrast, a 5' PPP was not required for PKR activation by the canonical PKR ligand, long dsRNA, suggesting that PPP recognition of ssRNA constitutes a distinct mechanism from dsRNA recognition.

Further investigation showed that 47 and 110 nt PPP-ssRNA with limited secondary structure (e.g. 5 bp stem-loop optimally placed 21-46 nucleotides away from the 5' end), could still activate PKR (Nallagatla et al., 2007). PPP-ssRNA dependent activation appears to be sequence non-specific, but shows a small degree of selectivity for the 5' nucleotide (Toroney et al., 2012). The structural basis for PPP recognition is unknown, but this mechanism is thought to underlie PKR activation by Influenza B virus ribonucleoprotein (Dauber et al., 2009).

1.2 The interferon induced proteins with tetratricopeptide repeats (IFITs)

Owing to its rapid induction and high expression levels following IFN treatment of various human cell lines, mRNA 561 was one of the first interferon inducible mRNAs to be cloned and characterized (Chebath et al., 1983; Kusari and Sen, 1986; Wathelet et al., 1986). The corresponding gene, which encodes a 56 kDa protein product, would later be known as IFI-56, ISG56, and IFIT1, thus becoming the founding member of the IFIT gene family. The mRNA encoding a related interferon-inducible gene, human IFIT2, was identified and characterized soon afterwards (Levy et al., 1986; Ulker and Samuel, 1987; Ulker et al., 1987; Wathelet et al., 1988). It wasn't until 1994 that the homologous interferon-inducible genes in mice were isolated (Bluyssen et al., 1994), and genes for human IFIT3, human IFIT5, and mouse IFIT3 were also identified and characterized around the same time (de Veer et al., 1998; Lee et al., 1994; Niikura et al., 1997; Smith and Herschman, 1996; Yu et al., 1997). An IFIT1-like gene, IFIT1B, was also identified in humans, but it remains poorly characterized (Wathelet et al., 1988). IFIT protein products are all cytosolic, ~ 55 kDa in size, and in humans and mice, share 34-61% sequence identity at the amino acid level (**Fig. 1.3 A**). They are all characterized by the presence of multiple tandem copies of the tetratricopeptide repeat motif, a helix-turn-helix motif implicated in mediating protein-protein interactions (**Fig. 1.3 B**). Despite being among the most potently induced ISGs and the subject of numerous studies since their identification between 1983 and 1994, the molecular mechanisms underlying IFIT antiviral activity are only just being unravelled.

1.2.1 The IFIT gene family

IFIT genes originated in an ancestral jawed vertebrate approximately 450 million years, and have been co-evolving with viruses and the IFN system since (Daugherty et al., 2016; Liu et al., 2013). Due to the unique viral pressures encountered by the different vertebrates, the IFIT gene family has undergone lineage-specific expansion and contractions (i.e. gene duplication and deletions), resulting in a distinct complement of IFITs among each vertebrate class (Liu et al., 2013). For instance, whereas zebrafish encode 10 IFIT-like genes, the genomes of most placental mammals, including that of humans, contain five: IFIT1, IFIT1B, IFIT2, IFIT3, and IFIT5 (Daugherty et al., 2016). Many mammals also possess species-specific deletions and duplications (e.g. mice, described below), and thus the number and composition of IFITs varies from mammal to mammal. Based on their phylogeny, the five IFITs likely arose from successive gene duplications of a single,

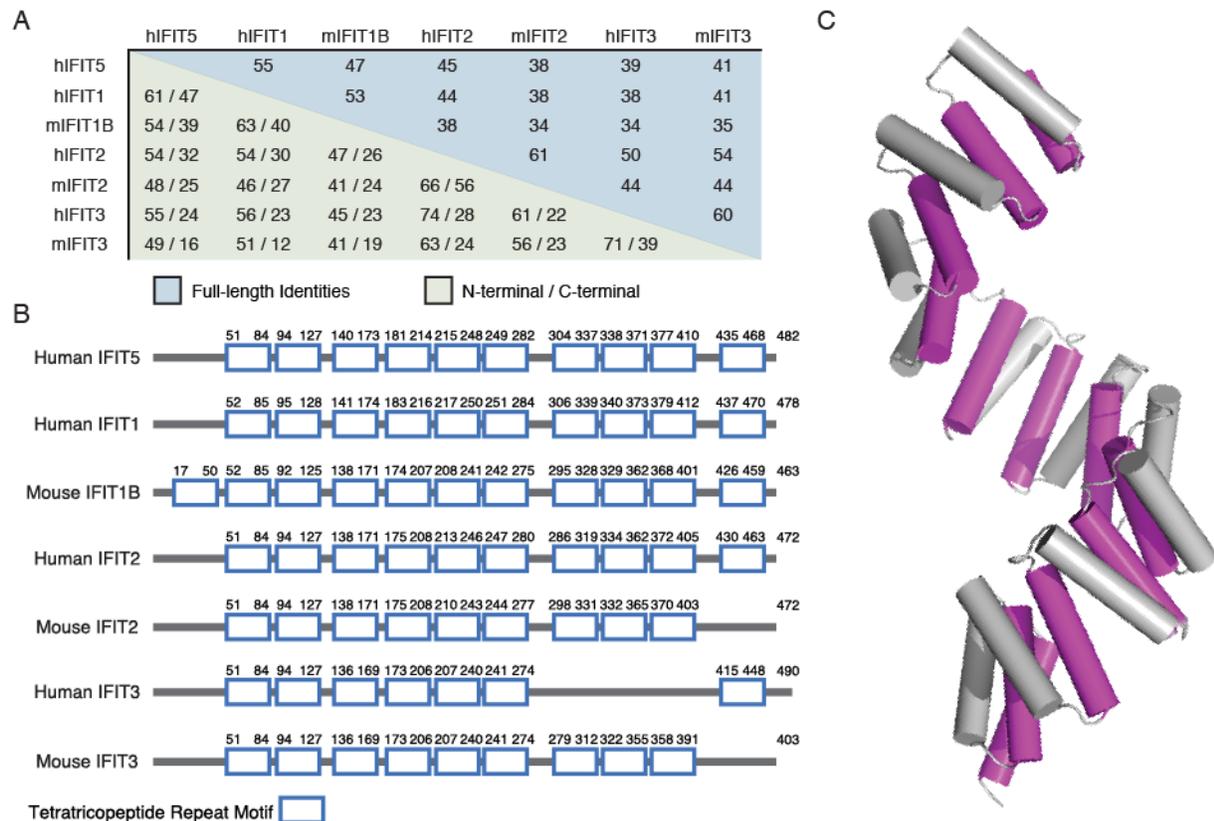


Figure 1.3 Relationship between IFIT proteins and primary structure

(A) Pairwise percent identity matrix of well-characterized human (h) and mouse (m) IFIT proteins. The N-terminal two-thirds of IFIT proteins are generally more conserved than the C-terminus. (B) IFIT proteins contain multiple copies of the tetratricopeptide repeat motif (TPR). (C) Crystal structure of the TPR domain of OGT (Jinek et al., 2004), the first multi-TPR containing protein crystallized showing a single, continuous super-helix.

mammalian precursor, and altogether they can be classified into 3 subgroups: IFIT5-like genes, IFIT1/IFIT1B-like, and IFIT2/IFIT3-like (Daugherty et al., 2016; Liu et al., 2013).

In humans, the five IFITs are clustered in a single locus on chromosome 10, with an additional non-transcribed pseudogene, IFIT1P, present on chromosome 13 (Fensterl and Sen, 2011). Mouse IFITs are also clustered together (on chromosome 19), and consist of three well-characterized genes: *Ifit1*, *Ifit2*, and *Ifit3*, and three more that are poorly characterized: *Ifit1b*, *Ifit1c*, and *Ifit3b* (Fensterl and Sen, 2011). Interestingly, mouse *Ifit1* was thought to be the murine ortholog of human IFIT1 for over two decades, but was recently shown to be, in fact, the murine ortholog of human IFIT1B (Daugherty et al., 2016). This discovery was made possible by a rigorous phylogenetic analysis which, by taking into account gene conversion events, successfully deconvoluted the complex evolutionary relationship of mammalian IFIT sequences (Daugherty et

al., 2016). Thus, mice have deleted the IFIT1 gene^{*}, and instead encode three copies of IFIT1B-like genes: *Ifit1*, *Ifit1b*, and *Ifit1c*; mice have also deleted the IFIT5 gene (Daugherty et al., 2016).

1.2.2 IFIT expression patterns *in vitro* and *in vivo*

Human and mouse IFITs all share a common gene architecture and, with some exceptions, contain 2 functional ISREs in their promoter regions (Fensterl and Sen, 2011). This makes them sensitive to a wide variety of stimuli that signal via IRFs (reviewed in (Diamond and Farzan, 2013; Fensterl and Sen, 2011; 2015; Sen and Peters, 2007)). Chief among these are Type I IFNs (α/β), which trigger ISRE-dependent induction through IRF9 (i.e. ISGF3) (Sen and Peters, 2007). IFN- γ can also upregulate IFITs, but not to the same levels as IFN- α/β (Der et al., 1998; Wathelet et al., 1987). Notably, IFITs can be induced in an IFN-independent manner, downstream of PRR recognition of viral PAMPs (e.g. by RLR and TLR) (Bandyopadhyay et al., 1995; Guo et al., 2000b; Tiwari et al., 1987). This direct activation is usually through IRF3 or IRF7, and places IFITs within a small subset of ‘early-responder’ ISGs that are induced directly following viral infection, and prior to IFN induction (Grandvaux et al., 2002; Sen and Peters, 2007). Other stimuli, such as LPS and retinoic acid, have been reported to upregulate IFITs, although the latter appears to do so indirectly through IFN- α (Lee et al., 1994; Niikura et al., 1997; Ovstebø et al., 2008; Smith and Herschman, 1996; Yu et al., 1997). In contrast to its human paralogs, IFIT1B lacks any ISREs and has yet to demonstrate responsiveness to IFN (Fensterl and Sen, 2011).

Under basal conditions, most cells do not express detectable IFIT mRNA or protein, but following viral infection or immunostimulation, they are rapidly induced to high levels. In fact, microarray analysis shows that IFITs usually rank among the most potently induced genes following interferon treatment or nucleic acid stimulation (Der et al., 1998; Goulet et al., 2013). At the protein level, one study showed that IFIT1 can reach > 2 million copies per cell after IFN treatment (Pichlmair et al., 2011). IFIT expression is transient, with mRNA levels peaking at 4-8 hours, and turning over within 12-24 hours post induction, although there are some instances of sustained IFIT levels (Chebath et al., 1983; Goulet et al., 2013; Guo et al., 2000b; Kusari and Sen,

^{*} Daugherty et al. proposed to rename mouse *Ifit1* to mouse IFIT1B, but at the time of writing of this thesis, the new nomenclature has not been adopted by official databases. Thus, the standard nomenclature of mouse *Ifit1* is retained in this thesis, even though it is referring to an IFIT1B-like protein. Work published prior to 2017 does not make a clear distinction between human IFIT1 and mouse *Ifit1*. In this thesis (excluding chapter 2), mouse *Ifit1* will be considered as an IFIT1B-like protein.

1986; Levy et al., 1986; Terenzi et al., 2006; 2005). Interestingly, IFIT5 has been shown to be constitutively expressed in several cell lines (Niikura et al., 1997; Pichlmair et al., 2011), and similarly IFIT1 has been reported to be upregulated in primary fibroblasts (Moll et al., 2011), livers of HCV infected individuals (Patzwahl et al., 2001), and CD34+ cells of patients with a hematopoietic malignancy (Pellagatti et al., 2006). Importantly, the expression kinetics of individual IFITs are context dependent, as they have been shown to vary *in vitro* with the type of cells and inducers (e.g. IFN/RNA/virus) being studied (Bandyopadhyay et al., 1995; 1992; Guo et al., 2000b; Terenzi et al., 2005; 2006; Wathelet et al., 1987). *In vivo* analysis of IFIT expression patterns in mice also shows striking tissue- and inducer-specific induction of individual IFIT genes, with several studies demonstrating selective up-regulation of only one or two of the IFIT proteins (Fensterl and Sen, 2015; Fensterl et al., 2008; Terenzi et al., 2007; Wachter et al., 2007). These distinct induction patterns suggest that IFITs may exert distinct and non-redundant antiviral effects in different cell types; however, their concomitant expression in many cell types also hints at possible synergistic or complementary activities.

1.2.3 Primary structure of IFIT proteins: The Tetratricopeptide Repeat motif

Unlike some of the highly expressed ISGs identified early on (e.g. PKR), IFIT proteins did not harbor any enzymatic activity. Their primary structure revealed the presence of multiple copies of the tetratricopeptide repeat (TPR) motif (**Fig. 1.3 B**) (Smith and Herschman, 1996), which at the time was known to be a degenerate, 34-amino-acid sequence characterized by a pattern of alternating small and large hydrophobic residues. No position is invariant, but the motifs can be identified by the presence of 8 loosely-conserved positions forming a TPR consensus sequence (Lamb et al., 1995; Main et al., 2003). Another defining characteristic of TPRs that facilitates their identification is their tendency to be arranged in tandem in a protein sequence, with clusters ranging in size from 3 to 16 motifs (Blatch and Lassel, 1999). Although, it is not uncommon to find individual blocks of one or more within a larger array. TPRs are found in all kingdoms of life, and are implicated in regulating diverse biological processes, usually through protein-protein interactions and the formation of multi-protein assemblies (Blatch and Lassel, 1999; D'Andrea and Regan, 2003).

The first crystal structure of a TPR containing protein was the TPR domain of protein phosphatase 5 (PP5), which is made up of three tandem repeats (Das et al., 1998). It showed that each motif folds into a pair of antiparallel α -helices, and adjacent TPRs are arranged in parallel

resulting in a regular repeat of antiparallel α -helices. Together, the three TPRs folded into a right-handed super-helix with a concave surface or continuous groove that can accommodate a target peptide or alpha-helix from an interacting partner. Based on the 3 repeats of PP5, a structural model for 12 tandem repeats was constructed, which suggested that multiple adjacent repeats could assemble in an open-ended manner to form an elongated, solenoid-like, right-handed super-helix (Das et al., 1998). The crystal structure of N-terminal O-linked GlcNAc transferase (OGT), which is made up of 11.5 TPR units, was the first of its kind and confirmed the model proposed by Das *et al.* (**Fig. 1.3 C**) (Jínek et al., 2004).

As TPR motif sequences are highly divergent, algorithms to delineate the individual TPRs in a protein sequence used to perform with limited accuracy (Karpenahalli et al., 2007). Whereas the manual analysis carried out by Smith *et al.* identified the presence of 10 putative TPR motifs in IFIT1 and IFIT2, subsequent studies relying on web-based resources such as Pfam incorrectly predicted only 4 to 6 (e.g. (Fensterl and Sen, 2011)). TPRpred is a specialized tool published relatively recently which provides a more accurate prediction for IFIT proteins (**Fig. 1.3 B**) (Karpenahalli et al., 2007). It employs a less stringent search algorithm and takes into account the tendency of TPRs to be found in tandem. Nevertheless, despite improvements in motif prediction, there remains several obstacles to predicting the 3-D structure of TPR-containing proteins based on existing models. This was evident in early attempts at modelling IFIT1 and IFIT2 using OGT as a template which, unsurprisingly, predicted one continuous super-helix (Fensterl and Sen, 2011; Pichlmair et al., 2011). However, as we shall see later, this template based homology modelling gave rise to errors as it could not predict the impact of residues found between adjacent TPRs, which can affect the helical stacking angles between successive motifs leading to changes in TPR super-helical parameters and tertiary fold (Main et al., 2003; Zeytuni and Zarivach, 2012).

1.2.4 IFIT biochemical activity

The expression of IFIT proteins has been linked to reduced viral replication in cell culture (e.g. (Raychoudhuri et al., 2011; Schmeisser et al., 2010; Terenzi et al., 2008; Zhang et al., 2007)). They are thought to exert an antiviral effect through modulating diverse biological processes (reviewed in (Diamond and Farzan, 2013; Fensterl and Sen, 2011)). The most well characterized of these is the ability of IFIT proteins to bind viral RNA and interfere with viral replication or viral translation (more details below and section 1.4). Other cellular processes in which IFITs are implicated in

regulating include cell proliferation (Feng et al., 2014; Hsu et al., 2011; Stawowczyk et al., 2011; Xiao et al., 2006), cell motility and migration (Lai et al., 2013; 2008), cell differentiation (Huang et al., 2008), and cytokine signalling and secretion (Berchtold et al., 2008; Huang et al., 2008; Li et al., 2009; Liu et al., 2011; McDermott et al., 2012). Until recently, IFIT proteins were also thought to inhibit cellular and viral cap dependent translation by interfering with 43S ribosomal complex formation, apparently through protein-protein interactions with subunits of the eukaryotic translation initiation factor 3 (eIF3) (Guo and Sen, 2000; Guo et al., 2000a; Hui et al., 2003; Terenzi et al., 2005; 2006; Wang et al., 2003)

1.2.5 IFIT1 and IFIT5 are RNA binding proteins that recognize viral 5' PPP RNA

To prevent aberrant recognition of self RNA by PRRs or ISGs, the innate immune system usually employs several safeguards at the level of PAMP recognition. For instance, potent activation of IFN signalling via RIG-I requires that the activating dsRNA display several features: a 5' PPP, a blunt-ended structure, and a minimal base-pair length of 10-20 nucleotides at the 5' end. These criteria help distinguish virus-derived nucleic acids from endogenous RNA, such as microRNA, which are short dsRNA that contain a 5' monophosphate and 3' overhangs (Marques et al., 2006). Activation of effector ISGs like PKR also requires the presence of virus-specific nucleic acids, such as dsRNA longer than ~ 33 bp in length, or ssRNA with a 5' PPP and limited secondary structure (Hull and Bevilacqua, 2016).

In 2008, Pichlmair *et al* hypothesized that other effectors ISGs require viral PAMP presence before they can realize their full antiviral potential. To test this, they took an unbiased proteomics approach to characterize ISGs which interact with PPP-RNA (Pichlmair et al., 2011). HEK293 cells were treated with IFN to upregulate antiviral molecules, and using affinity proteomics with PPP-RNA coated beads as 'bait', they pulled-down and identified endogenous human IFIT proteins (IFIT1, 2, 3 and 5) as PPP-RNA interactors. When recombinant IFITs were individually overexpressed in human or bacterial cells, pull-downs showed that only human IFIT1 and IFIT5 strongly bound to the PPP-RNA, whereas IFIT2 and IFIT3 interact indirectly through IFIT1. Further experiments using agarose gel shifts, enzyme-linked immunosorbent assays, and surface-plasmon resonance all confirmed that the interaction between IFIT1 and PPP-RNA was direct and specific, with an estimated $K_d \approx 250$ nM.

The antiviral activity of IFIT1 against several viruses was then investigated. Knockdown of human IFIT1 enhanced the replication of vesicular stomatitis virus (VSV), Rift valley fever

virus, and influenza A virus (IAV), all of which are negative-strand ssRNA viruses capable of generating PPP-RNA during their lifecycle. In contrast, the replication of encephalomyocarditis virus, a picornavirus that does not generate detectable PPP-RNA, was not affected by the knockdown treatment. When IFIT1 was pulled-down from cells infected with VSV or IAV, it was found to co-purify with viral RNA. Mouse Ifit1 was similarly shown to bind PPP-RNA, and its knockout resulted in enhanced VSV replication in mouse embryonic fibroblasts as well as decreased mouse survival following VSV infection. The authors therefore proposed that human IFIT1 and mouse Ifit1 can sequester viral PPP-RNA during infection to inhibit the replication of negative-strand ssRNA viruses (**Fig. 1.4**).

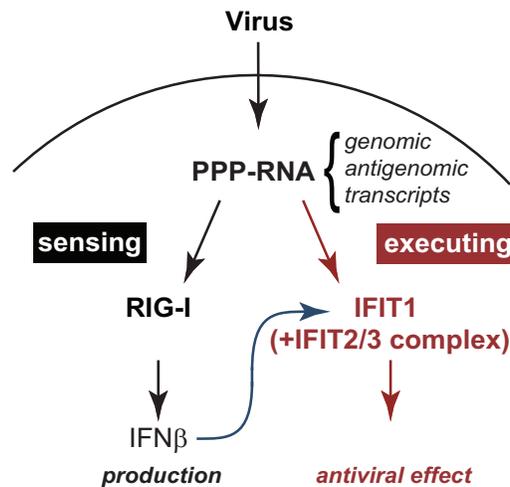


Figure 1.4 Working model for IFIT antiviral function on PPP-RNA

During viral infection, PPP-RNAs are generated which trigger RIG-I activation and subsequent IFIT upregulation. IFITs are part of an ‘executive’ branch of innate immunity, and their effector activity is targeted towards viral PPP-RNAs. Viral PPP-RNAs can be found on viral genomic and antigenomic (intermediate) RNA, or transiently present on viral mRNA transcripts before they are capped. Figure adapted from (Pichlmair et al., 2011).

1.3 Eukaryotic and viral mRNA cap synthesis

The discovery and characterization of a ‘blocked and methylated’ structure at the 5′ end of viral and cellular messenger RNA, or mRNA ‘cap’, was the result of the parallel combined efforts and pioneering work from numerous groups in the 1970s. As with many important findings at the time, studying viral systems played a key role in deciphering the cap structure, which consists of an *N*7-methylguanosine (m7G) moiety connected to the 5′ ends of RNA via a 5′-5′ triphosphate (ppp) linkage (Banerjee, 1980; Furuichi and Shatkin, 2000). Afterwards, considerable biochemical and genetic studies were devoted to investigating the eukaryotic and viral cap synthesis machinery which, despite utilizing diverse mechanisms, produce a chemically identical structure. The early work also identified additional modifications accompanying the mRNA cap which were restricted to higher eukaryotes (vertebrates and insects) and the viruses infecting them. Thus, in contrast to the mRNAs of yeast, plants, and other lower eukaryotes, which contain a minimal m7Gppp cap structure, the mRNA of vertebrates, insects, and their viruses are further modified by ribose 2′-O methylation at the first and sometimes second cap-proximal nucleotides (N1 and N2, where N is any nucleotide), to produce m7GpppNmN- and m7GpppNmNm-RNAs. *N*7 methylation has long since been known to play an essential role in mRNA processing, export, stability, and translation, but the function of ribose 2′-O methylation in the mRNA lifecycle, until recently, was poorly characterized.

1.3.1 A brief note on nomenclature

To distinguish between the different forms of cap, as well as the different modifications, the following nomenclature will be used throughout this text (see also **Fig. 1.5** for a schematic). All RNA base atoms will appear in italics (e.g. guanosine *N*7 or adenosine *N*6) to differentiate them from nucleotide positions along the mRNA chain (e.g. N1 and N2 for the penultimate and antepenultimate nucleotides that follow the cap). In the literature, the canonical nomenclature established early on referred to mRNA caps that are methylated on only guanosine *N*7 as Cap0 structures, while caps that contain additional ribose methylation at N1 only, or N1+N2, were described as Cap1 and Cap2 structures, respectively. These were also referred to as type I or type II structures. In this text, whereas Cap0/Cap1/Cap2 ‘structures’ refers to these different forms of cap that normally exist in nature, Cap0/Cap1/Cap2 ‘methyl’ or ‘methylation’ refers to the individual *N*7/*N*1/*N*2 methyl group or the act of modifying mRNA itself. The terms Cap1 methyl

and N1 ribose methyl, or Cap2 and N2 ribose methyl, may be used interchangeably. The term Cap0^{N2Me} or N2Me will also be used later to indicate capped RNAs that contain guanosine *N*7 methylation and ribose N2 methylation only, to distinguish them from Cap1 and Cap2 RNAs.

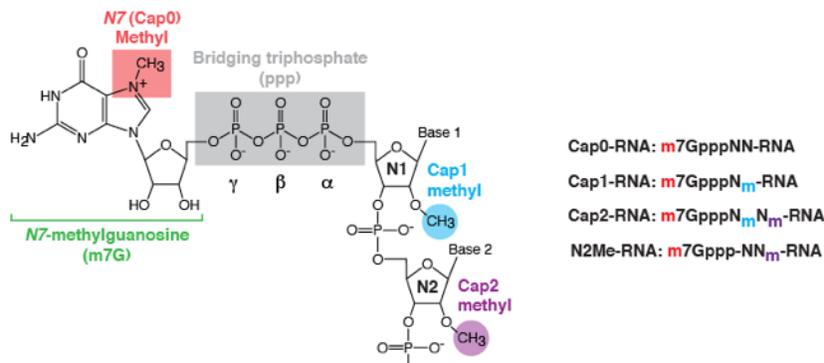


Figure 1.5 Schematic of differentially methylated mRNA cap structures

1.3.2 mRNA cap recognition in cap-dependent translation initiation

The mature and processed form of eukaryotic mRNA is made up of an open reading frame that is flanked at one end by the m7Gppp cap and 5' untranslated region (UTR), and at the other end by a 3' UTR and poly(A) tail. The mRNA cap promotes translation initiation in all eukaryotes, and thus its recognition by the eukaryotic translation initiation factor (eIF) 4F is a critical early step in cap-dependent translation initiation (Topisirovic et al., 2010). The large majority of viral mRNAs are also translated in this manner. eIF4F is a multi-subunit complex composed of the scaffolding protein eIF4G, the DEAD-box RNA helicase eIF4A, and the canonical cap binding protein eIF4E. During translation initiation, the eIF4F bound mRNAs are recruited to 43S ribosomal pre-initiation complexes (PIC) through an interaction between eIF4G and eIF3, resulting in a 48S ribosomal initiation complex which then proceeds to scan along the 5' UTR until it reaches an initiation codon, signalling the beginning of an open reading frame (Topisirovic et al., 2010).

*N*7 methylation of the cap is required for efficient cap-dependent translation, as the minimal structure recognized by eIF4E is m7Gpp (Marcotrigiano et al., 1997). The loss of *N*7 methylation results in > 100-fold reduction in affinity between eIF4E and cap analogs (Fechter and Brownlee, 2005). In contrast, the presence of an additional nucleotide (m7GpppN) and longer RNAs have only a modest effect on eIF4E-cap interactions (Fechter and Brownlee, 2005). The co-crystal structure of eIF4E with m7GDP revealed an α/β fold resembling a cupped-hand, with a narrow cap-binding slot on the concave basal surface of the protein (Marcotrigiano et al., 1997).

Specific engagement of an *N*7 methylated form of the cap, which contains a delocalized positive charge that is secondary to methylation, is mediated by the aromatic nature of the cap-binding slot (Hodel et al., 1997; Marcotrigiano et al., 1997), which is composed of two electron rich tryptophan residues that are highly selective for an electron deficient, positively-charged m7G (Hu et al., 2003). This mode of molecular recognition is known as the cation- π sandwich. Additional selectivity for a guanine cap is mediated by sequence specific hydrogen bonds directed towards the Watson-Crick face of m7G (*N*2, *N*1, and *O*6 groups) (Marcotrigiano et al., 1997). Positively charged residues that sit outside the cap-binding slot facilitate recognition of the bridging triphosphate.

The eIF4F complex has additional roles in translation initiation beyond direct m7G binding (Pelletier and Sonenberg, 1988). eIF4G harbors RNA binding activity, which helps stabilize the eIF4F mRNA complex. eIF4G also contains binding sites for the poly(A) binding protein (PABP), which therefore promotes synergy between the 5' cap and 3' poly(A) tail during translation. The helicase activity of eIF4A, assisted by eIF4B and eIF4H, unwinds local secondary structure and thus facilitates ribosomal binding and scanning along the 5' UTR in the 5' to 3' direction (Topisirovic et al., 2010).

1.3.3 Eukaryotic mRNA capping mechanisms

Capping is the first step in pre-mRNA processing (Martinez-Rucobo et al., 2015). It occurs co-transcriptionally in the nucleus, and is required to direct the downstream mRNA processing events such as pre-mRNA splicing, poly(A) tailing, and nuclear export (Aregger and Cowling, 2017). Cap synthesis in eukaryotes relies on three enzymatic activities (**Fig 1.6 A**) (Decroly et al., 2011b). First, an RNA 5' triphosphatase (RTPase) hydrolyzes the γ -phosphate of nascent ppp-RNA to produce pp-RNA. Next, a guanylyltransferase (GTase) reacts with GTP to form a covalent enzyme lysyl-N ϵ -GMP intermediate, and the GMP moiety is then transferred onto pp-RNA, yielding Gppp-RNA. Finally, an RNA *N*7-methyltransferase (*N*7-MTase) catalyzes *N*7-methylation of the cap using S-adenosyl-L-methionine (SAM) as a donor, giving rise to m7Gppp-RNA. In budding yeast, the first two activities are carried out by Cet1 and Ceg1, which together form a stable complex known as Capping Enzyme (CE), while the third is performed by Abd1 (Martinez-Rucobo et al., 2015). In metazoans, RTPase and GTase activities are encoded in a single bifunctional Mammalian Capping Enzyme (Mce), also called RNGTT (RNA guanylyl transferase

and 5' triphosphatase), while *N7*-methylation is catalyzed by RNMT (RNA (guanine-*N7*-) methyltransferase) (Cowling, 2010).

The capping process is functionally coupled to transcription by RNA pol II, the multi-subunit enzyme responsible for pre-mRNA synthesis (Martinez-Rucobo et al., 2015). The C-terminal domain (CTD) of RNA pol II contains heptad repeats that are differentially phosphorylated at serine/threonine residues throughout the transcription cycle to recruit factors, and to spatially and temporally coordinate pre-mRNA modifications (Cowling, 2010). Pol II is initially recruited to gene promoters in its unphosphorylated form, and upon initiation, will be phosphorylated at its CTD to promote transcriptional pausing and form a platform for binding of CE and other factors (Cowling, 2010). Transcriptional pausing is thought to provide a 'window of opportunity' for successful pre-mRNA capping (Adelman and Lis, 2012)

Electron microscopy studies of yeast RNA pol II showed that CE docks against one wall of pol II, spanning the RNA exit tunnel. Nascent ppp-RNA emerges from the tunnel once 17 nucleotides have been incorporated, thus triggering capping activity as CE is ideally placed for immediate substrate recognition (Martinez-Rucobo et al., 2015). A change in the CTD phosphorylation state promotes CE release and recruitment of Abd1 (Lidschreiber et al., 2013), which will then methylate the cap at *N7*. In mammals, recruitment of RNMT (the *N7*-MTase) is also mediated by phosphorylation of RNA pol II CTD, but the mechanism is not as well characterized as it is in yeast (Aregger and Cowling, 2013). The nuclear Cap Binding Complex (CBC) can now recognize an *N7*-methylated cap. Binding of CBC helps recruit kinases which promote transcriptional elongation, thus relieving RNA pol II from its transcriptionally paused state (Lidschreiber et al., 2013). CBC remains bound to the mRNA (until it is replaced by eIF4E) and plays an active role in further pre-mRNA processing and nuclear export.

CBC is a heterodimer composed of two polypeptides, CBP20 and CBP80, which bind m⁷Gppp-RNA through an induced fit mechanism (Calero et al., 2002; Mazza et al., 2002). Despite adopting an evolutionary divergent fold, CBP20 houses an aromatic cap binding slot analogous to that of eIF4E, and thus preferentially recognizes *N7*-methylated cap. Failure to *N7*-methylate the cap prevents CBC binding and leaves pre-mRNA exposed and susceptible to degradation through mRNA surveillance mechanisms in the nucleus (Jiao et al., 2010). In mammals, this pre-mRNA quality control is mediated by Dxo/Dom3Z, a nuclear enzyme which possesses decapping, 5' pyrophosphohydrolase, and 5' to 3' exonuclease activity (Jiao et al., 2013). Dxo/Dom3Z can

recognize Gppp- and ppp-RNA, two products of defective pre-mRNA capping, convert them to p-RNA, and along with XRN2 (5' to 3' exonuclease) target p-RNAs for degradation (Jiao et al., 2013).

1.3.4 Eukaryotic mRNA ribose 2'-O methylation at N1 (Cap1) and N2 (Cap2)

The mRNAs of higher eukaryotes are further modified by ribose Cap1 and Cap2 methylation on the first and second cap-proximal nucleotides. The enzymes responsible for catalyzing these modifications were isolated from HeLa cell fractions and partially characterized by Langberg et al in 1981 (Langberg and Moss, 1981), but the genes encoding each enzyme were not identified until much later (Bélanger et al., 2010; Haline-Vaz et al., 2008; Werner et al., 2011). Ribose 2'-O methylation at N1 and N2 are catalyzed by CMTr1 and CMTr2, respectively (Cap-specific mRNA (nucleoside-2'-O-)-methyltransferase 1 and 2) (Bélanger et al., 2010; Werner et al., 2011). They are both composed of multiple domains, and have in common a related catalytic domain belonging to the Rossmann-fold MTase (RFM) family, which possesses a conserved K-D-K-E catalytic tetrad characteristic of ribose 2'-O MTases (Werner et al., 2011). Phylogenetic analysis of their catalytic RFM domains indicates that the two are paralogous and likely arose from gene duplication of an early CMTr1-like precursor (Werner et al., 2011). Homologs are found throughout metazoan evolution but not in lower eukaryotes (Werner et al., 2011), consistent with the experimentally observed Cap1 and Cap2 structures in mammalian and insect mRNA, but not in that of yeast and plant (Banerjee, 1980). Homologs of CMTr1 are also present in protists (e.g. *Trypanosoma brucei*) and in many viral genomes.

1.3.4.1 The Cap1 methyltransferase CMTr1

CMTr1 is composed of a G-patch domain (which may be involved in protein-RNA interactions), a catalytic RFM domain, a non-catalytic GTase-like domain, and a C-terminal WW domain (Bélanger et al., 2010; Werner et al., 2011). Functional analysis of the isolated RFM domain shows that it retains only 60% activity *in vitro*, suggesting a role for the other domains in substrate binding or allosteric activation. The crystal structure of the isolated RFM domain in complex with capped RNA revealed a positively charged channel for RNA binding with two narrow clefts for binding the m7G moiety and SAM (Smietanski et al., 2014). Notably, cap recognition does not utilize a cation- π sandwich (which normally imparts high specificity for N7-methylation), which explains CMTr1's ability to modify both Gppp- and m7Gppp-RNA (Bélanger et al., 2010; Langberg and

Moss, 1981; Smietanski et al., 2014; Werner et al., 2011). Furthermore, the structure shows no sequence-specific interactions, explaining the broad substrate specificity of CMTr1 and the widespread prevalence of Cap1 structures *in vivo* (Banerjee, 1980).

CMTr1 also encodes an N-terminal nuclear localization signal which confines it to the nucleus (Bélanger et al., 2010; Haline-Vaz et al., 2008; Werner et al., 2011). Accordingly, *in vitro* Cap1 MTase activity has been purified from HeLa cell nuclear extracts, and mRNA isolated from mammalian cell nuclei all bear Cap1 type structures (Langberg and Moss, 1981; Perry and Kelley, 1976). Therefore, *in vivo* Cap1 methylation occurs either co-transcriptionally in the nucleus, or shortly afterwards. It is unclear how CMTr1 activity is controlled, or what ensures that all cellular mRNAs are modified before export. As with capping, it is possible that CMTr1 activity is linked to pre-mRNA processing at transcription initiation, as yeast two-hybrid and co-immunoprecipitation assays showed that CMTr1 interacts with the CTD of RNA pol II (Haline-Vaz et al., 2008). It is unclear yet if Cap1 modification precedes or follows *N7*-methylation during mRNA processing, as CMTr1 activity *in vitro* is not dependent on *N7*-methylation (Bélanger et al., 2010; Langberg and Moss, 1981; Smietanski et al., 2014; Werner et al., 2011), and Cap1 methylation does not interfere with RNMT *N7*-MTase activity *in vitro* (Thillier et al., 2012).

1.3.4.2 The Cap2 methyltransferase CMTr2

Unlike CMTr1, CMTr2 is composed of only two domains: a catalytic RFM domain, and a second inactive RFM-like domain which lacks the conserved SAM binding site and catalytic tetrad residues (Werner et al., 2011); deletion analysis shows that both domains are required for Cap2 methylation (Smietanski et al., 2014; Werner et al., 2011). Cap2 methylation is performed in the cytoplasm, as HeLa cell extract fractionation showed that Cap2 MTase activity was almost exclusively cytoplasmic (Langberg and Moss, 1981), and kinetic analysis of cap formation indicates that Cap2 structures arise from secondary methylation of Cap1 structures after mRNA has entered the cytoplasm (Perry and Kelley, 1976). However, a recent study showed that CMTr2 exhibits cytoplasmic and nuclear subcellular localization (Werner et al., 2011). Although this may be an artefact of protein overexpression, it suggests that CMTr2 may exist in the nucleus in an inactive form (Werner et al., 2011).

CMTr2 activity does not depend on *N7*-methylation or Cap1 methylation, but prior Cap1 modification has been shown to stimulate it (Werner et al., 2011). CMTr2 probably only modifies Cap1-structures *in vivo*, as mRNA with only m7G and ribose 2'-O methylation at N2 have not

been reported (Banerjee, 1980). This probably reflects the quantitative activity of CMTr1 in the nucleus and the cytoplasmic localization of CMTr2. However, whereas all mRNAs contain Cap1 methylation, only a subset of mRNA receive additional Cap2 methylation (Banerjee, 1980). In HeLa and Vero cells, only ~ 50% of mRNAs contained Cap2 structures (i.e. Cap1+Cap2 methylation) (Cleaves and Dubin, 1979; Furuichi et al., 1975). Analysis of the Cap2 containing mRNA across several mammalian cell lines does not reveal any patterns of sequence specific CMTr2 activity (Banerjee, 1980; Furuichi and Shatkin, 2000). As of yet, it is not clear how or why only a subset of mRNA contains Cap2 structures.

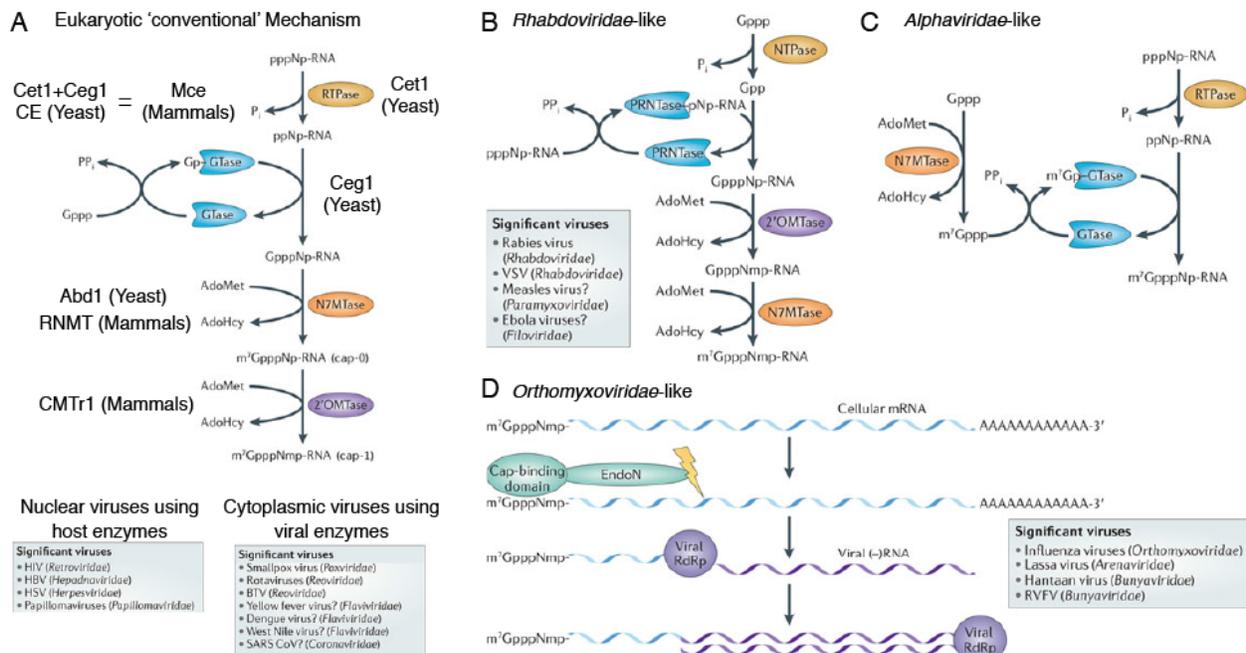


Figure 1.6 Eukaryotic and viral capping mechanisms

(A) Host nuclear 'conventional' capping pathway, also exploited by nuclear residing viruses and mimicked by cytoplasmic viruses encoding viral capping machinery. (B-D) 'Unconventional' viral capping mechanisms. Figure was adapted from (Decroly et al., 2011b).

1.3.5 Viral mRNA capping and 2'-O methylation

Viruses have adopted one of three general strategies to cap and methylate their mRNA. The first is utilized by nuclear retroviruses and DNA viruses, which do not encode capping enzymes and therefore exploit the host machinery (Fig. 1.6 A); the second is utilized by cytoplasmic RNA and DNA viruses, and relies on the action of virally-encoded capping enzymes (Fig. 1.6 A-C); the third employs a 'cap-snatching' mechanism unique to segmented negative-strand ssRNA viruses,

which steals capped RNA from the cellular pool (**Fig. 1.6 D**) (Decroly et al., 2011b). For the cytoplasmic RNA and DNA viruses, their capping mechanisms are referred to as ‘conventional’ if they mimic that of the eukaryotic host (**Fig. 1.6 A**), otherwise they are referred to as ‘unconventional’ (**Fig. 1.6 B-D**).

1.3.5.1 Retrovirus and DNA virus capping in the nucleus

Viruses replicating in the nucleus include the *Retroviridae* family (group VI positive-strand ssRNA viruses such as HIV) and most DNA viruses (group I dsDNA, group II ssDNA, and group VII retrotranscribing dsDNA). These viruses do not encode capping enzymes or cap-modifying MTases, and so they have come to rely on the host for mRNA synthesis (Decroly et al., 2011b). As such, their RNA pol II transcribed mRNAs are capped and *N*⁷-methylated in the nucleus by Mce and RNMT, respectively. Their mRNAs are also modified by CMTr1 in the nucleus, since the mRNAs or genomes of several positive-strand ssRNA retroviruses (e.g. rous sarcoma virus, avian sarcoma virus, and feline leukemia virus), and the mRNAs of several dsDNA viruses (adenovirus, simian virus 40, herpes simplex virus 1, and polyoma virus) all contain Cap1 methylation (Banerjee, 1980). The mRNA from adenovirus and herpes simplex virus 1 (HSV-1) infected cells were also shown to contain additional Cap2 methylation (Hashimoto and Green, 1976; Moss and Koczot, 1976; Moss et al., 1977), likely catalyzed by cytosolic CMTr2. In both cases, Cap2 structures appear to be less abundant than Cap1, and for HSV-1, described as a ‘minor’ component which is more prominent during infection at lower temperatures.

1.3.5.2 Vaccinia Virus and Reovirus ‘conventional’ capping in the cytoplasm

Most RNA and DNA viruses replicating in the cytosol cannot access the nuclear capping machinery, and therefore encode enzymes to synthesize mRNA caps. The viral capping machinery is highly diverse in terms of its genetic composition and domain organization. Nevertheless, many viruses follow the same sequence of reactions catalyzed by the eukaryotic enzymes, which is through sequential RTPase, GTase, *N*⁷-MTase, and 2'-O MTase activity; these mechanisms are referred to as ‘conventional’ (Decroly et al., 2011b).

The capping enzyme of Vaccinia virus, a member of the dsDNA *Poxviridae* family, is one of the first to be purified and characterized. It is a heterodimer composed of D1 and D12, which together are capable of performing all the steps necessary for Cap0 synthesis (Kyrieleis et al., 2014). Vaccinia capping enzyme is associated with the vaccinia RNA polymerase to ensure

efficient co-transcriptional capping of its mRNA (Hagler and Shuman, 1992). Vaccinia viral protein 39 (VP39) is a bifunctional enzyme which functions as a processivity factor for the viral poly(A) polymerase and also catalyzes 2'-O methylation to complete Cap1 synthesis (Hodel et al., 1996). Like eIF4E and CBC, VP39 contains a fully aromatic cap-binding slot which engages m7G in a cation- π sandwich, and is therefore highly specific for *N*7-methylated caps (Hodel et al., 1997). VP39 interacts with 6 cap-proximal nucleotides through a solvent exposed channel; RNA binding is sequence-independent and relies on conformation-dependent interactions with the RNA backbone (Hodel et al., 1998). Vaccinia virus mRNAs are substrates for CMTr2, as viral mRNA isolated early during HeLa cell infection (at 105 mins post infection) contains roughly equal proportions of Cap1 and Cap2 structures, whereas mRNA isolated later (5 – 7 hours post infection) were largely Cap1, with ~ 28 % Cap2 (Boone and Moss, 1977).

Viruses of the *Reoviridae* family contain a segmented dsRNA genome that remains encapsidated during infection within viral 'cores' that shelter the RNA and organize transcriptional and capping activities (Reinisch et al., 2000; Sutton et al., 2007). Orthoreovirus encodes a pentameric capping enzyme, λ 2, which forms a turret-like projection on the viral core. Inside λ 2 is a hollow cylindrical cavity lined with several inward-facing active sites to process mRNA as it emerges from the encapsidated polymerase (Reinisch et al., 2000). The spatial organization of active sites in the pentamer is consistent with a sequential mechanism in which *N*7-methylation of Gppp-RNA precedes 2'-O methylation (Furuichi et al., 1976; Reinisch et al., 2000). The pentamer is gated at its exit by a C-terminal flap, which is thought to retain the nascent RNA and ensure efficient capping and methylation (Reinisch et al., 2000). Supporting that notion, reovirus mRNA isolated from mouse fibroblasts during infection shows complete Cap1 methylation; these studies also showed that about half of the viral mRNA is further methylated by CMTr2 (Desrosiers et al., 1976). The capping enzyme VP4 of bluetongue virus (another member of the *Reoviridae* family) also utilizes an analogous enzyme to carry out Cap1 synthesis. In contrast to orthoreovirus, the capping enzyme VP4 of bluetongue virus lies inside the capsid and its structure shows a linear arrangement of multiple active sites indicative of substrate channeling. VP4 is tightly bound to the viral polymerase to allow for immediate capping and methylation of the emerging chain (Sutton et al., 2007).

1.3.5.3 Flavivirus, coronavirus, and alphavirus capping in the cytoplasm

The capping machinery of the flaviviruses, coronaviruses, and alphaviruses (all positive-strand ssRNA viruses) are not as extensively characterized as vaccinia virus, partly because all the components have not been experimentally identified or structurally characterized. *Flavivirus* and *Coronaviridae* capping machinery have gained attention recently, as flaviviruses such as West Nile virus (WNV), Dengue virus (DENV), and Zika virus, and coronaviruses (CoV) such as SARS-CoV and MERS-CoV, are associated with emerging diseases and outbreaks.

Flavivirus genome replication and mRNA synthesis is carried out by a membrane-bound Replication Complex (RC) composed of viral non-structural (NS) proteins and unknown host factors (Klema et al., 2015). NS3 and NS5 are two components of RC that are associated with each other and contain all the enzymatic activities necessary for mRNA synthesis, capping, and methylation (Saeedi and Geiss, 2013). NS3 contains protease activity in its N-terminus and helicase and RTPase activity in its C-terminal domain; NS5 encodes *N7*-MTase and 2'-O MTase in its N-terminal domain, and RNA-dependent RNA polymerase activity at its C-terminus (Saeedi and Geiss, 2013). Recent evidence suggests that the NS5 NTD also harbors GTase activity (Bollati et al., 2009; Egloff et al., 2007; Issur et al., 2009), which is striking considering that it is only ~ 30 kDa and already encodes both *N7*- and 2'-O MTase functionalities. The NS5 NTD from several virus species, including DENV and WNV, has been the subject of numerous structural and functional studies (reviewed in (Dong et al., 2014)). Unlike other viral 2'-O MTases, flavivirus MTase activity is regulated by RNA sequence and structure. *N7*-methylation requires a conserved RNA stem-loop structure found at the 5' end of flavivirus genomes, and specific nucleotides at position 2 and 3, while 2'-O methylation requires a specific 5' sequence, m7GpppAG, that is conserved among flaviviruses, and a minimum length of 20 nucleotides for activity (Dong et al., 2010; 2007). The structures of NS5 NTD show a typical RFM fold with a positively charged RNA binding channel and a single SAM binding site (Zhou et al., 2007). The presence of one SAM binding site, and the preference of NS5 for 2'-O methylating m7Gppp-RNA over Gppp-RNA, is consistent with a sequential activity in which *N7*-methylation precedes 2'-O methylation (Dong et al., 2010; Ray et al., 2006). The *in vivo* cap structure of DENV-2 during infection of Vero cells was shown to be Cap1, indicating that it is not further modified by CMTr2 (Cleaves and Dubin, 1979).

Coronavirus mRNA synthesis is also carried out by a membrane bound complex called replication-transcription complex (RTC), made up of 16 non-structural proteins (nsp), as well as other viral and cellular proteins (Sawicki et al., 2007). RTPase and *N7*-MTase activities are encoded by nsp13 and nsp14, respectively (Bouvet et al., 2010; Chen et al., 2009; Ivanov et al., 2004). GTase activity has yet to be isolated from any viral proteins, and enzyme linked GMP intermediates have not been observed (Sevajol et al., 2014). 2'-O methylation is performed by nsp16, which requires binding of a viral protein cofactor, nsp10, to stabilize the SAM binding pocket and form an extended RNA binding surface (Chen et al., 2011; Decroly et al., 2011a). In contrast to *N7*-methylation by nsp14, 2'-O methylation by nsp16/nsp10 shows sequence specificity, requiring an adenine at N1 (Chen et al., 2011; 2013). 2'-O methylation by nsp16 requires m⁷Gppp-RNA, suggesting that *N7*-methylation precedes 2'-O methylation *in vivo* (Bouvet et al., 2010). Purification of murine hepatitis virus (a mouse coronavirus) mRNA from mouse fibroblasts following infection reveals the presence of Cap1 and Cap2 structures on the viral mRNA (Lai et al., 1982), indicative of further processing by CMTr2.

Whereas flavivirus and potentially coronavirus capping follows the 'conventional' pathway, alphavirus mRNA capping follows an unconventional mechanism (**Fig. 1.6 C**). The cellular and viral GTase activities described thus far proceed through a covalently linked enzyme•GMP intermediate, but the alphavirus GTase (nsP1) first methylates GTP at the *N7* position and then forms a covalently linked nsP1•m⁷GMP intermediate (Ahola and Kääriäinen, 1995), which is then transferred onto nascent pp-RNA that was processed by the RTPase activity of alphavirus nsP2 (Ahola et al., 1997; Vasiljeva et al., 2000). Alphaviruses are also distinct from most viruses in that they do not encode any 2'-O MTases. Consistent with that, viral genomic RNA and mRNA isolated from cells infected with two alphaviruses, Semliki Forest virus and Sindbis virus, shows that they contain predominantly Cap0 structures; interestingly a small fraction of mRNA contains additional methylations on the *N2* position of the guanine cap (Dubin et al., 1977; Hsueh and Dubin, 1976; van Duijn et al., 1986). The roles of the resultant dimethylguanosine (m^{2,7}G-) and trimethylguanosine (m^{2,2,7}G-) caps on alphavirus RNA are unknown.

1.3.5.4 *Rhabdoviridae*-like capping in the cytoplasm

Vesicular stomatitis virus (VSV) is a non-segmented, negative-strand ssRNA virus (nsNSV) belonging to the *Rhabdoviridae* family, which is a member of the *Mononegavirales* order that encompasses other nsNSV families including *Filoviridae* (e.g. Ebola virus) and *Paramyxoviridae*

(e.g. measles virus, Newcastle disease virus, human parainfluenza virus, respiratory syncytial virus). They all encode a multifunctional L protein responsible for mRNA synthesis, capping, methylation, and polyadenylation, for which the L protein from VSV is the prototype (Liang et al., 2015).

VSV L protein is composed of several catalytic domains: an RNA-dependent RNA polymerase domain, a capping domain, and an MTase domain, as well as two others, a connector domain and C-terminal domain, which appear to play a structural role (Liang et al., 2015). VSV and potentially other nsNSVs follow an unconventional mechanism for capping and methylation (Decroly et al., 2011b), differing from eukaryotic and most viral capping mechanisms in two ways (**Fig. 1.6 B**). First, the capping domain harbors polyribonucleotidyl transferase (PRNTase) instead of GTase activity, and reacts with nascent pppRNA to form a covalent, enzyme-linked monophosphate RNA (E•pRNA) intermediate; the pRNA is then ligated to GDP, derived from GTP, to form GpppRNA (Liang et al., 2015; Ogino and Banerjee, 2007). Second, the MTase domain performs both *N*7- and 2'-O methylation using a single SAM binding site (Li et al., 2006), which is not unlike flavivirus NS5, but the sequence of events is different: 2'-O methylation by L protein precedes and facilitates *N*7-methylation (Paesen et al., 2015; Rahmeh et al., 2009). Capping and methylation of VSV mRNA is also tightly linked to mRNA synthesis. The nascent chain is capped and methylated only after 31 nucleotides have been incorporated (Tekes et al., 2011), and a failure to cap or methylate viral mRNA results in premature transcription termination and hyperpolyadenylation, respectively (Liang et al., 2015; Ogino, 2014). In this way, VSV L polymerase ensures efficient capping and 2'-O methylation of viral mRNA. A small fraction of VSV mRNA appears to be modified by CMTr2, as ~ 20% of VSV specific mRNA isolated from infected cells contained Cap2 structures (Rose, 1975).

Although L protein from viruses belonging to the *Paramyxoviridae* family are thought to use a VSV-like mechanism and contain a conserved Cap1 MTase domain, they do not all uniformly Cap1 methylate their mRNAs. For instance, mRNAs isolated during Newcastle disease virus infection are Cap0 type (Colonno and Stone, 1976), and respiratory syncytial virus mRNAs generated *in vitro* are also lacking Cap1 methylation (Barik, 1993). In contrast, measles virus mRNAs are properly Cap1 methylated (Yoshikawa et al., 1986). These results demonstrate that even though a virus may encode a Cap1 MTase, it may not be functional. They also show that 2'-

O methylation and *N*7-methylation are not coupled during *Paramyxoviridae* mRNA synthesis, which contrasts with the prototypical VSV mechanism.

1.3.5.5 Cap snatching by segmented negative strand RNA viruses in the nucleus or cytoplasm

Cap-snatching is a primer-dependent mechanism of mRNA synthesis that is restricted to the segmented negative-strand RNA virus (sNSV) group, which is composed of three families: *Orthomyxoviridae*, *Bunyaviridae*, and *Arenaviridae* (Reguera et al., 2016a). These viruses encode RNA-dependent RNA polymerases (RdRp) which synthesize viral mRNA using capped-primers ‘snatched’ from host cellular RNA (**Fig. 1.6 D**) (Decroly et al., 2011b; Reguera et al., 2014). The most well characterized of these is the cap-snatching mechanism of the nuclear residing Influenza virus (family *Orthomyxoviridae*). Its heterotrimeric RdRp, composed of PA, PB1, and PB2 subunits, associates with the CTD of actively transcribing RNA pol II near transcription start sites (Engelhardt et al., 2005). The PB2 subunit contains cap binding activity, and directs the RdRp to capped RNA pol II transcripts, while the PA subunit cleaves the RNA 10-15 nucleotides from the cap, thereby releasing capped fragments to be used as primers for mRNA synthesis (Dias et al., 2009; Guilligay et al., 2008; Plotch et al., 1981; Reich et al., 2014). The Influenza RdRp was initially thought to target primarily mRNA or nascent pre-mRNA for endonuclease cleavage and cap acquisition (Bouloy et al., 1978; Engelhardt et al., 2005; Plotch et al., 1981), but recent evidence suggests that more than half of the snatched sequences are derived from capped non-coding RNAs including U1/U2 small nuclear RNAs (which are involved in pre-mRNA processing and are also 2'-O methylated) (Gu et al., 2015). Notably, the presence of a Cap1 structure on donor mRNA dramatically enhances the priming activity of Influenza polymerases by up to 14-fold compared to a Cap0 structure (Bouloy et al., 1980; Wakai et al., 2011), which is consistent with early studies showing that Influenza mRNA contain Cap1 structures only (Krug et al., 1976). As Influenza and other *Orthomyxoviridae* are nuclear, it is unlikely that they encounter and acquire Cap2 structures for cap-snatching.

Viruses belonging to the *Bunyaviridae* and *Arenaviridae* family reside in the cytosol and encode a single chain RdRp known as L protein (Gerlach et al., 2015). Several structural and functional analyses showed that L proteins from *Bunya*- and *Arenaviridae* harbor endonuclease activity in their N-terminal region that is required for their cap-snatching and transcriptional activities, analogous to that of Influenza virus, but the location of the cap binding domain remains elusive (Gerlach et al., 2015; Lehmann et al., 2014; Reguera et al., 2016b; 2010). For some

Arenaviridae, studies have suggested that the cap-binding domain resides within polymerase associated accessory proteins (Kranzusch et al., 2010), or within the C-terminal region of L protein (Lehmann et al., 2014). Hantaviruses (family *Bunyaviridae*) encode a nucleocapsid (N) protein which can bind capped RNAs *in vitro* (Mir et al., 2008); in cells, Hantavirus N protein has been shown to recognize capped mRNAs containing premature termination codons (Cheng and Mir, 2012; Mir et al., 2008). These defective mRNAs would normally be targeted for degradation in cytosolic processing (P) bodies by the nonsense mediated decay pathway, but instead are protected at their 5' ends from cytosolic decapping enzymes by N, resulting in ~ 180 nucleotide capped fragments that are recovered from P bodies and serve as donors for cap-snatching by Hantavirus L protein (Cheng and Mir, 2012; Mir et al., 2008). It is unclear if *Arena*- and *Bunyaviridae* viruses preferentially acquire Cap1 mRNA in the same manner as Influenza virus. The 5'-cap structures of representative viral mRNAs have yet to be characterized, and the impact of ribose 2'-O methylation on their cap binding, endonuclease, and priming activities has not been systematically investigated. However, as both virus families replicate in the cytosol, where they would encounter only Cap1 or Cap2 structures, it is likely that their mRNA cap methylation status will reflect that of the cytosolic host mRNA.

1.3.6 Uncapped viral mRNA 5' ends

Some positive-strand ssRNA viruses do not rely on cap-dependent translation initiation, and instead recruit ribosomes or initiation factors directly via structured elements in the 5' UTR of their genomes or mRNAs. These elements, called internal ribosome entry sites (IRES), form a ribosomal landing pad that obviates the need for an mRNA cap (Pelletier and Sonenberg, 1988). Viruses relying on IRES-mediated translation include members of the *Hepacivirus* genus from the *Flaviviridae* family, such as hepatitis C virus (HCV), and viruses of the *Picornaviridae* family such as polio virus and encephalomyocarditis virus (EMCV). HCV genomes and mRNA are not 5' processed and thus retain a triphosphate (Kell et al., 2015). In contrast, *Picornaviridae* genome 5' ends are covalently attached to VPg (viral protein genome linked) which functions as a protein primer for RNA synthesis during genome replication (Lin et al., 2009). *Picornaviridae* mRNAs, however, are not capped with VPg and instead contain a 5' monophosphate (Banerjee, 1980; Nomoto et al., 1977). Interestingly, viruses of the *Caliciviridae* family (also positive-strand ssRNA viruses) covalently link their RNA to a VPg-like protein that is not related to the picornaviral VPg,

but instead functions as a proteinaceous cap substitute that interacts with eIF4E directly to promote translation of its viral mRNA (Goodfellow, 2011).

1.3.7 Potential biological roles of mRNA ribose 2'-O methylation

Both higher eukaryotes and their viruses utilize various mechanisms to ensure proper capping and methylation. In the eukaryotic nucleus, host and viral mRNA capping and subsequent processing are tightly linked to initiation, with quality control mechanisms in place to avoid failures of capping. The Cap1 methylation process is also potentially linked to pre-mRNA processing in the nucleus, although it is unclear if there are mechanisms ensuring proper and complete Cap1 methylation before mRNA export. Viruses ensure proper capping and ribose methylation through various means, for instance by relying on encapsidated assembly lines that are in close proximity or tightly bound to viral polymerases (e.g. *Reoviridae*), or the use of multi-component assemblies and multifunctional enzymes that encode all the activities required for mRNA synthesis, capping, and methylation. In some cases, Cap1 methylation is coupled to mRNA synthesis (e.g. VSV), and in others, a pre-requisite (e.g. Influenza cap-snatching). Except for alphaviruses, virtually all animal viral mRNA purified from infected cells contain Cap1 methylation; some also contain partial Cap2 methylation. Altogether, these mechanisms support the existence of a critical role for Cap1 methylation in the mRNA lifecycle. Yet, despite being discovered almost 40 years ago, a clear role for mRNA ribose 2'-O methylation remained enigmatic for nearly 35 years.

Studies addressing the function of Cap1 methylation during translation were limited in number, and showed only a minor impact on translation, if at all. *In vitro*, Cap1 methylation had a modest effect on binding to ribosomes in wheat germ extracts, but resulted in a 5-fold enhancement on ribosome binding in rabbit reticulocyte lysates, although it should be noted that ribosome binding in the latter case was not efficient to begin with (Muthukrishnan et al., 1976; 1978). In mouse L cells, Cap2-mRNA appeared to have a slower turnover rate compared to Cap1-mRNA (Perry and Kelley, 1976), but it is not clear if this was a result of methylation, or simply due to modification of inherently stable mRNA by the mouse L cell MTases. In *Xenopus* oocytes, progesterone treatment following injection of c-mos mRNA induced Cap1/Cap2 methylation of the injected mRNA, which then triggered oocyte maturation by enhancing Mos protein synthesis (Kuge et al., 1998). Although ribose 2'-O methylation was found to enhance translation under the specific circumstance described above, knockdown of CMTr1 in HeLa cells did not significantly affect translation nor impact cell viability (Bélanger et al., 2010).

1.4 mRNA Cap1 methylation is a molecular signature of self

As described above, the mRNAs of higher eukaryotes and many viruses which infect them display both Cap0- and Cap1-methylations. Whereas guanosine *N*7-methylation (Cap0) has an essential role in mRNA processing and promoting efficient cap-dependent translation initiation, a clear biological role for ribose 2'-O methylation at N1 (Cap1), on both viral and eukaryotic mRNA, eluded researchers for several decades.

1.4.1 Ribose 2'-O methylation of viral mRNA enhances virulence and evades IFIT activity

One of the earliest clues to the contribution of Cap1 methylation towards viral infection was uncovered in 2007. Through mutations in the KDKE tetrad of the West Nile virus (WNV) NS5 MTase domain, Zhou *et al.* generated WNV mutants lacking either *N*7-MTase or 2'-O MTase activity, and tested their infectivity (Zhou *et al.*, 2007). They showed that, whereas abrogating viral *N*7-methylation resulted in aborting viral infection, disrupting viral 2'-O methylation produced viruses that were partially attenuated in cell culture, but could still replicate efficiently under permissive conditions. Strikingly, in comparison to the wild-type virus, the 2'-O MTase mutant of WNV was highly attenuated *in vivo*, and mice infected with this mutant virus[†] showed decreased morbidity and mortality, suggesting that viral 2'-O methylation is required for virulence. Subsequently, Daffis *et al.* and Züst *et al.* discovered the underlying mechanism for this decreased virulence, showing that 2'-O MTase mutants of WNV, mouse hepatitis virus (MHV), human coronavirus (HCoV), and vaccinia virus all exhibit enhanced sensitivity to the antiviral effects of IFN, and in particular, IFIT proteins (Daffis *et al.*, 2010; Züst *et al.*, 2011).

Later studies also showed that 2'-O MTase mutants of Dengue virus (DENV), SARS coronavirus, Japanese encephalitis virus (JEV), and human metapneumovirus are more susceptible than their wild-type counterparts to IFN and/or IFITs (Kimura *et al.*, 2013; Li *et al.*, 2013; Menachery *et al.*, 2014; Züst *et al.*, 2013), revealing potentially broad antiviral activity of IFITs in this process. Together, these results rationalized the existence of eukaryotic and viral Cap1 ribose 2'-O methylation, suggesting that it evolved as a marker of 'self' to prevent IFIT recognition. Whereas it protects host mRNA from aberrant IFIT activity, many viruses have acquired mechanisms to generate Cap1 mRNA to mimic the host and escape IFIT-mediated restriction.

[†] The term 'mutant virus' will henceforth refer to viruses with deficient 2'-O MTase activity, and thus only generate Cap0-mRNAs. This term applies only to viruses that normally encode Cap1 MTase functionality in their genomes.

1.4.2 IFIT1 and IFIT1B selectively inhibit translation of viral Cap0-mRNA by competing with eIF4E/eIF4F

Mouse knock-outs experiments showed that mouse *Ifit1* was primarily responsible for these enhanced antiviral effects towards mutant viruses (Cho et al., 2013; Daffis et al., 2010; Kimura et al., 2013; Szretter et al., 2012; Züst et al., 2011). In human cells, the role of each IFIT in this process was less clear. To shed light on the contribution of IFIT proteins in targeting viral mRNA, Habjan *et al.* performed pull-downs from human and mouse cells and showed that only human IFIT1 and mouse *Ifit1* could interact directly with capped RNA (Habjan et al., 2013). Similarly, using primer-extension foot-printing assays with purified proteins, Kumar *et al.* showed that human IFIT1, rabbit IFIT1 and rabbit IFIT1B could specifically recognize capped RNAs (Kumar et al., 2014). Using the foot-printing assay, Kumar *et al.* estimated the affinity between IFIT1 or IFIT1B proteins and capped RNAs to be on the order of 10-20 nM, approximately 10-fold better than the interaction between IFIT1 and PPP-RNA (Pichlmair et al., 2011). Additional experiments from other groups confirmed IFIT1/IFIT1B preference for capped-RNA over PPP-RNA (Pinto et al., 2015). Importantly, all studies showed that 2'-O methylation of the capped RNA at N1 (i.e. Cap1 methylation) interfered with IFIT1 and IFIT1B RNA binding. Thus, IFIT1 and IFIT1B proteins specifically recognize capped mRNAs lacking ribose N1 methylation.

Using luciferase reporter mRNA transfected into mouse embryonic fibroblasts (MEFs), Kimura *et al.* showed that mouse *Ifit1* preferentially inhibited translation of reporter mRNA lacking Cap1 methylation (Kimura et al., 2013). Similarly, during MHV infection of MEFs, Habjan *et al.* performed pulsed stable-isotope labelling in cell culture (SILAC) and subsequent whole proteome analysis by LC-MS/MS, to show that mouse *Ifit1* inhibited synthesis of viral proteins only in the absence of viral Cap1 methylation (Habjan et al., 2013). Importantly, synthesis of cellular proteins in this assay was unaffected by mouse *Ifit1*, confirming selectivity towards viral Cap0-mRNAs. Mass spectrometry, pull-downs, and foot-printing assays all showed that IFIT1 and IFIT1B proteins could compete with eIF4E and eIF4F for binding to Cap0-RNA (Habjan et al., 2013; Kumar et al., 2014). Accordingly, IFIT1/IFIT1B proteins could inhibit 48S complex formation on Cap0-mRNA (Kumar et al., 2014). Of note, although IFIT1/IFIT1B proteins could associate weakly with 40S ribosomes, and interact non-specifically with tRNAs (e.g. initiator tRNA_i^{Met}), these activities did not interfere with 43S ribosomal complex formation (Kumar et al., 2014), in contrast to previous work by Sen and colleagues. Taken together, these

results showed that the antiviral activity of IFIT1 and IFIT1B is largely determined by their ability to selectively bind to Cap0 structures, and prevent translation of viral mRNAs lacking 2'-O methylation (**Fig. 1.7**).

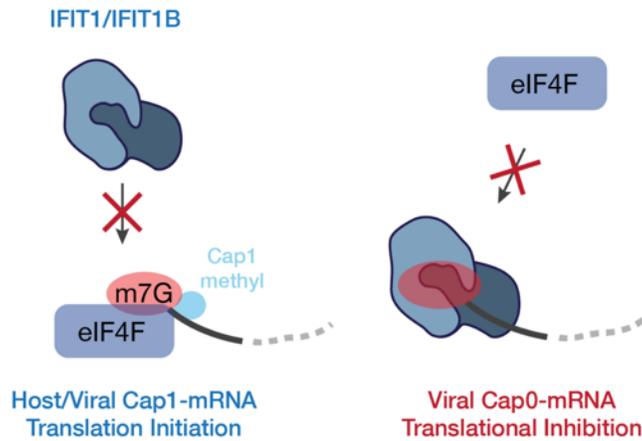


Figure 1.7 Model of mRNA self vs non-self discernment by IFIT1 and IFIT1B

During infection, IFIT1 and IFIT1B are upregulated and can selectively target viral Cap0-mRNA for translational inhibition. Host or viral mRNA displaying Cap1 methylation are protected from IFIT1/IFIT1B recognition.

1.4.3 Cap1 methylation also interferes with RIG-I Like Receptor signalling

In addition to evading IFIT antiviral activities, Cap1 methylation of viral mRNA also escapes detection by RIG-I-Like receptors. The loss of viral Cap1 methylation was reported to enhance both MDA5 and RIG-I dependent responses (Zust et al., 2011). Structures of RIG-I with capped and uncapped dsRNA show that a conserved histidine residue in its C-terminal regulatory domain (H830) is singularly responsible for sensing ribose 2'-O methylation at N1. When this residue was mutated (H830A), RIG-I could no longer discriminate between Cap0 and Cap1 RNAs, and therefore H830A restored RIG-I signaling in response to viral and Cap1 methylated dsRNA (Devarkar et al., 2016; Schuberth-Wagner et al., 2015).

Interestingly, Schuberth-Wagner *et al.* discovered that this histidine residue promotes immune-tolerance to self, as exogenous expression of the H830A mutant (that can bind Cap1 RNAs) results in immune signaling in response to endogenous RNA. They also showed that knockdown of hCMTr1 results in a RIG-I-dependent type I IFN response, demonstrating for the first time that the endogenous Cap1 MTase activity of hCMTr1 is required to suppress auto-immune responses. However, as RIG-I requires base-paired RNA with a 5' blunt end for optimal signaling, the nature of this RIG-I-activating endogenous ligand remains unclear.

1.5 Overview of thesis and objectives

IFITs are among the most potently induced genes following viral infection or interferon stimulation. Their primary mode of action was believed to be through disruptive protein-protein interactions, for instance, inhibiting translation by targeting subunits of eIF3. However, the work of Pichlmair *et al* and Daffis *et al* challenged that notion by uncovering a more direct role for IFIT proteins in antiviral innate immunity. Pichlmair *et al* showed that IFITs, namely IFIT1 and IFIT5, are RNA binding proteins which can interact with the 5' triphosphate commonly found at the ends of viral RNA. At the same time, Daffis *et al* discovered a role for Cap1 ribose 2'-O methylation of viral mRNA in evading restriction by IFIT proteins, although no direct interaction between IFITs and capped RNA was demonstrated at the time.

Thus, it was clear that IFIT proteins played a prominent role in the early host response to viral infection. Through self vs non-self discernment of nucleic acids, IFITs could target viral RNAs and selectively block viral processes. Little was known about the structure of IFIT proteins, except that their sequence is composed of multiple copies of the tetratricopeptide repeat motif, a helix-turn-helix motif with the propensity to form extended super-helical structures. Moreover, TPRs are a classical protein-protein interaction module, and as such there remains relatively few biochemical studies of TPR-nucleic-acid interactions, and no other structural studies to date. Therefore, how IFITs interact with RNA was unclear, although because of our early collaboration with Dr. Superti-Furga and Dr. Andreas Pichlmair, we knew only that a 5' triphosphate could stimulate RNA binding to IFIT1 and IFIT5.

In the first part of this thesis, I describe the structure of human IFIT5, its complex with 5' triphosphate RNAs, and an N-terminal fragment of human IFIT1. This work revealed a novel fold specific to IFIT proteins, uncovered the structural basis for recognizing the 5' triphosphate, demonstrated a preference for single stranded RNA by IFIT1 and IFIT5, and suggested a mechanism for potential sequence non-specific binding of RNA. The crystal structures, combined with small-angle X-ray scattering and limited protease digestion in solution, also uncovered a role for conformational changes in RNA binding. Finally, structure-guided mutational analysis and viral infection assays performed by our collaborators showed that RNA binding was required for the full-antiviral activity of IFIT1 and IFIT5.

Further studies on IFIT1 by several groups showed that its primary mode of action was to selectively inhibit translation of viral mRNA by binding to capped mRNA. In the second part of

my thesis, I determined the high resolution structure of monomeric human IFIT1 bound to capped RNA, which revealed that IFIT1 forms a water-filled tunnel with a relatively hydrophobic cap binding pocket that can recognize multiple forms and conformations of the cap. Structural and functional analyses of the interaction between IFIT1 and the cap-proximal region of RNA showed that, in addition to Cap1 methylation on N1, Cap2 methylation on N2 is another molecular signature of ‘self’ that could protect endogenous mRNA from IFIT1 recognition. Although RNA binding is generally sequence non-specific, I show evidence that IFIT1 could overcome Cap1 or Cap2 methylations in an RNA-dependent manner. This work helps to expand the molecular determinants of self vs non-self that govern the antiviral activity of IFIT1, and suggests an adaptable mechanism for its viral mRNA recognition which is important for its broad antiviral activity.

CHAPTER 2: STRUCTURAL BASIS FOR VIRAL 5'-PPP-RNA RECOGNITION BY HUMAN IFIT PROTEINS

Abbas, Y.M., Pichlmair, A., Gónna, M.W., Superti-Furga, G., and Nagar, B. (2013). Structural basis for viral 5'-PPP-RNA recognition by human IFIT proteins. *Nature* **494**, 60-64.

2.1 Abstract

Interferon-induced proteins with tetratricopeptide repeats (IFITs) are innate immune effector molecules that are thought to confer antiviral defense through disruption of protein-protein interactions in the host translation-initiation machinery. However, it was recently discovered that IFITs can directly recognize viral RNA bearing a 5'-triphosphate group (PPP-RNA), which is a molecular signature that distinguishes it from host RNA. Here we report crystal structures of human IFIT5, its complex with PPP-RNAs, and an amino-terminal fragment of IFIT1. The structures reveal a new helical domain that houses a positively charged cavity designed to specifically engage only single-stranded PPP-RNA, thus distinguishing it from the canonical cytosolic sensor of double-stranded viral PPP-RNA, retinoic acid-inducible gene I (RIG-I). Mutational analysis, proteolysis and gel-shift assays reveal that PPP-RNA is bound in a non-sequence-specific manner and requires a 5'-overhang of approximately three nucleotides. Abrogation of PPP-RNA binding in IFIT1 and IFIT5 was found to cause a defect in the antiviral response by human embryonic kidney cells. These results demonstrate the mechanism by which IFIT proteins selectively recognize viral RNA, and lend insight into their downstream effector function.

2.2 Introduction

The innate immune system relies on several germ-line-encoded receptors to distinguish self from non-self molecules in order to mount an appropriate early defense response. During viral infection, non-self molecules are derived from viral genomes generally in the form of double-stranded RNA (dsRNA) or PPP-RNA. The canonical host proteins responsible for sensing or interacting with these foreign nucleic acids include the Toll-like receptors and RIG-I-like receptors (Barbalat et al., 2011). Recently, an unbiased proteomics approach discovered that the IFITs could also directly engage PPP-RNA (Pichlmair et al., 2011).

IFITs are among the most potently expressed proteins of a group of interferon-stimulated genes (ISGs) (Fensterl and Sen, 2011), which are the culmination of virally triggered signaling pathways that lead to the production of interferon (IFN)- α , IFN- β and other cytokines. They are evolutionarily conserved from mammals to fish, with four well-characterized paralogues in humans: IFIT1, IFIT2, IFIT3 and IFIT5, ranging in mass from 54 to 56 kDa. IFITs are composed of tetratricopeptide repeats (TPRs), degenerate helix–turn–helix motifs of 34 amino acids in length, which are usually present in multiple copies as tandem arrays that generate solenoid-type scaffolds well-suited for mediating protein–protein interactions (Main et al., 2003).

IFITs have been implicated in modulating several biological processes, including inhibition of translation initiation, cell proliferation, and migration, in addition to mediating antiviral effects (Fensterl and Sen, 2011). Most of these functions are thought to occur through disruptive protein–protein interactions between IFITs and host cellular factors. Through their TPRs, human IFIT1 and IFIT2 were shown to inhibit key steps during translation initiation by interacting with the ‘e’ or ‘c’ subunits of eIF3 (Guo et al., 2000a; Terenzi et al., 2006). However, the unexpected finding that IFITs can bind RNA suggested a more direct role: after infection or interferon stimulation, it was found that IFITs form large multiprotein complexes with other family members and several different RNA-binding proteins, leading to viral clearance (Pichlmair et al., 2011). Like RIG-I, productive binding of both IFIT1 and IFIT5 were shown to depend on the presence of cytosolic PPP-RNAs (Hornung et al., 2006; Pichlmair et al., 2006; 2011). However, crystallographic and biochemical analyses of RIG-I bound to RNA revealed that it is a dsRNA-specific translocase (Myong et al., 2009), which optimally interacts with blunt-ended PPP-RNA (Jiang et al., 2011; Kowalinski et al., 2011; Luo et al., 2011; Schlee et al., 2009; Shu et al., 2013; Wang et al., 2010). The mechanism by which IFITs recognize PPP-RNA is unknown.

We describe here the crystal structure of full-length human IFIT5 with and without PPP-RNAs, as well as an N-terminal, protease-resistant fragment of human IFIT1 (nIFIT1). The structures reveal a novel arrangement of TPR domains that directly bind PPP-RNA in a non-sequence-specific manner and, to our knowledge, represent the first example of a TPR protein bound to a nucleic acid ligand. Structure-guided biochemical analysis of IFIT5 and IFIT1 indicated that only single-stranded RNA (ssRNA) can be accommodated within the protein, which undergoes a compaction upon binding. Finally, functional analysis in human embryonic kidney

(HEK) cells reveals a reduction of viral replication only in the presence of proper PPP-RNA binding by IFIT1 or IFIT5.

2.3 Results

2.3.1 Crystal structures of IFIT5 and nIFIT1

We crystallized and determined the structures of full-length human IFIT5 (residues 1–482) at 2.1 Å resolution and an N-terminal fragment of IFIT1 (residues 7–279) at 1.9 Å resolution using single-wavelength anomalous diffraction (**Table 2.1**). The structure of IFIT5 reveals a helical domain with approximate dimensions of 80 Å × 55 Å × 40 Å (**Fig. 2.1 A and B**). In most multi-TPR-containing proteins, such as O-linked N-acetylglucosamine transferase (OGT), the relationship between successive TPRs is regular and repeating, such that they form open-ended superhelical structures with distinct convex and concave surfaces (Jínek et al., 2004; Main et al., 2005). In IFIT5, of its total 24 α -helices, 18 form canonical TPRs (TPRs 1–9; **Fig. 2.1 A**), whereas the remaining 6 helices intervene between the TPRs such that the regular repeating relationship between them is disrupted. This results in the formation of three distinct bundles of TPRs (subdomains I, II and III) oriented with respect to one another to give the overall protein a relatively closed clamp-shaped structure (**Fig. 2.1 A and 2.5 A**).

The topology of subdomain I is unusual in that its two canonical TPRs (α 3 to α 6) are capped off on both ends by helices α 1 and α 2, preventing its further propagation into a superhelix. This is facilitated by a connecting 17-residue loop (L1) containing a highly conserved Cys-His-Phe-Thr-Trp pentapeptide motif that is invariant among nearly all of the IFIT proteins (**Fig. 2.2**), and forms a single turn of a helix that packs against the concave inner face of subdomain I (**Fig. 2.3 A**). This same arrangement of subdomain I is also found in the structure of nIFIT1 (root mean squared deviation, 1.4 Å; **Fig. 2.4**) and is probably a defining characteristic of all IFIT proteins given the high TPR and sequence conservation in subdomains I and II (**Fig. 2.2**).

The remainder of the IFIT5 structure forms a superhelix encompassing subdomains II and III, as well as a pair of extended non-TPR helices (α 15 and α 16) that form a pivot point between the latter two subdomains (**Fig. 2.1 B**). Subdomain II forms a canonical four-TPR-repeat domain in which, notably, its first helix (α 7) interacts with subdomain I in a manner reminiscent of TPR protein–ligand interactions observed previously (Zhang and Chan, 2007) (**Fig. 2.3 B and C**). This leads to the concave surface of subdomain II forming one wall of a large cavity in the center of the

protein closed off at its base by helix $\alpha 2$ (**Fig. 2.1 B**). The same TPR–ligand relationship between subdomains I and II is also maintained in the nFIT1 structure (**Fig. 2.4**).

The rest of the cylindrical cavity is created by the intervening pivot helices and the N-terminal TPRs from subdomain III. Subdomain III begins with two typical TPRs followed by an interrupting helix ($\alpha 21$), which inverts the direction of the final TPR9 such that it forms an S-shaped appendage at the carboxy terminus with two potential ligand-interacting concave surfaces (**Fig. 2.1 B and Fig. 2.5**). The deep pocket formed by this atypical arrangement of TPRs is approximately 28 Å deep by 15 Å wide, and is lined with an expansive collection of positively charged residues well-suited for the accommodation of nucleic acid (**Fig. 2.1 C**).

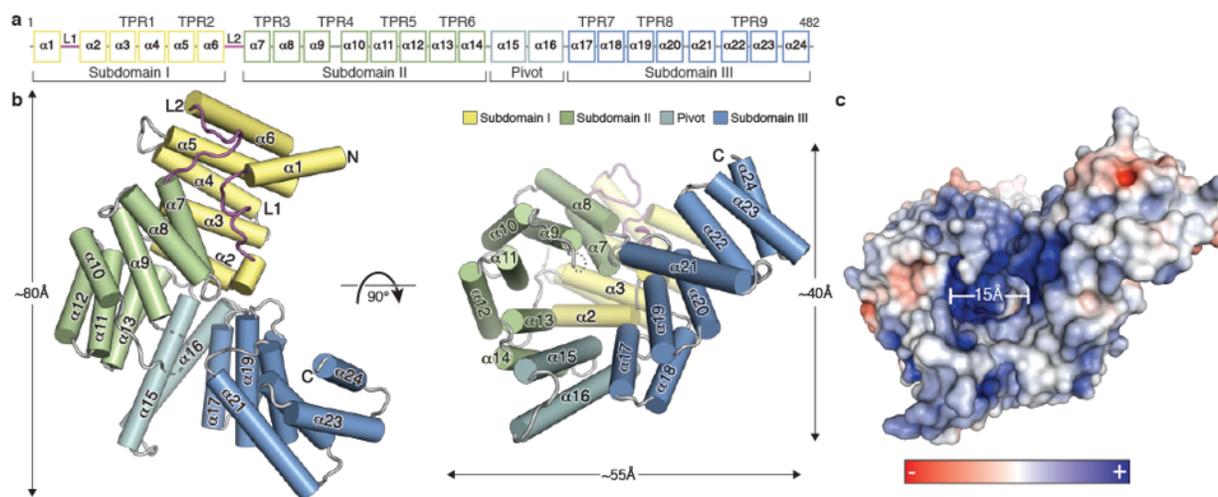


Figure 2.1 Structural overview of human IFIT5

(A) Secondary structure, TPR motif and subdomain organization of IFIT5. (B) Orthogonal views of IFIT5 with helices represented as cylinders. (C) Surface representation of IFIT5 coloured by electrostatic potential (using APBS) from $-5kT e^{-1}$ (red) to $+5kT e^{-1}$ (blue).

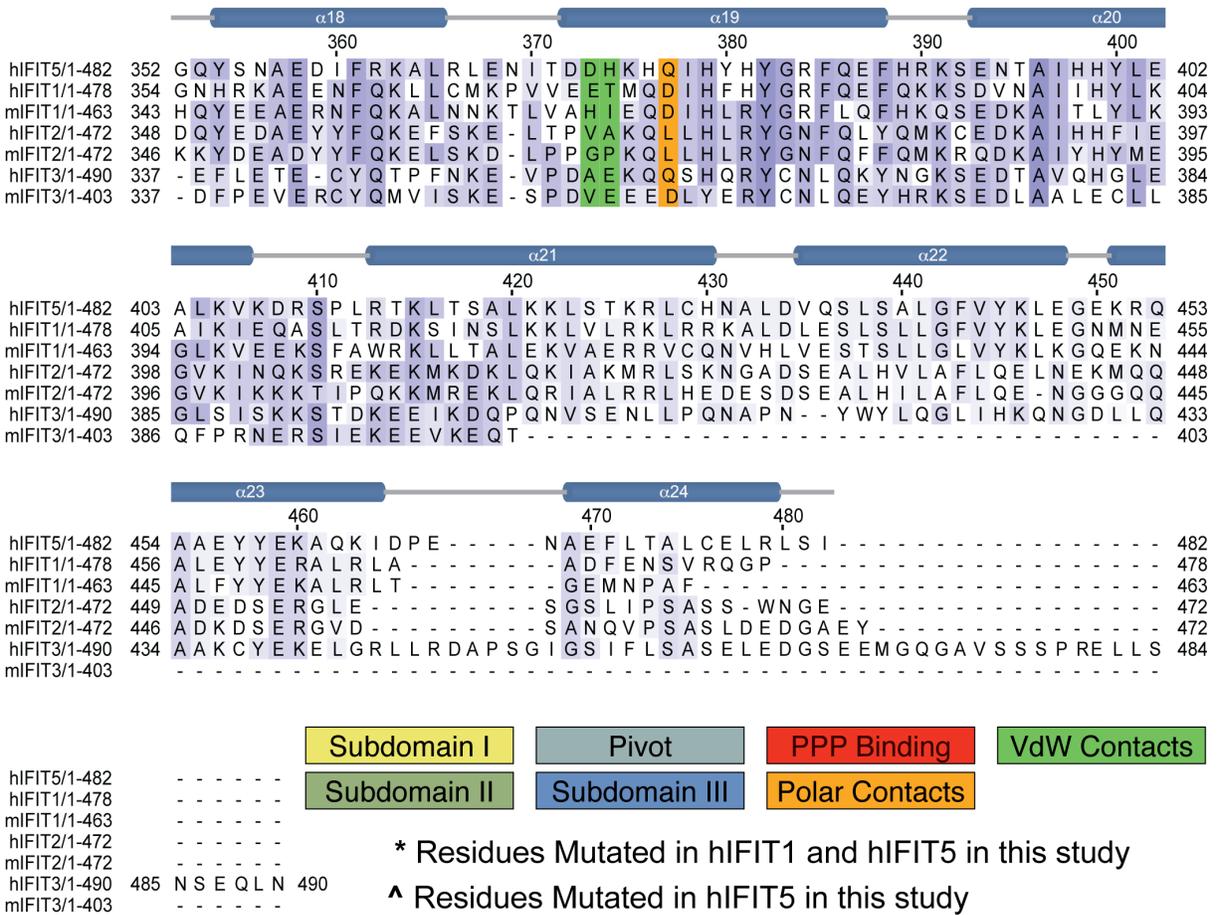


Figure 2.2 Sequence alignment coloured by BLOSUM62 conservation score

The secondary structure of IFIT5 is depicted, and coloured as in figures 1 and 2 in the main text. The numbering above is that of human IFIT5. RNA interacting residues forming polar contacts (hydrogen bonds) are shaded in orange, and those forming hydrophobic interactions (VdW, Van der Waals) are coloured green. Residues critical for PPP recognition are coloured red.

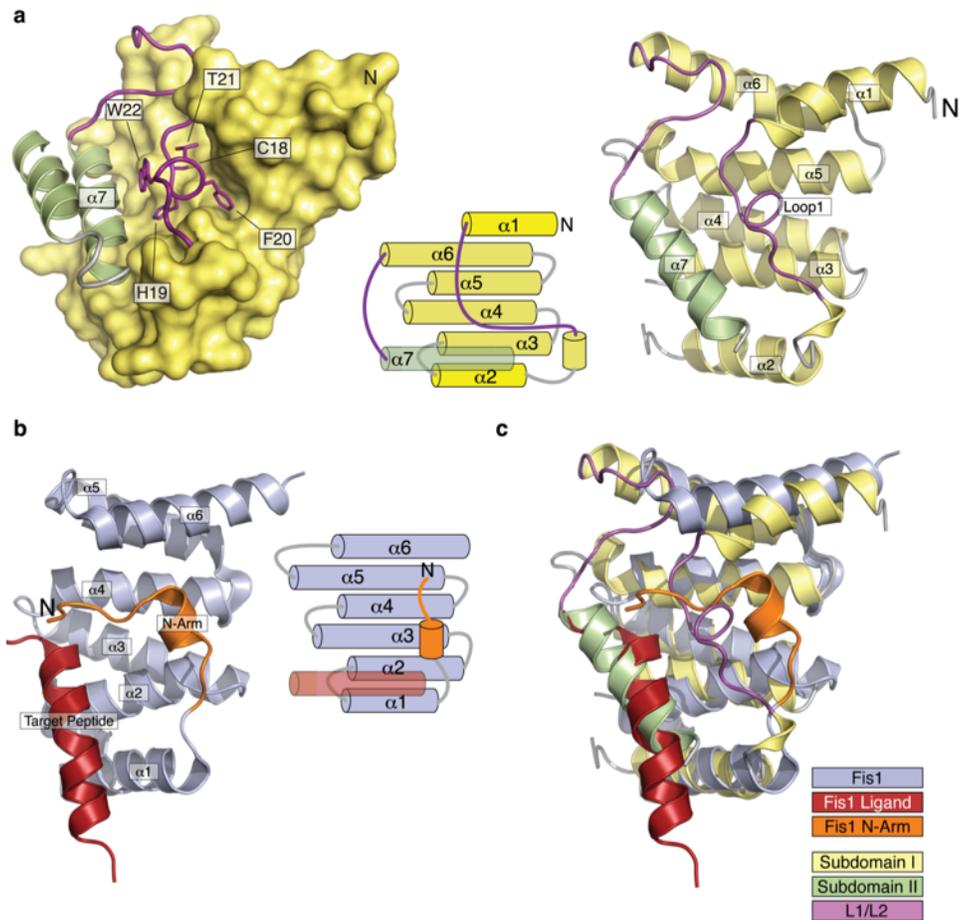


Figure 2.3 Comparison of IFIT5 Subdomain I to other TPR proteins

(A) Left, close-up view of the interactions between Subdomain I (yellow molecular surface) to Subdomain II (green ribbons). Residues from the CHFTW motif are shown as purple sticks. Middle, topology diagram of Subdomain I; Right, Cartoon representation of Subdomain I. (B) Cartoon representation of Fis1, a TPR containing protein involved in mitochondrial fission (Zhang and Chan, 2007). Both Subdomain I and Fis1 are composed of a 6-helix bundle with two central canonical TPRs (α 3- α 6 in IFIT5 and α 2- α 5 in Fis1) flanked on either side by a capping helix. Additionally, both structures have an N-terminal motif (N-arm in Fis1 and Loop1 in IFIT5) that sits in the concave binding surface, mediating an interaction with another α -helix. In Fis1, the capping helices are the first and last helices of the bundle (α 1/ α 6), whereas in IFIT5 the capping helices are the first and second helices in the primary sequence (α 1/ α 2). (C) Superposition of Subdomain I and Fis1 showing the similarities in target recognition. Backbone RMSD is 2.3 Å.

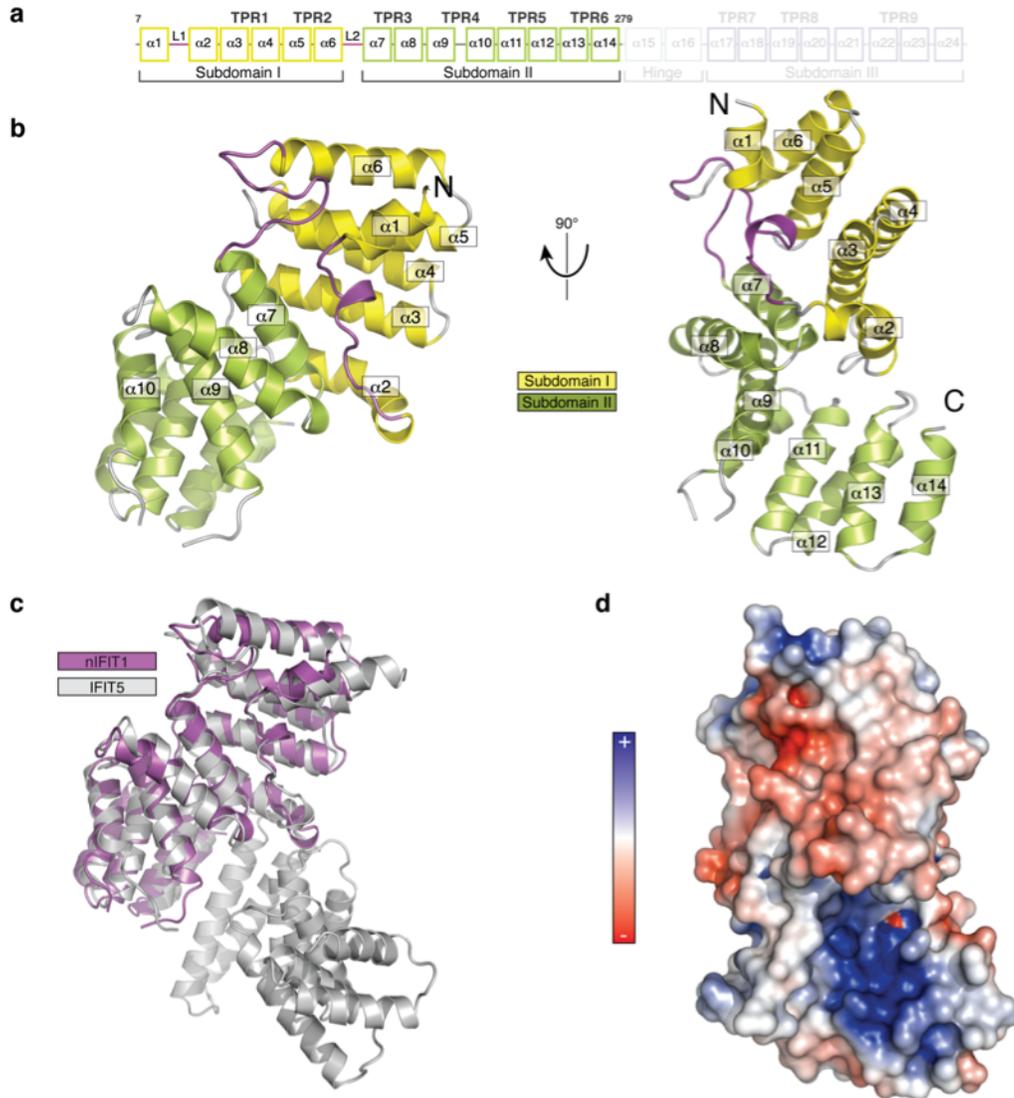


Figure 2.4 Crystal structure of nIFIT1 (IFIT1 residues 7-279)

(A) Secondary structure, TPR motif and subdomain organization of IFIT1. Faded boxes indicate the corresponding region from IFIT5 that is missing in nIFIT1. (B) Ribbon diagram of the nIFIT1 fragment. Although full-length IFIT1 is a dimer in solution, nIFIT1 migrated as a monomer on gel filtration. The crystal structure, however, did contain two molecules in the asymmetric unit. (C) Structural alignment of nIFIT1 and IFIT5. (D) Surface representation of nIFIT1 coloured by electrostatic potential. Like IFIT5, the concave surface of Subdomain II (which would form one wall of the RNA binding pocket) is positively charged. The orientation is identical as in (B) right.

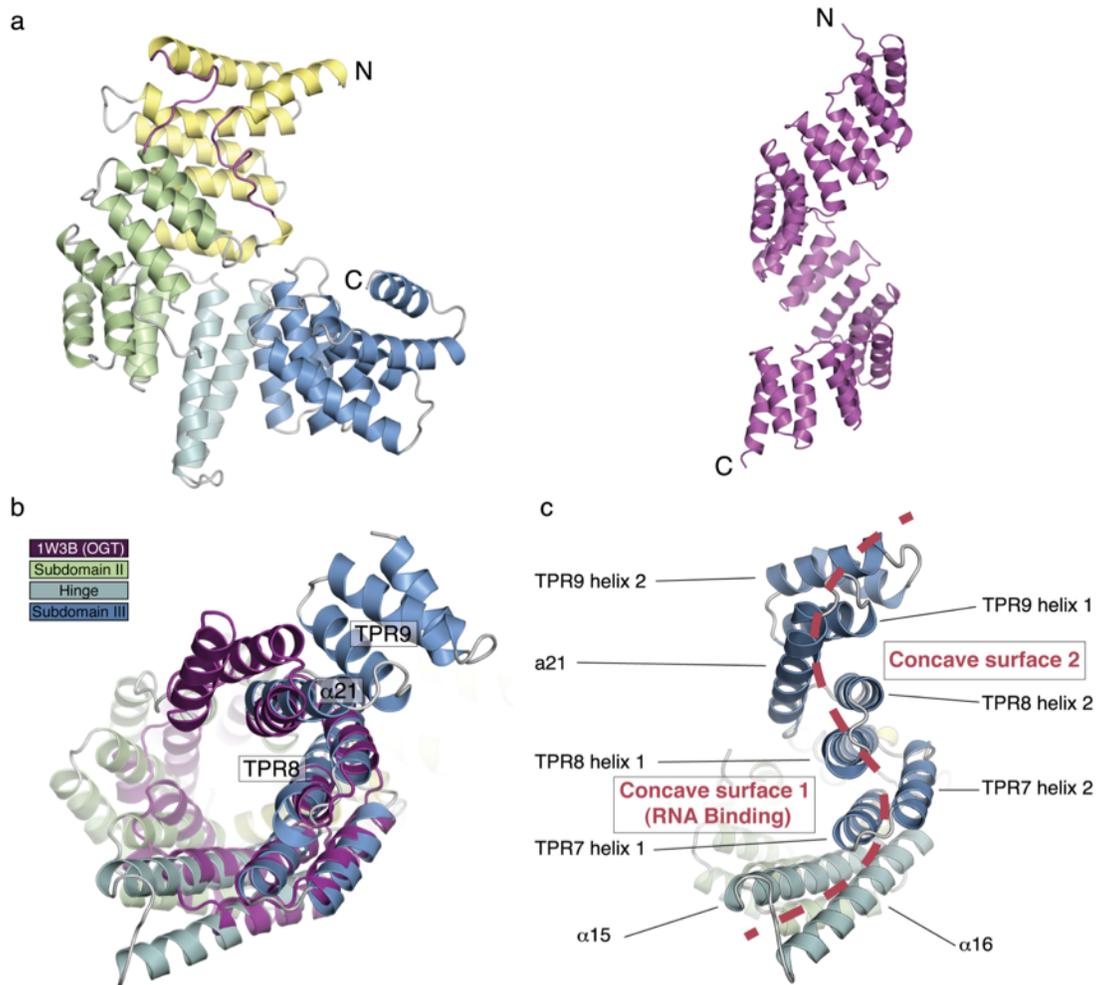


Figure 2.5 Comparison of the IFIT5 superhelix to O-linked GlcNac transferase (OGT)

(A) Comparison of IFIT5 and the TPR domain of OGT (purple), the canonical multi-TPR containing protein previously used for homology modelling (pdb 1w3b, ((Jínek et al., 2004; Main et al., 2005)). (B) OGT (residues 180-365) was aligned against IFIT5 residues 214-431 with an RMSD of 2.8 Å. The superhelix of IFIT5 is interrupted by $\alpha 21$ between TPR8 and TPR9. (C) S-hook shape at the C-terminal end of IFIT5.

2.3.2 IFIT5 specifically binds PPP-RNAs

To understand the structural basis for RNA binding by IFIT5, we *in vitro* transcribed 5'-triphosphate-bearing, short oligonucleotides of cytidine, uridine and adenosine, purified each PPP-RNA in complex with IFIT5 and determined their structures at resolutions of 1.86 Å (oligo-C), 2.0 Å (oligo-U) and 2.5 Å (oligo-A) using molecular replacement with the unliganded structure (**Table 2.2**). All of the structures were similar, and therefore we initially describe the general features of the IFIT5-oligo-C complex as it was the highest resolution structure. Difference Fourier maps revealed strong positive electron density within the central positively charged pocket from which the 5'-triphosphate and the first four nucleotides of the RNA could be reliably modelled (**Fig. 2.6 A and Fig. 2.7 A**). The 5'-triphosphate group is nestled deep within the pocket and makes a multitude of electrostatic interactions with protein side chains from helix α_2 (Glu 33, Thr 37 and Gln 41) located at the very base of the pocket, and residues from the concave inner surface of subdomain II (Lys 150, Tyr 250 and Arg 253) (**Fig. 2.6 C, left**). Arg 187 in IFIT1 was previously identified to be required for RNA interaction (Pichlmair et al., 2011); the homologous residue in IFIT5 (Arg 186) makes a weak salt-bridge with the α - and β -phosphates, and van der Waals contacts with the first ribose moiety. These RNA-interacting residues are for the most part conserved in sequence and structure between IFIT5 and IFIT1 (**Fig. 2.2 and Fig. 2.8**), the only IFITs that have been shown to bind PPP-RNA with strong affinity (Pichlmair et al., 2011). One notable exception is Thr 37, which is replaced in IFIT1 by Arg 38, suggesting slight differences in RNA recognition between IFIT1 and IFIT5. Conversely, in IFIT3, which is known to not bind PPP-RNA (Pichlmair et al., 2011), Tyr 250 is substituted with a negatively charged residue, Asp 242, and Arg 186 with His 182 (**Fig. 2.2**), both of which would interfere directly with RNA binding.

Interestingly, a metal ion that bridges the α - and γ -phosphates also seems to be an integral part of PPP-RNA recognition, and is coordinated Glu 33 (**Fig. 2.6 C, left and Fig. 2.9**). On the basis of ligand distances and geometry, the ion is probably Mg^{2+} from the *in vitro* transcription reactions, but could potentially also be Na^+ (a component of the crystallization buffer). It is unlikely that capped messenger RNA (mRNA) can be accommodated within this pocket owing to size constraints. In addition, given the critical interactions made with the γ -phosphate and the metal ion, the pocket is unlikely to accept 5'-monophosphorylated or 5'-hydroxylated RNA with

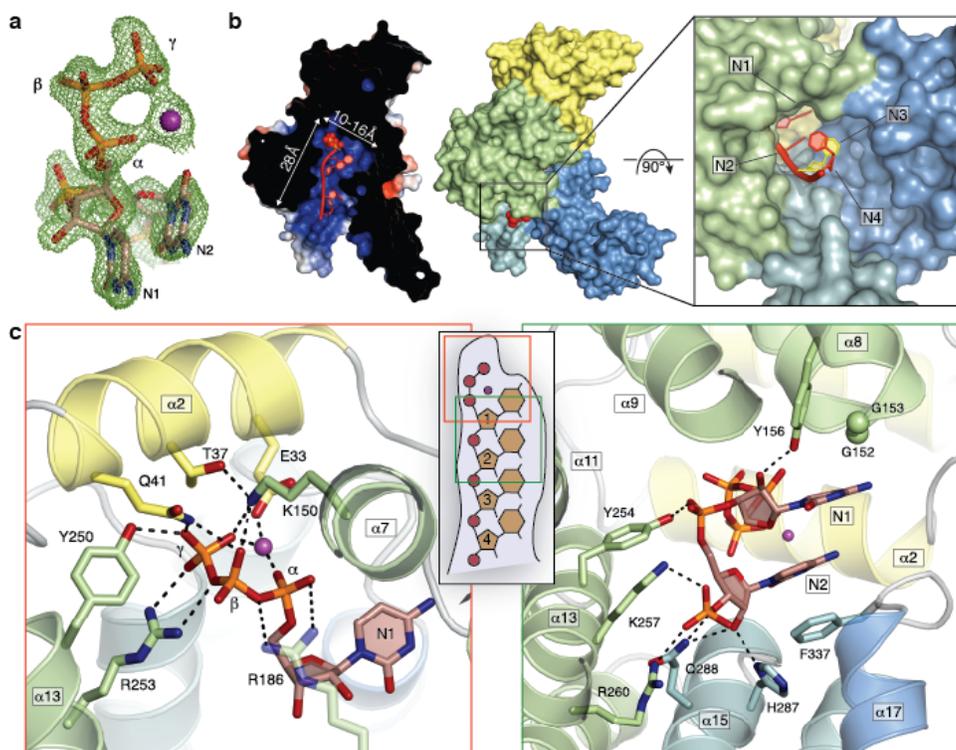


Figure 2.6 Structure of IFIT5 bound to PPP-RNA

(A) F_o-F_c electron density map of the triphosphate and first two nucleotides contoured at 3.5σ before inclusion of RNA into the model. The metal ion is indicated with a purple sphere. (B) Left: cross-section of the complex coloured by surface electrostatic potential. The triphosphate is shown as spheres and RNA nucleotides are shown in red. Middle: surface representation of IFIT5 bound to PPP-RNA coloured by subdomain. Protruding RNA is shown as red spheres. Right: close-up view looking down the axis of the RNA binding pocket. (C) Close-up view of the residues making specific contacts with the triphosphate group (left) and the first two nucleotides, N1 and N2 (right). Helices are coloured according to the subdomain to which they belong.

considerable affinity. Thus, the structure of the IFIT5 TPR domains have evolved to specifically engage PPP-RNA, and in doing so, distinguish between self and non-self nucleic acids.

Following the 5'-triphosphate end of the RNA, the first two nucleotides (N1, N2) are stably bound along the pocket before the third and fourth nucleotides (N3, N4) begin to protrude from the mouth of the pocket (**Fig. 2.6 B, right**). Well-defined density is observed for the phosphodiester backbone and ribose sugars (**Fig. 2.7 A**), which also form several specific interactions with the protein (**Fig. 2.6 C, right**). In particular, the 5'-phosphate of N2 hydrogen bonds with Tyr 254, and the 5'-phosphate of N3 makes a salt-bridge with Arg 260 and Lys 257, and hydrogen bonds with Gln 288. The 5'-phosphate of N4 interacts with Arg 294, and weak electron density was observed for the 5'-phosphate of a fifth nucleotide (**Fig. 2.7 A**). The 2'-hydroxyl of the ribose sugars also make specific interactions with the protein, but in this case,

interactions that are dependent on the sugar pucker. N1 adopts a *C2'-endo* conformation (commonly found in B-form double-stranded DNA; **Fig. 2.10**) and hydrogen bonds with Tyr 156 (**Fig. 2.6 C, right**), whereas N2 and N3 are *C3'-endo* (as found in A-form dsRNA) and interact through their 2'-hydroxyls with His 287 and Gln 288, and Arg 294 and Asp 343, respectively (**Fig. 2.6 C, right and Fig. 2.11 A**).

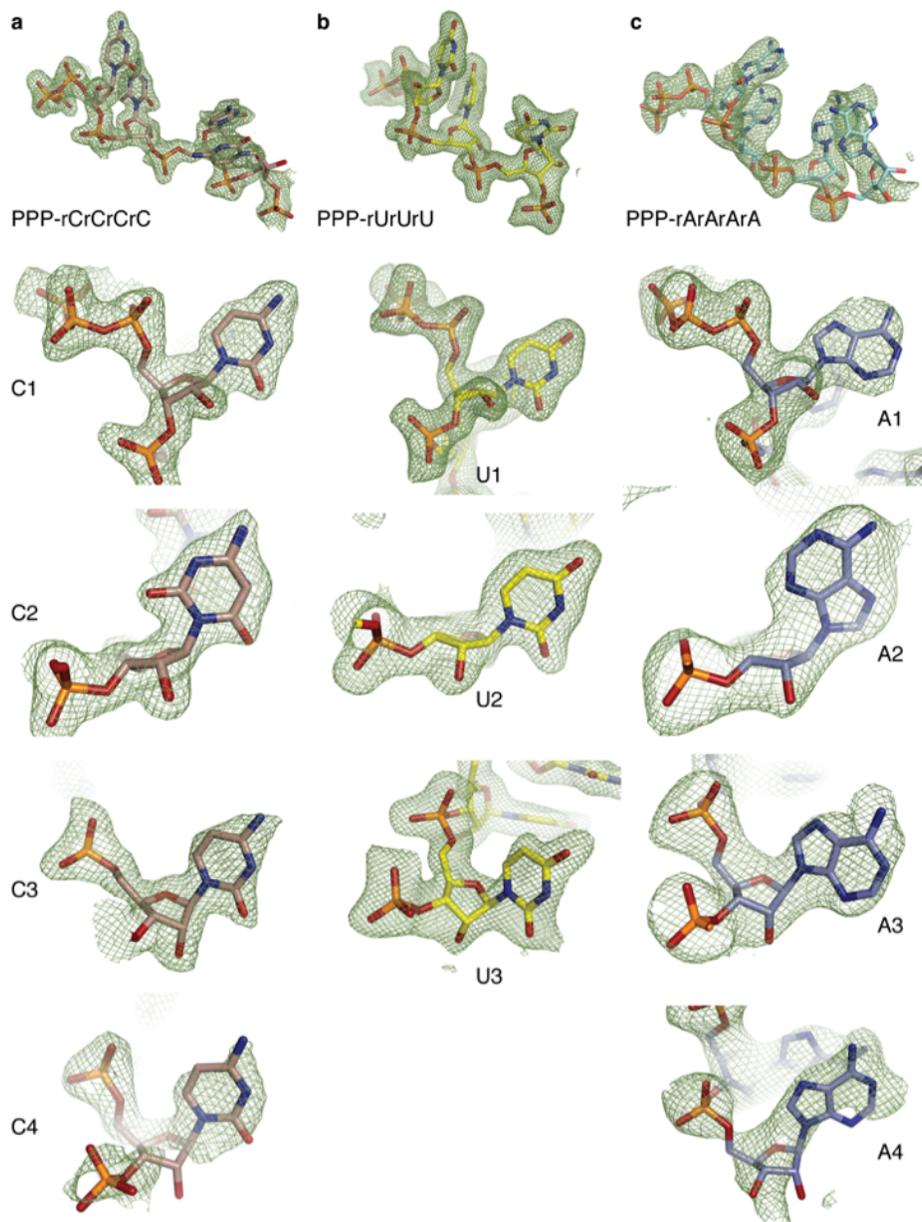


Figure 2.7 F_0 - F_c maps of RNA in the three IFIT5-PPP-RNA co-crystal structures

F_0 - F_c maps of the nucleotides used to build the oligo-C (a), oligo-U (b), and oligo-A (c) models contoured at 2σ before inclusion of any RNA in the model. Note that the base of C2 adopts both *syn* and *anti* conformations, and the base of A2 adopts a *syn* conformation with respect to the sugar.

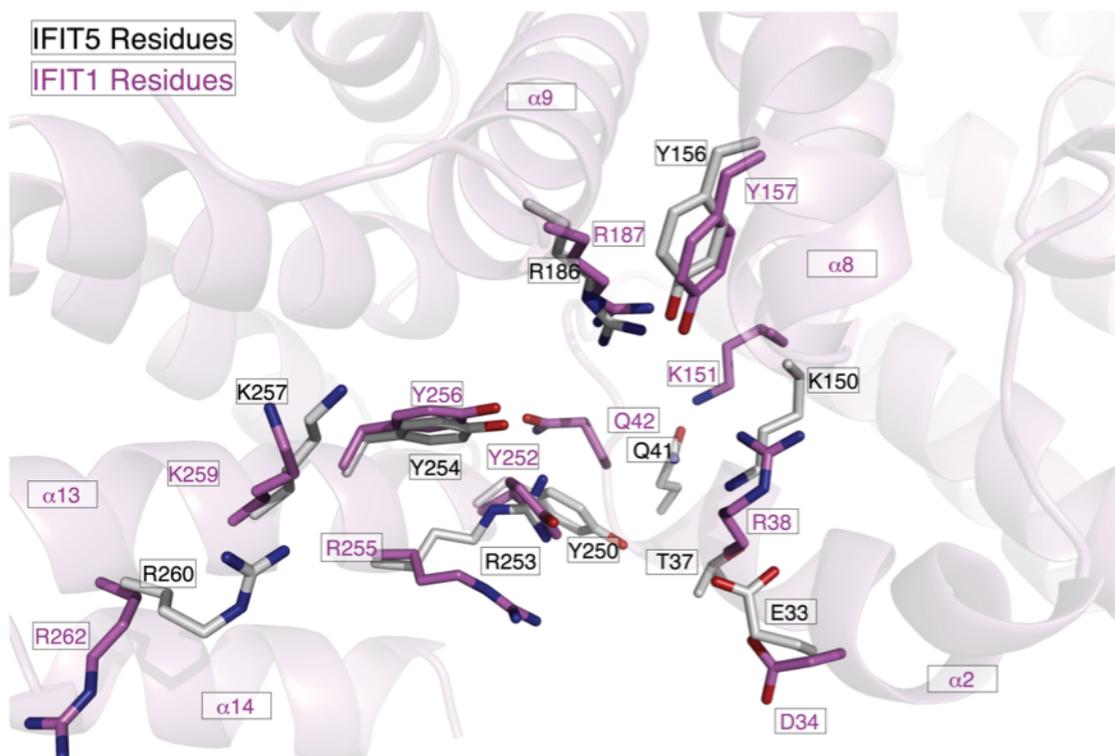


Figure 2.8 Structural alignment of nIFIT1 and PPP-RNA bound IFIT5

Structural alignment of potential PPP-RNA interacting residues from IFIT1 using binding site residues from IFIT5 as the reference. The crystal structure of nIFIT1 is depicted in transparent cartoon, with the helices labelled as in Fig. 2.5. For clarity, only the corresponding residues from IFIT5 are displayed, with the main chain hidden. The orientation is similar to Fig. 2.6 C, right. The PPP-RNA from IFIT5 is not shown for clarity.

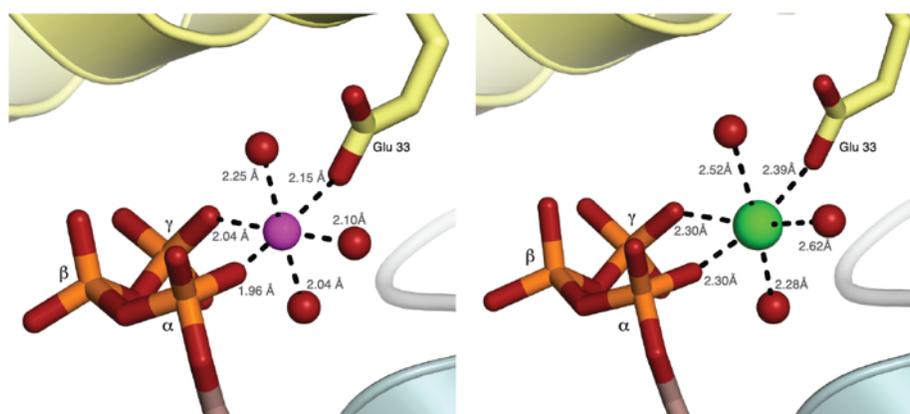


Figure 2.9 Close-up view of the metal ion-binding site in IFIT5

Six atoms – two oxygen atoms from the α and γ phosphates, one carboxylate oxygen from Glu 33 and 3 waters (red spheres) – ligate the ion (Mg^{2+} , purple and Na^{+} , green) in an octahedral geometry. The refined distances between the ion and its six ligands suggest that the metal is likely magnesium in the oligo-C complex (purple, top), and sodium in the oligo-U complex (green, bottom). Assignment of the metal was based on typical metal-ligand distances of magnesium (~ 2.1 Å) and sodium (~ 2.4 Å) (Harding and Nowicki, 2010).

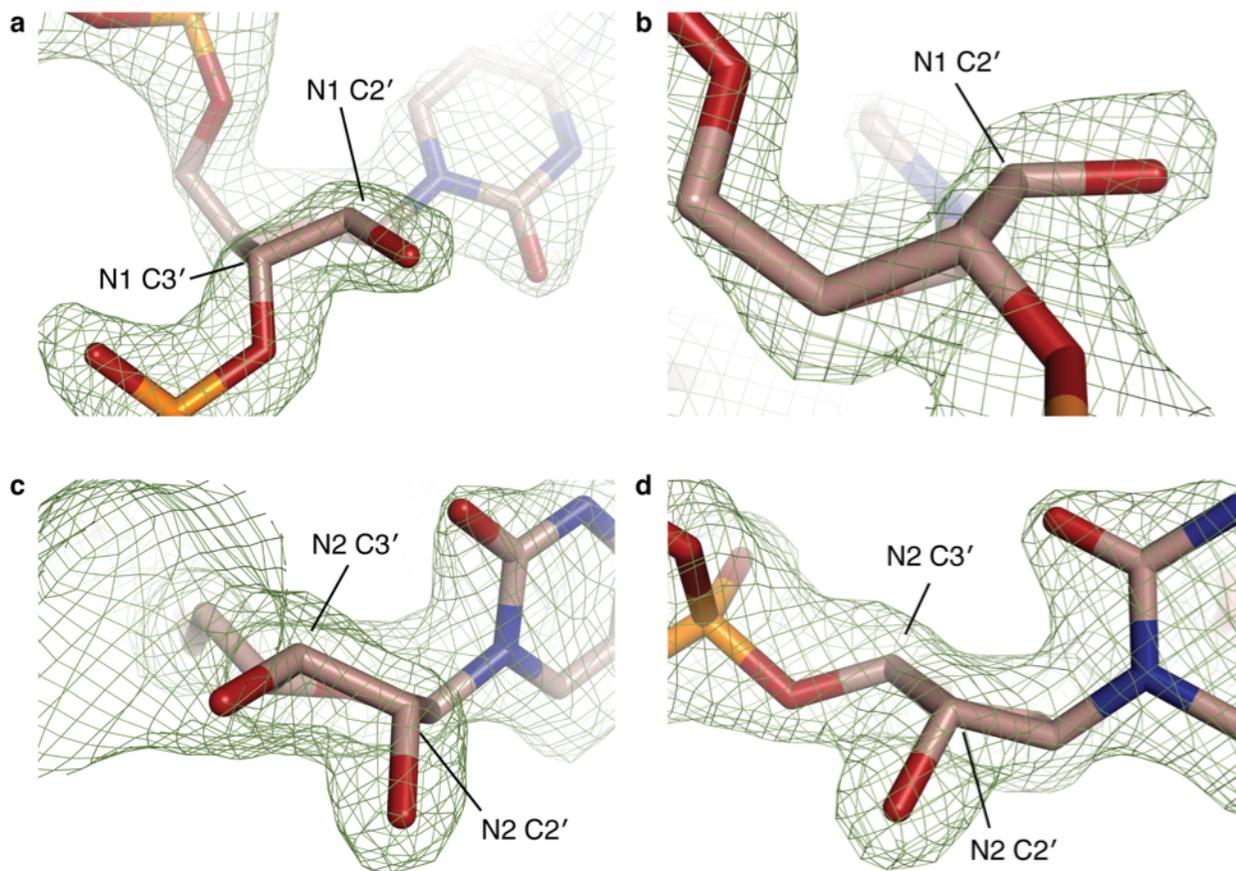
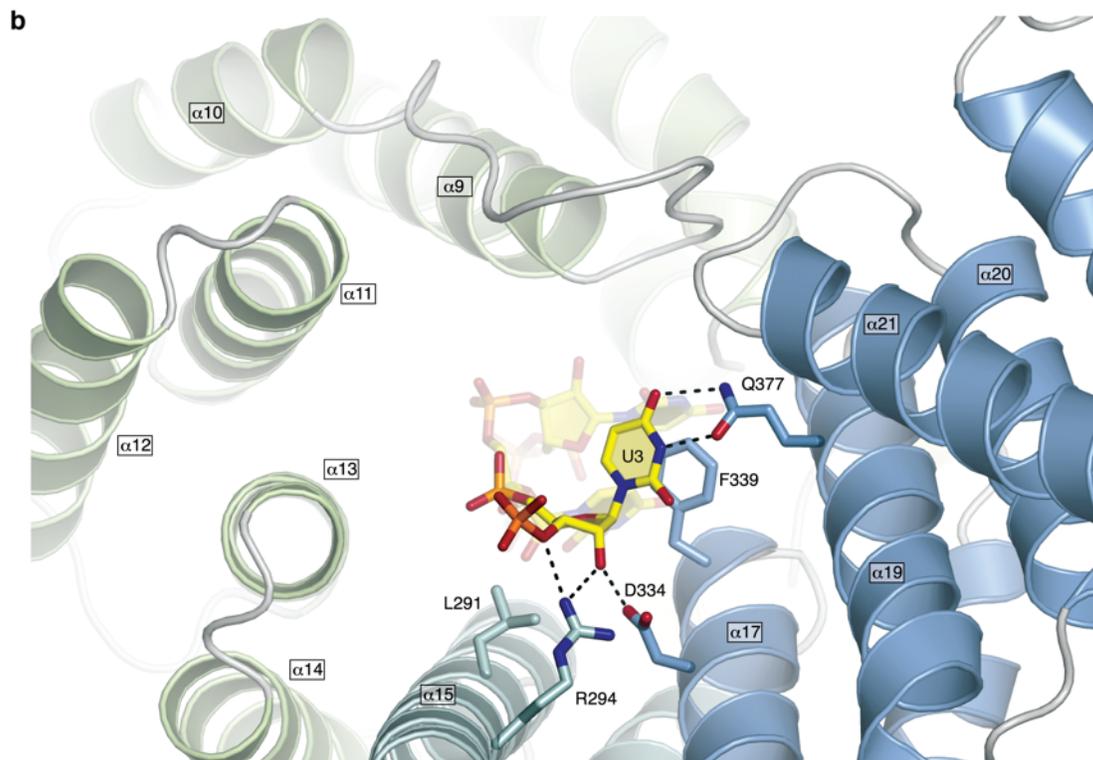
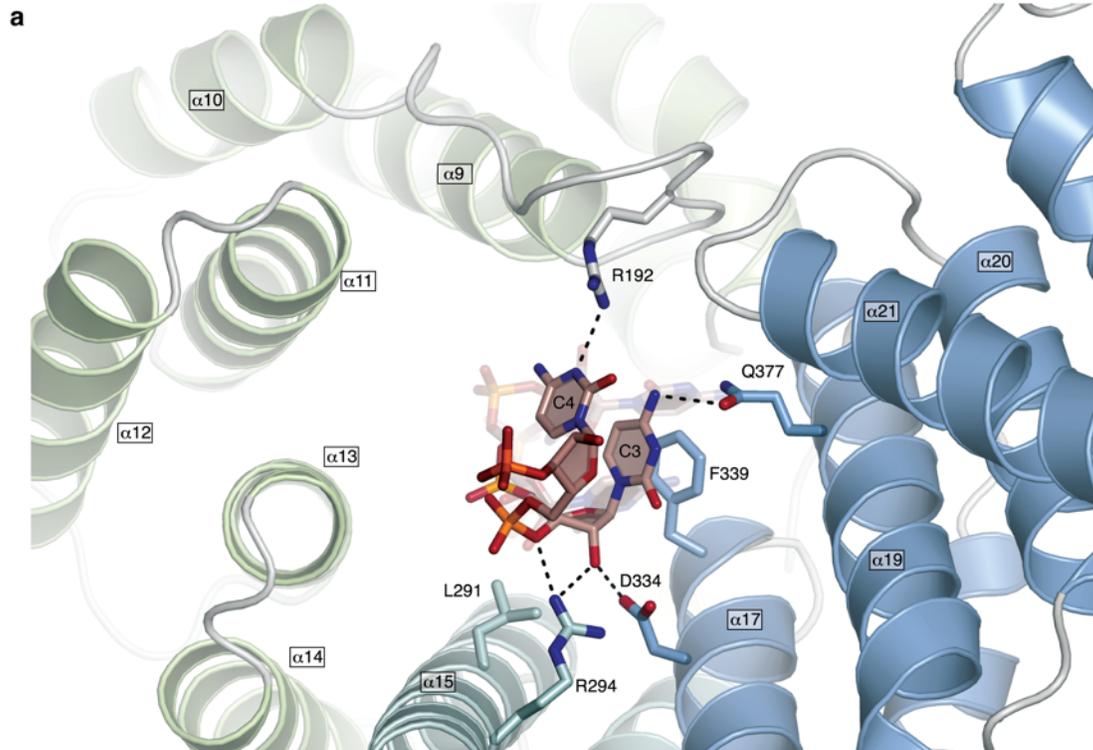


Figure 2.10 Sugar pucker of nucleotides 1 and 2 (N1 and N2)

(A and B) Orthogonal views of the ribose moiety at N1. **(C and D)** Orthogonal views of the ribose moiety at N2. The conformation of the first nucleotide of the PPP-RNA is the less common, DNA-like conformation. Thermodynamically stable conformations of the ribose sugar involve 4 atoms in a plane, and one atom usually out of the plane. In RNA, the 3'-carbon is usually out of the plane (*C3'-endo*), whereas here at N1, it is found in the *C2'-endo* conformation. In contrast, the sugar pucker of N2 is the typical *C3'-endo*. In (C), the 5'-phosphate of N3 is not shown for clarity.



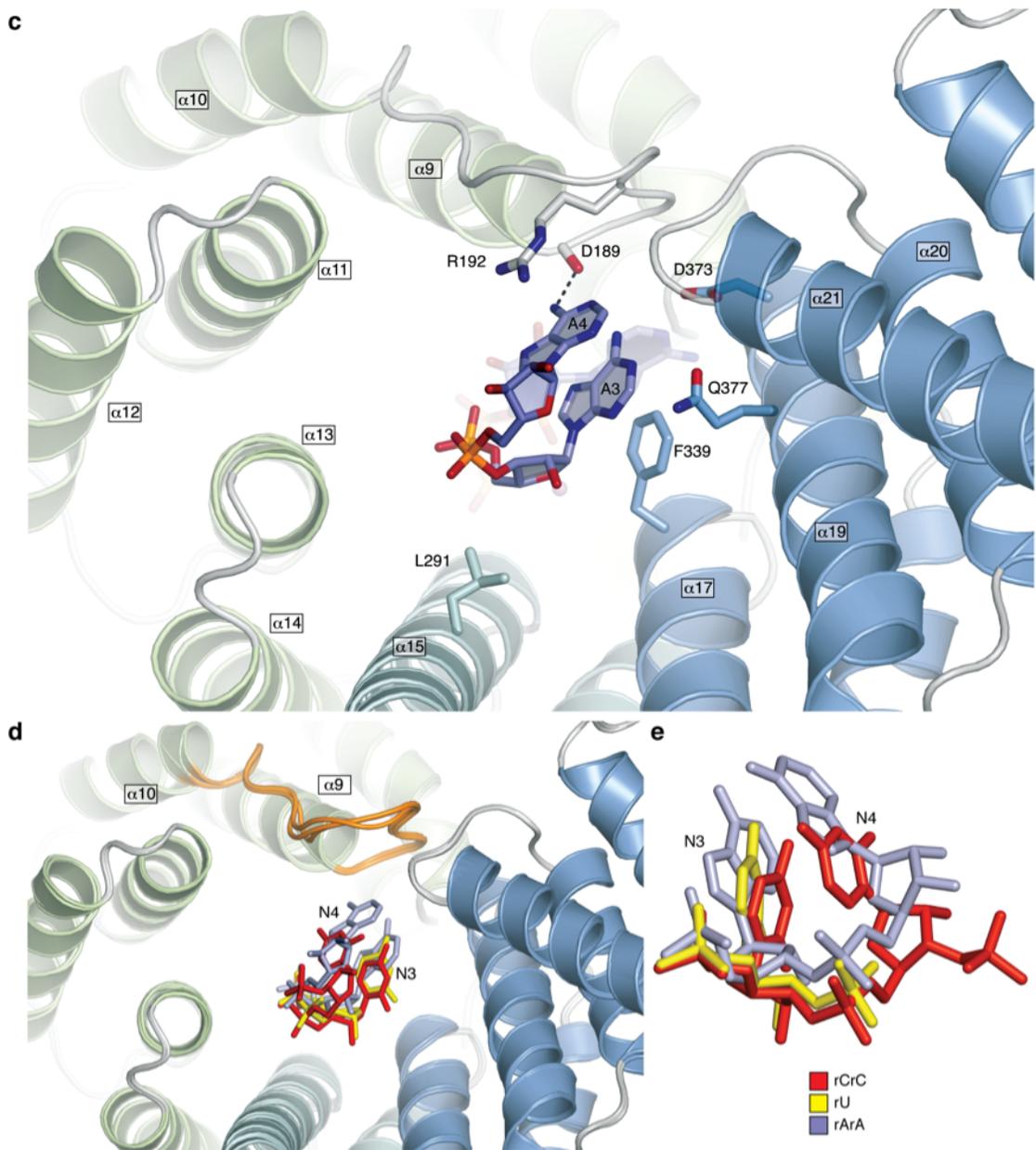


Figure 2.11 Interaction between IFIT5 and PPP-RNA at N3 and N4

IFIT5 in complex with (A) oligo-C, (B) oligo-U, and (C) oligo-A complexes. (D) The mobile loop (orange) between $\alpha 9$ and $\alpha 10$ that interacts with the base of N3 and becomes more ordered upon RNA binding. The residues in this loop adopt different conformations depending on the base identity, enabling IFIT5 to accommodate variable RNA sequences at these positions. (E) Superposition of RNA bases at positions 3 and 4.

2.3.3 PPP-RNA recognition is non-sequence specific

To investigate the potential for sequence-specific interactions at the 5'-end, we compared the crystal structures of IFIT5 in complex with the different RNAs. In both the oligo-C and oligo-U complexes, the pyrimidine base at position 1 is abutted from the top by van der Waals interactions with Tyr 156 and two glycine residues from the loop of TPR3 (between $\alpha 7$ and $\alpha 8$), and from the bottom by non-specific stacking interactions against the second base, which in turn stacks with Phe 337 (**Fig. 2.6 C, right**). Notably, the first two bases do not make any specific hydrogen bonds with protein residues and there is ample space adjacent to the pyrimidine ring edges, suggesting that the larger purine bases can also be easily accommodated (**Fig. 2.12 B**). The structure of the oligo-A complex confirms this notion and reveals that the adenine rings reach further out into the periphery making additional non-specific van der Waals contacts with Thr 371, His 374 and Phe 339, which were absent with the pyrimidine bases (**Fig. 2.12 B-D**).

The remaining bases stack against Phe 339 in a manner analogous to that observed for the first two bases and interact with a mobile loop from TPR4 (**Fig. 2.11**). Thus, IFIT5 seems to have evolved the capacity to accommodate any 5'-PPP-RNA sequence that may potentially be present in a viral genome.

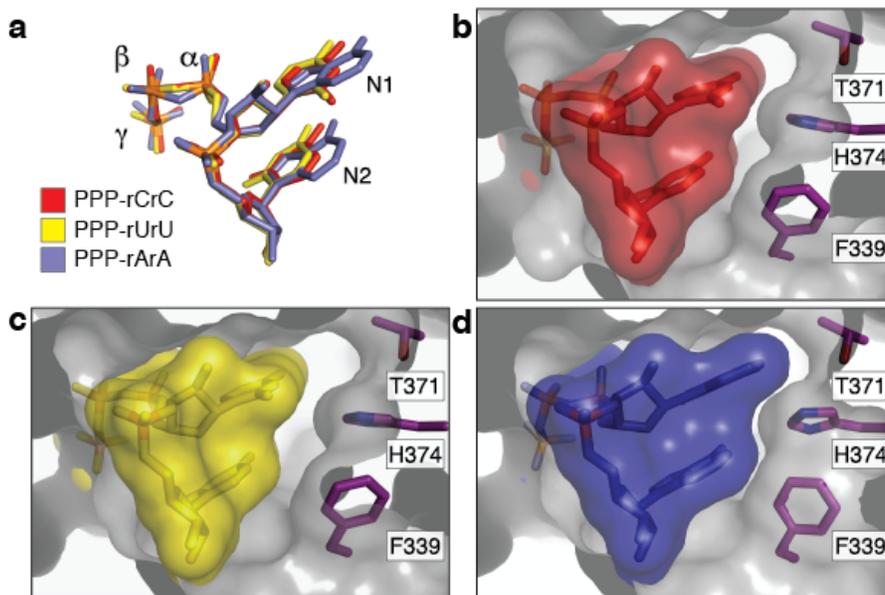


Figure 2.12 The interaction between IFIT5 and PPP-RNA is non-sequence specific

Close up of the RNA binding pocket in an orientation similar to that of Fig. 2.6 B, right. **(A)** Alignment of the first two nucleotides from the three IFIT5-RNA complexes. **(B and C)** Surface and stick representation of the first two nucleotides within the IFIT5-oligo-C and IFIT5-oligo-U complexes. The protein surface is depicted as a transparent grey cutaway. **(D)** IFIT5-oligo-A complex.

2.3.4 PPP-RNA binding involves a conformational change

Because the RNA-binding site in IFIT5 is a deep and narrow pocket (**Fig. 2.6 B, left**), the means by which RNA enters is unclear. Superposition of the RNA-bound and -free forms of IFIT5 reveal that the RNA-bound state is more compact, with the largest motions occurring at the pivot helices between subdomain III and the rest of the protein (**Fig. 2.13 A and Fig. 2.14**). These motions position several key residues from the different subdomains for optimal interaction with the RNA. Moreover, limited proteolysis of IFIT5 in the presence and absence of RNA supports the notion of compaction and stabilization of the protein in the RNA-bound form (**Fig. 2.13 B**).

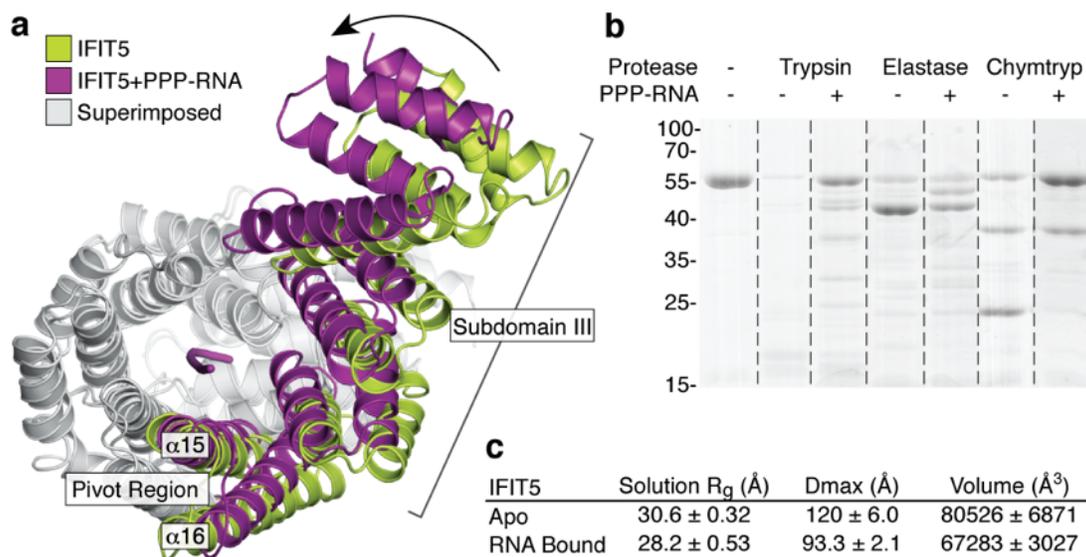


Figure 2.13 IFIT5 undergoes a conformational change upon binding PPP-RNA

(A) Comparison of IFIT5 bound to PPP-RNA (magenta) and the unbound form (green). Superimposed regions are coloured light grey. (B) SDS-PAGE gel of limited protease digestion of IFIT5 in the absence and presence of RNA taken from experiments at the 15 min time point. Dashed lines are to show that these lanes are cropped from a larger set of gels. (C) Summary of SAXS results. Data are the average of 3 measurements (at 3 different concentrations) ± standard deviation. Complete SAXS data is presented in Fig. 2.15.

To better understand the nature of the conformational change upon RNA binding, we used small-angle X-ray scattering (SAXS) measurements, which provide information on macromolecular size, state and flexibility directly in solution (Putnam et al., 2007; Rambo and Tainer, 2011). SAXS analysis revealed reductions in the radius of gyration (R_g , ~2.5 Å), the maximum dimension (D_{max} , ~25 Å) and the volume (~14,000 Å³) of the protein upon addition of RNA (**Fig. 2.13 C and Fig. 2.15**). The scattering curves show good agreement between solution (R_g , 28.2 Å) and crystal structure (R_g , 27.5 Å) for the RNA-bound form (**Fig. 2.15 H**), in contrast to the unliganded form, which displays considerable differences (solution R_g , 30.6 Å; crystal

structure Rg, 28.3 Å). This suggests that in solution, the unliganded protein is either more open or possibly flexible. To discern between these possibilities, we subjected the SAXS data to a Porod–Debye analysis, which provides information on the degree of flexibility present in the scattering sample (Rambo and Tainer, 2011). For both unliganded and RNA-bound IFIT5, the Porod–Debye plot showed characteristic plateaus that indicate the presence of distinct conformations for both species (**Fig. 2.15 K and L**). Thus, unliganded IFIT5 probably exists in a more open conformation in solution than that observed in the crystal structure, facilitating RNA entry.

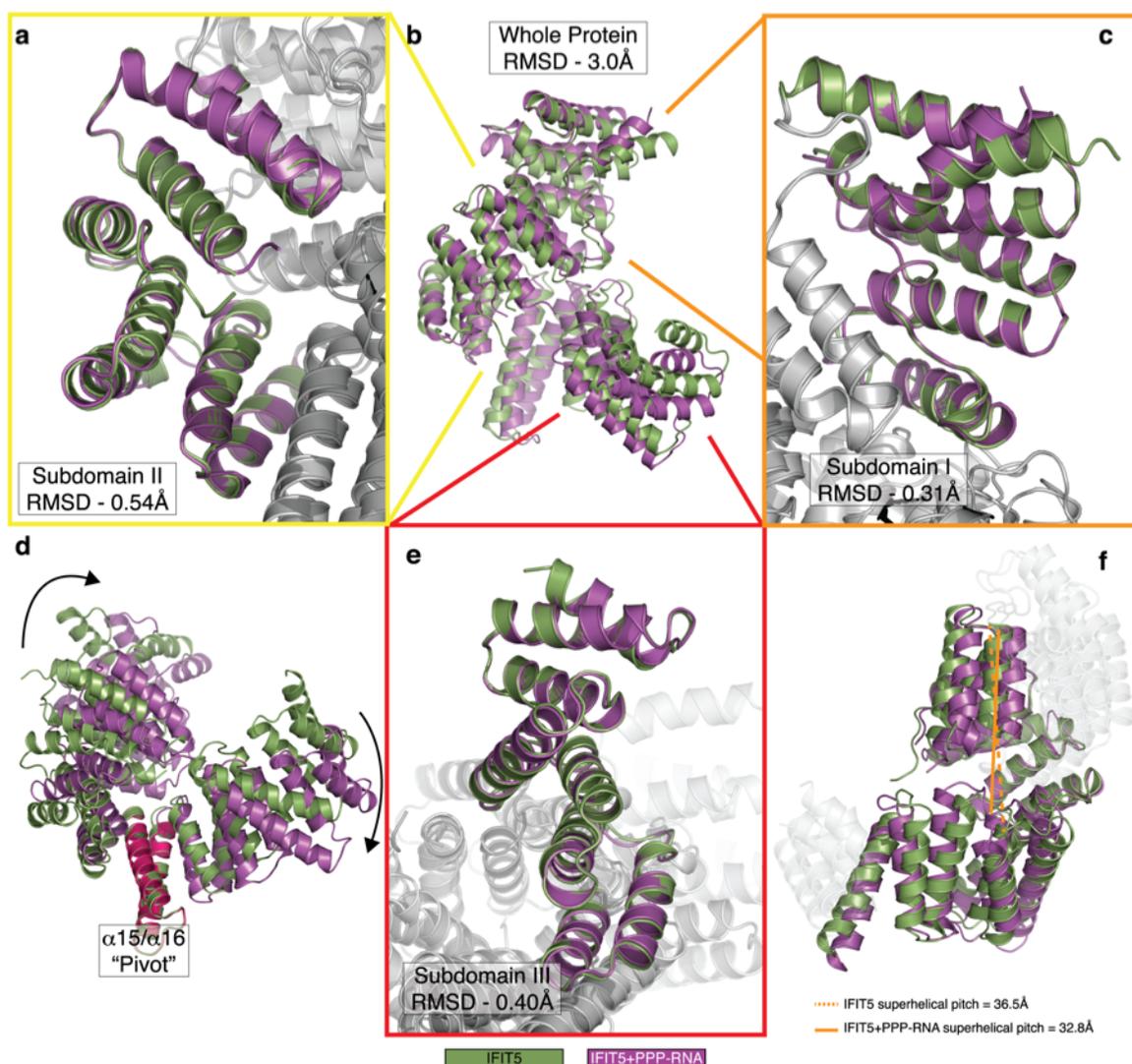
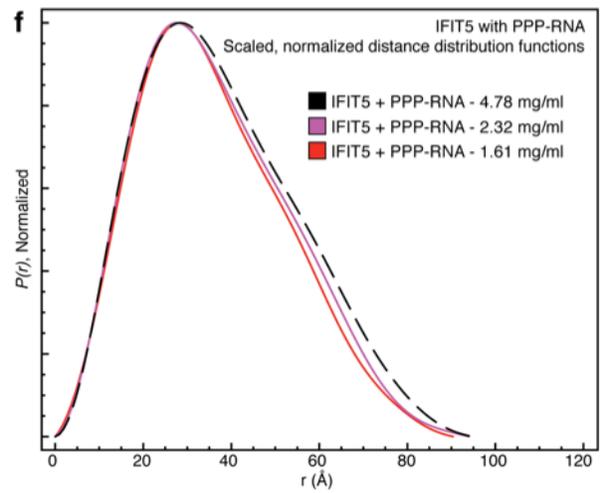
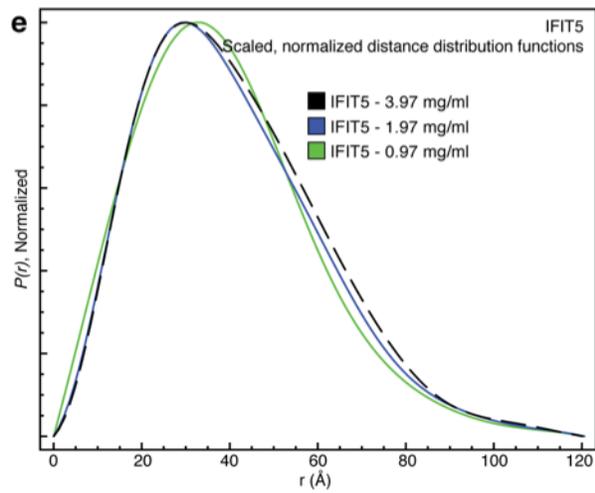
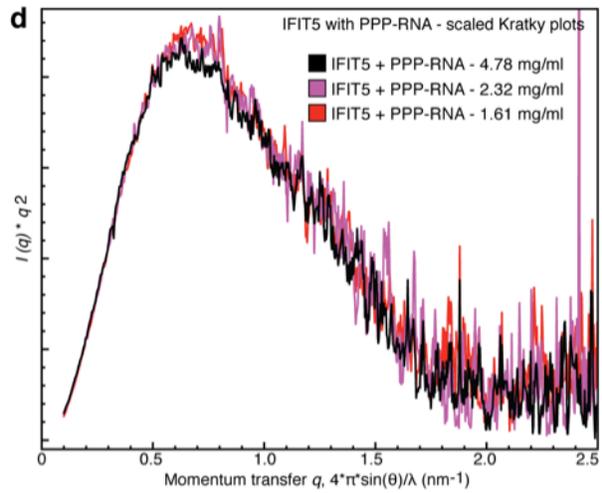
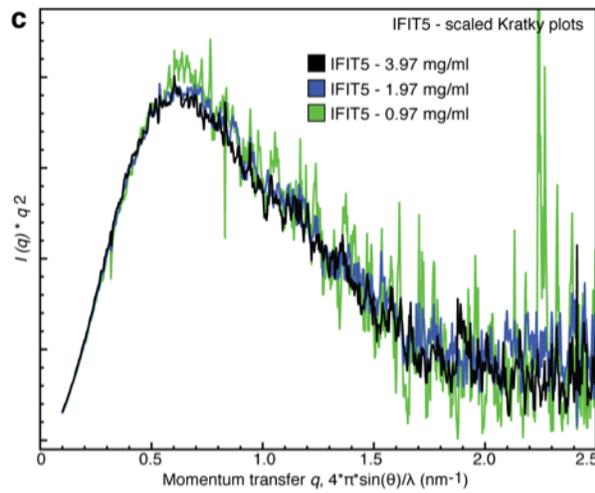
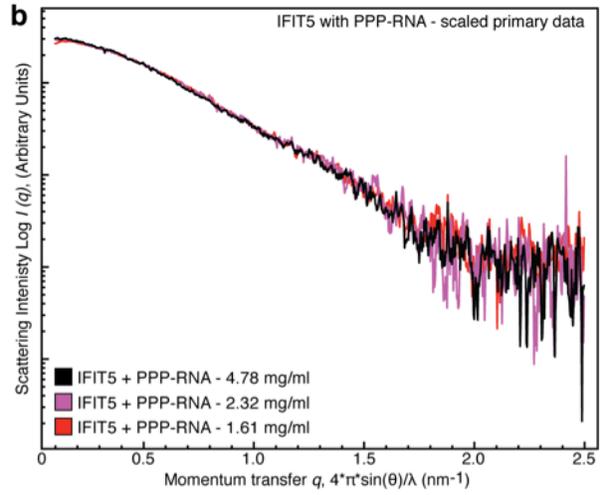
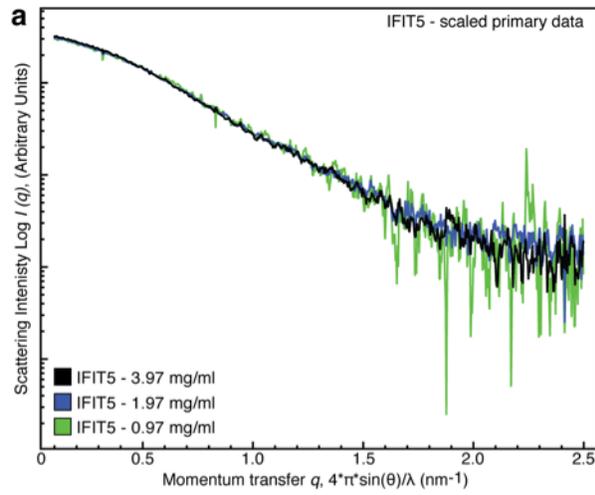


Figure 2.14 Comparison between RNA-free IFIT5 and RNA-bound IFIT5

(A-E) Superposition of the two forms using the different Subdomains as reference regions for the alignment. (F) The superhelical pitch (orange lines) decreases from 36.5 Å to 32.8 Å upon binding PPP-RNA



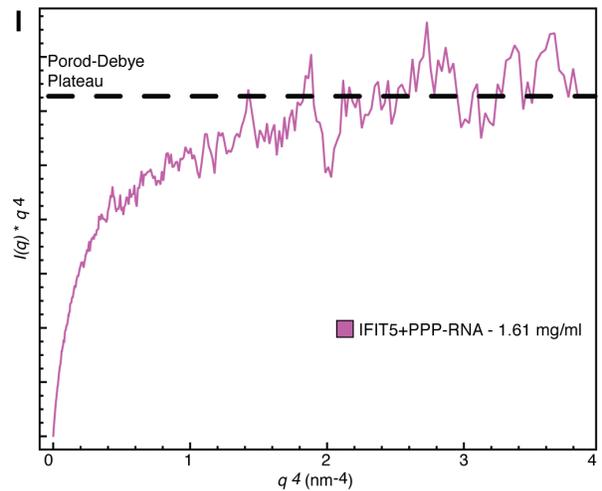
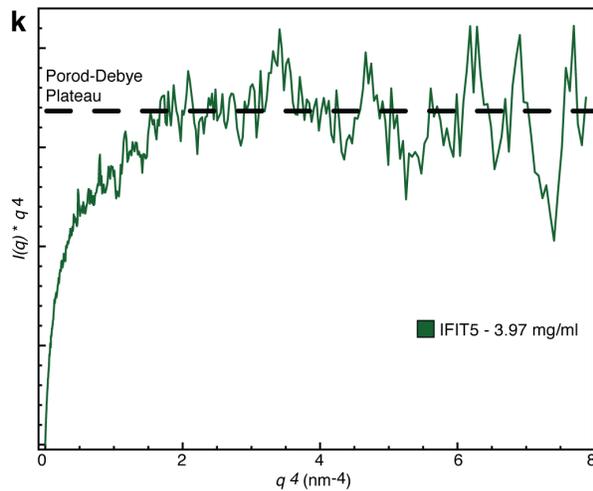
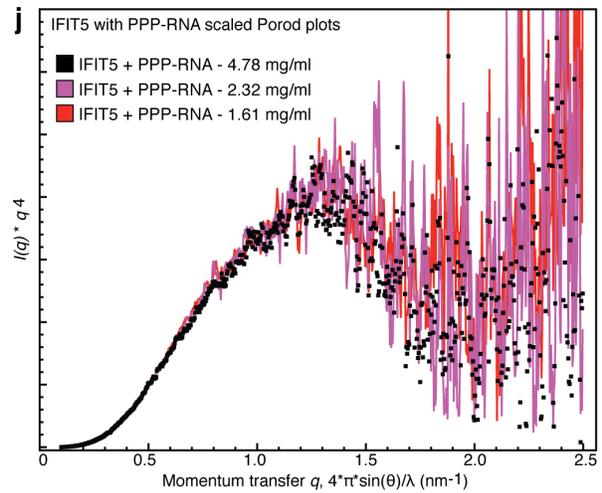
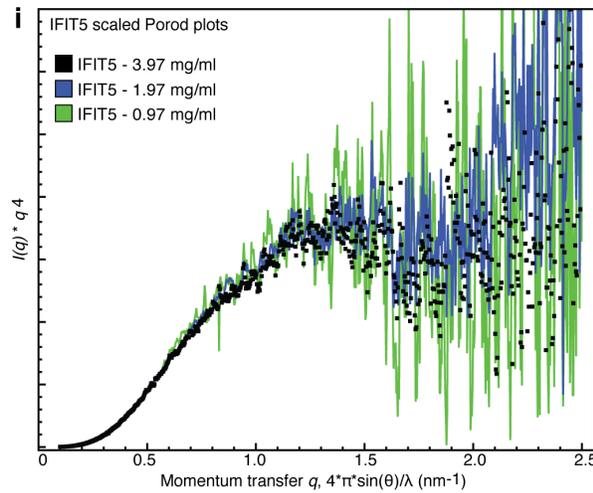
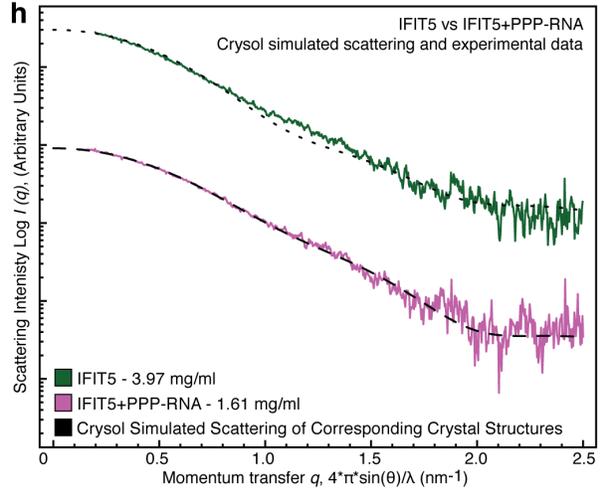
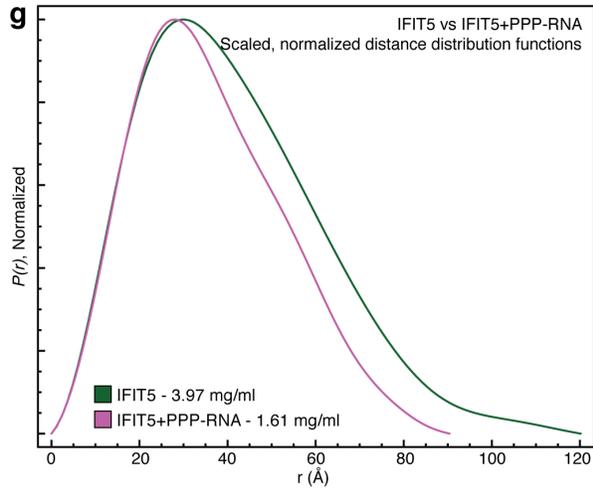


Figure 2.15 Small-angle X-ray scattering data NRA-free IFIT5 and RNA-bound IFIT5

1-D solution scattering profiles of **(A)** RNA-free IFIT5 and **(B)** IFIT5 with PPP-RNA. **(C and D)** Kratky transformation ($I(q) * q^2$ vs q) of the data, where the presence of a single peak that tends towards zero is indicative of a folded domain in both RNA-free IFIT5 and IFIT5 with PPP-RNA (Rambo and Tainer, 2011). **(E and F)** Distance distribution functions ($P(r)$) of each data set, determined using GNOM. The point at which each curve meets the x-axis is the determined D_{\max} (maximum particle dimension) that is reported in table 2.3 and Fig. 2.13C. **(G)** Comparison of the distance distribution functions of RNA-free IFIT5 and IFIT5 with PPP-RNA (curves normalized against the peak maximum). IFIT5 with PPP-RNA has a smaller D_{\max} than RNA-free IFIT5. **(H)** Comparison of the simulated scattering calculated for each crystal structure (using CRY SOL), against the solution scattering of the corresponding form. There is a lack of agreement between solution IFIT5 (green) and the crysol calculated scattering of RNA-Free IFIT5 (dotted curve), particularly around $q = 1.0 - 1.5 \text{ nm}^{-1}$. **(I and J)** Porod transformation of the data ($I(q) * q^4$ vs q) and **(K and L)** Porod-Debye transformation ($I(q) * q^4$ vs q^4). The Porod-Debye plateau was determined as in reference (Rambo and Tainer, 2011). Data in **(A-G and I-J)** were scaled by multiplying the $I(q)$ of each data set by a scale factor (= I_0 of the highest concentration measurement divided by the I_0 of the corresponding measurement).

2.3.5 IFIT5 and IFIT1 bind only PPP-ssRNAs

The internal diameter of the RNA-binding pocket in IFIT5 is roughly 15 Å, leaving no room to accommodate dsRNA, which would require a diameter of greater than 21 Å. Moreover, at least three bases are necessary to span the length of the pocket, suggesting that IFIT5 is potentially a sensor for PPP-ssRNA, or base-paired PPP-RNA with a minimum three-nucleotide overhang. By contrast, foreign PPP-RNA species in the cytosol that optimally activate RIG-I seem to require blunt-ended RNAs, which are thought to be the most potent immune-stimulant of RIG-I (Schlee et al., 2009).

To assess the recognition of distinct PPP-RNA species by IFITs, we used gel-shift assays. A 44-nucleotide ssRNA with no predicted secondary structure within the 5' 22 nucleotides was *in vitro* transcribed, to which complementary RNA strands of 15–20 nucleotides were annealed to generate dsRNA with blunt ends and various 5'-overhangs. Consistent with the crystal structure, we found that IFIT5 could shift both PPP-ssRNA and PPP-dsRNA with at least three-nucleotide overhangs, but could not efficiently shift blunt-ended PPP-RNA or PPP-dsRNA with 1–2-nucleotide overhangs (**Fig. 2.16 A and Fig. 2.17 A**). Similarly, IFIT1 could only shift PPP-ssRNA or PPP-dsRNA with at least five-nucleotide overhangs (**Fig. 2.17 B**). As a negative control we used IFIT3, which could not shift any species of RNA (**Fig. 2.16 A and Fig. 2.17 D**). Thus, owing to the limitations imposed by their RNA-binding pockets, IFIT5 and IFIT1 can engage only PPP-RNAs that have single-stranded 5'-ends.

2.3.6 Functional validation of PPP-RNA binding to IFITs

To examine the functional relevance of residues involved in binding PPP-RNA, we used PPP-RNA-coated beads to pull down c-Myc-tagged wild-type and mutant IFIT5 and IFIT1 expressed in HEK293 cells. We began by first corroborating that IFIT5, like IFIT1, could be pulled down by RNA only when it is triphosphorylated at the 5'-end (**Fig. 2.18 A**), and that replacing the triphosphate with 5'-cap, 5'-monophosphate or 5'-hydroxyl diminishes the binding (**Fig. 2.16 B**). The affinity of PPP-RNA for IFIT5 is between 250–500 nM (**Fig. 2.16 B**), similar to that found previously for IFIT1 (Pichlmair et al., 2011).

Next, we mutated key RNA contacts within the pocket and found that in most cases, a single-residue substitution was sufficient to abolish RNA binding *in vitro* (**Fig. 2.16 C**). In IFIT5, all residues recognizing the PPP-RNA were critical for binding, except for Glu 33, Tyr 156 and

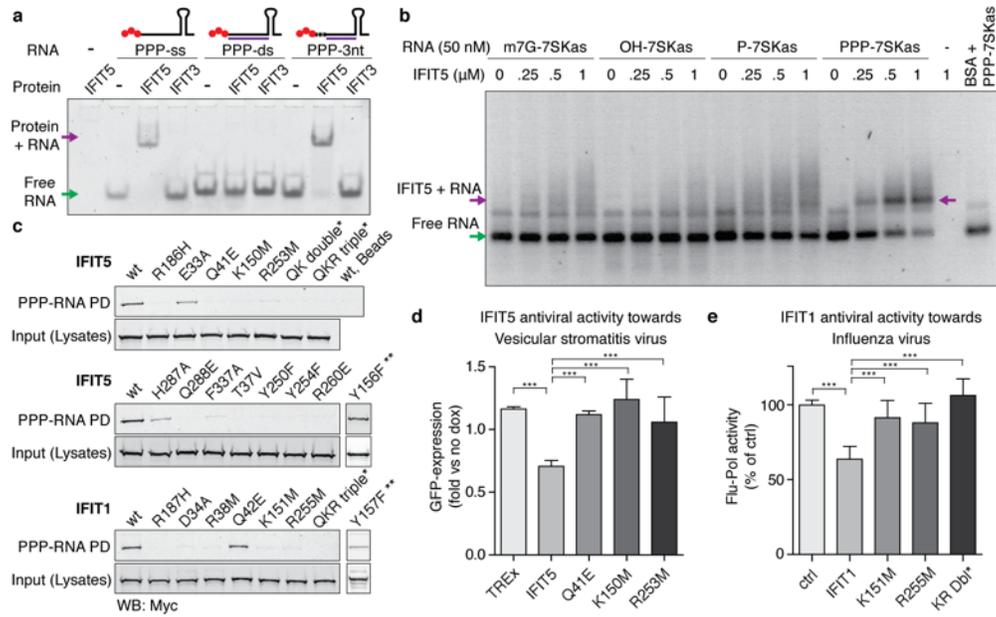


Figure 2.16 Functional analyses of IFIT binding to PPP-RNA

(A) Mobility shift assay between IFIT5/IFIT3 and ssRNA, dsRNA with blunt ends, or dsRNA with a 3-nucleotide overhang as indicated by the schematics above each set of lanes (PPP, red spheres; *in vitro* transcribed top strand, black line; synthetic complementary RNA, purple). (B) Agarose gel shift assay between IFIT5 and various RNAs indicated. (C) Biotinylated RNA pull-downs (PD) of wild-type (wt) and mutant IFIT1 and IFIT5 from HEK293 cell lysates. * QK Double is Q41E/K150M and QKR Triple is Q21E/K150M/R253M. ** Y156F and Y157F were carried out separately, and the appropriate positive and negative controls are in Fig. 2.18. (D and E) PPP-RNA binding is required for antiviral activities of IFIT5 and IFIT1. (D) Replication of Vesicular stomatitis virus expressing GFP in doxycycline (dox) inducible HEK-Flp-In cells expressing IFIT5 (and mutants). Average fold change (+/- SD) in dox-treated versus untreated cells of ten measurements. (E) Influenza virus in 293T cells transfected with IFIT1 (and mutants). Average percentage (+/- SD) of influenza polymerase activity as compared to control (ctrl) of four independent experiments done in duplicate measurements. *** = $p < 0.001$ (1 way ANOVA, Tukey's Multiple Comparison Test).

His 287. Homologous mutations in IFIT1 also lead to abrogation of RNA binding. Interestingly, IFIT1 Q42E, which would disrupt interactions with the β - and γ -phosphates (based on IFIT5), did not lead to loss of binding in IFIT1. This may be related to the natural substitution of nearby Thr37 in IFIT5 for arginine (R38) in IFIT1, which possibly replaces the Gln 41 interaction with the 5'-triphosphate group. Regardless, the PPP-RNA-binding pocket identified here is probably involved in a similar mode of recognition in other IFIT family members.

Finally, to investigate whether the RNA-interacting residues are important for the antiviral activity of IFIT5 against virus infection, we used HEK293 Flp-In TREx cells that inducibly express IFIT5, and IFIT5 mutants that have lost their ability to bind PPP-RNA. Consistent with the

mutational analysis, IFIT5 mutants were impaired in their ability to restrict growth of vesicular stomatitis virus compared to wild-type IFIT5 (**Fig. 2.16 D**). Similarly, IFIT1 lacking the ability to bind PPP-RNA was not able to inhibit the activity of an influenza virus polymerase (**Fig. 2.16 E**), consistent with the notion that binding to PPP-RNA is critical for the antiviral activity of IFIT1.

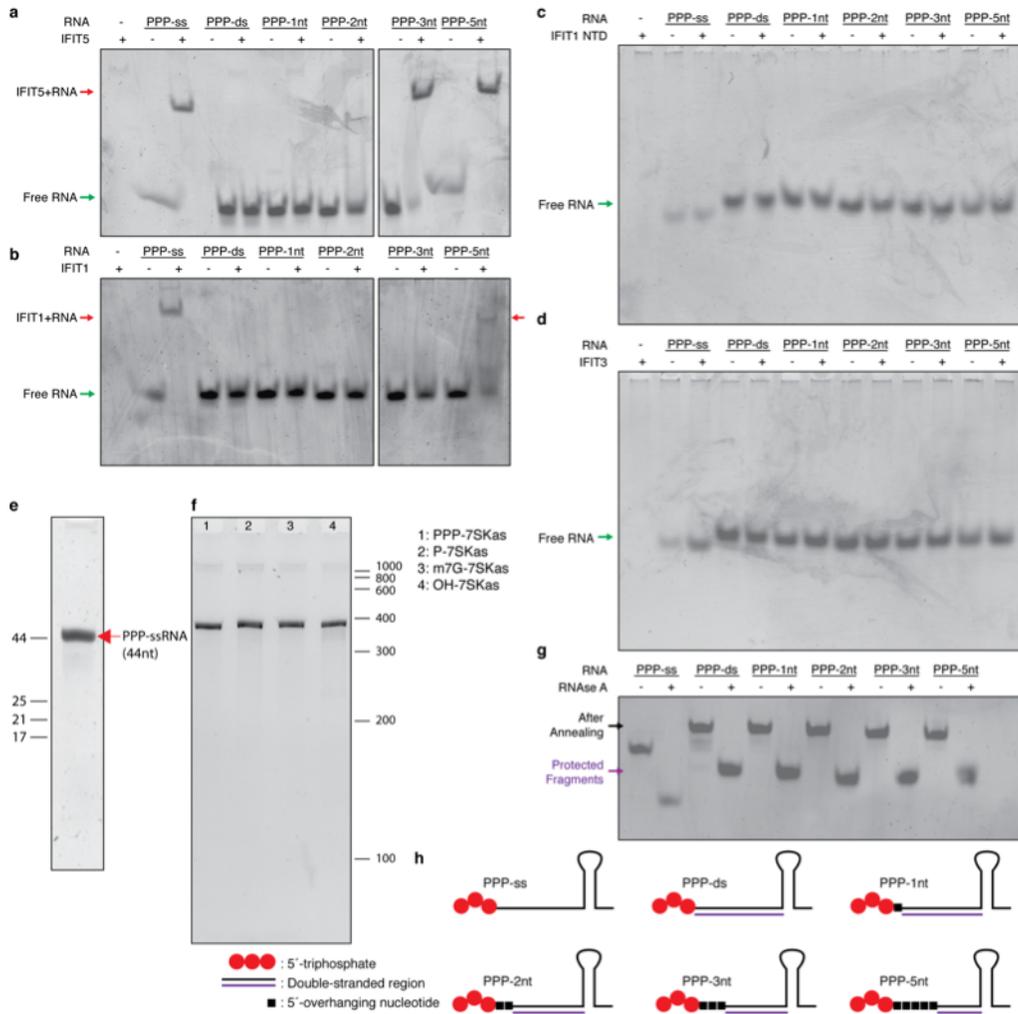


Figure 2.17 IFIT1 and IFIT5 preferentially bind ssRNA or base-paired RNA with 5' overhangs

(A and B) IFIT1 and IFIT5 preferentially bind ssRNA (C) The N-terminal domain of IFIT1 used for crystallization has little or no affinity towards any PPP-RNA. (D) IFIT3 is used as a negative control and cannot shift any of the PPP-RNAs tested. (E) 15% denaturing PAGE in 1X TBE (19:1 acrylamide:bisacrylamide) stained with SyBr gold. The first two lanes are marker lanes with OH-RNA. The third lane is the *in vitro* transcribed 44mer used for gel shift analysis. (F) 5% denaturing gel analysis of 7SK-as RNAs used in Fig. 2.16. (G) Gel shift to validate proper annealing of the bottom strands to generate blunt-ended dsRNA and dsRNA with various overhangs. RNase A degradation was carried out by mixing 1 pmol of PPP-RNA with 500 ng of RNase A for 30 min at 4 °C and run on 12% native PAGE in 1X TAE. (H) Predicted structure of the PPP-RNAs used in this experiment.

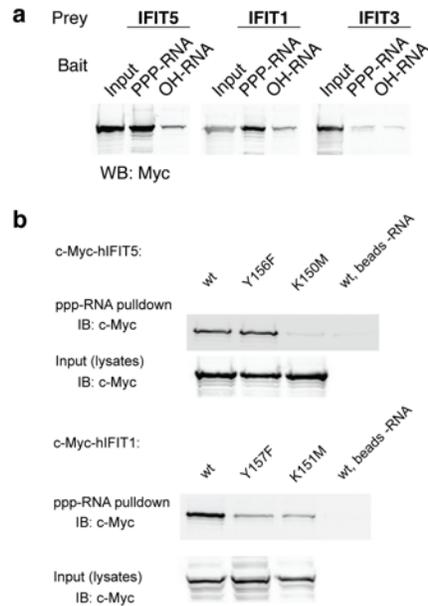


Figure 2.18 Pull-downs between myc-tagged IFITs and RNA-coated beads

(A) Pull-downs of IFIT1, IFIT3 and IFIT5 with PPP-RNA and OH-RNA from HEK293 cell lysates. (B) Additional tyrosine mutation (IFIT5 Y156F and IFIT1 Y157F). The pull-down alongside appropriate positive and negative controls is shown here. The second lane of each blot was spliced out and displayed adjacent to the gel in main text Fig. 2.16 C

2.4 Discussion

The structural basis for IFIT recognition of foreign RNA described here validates the new paradigms put forth for how this family of interferon-stimulated genes carry out their effector functions, and brings to the forefront the versatility of the TPR motif in recognizing diverse ligands, paralleling established receptors of the innate immune system such as those containing leucine-rich repeats (Barbalat et al., 2011).

In addition to using protein–protein interactions to confer downstream antiviral activity (Fensterl and Sen, 2011; Pichlmair et al., 2011), the principal molecular role of IFITs seems to be initiated by direct recognition of foreign PPP-ssRNAs. PPP-RNAs are found within the genome of negative-sense ssRNA viruses such as influenza and vesicular stomatitis virus. Other RNA viruses, such as positive-sense viruses which have 5'-capped genomes, can also generate cytosolic PPP-RNAs as replicative intermediates during their life cycle. Hence, the evolution of a binding site to specifically recognize PPP-RNA allows IFITs to distinguish self from non-self RNAs, as cytosolic host ssRNAs bear a 5'-monophosphate (on ribosomal RNA and transfer RNA) (Hornung

et al., 2006) or are 5'-capped (in the case of mRNA). In doing so, one possible mechanism for IFIT function may be to latch onto the ends of viral RNA, preventing it from being properly replicated or packaged into progeny virions.

Recent studies have suggested a role for IFIT proteins in sensing the 5'-cap methylation status of some viral RNA (for example, West Nile virus, poxvirus and coronavirus) (Daffis et al., 2010; Szretter et al., 2012; Zust et al., 2011). A 5'-cap is present on positive-sense virus genomes, but most viruses also have the ability to either hijack a cap from host mRNA or encode machinery to add a 5'-cap structure to their mRNA, thereby potentially circumventing IFIT recognition (Decroly et al., 2011b). Although a 5'-cap cannot be accommodated within the RNA-binding pocket of IFIT5 identified here, we do not preclude the possibility that other IFITs may be able to recognize capped viral RNAs. Taken together, it is clear that unravelling the structural details that underlie IFIT biology will improve our understanding of the complex interplay between pathogens and host innate immunity, and hopefully pave the way for the development of new immunotherapeutics.

2.5 Materials and Methods

2.5.1 Protein expression and purification

For crystallization and gel shift assays, full-length IFITs (1, 3 and 5), and nIFIT1 (residues 7-279) were cloned into a pSMT3 vector (Mossessova and Lima, 2000) between BamHI and NotI sites. The fusion proteins contained an N-terminal, Ulp1-cleavable 6xHis-Sumo tag. All proteins were expressed in BL21 (DE3) cells using standard protocols, and purified with a two-step Ni-affinity chromatography followed by cleavage of the tag. RNA contaminated samples of IFIT1 and IFIT5 were passed over a Mono Q 4.6/100 PE (GE Healthcare) or HiTrap Q HP 5ml (GE Healthcare) in 25 mM Tris-HCl pH 8.0 and eluted over a shallow salt gradient. To recover IFIT5 that co-purified with RNA from *E. coli*, the contaminated fractions were incubated in 50 mM Tris-HCl pH 8.0 with 5 M NaCl, and buffer exchanged in an Amicon Ultracel (30kDa cut-off) concentrator several times until the bound RNA flowed through (purity was determined by $A_{260}:A_{280}$ ratios). A final gel filtration step using Superdex75 or Superdex200 (GE Healthcare) columns was carried out in 25 mM Tris-HCl pH 8.0, 150 mM NaCl, 5% glycerol and 3 mM DTT (IFITs 1, 3 and 5) or 20 mM NaH_2PO_4 , 150 mM NaCl and 3 mM DTT (nIFIT1). Selenomethionyl (SeMet) derivative proteins were expressed by inhibition of methionine biogenesis pathways, and purified as the native. For

the gel shift with 7SK-as, IFIT5 was expressed from a pETG10A-hIFIT5 plasmid and purified on a HisTrap column as previously described (Pichlmair et al., 2011), followed by gel filtration on a Superdex 200 (GE Healthcare) in 2x PBS and 0.5mM TCEP. For pulldowns, c-myc-hIFIT1 or c-myc-hIFIT5 constructs (wild-type or mutant) were expressed by transfection of pCS2-6myc-based vectors into HEK293 cells. Cells were cultured in DMEM (PAA) supplemented with 10% (vol/vol) FCS (Invitrogen) and antibiotics (100 U/ml penicillin and 100 µg/ml streptomycin) and lysed by incubation in TAP buffer (50 mM Tris, pH 7.5, 100 mM NaCl, 5% (vol/vol) glycerol, 0.2% (vol/vol) Nonidet-P40, 1.5 mM MgCl₂ and protease inhibitor ‘cocktail’ (Complete; Roche)).

2.5.2 RNA and *in vitro* transcription

In vitro transcription protocols were adapted from (McKenna et al., 2007). Briefly, T7 polymerase was made recombinantly or purchased from NEB and used to transcribe the templates. To generate the various PPP-RNAs for crystallization, we used several dsOligo templates (oligoCT₇, oligoUT₇, oligoAT₇, BioCorp DNA, Table 2.4), which encoded 3 templated positions of C, U or A. By ensuring that each reaction contained only CTP, UTP or ATP, the final product was guaranteed to have sequence homogeneity. The typical non-templated $n + 1$, $n + 2$ products of T₇ transcription were observed (see crystallization section). The reactions were cleaned up by phenol/chloroform extraction, and precipitated with ethanol. For gel-shift assays, a 44nt PPP-ssRNA (Table 2.5) was cloned into pGEX-6P-1 between BamHI and EcoRI. The template was linearized with EcoRI (NEB) prior to run-off transcription, and the transcript purified on a Superdex75 column. Synthetic RNAs (Table 2.5) were ordered from IDT Technologies to generate the double stranded RNA. Biotinylated PPP-RNA was produced *in vitro* using SP6 MegaScript kit from Ambion, with addition of Biotin-16-UTP (Biozym/Epicentre BU6105H) and using a plasmid encoding antisense 7SK RNA (7SK-as, Table 2.5) as template. Biotinylated RNA was purified from the *in vitro* transcription reaction using RNeasy kit (Qiagen). 7SK-as RNA for gel shifts was prepared using the SP6 MEGAscript kit (Invitrogen). 5'-monophosphorylated RNA was obtained by adding guanosine 5'-monophosphate (Sigma-Aldrich) at a 5:1 ratio to GTP in the *in vitro* transcription reaction, and capped RNA by adding cap analog (m7G(5')ppp(5')G, Epicentre) at a 4:1 ratio. 5'-OH RNA was prepared by CIP (New England Biolabs) treatment of ppp-RNA. RNA was purified using the RNeasy kit (Qiagen).

2.5.3 Crystallization and structure determination

Prior to crystallization, proteins were buffer exchanged into their respective gel filtration buffers supplemented with 1-10 mM TCEP. Crystals of apo IFIT5 were obtained in 0.1 M HEPES pH 7.5, 5-10% PEG 3350 and 0-7.5% glycerol, and used as seeds for SeMet IFIT5 crystallization. Single crystals of SeMet IFIT5 (~0.7 mm x 0.2 mm x 0.2 mm) grew at 4 °C at 4 mg/ml in 10% PEG 3350, 0.1 M HEPES pH7.5, 2% ethylene glycol. The structure of apo IFIT5 was solved by single-wavelength anomalous diffraction (SAD) using ShelX to determine the heavy atom substructure (Sheldrick, 2007), and refined with Arp/Warp (Langer et al., 2008), Coot (Emsley et al., 2010), and Phenix (Adams et al., 2010). The final model contained residues 2-189 and 193-481 from the full-length construct.

To crystallize a 1:1 complex of IFIT5:PPP-RNA, pellets of the RNA oligos (see RNA and *in vitro* transcription) were resuspended in a solution of IFIT5, incubated for at least 1 hour, and purified on Mono Q and Superdex200 columns as described above. Crystals of the complex were obtained at 22 °C between 5 and 20 mg/ml in 5-10% ethanol and 0.1 M Tris pH 7 - 8. The structures were solved using the apo-IFIT5 structure broken up into two search models (residues 1-282 and residues 283-481) for molecular replacement. For the oligo-C structure, four nucleotides were modelled in the electron density and the 5'-phosphate of a fifth nucleotide was also modelled with no electron density visible for the fifth sugar and base. The second base was found in both *syn*- and *anti*- conformations and both were modelled with occupancy 0.5 for each (Fig. 2.7 A). For the oligo-U structure, RNA could be modelled up to the 5'-phosphate of the fourth nucleotide (Fig. 2.7 B). For the oligo-A structure, 4 nucleotides could be modelled within the electron density and the second base was found in the *anti*- conformation (Fig. 2.7 C). For the metal ion, both sodium and magnesium could be refined with acceptable temperature factors, but the distances after refinement more closely matched the ligation geometry of magnesium, in the oligo-C and oligo-A complexes. The distances of the metal within the oligo-U complex were more consistent with sodium. The final models contained residues 1-482 (in the oligo-C and oligo-U complex) and residues 6-481 (in the oligo-A structure).

Crystals of SeMet nIFIT1 were obtained at 4 °C in 17-20% PEG 3350, 0.25-0.3 M KSCN and 6% glycerol at 5-15 mg/ml. The structure was solved by the SAD method using SOLVE (Terwilliger, 2003) and refined as above. The final model contained two molecules per asymmetric

unit (IFIT1 residues 10-84, 91-195, and 198-278 in chain A; residues 9-27, 46-83, 91-193 and 197-278 in chain B). Chain A was used for structural analyses in the main text.

SAD data were collected at the CLS 08ID-1 beamline using 0.979 Å synchrotron radiation under a nitrogen cryostream. IFIT5:PPP-RNA complex crystals were collected using a Rigaku MicroMax-007 HF (rotating copper anode) and 1.54 Å radiation under a nitrogen cryostream.

APBS was used to calculate surface electrostatic potential (Baker et al., 2001), and PyMol to generate all molecular figures (<http://www.pymol.org>).

2.5.4 Small Angle X-ray Scattering

IFIT5 and IFIT5:PPP-RNA were purified as above and dialysed into SAXS buffer (25 mM Tris-HCl pH 8.0, 150 mM NaCl, and 1 mM TCEP). Measurements were made on an Anton Paar SAXSess mc2 equipped with a PANalytical PW3830 X-ray generator and a Roper/Princeton CCD Detector. The beam length was set to 16 mm and the beam profile recorded using an image plate for subsequent desmearing. 1-D data was collected along 10 mm of the CCD, with 10 sec exposure time per frame. For both proteins, data was collected at 4 °C with 3 different concentrations (to evaluate concentration dependent effects), and for a maximum amount of time before radiation damage was detectable. SAXSquant 3.5 (Anton Paar) was used for background correction, scaling, buffer subtraction and desmearing. The R_g and $I(0)$ were estimated from Guinier plots using PRIMUS (Konarev et al., 2003) in the regions between q min (Table 2.3) up to $qR_g < 1.3$. Simulated scattering curves of the crystal structures were computed with CRY SOL (Svergun et al., 1995), and distance distribution functions ($P(r)$) determined with GNOM (SvergunIUCr, 1992). Porod and Porod-Debye analyses were carried out as described (Rambo and Tainer, 2011).

2.5.5 Gel shift assays

Electrophoretic mobility shift assays were adapted from (Hellman and Fried, 2007). Protein and RNA were incubated in binding buffer (10 mM Tris pH 7.9, 100 mM NaCl, 1 mM TCEP, 5% v/v glycerol) at 4 °C at a ratio of 0.5 μM : 0.1 μM (protein:RNA) for 2-4 hours, and run on 7% PAGE, 1X TAE supplemented with 100 mM NaCl, in 1X TAE running buffer. The temperature during the run was maintained at < 10 °C. The RNA was visualized with SyBr Gold (Invitrogen) staining and scanned using a Typhoon variable mode imager or UV transillumination. To generate blunt-ended dsRNA and dsRNA with various overhangs, the 44-mer ssRNA was mixed with the complementary bottom strand (Fig. 2.17) at final concentrations of 1 μM and 1.1 μM, respectively

in annealing buffer (10 mM NaH₂PO₄/Na₂HPO₄ pH 7.5, 50 mM NaCl, 1 mM EDTA). Annealing was done by heating to 95 °C followed by slow cooling to room temperature. Proper annealing was verified on a 15% Native PAGE with an RNase A protection assay (Fig. 2.17). For the agarose gel shift with 7SK-as RNA, IFIT5 was diluted in PBS, and mixed at the indicated concentrations with 50 nM 7SK-as RNA. The reaction was supplemented with 5x loading buffer (250 mM DTT, 50% glycerol, 0.05% Bromophenol Blue, 2 x Tris-glycine) and incubated for 15 min at RT. The reactions were analyzed on a 0.8% agarose gel in 1x Tris-glycine running buffer and RNA was stained with SyBr Gold (Invitrogen).

2.5.6 Limited Proteolysis

In 20 µl reactions, 20 µg of IFIT5 or IFIT5:PPP-RNA (purified as above) were incubated with Elastase, Trypsin or Chymotrypsin at protease:protein ratios of 1:10, 1:10 and 1:100, respectively. At time points 0 (before addition of protease), 5 min, 15 min, and 30 min, 5 µl was removed, mixed with 1x SDS sample buffer, boiled at 95 °C and frozen at -20 °C until gel analysis.

2.5.7 Mutational analysis and pulldowns

Point mutations were introduced into pCS2-6myc-hIFIT1 or pCS2-6myc-IFIT5 using the Quick change II site-directed mutagenesis kit (Stratagene). For pull-downs on PPP-RNA, 1 µg 7SK-as RNA was added to streptavidin resin (Ultralink Immobilized Streptavidin Plus Gel, Pierce 53117), followed by incubation for 60 min with 3 mg HEK293 cell lysates. Beads were washed three times in TAP buffer (Pichlmair et al., 2011); for precipitation of IFIT5, the NaCl concentration in TAP buffer was raised to 250 mM. Proteins were eluted by boiling in SDS sample buffer and analyzed by SDS-PAGE. Staining for c-myc was done using IRDye-conjugated anti-c-Myc (600-432-381) antibody from Rockland.

2.5.8 Flu polymerase activity (IFIT1 antiviral activity assay)

To test the influence of IFIT1 PPP-RNA binding mutants on virus replication we used an influenza replication assay (Dittmann et al., 2008; Pichlmair et al., 2011). 293T cells were co-transfected with 125 ng of pHH21-Seg.4-FFLuc (a kind gift of Georg Kochs), coding for an influenza polymerase template expressing firefly luciferase, 25 ng of renilla luciferase expression control plasmid (pRL-RK, Promega), 250 ng of plasmids coding for the indicated siRNA-resistant Myc-tagged IFIT1 versions or the control plasmid Myc-IFIT3, and IFIT1 siRNA (final concentration

20 nM). 24h later cells were infected with influenza virus (strain A/PR/8/34) (multiplicity of infection: 10) and expression of firefly and renilla luciferase was analysed after over-night culture, and measured relative to control.

2.5.9 HEK-Flip-In (IFIT5 antiviral activity assay)

Isogenic HEK293 Flp-In TREx cells that inducibly express the indicated IFIT5 mutants were generated as before (Pichlmair et al., 2011). 1×10^5 cells/24-well-cavity were seeded, left untreated or were treated with $1 \mu\text{g/ml}$ doxycycline for 8h and infected with VSV-GFP (multiplicity of infection: 0.01) and GFP expression was tested in a spectrofluorimeter after 24h.

2.6 Tables

Table 2.1 Data collection and refinement statistics IFIT5 and nIFIT1

	IFIT5 (SeMet) (PDB 4HOQ)	nIFIT1 (SeMet) (PDB 4HOU)
Data collection		
Space group	P2 ₁ 2 ₁ 2 ₁	C2
Cell dimensions		
<i>a, b, c</i> (Å)	64.2, 71.4, 117.6	84.5, 177.0, 55.1
α, β, γ (°)	90, 90, 90	90, 130, 90
Resolution (Å)	50-2.07 (2.14-2.07)	50-1.95 (2.02-1.95)
R_{sym}	14.4 (40.6)	7.9 (52.8)
<i>I</i> / <i>I</i>	19.5 (3.2)	23.4 (2.6)
Completeness (%)	96.1 (78.7)	98.0 (95.4)
Redundancy	8.6 (5.0)	7.4 (5.0)
Refinement		
Resolution (Å)	37.06-2.07 (2.12-2.07)	30.52-1.93 (1.96-1.93)
No. reflections	32241	45388
$R_{\text{work}}/R_{\text{free}}$	16.5/21.3 (21.6/29.4)	19.8/23.0 (23.6/27.2)
No. atoms		
Protein	4154	4285
Ligand/ion	0	0
Water	258	238
B-factors		
Protein	35.6	46.2
Ligand/ion	-	-
Water	40.8	46.1
R.m.s deviations		
Bond lengths (Å)	0.006	0.007
Bond angles (°)	0.94	1.12

*Highest resolution shell is shown in parenthesis.

Table 2.2 Data collection and refinement statistics IFIT5 with PPP-RNAs

	IFIT5-oligo-C (PDB ID 4HOR)	IFIT5-oligo-U (PDB ID 4HOS)	IFIT5-oligo-A (PDB ID 4HOT)
Data collection			
Space group	P2 ₁	P2 ₁	P2 ₁
Cell dimensions			
<i>a, b, c</i> (Å)	54.5, 84.9, 61.5	54.5, 84.9, 60.9	54.6, 85.2, 60.7
α, β, γ (°)	90, 106.9, 90	90, 106.8, 90	90, 106, 90
Resolution (Å)	50-1.86 (1.93-1.86)	50-2.00 (2.07-2.00)	50-2.50 (2.59-2.50)
<i>R</i> _{sym}	8.6 (59.0)	11.9 (67.0)	12.6 (59.7)
<i>I</i> / <i>I</i>	25.1 (1.68)	17.4 (2.4)	15.3 (1.7)
Completeness (%)	95.5 (86.2)	99.9 (99.4)	97.9 (82.2)
Redundancy	10.3 (5.1)	7.3 (6.2)	6.9 (3.9)
Refinement			
Resolution (Å)	29.57-1.86 (1.89-1.86)	31.8-1.99 (2.02-1.99)	34.43-2.50 (2.60-2.50)
No. reflections	42808	36075	16270
<i>R</i> _{work} / <i>R</i> _{free}	16.4/20.1 (23.8/29.3)	17.5/20.7 (22.5/30.1)	18.0/23.0 (25.0/32.5)
No. atoms			
Protein	4059	3989	3953
Ligand/ion	114	74	98
Water	425	363	43
B-factors			
Protein	33.3	33.0	57.0
Ligand/ion	32.9	29.9	47.1
Water	38.3	36.7	42.6
R.m.s deviations			
Bond lengths (Å)	0.007	0.006	0.006
Bond angles (°)	1.03	0.96	1.02

*Highest resolution shell is shown in parenthesis

Table 2.3 Sequences of DNA templates used to generate the PPP-RNA for crystallization

Sequence Name	Sequence 5' -> 3'
oligoAT ₇ Fw	taatac gactcact ataaaa
oligoAT ₇ RV	TTT tatag tgagtc tatta
oligoUT ₇ Fw	taatac gactcact atattt
oligoUT ₇ RV	AAA tatag tgagtc gtatta
oligoCT ₇ Fw	taatac gactcact ataccc
oligoCT ₇ RV	GGG tatag tgagtc gtatta

Bold, capitalized positions indicate the template region to be transcribed

CONNECTING TEXT

In addition to recognizing viral PPP-RNAs, IFIT1 has been shown to bind viral mRNA cap structures lacking ribose 2'-O methylation on the first and second transcribed nucleotides. This property allows IFIT1 to distinguish between virus and host mRNAs, the latter of which are normally methylated at these two positions. The structure of IFIT5 with PPP-RNA revealed the structural basis for IFIT RNA interactions, but could not explain how IFIT1 accommodates the mRNA cap. I therefore turned my attention to structurally characterizing the interaction between IFIT1 and capped RNA.

CHAPTER 3: STRUCTURE OF HUMAN IFIT1 WITH CAPPED RNA REVEALS ADAPTABLE MRNA BINDING AND MECHANISMS FOR SENSING N1 AND N2 RIBOSE 2'-O METHYLATIONS

Abbas, Y.M., Laudenbach, B.T., Martínez-Montero, S., Cencic, R., Habjan, M., Pichlmair, A., Damha, M.J., Pelletier, J., and Nagar, B. (2017). Structure of human IFIT1 with capped RNA reveals adaptable mRNA binding and mechanisms for sensing N1 and N2 ribose 2'-O methylations. *Proc Natl Acad Sci USA* **114**, E2106-E2115.

3.1 Abstract

IFIT1 is an effector of the host innate immune antiviral response that prevents propagation of virus infection by selectively inhibiting translation of viral mRNA. It relies on its ability to compete with the translation initiation factor eIF4F to specifically recognize foreign capped mRNAs, while remaining inactive against host mRNAs marked by ribose 2'-O methylation at the first cap-proximal nucleotide (N1). We report here several crystal structures of RNA-bound human IFIT1, including a 1.6 Å complex with capped RNA. IFIT1 forms a water-filled, positively charged RNA-binding tunnel with a separate hydrophobic extension that unexpectedly engages the cap in multiple conformations (*syn*- and *anti*-) giving rise to a relatively plastic and non-specific mode of binding, in stark contrast to eIF4E. Cap-proximal nucleotides encircled by the tunnel provide affinity to compete with eIF4F while allowing IFIT1 to select against N1 methylated mRNA. Gel-shift binding assays confirm that N1 methylation interferes with IFIT1 binding, but in an RNA-dependent manner, while translation assays reveal that N1 methylation alone is not sufficient to prevent mRNA recognition at high IFIT1 concentrations. Structural and functional analysis show that 2'-O methylation at N2, another abundant mRNA modification, is also detrimental for RNA binding, thus revealing a novel and potentially synergistic role for it in self versus non-self-mRNA discernment. Finally, structure-guided mutational analysis confirms the importance of RNA binding for IFIT1 restriction of a human coronavirus mutant lacking viral N1 methylation. Our structural and biochemical analysis sheds new light on the molecular basis for IFIT1 translational inhibition of capped viral RNA.

3.2 Introduction

Infection by a virus relies on its ability to exploit the host's translational machinery to convert its genome into protein products that can ultimately be used to assemble new viral particles. In eukaryotes, endogenous mRNA is protected by a highly conserved 5' cap structure consisting of an *N*7-methylguanosine triphosphate (m7Gppp/Cap0) moiety. This is recognized by the eukaryotic translation initiation factor 4E to promote cap-dependent translation (eIF4E together with eIF4G and eIF4A comprise eIF4F) (Pelletier et al., 2015). In higher eukaryotes, the mRNA cap is further modified by ribose 2'-O methylation on the first and sometimes second cap-proximal nucleotides (N1 and N2, where N is any nucleotide, **Fig. 3.2 A**), resulting in Cap1- (m7GpppNmN) or Cap2- (m7GpppNmNm) mRNA (Bélanger et al., 2010; Werner et al., 2011). N1 methylation was recently shown to serve as a molecular signature of 'self' which can subvert mammalian antiviral responses (Daffis et al., 2010; Zust et al., 2011). As such, many viruses also produce Cap1-mRNA, either through the action of host- or virally-encoded 2'-O methyltransferases (MTases) or through viral 'cap-snatching' enzymes (Decroly et al., 2011b; Hyde and Diamond, 2015). Hence, Cap0-mRNAs (along with other virus-derived RNAs) are marked as 'non-self' and can trigger responses such as the Type I Interferon (IFN) antiviral program (Devarkar et al., 2016; Schuberth-Wagner et al., 2015; Zust et al., 2011) which culminates in the induction of hundreds of IFN-Stimulated Genes (ISGs) (Fensterl et al., 2015).

Among the most potently induced of the ISGs are the IFITs (IFN induced proteins with Tetratricopeptide Repeats), a family of antiviral effectors whose expression can also be triggered downstream of IFN-independent signaling (Fensterl and Sen, 2015). They are conserved throughout vertebrate evolution with humans and most mammals encoding five paralogues: IFIT1, IFIT1B, IFIT2, IFIT3 and IFIT5, although many also possess species-specific duplications and deletions. For instance, mice lack IFIT5 and were only recently discovered to have also lost IFIT1 (Daugherty et al., 2016). Therefore, what is currently known as mouse *Ifit1* (54 % sequence identity with human IFIT1) is actually an ortholog of human IFIT1B. In humans, IFIT1B (67 % sequence identity with human IFIT1) is not known to be IFN-inducible (Fensterl and Sen, 2015), and recent data suggests that it may be non-functional (Daugherty et al., 2016). IFITs are structurally related and are composed of tandem copies of the tetratricopeptide repeat (TPR), a helix-turn-helix motif. Structures of several IFITs have shown that their TPRs coalesce into

distinct super-helical subdomains that form clamp-shaped structures (Abbas et al., 2013; Feng et al., 2013; Katibah et al., 2013; Yang et al., 2012).

Recently, it was discovered that IFITs play a prominent role in impeding viral replication by directly binding the 5' end of viral RNA (Pichlmair et al., 2011). Thus, IFIT1 and IFIT1B can compete with eIF4F to selectively bind and sequester viral Cap0-mRNA resulting in its translational inhibition (Habjan et al., 2013; Kimura et al., 2013; Kumar et al., 2014). In this manner, mouse *Ifit1* has been shown to restrict a broad spectrum of wild-type and mutant viruses lacking 2'-O MTase activity, including alphaviruses, coronaviruses, flaviviruses, and vaccinia virus (Daffis et al., 2010; Habjan et al., 2013; Hyde et al., 2014; Menachery et al., 2014; Reynaud et al., 2015; Szretter et al., 2012; Zust et al., 2013), while mutating viral N1 methylation enhanced coronavirus and flavivirus sensitivity to human IFIT1 (Habjan et al., 2013; Menachery et al., 2014; Pinto et al., 2015; Zust et al., 2013). In contrast, host cellular mRNA is not targeted as it bears N1 2'-O methylation, which interferes with IFIT1 and IFIT1B binding (Habjan et al., 2013; Kimura et al., 2013; Kumar et al., 2014). That many cytoplasmic virus families have adapted by acquiring 2'-O MTases to generate their own Cap1-mRNA thereby potentially escaping IFIT1/IFIT1B restriction, underscores the importance of these proteins in this process (Hyde and Diamond, 2015). Furthermore, alphaviruses, which display only Cap0-mRNA, can still subvert mouse *Ifit1* activity by encoding cap-proximal structural elements (Hyde et al., 2014; Reynaud et al., 2015), which has also been shown to interfere with RNA binding and enhance pathogenicity.

IFIT1, along with IFIT5, can also recognize uncapped viral PPP-RNA (another 'non-self' marker of infection) to potentially inhibit the replication of some negative-sense single-stranded (ss) RNA viruses (Abbas et al., 2013; Pichlmair et al., 2011). The crystal structure of human IFIT5 bound to uncapped PPP-RNA revealed that the RNA sits in a narrow, positively-charged tunnel at the core of the protein, with a network of electrostatic interactions specifically recognizing the PPP moiety (Abbas et al., 2013). Up to four nucleotides are also stably bound within the tunnel in a sequence non-specific manner. The sequence identity between IFIT5 and IFIT1 (55 %) and a structure of the N-terminal region of IFIT1 suggested that IFIT1 accommodates capped RNA in a similar fashion. However, IFIT5 cannot bind capped mRNA (Abbas et al., 2013; Habjan et al., 2013; Kumar et al., 2014), and indeed, protein residues at the base of the tunnel would block any further progression beyond the PPP moiety. Thus, how IFIT1-like proteins can accommodate Cap0-mRNA remains unclear.

We report here several crystal structures of RNA-bound human IFIT1. The structures reveal that the positively-charged RNA-binding tunnel of IFIT1 is distinct from that found in IFIT5 and further extended to allow binding of both capped and uncapped RNAs. Strikingly, mRNA binding and cap recognition by IFIT1 appears to be adaptable and its mechanism is evolutionarily divergent from eIF4E and other cap binding proteins. The shape of the tunnel in the vicinity of the 2'-hydroxyls of N1 and N2 sterically occludes RNA methylated at these positions. A comprehensive analysis of the interaction between human IFIT1 and differentially methylated capped RNAs corroborates the structural findings, revealing that either N1 or N2 methylation alone interferes with IFIT1 binding, but in an RNA-dependent manner. Combining N1 and N2 methylation resulted in an additive and potentially synergistic effect in inhibiting IFIT1 activity, particularly towards susceptible RNA sequences and at high IFIT1 concentrations. Our structural and biochemical analysis therefore sheds light on IFIT1 antiviral activity and reveals a previously uncharacterized role for N2 ribose methylation and Cap2 structures as signatures of 'self' mRNA.

3.3 Results

3.3.1 RNA binding and inhibition of *in vitro* translation by human IFIT1

The interaction between IFIT1 and capped-RNA is well established, but the precise structural determinants of the viral RNA important for binding are as yet unclear. Thus, we began by carrying out electrophoretic mobility shift assays (EMSAs) between human IFIT1 and two 5' capped sequences derived from genomes of coronaviruses known to be restricted by human IFIT1 or mouse Ifit1: human coronavirus strain 229E (HCoV) and murine hepatitis virus strain A59 (MHV) (Habjan et al., 2013). Human IFIT1 bound the capped-RNAs with apparent affinities of ~ 250 nM and < 100 nM, respectively (**Fig. 3.1 A**). Binding strength decreased as the stability and proximity of RNA secondary structure to the 5'-end increased (**Fig. 3.1 A and Fig. 3.2 B**), confirming the preference for ssRNA as previously demonstrated for human IFIT1 and mouse Ifit1 (Abbas et al., 2013; Hyde et al., 2014). As with IFIT5, RNA binding is generally sequence non-specific and replacing the first 3 nts of HCoV (sequence ACU) with GGG resulted in only a modest enhancement of binding (**Fig. 3.1 A, right**). IFIT1 also binds uncapped PPP-RNA, but this is inherently weaker and more sensitive to the presence of predicted secondary structure at the 5'-end (**Fig. 3.2 C**). This is in contrast to IFIT5, which binds PPP-RNA but cannot accommodate capped

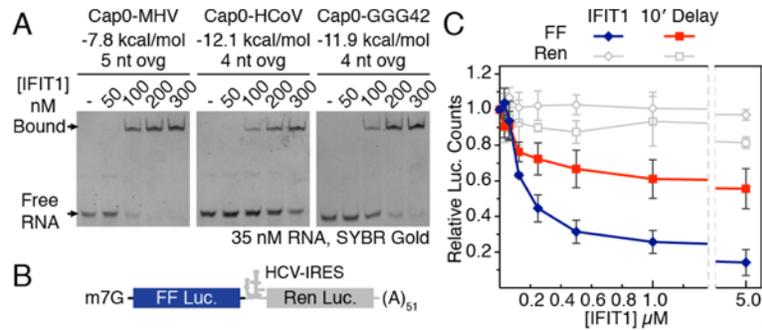


Figure 3.1 RNA binding and inhibition of *in vitro* translation by human IFIT1

(A) EMSAs between human IFIT1 and capped-RNA visualized by SYBR Gold staining. Cap0-MHV, first 41 nts of MHV strain A59; Cap0-HCoV, first 42 nts of HCoV strain 229E; Cap0-GGG42, ACU to GGG modification of HCoV. The RNA secondary structure minimum free energy (kcal/mol) and 5'-overhang length (ovg) are described (See also Fig. 3.2 B). (B) Schematic of bicistronic mRNA reporter. (C) Translation assay with IFIT1 titrated into Krebs extracts programmed with Cap0/m7Gppp reporter, and titration following a 10-minute pre-incubation of the reporter with extracts. FF and Ren luciferase (luc.) activities at each concentration were normalized against buffer control, which was set to 1. Data represent the mean of 2 independent measurements performed in duplicate \pm standard deviation.

RNA, as shown by its crystal structure and a variety of biochemical assays from several groups (Abbas et al., 2013; Habjan et al., 2013; Kumar et al., 2014).

To understand the contribution of IFIT1 binding to capped RNA in a more physiological context, we employed an *in vitro* translation system to assess the effect on translation initiation. The system consists of Krebs extracts programmed with a bicistronic Cap0-mRNA reporter (Novac et al., 2004). The 5'-cistron expresses a Firefly luciferase (FF) reporter that is translated in a cap-dependent manner, while the 3'-cistron expresses a Renilla luciferase (Ren) reporter under the control of an internal ribosome entry site (IRES) from hepatitis C virus (HCV, **Fig. 3.1 B**). Ren expression serves as an internal control for non-specific translational inhibition by IFIT1. Titrating human IFIT1 into these extracts at concentrations ranging from \sim 30 nM to 5 μ M showed that IFIT1 could inhibit Cap0-dependent translation with IC_{50} values of \sim 50-200 nM (**Fig. 3.1 C**, and other figures herein). Interestingly, addition of IFIT1 after the reporter was pre-incubated with translation extracts for 10 minutes resulted in its inhibitory activity being reduced by more than an order of magnitude ($IC_{50} > 5 \mu$ M), presumably due to the formation of a closed-loop mRNP that facilitates ribosome re-initiation (Amrani et al., 2008). This suggests that optimal IFIT1 activity *in vivo* is probably only realized in cells that are already expressing the protein before infection, as might be the case for cells activated by IFN signaling in a paracrine manner.

3.3.2 Overall structure of full-length human IFIT1.

To gain insight into the mechanisms of viral RNA binding by IFIT1, we initially crystallized RNA-bound, full-length, wild-type human IFIT1 (residues 1-478) in complex with short PPP- and m7Gppp-containing oligoadenosines. The IFIT1-RNA complex purified and crystallized as a dimer with two molecules in the asymmetric unit, but diffracted X-rays to only ~ 2.7 Å. To improve the resolution, we mutated the dimerization interface at the C-terminal end of the protein to produce a monomeric version that crystallized in a different space group and diffracted X-rays to 1.58 Å resolution (Table S1). The overall folds of the wild-type and monomeric mutant are essentially the same (r.m.s.d. 0.35 Å). Henceforth, we describe only the high-resolution structures with respect to RNA binding, while all functional assays were performed with wild-type protein. The structural and functional analysis of IFIT1 dimerization will be described elsewhere.

Human IFIT1 is made up of 23 α -helices, 18 of which form 9 TPR motifs that together form three distinct subdomains interrupted by non-TPR structural elements (**Fig. 3.3 A and B and Fig. 3.4 A-C**). The overall structure is similar to the previously determined RNA bound structure of human IFIT5 (r.m.s.d. 1.9 Å, **Fig. 3.4D**) and the N-terminal region of human IFIT1 (r.m.s.d. 0.8 Å, **Fig. 3.4 E**) (Abbas et al., 2013). The subdomains are arranged to form a clamp-shaped structure with a central RNA-binding tunnel that is approximately 30-40 Å in length and 12-19 Å in width,

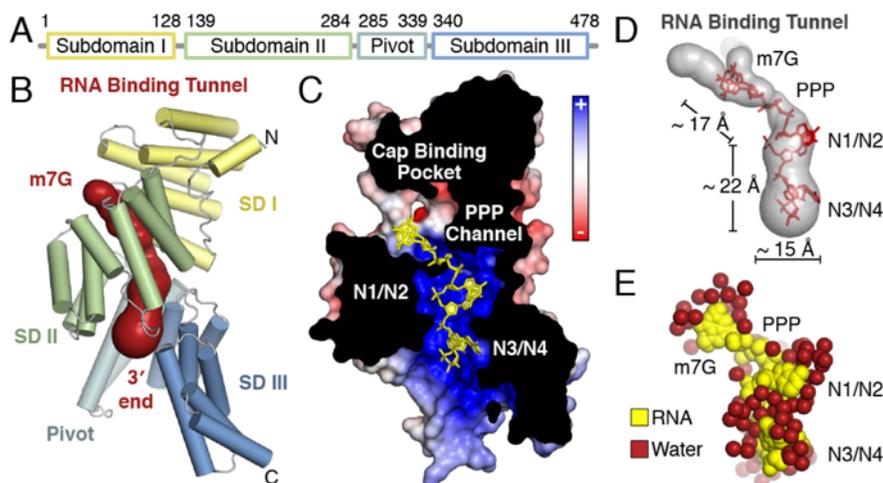


Figure 3.3 Overall structure of monomeric, RNA-bound human IFIT1

(A) Schematic of IFIT1 subdomains. (B) Cartoon representation of human IFIT1 colored by subdomain (SD) and surface representation of the tunnel (dark red) determined by CAVER. (C) Cross-section of IFIT1 colored by surface electrostatic potential from negative (-10 kTe $^{-1}$; red) to positive ($+10$ kTe $^{-1}$; blue) with capped-RNA (yellow sticks). (D) Dimensions of the IFIT1 tunnel (gray surface) and capped RNA (red sticks). (E) Waters surrounding the RNA inside the tunnel.

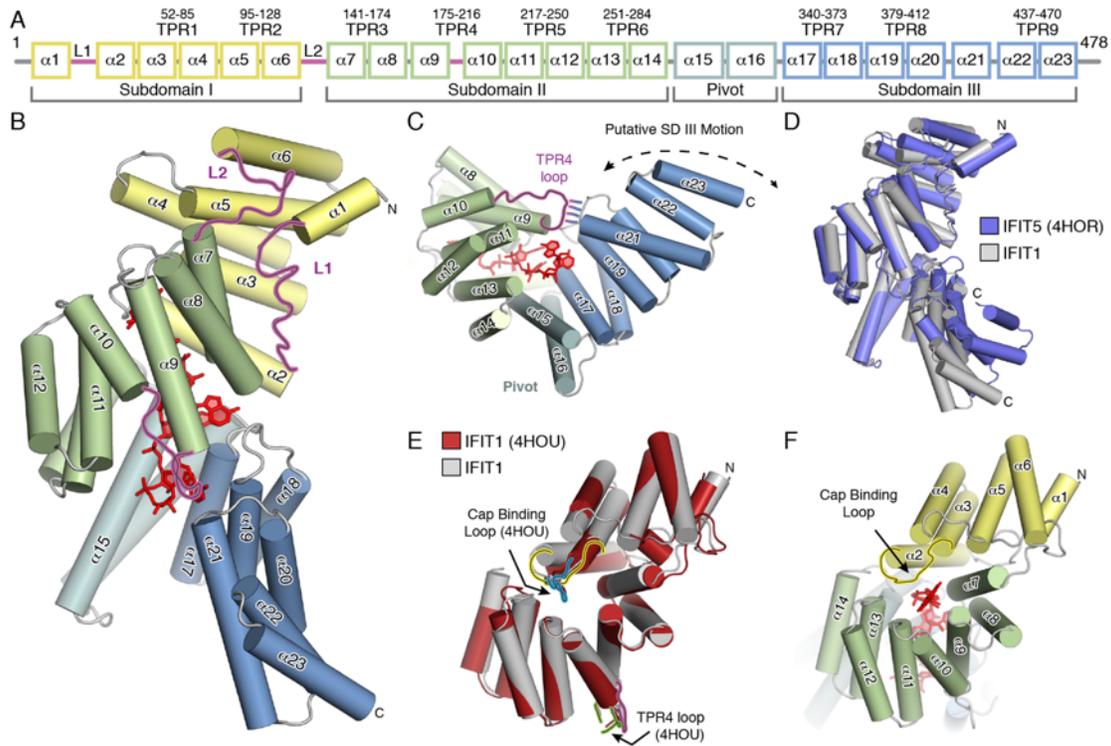


Figure 3.4 Overall structure of monomeric, RNA-bound human IFIT1 cont'd

(A) Secondary structure, TPR motif, and subdomain organization of human IFIT1, with TPR motif sequence numbers above. Note that due to an insertion between $\alpha 9$ and $\alpha 10$, TPR4 was annotated based on sequence and structure rather than sequence alone. All other TPR motifs were predicted based on sequence using TPRpred (<https://toolkit.tuebingen.mpg.de/tprpred>). (B) Overall structure of monomeric human IFIT1 with subdomains colored according to (A). The non-TPR helices $\alpha 1$ and $\alpha 2$ are bridged by Loop L1, which houses a highly conserved CHFTW motif (residues 19-23, see Fig. 3.10) that mediates contacts with $\alpha 7$. Subdomains I and II are linked by loop L2. The two non-TPR pivot helices ($\alpha 15/\alpha 16$) connect subdomains II and III. (C) TPR4 ($\alpha 9/\alpha 10$) has an unusually long intra-TPR loop (purple) which forms a lid over the 3' exit of the RNA binding tunnel. Additionally, it mediates contacts between subdomains II and III (indicated by purple/blue dashes). The sequence of this loop is not conserved, but the insertion is found in almost all IFIT1- and IFIT5-like genes. The double headed arrow indicates putative subdomain III motions between RNA-free and RNA-bound forms of IFIT1, based on the structural analysis of human IFIT5. (D) Superposition of RNA-bound IFIT1 over RNA-bound IFIT5 (PDB 4HOR). The two structures share a high degree of global similarity. The C-terminal regions appear to diverge, but that is due to inter-subdomain angle differences between IFIT1 and IFIT5. (E) Superposition of RNA bound IFIT1 over the structure of RNA free, N-terminal IFIT1 (PDB 4HOU). In the partial structure of RNA-free, N-terminal IFIT1, the cap binding loop (blue) adopts a different conformation than in full-length (yellow, see also (F)). However, in one of the two molecules of the asymmetric unit (of 4HOU), this loop has high B-factors (loop average 61 \AA^2 , protein average 46 \AA^2), while in the other it is disordered, indicating that this loop is mobile in the absence of RNA. The TPR4 loop (green) is also disordered in the structure of N-terminal IFIT1 (4HOU), likely because of the lack of RNA and missing subdomain contacts. (F) The cap binding pocket is formed at the interface of subdomains I and II, with one wall formed by helix $\alpha 2$ and the cap binding loop (connecting $\alpha 2$ and $\alpha 3$), and the other by subdomain II.

accommodating only ssRNA with a total of five nucleotides (cap + four RNA nucleotides, **Fig. 3.3 B-D**). As with IFIT5, a pair of long non-TPR pivot helices connect the second and third subdomains and likely function in an analogous fashion to regulate closure of the protein around the RNA (Abbas et al., 2013) (**Fig. 3.4 C**). About 30-40% of the tunnel volume is occupied by bound water molecules (**Fig. 3.3 E**), which appears to be an important facet for recognition of different RNA sequences and structures (discussed below). We demarcate four distinct regions of the tunnel according to their role in RNA binding: 1) the cap binding pocket, which houses the *N*⁷-methylguanosine moiety, 2) the triphosphate channel, which links the cap binding pocket to the 5'-end of the RNA, 3) the first dinucleotide (N1 and N2), where the presence of 2'-O methylation is sensed, and 4) the second dinucleotide (N3 and N4), where the requirement for single stranded 5'-ends is reinforced.

3.3.3 The IFIT1 RNA-binding tunnel houses a functionally distinct cap binding pocket.

IFIT1 and IFIT5 were previously characterized as PPP-RNA binding proteins (Pichlmair et al., 2011), although more recent evidence revealed that the primary role of IFIT1 is in binding capped-RNA. Conversely, the role of IFIT5 remains restricted to recognition of 5'-phosphorylated RNAs (Abbas et al., 2013; Feng et al., 2013; Katibah et al., 2013). The structure of IFIT1 bound to PPP-RNA revealed that, like IFIT5, the PPP moiety is ligated by numerous specific electrostatic interactions from protein side chains (**Fig. 3.6 A-C**). However, there are some key differences. IFIT5 recognition of PPP-RNA utilizes a positively charged metal ion bound between the α - and γ -phosphates, which stabilizes a bent conformation of the PPP facilitated by T37 at the base of the tunnel in IFIT5 (**Fig. 3.5 A**). The corresponding position in IFIT1 is occupied by an arginine (R38), and an ion is no longer part of its PPP binding. This results in a more extended conformation of the PPP that allows it to reach towards the entrance of a neighboring unoccupied pocket.

The crystal structure of IFIT1 bound to m7Gppp-RNA revealed that this adjacent pocket harbors the cap moiety. Whereas most of the RNA binding tunnel is positively charged, the cap binding pocket is generally more hydrophobic and interactions with the cap occur predominantly through non-specific van der Waals contacts (**Fig. 3.4 C and 3.5 B**). Surprisingly, we found that the m7G base adopts both *syn*- and *anti*- conformations with approximately equal occupancies (**Fig. 3.5 C and Fig. 3.6 D**, discussed in detail below). In either conformation, m7G sits atop a tryptophan residue (W147) making π - π stacking interactions, reminiscent of other cap binding

proteins such as eIF4E (Quiocho et al., 2000) (**Fig. 3.5 D and E**). Additionally, the base is abutted by I183 from the same side as W147, and on the other side by L46 and T48 emanating from a flexible loop that forms the outer wall of the pocket, which we term the ‘cap binding loop’ (**Fig. 3.5 D and E and Fig. 3.4 F**). The ribose of m7G is similar in the *syn*- and *anti*- modes of binding, adopting an S-type conformation that is stabilized by an intramolecular hydrogen bond between the ribose 3'-OH and the bridging β -phosphate (**Fig. 3.6 D**). It sits in a pocket formed by Q42, L46, R187, Y218, I183 and L150 (**Fig. 3.5 E**), and two ordered water molecules – the first coordinated by Q42, and the second bridging the ribose 3'-OH to the backbone carbonyl of W147 (**Fig. 3.6 F**).

As in PPP-RNA bound IFIT1, the bridging triphosphate in the cap bound structure is in an extended conformation stabilized by numerous electrostatic interactions, although pulled slightly towards the cap binding pocket (**Fig. 3.5 F and Fig. 3.6 E**). The γ -phosphate interacts with K151,

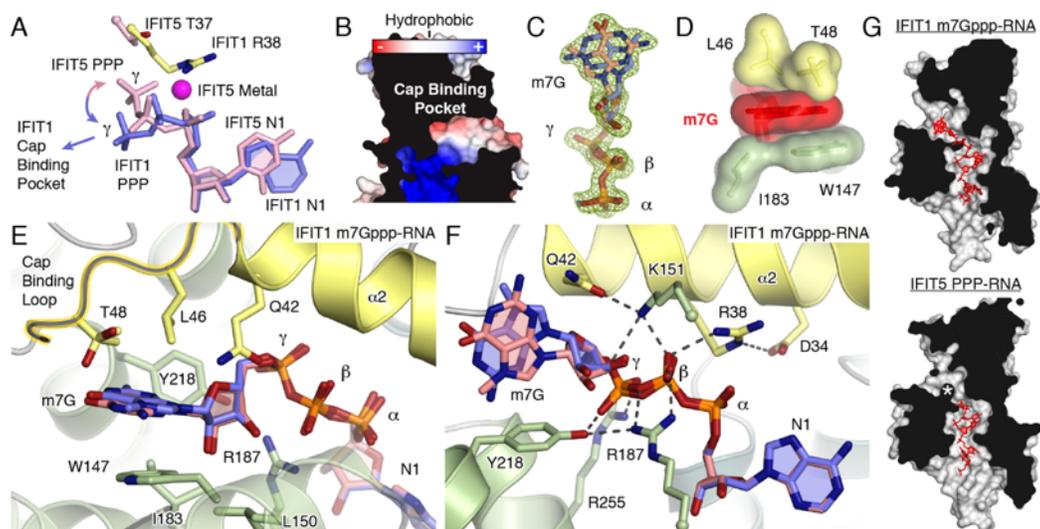


Figure 3.5 IFIT1 mRNA cap binding mechanism

(A) The IFIT1 PPP (blue) adopts an extended conformation compared to the ‘bent’ IFIT5 PPP (pink). The γ -phosphate from PPP-RNA-bound IFIT1 points towards the nearby unoccupied cap binding pocket. (B) Protein cross-section and close-up the cap binding pocket. This view is rotated by $\sim 180^\circ$ compared to Fig. 3.4 C. (C) Simulated annealing $2F_o-F_c$ omit map of the m7Gppp- moiety contoured at 1σ . *Syn*- and *anti*-configuration carbons are colored light blue and salmon, respectively. (D) Surface/stick representation of residues (colored by subdomain) abutting the m7G base moiety from above and below. The inter-planar distance between m7G and Trp 147 is 3.4-3.7 Å. (E and F) Cartoon/stick representation of residues interacting with the m7Gppp- moiety. Shown are both conformations of the m7GpppA dinucleotide, which was modeled as a single residue during model building and refinement. (G) Cross sections of the IFIT1 and IFIT5 RNA binding tunnels (grey/black) with m7Gppp- or PPP-RNA (red sticks). The * shows where K48 and Q41 block the IFIT5 putative cap binding pocket (see Fig. 3.6 H-K).

R255, Y218, and R187, and the β -phosphate is coordinated by K151, R187, and R38. Additionally, 6 highly ordered water molecules mediate hydrogen bonds with the α - and γ -phosphates (**Fig. 3.6 F**). Finally, m⁷Gppp binding is facilitated by a high degree of inter-residue coordination, for instance R38 is held in place by D34, while W147 is coordinated by E176 (**Fig. 3.6 G**).

In IFIT5, although most of these cap binding residues are conserved, substitutions at a few key positions render it unable to bind cap productively. As described above, replacement of R38 in IFIT1 with T37 in IFIT5 causes it to recognize a more compact conformation of the PPP in a metal dependent manner. This positions the γ -phosphate away from the putative cap binding pocket, which draws in several residues such as Q41 (from helix α 2) and K48 (from the putative cap binding loop) causing them to block access to its putative cap binding pocket (**Fig. 3.5 G and Fig. 3.6 H and I**). Therefore, the formation of a positively charged RNA binding tunnel with an accessible and spatially separated cap binding pocket in IFIT1 explains, at least in part, why it can bind capped-RNA whereas IFIT5 cannot. The preference for an arginine or threonine on helix α 2 is highly conserved amongst IFIT1/1B-like and IFIT5-like sequences, respectively (**Fig. 3.6 J**), and the identity of this PPP bridging residue (Arg or Thr) appears to play a major role in determining the 5' specificities of IFIT1/1B- or IFIT5-like genes. Interestingly, a small group of mammalian IFIT5-like genes retain an arginine at this position; these sequences probably have a hybrid IFIT1/IFIT5 character, and possibly resemble an ancestral IFIT1/IFIT5 precursor gene, since they all belong to non-placental mammals (e.g. opossum and platypus, **Fig. 3.6 K**).

Figure 3.6 IFIT1 PPP- and m7Gppp- binding mechanism

(A) Simulated annealing $2F_o - F_c$ omit map of the PPP moiety contoured at 1σ . (B) Cartoon/stick representation of residues making specific contacts with the triphosphate group from PPP-RNA bound IFIT1. (C) Superposition of PPP-RNA bound IFIT1 residues (blue carbons with orange phosphates) over PPP-RNA bound IFIT5 residues (pink carbons with yellow-orange phosphates). The metal from PPP-RNA bound IFIT5 is represented with a magenta sphere. The PPP from IFIT1 is in a different conformation, and makes additional H-bonds with Y256, Y218 and R187 not present in IFIT5. R38/D34 from IFIT1 are replaced with T37/E33 in IFIT5. Depending on the region, IFIT5 residue positions are offset by -1 or -2 compared to IFIT1 residue positions. (D) Simulated annealing $2F_o - F_c$ omit map of the m7Gppp moiety contoured at 1σ . *Syn*- and *anti*-configuration carbons are colored light blue and salmon, respectively. The hydrogen bond between the β -phosphate and the 3'-OH is indicated with the dashed line. Due to the constraints imposed on the base and bridging PPP binding, the ribose sugar pucker adapts by switching from a C2'-*endo* conformation in the *anti*-configuration, to C1'-*exo* in the *syn*-configuration. The final refined occupancies are ~ 0.5 for each conformation. (E) Superposition of PPP-RNA bound IFIT1 (grey) over m7Gppp-RNA bound IFIT1 (green side-chain carbons and blue RNA carbons). The extended PPP conformation is common to both PPP- and m7Gppp-RNA binding, except that the γ -phosphate of m7Gppp- is repositioned at the entrance of the cap binding pocket, away from Y256 and towards Y218, for more optimal m7G binding inside the pocket. (F) Water mediated hydrogen bonds between IFIT1 and the m7Gppp moiety. Starting with the water at Q42 (*) and going clockwise, the isotropic B-factors of these waters are 27.6, 18.2, 17.6, 18.0, 19.4, 24.1, 15.3, 17.8 \AA^2 ; most of which are lower than the crystal isotropic B factor average (30.65 \AA^2). (G) Coordination between residues involved in m7Gppp-RNA binding. R187 is coordinated by Y218 and Y157; R38 by D34; K151 by R38 and Q42; and W147 by E176. (H) Superposition of m7Gppp-RNA-bound IFIT1 over IFIT5 PPP-RNA. In IFIT5, T37 facilitates recognition of a bent/compact PPP conformation that is further stabilized by a metal bound to the α - and γ -phosphates and coordinated by E33. This draws in K150 (which is H-bonded to T37), Q41 (which interacts with the PPP directly), and K48 (which is coordinated by Q41). (I) In IFIT5, K48 sits atop the cap binding residue to block access to the cap binding pocket as indicated by the attempt to model m7Gppp-. In IFIT1, K49 from its cap binding loop is facing the solvent. Also, in IFIT1, Q42 is not involved in binding the PPP directly. (J) WebLogo (<http://weblogo.berkeley.edu/>) sequence consensus of helix $\alpha 2$ from 86 IFIT1- and IFIT1B-like genes, and 60 IFIT5-like genes. (K) The mammalian IFIT5 sequences were split into placentals and non-placentals (i.e. tasmanian devil, opossum, platypus). Only non-placental IFIT5-like sequences have an arginine at position 37 (human IFIT5 numbering).

3.3.4 IFIT1 can non-specifically accommodate multiple forms and conformations of the cap.

The high-resolution structure of the monomeric IFIT1 mutant (1.58 Å) allowed us to unambiguously build two conformations for the m7G base that are consistent with the electron density, which has a relatively symmetric shape due to an $\sim 180^\circ$ rotation about the *N*-glycosidic bond connecting the base to the ribose. Multiple base conformations were also evident in the lower resolution wild-type m7Gppp-RNA co-crystals. This results in the interactions at the base periphery being quite distinct in the two conformations (**Fig. 3.7 A**). Notably, there are no direct hydrogen bonds from the protein towards the base in either conformation, but there are a small number of water-mediated interactions. In the *anti*-orientation, *N3* of the base is weakly hydrogen bonded to a water molecule that is coordinated by Q42. The *N7*-methyl (*C7*) and *O6* groups make van der Waals contacts with N216, and the remainder of the base is partially oriented towards water molecules near the proximal opening of the tunnel leading to bulk solvent. In the *syn*-orientation, the *N7*-methyl and *O6* are instead pointing towards the bulk solvent while *N2* is nestled between Y218 and N216.

Because of the presence of the *N7*-methyl group on m7G, it acquires a delocalized positive charge on its imidazole ring that could in principle enhance the stacking with W147 through additional cation- π interactions (Dougherty, 2013; Quijcho et al., 2000). However, the geometry of cation- π stacking changes with the base orientation (**Fig. 3.7 A**). Whereas the *anti*-conformation places the positive charge at an angle away from W147, the *syn*-conformation places it directly over the indole ring of W147. To test whether *N7*-methylation and associated positive charge controls cap orientation, we determined the structure of Gppp-RNA (lacking the *N7*-methyl group) bound to IFIT1 (1.7 Å, **Fig. 3.7 B and Fig. 3.8 A**). Here, the base exists only in the *anti*-conformation, indicating that the ability of the base to adopt two conformations depends, at least in part, on *N7*-methylation. The presence of two base conformations is also determined by the chemical environment surrounding the base, since the structure of an N216 mutant of monomeric IFIT1 (N216A) bound to m7Gppp-RNA also resulted in the base adopting only the *anti*-conformation (**Fig. 3.8 B**).

Despite these observations, IFIT1 appears not to be selective for *N7*-methylation. In fact, gel shift assays suggest that Gppp-RNA binding is, in some cases, more efficient than m7Gppp-RNA binding (**Fig. 3.8 C** and (Habjan et al., 2013; Kumar et al., 2014)), in stark contrast to the

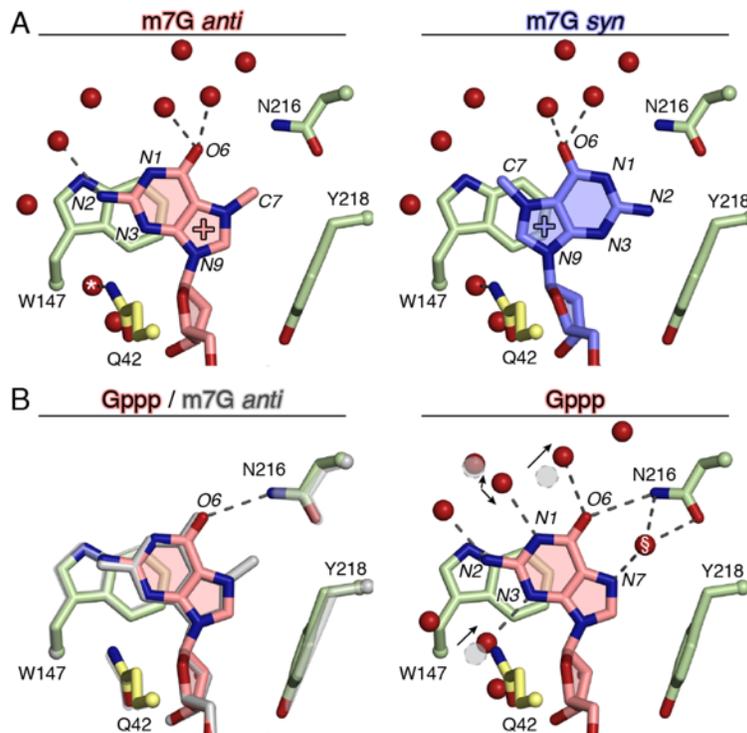


Figure 3.7 IFIT1 can accommodate multiple forms and conformations of the cap

(A) m7G base interactions at its periphery in the *anti*- or *syn*-modes of binding. The water H-bonded to Q42 (*) is 3.3 Å away from N3 in the *anti*-mode. (B) Left, the Gppp- moiety adopts only *anti*, and approaches N216 to form a weak H-bond through O6; m7Gppp-RNA bound IFIT1 is superposed in gray. Right, the water structure surrounding the G moiety changes compared to m7G, and satisfies almost all H-bond donor and acceptor groups. The arrows depict the movement of waters from the m7Gppp- bound form (gray circles) to the Gppp-bound form (red spheres). The water molecule H-bonded to Q42 and N3 becomes more ordered in the Gppp-RNA structure (same resolution and crystal form but B-factor decreases from 28 Å² to 15.3 Å²). The water at N7 (§) appears only in the Gppp-bound form.

case with eIF4E, whose cap binding is strongly dependent on proper methylation (Quiocho et al., 2000). In the absence of the methyl group, the guanine ring moves closer to N216 making a hydrogen bond with it through O6 (Fig. 3.7 B, left), although removing this hydrogen bond through mutation (N216A or N216D) does not weaken Gppp-RNA binding (Fig. 3.8 D). Interestingly, the water structure surrounding Gppp-RNA changes compared to m7Gppp-RNA, such that almost all hydrogen bond donor and acceptor sites on the unmethylated base are now satisfied (Fig. 3.7 B, right), and this may be a contributing factor for maintaining strong Gppp-RNA binding (relative to m7Gppp-RNA).

The physiological relevance of Gppp-RNA binding by IFIT1 is unclear, but one possibility is that it may facilitate targeting of transient intermediates formed during viral mRNA capping. We therefore tested IFIT1 activity in extracts programmed with a Gppp-capped reporter (Fig.

3.8E). In this system, although overall translation is less efficient than in Cap0-programmed extracts, translation initiation still proceeds through binding of the cap-proximal nucleotides via eIF4G (De Gregorio et al., 1998). IFIT1 titration resulted in translational inhibition similar to m7Gppp-capped mRNA ($IC_{50} \approx 200$ nM), highlighting the importance of binding not only the cap, but also the proximal nucleotides to provide additional affinity to allow competition with eIF4F (Kumar et al., 2014).

Since cap binding does not rely on any guanine specific hydrogen bonds, we wondered whether IFIT1 could also recognize Appp-capped RNAs. Indeed, IFIT1 can bind Appp-RNA (**Fig. 3.8 F**) and inhibit translation initiation from an Appp-capped reporter ($IC_{50} \approx 500$ nM, **Fig. 3.8 E**). Thus, it appears that IFIT1 has evolved to recognize not only canonically capped mRNA, but rather

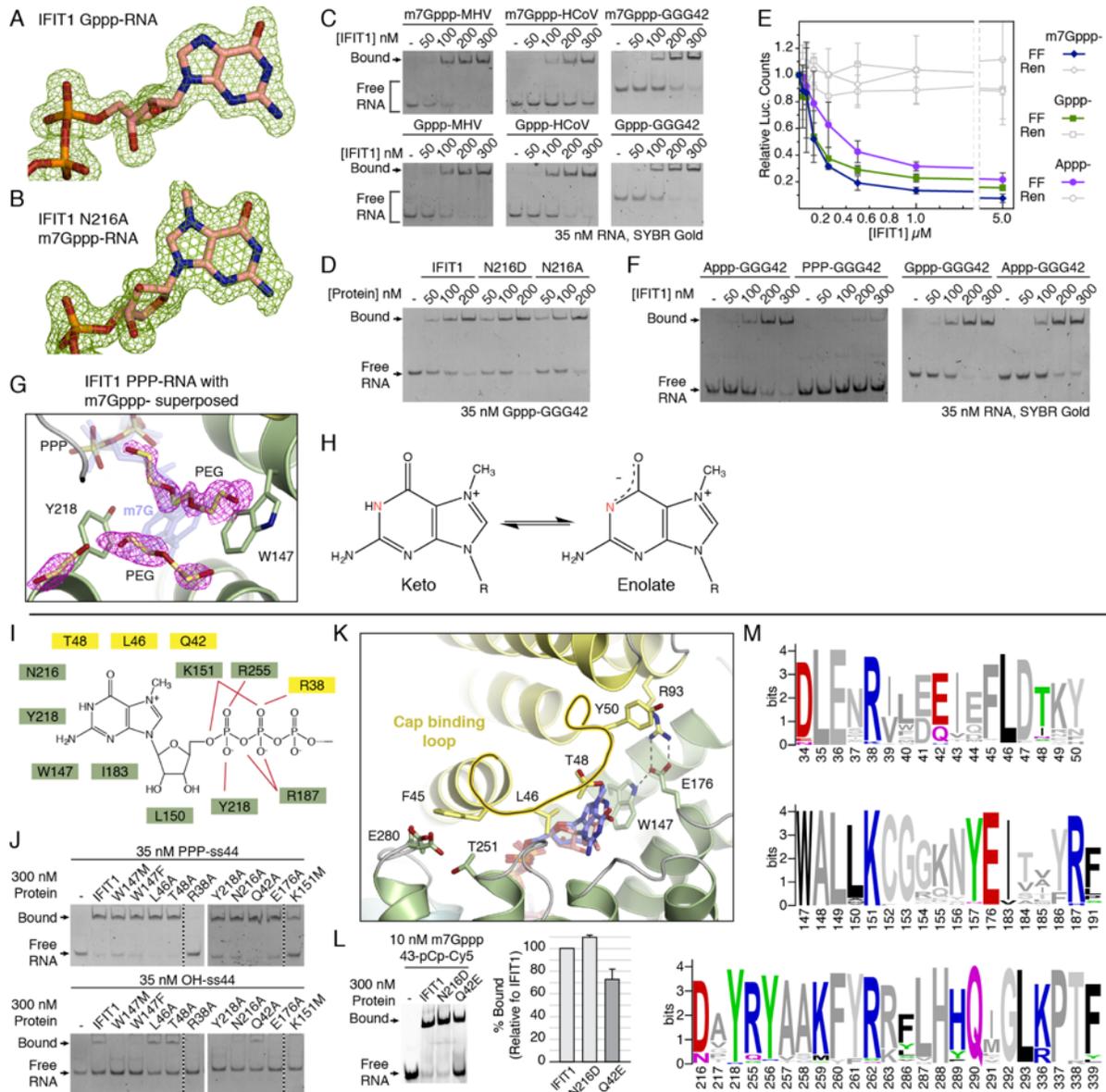


Figure 3.8 IFIT1 can accommodate multiple forms and conformations of the cap and mutational analysis of cap recognition

(A) Simulated annealing $2F_o-F_c$ omit map of the Gppp moiety contoured at 1σ . (B) Simulated annealing $2F_o-F_c$ omit map of m7Gppp moiety bound to monomeric IFIT1 N216A, contoured at 1σ . (C) EMSA between IFIT1 and m7Gppp- or Gppp-RNA. For HCoV, the lack of *N7*-methyl appears to enhance RNA binding. Whether this is a common feature of all IFIT1-like proteins remains to be validated. (D) EMSA between Gppp-GGG42 and IFIT1, IFIT1 N216D, or IFIT1 N216A. (E) *In vitro* translation assay with extracts programmed with Cap0-, Gppp-, or Appp-capped reporter. Data represent the mean of 2 independent measurements performed in duplicate \pm standard deviation. (F) EMSA between IFIT1 and Appp- or PPP-GGG42 (left), and between IFIT1 and Gppp- or Appp-GGG42 (right). (G) Simulated annealing $2F_o-F_c$ omit map of the PEG molecules inside the cap binding pocket in PPP-RNA bound IFIT1, contoured at 1σ . The m7G from m7Gppp-RNA bound IFIT1 is superposed in light blue. The PEG molecules interact non-specifically with residues lining the cap binding pocket. (H) At physiological pH, the m7G moiety exists in an equilibrium between a cationic ‘Keto’ tautomer, and a zwitterionic ‘Enolate’ tautomer in which *NI* (red atom) is deprotonated. (I) Schematic of residues lining the cap binding pocket and PPP channel. Residues are colored by subdomain, and red lines indicate hydrogen bonds or salt bridges. (J) EMSA between IFIT1 mutants and PPP-ss44 or OH-ss44. Except for R38A and K151M (which target the bridging triphosphate), all mutants retain binding to PPP-ss44, confirming that the protein fold is not affected by the mutations. Dashed lines indicate lanes that were cropped out. (K) Y50 and F45 from the cap binding loop are distal from the m7G moiety. F45 from subdomain I packs against E280 and T251 from subdomain II; while Y50 stacks against R93, which in turn is salt bridged to E176. These interactions may be important for subdomain contacts, and/or pre-organizing the cap binding loop. (L) EMSA between pCp-Cy5 labeled m7Gppp-43 and IFIT1, IFIT1 N216D, or IFIT1 Q42E, similar to Fig. 5A. (M) Composite WebLogo sequence consensus from 86 IFIT1- and IFIT1B-like genes showing residues involved in m7Gppp-RNA binding. See also Fig. 3.10.

diverse 5'-5' linked base modifications of the mRNA through relatively non-specific interactions in the pocket. The notion of lack of specificity is underscored by the structure of IFIT1 with PPP-RNA, where the cap binding pocket is occupied by PEG (polyethylene glycol) molecules from the crystallization solution, which form interactions that mimic cap binding (Fig. 3.8 G).

3.3.5 Mutational analysis of cap recognition.

To test whether our structural findings are functionally valid, we mutated several residues involved in cap binding and assayed them in fluorescent gel-shift binding assays (with m7Gppp-43 RNA, Table S2) and the translational inhibition assays described above (Fig. 3.9 A and B and Fig. 3.8 I). Both R38A and K151M are critical for binding, and K151M reduces IFIT1 inhibitory activity by 1-2 orders of magnitude ($IC_{50} > 5 \mu M$). Y218A and Q42A weakened binding to capped RNA and reduced IFIT1 inhibitory activity. N216A retains full binding to m7Gppp-RNA, indicating that the N216-cap interactions (e.g. with the *N7*-methyl) are dispensable. W147 is perhaps the most important residue inside the cap binding pocket as W147M essentially abolished m7Gppp-RNA

binding, while mutation to another aromatic residue (W147F) largely retained binding. Accordingly, W147M reduced IFIT1 translation inhibition ~ 40 -fold, and W147F mostly retained inhibitory activity (compared to W147M). Mutation of E176, which coordinates W147, had similar effects as W147F. From the cap binding loop, T48 was deemed dispensable, but L46 was required for optimal binding and translational inhibition. All cap binding pocket mutants tested here retained binding towards PPP-RNA, except for R38A and K151M (which target the PPP moiety), indicating that the protein fold was not disrupted by the mutations (**Fig. 3.8 J**). Taken together, the mutational analysis confirms the importance of cap binding and the role of the RNA binding tunnel in mediating translational inhibition by IFIT1.

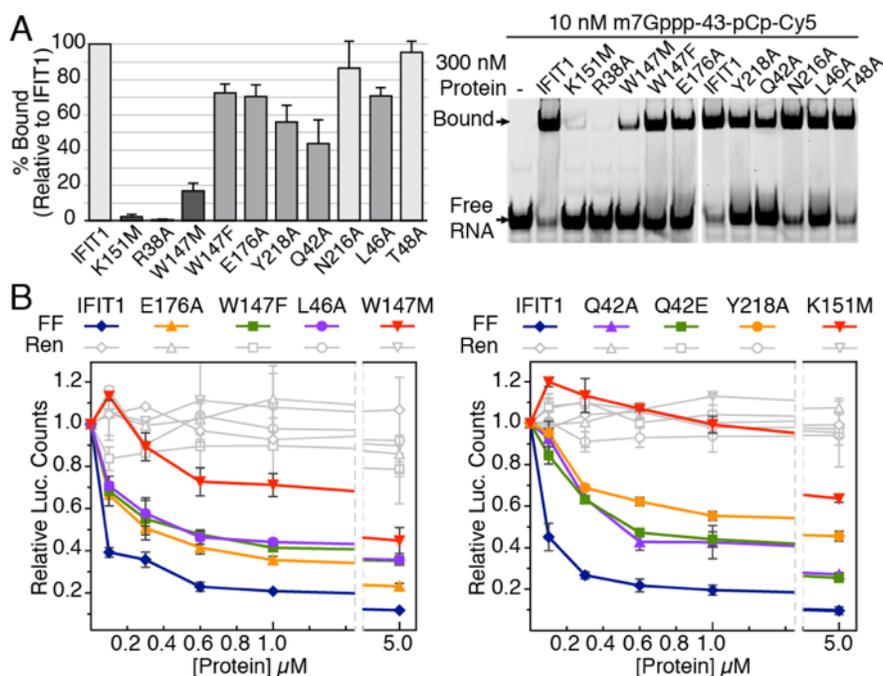


Figure 3.9 Functional validation of cap recognition

(A) Mutational analysis of cap binding residues investigated by fluorescent EMSA with 3'-end-labelled (pCp-Cy5) RNA. Left, quantification of % bound (upper band from right panel) for each mutant normalized against IFIT1. Data represent the mean of 3 independent measurements \pm standard deviation. (B) *In vitro* translation assays with RNA binding mutants and Cap0 reporter. Data represent the mean of 2 measurements \pm standard deviation.

Our results are in agreement with previous mutational binding assays based on *in silico* modeling (Kumar et al., 2014). In this model, Phe 45 and Tyr 50 from the cap binding loop were also predicted to interact with the base, however, our structures reveal that these two residues are distal from the m7G moiety and are probably important for maintaining subdomain contacts, or helping pre-organize the cap binding loop (**Fig. 3.8 K**).

3.3.6 The cap binding mechanism is conserved in IFIT1 and IFIT1B proteins across mammalian evolution.

The mode of cap-binding identified here likely applies to all mammalian IFIT1- and IFIT1B-like genes, as the residues involved in *N*7-methylguanosine triphosphate recognition are highly conserved (**Fig. 3.9 M and Fig. 3.10**). Two notably prevalent differences in cap binding residues compared to human IFIT1 include Q42 and N216, which are replaced with a glutamate and aspartate, respectively, in many of the orthologs and paralogs (including human IFIT1B). Both substitutions are conservative, since neither would disrupt hydrogen bonding patterns nor interfere with the van der Waals interactions with the cap. We tested this by carrying out an EMSA between m7Gppp-RNA and IFIT1 N216D or IFIT1 Q42E (**Fig. 3.9 L**). Whereas N216D had no impact on RNA binding, Q42E weakened the interaction and in translation assays, Q42E reduced IFIT1 activity similarly to Q42A (**Fig. 3.8 B, right**). In other IFIT1-like genes, such as rabbit IFIT1 and rabbit IFIT1B (both of which bind m7Gppp-RNA with ~ 20 and 10 nM affinity, respectively (Kumar et al., 2014)), the natural Q42E variation is likely overcome by compensatory interactions.

Unlike IFIT1B from other species, human IFIT1B lacks an apparent function in RNA binding (Daugherty et al., 2016). Sequence comparison shows that, along with Q42E, human IFIT1B has acquired additional substitutions that could impact RNA recognition: L150 is replaced with an Ala, which would affect cap ribose interactions, and R255 with Gln, which would disrupt a salt-bridge with the γ -phosphate (**Fig. 3.5 E and F**). Supporting this, mutation of R255 in human IFIT1 (R255M) was shown to disrupt capped- and PPP-RNA binding (Abbas et al., 2013; Habjan et al., 2013). On the other hand, mice lack a *bona fide* IFIT1 ortholog and instead contain three copies of IFIT1B-like genes (Daugherty et al., 2016), currently annotated as mouse *Ifit1*, mouse *Ifit1b* and mouse *Ifit1c* (**Fig. 3.10**). Mouse *Ifit1b* and *Ifit1c* harbor several substitutions that would disrupt RNA binding, such as R255G and Q42T in both, and R187H in *Ifit1b* (**Fig. 3.10**). Consistent with this, pull-downs showed that mouse *Ifit1c* cannot bind capped RNA directly (Habjan et al., 2013).

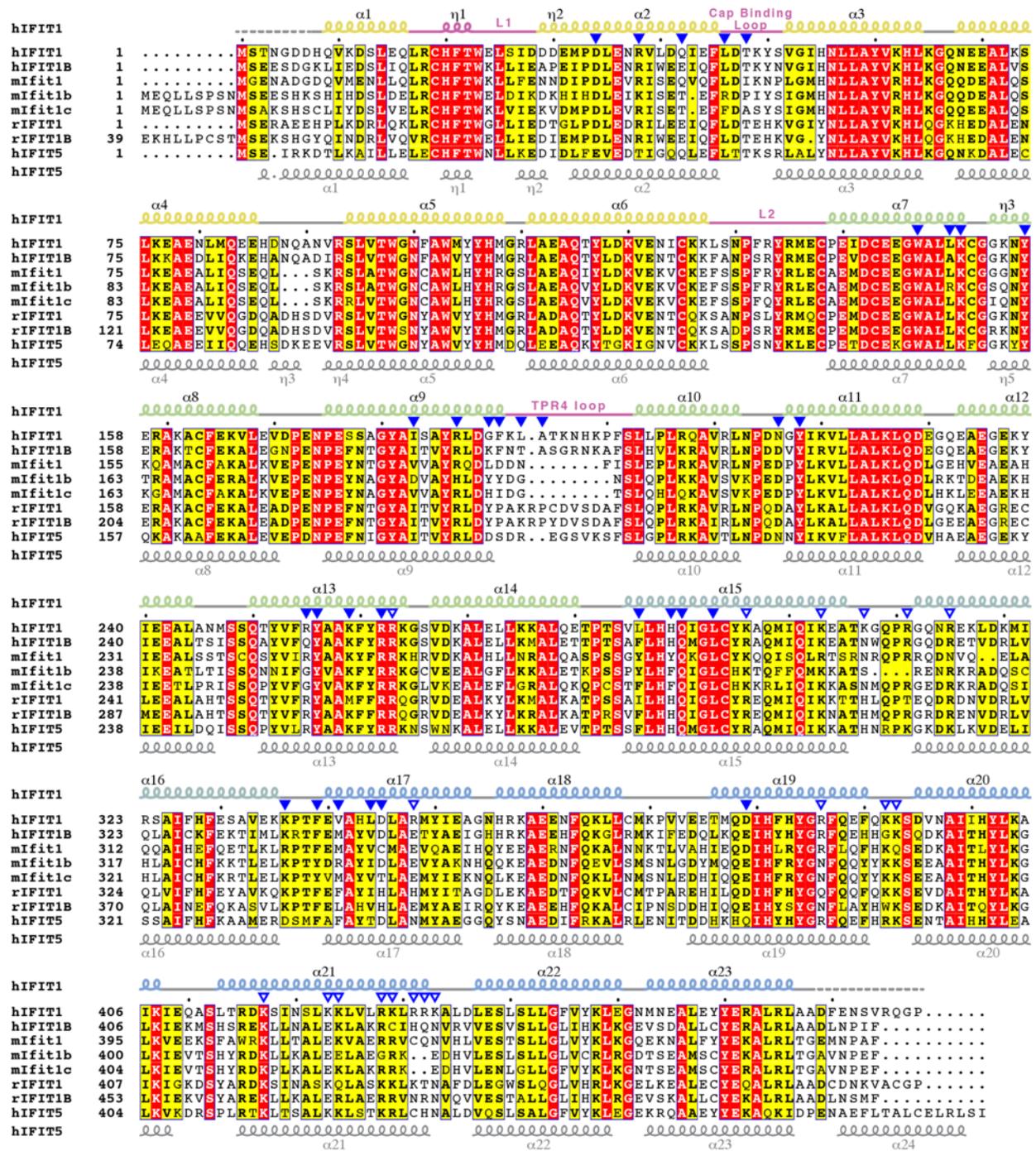


Figure 3.10 ClustalO sequence alignment of select IFIT genes from humans, mice, and rabbits

Filled triangles indicate residues directly in contact with the RNA. Open triangles are positively charged residues lining the protein surface outside of the tunnel (in the groove). Secondary structure elements are indicated for human IFIT1 (colored by subdomain) and human IFIT5 (4HOR). This figure was prepared using ESPrift (<http://esprift.ibcp.fr/ESPrift/ESPrift/>). NCBI reference numbers for mFit1, mFit1b, and mFitc are NP_032357.2, NP_444447.1, and NP_001103987.1, respectively

3.3.7 IFIT1 cap binding is distinct from canonical cap binding proteins.

The IFIT1 cap binding mechanism described here is quite distinct from canonical cap binding proteins such as eIF4E (Marcotrigiano et al., 1997), CBC (Cap Binding Complex (Mazza et al., 2002)), and VP39 (vaccinia virus N1 2'-O MTase, (Hodel et al., 1997)). Through convergent evolution these proteins evolved a highly specific cap binding slot between two aromatic side-chains that engage the methylated guanine in a cation- π sandwich (Quiocho et al., 2000). Charge-charge interactions with the delocalized positive charge and van der Waals contacts with the *N7*-methyl also play a role (Quiocho et al., 2000). In these proteins, the absence of *N7*-methylation and associated positive charge on the base results in > 100-fold loss in binding affinity (Fechter and Brownlee, 2005). These proteins also rely on hydrogen bonds targeting groups at the m7G base periphery. The cumulative effect of these restrictions results in highly specific recognition of the cap in a single *anti*- conformation of m7G (Fig. 3.11).

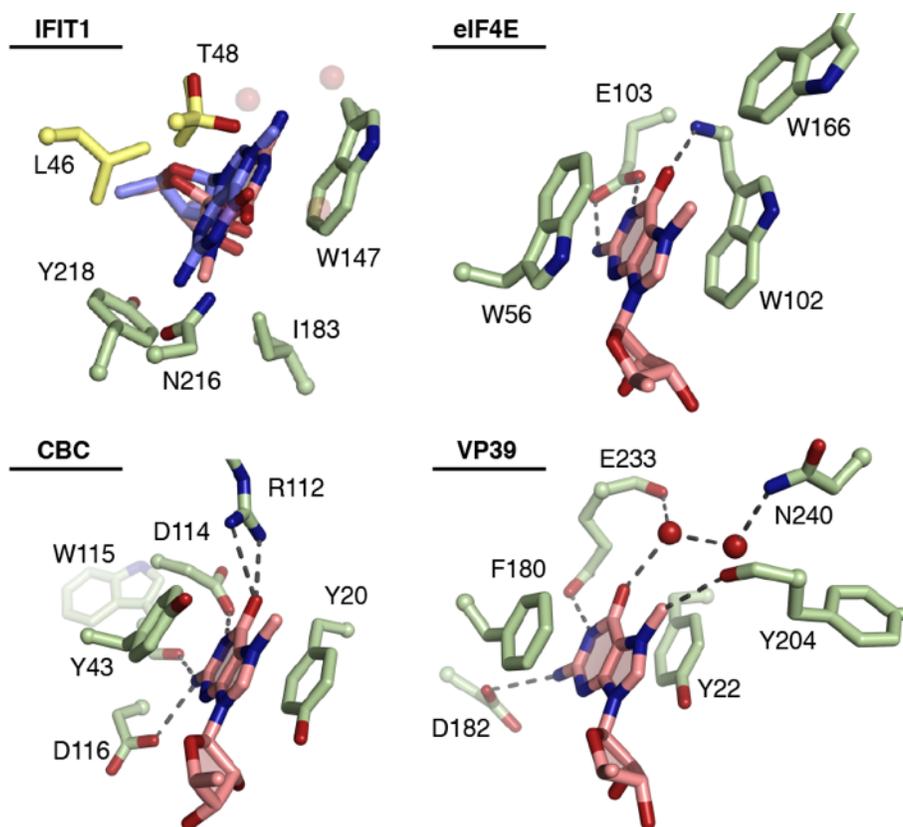


Figure 3.11 Comparison to the canonical cap binding proteins

eIF4E (PDB 1EJ1), CBC (PDB 1H2T), and VP39 (1V39). In IFIT1, some of the nearby water molecules are shown as transparent red spheres, and T48 adopts two conformations.

By contrast, IFIT1 engagement of the cap is relatively less specific, permitting both *syn*- and *anti*-base orientations as described above. Although IFIT1 does utilize one aromatic residue for cap stacking, the remainder of its sandwich is formed by aliphatic side chains (Leu 46, Thr 48, and Ile 183) rather than another aromatic residue. Therefore, the lack of an electron-rich, aromatic cap binding slot reduces the dependence on an electron deficient, *N7*-methylated base. Additionally, protein contacts with the *N7*-methyl are dispensable (e.g. the N216A mutant), and they are altogether absent when the cap is in the *syn*- configuration. Importantly, IFIT1 lacks any sequence specific hydrogen bonding from protein residues, and instead utilizes a more plastic, water-mediated hydrogen bonding network for base recognition.

Finally, at physiological pH m7G exists in equilibrium between two forms: a positively charged ‘keto tautomer’ and a zwitterionic ‘enolate tautomer’ (in which *N1* is deprotonated, **Fig. 3.8 H**, (34)). The canonical cap binding proteins are highly selective for guanine as the base and in particular, its keto form. These aspects are enforced by two elements, 1) the cation– π sandwich, which is only compatible with an electron-deficient, positively-charged keto tautomer (Hu et al., 2003), and 2) Asp or Glu residues at one end of the cap binding slot, which hydrogen bond with a protonated *N1* and the *N2*-amino group. Conversely, IFIT1 does not form any keto- or enolate-specific interactions with the base, suggesting that IFIT1 is not selective for the tautomerization state, reinforcing the lack of guanine specificity.

3.3.8 Binding of cap-proximal nucleotides.

Recognition of the four RNA nucleotides following the cap is conformation specific and can be divided into two distinct dinucleotide groups diverging between N2 and N3 (**Fig. 3.2 C-E**). The RNA backbone lies along the superhelical axis of the protein, and is recognized by specific hydrogen bonds and salt bridges from protein residues targeting the 5'-phosphates and 2'-hydroxyls of N2 and N3 (**Fig. 3.12 A and B**). By contrast, recognition of the bases is predominantly through sequence non-specific van der Waals and stacking interactions. The first dinucleotide (N1 and N2) adopts geometry similar to CpG dinucleotides found within Z-form RNA and UUCG tetraloops (D'Ascenzo et al., 2016), and is tightly sandwiched between multiple protein residues (**Fig. 3.12 A and C**). The second dinucleotide (N3 and N4) adopts A-form helical geometry with the bases also stacked upon each other and abutted by protein residues from above and below (**Fig. 3.12 B and D**). The large water network inside the tunnel interacts with all groups of the RNA, and mediates both intermolecular protein-RNA and intramolecular RNA-RNA

interactions (**Fig. 3.12 E**). Interestingly, a large part of this extensive water network is involved in protein-base contacts, allowing IFIT1 to recognize a wide variety of RNA sequences that may exist at the 5'-end of viral RNA. A small degree of sequence dependent binding affinity variation may exist since there are two adenine-specific hydrogen bonds at N2 and N4 (**Fig. 3.12 A and B**), although RNA binding assays show that adenosines are not strictly required at these positions.

The 3'-end of the RNA (N4) emerges from the C-terminal opening of the tunnel and points towards a positively-charged, solvent-exposed groove formed by the pivot helices and the third subdomain (**Fig. 3.13**). This surface is contiguous with the RNA binding tunnel and also appears to contribute to RNA interactions as the analogous region in IFIT5 can apparently bind tRNA (Katibah et al., 2013). In IFIT1, the groove does play some role in RNA binding, as primer-extension toe-printing assays suggested that IFIT1 has a 6-8 nt footprint at the 5'-end of mRNA, and mutational analysis of this region had an impact on mRNA binding (Kumar et al., 2014).

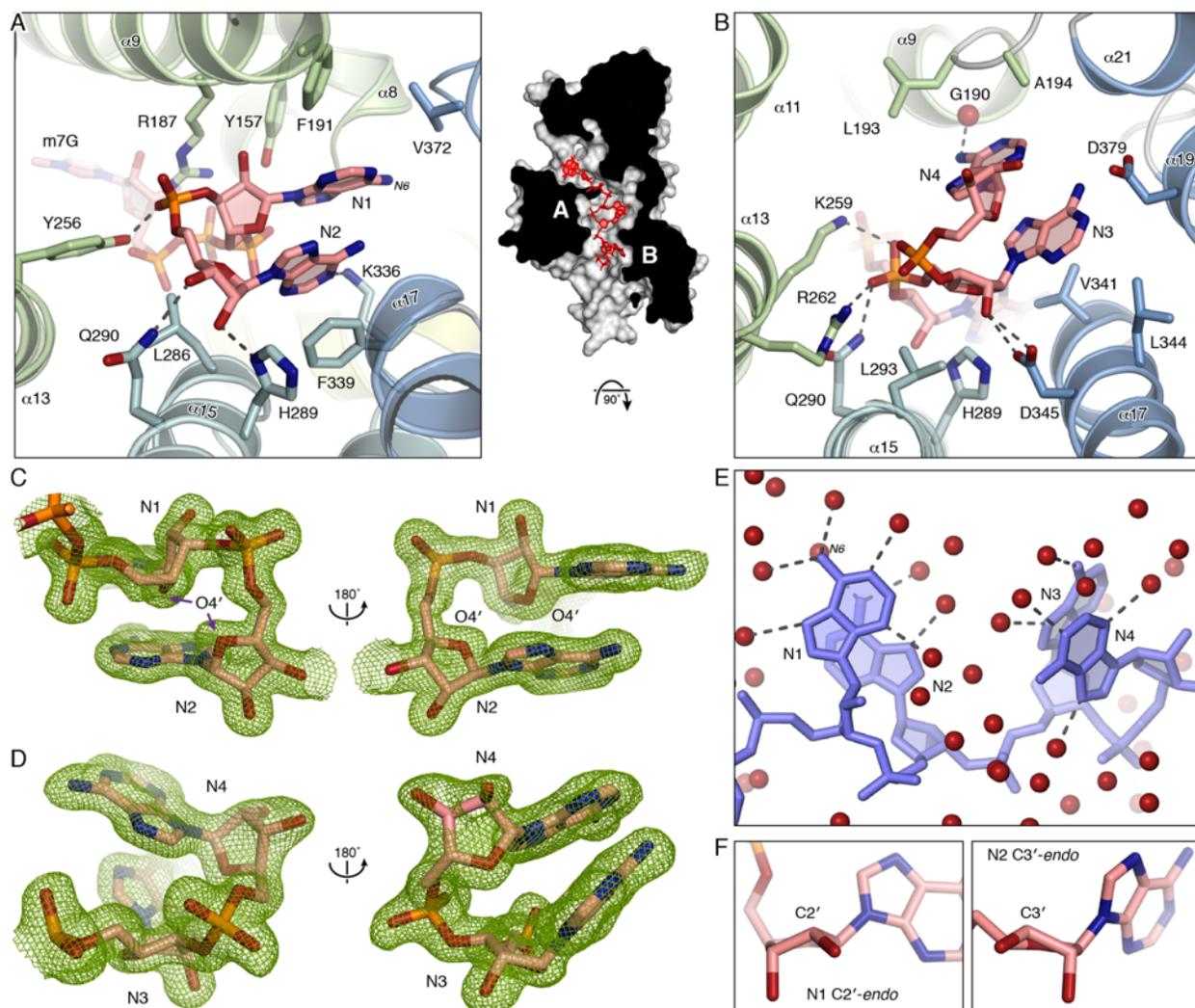


Figure 3.12 RNA binding mechanism at N1-N4

(A) Interactions between IFIT1 and RNA at N1/N2 and (B) at N3/N4. The protein is oriented such that we are looking along the RNA binding tunnel from the 3'-exit (similar orientation as Fig. 3.4 C). Note that K336 and the backbone carbonyl of G190 are H-bonded to N2 and N4 respectively. Concerning Fig. 3.14 and the multiple roles of R187, Y157, H289, and Q290: R187 packs against the ribose of N1, coordinates Y218 and Y157 (Fig. 3.6 G), and interacts with m7G and bridging PPP (Fig. 3.5 E and F); Y157 contacts the N1 ribose and N1 adenine through Van der Waals; Q290 is H-bonded to the 3' oxygen of N2 and the inter-nucleotide phosphate between N2 and N3; H289 is H-bonded to the 2'-OH of N2, and contacts the ribose of N3 via Van der Waals. V372 could potentially clash with methylation at the *N6* position of the first adenosine (see discussion). (C) Simulated annealing $2F_o - F_c$ omit map of N1/N2 contoured at 1σ . Two views rotated by 180° are shown. The dinucleotide conformation here resembles that found in CpG dinucleotides of Z-form RNA and UUCG tetraloop sequences. The defining feature of this rare dinucleotide motif is the antiparallel arrangement of the two riboses (with their respective O4' atoms pointing towards each other), and a lone pair- π stack between the O4' atom of N1 and the base of N2. (D) Simulated annealing $2F_o - F_c$ omit map of N3/N4 contoured at 1σ . N3/N4 adopt standard A-form helical geometry. (E) Hydrogen-bonds between the RNA bases and the waters (red spheres) inside the tunnel. *N6* methylation of the first adenosine could disrupt water mediated interactions at the first nucleotide (see discussion). (F) N1 and N2 adopt C2'-endo and C3'-endo conformations respectively. RNA nucleotides are typically in an equilibrium between the two, but generally favor the C3'-endo conformation.

However, co-crystal structures of IFIT1 with longer oligonucleotides (6-8 nt in length) revealed extra electron density for only the 5'-phosphate of a fifth nucleotide, as was shown for IFIT5 (Abbas et al., 2013), suggesting that only the first four nucleotides are stably bound by IFIT1, while residues in the positively charged groove probably contribute to non-specific RNA binding.

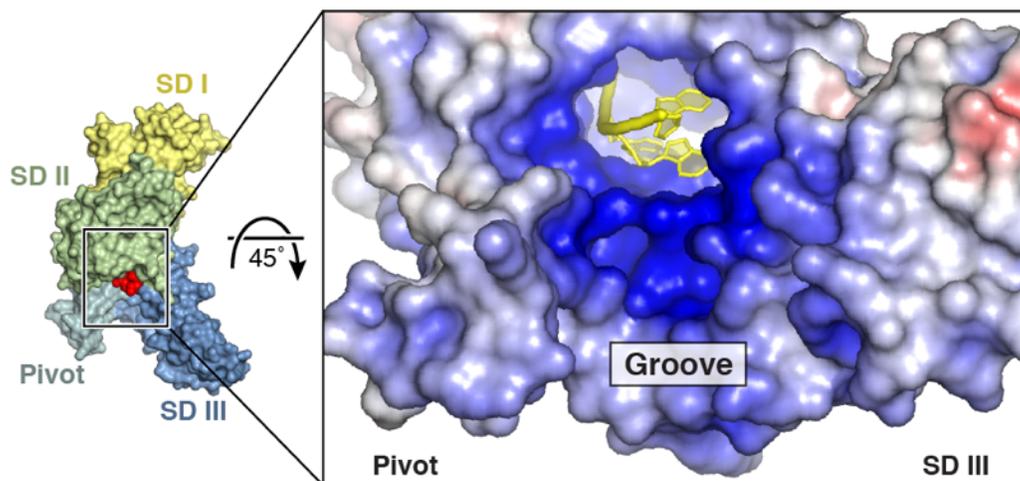


Figure 3.13 IFIT1 forms a positively-charged, solvent-exposed RNA binding groove

See Fig. 3.10 for residues in this region.

3.3.9 IFIT1 senses ribose 2'-O methylation at N1 and N2.

The mRNA of higher eukaryotes is normally modified by ribose 2'-O methylation at N1 and N2 (Banerjee, 1980). Whereas all cellular mRNAs are methylated at N1 in the nucleus by the endogenous Cap1-methyltransferase (CMTr1) (Bélanger et al., 2010), ribose methylation at N2 arises from secondary methylation in the cytoplasm through the action of CMTr2 (Perry and Kelley, 1976; Werner et al., 2011), and accompanies N1 methylation on up to 50% of cellular mRNAs (Cleaves and Dubin, 1979; Wei and Moss, 1975). N1 ribose methylation is a molecular determinant of self which can protect mRNA from IFIT1/IFIT1B recognition (Habjan et al., 2013; Kumar et al., 2014), but the role of N2 methylation in this process is unknown. To gain additional insight into self- vs non-self-mRNA discernment by IFIT1, and to explore the uncharacterized role of ribose N2 methylation in this process, we examined the interaction between human IFIT1 and differentially methylated RNA. Note that, to distinguish the naturally occurring Cap1 and Cap2 structures (m7GpppNmN- and m7GpppNmNm-) from capped RNAs that contain ribose N2 methylation only (m7GpppNNm-), we refer to the latter as Cap0^{N2Me}-RNA.

When bound to IFIT1, N1 and N2 adopt a rare Z-RNA like conformation that is dependent on their respective ribose conformations (**Fig. 3.12 F**) (D'Ascenzo et al., 2016). Whereas N2 is in the favorable C3'-*endo* conformation, N1 adopts a C2'-*endo* conformation and places its 2'-OH in close proximity to the side chains of two highly conserved residues, R187 and Y157 (**Fig. 3.14 A and Fig. 3.12 A**). Modeling of ribose 2'-O methylation on N1 to mimic Cap1-mRNA shows that the methyl group would clash with these protein residues (**Fig. 3.14 A**). Rotating the methyl group away introduces a steric clash with the RNA itself and interferes with the water network. Interestingly, N2 ribose methylation is also predicted to disrupt RNA binding to IFIT1, because of hydrogen bonds with H289 and steric hindrance by Q290 (**Fig. 3.14 A and Fig. 3.12 B**). Thus, the IFIT1 tunnel is restricted to interact with RNAs not methylated at these 2'-hydroxyls.

Consistent with this, either N1 or N2 methylation of HCoV and GGG42 RNAs is sufficient to disrupt binding with up to 2.5 μ M IFIT1 (**Fig. 3.14 B and Fig. 3.15 A and B**). Surprisingly, at the same concentrations, individual N1 or N2 methylation only partially reduced the interaction between IFIT1 and MHV RNA, and combining both was required to fully abolish binding for this sequence (**Fig. 3.14 B and Fig. 3.15 A-C**). This RNA-dependent effect is likely due to the longer overhang and decreased secondary structure stability of the MHV sequence compared to HCoV/GGG42 (**Fig. 3.1 A and Fig. 3.2 B**), allowing it to maintain relatively strong binding to

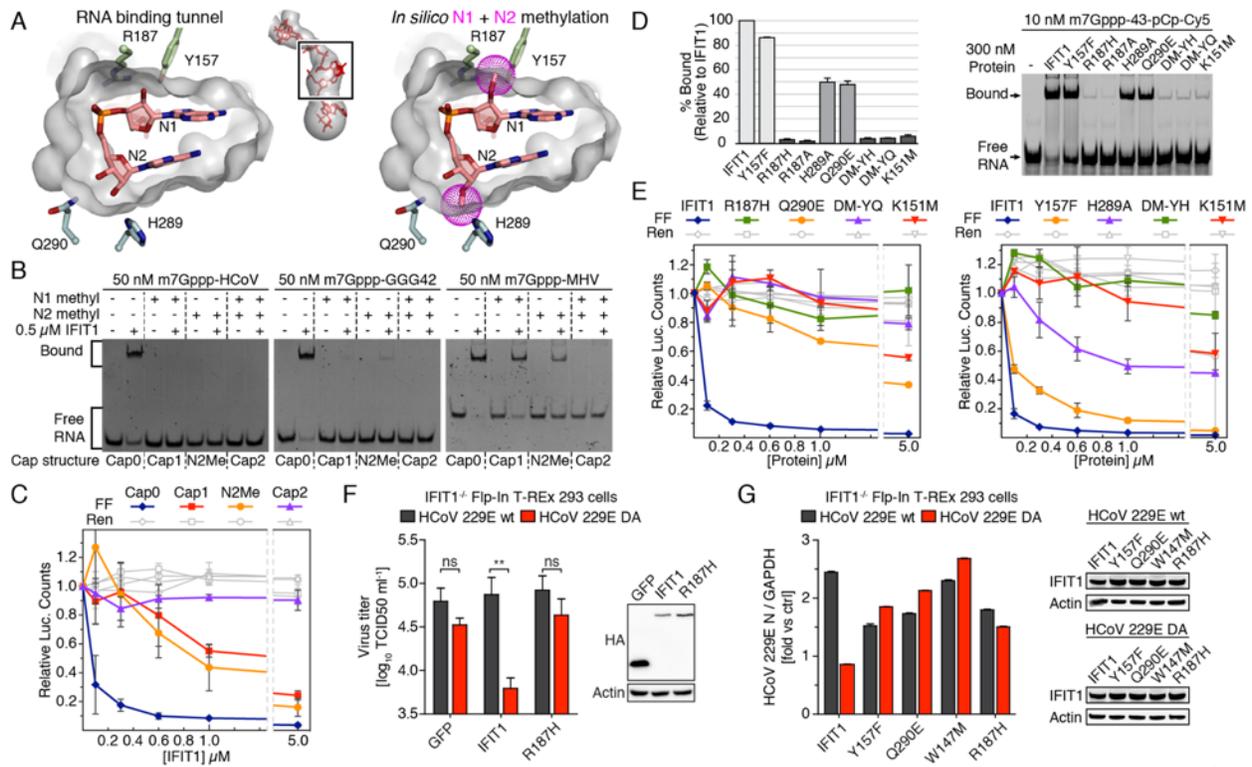


Figure 3.14 Sensing of N1 and N2 ribose 2'-O methylation by IFIT1

(A) Left, cross-section of the IFIT1 tunnel van der Waals surface (grey) and residues predicted to clash with N1 and N2 methylations. Right, *in silico* rigid body modeling of N1 and N2 methylations (purple dots). (B) SYBR Gold stained EMSAs between 0.5 μ M IFIT1 and differentially methylated m7Gppp-RNA. The dashed lines demarcate lanes with different cap structures, as indicated by the labels below the gel. See Fig. 3.15 A-E for additional gel shifts. (C) Translation assay with differentially methylated reporter mRNA. Data represent the mean of 3 independent measurements performed in duplicate \pm standard deviation. (D) Mutational analysis of 2'-O methyl sensing residues investigated by fluorescent EMSA similar to Fig. 3.9 A. (E) *In vitro* translation assays with 2'-O methyl sensing mutants and Cap0 reporter. Data represent the mean of 2 measurements \pm standard deviation. (F) Flp-In T-REx 293 IFIT1 knockout cells were co-transfected with expression plasmids for CD13 and either GFP control, wild-type IFIT1, or IFIT1 R187H and infected with HCoV 229E wt or HCoV 239E DA. Virus titers in supernatants were determined by TCID50 18 h post infection. Data represent the mean of three independent experiments \pm standard deviation. **** $P < 0.01$** as analysed by two-way ANOVA with Bonferroni post-test. The westerns show expression of proteins at the time of infection. (G) Similar to (F), except IFIT1 knockout cells were reconstituted with the indicated IFIT1 constructs or GFP, and virus growth determined by quantitative PCR. Data represent the mean fold change \pm standard deviation of triplicate measurements of the viral N-gene signal, relative to GFP control (ctrl). One representative experiment of three is shown. Western blots show protein expression at the end of the experiment.

IFIT1 even when modified with a single ribose methylation, and requiring the additive effect of multiple methylations (N1+N2) to avoid IFIT1 recognition. However, we cannot rule out the existence of RNA-sequence or other structural elements within MHV that intrinsically enhance its

affinity for IFIT1. At higher protein concentrations (5 μ M IFIT1) where non-specific interactions may play a role, the additive effects of N1+N2 ribose methylations also became apparent for the GGG42 RNA (**Fig. 3.15 D and E**).

Consistent with the above, single N1 or N2 methylation of the reporter mRNA reduced IFIT1 inhibitory activity in translation assays by ~ 10 -fold ($IC_{50} \approx 1 \mu$ M, **Fig. 3.14 C**), while translation of Cap1- and Cap0^{N2Me}-mRNA was still strongly inhibited by 5 μ M IFIT1. This intriguingly suggests that N1- or N2-methyl steric hindrance can be overcome at very high IFIT1 concentrations, possibly from non-specific RNA interactions contributed by the solvent exposed groove of IFIT1. As before, combining N1 and N2 methylations (Cap2 reporter) resulted in a striking rescue of translational inhibition, restoring FF levels to 90% even in the presence of 5 μ M IFIT1 (**Fig. 3.14 C**). Taken together, our combined structural and functional analysis confirms the role of N1 methylation in interfering with IFIT1 inhibitory activity, and reveals an analogous function for N2 methylation. Importantly, our data suggest that the combination of N1 and N2 methylation, as found in nearly half of endogenous mRNAs, produces an additive and potentially synergistic protective effect against IFIT1 recognition, which is particularly evident under circumstances where IFIT1 can overcome individual 2'-O methylation in an RNA-dependent or protein concentration-dependent manner.

To further confirm the importance of 2'-O methyl sensing for IFIT1 activity, we mutated the residues predicted to clash with N1 or N2 ribose methylations, and tested their impact on RNA binding and translational inhibition (**Fig. 3.14 D and E**). At N1, Y157F had only a minor effect on capped RNA binding, but both R187H and R187A abolished the interaction. At N2, mutating either H289 (H289A) or Q290 (Q290E) partially reduced binding, and combining either mutant with Y157F (DM-YH, Y157F/H289A; or DM-YQ, Y157F/Q290E) completely disrupted binding. Translation assays also showed reductions in IFIT1 inhibitory activity for all mutants, with R187H and the two double mutants having the greatest effect. It should be noted, however, that these residues are highly conserved (**Fig. 3.8 M**), and as such, play an integral role in general RNA binding that extends beyond 2'-O methyl sensing (**Fig. 3.12 A and B**). Thus, capped-RNA recognition and 2'-O methyl sensing by IFIT1 are two tightly linked processes that have likely co-evolved.

Figure 3.15 Sensing of N1 and N2 ribose 2'-O methylation by IFIT1 cont'd

(A) IFIT1 senses N1 and N2 ribose 2'-O methylation (A and B) EMSAs between 1 μ M or 2.5 μ M IFIT1 and differentially methylated m7Gppp-RNA (C) Comparison of IFIT1 binding to 35 nM Cap0-MHV, Cap1-MHV, or N2Me-MHV. Cap1 methylation reduces the apparent affinity to m7Gppp-MHV by \sim 4 fold, as the apparent K_d for Cap1-MHV is in the 200-300 nM range, whereas the apparent K_d for Cap0-MHV is \sim 75 nM (see Fig. 3.1 A). (D) EMSAs between 5 μ M IFIT1 and differentially methylated m7Gppp-RNA. Note that at these protein concentrations, a second band is more prominent at the top of the gel, which is likely non-specific or higher-order interactions. For Cap2-GGG42 and Cap2-MHV, the red arrowhead points to the smeared band corresponding to unbound RNA. (E) Similar to (D), except a higher concentration of RNA was used to improve the staining sensitivity. This gel shift demonstrates a clear additive effect of N1+N2 for GGG42. (F) SYBR Gold stained EMSA between IFIT1 mutants targeting 2'-O methyl sensing residues and Cap0-HCoV RNA. (G) 1X TBE, 17% 8M urea denaturing PAGE of RNA used in this study. (H-I) Comparison between IFIT1 and RIG-I self vs non-self discernment of capped RNAs (H) In addition to recognizing base-paired, blunt-ended RNAs with a 5'-PPP, RIG-I is activated by m7Gppp/Cap0 containing RNAs. Recognition of viral, self, or synthetic Cap0-RNAs by RIG-I leads to IFN production, and IFIT1 and other ISG upregulation. In contrast, IFIT1 targets single-stranded mRNAs, or structured mRNAs if they have an \sim 4 nt overhang, leading to inhibition of mRNA translation. (I) N1 methylation (i.e. Cap1-RNA) has been shown to abolish RIG-I activity, whereas N2 methylation alone has a partial effect; the combination of N1+N2 (i.e. Cap2-RNA) has not been tested with RIG-I. In contrast, N1 and N2 methylation alone have a similar effect on IFIT1 RNA binding, and N1+N2 methylation has an additive and potentially synergistic effect in preventing RNA binding, which may be important for protecting inherently susceptible self-mRNAs from IFIT1 recognition. (J) RIG-I is a multi-domain protein that binds RNA through a central helicase domain (not drawn) and a C-terminal Regulatory domain (RD). The RIG-I RD houses a basic cleft which recognizes the blunt-end, the bridging PPP, and the 2'-OH of N1, but makes no contacts with the m7G moiety. From the RD, H830 hydrogen-bonds with the N1 2'-OH to prevent binding of Cap1-RNA, while C829 makes minimal contacts with N2 but is predicted to clash with methylation at this position. Mutation of H830A does not affect Cap0-RNA binding but impairs RIG-I ability to discern self (Cap1) from non-self (Cap0) RNA. IFIT1 cap recognition differs, utilizing a narrow tunnel that makes contacts with the m7G, the bridging PPP, and the cap-proximal nucleotides. Several conserved residues mediate 2'-O methyl sensing and general RNA binding, and their mutation is deleterious for IFIT1 activity. These differences between IFIT1 and RIG-I reflect their complementary roles in self vs non-self-mRNA discernment.

3.3.10 Functional validation of IFIT1 activity against 2'-O methyltransferase deficient human coronavirus.

Human IFIT1 has been shown to inhibit replication of viruses lacking N1 ribose 2'-O methylation, such as HCoV 229E bearing a D129A mutation in its viral 2'-O MTase gene (HCoV 229E DA, (Habjan et al., 2013)). Therefore, to functionally validate our results in a biological context, we tested the antiviral activity of IFIT1 RNA-binding mutants against wild-type (wt) HCoV 229E and HCoV 229E DA. First, we verified that the IFIT1 mutants utilized in cell-based assays (R187H, W147M, Y157F, and Q290E) disrupted the interaction with Cap0-HCoV RNA (Fig. 3.15 F). Next,

we reconstituted Flp-In T-REx 293 IFIT1 knockout cells with human IFIT1 or IFIT1 mutants, and assayed HCoV growth in these cells (**Fig. 3.14 F-G**). While expression of a control protein (GFP) led to comparable accumulation of both wt and DA virus in the supernatant of infected cells, expression of IFIT1 significantly reduced growth of the DA mutant virus, but not wt virus (**Fig. 3.14 F**). In contrast, IFIT1 R187H, which disrupts interactions with the cap ribose and bridging triphosphate (**Fig. 3.5 E-F**), was unable to impair HCoV 229E DA virus growth (**Fig. 3.14 F**). Similarly, W147M, which disrupts cap recognition, or Y157F and Q290E, which impair binding to Cap0-HCoV RNA (**Fig. 3.15 F**), lost their antiviral activity against HCoV 229E DA (**Fig. 3.14 G**). Thus, IFIT1 binding to 2'-O unmethylated viral RNA is required for its antiviral properties.

3.4 Discussion

The ability of many viruses to cap their mRNA and mimic the host's allows them to hijack a cell's translational machinery and replicate new virus particles. To counteract this, host cells have evolved as part of their antiviral program, the IFIT proteins. By competing with eIF4E/eIF4F for binding to capped RNA, IFIT1 can prevent viral propagation by latching onto the ends of mRNA and preventing assembly of ribosomal initiation complexes (Habjan et al., 2013; Kumar et al., 2014). Whereas recognition of the cap by eIF4E and other cap binding proteins occurs in a highly specific manner (Fechter and Brownlee, 2005; Quijcho et al., 2000), we surprisingly found that recognition of the cap moiety by IFIT1 is instead non-specific with regards to both sequence and structure. Through its highly water-filled cap binding pocket, IFIT1 can accommodate not only *bona fide* cap in different orientations, but also unmethylated cap, adenine cap, and presumably other structures too. This built-in plasticity may in part be to allow IFIT1 to maintain a broad spectrum of antiviral activity, and to thwart the ability of viral structures to rapidly evolve. Another possibility is that IFIT1 genes simply have not had enough time to evolve exquisite cap specificity, since they emerged relatively recently in evolution (jawed vertebrates (Daugherty et al., 2016)) compared to eIF4E and CBC, which are essential genes in all eukaryotes (Marcotrigiano et al., 1997; Mazza et al., 2002). Regardless, the penalty for this plasticity is likely a reduction in affinity for the cap moiety and in this respect, the recognition of nucleotides beyond the cap provides IFIT1 the additional affinity required to compete with an otherwise very tight eIF4F-5'-cap complex.

The recognition of cap-proximal nucleotides by IFIT1 also plays a critical role in discerning self from non-self. Our structural analysis revealed that IFIT1 forms a tight interacting

surface around the ribose 2'-hydroxyls of N1 and N2, thus preventing recognition of endogenous mRNAs methylated at these positions and restricting IFIT1 activity to unmethylated viral mRNAs. This is supported by a comprehensive gel shift analysis which showed a preference for recognizing Cap0 structures over N1 or N2 methylated RNA, *in vitro* translation assays which showed a reduction in IFIT1's ability to inhibit translation of N1 or N2 methylated mRNA reporters, and human coronavirus infectivity assays which showed enhanced IFIT1 antiviral activity when viral N1 methylation was mutated. In this way, IFIT1 effector function complements RIG-I receptor activity (summarized in **Fig. 3.15 H-J**), as RIG-I detects blunt-ended, base-paired PPP- and Cap0-RNAs to upregulate IFIT1 and other ISGs (Devarkar et al., 2016; Schuberth-Wagner et al., 2015).

Cellular N1 methylation was generally thought to be the primary determinant of 'self' protecting endogenous mRNA from IFIT1 recognition, but we discovered that the ability to discern between Cap0 and Cap1 structures is diminished for one of our RNAs (MHV, **Fig. 3.14 B**), an effect which is possibly linked to its 5'-sequence or secondary structure. This finding is in accordance with a recent study by Daugherty *et al.*, which demonstrated that human IFIT1 can target both Cap0 and Cap1 mRNAs when overexpressed in a yeast system (Daugherty et al., 2016). Furthermore, while this manuscript was under revision, Young *et al.* similarly showed that N1 methylation of a reporter gene only partially reduced its sensitivity to human IFIT1 in an *in vitro* translation system (Young et al., 2016). Interestingly, using viral mRNA in the same system, Young *et al.* also noted potential RNA-dependent effects for N1 methylation. Altogether, these observations lead to the conclusion that N1 methylation alone may not be enough to protect all endogenous mRNA from IFIT1, and that there are other determinants of 'self' that govern IFIT1 activity.

Our structural and functional analysis reveals that N2 methylation by the Cap2-MTase (CMTr2) could fulfil this role, providing an additional safeguard against aberrant recognition of mRNAs that are otherwise susceptible to IFIT1 (**Fig. 3.14 B**). However, as Cap2 structures are not as ubiquitous as Cap1 (Banerjee, 1980), other elements may prevent self-recognition. For instance, actively translating mRNAs are generally found in pre-formed mRNP complexes and, as indicated by our order of addition experiment (**Fig. 3.1 C**), would be protected from IFIT1 competition. Similarly, newly synthesized mRNAs undergo a pioneer round of translation directed by the CBC (Maquat et al., 2010), which may also offer protection from IFIT1. Adenosine N6-methylation of the first transcribed nucleotide is another modification which accompanies ribose N1 methylation

on ~ 20-30 % of cellular mRNAs, in the form of *N*6,O2'-dimethyladenosine (m6Am) (Wei et al., 1976; 1975). Our structure suggests that m6Am could protect self-mRNAs by disrupting water-mediated interactions and impinging on nearby residues (**Fig. 3.12 A and E**), and thus merits further investigation. Finally, cap-proximal secondary structure could combine with mRNA modifications to further prevent self-recognition.

To what extent do viruses exploit these mechanisms to alter IFIT1 activity? Our observations support the model whereby viral N1 methylation evades or dampens IFIT1 activity (**Fig. 3.14 F and G**), and is consistent with previous studies showing enhanced sensitivity of coronaviruses and flaviviruses to IFIT1 when viral N1 methylation was mutated (Habjan et al., 2013; Menachery et al., 2014; Pinto et al., 2015; Zust et al., 2013). Similarly, Young *et al.* recently showed that Parainfluenza virus type 5 (PIV5) was more sensitive to human IFIT1-mediated restriction than PIV3, partly because PIV5 mRNAs were not completely N1 methylated during infection (Young et al., 2016). On the other hand, vesicular stomatitis virus (VSV) has been shown to uniformly N1-methylate its mRNAs *in vivo* (Liang et al., 2015; Rose, 1975), yet, it remains sensitive to IFIT1 restriction (Daugherty et al., 2016). One explanation is that the short and potentially unstructured 5'-untranslated regions of VSV mRNAs could allow IFIT1 to overcome N1 methylation (as described here for MHV RNA). Alternatively, VSV mRNAs may display another pattern specifically recognized by IFIT1 (as proposed by Daugherty *et al.*), such as RNA-sequence or -structural elements (Daugherty et al., 2016). Further work is needed to validate either notion, and to determine if this ability to overcome viral N1 methylation is an adaptation that allows IFIT1 to target other Cap1-containing viruses, or if it is restricted to VSV and related viruses.

IFIT1B (which includes mouse *Ifit1*) is the only other IFIT family member known to specifically recognize capped RNAs to inhibit viral replication (Habjan et al., 2013; Kumar et al., 2014). Our analysis supports the notion that both proteins utilize a similar mode of cap-recognition, and thus should display overlapping antiviral activities. However, recent evidence suggested otherwise (Daugherty et al., 2016). Based on our structural and functional data, we propose that both IFIT1 and IFIT1B can target Cap0-containing viruses, but they may differ in their sensitivity to cap-proximal modifications such as methylation, RNA-sequence, or RNA-structure. These differences and underlying molecular mechanisms are not entirely clear yet, but one possibility is that the ability to overcome N1 methylation in an RNA-dependent manner could distinguish IFIT1

from IFIT1B proteins, and may explain why IFIT1 overexpression inhibited wild-type VSV replication (a Cap1-containing ssRNA virus), whereas IFIT1B overexpression did not (Daugherty et al., 2016). Regardless, in humans IFIT1B appears to be non-functional, and has been deleted or pseudogenized in several other mammals, consistent with the notion that widespread viral evasion strategies (e.g. N1 methylation) have generally defeated IFIT1B (Daugherty et al., 2016), whereas IFIT1 was retained possibly due to its adaptable nature.

Taken together, through a relatively non-specific cap binding pocket, and a potentially plastic RNA binding mechanism, IFIT1 appears to have grafted adaptability onto an otherwise germ-line encoded member of the innate immune system, to broadly defend against rapidly evolving viral pathogens. At the same time, the host evolved multiple mechanisms which combine to limit detrimental IFIT1 activity against endogenous mRNAs. Clearly, further work is needed to validate the physiological relevance of these notions, particularly with respect to understanding how IFIT1 can overcome N1 methylation in an RNA-dependent manner, the differential specificities of IFIT1 and IFIT1B proteins, and if viruses exploit CMTr2 and other enzymes to modify their mRNA and evade IFIT1. Finally, it has been established that human IFIT1 can form complexes with other IFIT family members (IFIT2 and IFIT3) and several host factors (Pichlmair et al., 2011), which could play a role in modulating self- vs non-self-mRNA recognition and translational inhibition. Our structural and functional analysis of IFIT1 capped RNA interactions will provide a framework for future structure-guided studies of IFIT function. Moreover, these efforts will provide important contributions to the development of mRNA therapeutics and to vaccine design, as emerging research suggests that rendering viruses more susceptible to IFIT1-like antiviral responses, by inhibiting their mRNA 2'-O methylation or modifying their 5'-secondary structure, is a strategy for the rapid development of live, attenuated vaccines (e.g. refs. (Hyde et al., 2014; Menachery et al., 2014; Reynaud et al., 2015; Zust et al., 2013)).

3.5 Materials and Methods

3.5.1 Cloning, protein expression, and purification of IFIT proteins.

IFIT1 (UniProt accession number P09914) and IFIT5 (UniProt accession number Q13325) were cloned into a pSMT3 vector (pET-28a backbone) between BamHI and NotI sites, resulting in fusion proteins encoding an N-terminal, Ulp1-cleavable 6xHis-Sumo tag (Mossessova and Lima, 2000). Point mutants were generated by site-directed mutagenesis using iProof High-Fidelity DNA

polymerase (Bio-Rad). All IFITs were expressed in BL21 (DE3) or Rosetta 2 (DE3) pLysS cells using standard protocols, and purified by Ni-affinity chromatography followed by cleavage of the tag. Cleaved proteins carry over an N-terminal serine residue encoded by the BamHI site. For IFIT1 and mutants, cleaved proteins were further purified by passing over a 5 ml HiTrap SP HP column (GE Healthcare) equilibrated with 50 mM HEPES pH 7.5 and 1 mM DTT, and eluted with a salt gradient between 100 and 400 mM NaCl. For IFIT5, a second Ni column was performed followed by a Mono Q 4.6/100 PE (GE Healthcare) column in 25 mM Tris pH 8.0 and 1 mM DTT, and eluted over a shallow salt gradient between 100 and 200 mM NaCl. All proteins were further purified by gel filtration using superdex 200 10/300 or 16/60 columns (GE Healthcare) in 20 mM Tris pH 7.6, 150 mM NaCl and 1 mM DTT (gel filtration buffer). All pure proteins had a final $A_{260}/_{280}$ ratio ~ 0.5 , indicating no RNA contamination.

3.5.2 Cloning, protein expression, and purification of RNA modifying enzymes.

The Human *N7*-methyltransferase (RNMT, a gift from Dr. Stewart Shuman, Sloan Kettering Institute) residues 2-476 was cloned into pSMT3 between BamHI and NotI, expressed in Rosetta 2 (DE3) pLysS and the purification was adapted from ref. (Thillier et al., 2012). RNMT was purified by Ni-affinity chromatography followed by cleavage of the tag as above, then passed over a 5 ml HiTrap SP HP column (GE Healthcare) equilibrated with 20 mM Tris pH 7, 10 % glycerol and 5 mM β -ME, and eluted with a salt gradient between 50 and 500 mM NaCl. RNMT was further purified on a superdex 200 16/60 column in 20 mM Tris pH 7.6, 300 mM NaCl, 10 % glycerol and 5 mM β -ME. RNMT protein containing fractions were dialyzed overnight against 20 mM Tris pH 7.6, 300 mM NaCl, 50 % glycerol, and 5 mM β -ME, and stored at $-20\text{ }^{\circ}\text{C}$ at 0.8 mg/ml. A synthetic gene encoding *Trypanosoma brucei brucei* Cap2 methyltransferase (TbMTr2, also called TbCom1, reference gene name Tb11.02.2500, UniProt Q385S9) was ordered as an *E. coli* codon optimized gBlock Gene Fragment (IDT) and cloned into pProEX HTb between BamHI and NotI, resulting in a fusion protein with a TEV protease cleavable 6xHis N-terminal tag. TbMTr2 was expressed in BL21 (DE3) and purified by two-step Ni affinity chromatography using standard protocols. The second Ni flow-through was applied to a 5 ml HiTrap Q HP column (GE Healthcare) equilibrated with 25 mM Tris pH 8.0 and 1 mM DTT, and the column washed with the same buffer containing 50 mM NaCl. Cleaved TbMTr2 was present in the Q column flow-through and low salt wash, and was applied to a 5 ml HiTrap SP HP column equilibrated with 50

mM HEPES pH 7.5, 1 mM DTT, and eluted with a salt gradient between 50 and 500 mM NaCl. Protein containing fractions were concentrated and injected onto a superdex 200 10/300 column equilibrated with gel filtration buffer (see above). Fractions corresponding to TbMTr2 dimers were pooled and an aliquot stored at 3.84 mg/ml (80 μ M) at -20°C in 50 mM Tris pH 7.4, 5 mM DTT, 50% glycerol and 25 mM NaCl. Human Cap2 methyltransferase (CMTr2/FTSJD1, UniProt Q8IYT2) was ordered from DNASU and cloned into pFastBac HTb between BamHI and NotI. Recombinant baculovirus was generated according to the Bac-to-Bac Baculovirus Expression System protocol (Invitrogen) using DH10MultiBac cells (Berger et al., 2004). Virus was added to *Sf9* cells grown in I-Max medium (Wisent Bioproducts), and CMTr2 was expressed at 27°C for 64-72 h. Cells were pelleted at 1000x g at 4°C for 10 mins, and pellets were resuspended in 20 mM Tris pH 8, 10 mM imidazole, 500 mM NaCl, 10 % glycerol and 5 mM β -ME (Ni A Buffer), supplemented with cOmplete EDTA-free Protease Inhibitor Cocktail (Roche) and 1 mM PMSF. Cells were lysed by sonication on ice (10 sec ON, 20 sec OFF, 50% amplitude, 6 total pulses), and clarified by ultracentrifugation at 41,000 rpm using a Type 70 Ti rotor. Supernatants were applied to Ni resin, washed 3 times with 10 bed volumes Ni A, and eluted with 4 bed volumes Ni A with 350 mM imidazole. Eluted protein was cleaved overnight with TEV protease in Ni A with 2 mM DTT instead of β -ME. A second Ni bead purification was performed to remove TEV protease, and the flow-through was concentrated and injected onto superdex 200 10/300 equilibrated in 20 mM Tris pH 7.6, 150 mM NaCl and 3 mM DTT. Fractions corresponding to monomeric protein were pooled, concentrated to 4 mg/ml, and stored at 4°C .

3.5.3 Crystallization and data collection.

The monomeric mutants of IFIT1 (L457E/L464E) or IFIT1 (N216A/L457E/L464E) were mixed with molar excess chemically synthesized oligos (see below for RNA preparation) either directly before crystallization or prior to gel filtration. The protein buffer was 20 mM Tris pH 7.6, 150 mM NaCl and 1 mM TCEP. The co-crystals were obtained with 5-10 mg/ml protein drops mixed 1:1 with reservoir solution containing 27-32 % PEG 200 (Sigma), 0.1 M Tris pH 8.1, and 200 mM CaCl_2 using the hanging drop vapour diffusion method at 4°C . Crystals were flash frozen in liquid nitrogen without additional cryo-protection. All data were collected at 100K with 0.979 \AA X-rays on beamline 08ID-1 at the CLS, which is equipped with a Mar300CCD detector (Grochulski et al., 2011), and the images integrated, scaled, and merged using the HKL2000 suite (43). Data were

truncated with Ctruncate (ccp4 suite (Winn et al., 2011)) and 5 % of reflections were set aside for the R_{free} set. Only the highest resolution dataset was used as a master file to generate all R_{free} sets. Data collection statistics from HKL2000 are in Table 3.1.

3.5.4 Structure determination, model building, and refinement.

The initial crystal structure was determined by molecular replacement using the PHASER program in Phenix (McCoy et al., 2007). The structure of wild-type human IFIT1 was used as a search model. Subsequent structures were determined with rigid body refinement with the protein only. Restraints for m7GpppA and GpppA, as well as bound PEG molecules, were calculated using Phenix eLBOW (Moriarty et al., 2009). The structure was refined iteratively using Phenix (Adams et al., 2010) with manual model building in Coot (Emsley et al., 2010). The refinement strategy included all-isotropic B-factor refinement, occupancy refinement, and TLS. All final models contained protein residues 8-467. Structure validation was performed with MOLPROBITY in Phenix (Chen et al., 2010). Ramachandran statistics are as follows: IFIT1 L457E/L464E with m7Gppp-RNA, 98.5% favored, 0% outliers, 0.2% rotamer outliers; IFIT1 L457E/L464E with Gppp-RNA: 98.9% favored, 0% outliers, 0% rotamer outliers; IFIT1 L457E/L464E with PPP-RNA: 98.9% favored, 0.2% outliers, 0.2% rotamer outliers; IFIT1 N216A/L457E/L464E with m7Gppp-RNA: 98.7% favored, 0% outliers, 0.2% rotamer outliers. The MOLPROBITY overall score for each was 1.00, 0.99, 0.93, and 0.92 respectively. Refinement statistics are in Table 3.1.

3.5.5 Sequence, structure and RNA analysis.

APBS was used to calculate the surface electrostatic potential (Baker et al., 2001), and PyMol to generate all molecular figures (<https://www.pymol.org/>). For surface electrostatic analysis, all surfaces are colored by electrostatic potential from negative (-10 kTe^{-1} ; red) to positive ($+10 \text{ kTe}^{-1}$; blue). CAVER (Pavelka et al., 2016) was used to draw the protein tunnel in Fig. 3.3. ESPript was used to generate the sequence alignment in Fig. 3.10 (Robert and Gouet, 2014). RNA secondary structure prediction was performed with Mfold (Zuker, 2003). *In silico* N1/N2 methylation for Fig. 3.14 was performed by rigid body docking. For sequence conservation analysis, all mammalian IFIT1-like and IFIT5-like annotated sequences were retrieved from the non-redundant protein sequences database using human IFIT1 as a BLAST query. The results were manually curated to remove highly related isoforms, and to remove sequences with large N- and C-terminal deletions (for e.g. those missing R38/T37). The final data set contained 86 IFIT1- or

IFIT1B-like sequences, and 60 IFIT5-like genes. Sequences were aligned with ClustalOmega, and edited in Jalview (<http://www.jalview.org/>) to remove columns arising from insertions in non-human IFIT1 or IFIT5 genes. This was done to maintain human IFIT1 or human IFIT5 residue numbering. The aligned and edited sequences were used to generate the sequence logos with WebLogo (Schneider and Stephens, 2004), and only blocks related to m7Gppp- binding residues are shown.

3.5.6 Enzymatic preparation and purification of RNA for EMSAs.

All PPP-RNAs used for EMSAs were prepared enzymatically by run-off transcription using T7 RNA polymerase. RNA sequences are listed in Table 3.2. Single stranded oligo templates (also in Table 3.2) were mixed 1:1 with the appropriate promoter sequence at a final concentration of 10 μ M (each), heated to 95 degrees for 2-3 mins, and slow cooled on the bench. Transcription was carried out at 37 °C with 1 μ M template (bottom strand) in 1X buffer (NEB) supplemented with 10 mM DTT, 2 mM ATP, 2 mM UTP, 2 mM CTP, 2 mM GTP, 2 U/ml pyrophosphatase (*E. coli*, NEB), RNasin (Promega), and 2.5 U/ μ L T7 RNA polymerase (NEB), in 0.5-1 ml final volume. MgCl₂ concentrations between 5 and 40 mM were used and optimized for each template. Appcapped GGG42 was prepared by replacing the GTP with 8:1 A(5')ppp(5')G:GTP. All RNAs were purified by 10 or 12% denaturing PAGE. The ss44, MHV, HCoV, and GGG42 transcribed RNA were validated by MS to confirm the purity and sequence of the final product, all of which contained minor amounts of n+1 contamination carried over from the transcription and purification, and to confirm the presence of an intact triphosphate or Appp- cap (details below). 3'-end labeling of m7Gppp-43 with pCp-Cy5 (Jena Biosciences) was performed with T4 RNA ligase (NEB), and the labeled product purified by gel extraction on 8% denaturing PAGE. OH-RNA was prepared by treating the PPP-RNA with CIP (NEB). m7Gppp- and Gppp-RNA were prepared by post-transcriptional capping of PPP-RNA with the vaccinia capping system (NEB) with or without S-adenosylmethionine (SAM). Cap1-RNA was prepared by including the Cap1 2'-O-Methyltransferase (NEB) in the capping reactions. NEB reactions were performed according to the manufacturer's instructions. Ribose N2 methylations were performed with purified CMTr2 or TbMTr2 as indicated in Table 3.3. For N2 methylation with TbMTr2, the protocol was adapted from ref (Hall and Ho, 2006). 200 pmol of Cap0- or Cap1-RNAs were incubated with 800 pmol TbMTr2 in 200 μ L reactions containing 50 mM tris pH 7.5, 5 mM DTT, 2 mM SAM and RNasin. Reactions were carried out for 2 hours at 27 °C. For CMTr2 N2 methylation, the protocol was

adapted from ref (Werner et al., 2011). 150-200 pmol of Cap0- or Cap1-RNAs were incubated with 500 pmol CMTr2 in 100 μ L reactions containing 30 mM Tris pH 7.4, 50 mM KCl, 10 mM EDTA, 2.5 mM SAM and RNasin. Reactions were carried out for 2 hours at 37 °C. Modified RNAs were purified by phenol/chloroform extraction and ethanol precipitation. The purity of RNA used in this study was also confirmed by 17% denaturing PAGE (Fig. 3.15 G). All short RNA used for Cap0/Cap1/Cap02/Cap2 binding analysis were validated using LC-MS to confirm > 90-95% homogeneity of 5' modification (details below).

3.5.7 RNA electromobility gel shift assay (EMSA).

10X stocks of purified protein were prepared in 25 mM Tris pH 7.6, 150 mM NaCl, 50 % glycerol, 5 mM β -ME, 0.5 mM PMSF and 1 mM EDTA, and stored at -20 °C. RNAs were diluted in MilliQ water at a final concentration of 1 μ M and heat refolded prior to use. For capped-RNA, the RNA was heated at 95 °C for 2-3 mins and slow cooled on the bench. For PPP-RNA, heating was at 65 °C for 10 mins with slow cooling on the bench. Protein and RNA were mixed with 10X buffer (500 mM Tris pH 7.4, 1.25 M NaCl, 10 mM EDTA, and 0.1 mg/ml Heparin) at a final volume of 20 μ L and the binding proceeded on the bench for at least 15 mins. For Fig. 3.15 E, only 12 μ L was loaded in each well. For EMSAs with pCp-Cy5 labeled RNA, the final volume was 10 μ L. The final reactions all contained 5 % glycerol (carried over from the protein storage buffer) and no loading dyes were used. The final RNA and protein concentrations are indicated in each figure. Bound and unbound RNA were resolved on 1X TBE, 10% 19:1 acrylamide:bisacrylamide native gels supplemented with 25 mM NaCl. The gels were pre-run in a cold room in 0.5X TBE running buffer for at least 30 minutes, loaded with samples, and run at 120-180 V for approx. 45-60 minutes. The temperature of the system remained 8-14 °C during the run. For EMSAs with fluorescent RNA, the gel apparatus was covered with aluminum foil to minimize light exposure. Gels were stained with SYBR Gold (Invitrogen) and visualized with a UV trans-illuminator equipped with a SYBR Green filter. For Fig. 3.9 A, 3.8 L, and 3.14 D, pCp-Cy5 labeled RNA was visualized with a Typhoon R3 imager (GE Healthcare). Band densitometry was performed with ImageJ (<https://imagej.nih.gov/ij/>).

3.5.8 Preparation of reporter mRNA.

The bicistronic reporter, pSP-(CAGless)/FF/HCV/Ren was derived from pSP(CAG)33/FF/HCV/Ren·pA51 by exchanging the EcoRI/SacI fragment with an

oligonucleotide (5'GAATTCACAATTCGAGCTC3', restriction sites are underlined) (Robert et al., 2006). The 5' UTR of (CAGless)-FF/HCV/Ren is described in Table 3.2. For *in vitro* transcription, pSP-(CAGless)/FF/HCV/Ren was linearized with *Bam*HI. *In vitro* transcriptions were performed with SP6 RNA polymerase in the presence of m7G(5')ppp(5')G, G(5')ppp(5')G or A(5')ppp(5')G RNA Cap Analog (NEB). N1 and N2 methylations were performed post transcriptionally using mRNA Cap1 2'-O-Methyltransferase (NEB) or purified TbMTr2. Reporter mRNA was heated at 65 °C for 5 min, then placed on ice for 5 min before reactions were set up. To generate Cap1 reporter, N1 methylation was performed according to NEB protocols, except the reaction was performed at 30 °C for 2 hours. N2 methylation of Cap0- and Cap1-reporter were performed as described above for the short RNA. Modified mRNA was purified by phenol/chloroform extraction, ethanol precipitation, and desalting columns.

3.5.9 *In vitro* translation assay with Krebs extracts.

Preparation of the mouse-derived Krebs-2 extracts for cell-free translation is described in ref. (56). *In vitro* translations were set up with a final volume of 10 µl with 4 ng/µL reporter mRNA (~ 4 nM final) and 1 µL of purified protein (in gel filtration buffer) in untreated Krebs-2 extracts as described previously (Novac et al., 2004). The reactions were set up on ice and translation was allowed to proceed for 1 hr at 30 °C. The reactions were stopped on ice, and FF and Ren luciferase activities (RLU) were measured on a Berthold Lumat LB 9507 luminometer. Values obtained were normalized against buffer control, which was set at 1. In Fig. 3.1 C, 3.2 D and 3.8 E, the following final protein concentrations were used: 0.031, 0.062, 0.125, 0.250, 0.5, 1, and 5 µM. For Fig. 3.9 B, 3.14 C and 3.14 E, 0.1, 0.3, 0.6, 1, and 5 µM final protein concentration was used. Except for Fig. 5B and 8E, all data are represented by the mean of 2-3 independent measurements performed in duplicate (as indicated in each figure legend) ± standard deviation.

3.5.10 Mammalian constructs, cells, reagents and viruses.

Mammalian IFIT1 expression constructs were generated by targeted mutagenesis PCR and cloned into pTO-SII-HA-GW (Pichlmair et al., 2011) using gateway cloning (Invitrogen). Sequences were verified by Sanger sequencing. Huh7 cells were a gift from Georg Kochs (University of Freiburg), Flp-In T-REx 293 IFIT1^{-/-} cells were a gift from Giulio Superti-Furga (CeMM, Vienna). Cells were maintained in DMEM (PAA Laboratories) containing 10% fetal calf serum (GE Healthcare) and antibiotics (100 U/ml penicillin, 100 µg/ml streptomycin). Antibodies used to test

expression levels were actin-HRP (Santa Cruz: sc-47778), and HA-HRP (Sigma: H6533). Interferon- β was a gift from Peter Stäheli (University of Freiburg). Wild-type and 2'-O-methyltransferase-deficient recombinant coronaviruses 229E (HCoV-229E) have been described previously (Habjan et al., 2013).

3.5.11 Virus infections and determination of virus titers.

To determine the impact of IFIT1 mutants on virus growth, Flp-In T-REx 293 IFIT1^{-/-} cells were co-transfected with plasmids for IFIT1 and CD13 (HCoV receptor). Cells were treated with 20 units of Interferon- β or left untreated and 24 h later infected with HCoV 229E wt and HCoV 229E DA with a multiplicity of infection of 1 and 1.25, respectively. 18 h post infection, cells were harvested for qRT-PCR analysis and supernatant was harvested to test for virus accumulation by TCID50 on Huh7 cells.

3.5.12 Real-time RT-PCR.

Total RNA was isolated using the NucleoSpin RNA Plus kit (Machery-Nagel). 200-500 ng of RNA was reverse transcribed with PrimeScript RT Master Mix (TAKARA) and quantified by real-time PCR using the QuantiFast SYBR Green RT-PCR Kit (Qiagen) and a CFX96 Touch Real-Time PCR Detection System (BioRad). Each cycle consisted of 10 sec at 95 °C and 30 sec at 60 °C, followed by melting curve analysis. Primer sequences are in Table 3.2.

3.5.13 Chemical synthesis of PPP-AAAA.

Phosphoramidite solid phase synthesis conditions were used for the synthesis of RNA in a twist oligonucleotide synthesis column (Glen Research) and adapted from refs (Zlatev et al., 2010; 2012). The 5'pppAAAA3' strand was synthesized at the 1 μ mol scale, using a chain alkylamine controlled-pore glass as a solid support derivatized with *N*-phenoxyacetyl adenosine (ChemGenes). (5'-ODMTr-2'-TBDMS-A_{Pac})-3'-*O*-(2-cyanoethyl-N,N-diisopropyl)phosphoramidite was prepared as a 0.15 M solution in acetonitrile. 5-Ethylthiotetrazole was used as an activator, and 3% trichloroacetic acid in dichloromethane was used to detritylate. Capping was carried out using phenoxyacetic anhydride in THF and 16% *N*-methylimidazole in THF. 0.1 M I₂ in 1:2:10 pyridine/water/THF was used for oxidation. The RNA phosphoramidite was coupled for 600 s. Installation of the triphosphate moiety (or cap structure)

at the 5'-end of the RNA was performed in the same twist columns in a fully automated procedure integrated within the synthesis program of the ABI-394 (Zlatev et al., 2010; 2012).

For 5'-H-phosphitylation of solid-supported RNA, a solution of 3 mL diphenyl phosphite in 12 mL dry pyridine was prepared and applied to the column three times (delivery time 5 sec, wait time 300 s each). The support was washed with ACN for 3 min. A 0.1 M solution of triethylammonium buffer was delivered six times to the column (delivery time 5 s, wait times $60 \times 2 + 450 \times 4$ s).

To prepare solid-supported 5'-phosphoroimidazolide RNA, the oxidation solution was prepared as follows. To a solution of imidazole (900 mg) in *N,O*-bistrimethylsilylacetamide (3 mL), ACN (6 mL), BrCl_3C (6 mL), and triethylamine (1.2 mL) were added. Activated 3 Å molecular sieves were added and the solution was stored for 24 h at 4 °C prior to use. The mixture was delivered to the column 5 times (delivery time 5 sec, waiting time 1800 s each). The solid support was then washed with ACN for 3 min.

For substitution of imidazole by pyrophosphate moiety, tributylammonium pyrophosphate salt required for phosphorylation reactions was first prepared from tetrasodium pyrophosphate. Tetrasodium pyrophosphate decahydrate (2.3 g) was dissolved in 20 mL of sterile nuclease-free water and eluted through a glass column filled with 40 mL of wet DOWEX-50WX8 resin, H⁺ form. The eluted product was collected in a flask containing 40 mL of ethanol and 2.5 mL of tributylamine while stirring at 0 °C. The resin was washed with additional 20 mL of water. Solvents were evaporated and the residue coevaporated four times with dry dioxane to yield 2.7 g of tributylammonium pyrophosphate that was dissolved in a mixture of 3.3 mL dry DMF and 3.3 mL of anhydrous ACN. This solution was delivered five times to the column (delivery time 3 s, wait time $300 \times 4 + 18000$ s). The solid support was then washed with ACN and dried under argon.

For deprotection and release from the support, the twist column was removed from the synthesizer and the CPG was transferred to a screw cap eppendorf. Deprotection and cleavage from the solid support was achieved with 3:1 aqueous $\text{NH}_4\text{OH}/\text{EtOH}$ for 3 h at room temperature. After decanting to remove the CPG, the deprotection solution was removed under vacuum in a speedvac lyophilizer. Desilylation was achieved by adding 150 μL of 0.1 M solution of TBAF in THF and shaking at room temperature for 24 h. After addition of 850 μL of autoclaved water, the crude sample was desalted using NAP-25 sephadex size exclusion columns according to manufacturer protocol. Purification of PPP-AAAA is described below.

3.5.14 Chemical synthesis of Gppp-AAAA.

The synthesis of Gppp-AAAA was performed following a similar procedure to PPP-AAAA except that the substitution reaction on the solid-supported 5'-phosphoroimidazolide was carried out by reaction with guanosine-5-diphosphate instead of pyrophosphate (Thillier et al., 2012).

For the capping reaction with GDP, the bis(tri-*n*-butylammonium) salt of guanosine-5-diphosphate (GDP) was prepared from the commercially available sodium salt purchased from Sigma-Aldrich. GDP sodium salt (0.5 g) was dissolved in 25 mL of sterile nuclease-free water and eluted through a glass column filled with 20 mL of wet DOWEX-50WX8 resin, H⁺ form. The eluted product was collected in a flask containing 40 mL of ethanol and 2.5 mL of tributylamine while stirring at 0 °C. The resin was rinsed with addition of 20 mL of water. The solvents were evaporated and the residue was coevaporated two times with absolute ethanol and three times with dry DMF to yield a hygroscopic powder (0.90 g). 103 mg of this solid were transferred to a glass vial and dissolved in 0.5 mL of dry DMF. 28 mg of ZnCl₂ were also added to the mixture. The resulting solution was applied to the column containing the solid-supported 5'-phosphoroimidazolide RNA and was allowed to react for 24 h at room temperature. The solution was removed, and the support was washed with MeCN and DMF. The column was dried for 1 min by blowing argon. The deprotection and purification steps were similar to those described above for PPP-AAAA. Note that this procedure differs from the one described in ref. (Thillier et al., 2012). The work by Debart et col. reports the synthesis of capRNA using 2'-PivOM protected amidites meanwhile herein we use the classical 2'-*O*-TBDMS chemistry instead of only TBDMS. Similar results were obtained. Purification of Gppp-AAAA is described below.

3.5.15 Purification of chemically synthesized PPP-AAAA and Gppp-AAAA.

Purification was performed on an HPLC using a Protein Pak DEA 5PW analytical anion exchange HPLC column. A stationary phase of Milli-Q water and a mobile phase of 1 M LiClO₄ in water was used for analysis and purification using a gradient of 0-50 % over 46 min. Following purification, excess LiClO₄ salts were removed using NAP-25 sephadex size exclusion columns. The oligos were characterized by LC-MS (details below).

3.5.16 Preparation of m7Gppp-AAAA.

m7Gppp-AAAA was prepared by enzymatically modifying Gppp-AAAA according to ref. (Thillier et al., 2012). Reaction conditions contained 40 mM Tris pH 8.0, 5 mM DTT, and 4 mM

SAM. For every 3.3 nmol of Gppp-AAAA, 20 pmol RNMT in an 80 μ L reaction was utilized. The reactions were scaled up according to crystallization needs. The modified oligo was mixed with purified protein and the complex cleaned up by gel filtration prior to crystallization. LC-MS was used to verify that the modification was quantitative (see below).

3.5.17 Characterization of RNA by LC-MS.

RNA were analyzed by Liquid Chromatography Mass Spectrometry using a Dionex Ultimate 3000 coupled to a Bruker Maxis Impact QTOF in negative ESI mode. Samples were run through an Acclaim RSLC 120 C18 column (2.2 μ M 120 \AA 2.1 x 50 mm) using a gradient of 98% mobile phase A (100 mM HFIP and 5 mM TEA in H₂O) and 2% mobile phase B (MeOH) to 40% mobile phase A and 60% mobile phase B in 8 minutes. The data was processed and deconvoluted using the Bruker DataAnalysis software version 4.1. MS data are summarized in Table 3.3.

3.6 Tables

Table 3.1 Data collection and refinement statistics IFIT1 with capped and uncapped RNAs

Property	Value			
	IFIT1 (L457E/L464E) m7Gppp-AAAA	IFIT1 (L457E/L464E) Gppp-AAAA	IFIT1 (L457E/L464E) PPP-AAAA	IFIT1 (N216A/L457E/L464E) m7Gppp-AAAA
Data collection				
PDB ID code	5UDI	5UDJ	5UDK	5UDL
Space group	P4 ₂ 22	P4 ₂ 22	P4 ₂ 22	P4 ₂ 22
Mol. per ASU	1	1	1	1
Cell dimensions				
a, b, c, Å	111.7, 111.7, 93.2	112.5, 112.5, 93.5	112.4, 112.4, 93.2	111.8, 111.8, 93.0
α, β, γ, °	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution, Å	50-1.58 (1.61-1.58)	50-1.69 (1.72-1.69)	50-1.65 (1.68-1.65)	50 - 1.65 (1.68-1.65)
Completeness, %	100 (100)	100 (100)	99.7 (95.2)	100 (99.9)
R _{pim} , %	2.4 (43.8)	4.1 (52.6)	2.4 (36.8)	2.8 (39.6)
<I/σ(I)>	33.6 (1.74)	22.2 (1.9)	30.3 (1.96)	27.4 (1.78)
Redundancy	14.5 (12.4)	14.5 (12.2)	14.2 (9.4)	14.6 (12.3)
Refinement				
Resolution, Å	40.1-1.58 (1.6-1.58)	40.3-1.69 (1.71-1.69)	44.3-1.65 (1.67-1.65)	40.1-1.64 (1.67-1.64)
No. reflections	81063 (7992)	67665 (6640)	71917 (6888)	71805 (6876)
R _{work} /R _{free}	15.7/17.8 (25.0/27.6)	15.9/18.1 (24.8/30.3)	16.4/18.4 (24.2/26.8)	16.4/18.2 (24.3/24.6)
No. of non-H atoms				
Protein	3842	3852	3833	3816
Ligand - RNA	168	116	97	117
Ligand - PEG	10	10	41	10
Ligand - ion	2	2	2	2
Solvent	394	382	335	313
Avg. B factors, Å²				
All atoms	30.65	27.64	30.61	30.86
Protein	30.7	27.6	30.4	30.9
Ligand - RNA	18.8	17.0	19.3	22.2
Ligand - PEG	49.2	48.5	57.4	62.6
Ligand - ion	32.1	30.8	26.6	39.5
Solvent	34.8	30.8	32.6	32.6
R.m.s. deviations				
Bond lengths, Å	0.017	0.014	0.012	0.008
Bond angles, °	1.579	1.283	1.166	1.005

*Highest resolution shell in parenthesis.

Table 3.2 Sequences of *in vitro* transcribed RNA, DNA oligos templates for *in vitro* transcription, and primers for quantitative real-time PCR.

Name	Sequence
MHV RNA	GUAUAAGAGUGAUUGGCGUCCGUACGUACCCUCUCAACUCU
HCoV RNA	ACUUAAGUACCUUAUCUAUCUACAGAUAGAAAAGUUGCUIUUU
GGG42 RNA	GGGUAAGUACCUUAUCUAUCUACAGAUAGAAAAGUUGCUIUUU
ss44 RNA	GGGAGAGAGAGAGAGAGAGUAAGGGCGUCGUCGCCCCGAGAAUU
43 RNA*	GGGUAAGUACCUUAUCUAUCUACAGAUAGAAAAGUUGCUIAAAA
Reporter 5'-UTR [†]	GAAUACACGGAAUUCAGCAAUUCGAGCUCGCCCCGGGAUCUGCGAUCUAAGUAAGCU UGGCAUUCGGUACUGUUGGUAAGCCACCAUG
T7 promoter top [‡]	TAATACGACTCACTATAG
T7Φ2.5 promoter	TAATACGACTCACTATTA
MHV template	AGAGTTGAGAGGGTACGTACGGACGCCAATCACTCTTATACTATAGTGAGTCGTATTA
HCoV template	AAAAGCAACTTTTCTATCTGTAGATAGATAAGGTACTTAAGTAATAGTGAGTCGTATTA
GGG42 template	AAAAGCAACTTTTCTATCTGTAGATAGATAAGGTACTTACCCTATAGTGAGTCGTATTA
ss44 template	AATTCTCGGGGCGACGACGCCCTTACTCTCTCTCTCTCCCTATAGTGAGTCGTATTA
43 template	TTTTAGCAACTTTTCTATCTGTAGATAGATAAGGTACTTACCCTATAGTGAGTCGTATTA
Human GAPDH F	GATCCACCCATGGCAAATTC
Human GAPDH R	AGCATCGCCCCACTTGATT
HCoV 229E N F	CAGTCAAATGGGCTGATGCA
HCoV 229E N R	AAAGGGCTATAAAGAGAATAAGGTATTCT

This RNA is similar to GGG42, except that it is designed to have a 5 nt overhang to allow strong IFIT1 binding, and a free 3'-end with sequence AAA to improve pCp-Cy5 labeling efficiency with T4 RNA ligase.

[†]Start codon underlined, this transcript was prepared with a plasmid and not a DNA oligo.

[‡]Promoter top strand for all sequences except HCoV and mRNA reporter.

[§]Promoter top strand for HCoV.

Table 3.3 Summary of Mass Spec data

Sequence	Expected (Da)	Detected (Da)
PPP-AAAA	1516.14 (Na adduct)	1516.12 (Na adduct)
Gppp-AAAA	1759.23	1759.18
m7Gppp-AAAA	1773.25	1773.19
PPP-MHV	13300.71 (Na Adduct)	13300.05 (Na adduct)
m7Gppp-MHV	13557.99	13556.99
Cap1-MHV	13572.02	13570.97
N2Me-MHV	13572.02	13571.42
Cap2-MHV	13586.04	13585.45
Cap2-MHV (TbMTr2 modified)	13586.04	13584.69
PPP-HCoV	13569.86 (Na Adduct)	13569.42 (Na adduct)
m7Gppp-HCoV	13827.14	13825.91
Cap1-HCoV	13841.16	13840.38
N2Me-HCoV	13841.16	13840.61
Cap2-HCoV (TbMTr2 modified)	13855.08	13854.48
PPP-GGG42	13664.93 (Na Adduct)	13664.47 (Na adduct)
Gppp-GGG42	13908.18	13908.09
Appp-GGG42	13892.18	13892.22
m7Gppp-GGG42	13922.21	13921.55
Cap1-GGG42	13936.23	13935.91
N2Me-GGG42	13936.23	13934.95
Cap2-GGG42 (TbMTr2 modified)	13950.25	13949.29
PPP-ss44	14705.73 (Na Adduct)	14704.64 (Na adduct)

All values are reported as molecular weight, except for PPP-AAAA, Gppp-AAAA, and m7Gppp-AAAA, which are reported as exact monoisotopic neutral mass. For all RNA except the short oligoadenosines, only the most intense peak is reported, as they all contained Na or K adducts. All PPP-RNA contain variable amounts of PP- and P-RNA, likely due to hydrolysis of the PPP under the alkaline conditions of LC-MS. For all m7Gppp-, Cap1-, N2Me-, and Cap2-RNA, expected values are calculated based on the zwitterionic enolate form of m7G (in which N1 is deprotonated, See Fig. 3.8 H). Unless otherwise specified, N2 methylations to generate N2Me- or Cap2-RNA were performed with the human CMTr2 protein.

3.7 Acknowledgements

We thank the CMCF staff for X-ray data collection (Canadian Light Source); Dr. A. Wahba for LC-MS (Dept. of Chemistry, McGill University); A. Gorelik, Dr. N. Siddiqui, and Dr. M.W. Görna for feedback and reading of the manuscript; A. Gorelik and K. Illes for *Sf9* expression; P. Sénéchal for assistance with translation assays. Support was provided by the CIHR Strategic Training Initiative in Chemical Biology and the NSERC CREATE Training Program in Bionanomachines (Y.M.A.); a Canada Research Chair (B.N.); Canadian Institutes of Health Research Grants (MOP-133535, B.N. and MOP-115126, J.P.); Discovery Grant from the Natural Sciences and

Engineering Research Council of Canada (M.J.D.); McGill CIHR Drug Development Training Program (S.M-M.); the Max-Planck Free Floater program, the ERC (StG iVIP, 311339), InfectERA (ERASE) and the German Federal ministry of Education and Research (A.P.).

CHAPTER 4: GENERAL DISCUSSION

4.1 Self vs non-self mRNA discernment by IFIT1

The regulation of mRNA translation is a critical nexus in the host-virus arms race. This is evident from the numerous host effector molecules that target mRNA translation as part of their antiviral program (Li et al., 2015), and the diverse strategies viruses have evolved to shut-down host protein synthesis to favor viral mRNA translation (Walsh and Mohr, 2011). A well-studied host antiviral effector is protein kinase R, which upon infection will recognize viral PAMPs and inhibit cellular protein synthesis by phosphorylating eIF2 α (Hull and Bevilacqua, 2016). Human IFIT1-like proteins also inhibit translation initiation, but despite being identified and characterized in the 80s, their mode of viral inhibition remained enigmatic for almost 3 decades, and is still a subject of debate. Sen and colleagues originally discovered that human IFIT1 and mouse Ifit1 are inhibitors of translation, but they attributed this to an apparent ability of these proteins to interact with subunits of eIF3 and disrupt assembly of 43S ribosomal complexes (Guo et al., 2000a; Hui et al., 2003; 2005). Thus, they proposed that IFIT1 promoted an antiviral state in cells by decreasing general protein synthesis and reducing cell proliferation (Geiss et al., 2001; Guo et al., 2000a). However, several studies were unable to observe interactions between IFITs and eIF3 subunits (Pichlmair et al., 2011; Siegfried et al., 2013; Stawowczyk et al., 2011), suggesting that translational inhibition by IFIT1 and IFIT1B proteins was through an alternate mechanism.

Indeed, a more direct antiviral role for IFIT1 and IFIT1B was recently discovered. By interacting with the mRNA cap and cap-proximal regions, they can compete with eIF4F for mRNA binding and inhibit 48S ribosomal complex formation (Habjan et al., 2013; Kumar et al., 2014). Through sensing the mRNA cap-proximal methylation status, IFIT1 and IFIT1B can distinguish host methylated from viral unmethylated mRNA, and selectively target viral mRNA for translational inhibition (Daffis et al., 2010; Habjan et al., 2013; Kimura et al., 2013; Kumar et al., 2014; Züst et al., 2011). These studies therefore uncovered an elegant mechanism governing IFIT1 and IFIT1B activities, and rationalized the existence of cap-proximal methylations – another long-standing puzzle that remained unsolved for almost 35 years (Daffis et al., 2010; Furuichi et al., 1975). Despite these advances in our knowledge, a complete understanding of human IFIT1 antiviral activity has been hampered by the paucity of structural information, and a complicated evolutionary history which has convoluted the relationship between human IFIT1 and the related

mouse gene on which much of our *in vivo* knowledge was based upon (Daugherty et al., 2016). The crystal structure of IFIT1 with capped-RNA presented in this thesis, which is complemented with mutational and *in vitro* functional assays, validates that IFIT1 antiviral function is rooted in its ability to encircle the mRNA cap and cap-proximal regions. To better appreciate the role of IFIT1 at the interface of host-virus interactions, we must first consider the host mechanisms limiting self-recognition, and then examine how the virus exploits these mechanisms for its benefit.

4.1.1 mRNA cap recognition by human IFIT1

The structures of capped-RNA bound IFIT1 revealed a surprisingly adaptable and non-specific mode of mRNA cap recognition that is evolutionarily distinct from canonical cap binding proteins. While this non-specific cap binding mechanism could be due to the relatively recent emergence of IFIT proteins in evolution, it may have evolved as such to limit aberrant self-recognition of host mRNAs. The lack of sequence specific hydrogen bonds for the cap sacrifices affinity for this moiety, and thus IFIT1 compensates by forming a tight molecular interface between the protein and the cap-proximal region, to compete with eIF4F. In doing so, IFIT1 can more effectively distinguish between host and viral mRNAs based on cap-proximal methylation status. As IFIT1 is highly inducible, with protein levels on the order of 1-2 million copies per cell (Pichlmair et al., 2011), it must strike a delicate balance between high affinity for foreign mRNAs, and indiscriminate binding to all mRNAs – viral or otherwise. The decreased affinity for cap would prevent the latter scenario from dictating IFIT1 activity and thus prevent self mRNA recognition. Another possibility to consider is that the increased emphasis on cap-proximal nucleotides allows IFIT1 to retain binding towards PPP-RNAs, and thus engenders IFIT1 with a secondary antiviral mechanism targeting another marker of viral infection (discussed in section 4.2.2).

4.1.2 CMTr1, CMTr2, and other enzymes protecting endogenous mRNA from IFIT1

CMTr1 is a nuclear enzyme that appears to uniformly modify all host mRNAs with N1 ribose methylation before their export to the cytoplasm (Banerjee, 1980; Langberg and Moss, 1981; Perry and Kelley, 1976). The role of N1 methylation in preventing IFIT1 and IFIT1B binding *in vitro* has been extensively characterized by several groups (Habjan et al., 2013; Kimura et al., 2013; Kumar et al., 2014; Pinto et al., 2015). This is further confirmed by the functional analysis performed herein, and is in agreement with the structural analysis showing that N1 methylation would clash with two highly conserved residues, one of which is critical for RNA binding (R187).

As CMTr1 has yet to be implicated in general translational mechanisms (Bélanger et al., 2010), it may have evolved, in part, to protect host mRNAs from the antiviral innate immune system. Supporting this, knockdown of CMTr1 has been shown to promote an auto-immune RIG-I-dependent response (Schuberth-Wagner et al., 2015), another sensor of N1 methylation. However, the role of CMTr1 in protecting host mRNAs from IFIT1 *in vivo* requires further experimentation. One experiment that could be performed is CMTr1 knockdown in the presence of IFIT1, and investigating the effect on global or transcript specific mRNA translation. However, such an experiment must be carefully interpreted as the lack of N1 methylation could influence the activity of other mRNA modifying enzymes (e.g. CMTr2 and the m6Am MTase – discussed further below).

Another question to address is how does the host ensure all mRNAs are modified by CMTr1? CMTr1 has been shown to associate with the CTD of RNA pol II (Haline-Vaz et al., 2008), suggesting that it may modify mRNAs co-transcriptionally. Interestingly, enzymatic assays showed that CMTr1 is not as efficient as the Vaccinia virus MTase at catalyzing N1 methylation (Werner et al., 2011). Thus, its association with RNA pol II may also function to stimulate its activity. Finally, although constitutively expressed, CMTr1 is an interferon stimulated gene (Haline-Vaz et al., 2008), suggesting that it is upregulated during antiviral responses to further protect host mRNAs from aberrant self-recognition.

RNA binding and translation assays performed here, and functional assays performed by other groups, indicate that N1 methylation may not universally protect mRNAs from IFIT1 recognition (Daugherty et al., 2016; Young et al., 2016). Thus, N2 methylation by CMTr2 in the cytoplasm, which normally follows N1 methylation *in vivo*, could provide an additional safeguard to protect mRNAs that may otherwise be susceptible to IFIT1. The RNA determinants of this susceptibility to IFIT1 are not clear yet, and require future assays to systematically characterize the complete RNA binding spectrum of human IFIT1. Nevertheless, the discovery that N2 methylation alone can protect RNA from IFIT1 recognition, and the striking synergy observed between N1+N2 methylation in rescuing mRNA translation from IFIT1 inhibition, will hopefully prompt future analysis of CMTr2 structure and function. Of note, the RNA-sequence, RNA-structure, and transcript specificity of CMTr2 need to be investigated. Thus far, we know only that N2 methylation accompanies N1 on about half of cellular mRNAs in HeLa and Vero cells (Cleaves and Dubin, 1979; Wei and Moss, 1975), although lower numbers have been reported in other cell

types; prior N1 methylation enhances CMTr2 activity (Smietanski et al., 2014); and N2 methylation requires full-length CMTr2 (Smietanski et al., 2014).

Another modification which could contribute to protecting self mRNA is adenosine *N6*-methylation. Bulk analysis of HeLa cell and viral mRNA caps in the 70s revealed that when the first nucleotide (N1) of a transcript is an adenosine, it is often methylated at the *N6* position resulting in *N6,O2'*-dimethyladenosine (m6Am) (Wei et al., 1976; 1975). The function of m6Am in translation is poorly defined, and the as-yet-unknown enzyme responsible for *N6*-methylation at N1 has only been partially purified from HeLa cells (Keith et al., 1978a; 1978b). Using the partially purified enzyme, Keith *et al.* showed that it preferentially modifies Cap1 mRNAs (Keith et al., 1978a). As 20-30% of cellular mRNAs contain m6Am at N1 (Wei et al., 1976; 1975), addressing its role in limiting IFIT1 self-recognition warrants further investigation.

4.1.4 Viral evasion of IFIT1

4.1.4.1 Viral N1 methylation

A large majority of viruses have evolved mechanisms to N1 methylate their mRNA, thereby potentially escaping human IFIT1 detection. Indeed, several studies have shown that mutants of SARS-CoV, Human CoV, Dengue Virus, and West Nile virus are more sensitive to human IFIT1 than their wild-type counterparts (Habjan et al., 2013; Menachery et al., 2014; Pinto et al., 2015; Zust et al., 2013). Altogether, these studies are consistent with the structural and functional analysis performed in this thesis.

One recent interesting study focused on viruses from the *Paramyxoviridae* family (Young et al., 2016). They showed that two members of the *Rubulavirus* genus, Parainfluenza virus type 5 (PIV5) and PIV2, are more sensitive to IFIT1-mediated restriction than PIV3, a member of the *Respirovirus* genus. By extracting viral mRNA from infected cells, they showed that PIV5 and other rubulavirus mRNA were sensitive to IFIT1 translational inhibition, whereas PIV3 and other non-rubulavirus mRNA were resistant. *In vitro* CMTr1 treatment of rubulavirus mRNA isolated from infected cells partially rescued them from IFIT1 inhibition, suggesting that incomplete *in vivo* methylation explains, in part, their enhanced sensitivity to IFIT1-mediated restriction.

Young *et al.* made several critical observations in their study. First, they showed that incomplete methylation of viral mRNA only partially explained the enhanced sensitivity of PIV5 to IFIT1, suggesting that IFIT1's ability to overcome N1 methylation can also contribute to

restricting PIV5. Second, *in vitro* N1 methylation rescued viral nucleoprotein mRNA, but not viral matrix protein mRNA from translational inhibition, which is consistent with the RNA-dependent effects observed in our gel-shift assays (section 3.3.9). Third, even though they encode a conserved Cap1-MTase domain in their L protein, the rubulaviruses tested do not uniformly N1 methylate their mRNA. Interestingly, incomplete methylation has also been described for other members of the *Paramyxoviridae* family, such as Newcastle disease virus (Colonno and Stone, 1976) and respiratory syncytial virus (Barik, 1993), but not measles virus (Yoshikawa et al., 1986). Thus, the presence of a Cap1 MTase in a viral genome should not preclude IFIT1 inhibition, as the MTase could be inactive or inefficient.

4.1.4.2 Viral 5' RNA structure and other modifications

The tight interface formed by the IFIT1 tunnel is important for its ability to sense cap-proximal modifications and compete with eIF4F, but also confers selectivity for RNA with unstructured ends. Two studies have shown that alphaviruses (family *Togaviridae*) can exploit this to their advantage. Alphaviruses replicate in the cytoplasm and do not encode an N1 MTase, therefore producing mainly Cap0-mRNAs. Despite the lack of N1 methylation, pathogenic alphaviruses can antagonize mouse Ifit1 activity *in vivo* by encoding secondary structure at their genomic 5'-ends (Hyde et al., 2014). Experiments in which 5' RNA structures were mutated to reduce their stability and decrease their proximity to the 5' end enhanced alphavirus susceptibility to mouse Ifit1 *in vitro* and *in vivo* (Hyde et al., 2014; Reynaud et al., 2015).

Other virus families also encode stable 5' secondary structure (Hyde and Diamond, 2015), but it is not clear if they impact mouse Ifit1 in a similar manner. It is also unclear if this mechanism participates in escaping human IFIT1 activity. *In vitro* binding analysis shows that base-pairing and cap-proximal secondary structure interferes with human IFIT1 binding (chapter 2 and 3), consistent with the structure, but *in vivo* activity could be influenced by factors such as IFIT1 protein interacting partners.

Other modifications potentially exploited by viruses include N2 ribose methylation and m6Am, both of which have been detected in viral mRNA isolated from infected cells (reviewed in (Banerjee, 1980)). These modifications are likely catalyzed by the endogenous mammalian enzymes as viruses do not encode the corresponding proteins. For the most part, N2 methylation and m6Am are not as prevalent as N1 ribose methylation, but they could potentially contribute to

viral evasion strategies, especially as N2 methylation alone, or N2 in combination with N1, has a strong effect on preventing human IFIT1 binding.

4.1.4.3 New perspectives on IFIT1 activity

A recently published study challenged the notions that IFIT1 targets Cap0-mRNAs during infection, and that IFIT1 distinguishes between viral Cap0- and Cap1-mRNAs (Daugherty et al., 2016), which is at odds with the structural and functional work performed herein, and the multiple *in vitro* binding studies and viral infectivity assays already published (see above). The conclusions of Daugherty *et al.* were based on several experiments, which I will attempt to reinterpret in light of the structural and functional observations presented in this thesis. In the first experiment, they overexpressed IFIT1 or IFIT1B in yeast cells, whose mRNAs have only Cap0-structures, and showed that both proteins inhibit yeast growth. Co-expression of CMTr1 rescued the yeast cells from IFIT1B-mediated inhibition, but not from IFIT1. This experiment confirmed that IFIT1B proteins (which includes mouse Ifit1) are highly sensitive to N1 methylation, but suggest that IFIT1 is not distinguishing between Cap0- and Cap1- structures. Daugherty *et al.* proposed that IFIT1 is targeting another ‘non-self’ pattern that is displayed by yeast mRNAs. While this could be the case, a more likely explanation that is consistent with our data is that IFIT1 is overcoming the steric hindrance imposed by N1 methylation in an RNA-dependent manner, targeting one or more essential yeast mRNAs that are not properly protected.

To discern between these possibilities, several experiments should be performed. Most importantly, the yeast genetic assay needs to be validated using IFIT1 mutants described herein to show if the observed inhibition of growth is mediated by cap-recognition, general RNA binding inside the tunnel or outside, or another unknown mechanism not related to RNA binding. Ribosomal profiling of yeast mRNA in the presence of IFIT1 and CMTr1 can be performed to ascertain which mRNAs are inhibited (Ingolia, 2016), and thereby uncover RNA sequence and structure motifs that are susceptible to IFIT1. High-throughput analysis of IFIT1 RNA-binding can be performed to completely characterize its RNA-binding spectrum, for instance using an *in vitro* evolution approach or by generating a large pool of RNA with all possible combinations of nucleotides at the first four positions. Finally, co-crystallization attempts between Cap1-oligos and human IFIT1 can be attempted, to uncover how IFIT1 may overcome N1 methylation. Of note, Cap1-oligos do not co-crystallize with IFIT1 under the same conditions which allowed co-

crystallization with PPP-RNA and m7Gppp-RNA. Thus, *in silico* modelling could be performed instead.

In a second set of experiments, Daugherty *et al.* overexpressed IFIT1 and IFIT1B in mammalian cells and showed that only IFIT1 proteins inhibited growth of vesicular stomatitis virus (VSV), a virus that is known to uniformly N1 methylate its mRNAs *in vivo* (Rose, 1975). These observations are more challenging to interpret without knowing the mode of inhibition by IFIT1, but they could be rationalized by the existence of another unknown ‘non-self’ pattern on VSV mRNA, as proposed by Daugherty *et al.*, or by an ability to sterically overcome VSV N1 methylation, as suggested in chapter 3. Of note, VSV encodes only 5 mRNAs, all of which have relatively short 5′ UTRs (14-49 nt) that are potentially unstructured and not dependent on eIF4F levels for translation (Connor and Lyles, 2002). These properties could confer VSV with high susceptibility to human IFIT1 translational inhibition. Interestingly, all 5 VSV mRNAs have a consensus sequence at their transcription start site, 5′ AACAGnnAUC (where n is any nucleotide), and this sequence could contribute to overcoming Cap1 steric hindrance as in the structures of RNA-bound IFIT1 we observed adenine specific hydrogen bonds at N2 and N4. These notions will hopefully form the basis for future experiments to test their validity.

The third set of experiments performed by Daugherty *et al.* arguing against the prevailing model of IFIT1 activity was performed with vaccinia virus (VACV) infection in the presence of IFIT1 or IFIT1B. Wild-type VACV infection was resistant to both sets of proteins, but mutant VACV (lacking viral N1 methylation) was rendered susceptible to IFIT1B and not IFIT1. These observations can be explained by the presence of other ‘self’ mechanisms preventing IFIT1 recognition, as presented in this thesis. Whether this is due to N2 methylation or m6Am remains to be seen. Both modifications accompany N1 methylation of VACV mRNA *in vivo* (Boone and Moss, 1977), but *in vitro* assays have shown that the enzymes catalyzing N2 ribose methylation (CMTr2) and m6 adenine methylation (at N1) are enhanced by prior N1 ribose methylation (Keith *et al.*, 1978a; Werner *et al.*, 2011). Therefore, in absence of viral N1 methylation, it is not clear if mutant VACV mRNAs are substrates of endogenous CMTr2 and the m6Am MTase.

4.1.4.4 The differing roles of IFIT1 and IFIT1B

The role of N1 methylation in escaping mouse *Ifit1* restriction has been extensively characterized *in vivo*. Much of our understanding of human IFIT1 in this antiviral process was based on work with mouse *Ifit1*, and these comparisons were predicated on the assumption that human IFIT1 and

mouse Ifit1 are orthologs, which is incorrect as mouse Ifit1 belongs to the IFIT1B subfamily. That human IFIT1 and mouse Ifit1 belong to distinct subfamilies which diverged sometime near the origin of placental mammals was only recently brought to light (Daugherty et al., 2016).

Based on the functional assays which I summarized above, Daugherty *et al.* proposed that IFIT1 and IFIT1B proteins have distinct 5' specificities which give rise to distinct antiviral activities. This is probably only partially correct, since *in vitro* binding assays show that human IFIT1 and mouse Ifit1 have similar affinity to Cap0-mRNAs, and are similarly impacted by Cap1 methylation (Habjan et al., 2013; Kumar et al., 2014). Furthermore, virus infectivity assays reveals common viral targets (Habjan et al., 2013). Thus, IFIT1 and IFIT1B proteins may have overlapping 5' specificity.

A more accurate description might be that IFIT1 and IFIT1B have evolved different RNA-sequence or RNA-structure preferences that confers differential sensitivity to 5'-modifications. For instance, whereas both are sensitive to N1 methylation on most sequences, IFIT1 may overcome N1 methylation in an RNA-dependent manner, leading to greater antiviral potential towards some Cap1 generating viruses (for e.g. VSV). Regardless, the idea that IFIT1B proteins have been defeated by most viral evasion strategies, whereas IFIT1 has somewhat adapted, is supported by observations that IFIT1B has been pseudogenized or deleted multiple times during mammalian evolution (Daugherty et al., 2016). Although human genomes have retained IFIT1B, it may not be interferon-inducible (Fensterl and Sen, 2011), and our sequence analysis shows that it acquired several mutations in key m7Gppp- interacting residues that impact binding, consistent with its lack of activity in functional assays (Daugherty et al., 2016).

4.2 Viral PPP-RNA recognition

IFIT proteins can also interact with single-stranded PPP-RNA. In humans, this is mediated by IFIT1 and IFIT5, potentially allowing them to sequester viral PPP-RNAs during infection and block viral replication (Pichlmair et al., 2011). Crystal structures of PPP-RNA bound IFIT1 and IFIT5 support that notion, and reveal a common mechanism for the specific recognition of single-stranded PPP-RNA.

4.2.1 PPP-RNA recognition by IFIT5

Knockdown and overexpression assays support a role for IFIT5 in limiting replication of VSV, a non-segmented negative-strand ssRNA virus (nsNSV) (this work and (Feng et al., 2013; Pichlmair

et al., 2011). In contrast, knockdown of IFIT5 did not enhance replication of Rift Valley fever virus (RVFV) or Influenza A virus (IAV), both of which are segmented negative-strand RNA viruses (Pichlmair et al., 2011). These differences suggest potential virus- or virus-family specific mode of action for IFIT5.

The precise nature of PPP-containing RNAs targeted by IFIT5 during VSV infection is not clear, and thus the stage of viral replication inhibited by IFIT5 is unknown. IFIT5 could potentially target incoming viral ribonucleoprotein (vRNPs), positive-strand antigenomic template RNA, or negative-strand progeny vRNPs, all of which would display 5'-triphosphates. IFIT5 could also target the transient PPP on nascent mRNA, which during transcription remain uncapped until the VSV polymerase has incorporated 31 nucleotides (Tekes et al., 2011). Finally, VSV and other nsNSVs are known to produce large amounts of PPP-containing leader RNA, which are ~ 50 nt RNAs generated early during transcription (Grinnell and Wagner, 1984). These leader RNAs could be potential targets of IFIT5, or could be a part of a mechanism to antagonize IFIT5 activity (Pichlmair et al., 2011).

Further experiments are required to discern between these possibilities. For instance, PCR amplification of viral RNA during infection could tell us at which step viral replication is blocked, and IFIT5 pull-downs during infection followed by sequencing would show us which viral RNA species it recognizes. Other potential experiments could assay for co-localization between IFIT5 and viral RNPs in cells, or investigate the interaction between IFIT5 and purified viral RNPs by electron microscopy.

Given the tendency of innate immune sensors to require more than one signature on their cognate PAMPs for activity (e.g. RIG-I requires PPP and a blunt end), it would be interesting to see if IFIT5 can synergistically bind PPP-RNAs containing other motifs. Indeed, PPP-RNAs (47-110 nts in length) have been shown to activate PKR if they contain short stem loops near the middle (Nallagatla et al., 2007), and it is conceivable that IFIT5 may target similar motifs. *In vitro* evolution of ligands, analogous to those performed with PKR (Zheng and Bevilacqua, 2004), could be used to discover novel PPP-containing motifs recognized by IFIT5.

4.2.2 PPP-RNA recognition by IFIT1

The role of IFIT1 in recognizing viral PPP-RNAs remains a topic of debate. Several experiments show that human IFIT1 preferentially recognizes capped-RNAs over PPP-RNAs (chapter 3 and (Habjan et al., 2013; Kumar et al., 2014), indicating that its primary function is to inhibit translation

of viral mRNAs. However, it is well-established that IFIT1 is at the center of a large multi-protein complex that includes IFIT2, IFIT3, and other host factors, and therefore a role for IFIT1 in recognizing viral PPP-RNA in this context should be considered (Pichlmair et al., 2011).

IFIT1 has been shown to inhibit replication of VSV, IAV, and RVFV (Pichlmair et al., 2011). As with IFIT5, it is unknown at which stage IFIT1 would recognize the PPP-RNAs. Both IAV and RVFV ‘circularize’ their genomes through complementarity between their 5’ and 3’ ends, which is facilitated by nucleocapsid proteins coating the viral RNA and the association of viral polymerase with the genome ends (Moeller et al., 2012; Weber et al., 2013). Thus, it is unclear how IFIT1 could access the 5’-PPP on genomic RVFV and IAV RNA. One possibility is that IFIT1 displaces the polymerase and accesses the 5’-end in a stochastic manner. IFIT1 could also target intermediates of RVFV replication in the cytosol. As IAV resides in the nucleus, it is unlikely that IFIT1 can access its PPP-RNA intermediates. Regardless, one study utilized IFIT1 knockout A549 cells and showed that IAV replication is unaffected (Pinto et al., 2015), arguing against a role for IFIT1 in inhibiting IAV replication. Finally, it is altogether possible that the antiviral effect of IFIT1 observed in these assays is due to viral Cap1-mRNA binding and translational inhibition, as described above for VSV.

The structure-based mutants described in chapter 3 show that it is possible to disrupt capped-RNA binding by IFIT1 without impacting its PPP-RNA binding activities. Therefore, future structure-guided mutations could be performed to further engineer IFIT1 mutants that only bind capped-RNA, or only bind PPP-RNA. These mutants could be used in cell-based assays to determine which of its two activities mediate an antiviral effect towards different viruses.

4.2.3 Diverse *in vitro* nucleic acid binding by IFIT5

RNA binding and functional assays performed by several groups have revealed unexpected diversity in nucleic acid recognition by human IFIT5, which in addition to PPP-RNAs has been shown to bind 5’ monophosphate RNAs (P-RNA), cellular tRNAs, and AT-rich dsDNA (Feng et al., 2013; Katibah et al., 2014; 2013). One study has even shown binding between IFIT5 and capped-RNA (discussed further below). The structural basis for these observations are unknown, although binding of some of these nucleic acids could be mediated by regions outside the IFIT5 tunnel, as demonstrated for IFIT5 and tRNA binding (Katibah et al., 2013). The functional relevance of these interactions has not been explored or validated, but it implicates IFIT5 in

multiple modes of viral nucleic acid binding, and suggests a potential role for IFIT5 in tRNA biology (Katibah et al., 2014).

4.2.3.1 Does IFIT5 specifically bind capped-RNA?

As IFIT1 and IFIT5 genes diverged very early during mammalian evolution (Daugherty et al., 2016), one would expect non-redundant antiviral activities. The prevailing model suggests that IFIT5 antiviral activity is based on its ability to recognize viral PPP-RNAs, while IFIT1 targets primarily capped viral mRNAs to block their translation. This is supported by structural analysis of human IFIT5, which showed that, in comparison to IFIT1, its putative cap binding pocket is spatially restricted and cannot accommodate m7G. Furthermore, its PPP recognition mechanism is not compatible with cap binding, as the γ -phosphate is oriented in the opposite direction of the putative cap-binding pocket. Binding assays using multiple techniques performed by several groups support the notion that IFIT5 preferentially binds PPP-RNAs over capped-RNAs (this work and (Habjan et al., 2013; Kumar et al., 2014)). Additionally, our *in vitro* translation assays validate that IFIT5 cannot appreciably inhibit translation of capped RNAs.

Considering these multiple lines of evidence, the finding by Katibah *et al.* that IFIT5 binds capped-RNA in polyacrylamide gel-shift assays is unexpected (Katibah et al., 2014). Even more so, the affinity between IFIT5 and the capped-RNA in their assay was ~ 1.7 nM. However, in the same experiment, they showed that IFIT5 binds PPP-RNA and P-RNA with similar affinity as capped-RNA (1.4 and 0.53 nM). This suggests a lack of specific binding for 5' modifications under their conditions, as additional contacts to recognize two extra phosphates (PPP vs P), or two extra phosphates and a cap moiety (m7Gppp vs P), would be expected to yield greater affinity for capped and PPP-RNA compared to P-RNA.

The affinity for PPP-RNA observed by Katibah *et al.* is 2 orders of magnitude stronger than other reported values, which are in the range of 200-500 nM (this work and (Kumar et al., 2014)). While I cannot rule out that our assay conditions and that of Kumar *et al.* might be underestimating the IFIT5-PPP-RNA binding strength, these discrepancies suggest that the polyacrylamide gel-shift conditions utilized by Katibah *et al.* favor non-specific binding, possibly explaining why they observed an interaction between IFIT5 and capped-RNA. Indeed, examining their assay condition shows that it lacks any competitive inhibitors for non-specific binding, such as heparin, a standard component when performing gel shifts (Hellman and Fried, 2007). Of note,

polyacrylamide has been shown to alter dissociation kinetics of protein-nucleic acid interactions, thereby stabilizing complexes during polyacrylamide gel shifts (Fried and Liu, 1994). Thus, to further understand the relative contribution of IFIT5 and IFIT1 to cap recognition *in vivo*, future experiments using isothermal titration calorimetry, a label-free and true equilibrium binding assay, should be performed to clarify these contradictory results.

4.3 Additional differences between human and mouse IFIT genes

Our understanding of human IFIT antiviral activity relies heavily on mouse *in vitro* and *in vivo* work. However, the distinct complement of IFIT genes in the two species, which differs both in composition and potential RNA-binding specificities, could have an impact on our ability to extrapolate human IFIT function based on mouse data. These have been partially discussed above. The first major difference is that mice have deleted IFIT1, and their antiviral cap recognition activity resides mainly in mouse Ifit1 (which is an IFIT1B protein) (Daugherty et al., 2016). Second, unlike IFIT1B from mouse and other species, human IFIT1B may not be interferon-inducible (Fensterl and Sen, 2011), and potentially lacks a function in RNA binding. As proposed by Daugherty *et al.*, this loss of function may be a consequence of the numerous viral evasion strategies that have effectively rendered IFIT1B antiviral activity limited (Daugherty et al., 2016). The difference between human IFIT1 and mouse Ifit1 is not entirely clear yet, but their differential sensitivity to 5' modifications may give rise to distinct antiviral activities *in vivo*. Third, mice have additional copies of IFIT1B (Daugherty et al., 2016; Fensterl and Sen, 2011), which may have emerged to functionally compensate for the lack of IFIT1. However, these IFIT1B-like proteins (mouse Ifit1b and mouse Ifit1c) have both mutated key m7Gppp- interacting residues, suggesting that they cannot support RNA binding on their own. Indeed, pull-downs did not detect a direct interaction between capped RNA and mouse Ifit1c (Habjan et al., 2013).

The fourth major difference lies within the potential to form IFIT complexes. Human IFITs are known to form a multi-protein complex made up of IFIT1, IFIT2, and IFIT3 (Pichlmair et al., 2011; Stawowczyk et al., 2011). The function of this complex is unknown, but its ability to recruit multiple host factors involved in RNA processing suggests a function in RNA biology (Pichlmair et al., 2011). In mice, the composition of this large complex is potentially different or altogether absent. While an interaction between mouse IFIT2 and mouse IFIT3 has been detected (Siegfried et al., 2013), neither protein interacts with mouse Ifit1 (Habjan et al., 2013; Siegfried et al., 2013).

Instead, pull-downs from mouse cell lysates using capped-RNA showed that mouse Ifit1c co-purifies with mouse Ifit1 (Habjan et al., 2013). Therefore, instead of binding RNA directly, mouse Ifit1c is potentially forming a heterodimer with mouse Ifit1 that could functionally compensate for the lack of an IFIT1-IFIT2-IFIT3 complex. This heterodimer may even compensate for the absence of IFIT1.

Finally, and for reasons unknown, IFIT5 is deleted from mouse, rat and other rodent genomes (Fensterl and Sen, 2011). The evolutionary forces which have led to the differing IFIT repertoires in humans and mice are not altogether apparent, but the potential for forming distinct IFIT-RNA interactions or IFIT-IFIT complexes could shape their overall antiviral responses.

4.4 Therapeutic potential for targeting IFIT1

A better understanding of IFIT1 RNA-binding activity could impact drug design and development. Two potential avenues will be discussed, mRNA therapeutics and viral vaccine design. For mRNA therapeutics, the objective would be to avoid unwanted IFIT1 recognition, whereas for vaccine design, the aim is to rationally engineer viruses to enhance their susceptibility to IFIT1 and effectively attenuate them.

4.4.1 mRNA therapeutics

In vitro transcribed mRNAs (IVTmRNAs) are a class of drugs that have potential therapeutic applications in cancer immunotherapy, protein-replacement therapy, and genome engineering, among others (Sahin et al., 2014). To maximize their benefit, factors such as half-life, stability, immunostimulatory potential, and translational efficiency of the IVTmRNAs need to be examined (Sahin et al., 2014). IFIT1 could hinder the efficacy of IVTmRNAs delivered *in vivo* or *ex vivo* by inhibiting their translation. Although IFIT1 is not normally expressed in most cells, IVTmRNAs could lead to unwanted IFIT1 upregulation if they are capable of triggering RNA sensing pathways, or if their protein products are immunogenic. Additionally, IFIT1 upregulation has been reported in primary human fibroblasts (Moll et al., 2011), some sarcoma cell lines (Berchtold et al., 2013), during malignancies (e.g. in CD34+ cells of patients with myelodysplastic syndromes (Pellagatti et al., 2006)), during chronic viral infection (e.g. in livers of HCV individuals (Patzwahl et al., 2001)), and in diseases associated with type I interferonopathies (e.g. systemic lupus erythematosus (Ye et al., 2003)).

The pharmacokinetic profile of IVTmRNAs can be fine-tuned in several ways, for example, by altering the cap structure (Sahin et al., 2014). Analogs of cap dinucleotide (i.e. modifications on m⁷GpppG) can be added to *in vitro* transcription reactions and are efficiently incorporated at the 5' end of mRNAs (Stepiński et al., 2001). These have been shown to enhance IVTmRNA stability and translation (Kuhn et al., 2010). However, they could potentially be targeted by IFIT1 as its cap recognition mechanism is relatively plastic and not specific for *bona fide* mRNA cap. For instance, one commonly used cap analog that enhances cap-dependent translation is 7-methyl(3'-O-methyl)GpppG (or "anti-reverse" cap analog, ARCA); it does so by preventing reverse incorporation of m⁷GpppG during *in vitro* transcription (Stepiński et al., 2001). Preliminary data suggests that ARCA-containing mRNAs are inhibited by human IFIT1 during *in vitro* translation assays (unpublished observations), supporting the need to test these and other cap analogs for IFIT1 binding. Regardless, the cap recognition mechanisms of IFIT1 and eIF4E are highly divergent and, in principle, these differences could be harnessed by cap modifications engineered to maintain strong eIF4E binding while avoiding IFIT1.

Another way to enhance IVTmRNA efficiency is through 5'-UTR modifications (Sahin et al., 2014). Changes in cap-proximal secondary structure can modulate translation efficiency, but they also influence IFIT1 mRNA binding. For example, whereas decreased structure at the 5' end could promote IVTmRNA translation (Babendure et al., 2006), it could render them more susceptible to unwanted IFIT1 recognition. In that regard, it would be useful to incorporate multiple cap-proximal modifications such as N1+N2 methylations (as demonstrated in this work), to minimize IFIT1 binding.

4.4.2 Implications for vaccine design

Recent viral outbreaks highlight the need for rapid, vaccine development. An emerging strategy is to generate live, attenuated viruses by mutating their Cap1 methyltransferase machinery, therefore rendering them more susceptible to IFN and IFIT1/IFIT1B activity (Li et al., 2013; Menachery et al., 2014; Zhang et al., 2014; Zust et al., 2013). These mutant viruses are stable in permissive cell culture conditions but when infected into animal models are cleared rapidly, although not before eliciting a robust immune response that confers protection against future challenge with pathogenic, wild-type virus. Importantly, the recombinant viruses do not show signs of reversion to wild-type after multiple passages *in vitro* (Li et al., 2013; Zust et al., 2013); the recombinant viruses also retain the engineered mutation *in vivo* (Li et al., 2013).

Another approach to design live, attenuated vaccines is through viral 5' RNA modifications (Hyde et al., 2014; Reynaud et al., 2015). As described above, alphaviruses are positive-strand ssRNA viruses that replicate in the cytoplasm but lack any Cap1-MTase, therefore producing mainly Cap0 structures (Dubin et al., 1977). Pathogenic alphaviruses can still antagonize mouse Ifit1 activity by encoding cap-proximal secondary structure (Hyde et al., 2014). Mutations in these structural elements, decreasing their stability and/or increasing their distance from the 5'-end, results in increased sensitivity to IFN and mouse Ifit1 in cell culture (Hyde et al., 2014; Reynaud et al., 2015), and a corresponding decrease of virulence in mice *in vivo* (Hyde et al., 2014). Notably, Hyde *et al.* have demonstrated that a single G-to-A point mutant at nucleotide 3 of Venezuelan equine encephalitis virus (VEEV), which reduces its 5' RNA-structure stability and increases sensitivity to mouse Ifit1, is largely responsible for the conversion of VEEV from a virulent strain into an attenuated vaccine strain (Hyde et al., 2014). The observations of Hyde *et al.* and Reynaud *et al.* are highly consistent with the structural and biochemical assays performed in this thesis, which showed a requirement for single stranded 5' ends and overhangs of approximately 3-5 nucleotides.

Both vaccine strategies described above could potentially benefit from the structural and biochemical analysis performed in this thesis, so as to rationally engineer viral vaccine strains and increase their sensitivity to human IFIT1 inhibition. It should be noted, however, that further work is needed to completely understand the RNA binding spectrum of human IFIT1. Moreover, some of the studies described above were performed with *in vivo* rodent models, or utilized rodent and (non-human) primate cell lines, all of which potentially encode a distinct complement of functional IFIT proteins when compared to humans (Daugherty et al., 2016). Thus, further structural and functional comparisons of IFIT1 and IFIT1B proteins from humans and these model organisms is required to fully appreciate the potential of this strategy for developing analogous human vaccines. The work described in this thesis will provide a framework for future structural and biochemical analysis addressing these questions.

REFERENCES

- Abbas, Y.M., Pichlmair, A., Góna, M.W., Superti-Furga, G., and Nagar, B. (2013). Structural basis for viral 5'-PPP-RNA recognition by human IFIT proteins. *Nature* *494*, 60–64.
- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* *66*, 213–221.
- Adelman, K., and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet* *13*, 720–731.
- Ahola, T., and Kääriäinen, L. (1995). Reaction in alphavirus mRNA capping: formation of a covalent complex of nonstructural protein nsP1 with 7-methyl-GMP. *Proc Natl Acad Sci USA* *92*, 507–511.
- Ahola, T., Laakkonen, P., Vihinen, H., and Kääriäinen, L. (1997). Critical residues of Semliki Forest virus RNA capping enzyme involved in methyltransferase and guanylyltransferase-like activities. *J Virol* *71*, 392–397.
- Amrani, N., Ghosh, S., Mangus, D.A., and Jacobson, A. (2008). Translation factors promote the formation of two states of the closed-loop mRNP. *Nature* *453*, 1276–1280.
- Anthony, S.J., Epstein, J.H., Murray, K.A., Navarrete-Macias, I., Zambrana-Torrel, C.M., Solovyov, A., Ojeda-Flores, R., Arrigo, N.C., Islam, A., Khan, S.A., et al. (2013). A Strategy To Estimate Unknown Viral Diversity in Mammals. *MBio* *4*, e00598–13–e00598–13.
- Aregger, M., and Cowling, V.H. (2013). Human cap methyltransferase (RNMT) N-terminal non-catalytic domain mediates recruitment to transcription initiation sites. *Biochemical Journal* *455*, 67–73.
- Aregger, M., and Cowling, V.H. (2017). Regulation of mRNA capping in the cell cycle. *RNA Biol* *14*, 11–14.
- Babendure, J.R., Babendure, J.L., Ding, J.-H., and Tsien, R.Y. (2006). Control of mammalian translation by mRNA structure near caps. *Rna* *12*, 851–861.
- Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences* *98*, 10037–10041.
- Baltimore, D. (1971). Expression of animal virus genomes. *Bacteriological Reviews* *35*, 235–241.
- Bandyopadhyay, S.K., Kumar, R., Rubin, B.Y., and Sen, G.C. (1992). Interferon-inducible gene expression in HL-60 cells: effects of the state of differentiation. *Cell Growth Differ* *3*, 369–375.

- Bandyopadhyay, S.K., Leonard, G.T., Bandyopadhyay, T., Stark, G.R., and Sen, G.C. (1995). Transcriptional induction by double-stranded RNA is mediated by interferon-stimulated response elements without activation of interferon-stimulated gene factor 3. *Journal of Biological Chemistry* 270, 19624–19629.
- Banerjee, A.K. (1980). 5'-terminal cap structure in eucaryotic messenger ribonucleic acids. *Microbiol. Rev.* 44, 175–205.
- Barbalat, R., Ewald, S.E., Mouchess, M.L., and Barton, G.M. (2011). Nucleic Acid Recognition by the Innate Immune System. *Annu Rev Immunol* 29, 185–214.
- Barik, S. (1993). The structure of the 5' terminal cap of the respiratory syncytial virus mRNA. *J. Gen. Virol.* 74, 485–490.
- Berchtold, S., Lampe, J., Weiland, T., Smirnow, I., Schleicher, S., Handgretinger, R., Kopp, H.G., Reiser, J., Stubenrauch, F., Mayer, N., et al. (2013). Innate Immune Defense Defines Susceptibility of Sarcoma Cells to Measles Vaccine Virus-Based Oncolysis. *J Virol* 87, 3484–3501.
- Berchtold, S., Manncke, B., Klenk, J., Geisel, J., Autenrieth, I.B., and Bohn, E. (2008). Forced IFIT-2 expression represses LPS induced TNF-alpha expression at posttranscriptional levels. *BMC Immunol.* 9, 75.
- Berger, I., Fitzgerald, D.J., and Richmond, T.J. (2004). Baculovirus expression system for heterologous multiprotein complexes. *Nat Biotechnol* 22, 1583–1587.
- Bélanger, F., Stepinski, J., Darzynkiewicz, E., and Pelletier, J. (2010). Characterization of hMTr1, a human Cap1 2'-O-ribose methyltransferase. *Journal of Biological Chemistry* 285, 33037–33044.
- Bin Wu, and Hur, S. (2015). How RIG-I like receptors activate MAVS. *Curr Opin Virol* 12, 91–98.
- Blatch, G.L., and Lassle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays* 21, 932–939.
- Blyussen, H.A., Vlietstra, R.J., Faber, P.W., Smit, E.M., Hagemeyer, A., and Trapman, J. (1994). Structure, chromosome localization, and regulation of expression of the interferon-regulated mouse Ifi54/Ifi56 gene family. *Genomics* 24, 137–148.
- Bollati, M., Milani, M., Mastrangelo, E., Ricagno, S., Tedeschi, G., Nonnis, S., Decroly, E., Selisko, B., de Lamballerie, X., Coutard, B., et al. (2009). Recognition of RNA Cap in the Wesselsbron Virus NS5 Methyltransferase Domain: Implications for RNA-Capping Mechanisms in Flavivirus. *J Mol Biol* 385, 140–152.
- Boone, R.F., and Moss, B. (1977). Methylated 5'-terminal sequences of vaccinia virus mRNA species made in Vivo at early and late times after infection. *Virology* 79, 67–80.

- Botos, I., Segal, D.M., and Davies, D.R. (2011). The Structural Biology of Toll-like Receptors. *Structure* *19*, 447–459.
- Bouloy, M., Plotch, S.J., and Krug, R.M. (1978). Globin mRNAs are primers for the transcription of influenza viral RNA in vitro. *Proc Natl Acad Sci USA* *75*, 4886–4890.
- Bouloy, M., Plotch, S.J., and Krug, R.M. (1980). Both the 7-methyl and the 2'-O-methyl groups in the cap of mRNA strongly influence its ability to act as primer for influenza virus RNA transcription. *Proc Natl Acad Sci USA* *77*, 3952–3956.
- Bouvet, M., Debarnot, C., Imbert, I., Selisko, B., Snijder, E.J., Canard, B., and Decroly, E. (2010). In Vitro Reconstitution of SARS-Coronavirus mRNA Cap Methylation. *PLoS Pathog* *6*, e1000863.
- Bruns, A.M., Leser, G.P., Lamb, R.A., and Horvath, C.M. (2014). The Innate Immune Sensor LGP2 Activates Antiviral Signaling by Regulating MDA5-RNA Interaction and Filament Assembly. *Mol Cell* *55*, 771–781.
- Calero, G., Wilson, K.F., Ly, T., Rios-Steiner, J.L., Clardy, J.C., and Cerione, R.A. (2002). Structural basis of m7GpppG binding to the nuclear cap-binding protein complex. *Nat Struct Biol* *9*, 912–917.
- Chebath, J., Merlin, G., Metz, R., Benech, P., and Revel, M. (1983). Interferon-induced 56,000 Mr protein and its mRNA in human cells: molecular cloning and partial sequence of the cDNA. *Nucleic Acids Res* *11*, 1213–1226.
- Chen, Q., Sun, L., and Chen, Z.J. (2016). Regulation and function of the cGAS–STING pathway of cytosolic DNA sensing. *Nat Immunol* *17*, 1142–1149.
- Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., IUCr (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* *66*, 12–21.
- Chen, Y., Cai, H., Pan, J., Xiang, N., Tien, P., Ahola, T., and Guo, D. (2009). Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proceedings of the National Academy of Sciences* *106*, 3484–3489.
- Chen, Y., Su, C., Ke, M., Jin, X., Xu, L., Zhang, Z., Wu, A., Sun, Y., Yang, Z., Tien, P., et al. (2011). Biochemical and Structural Insights into the Mechanisms of SARS Coronavirus RNA Ribose 2'-O-Methylation by nsp16/nsp10 Protein Complex. *PLoS Pathog* *7*, e1002294.
- Chen, Y., Tao, J., Sun, Y., Wu, A., Su, C., Gao, G., Cai, H., Qiu, S., Wu, Y., Ahola, T., et al. (2013). Structure-function analysis of severe acute respiratory syndrome coronavirus RNA cap guanine-N7-methyltransferase. *J Virol* *87*, 6296–6305.
- Cheng, E., and Mir, M.A. (2012). Signatures of host mRNA 5' terminus for efficient hantavirus cap snatching. *J Virol* *86*, 10173–10185.

- Cho, H., Shrestha, B., Sen, G.C., and Diamond, M.S. (2013). A role for ifit2 in restricting west nile virus infection in the brain. *J Virol* 87, 8363–8371.
- Civril, F., Bennett, M., Moldt, M., Deimling, T., Witte, G., Schiesser, S., Carell, T., and Hopfner, K.-P. (2011). The RIG-I ATPase domain structure reveals insights into ATP-dependent antiviral signalling. *EMBO Rep* 12, 1127–1134.
- Cleaves, G.R., and Dubin, D.T. (1979). Methylation status of intracellular dengue type 2 40 S RNA. *Virology* 96, 159–165.
- Colonno, R.J., and Stone, H.O. (1976). Newcastle disease virus mRNA lacks 2'-O-methylated nucleotides. *Nature* 261, 611-614.
- Connor, J.H., and Lyles, D.S. (2002). Vesicular stomatitis virus infection alters the eIF4F translation initiation complex and causes dephosphorylation of the eIF4E binding protein 4E-BP1. *J Virol* 76, 10177–10187.
- Cowling, V.H. (2010). Regulation of mRNA cap methylation. *Biochem J* 425, 295–302.
- Cui, S., Eisenächer, K., Kirchhofer, A., Brzózka, K., Lammens, A., Lammens, K., Fujita, T., Conzelmann, K.-K., Krug, A., and Hopfner, K.-P. (2008). The C-Terminal Regulatory Domain Is the RNA 5'-Triphosphate Sensor of RIG-I. *Mol Cell* 29, 169–179.
- D'Andrea, L.D., and Regan, L. (2003). TPR proteins: the versatile helix. *Trends Biochem Sci* 28, 655–662.
- D'Ascenzo, L., Leonarski, F., Vicens, Q., and Auffinger, P. (2016). “Z-DNA like” fragments in RNA: a recurring structural motif with implications for folding, RNA/protein recognition and immune response. *Nucleic Acids Res* 44, 5944–5956.
- Daffis, S., Szretter, K.J., Schriewer, J., Li, J., Youn, S., Errett, J., Lin, T.-Y., Schneller, S., Zust, R., Dong, H., et al. (2010). 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature* 468, 452–456.
- Das, A.K., Cohen, P.W., and Barford, D. (1998). The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions. *Embo J* 17, 1192–1199.
- Dauber, B., Martínez-Sobrido, L., Schneider, J., Hai, R., Waibler, Z., Kalinke, U., García-Sastre, A., and Wolff, T. (2009). Influenza B Virus Ribonucleoprotein Is a Potent Activator of the Antiviral Kinase PKR. *PLoS Pathog* 5, e1000473.
- Daugherty, M.D., Schaller, A.M., Geballe, A.P., and Malik, H.S. (2016). Evolution-guided functional analyses reveal diverse antiviral specificities encoded by IFIT1 genes in mammals. *Elife* 5, 60.
- De Gregorio, E., Preiss, T., and Hentze, M.W. (1998). Translational activation of uncapped mRNAs by the central part of human eIF4G is 5' end-dependent. *Rna* 4, 828–836.

de Veer, M.J., Holko, M., Frevel, M., Walker, E., Der, S., Paranjape, J.M., Silverman, R.H., and Williams, B.R. (2001). Functional classification of interferon-stimulated genes identified using microarrays. *J. Leukoc. Biol.* *69*, 912–920.

de Veer, M.J., Sim, H., Whisstock, J.C., Devenish, R.J., and Ralph, S.J. (1998). IFI60/ISG60/IFIT4, a New Member of the Human IFI54/IFIT2 Family of Interferon-Stimulated Genes. *Genomics* *54*, 267–277.

Decroly, E., Debarnot, C., Ferron, F., Bouvet, M., Coutard, B., Imbert, I., Gluais, L., Papageorgiou, N., Sharff, A., Bricogne, G., et al. (2011a). Crystal Structure and Functional Analysis of the SARS-Coronavirus RNA Cap 2'-O-Methyltransferase nsp10/nsp16 Complex. *PLoS Pathog* *7*, e1002059.

Decroly, E., Ferron, F., Lescar, J., and Canard, B. (2011b). Conventional and unconventional mechanisms for capping viral mRNA. *Nat Rev Microbiol* *10*, 51–65.

Der, S.D., Zhou, A., Williams, B.R., and Silverman, R.H. (1998). Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays. *Proc Natl Acad Sci USA* *95*, 15623–15628.

Desrosiers, R.C., Sen, G.C., and Lengyel, P. (1976). Difference in 5' terminal structure between the mRNA and the double-stranded virion RNA of reovirus. *Biochem Biophys Res Commun* *73*, 32–39.

Devarkar, S.C., Wang, C., Miller, M.T., Ramanathan, A., Jiang, F., Khan, A.G., Patel, S.S., and Marcotrigiano, J. (2016). Structural basis for m7G recognition and 2'-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. *Proceedings of the National Academy of Sciences* *113*, 596–601.

Diamond, M.S., and Farzan, M. (2013). The broad-spectrum antiviral functions of IFIT and IFITM proteins. *Nat Rev Immunol* *13*, 46–57.

Dias, A., Bouvier, D., Crépin, T., McCarthy, A.A., Hart, D.J., Baudin, F., Cusack, S., and Ruigrok, R.W.H. (2009). The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* *458*, 914–918.

Dittmann, J., Stertz, S., Grimm, D., Steel, J., Garcia-Sastre, A., Haller, O., and Kochs, G. (2008). Influenza A Virus Strains Differ in Sensitivity to the Antiviral Action of Mx-GTPase. *J Virol* *82*, 3624–3631.

Dong, H., Fink, K., Züst, R., Lim, S.P., Qin, C.F., and Shi, P.Y. (2014). Flavivirus RNA methylation. *J. Gen. Virol.* *95*, 763–778.

Dong, H., Chang, D.C., Xie, X., Toh, Y.-X., Chung, K.Y., Zou, G., Lescar, J., Lim, S.P., and Shi, P.-Y. (2010). Biochemical and genetic characterization of dengue virus methyltransferase. *Virology* *405*, 568–578.

Dong, H., Ray, D., Ren, S., Zhang, B., Puig-Basagoiti, F., Takagi, Y., Ho, C.K., Li, H., and Shi,

- P.-Y. (2007). Distinct RNA elements confer specificity to flavivirus RNA cap methylation events. *J Virol* *81*, 4412–4421.
- Dougherty, D.A. (2013). The Cation– π Interaction. *Acc. Chem. Res.* *46*, 885–893.
- Dubin, D.T., Stollar, V., Hsueh, C.-C., Timko, K., and Guild, G.M. (1977). Sindbis virus messenger RNA: the 5'-termini and methylated residues of 26 and 42 S RNA. *Virology* *77*, 457–470.
- Egloff, M.-P., Decroly, E., Malet, H., Selisko, B., Benarroch, D., Ferron, F., and Canard, B. (2007). Structural and Functional Analysis of Methylation and 5'-RNA Sequence Requirements of Short Capped RNAs by the Methyltransferase Domain of Dengue Virus NS5. *J Mol Biol* *372*, 723–736.
- Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* *411*, 494–498.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* *66*, 486–501.
- Engelhardt, O.G., Smith, M., and Fodor, E. (2005). Association of the influenza A virus RNA-dependent RNA polymerase with cellular RNA polymerase II. *J Virol* *79*, 5812–5818.
- Fechter, P., and Brownlee, G.G. (2005). Recognition of mRNA cap structures by viral and cellular proteins. *J. Gen. Virol.* *86*, 1239–1249.
- Feng, F., Yuan, L., Wang, Y.E., Crowley, C., Lv, Z., Li, J., Liu, Y., Cheng, G., Zeng, S., and Liang, H. (2013). Crystal structure and nucleotide selectivity of human IFIT5/ISG58. *Cell Res* *23*, 1055–1058.
- Feng, X., Wang, Y., Ma, Z., Yang, R., Liang, S., Zhang, M., Song, S., Li, S., Liu, G., Fan, D., et al. (2014). MicroRNA-645, up-regulated in human adenocarcinoma of gastric esophageal junction, inhibits apoptosis by targeting tumor suppressor IFIT2. *BMC Cancer* *14*, 633.
- Fensterl, V., and Sen, G.C. (2011). The ISG56/IFIT1 gene family. *Journal of Interferon & Cytokine Research* *31*, 71–78.
- Fensterl, V., and Sen, G.C. (2015). Interferon-induced Ifit proteins: their role in viral pathogenesis. *J Virol* *89*, 2462–2468.
- Fensterl, V., Chattopadhyay, S., and Sen, G.C. (2015). No Love Lost Between Viruses and Interferons. *Annual Review of Virology* *2*, 549–572.
- Fensterl, V., White, C.L., Yamashita, M., and Sen, G.C. (2008). Novel characteristics of the function and induction of murine p56 family proteins. *J Virol* *82*, 11045–11053.
- Flint, S.J., Enquist, L.W., Racaniello, V.R., and Skalka, A.A.M. *Principles of Virology*, 3rd

Edition, Volume I (ASM Press).

Fried, M.G., and Liu, G. (1994). Molecular sequestration stabilizes CAP–DNA complexes during polyacrylamide gel electrophoresis. *Nucleic Acids Res* 22, 5054–5059.

Furuichi, Y., and Shatkin, A.J. (2000). Viral and cellular mRNA capping: past and prospects. *Adv Virus Res* 55, 135–184.

Furuichi, Y., Morgan, M., Shatkin, A.J., Jelinek, W., Salditt-Georgieff, M., and Darnell, J.E. (1975). Methylated, blocked 5' termini in HeLa cell mRNA. *Proc Natl Acad Sci USA* 72, 1904–1908.

Furuichi, Y., Muthukrishnan, S., Tomasz, J., and Shatkin, A.J. (1976). Mechanism of formation of reovirus mRNA 5'-terminal blocked and methylated sequence, m7GpppGmpC. *Journal of Biological Chemistry* 251, 5043–5053.

Geiss, G., Jin, G., Guo, J., Bumgarner, R., Katze, M.G., and Sen, G.C. (2001). A comprehensive view of regulation of gene expression by double-stranded RNA-mediated cell signaling. *Journal of Biological Chemistry* 276, 30178–30182.

Gerlach, P., Malet, H., Cusack, S., and Reguera, J. (2015). Structural Insights into Bunyavirus Replication and Its Regulation by the vRNA Promoter. *Cell* 161, 1267–1279.

Goodfellow, I. (2011). The genome-linked protein VPg of vertebrate viruses - a multifaceted protein. *Curr Opin Virol* 1, 355–362.

Goubau, D., Schlee, M., Deddouche, S., Pruijssers, A.J., Zillinger, T., Goldeck, M., Schuberth, C., Van der Veen, A.G., Fujimura, T., Rehwinkel, J., et al. (2014). Antiviral immunity via RIG-I-mediated recognition of RNA bearing 5'-diphosphates. *Nature* 514, 372–375.

Goulet, M.-L., Olagnier, D., Xu, Z., Paz, S., Belgnaoui, S.M., Lafferty, E.I., Janelle, V., Arguello, M., Paquet, M., Ghneim, K., et al. (2013). Systems Analysis of a RIG-I Agonist Inducing Broad Spectrum Inhibition of Virus Infectivity. *PLoS Pathog* 9, e1003298.

Grandvaux, N., Servant, M.J., tenOever, B., Sen, G.C., Balachandran, S., Barber, G.N., Lin, R., and Hiscott, J. (2002). Transcriptional Profiling of Interferon Regulatory Factor 3 Target Genes: Direct Involvement in the Regulation of Interferon-Stimulated Genes. *J Virol* 76, 5532–5539.

Grinnell, B.W., and Wagner, R.R. (1984). Nucleotide sequence and secondary structure of VSV leader RNA and homologous DNA involved in inhibition of DNA-dependent transcription. *Cell* 36, 533–543.

Grochulski, P., Fodje, M.N., Gorin, J., Labiuk, S.L., Berg, R., IUCr (2011). Beamline 08ID-1, the prime beamline of the Canadian Macromolecular Crystallography Facility. *J Synchrotron Radiat* 18, 681–684.

Gu, W., Gallagher, G.R., Dai, W., Liu, P., Li, R., Trombly, M.I., Gammon, D.B., Mello, C.C., Wang, J.P., and Finberg, R.W. (2015). Influenza A virus preferentially snatches noncoding RNA

caps. *Rna* 21, 2067–2075.

Guilligay, D., Tarendeau, F., Resa-Infante, P., Coloma, R., Crépin, T., Sehr, P., Lewis, J., Ruigrok, R.W.H., Ortín, J., Hart, D.J., et al. (2008). The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat Struct Mol Biol* 15, 500–506.

Guo, J., and Sen, G.C. (2000). Characterization of the interaction between the interferon-induced protein P56 and the Int6 protein encoded by a locus of insertion of the mouse mammary tumor virus. *J Virol* 74, 1892–1899.

Guo, J., Hui, D.J., Merrick, W.C., and Sen, G.C. (2000a). A new pathway of translational regulation mediated by eukaryotic initiation factor 3. *Embo J* 19, 6891–6899.

Guo, J., Peters, K.L., and Sen, G.C. (2000b). Induction of the Human Protein P56 by Interferon, Double-Stranded RNA, or Virus Infection. *Virology* 267, 209–219.

Habjan, M., Andersson, I., Klingström, J., Schümann, M., Martin, A., Zimmermann, P., Wagner, V., Pichlmair, A., Schneider, U., Mühlberger, E., et al. (2008). Processing of genome 5' termini as a strategy of negative-strand RNA viruses to avoid RIG-I-dependent interferon induction. *PLoS ONE* 3, e2032.

Habjan, M., Hubel, P., Lacerda, L., Benda, C., Holze, C., Eberl, C.H., Mann, A., Kindler, E., Gil-Cruz, C., Ziebuhr, J., et al. (2013). Sequestration by IFIT1 Impairs Translation of 2'O-unnmethylated Capped RNA. *PLoS Pathog* 9, e1003663.

Hagler, J., and Shuman, S. (1992). A freeze-frame view of eukaryotic transcription during elongation and capping of nascent mRNA. *Science* 255, 983–986.

Haline-Vaz, T., Lima Silva, T.C., and Zanchin, N.I.T. (2008). The human interferon-regulated ISG95 protein interacts with RNA polymerase II and shows methyltransferase activity. *Biochem Biophys Res Commun* 372, 719–724.

Hall, M.P., and Ho, C.K. (2006). Functional characterization of a 48 kDa *Trypanosoma brucei* cap 2 RNA methyltransferase. *Nucleic Acids Res* 34, 5594–5602.

Harding, M.M., and Nowicki, M.W. (2010). Metals in protein structures: a review of their principal features. *Crystallography Reviews* 16, 247–302.

Hashimoto, S.I., and Green, M. (1976). Multiple methylated cap sequences in adenovirus type 2 early mRNA. *J Virol* 20, 425–435.

Hellman, L.M., and Fried, M.G. (2007). Electrophoretic mobility shift assay (EMSA) for detecting protein–nucleic acid interactions. *Nat Protoc* 2, 1849–1861.

Hodel, A.E., Gershon, P.D., and Quioco, F.A. (1998). Structural basis for sequence-nonspecific recognition of 5'-capped mRNA by a cap-modifying enzyme. *Mol Cell* 1, 443–447.

Hodel, A.E., Gershon, P.D., Shi, X., and Quioco, F.A. (1996). The 1.85 Å structure of vaccinia

protein VP39: a bifunctional enzyme that participates in the modification of both mRNA ends. *Cell* 85, 247–256.

Hodel, A.E., Gershon, P.D., Shi, X., Wang, S.M., and Quioco, F.A. (1997). Specific protein recognition of an mRNA cap through its alkylated base. *Nat Struct Biol* 4, 350–354.

Hornung, V., Ellegast, J., Kim, S., Brzózka, K., Jung, A., Kato, H., Poeck, H., Akira, S., Conzelmann, K.-K., Schlee, M., et al. (2006). 5'-Triphosphate RNA is the ligand for RIG-I. *Science* 314, 994–997.

Hsu, Y.-A., Lin, H.-J., Sheu, J.J.C., Shieh, F.-K., Chen, S.-Y., Lai, C.-H., Tsai, F.-J., Wan, L., and Chen, B.-H. (2011). A Novel Interaction Between Interferon-Inducible Protein p56 and Ribosomal Protein L15 in Gastric Cancer Cells. *DNA and Cell Biology* 30, 671–679.

Hsueh, C.-C., and Dubin, D.T. (1976). Di- and trimethylated congeners of 7-methylguanine in Sindbis virus mRNA. *Nature* 264, 190–191.

Hu, G., Tsai, A.L., and Quioco, F.A. (2003). Insertion of an N7-methylguanine mRNA Cap between Two Coplanar Aromatic Residues of a Cap-binding Protein Is Fast and Selective for a Positively Charged Cap. *Journal of Biological Chemistry* 278, 51515–51520.

Huang, X., Shen, N., Bao, C., Gu, Y., Wu, L., and Chen, S. (2008). Interferon-induced protein IFIT4 is associated with systemic lupus erythematosus and promotes differentiation of monocytes into dendritic cell-like cells. *Arthritis Res. Ther.* 10, R91.

Hui, D.J., Bhasker, C.R., Merrick, W.C., and Sen, G.C. (2003). Viral stress-inducible protein p56 inhibits translation by blocking the interaction of eIF3 with the ternary complex eIF2.GTP.Met-tRNAi. *Journal of Biological Chemistry* 278, 39477–39482.

Hui, D.J., Terenzi, F., Merrick, W.C., and Sen, G.C. (2005). Mouse p56 blocks a distinct function of eukaryotic initiation factor 3 in translation initiation. *Journal of Biological Chemistry* 280, 3433–3440.

Hull, C.M., and Bevilacqua, P.C. (2016). Discriminating Self and Non-Self by RNA: Roles for RNA Structure, Misfolding, and Modification in Regulating the Innate Immune Sensor PKR. *Acc. Chem. Res.* 49, 1242–1249.

Hulo, C., de Castro, E., Masson, P., Bougueleret, L., Bairoch, A., Xenarios, I., and Le Mercier, P. (2010). ViralZone: a knowledge resource to understand virus diversity. *Nucleic Acids Res* 39, D576–D582.

Hyde, J.L., and Diamond, M.S. (2015). Innate immune restriction and antagonism of viral RNA lacking 2'-O methylation. *Virology* 479-480, 66–74.

Hyde, J.L., Gardner, C.L., Kimura, T., White, J.P., Liu, G., Trobaugh, D.W., Huang, C., Tonelli, M., Paessler, S., Takeda, K., et al. (2014). A viral RNA structural element alters host recognition of nonself RNA. *Science* 343, 783–787.

- Ingolia, N.T. (2016). Ribosome Footprint Profiling of Translation throughout the Genome. *Cell* *165*, 22–33.
- Isaacs, A., and Lindenmann, J. (2015). Pillars Article: Virus Interference. I. The Interferon. *Proc R Soc Lond B Biol Sci.* 1957. *147*: 258-267. (American Association of Immunologists).
- Issur, M., Geiss, B.J., Bougie, I., Picard-Jean, F., Despins, S., Mayette, J., Hobdey, S.E., and Bisaiillon, M. (2009). The flavivirus NS5 protein is a true RNA guanylyltransferase that catalyzes a two-step reaction to form the RNA cap structure. *Rna* *15*, 2340–2350.
- Ivanov, K.A., Thiel, V., Dobbe, J.C., van der Meer, Y., Snijder, E.J., and Ziebuhr, J. (2004). Multiple Enzymatic Activities Associated with Severe Acute Respiratory Syndrome Coronavirus Helicase. *J Virol* *78*, 5619–5632.
- Jiang, F., Ramanathan, A., Miller, M.T., Tang, G.-Q., Gale, M., Patel, S.S., and Marcotrigiano, J. (2011). Structural basis of RNA recognition and activation by innate immune receptor RIG-I. *Nature* *479*, 423–427.
- Jiao, X., Chang, J.H., Kilic, T., Tong, L., and Kiledjian, M. (2013). A Mammalian Pre-mRNA 5' End Capping Quality Control Mechanism and an Unexpected Link of Capping to Pre-mRNA Processing. *Mol Cell* *50*, 104–115.
- Jiao, X., Xiang, S., Oh, C., Martin, C.E., Tong, L., and Kiledjian, M. (2010). Identification of a quality-control mechanism for mRNA 5'-end capping. *Nature* *467*, 608–611.
- Jínek, M., Rehwinkel, J., Lazarus, B.D., Izaurrealde, E., Hanover, J.A., and Conti, E. (2004). The superhelical TPR-repeat domain of O-linked GlcNAc transferase exhibits structural similarities to importin α . *Nat Struct Mol Biol* *11*, 1001–1007.
- Kapadia, S.B., Brideau-Andersen, A., and Chisari, F.V. (2003). Interference of hepatitis C virus RNA replication by short interfering RNAs. *Proc Natl Acad Sci USA* *100*, 2014–2018.
- Karikó, K., Bhuyan, P., Capodici, J., and Weissman, D. (2004). Small interfering RNAs mediate sequence-independent gene suppression and induce immune activation by signaling through toll-like receptor 3. *J Immunol* *172*, 6545–6549.
- Karki, S., Li, M.M.H., Schoggins, J.W., Tian, S., Rice, C.M., and MacDonald, M.R. (2012). Multiple Interferon Stimulated Genes Synergize with the Zinc Finger Antiviral Protein to Mediate Anti-Alphavirus Activity. *PLoS ONE* *7*, e37398.
- Karpenahalli, M.R., Lupas, A.N., and Söding, J. (2007). TPRpred: a tool for prediction of TPR-, PPR- and SEL1-like repeats from protein sequences. *BMC Bioinformatics* *8*, 2.
- Katibah, G.E., Qin, Y., Sidote, D.J., Yao, J., Lambowitz, A.M., and Collins, K. (2014). Broad and adaptable RNA structure recognition by the human interferon-induced tetratricopeptide repeat protein IFIT5. *Proceedings of the National Academy of Sciences* *111*, 201412842–12030.
- Katibah, G.E., Lee, H.J., Huizar, J.P., Vogan, J.M., Alber, T., and Collins, K. (2013). tRNA

binding, structure, and localization of the human interferon-induced protein IFIT5. *Mol Cell* 49, 743–750.

Kato, H., Takeuchi, O., Mikamo-Satoh, E., Hirai, R., Kawai, T., Matsushita, K., Hiiragi, A., Dermody, T.S., Fujita, T., and Akira, S. (2008). Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5. *J Exp Med* 205, 1601–1610.

Kato, H., Takeuchi, O., Sato, S., Yoneyama, M., Yamamoto, M., Matsui, K., Uematsu, S., Jung, A., Kawai, T., Ishii, K.J., et al. (2006). Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* 441, 101–105.

Keith, J.M., Ensinger, M.J., and Moss, B. (1978a). HeLa cell RNA (2'-O-methyladenosine-N6)-methyltransferase specific for the capped 5'-end of messenger RNA. *Journal of Biological Chemistry* 253, 5033–5039.

Keith, J.M., Muthukrishnan, S., and Moss, B. (1978b). Effect of methylation of the N6 position of the penultimate adenosine of capped mRNA on ribosome binding. *Journal of Biological Chemistry* 253, 5039–5041.

Kell, A., Stoddard, M., Li, H., Marcotrigiano, J., Shaw, G.M., and Gale, M., Jr (2015). Pathogen-Associated Molecular Pattern Recognition of Hepatitis C Virus Transmitted/Founder Variants by RIG-I Is Dependent on U-Core Length. *J Virol* 89, 11056–11068.

Kim, D.-H., Longo, M., Han, Y., Lundberg, P., Cantin, E., and Rossi, J.J. (2004). Interferon induction by siRNAs and ssRNAs synthesized by phage polymerase. *Nat Biotechnol* 22, 321–325.

Kimura, T., Katoh, H., Kayama, H., Saiga, H., Okuyama, M., Okamoto, T., Umemoto, E., Matsuura, Y., Yamamoto, M., and Takeda, K. (2013). Ifit1 inhibits JEV replication through binding to 5' capped 2'-O unmethylated RNA. *J Virol* 87, 9997–10003.

Klema, V., Padmanabhan, R., and Choi, K. (2015). Flaviviral Replication Complex: Coordination between RNA Synthesis and 5'-RNA Capping. *Viruses* 7, 4640–4656.

Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* 36, 1277–1282.

Kowalinski, E., Lunardi, T., McCarthy, A.A., Louber, J., Brunel, J., Grigorov, B., Gerlier, D., and Cusack, S. (2011). Structural Basis for the Activation of Innate Immune Pattern-Recognition Receptor RIG-I by Viral RNA. *Cell* 147, 423–435.

Kranzusch, P.J., Schenk, A.D., Rahmeh, A.A., Radoshitzky, S.R., Bavari, S., Walz, T., and Whelan, S.P.J. (2010). Assembly of a functional Machupo virus polymerase complex. *Proceedings of the National Academy of Sciences* 107, 20069–20074.

Krug, R.M., Morgan, M.A., and Shatkin, A.J. (1976). Influenza viral mRNA contains internal

- N6-methyladenosine and 5'-terminal 7-methylguanosine in cap structures. *J Virol* 20, 45–53.
- Kuge, H., Brownlee, G.G., Gershon, P.D., and Richter, J.D. (1998). Cap ribose methylation of c-mos mRNA stimulates translation and oocyte maturation in *Xenopus laevis*. *Nucleic Acids Res* 26, 3208–3214.
- Kuhn, A.N., Diken, M., Kreiter, S., Selmi, A., Kowalska, J., Jemielity, J., Darzynkiewicz, E., Huber, C., Türeci, Ö., and Sahin, U. (2010). Phosphorothioate cap analogs increase stability and translational efficiency of RNA vaccines in immature dendritic cells and induce superior immune responses in vivo. *Gene Ther* 17, 961–971.
- Kumar, P., Sweeney, T.R., Skabkin, M.A., Skabkina, O.V., Hellen, C.U.T., and Pestova, T.V. (2014). Inhibition of translation by IFIT family members is determined by their ability to interact selectively with the 5'-terminal regions of cap0-, cap1- and 5"ppp- mRNAs. *Nucleic Acids Res* 42, 3228–3245.
- Kusari, J., and Sen, G.C. (1986). Regulation of synthesis and turnover of an interferon-inducible mRNA. *Mol Cell Biol* 6, 2062–2067.
- Kyrieleis, O.J.P., Chang, J., la Peña, de, M., Shuman, S., and Cusack, S. (2014). Crystal structure of vaccinia virus mRNA capping enzyme provides insights into the mechanism and evolution of the capping apparatus. *Structure* 22, 452–465.
- Lai, K.C., Liu, C.J., Chang, K.W., and Lee, T.C. (2013). Depleting IFIT2 mediates atypical PKC signaling to enhance the migration and metastatic activity of oral squamous cell carcinoma cells. *Oncogene* 32, 3686–3697.
- Lai, K.-C., Chang, K.-W., Liu, C.-J., Kao, S.-Y., and Lee, T.-C. (2008). IFN-induced protein with tetratricopeptide repeats 2 inhibits migration activity and increases survival of oral squamous cell carcinoma. *Mol. Cancer Res.* 6, 1431–1439.
- Lai, M.M., Patton, C.D., and Stohlman, S.A. (1982). Further characterization of mRNA"s of mouse hepatitis virus: presence of common 5"-end nucleotides. *J Virol* 41, 557–565.
- Lamb, J.R., Tugendreich, S., and Hieter, P. (1995). Tetratricopeptide repeat interactions: to TPR or not to TPR? *Trends Biochem Sci* 20, 257–259.
- Langberg, S.R., and Moss, B. (1981). Post-transcriptional modifications of mRNA. Purification and characterization of cap I and cap II RNA (nucleoside-2'-)-methyltransferases from HeLa cells. *Journal of Biological Chemistry* 256, 10054–10060.
- Langer, G., Cohen, S.X., Lamzin, V.S., and Perrakis, A. (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* 3, 1171–1179.
- Lässig, C., Matheisl, S., Sparrer, K.M., de Oliveira Mann, C.C., Moldt, M., Patel, J.R., Goldeck, M., Hartmann, G., García-Sastre, A., Hornung, V., et al. (2015). ATP hydrolysis by the viral RNA sensor RIG-I prevents unintentional recognition of self-RNA. *Elife* 4, 1–33.

- Lee, C.G., Demarquoy, J., Jackson, M.J., and O'Brien, W.E. (1994). Molecular cloning and characterization of a murine LPS-inducible cDNA. *J Immunol* *152*, 5758–5767.
- Lehmann, M., Pahlmann, M., Jérôme, H., Busch, C., Lelke, M., and Günther, S. (2014). Role of the C terminus of Lassa virus L protein in viral mRNA synthesis. *J Virol* *88*, 8713–8717.
- Levy, D., Larner, A., Chaudhuri, A., Babiss, L.E., and Darnell, J.E. (1986). Interferon-stimulated transcription: isolation of an inducible gene and identification of its regulatory region. *Proc Natl Acad Sci USA* *83*, 8929–8933.
- Li, J., Wang, J.T., and Whelan, S.P.J. (2006). A unique strategy for mRNA cap methylation used by vesicular stomatitis virus. *Proc Natl Acad Sci USA* *103*, 8493–8498.
- Li, M.M.H., MacDonald, M.R., and Rice, C.M. (2015). To translate, or not to translate: viral and host mRNA regulation by interferon-stimulated genes. *Trends in Cell Biology* *25*, 320–329.
- Li, S.H., Dong, H., Li, X.F., Xie, X., Zhao, H., Deng, Y.Q., Wang, X.Y., Ye, Q., Zhu, S.Y., Wang, H.J., et al. (2013). Rational Design of a Flavivirus Vaccine by Abolishing Viral RNA 2'-O Methylation. *J Virol* *87*, 5812–5819.
- Li, Y., Li, C., Xue, P., Zhong, B., Mao, A.-P., Ran, Y., Chen, H., Wang, Y.-Y., Yang, F., and Shu, H.-B. (2009). ISG56 is a negative-feedback regulator of virus-triggered signaling and cellular antiviral response. *Proceedings of the National Academy of Sciences* *106*, 7945–7950.
- Liang, B., Li, Z., Jenni, S., Rahmeh, A.A., Morin, B.M., Grant, T., Grigorieff, N., Harrison, S.C., and Whelan, S.P.J. (2015). Structure of the L Protein of Vesicular Stomatitis Virus from Electron Cryomicroscopy. *Cell* *162*, 314–327.
- Lidschreiber, M., Leike, K., and Cramer, P. (2013). Cap completion and C-terminal repeat domain kinase recruitment underlie the initiation-elongation transition of RNA polymerase II. *Mol Cell Biol* *33*, 3805–3816.
- Lin, J.-Y., Chen, T.-C., Weng, K.-F., Chang, S.-C., Chen, L.-L., and Shih, S.-R. (2009). Viral and host proteins involved in picornavirus life cycle. *Journal of Biomedical Science* *2009* *16*:1 *16*, 103.
- Liu, L., Botos, I., Wang, Y., Leonard, J.N., Shiloach, J., Segal, D.M., and Davies, D.R. (2008). Structural Basis of Toll-Like Receptor 3 Signaling with Double-Stranded RNA. *Science* *320*, 379–381.
- Liu, X.Y., Chen, W., Wei, B., Shan, Y.F., and Wang, C. (2011). IFN-Induced TPR Protein IFIT3 Potentiates Antiviral Signaling by Bridging MAVS and TBK1. *The Journal of Immunology* *187*, 2559–2568.
- Liu, Y., Zhang, Y.-B., Liu, T.-K., and Gui, J.-F. (2013). Lineage-Specific Expansion of IFIT Gene Family: An Insight into Coevolution with IFN Gene Family. *PLoS ONE* *8*, e66859.
- Lu, C., Xu, H., Ranjith-Kumar, C.T., Brooks, M.T., Hou, T.Y., Hu, F., Herr, A.B., Strong, R.K.,

- Kao, C.C., and Li, P. (2010). The Structural Basis of 5' Triphosphate Double-Stranded RNA Recognition by RIG-I C-Terminal Domain. *Structure* *18*, 1032–1043.
- Luecke, S., and Paludan, S.R. (2016). Molecular requirements for sensing of intracellular microbial nucleic acids by the innate immune system. *Cytokine*.
- Luo, D., Ding, S.C., Vela, A., Kohlway, A., Lindenbach, B.D., and Pyle, A.M. (2011). Structural Insights into RNA Recognition by RIG-I. *Cell* *147*, 409–422.
- Luo, D., Kohlway, A., Vela, A., and Pyle, A.M. (2012). Visualizing the Determinants of Viral RNA Recognition by Innate Immune Sensor RIG-I. *Structure* *20*, 1983–1988.
- Main, E.R.G., Stott, K., Jackson, S.E., and Regan, L. (2005). Local and long-range stability in tandemly arrayed tetratricopeptide repeats. *Proc Natl Acad Sci USA* *102*, 5721–5726.
- Main, E.R.G., Xiong, Y., Cocco, M.J., D'Andrea, L., and Regan, L. (2003). Design of stable alpha-helical arrays from an idealized TPR motif. *Structure* *11*, 497–508.
- Mankan, A.K., Schmidt, T., Chauhan, D., Goldeck, M., Honing, K., Gaidt, M., Kubarenko, A.V., Andreeva, L., Hopfner, K.P., and Hornung, V. (2014). Cytosolic RNA:DNA hybrids activate the cGAS-STING axis. *Embo J* *33*, 2937–2946.
- Maquat, L.E., Tarn, W.-Y., and Isken, O. (2010). The pioneer round of translation: features and functions. *Cell* *142*, 368–374.
- Marcotrigiano, J., Gingras, A.C., Sonenberg, N., and Burley, S.K. (1997). Cocystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell* *89*, 951–961.
- Marq, J.-B., Kolakofsky, D., and Garcin, D. (2010). Unpaired 5' ppp-nucleotides, as found in arenavirus double-stranded RNA panhandles, are not recognized by RIG-I. *J. Biol. Chem.* *285*, 18208–18216.
- Marques, J.T., Devosse, T., Wang, D., Zamanian-Daryoush, M., Serbinowski, P., Hartmann, R., Fujita, T., Behlke, M.A., and Williams, B.R.G. (2006). A structural basis for discriminating between self and nonself double-stranded RNAs in mammalian cells. *Nat Biotechnol* *24*, 559–565.
- Martinez-Rucobo, F.W., Kohler, R., van de Waterbeemd, M., Heck, A.J.R., Hemann, M., Herzog, F., Stark, H., and Cramer, P. (2015). Molecular Basis of Transcription-Coupled Pre-mRNA Capping. *Mol Cell* *58*, 1079–1089.
- Matsumoto, M., Funami, K., Tanabe, M., Oshiumi, H., Shingai, M., Seto, Y., Yamamoto, A., and Seya, T. (2003). Subcellular localization of Toll-like receptor 3 in human dendritic cells. *J Immunol* *171*, 3154–3162.
- Mazza, C., Segref, A., Mattaj, I.W., and Cusack, S. (2002). Large-scale induced fit recognition of an m(7)GpppG cap analogue by the human nuclear cap-binding complex. *Embo J* *21*, 5548–5557.

- McCaffrey, A.P., Nakai, H., Pandey, K., Huang, Z., Salazar, F.H., Xu, H., Wieland, S.F., Marion, P.L., and Kay, M.A. (2003). Inhibition of hepatitis B virus in mice by RNA interference. *Nat Biotechnol* 21, 639–644.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., IUCr (2007). Phaser crystallographic software. *J Appl Crystallogr* 40, 658–674.
- McDermott, J.E., Vartanian, K.B., Mitchell, H., Stevens, S.L., Sanfilippo, A., and Stenzel-Poore, M.P. (2012). Identification and Validation of Ifit1 as an Important Innate Immune Bottleneck. *PLoS ONE* 7, e36465.
- McKenna, S.A., Kim, I., Puglisi, E.V., Lindhout, D.A., Aitken, C.E., Marshall, R.A., and Puglisi, J.D. (2007). Purification and characterization of transcribed RNAs using gel filtration chromatography. *Nat Protoc* 2, 3270–3277.
- Menachery, V.D., Yount, B.L., Josset, L., Gralinski, L.E., Scobey, T., Agnihothram, S., Katze, M.G., and Baric, R.S. (2014). Attenuation and restoration of severe acute respiratory syndrome coronavirus mutant lacking 2'-o-methyltransferase activity. *J Virol* 88, 4251–4264.
- Mir, M.A., Duran, W.A., Hjelle, B.L., Ye, C., and Panganiban, A.T. (2008). Storage of cellular 5' mRNA caps in P bodies for viral cap-snatching. *Proceedings of the National Academy of Sciences* 105, 19294–19299.
- Moeller, A., Kirchdoerfer, R.N., Potter, C.S., Carragher, B., and Wilson, I.A. (2012). Organization of the Influenza Virus Replication Machinery. *Science* 338, 1631–1634.
- Moll, H.P., Maier, T., Zommer, A., Lavoie, T., and Brostjan, C. (2011). The differential activity of interferon- β subtypes is consistent among distinct target genes and cell types. *Cytokine* 53, 52–59.
- Moriarty, N.W., Grosse-Kunstleve, R.W., and Adams, P.D. (2009). electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr D Biol Crystallogr* 65, 1074–1080.
- Moss, B., and Kocot, F. (1976). Sequence of methylated nucleotides at the 5'-terminus of adenovirus-specific RNA. *J Virol* 17, 385–392.
- Moss, B., Gershowitz, A., Stringer, J.R., Holland, L.E., and Wagner, E.K. (1977). 5'-Terminal and internal methylated nucleosides in herpes simplex virus type 1 mRNA. *J Virol* 23, 234–239.
- Mossessova, E., and Lima, C.D. (2000). Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol Cell* 5, 865–876.
- Muller, U., Steinhoff, U., Reis, L.F., Hemmi, S., Pavlovic, J., Zinkernagel, R.M., and Aguet, M. (1994). Functional role of type I and type II interferons in antiviral defense. *Science* 264, 1918–1921.

- Muthukrishnan, S., Morgan, M., Banerjee, A.K., and Shatkin, A.J. (1976). Influence of 5'-terminal m7G and 2'-O-methylated residues on messenger ribonucleic acid binding to ribosomes. *Biochemistry* *15*, 5761–5768.
- Muthukrishnan, S., Moss, B., Cooper, J.A., and Maxwell, E.S. (1978). Influence of 5'-terminal cap structure on the initiation of translation of vaccinia virus mRNA. *Journal of Biological Chemistry* *253*, 1710–1715.
- Myong, S., Cui, S., Cornish, P.V., Kirchhofer, A., Gack, M.U., Jung, J.U., Hopfner, K.P., and Ha, T. (2009). Cytosolic Viral Sensor RIG-I Is a 5'-Triphosphate-Dependent Translocase on Double-Stranded RNA. *Science* *323*, 1070–1074.
- Nallagatla, S.R., Hwang, J., Toroney, R., Zheng, X., Cameron, C.E., and Bevilacqua, P.C. (2007). 5'-Triphosphate-Dependent Activation of PKR by RNAs with Short Stem-Loops. *Science* *318*, 1455–1458.
- Netherton, C.L., and Wileman, T. (2011). Virus factories, double membrane vesicles and viroplasm generated in animal cells. *Curr Opin Virol* *1*, 381–387.
- Niikura, T., Hirata, R., and Weil, S.C. (1997). A Novel Interferon-Inducible Gene Expressed during Myeloid Differentiation. *Blood Cells, Molecules, and Diseases* *23*, 337–349.
- Nomoto, A., Kitamura, N., Golini, F., and Wimmer, E. (1977). The 5'-terminal structures of poliovirion RNA and poliovirus mRNA differ only in the genome-linked protein VPg. *Proc Natl Acad Sci USA* *74*, 5345–5349.
- Novac, O., Guenier, A.S., and Pelletier, J. (2004). Inhibitors of protein synthesis identified by a high throughput multiplexed translation screen. *Nucleic Acids Res* *32*, 902–915.
- Ogino, T. (2014). Capping of vesicular stomatitis virus pre-mRNA is required for accurate selection of transcription stop-start sites and virus propagation. *Nucleic Acids Res* *42*, 12112–12125.
- Ogino, T., and Banerjee, A.K. (2007). Unconventional Mechanism of mRNA Capping by the RNA-Dependent RNA Polymerase of Vesicular Stomatitis Virus. *Mol Cell* *25*, 85–97.
- Ohto, U., Shibata, T., Tanji, H., Ishida, H., Krayukhina, E., Uchiyama, S., Miyake, K., and Shimizu, T. (2015). Structural basis of CpG and inhibitory DNA recognition by Toll-like receptor 9. *Nature* *520*, 702–705.
- Ovstebø, R., Olstad, O.K., Brusletto, B., Møller, A.S., Aase, A., Haug, K.B.F., Brandtzaeg, P., and Kierulf, P. (2008). Identification of genes particularly sensitive to lipopolysaccharide (LPS) in human monocytes induced by wild-type versus LPS-deficient *Neisseria meningitidis* strains. *Infect. Immun.* *76*, 2685–2695.
- Paesen, G.C., Collet, A., Sallamand, C., Debart, F., Vasseur, J.-J., Canard, B., Decroly, E., and Grimes, J.M. (2015). X-ray structure and activities of an essential Mononegavirales L-protein domain. *Nat Commun* *6*, 8749.

- Patzwahl, R., Ramadori, G., and Mihm, S. (2001). Enhanced expression of interferon-regulated genes in the liver of patients with chronic hepatitis C virus infection: detection by suppression-subtractive hybridization. *J Virol* 75, 1332–1338.
- Pavelka, A., Sebestova, E., Kozlikova, B., Brezovsky, J., Sochor, J., and Damborsky, J. (2016). CAVER: Algorithms for Analyzing Dynamics of Tunnels in Macromolecules. *IEEE/ACM Trans Comput Biol Bioinform* 13, 505–517.
- Peisley, A., Wu, B., Yao, H., Walz, T., and Hur, S. (2013). RIG-I Forms Signaling-Competent Filaments in an ATP-Dependent, Ubiquitin-Independent Manner. *Mol Cell* 51, 573–583.
- Pellagatti, A., Cazzola, M., Giagounidis, A.A.N., Malcovati, L., Porta, M.G.D., Killick, S., Campbell, L.J., Wang, L., Langford, C.F., Fidler, C., et al. (2006). Gene expression profiles of CD34+ cells in myelodysplastic syndromes: involvement of interferon-stimulated genes and correlation to FAB subtype and karyotype. *Blood* 108, 337–345.
- Pelletier, J., and Sonenberg, N. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* 334, 320–325.
- Pelletier, J., Graff, J., Ruggero, D., and Sonenberg, N. (2015). Targeting the eIF4F translation initiation complex: a critical nexus for cancer development. *Cancer Res* 75, 250–263.
- Perry, R.P., and Kelley, D.E. (1976). Kinetics of formation of 5' terminal caps in mRNA. *Cell* 8, 433–442.
- Pichlmair, A., Schulz, O., Tan, C.P., Naslund, T.I., Liljestrom, P., Weber, F., and Reis e Sousa, C. (2006). RIG-I-Mediated Antiviral Responses to Single-Stranded RNA Bearing 5'-Phosphates. *Science* 314, 997–1001.
- Pichlmair, A., Schulz, O., Tan, C.P., Rehwinkel, J., Kato, H., Takeuchi, O., Akira, S., Way, M., Schiavo, G., and Reis e Sousa, C. (2009). Activation of MDA5 Requires Higher-Order RNA Structures Generated during Virus Infection. *J Virol* 83, 10761–10769.
- Pichlmair, A., Lassnig, C., Eberle, C.-A., Góna, M.W., Baumann, C.L., Burkard, T.R., Bürckstümmer, T., Stefanovic, A., Krieger, S., Bennett, K.L., et al. (2011). IFIT1 is an antiviral protein that recognizes 5'-triphosphate RNA. *Nat Immunol* 12, 624–630.
- Pinto, A.K., Williams, G.D., Szretter, K.J., White, J.P., Proença-Módena, J.L., Liu, G., Olejnik, J., Brien, J.D., Ebihara, H., Mühlberger, E., et al. (2015). Human and Murine IFIT1 Proteins Do Not Restrict Infection of Negative-Sense RNA Viruses of the Orthomyxoviridae, Bunyaviridae, and Filoviridae Families. *J Virol* 89, 9465–9476.
- Plotch, S.J., Bouloy, M., Ulmanen, I., and Krug, R.M. (1981). A unique cap(m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell* 23, 847–858.
- Plumet, S., Herschke, F., Bourhis, J.-M., Valentin, H., Longhi, S., and Gerlier, D. (2007). Cytosolic 5'-Triphosphate Ended Viral Leader Transcript of Measles Virus as Activator of the

RIG I-Mediated Interferon Response. *PLoS ONE* 2, e279.

Putnam, C.D., Hammel, M., Hura, G.L., and Tainer, J.A. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* 40, 191–285.

Quioco, F.A., Hu, G., and Gershon, P.D. (2000). Structural basis of mRNA cap recognition by proteins. *Curr Opin Struct Biol* 10, 78–86.

Rahmeh, A.A., Li, J., Kranzusch, P.J., and Whelan, S.P.J. (2009). Ribose 2'-O Methylation of the Vesicular Stomatitis Virus mRNA Cap Precedes and Facilitates Subsequent Guanine-N-7 Methylation by the Large Polymerase Protein. *J Virol* 83, 11043–11050.

Rambo, R.P., and Tainer, J.A. (2011). Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* 95, 559–571.

Ray, D., Shah, A., Tilgner, M., Guo, Y., Zhao, Y., Dong, H., Deas, T.S., Zhou, Y., Li, H., and Shi, P.-Y. (2006). West Nile virus 5'-cap structure is formed by sequential guanine N-7 and ribose 2'-O methylations by nonstructural protein 5. *J Virol* 80, 8362–8370.

Raychoudhuri, A., Shrivastava, S., Steele, R., Kim, H., Ray, R., and Ray, R.B. (2011). ISG56 and IFITM1 proteins inhibit hepatitis C virus replication. *J Virol* 85, 12881–12889.

Reguera, J., Cusack, S., and Kolakofsky, D. (2014). Segmented negative strand RNA virus nucleoprotein structure. *Curr Opin Virol* 5, 7–15.

Reguera, J., Gerlach, P., and Cusack, S. (2016a). Towards a structural understanding of RNA synthesis by negative strand RNA viral polymerases. *Curr Opin Struct Biol* 36, 75–84.

Reguera, J., Gerlach, P., Rosenthal, M., Gaudon, S., Coscia, F., Günther, S., and Cusack, S. (2016b). Comparative Structural and Functional Analysis of Bunyavirus and Arenavirus Cap-Snatching Endonucleases. *PLoS Pathog* 12, e1005636.

Reguera, J., Weber, F., and Cusack, S. (2010). Bunyaviridae RNA Polymerases (L-Protein) Have an N-Terminal, Influenza-Like Endonuclease Domain, Essential for Viral Cap-Dependent Transcription. *PLoS Pathog* 6, e1001101.

Rehwinkel, J., Tan, C.P., Goubau, D., Schulz, O., Pichlmair, A., Bier, K., Robb, N., Vreede, F., Barclay, W., Fodor, E., et al. (2010). RIG-I Detects Viral Genomic RNA during Negative-Strand RNA Virus Infection. *Cell* 140, 397–408.

Reich, S., Guilligay, D., Pflug, A., Malet, H., Berger, I., Crépin, T., Hart, D., Lunardi, T., Nanao, M., Ruigrok, R.W.H., et al. (2014). Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature* 516, 361–366.

Reinisch, K.M., Nibert, M.L., and Harrison, S.C. (2000). Structure of the reovirus core at 3.6 Å resolution. *Nature* 404, 960–967.

- Reynaud, J.M., Kim, D.Y., Atasheva, S., Rasaloukaya, A., White, J.P., Diamond, M.S., Weaver, S.C., Frolova, E.I., and Frolov, I. (2015). IFIT1 Differentially Interferes with Translation and Replication of Alphavirus Genomes and Promotes Induction of Type I Interferon. *PLoS Pathog* *11*, e1004863–31.
- Rigby, R.E., Webb, L.M., Mackenzie, K.J., Li, Y., Leitch, A., Reijns, M.A.M., Lundie, R.J., Revuelta, A., Davidson, D.J., Diebold, S., et al. (2014). RNA:DNA hybrids are a novel molecular pattern sensed by TLR9. *Embo J* *33*, 542–558.
- Robert, F., Kapp, L.D., Khan, S.N., Acker, M.G., Kolitz, S., Kazemi, S., Kaufman, R.J., Merrick, W.C., Koromilas, A.E., Lorsch, J.R., et al. (2006). Initiation of protein synthesis by hepatitis C virus is refractory to reduced eIF2.GTP.Met-tRNA(i)(Met) ternary complex availability. *Mol. Biol. Cell* *17*, 4632–4644.
- Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* *42*, W320–W324.
- Rose, J.K. (1975). Heterogeneous 5'-terminal structures occur on vesicular stomatitis virus mRNAs. *Journal of Biological Chemistry* *250*, 8098–8104.
- Saeedi, B.J., and Geiss, B.J. (2013). Regulation of flavivirus RNA synthesis and capping. *WIREs RNA* *4*, 723–735.
- Sahin, U., Karikó, K., and Türeci, Ö. (2014). mRNA-based therapeutics [mdash] developing a new class of drugs. *Nature Reviews Drug Discovery* *13*, 759–780.
- Sawicki, S.G., Sawicki, D.L., and Siddell, S.G. (2007). A contemporary view of coronavirus transcription. *J Virol* *81*, 20–29.
- Schlee, M., and Hartmann, G. (2016). Discriminating self from non-self in nucleic acid sensing. *Nat Rev Immunol* *16*, 566–580.
- Schlee, M., Roth, A., Hornung, V., Hagmann, C.A., Wimmenauer, V., Barchet, W., Coch, C., Janke, M., Mihailovic, A., Wardle, G., et al. (2009). Recognition of 5' Triphosphate by RIG-I Helicase Requires Short Blunt Double-Stranded RNA as Contained in Panhandle of Negative-Strand Virus. *Immunity* *31*, 25–34.
- Schmeisser, H., Mejido, J., Balinsky, C.A., Morrow, A.N., Clark, C.R., Zhao, T., and Zoon, K.C. (2010). Identification of alpha interferon-induced genes associated with antiviral activity in Daudi cells and characterization of IFIT3 as a novel antiviral gene. *J Virol* *84*, 10671–10680.
- Schmidt, A., Schwerd, T., Hamm, W., Hellmuth, J.C., Cui, S., Wenzel, M., Hoffmann, F.S., Michallet, M.-C., Besch, R., Hopfner, K.-P., et al. (2009). 5'-triphosphate RNA requires base-paired structures to activate antiviral signaling via RIG-I. *Proc Natl Acad Sci USA* *106*, 12067–12072.
- Schneider, T.D., and Stephens, R.M. (2004). Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* *18*, 1–4.

- Schoggins, J.W., Wilson, S.J., Panis, M., Murphy, M.Y., Jones, C.T., Bieniasz, P., and Rice, C.M. (2011). A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature* 472, 481–485.
- Schuberth-Wagner, C., Ludwig, J., Bruder, A.K., Herzner, A.M., Zillinger, T., Goldeck, M., Schmidt, T., Schmid-Burgk, J.L., Kerber, R., Wolter, S., et al. (2015). A Conserved Histidine in the RNA Sensor RIG-I Controls Immune Tolerance to N1-2'O-Methylated Self RNA. *Immunity* 43, 41–51.
- Schulz, O., Diebold, S.S., Chen, M., Näslund, T.I., Nolte, M.A., Alexopoulou, L., Azuma, Y.-T., Flavell, R.A., Liljeström, P., and Reis e Sousa, C. (2005). Toll-like receptor 3 promotes cross-priming to virus-infected cells. *Nature* 433, 887–892.
- Sen, G.C., and Peters, G.A. (2007). Viral stress-inducible genes. *Adv Virus Res* 70, 233–263.
- Sevajol, M., Subissi, L., Decroly, E., Canard, B., and Imbert, I. (2014). Insights into RNA synthesis, capping, and proofreading mechanisms of SARS-coronavirus. *Virus Research* 194, 90–99.
- Sheldrick, G.M. (2007). A short history of SHELX. *Acta Crystallogr A Found Crystallogr* 64, 112–122.
- Shu, C., Sankaran, B., Chaton, C.T., Herr, A.B., Mishra, A., Peng, J., and Li, P. (2013). Structural Insights into the Functions of TBK1 in Innate Antimicrobial Immunity. *Structure* 21, 1137–1148.
- Siegfried, A., Berchtold, S., Manncke, B., Deuschle, E., Reber, J., Ott, T., Weber, M., Kalinke, U., Hofer, M.J., Hatesuer, B., et al. (2013). IFIT2 is an effector protein of type I IFN-mediated amplification of lipopolysaccharide (LPS)-induced TNF- α secretion and LPS-induced endotoxin shock. *The Journal of Immunology* 191, 3913–3921.
- Sledz, C.A., Holko, M., de Veer, M.J., Silverman, R.H., and Williams, B.R.G. (2003). Activation of the interferon system by short-interfering RNAs. *Nat Cell Biol* 5, 834–839.
- Smietanski, M., Werner, M., Purta, E., Kaminska, K.H., Stepinski, J., Darzynkiewicz, E., Nowotny, M., and Bujnicki, J.M. (2014). Structural analysis of human 2'-O-ribose methyltransferases involved in mRNA cap structure formation. *Nat Commun* 5, 3004.
- Smith, J.B., and Herschman, H.R. (1996). The glucocorticoid attenuated response genes GARG-16, GARG-39, and GARG-49/IRG2 encode inducible proteins containing multiple tetratricopeptide repeat domains. *Arch. Biochem. Biophys.* 330, 290–300.
- Stawowczyk, M., Van Scoy, S., Kumar, K.P., and Reich, N.C. (2011). The interferon stimulated gene 54 promotes apoptosis. *Journal of Biological Chemistry* 286, 7257–7266.
- Stepiński, J., Waddell, C., Stolarski, R., Darzynkiewicz, E., and Rhoads, R.E. (2001). Synthesis and properties of mRNAs containing the novel "anti-reverse" cap analogs 7-methyl(3'-O-methyl)GpppG and 7-methyl(3'-deoxy)GpppG. *Rna* 7, 1486–1495.

- Strahle, L., Garcin, D., and Kolakofsky, D. (2006). Sendai virus defective-interfering genomes and the activation of interferon-beta. *Virology* 351, 101–111.
- Sun, Y., Jain, D., Koziol-White, C.J., Genoyer, E., Gilbert, M., Tapia, K., Panettieri, R.A., Jr, Hodinka, R.L., and López, C.B. (2015). Immunostimulatory Defective Viral Genomes from Respiratory Syncytial Virus Promote a Strong Innate Antiviral Response during Infection in Mice and Humans. *PLoS Pathog* 11, e1005122.
- Sutton, G., Grimes, J.M., Stuart, D.I., and Roy, P. (2007). Bluetongue virus VP4 is an RNA-capping assembly line. *Nat Struct Mol Biol* 14, 449–451.
- Svergun, D.I., IUCr (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* 25, 495–503.
- Svergun, D., Barberato, C., and Koch, M. (1995). CRY SOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* 28, 768–773.
- Szretter, K.J., Daniels, B.P., Cho, H., Gainey, M.D., Yokoyama, W.M., Gale, M., Virgin, H.W., Klein, R.S., Sen, G.C., and Diamond, M.S. (2012). 2'-O methylation of the viral mRNA cap by West Nile virus evades ifit1-dependent and -independent mechanisms of host restriction in vivo. *PLoS Pathog* 8, e1002698.
- Takahasi, K., Yoneyama, M., Nishihori, T., Hirai, R., Kumeta, H., Narita, R., Gale, M., Inagaki, F., and Fujita, T. (2008). Nonself RNA-sensing mechanism of RIG-I helicase and activation of antiviral immune responses. *Mol Cell* 29, 428–440.
- Tanji, H., Ohto, U., Shibata, T., Taoka, M., Yamauchi, Y., Isobe, T., Miyake, K., and Shimizu, T. (2015). Toll-like receptor 8 senses degradation products of single-stranded RNA. *Nat Struct Mol Biol* 22, 109–115.
- Tatematsu, M., Nishikawa, F., Seya, T., and Matsumoto, M. (2013). Toll-like receptor 3 recognizes incomplete stem structures in single-stranded viral RNA. *Nat Commun* 4, 1833.
- Tekes, G., Rahmeh, A.A., and Whelan, S.P.J. (2011). A Freeze Frame View of Vesicular Stomatitis Virus Transcription Defines a Minimal Length of RNA for 5' Processing. *PLoS Pathog* 7, e1002073.
- Terenzi, F., Hui, D.J., Merrick, W.C., and Sen, G.C. (2006). Distinct induction patterns and functions of two closely related interferon-inducible human genes, ISG54 and ISG56. *Journal of Biological Chemistry* 281, 34064–34071.
- Terenzi, F., Pal, S., and Sen, G.C. (2005). Induction and mode of action of the viral stress-inducible murine proteins, P56 and P54. *Virology* 340, 116–124.
- Terenzi, F., Saikia, P., and Sen, G.C. (2008). Interferon-inducible protein, P56, inhibits HPV DNA replication by binding to the viral protein E1. *Embo J* 27, 3311–3321.

- Terenzi, F., White, C., Pal, S., Williams, B.R.G., and Sen, G.C. (2007). Tissue-specific and inducer-specific differential induction of ISG56 and ISG54 in mice. *J Virol* *81*, 8656–8665.
- Terwilliger, T.C. (2003). SOLVE and RESOLVE: Automated Structure Solution and Density Modification. In *Macromolecular Crystallography, Part D*, (Elsevier), pp. 22–37.
- Thillier, Y., Decroly, E., Morvan, F., Canard, B., Vasseur, J.-J., and Debart, F. (2012). Synthesis of 5' cap-0 and cap-1 RNAs using solid-phase chemistry coupled with enzymatic methylation by human (guanine-N⁷)-methyl transferase. *Rna* *18*, 856–868.
- Tiwari, R.K., Kusari, J., and Sen, G.C. (1987). Functional equivalents of interferon-mediated signals needed for induction of an mRNA can be generated by double-stranded RNA and growth factors. *Embo J* *6*, 3373–3378.
- Topisirovic, I., Svitkin, Y.V., Sonenberg, N., and Shatkin, A.J. (2010). Cap and cap-binding proteins in the control of gene expression. *WIREs RNA* *2*, 277–298.
- Toroney, R., Hull, C.M., Sokoloski, J.E., and Bevilacqua, P.C. (2012). Mechanistic characterization of the 5'-triphosphate-dependent activation of PKR: Lack of 5'-end nucleobase specificity, evidence for a distinct triphosphate binding site, and a critical role for the dsRBD. *Rna* *18*, 1862–1874.
- Uchikawa, E., Lethier, M., Malet, H., Brunel, J., Gerlier, D., and Cusack, S. (2016). Structural Analysis of dsRNA Binding to Anti-viral Pattern Recognition Receptors LGP2 and MDA5. *Mol Cell* *62*, 586–602.
- Ulker, N., and Samuel, C.E. (1987). Mechanism of interferon action. II. Induction and decay kinetics of the antiviral state and protein P54 in human amnion U cells treated with gamma interferon. *Journal of Biological Chemistry* *262*, 16804–16807.
- Ulker, N., Zhang, X., and Samuel, C.E. (1987). Mechanism of interferon action. I. Characterization of a 54-kDa protein induced by gamma interferon with properties similar to a cytoskeletal component. *Journal of Biological Chemistry* *262*, 16798–16803.
- van Duijn, L.P., Kasperaitis, M., Ameling, C., and Voorma, H.O. (1986). Additional methylation at the N(2)-position of the cap of 26S Semliki Forest virus late mRNA and initiation of translation. *Virus Research* *5*, 61–66.
- Vasiljeva, L., Merits, A., Auvinen, P., and Kääriäinen, L. (2000). Identification of a novel function of the alphavirus capping apparatus. RNA 5'-triphosphatase activity of Nsp2. *Journal of Biological Chemistry* *275*, 17281–17287.
- Wacher, C., Müller, M., Hofer, M.J., Getts, D.R., Zabarar, R., Ousman, S.S., Terenzi, F., Sen, G.C., King, N.J.C., and Campbell, I.L. (2007). Coordinated regulation and widespread cellular expression of interferon-stimulated genes (ISG) ISG-49, ISG-54, and ISG-56 in the central nervous system after infection with distinct viruses. *J Virol* *81*, 860–871.
- Wakai, C., Iwama, M., Mizumoto, K., and Nagata, K. (2011). Recognition of cap structure by

influenza B virus RNA polymerase is less dependent on the methyl residue than recognition by influenza A virus polymerase. *J Virol* 85, 7504–7512.

Walsh, D., and Mohr, I. (2011). Viral subversion of the host protein synthesis machinery. *Nat Rev Microbiol* 9, 860–875.

Wang, C., Pflugheber, J., Sumpter, R., Sodora, D.L., Hui, D., Sen, G.C., and Gale, M. (2003). Alpha interferon induces distinct translational control programs to suppress hepatitis C virus RNA replication. *J Virol* 77, 3898–3912.

Wang, Y., Ludwig, J., Schubert, C., Goldeck, M., Schlee, M., Li, H., Juranek, S., Sheng, G., Micura, R., Tuschl, T., et al. (2010). Structural and functional insights into 5'-ppp RNA pattern recognition by the innate immune receptor RIG-I. *Nat Struct Mol Biol* 17, 781–787.

Wathelet, M.G., Clauss, I.M., Content, J., and Huez, G.A. (1988). The IFI-56K and IFI-54K interferon-inducible human genes belong to the same gene family. *FEBS Lett* 231, 164–171.

Wathelet, M.G., Clauss, I.M., Nols, C.B., Content, J., and Huez, G.A. (1987). New inducers revealed by the promoter sequence analysis of two interferon-activated human genes. *Eur J Biochem* 169, 313–321.

Wathelet, M., Moutschen, S., Defilippi, P., Cravador, A., Collet, M., Huez, G., and Content, J. (1986). Molecular cloning, full-length sequence and preliminary characterization of a 56-kDa protein induced by human interferons. *Eur J Biochem* 155, 11–17.

Weber, M., Gawanbacht, A., Habjan, M., Rang, A., Borner, C., Schmidt, A.M., Veitinger, S., Jacob, R., Devignot, S., Kochs, G., et al. (2013). Incoming RNA Virus Nucleocapsids Containing a 5'-Triphosphorylated Genome Activate RIG-I and Antiviral Signaling. *Cell Host Microbe* 13, 336–346.

Wei, C.M., and Moss, B. (1975). Methylated nucleotides block 5'-terminus of vaccinia virus messenger RNA. *Proc Natl Acad Sci USA* 72, 318–322.

Wei, C.M., Gershowitz, A., and Moss, B. (1976). 5'-Terminal and internal methylated nucleotide sequences in HeLa cell mRNA. *Biochemistry* 15, 397–401.

Wei, C.-M., Gershowitz, A., and Moss, B. (1975). N⁶, O^{2'}-dimethyladenosine a novel methylated ribonucleoside next to the 5' terminal of animal cell and virus mRNAs. *Nature* 257, 251–253.

Werner, M., Purta, E., Kaminska, K.H., Cymerman, I.A., Campbell, D.A., Mitra, B., Zamudio, J.R., Sturm, N.R., Jaworski, J., and Bujnicki, J.M. (2011). 2'-O-ribose methylation of cap2 in human: function and evolution in a horizontally mobile family. *Nucleic Acids Res* 39, 4756–4768.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67, 235–242.

- Xiao, S., Li, D., Zhu, H.-Q., Song, M.-G., Pan, X.-R., Jia, P.-M., Peng, L.-L., Dou, A.-X., Chen, G.-Q., Chen, S.-J., et al. (2006). RIG-G as a key mediator of the antiproliferative activity of interferon-related pathways through enhancing p21 and p27 proteins. *Proc Natl Acad Sci USA* *103*, 16448–16453.
- Yang, Z., Liang, H., Zhou, Q., Li, Y., Chen, H., Ye, W., Chen, D., Fleming, J., Shu, H., and Liu, Y. (2012). Crystal structure of ISG54 reveals a novel RNA binding structure and potential functional mechanisms. *Cell Res* *22*, 1328–1338.
- Ye, S., Pang, H., Gu, Y.-Y., Hua, J., Chen, X.-G., Bao, C.-D., Wang, Y., Zhang, W., Qian, J., Tsao, B.P., et al. (2003). Protein interaction for an interferon-inducible systemic lupus associated gene, IFIT1. *Rheumatology (Oxford)* *42*, 1155–1163.
- Yoneyama, M., Kikuchi, M., Natsukawa, T., Shinobu, N., Imaizumi, T., Miyagishi, M., Taira, K., Akira, S., and Fujita, T. (2004). The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol* *5*, 730–737.
- Yoshikawa, Y., Mizumoto, K., and Yamanouchi, K. (1986). Characterization of Messenger RNAs of Measles Virus. *Journal of General Virology* *67*, 2807–2812.
- Young, D.F., Andrejeva, J., Li, X., Inesta-Vaquera, F., Dong, C., Cowling, V.H., Goodbourn, S., and Randall, R.E. (2016). Human IFIT1 Inhibits mRNA Translation of Rubulaviruses but Not Other Members of the Paramyxoviridae Family. *J Virol* *90*, 9446–9456.
- Yu, M., Tong, J.H., Mao, M., Kan, L.X., Liu, M.M., Sun, Y.W., Fu, G., Jing, Y.K., Yu, L., Lepaslier, D., et al. (1997). Cloning of a gene (RIG-G) associated with retinoic acid-induced differentiation of acute promyelocytic leukemia cells and representing a new member of a family of interferon-stimulated genes. *Proc Natl Acad Sci USA* *94*, 7406–7411.
- Zeytuni, N., and Zarivach, R. (2012). Structural and Functional Discussion of the Tetra-Trico-Peptide Repeat, a Protein Interaction Module. *Structure* *20*, 397–405.
- Zhang, Y., and Chan, D.C. (2007). Structural basis for recruitment of mitochondrial fission complexes by Fis1. *Proceedings of the National Academy of Sciences* *104*, 18526–18530.
- Zhang, Y., Burke, C.W., Ryman, K.D., and Klimstra, W.B. (2007). Identification and Characterization of Interferon-Induced Proteins That Inhibit Alphavirus Replication. *J Virol* *81*, 11246–11255.
- Zhang, Y., Wei, Y., Zhang, X., Cai, H., Niewiesk, S., and Li, J. (2014). Rational design of human metapneumovirus live attenuated vaccine candidates by inhibiting viral mRNA cap methyltransferase. *J Virol* *88*, 11411–11429.
- Zhang, Z., Ohto, U., Shibata, T., Krayukhina, E., Taoka, M., Yamauchi, Y., Tanji, H., Isobe, T., Uchiyama, S., Miyake, K., et al. (2016). Structural Analysis Reveals that Toll-like Receptor 7 Is a Dual Receptor for Guanosine and Single-Stranded RNA. *Immunity* *45*, 737–748.
- Zheng, X., and Bevilacqua, P.C. (2004). Activation of the protein kinase PKR by short double-

stranded RNAs with single-stranded tails. *Rna* *10*, 1934–1945.

Zhou, Y., Ray, D., Zhao, Y., Dong, H., Ren, S., Li, Z., Guo, Y., Bernard, K.A., Shi, P.Y., and Li, H. (2007). Structure and Function of Flavivirus NS5 Methyltransferase. *J Virol* *81*, 3891–3903.

Zlatev, I., Lavergne, T., Debart, F., Vasseur, J.-J., Manoharan, M., and Morvan, F. (2010). Efficient Solid-Phase Chemical Synthesis of 5'-Triphosphates of DNA, RNA, and their Analogues. *Organic Letters* *12*, 2190–2193.

Zlatev, I., Manoharan, M., Vasseur, J.-J., and Morvan, F. (2012). Solid-phase chemical synthesis of 5'-triphosphate DNA, RNA, and chemically modified oligonucleotides. *Curr Protoc Nucleic Acid Chem Chapter 1*, Unit1.28–1.28.16.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* *31*, 3406–3415.

Zust, R., Cervantes-Barragan, L., Habjan, M., Maier, R., Neuman, B.W., Ziebuhr, J., Szretter, K.J., Baker, S.C., Barchet, W., Diamond, M.S., et al. (2011). Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat Immunol* *12*, 137–143.

Zust, R., Dong, H., Li, X.-F., Chang, D.C., Zhang, B., Balakrishnan, T., Toh, Y.-X., Jiang, T., Li, S.-H., Deng, Y.-Q., et al. (2013). Rational Design of a Live Attenuated Dengue Vaccine: 2'-O-Methyltransferase Mutants Are Highly Attenuated and Immunogenic in Mice and Macaques. *PLoS Pathog* *9*, e1003521–13.