Hydrogen/deuterium exchange mass spectrometry reveals distinct mechanisms of viral RNA recognition and activation by RIG-I like receptors

Zheng, Jie

2015

Zheng, J. (2015). Hydrogen/deuterium exchange mass spectrometry reveals distinct mechanisms of viral RNA recognition and activation by RIG-I like receptors. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/62943>

<https://doi.org/10.32657/10356/62943>

Hydrogen/Deuterium Exchange Mass Spectrometry reveals
distinct mechanisms of viral RNA recognition and
activation by RIG-I like Receptors

ZHENG JIE

SCHOOL OF BIOLOGICAL SCIENCES

2015

Hydrogen/Deuterium Exchange Mass Spectrometry reveals
distinct mechanisms of viral RNA recognition and
activation by RIG-I like Receptors

ZHENG JIE

School of Biological Sciences

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2015

Acknowledgements

In the course of study, I would like to thank all parties that have assisted me all the way. My heartfelt gratitude goes to my supervisors Dr. Tang Kai and Dr. Liu Chuanfa for their guidance and support throughout the PhD program. I especially acknowledge Dr. Luo Dahai for collaborations on RIG-I like receptors (RLRs). Obviously I could not have accomplished this thesis without their enlightening instruction, kindness and patience. Their strict requirements and rigorous attitudes towards science will continue to impact my research career in the future.

I would like to thank my parents and all my family members for their endless support and encouragement in the past twenty-seven years. I also wish to extend my sincere thanks to Dr. Richard Sugrue lab for collaborations on H1N1 project.

I acknowledge Bruce, Pascal from Dr. Patrick, Griffin lab (Scripps, FL) for software technical support of HDX Workbench.

I would like to thank all my wonderful lab members: Mr. Dong Xueming, Dr. Nantika Panutdaporn, Dr. Gao Xiang, Dr. Wang Han, Mr. John Williams, especially Mr. Dong Xueming for professional guidance on proteomic techniques and Hydrogen/Deuterium exchange mass spectrometry. I also would like to thank Ms. Huiyee Yong, Dr. Ho Chun Loong, Dr. Yang lifeng, Mr. Bi Xiaobao, Mr. Fung Tosing from other laboratories for assistance and support on my Ph.D projects. Especially, I wish to thank senior engineer Rinson Lai from alpha-analytical Inc, Vicky Tan from thermofisher scientific Inc for professional assistance on LTQ-Orbitrap mass spectrometer and nano-liquid chromatography, respectively. Moreover, I thank our school IT engineer Alexander Lim for all the support given these years.

Table of Contents

Acknowledgements.....	iii
Table of Contents.....	iv
List of figures.....	vii
List of tables.....	viii
Abbreviations.....	ix
Abstract.....	xii
Chapter 1: Introduction.....	1
1.1 RIG-I like receptors and innate immunity	1
1.1.1 Innate immunity receptors	1
1.1.2 RNA recognition features of RLRs	4
1.1.3 Interferon β antiviral pathways	7
1.1.4 RLR signaling pathways.....	10
1.2 Hydrogen/Deuterium Exchange Mass Spectrometry	14
1.2.1 Principle of HDX and its applications	15
1.2.2 HDX experimental design	17
1.3 Influenza A virus	21
1.3.1 Genome structure of influenza A virus.....	22
1.3.2 Influenza A viral proteins	24
1.3.3 Influenza A virus infectious cycles.....	26
1.4 Proteomics and applications on virology.....	29
1.4.1 Proteomic searching engines	31
1.5 Objectives and organization of this dissertation	34
Chapter 2: HDX-MS characterization of RIG-I activation mechanism	36
2.1 Introduction.....	36
2.2 Materials and methods	38
2.2.1 Materials	38
2.2.2 Cloning, protein expression and purification.....	38
2.2.3 RNA work.....	40
2.2.4 Hydrogen/Deuterium exchange	40
2.2.5 LCMS under quench condition.....	41
2.2.6 Structure modeling.....	43

2.3 Results.....	43
2.3.1 Optimization of RIG-I protein expression and purification.....	43
2.3.2 Human RIG-I maintains the auto-inhibited conformation through the intra-molecular CARD2:HEL2i interactions	45
2.3.3 RIG-I binds and responds to different RNA species (3p8L, 3p10L, and polyIC) differently	50
2.3.4 ATP allosterically modulates conformational rearrangements of RIG-I RNA complex.....	54
2.3.5 The recognition of 3p10L by LGP2 is similar as RIG-I: HEL1 and CTD dominantly mediate hairpin RNA binding.....	58
2.4 Discussion and conclusion.....	61
2.4.1 Differential behaviors in RNA recognition model HEL-CTD	61
2.4.2 CARD2-HEL2i interface governs the key switch of RIG-I molecular machinery.....	62
 Chapter 3: HDX-MS characterization of MDA5 activation mechanism.....	65
3.1 Introduction.....	65
3.2 Materials and methods	66
3.2.1 Materials	66
3.2.2 Cloning, protein expression and purification.....	67
3.2.3 RNA work.....	68
3.2.4 Hydrogen/Deuterium exchange	69
3.2.5 LCMS under quench condition.....	69
3.2.6 Structure modeling.....	70
3.3 Results.....	71
3.3.1 Protein expression and purification of MDA5.....	71
3.3.2 CARDS of apo MDA5 are exposed in solution	71
3.3.3 MDA5 binds polyIC through HEL and CTD	75
3.3.4 Highly cooperative binding of MDA5 to polyIC.....	77
3.3.5 ATP weakens interaction between MDA5 and polyIC but does not disrupt cooperativity	81
3.3.6 Recognition of PolyIC by LGP2.....	82
3.3.7 LGP2 forms different complex with PolyIC compared with MDA5:PolyIC filamentous structure	85
3.3.8 LGP2 stays no interactions with polyIC bound MDA5 in solution.....	87
3.4 Discussion and conclusion.....	89
 Chapter 4: protein analysis of Annexin A2 in influenza A virus particles	92
4.1 Introduction.....	92

4.2 Materials and methods	94
4.2.1 Materials	94
4.2.2 Virus production and purification	95
4.2.3 Protease treatment of virions	95
4.2.4 In gel digestion.....	96
4.2.5 One-dimensional LC-MS/MS	97
4.2.6 Database searching	98
4.2.6.1 Integrated proteomics pipeline searching engine.....	98
4.2.6.2 Global Proteome Machine database searching	99
4.2.7 SDS-PAGE and western blot analysis	99
4.2.8 Immunofluorescence microscopy	99
4.2.9 Transfections siRNA molecules	100
4.3 Results.....	100
4.3.1 Subtilisin digestion of bovine serum albumin and its compatibility with 3 % sodium dodecyl sulfate (SDS)	100
4.3.2 “Tube gel” in-gel digestion and LC-MS/MS analysis	102
4.3.3 Proteomic analysis of untreated IAV particles, subtilisin-treated IAV particles, and subtilisin&3% SDS treated IAV particles	103
4.3.4 Western blot validation of host proteins incorporated into IAV	108
4.3.5 Viral NP is partially associated with α -tubulin and is co-enriched with annexin A2 in the virus infected cells visualized by fluorescence microscopy	110
4.3.6 Knockdown of annexin A2 in DF-1 cells does not affect IVA virus productions.....	114
4.4 Discussion and conclusion.....	116
Chapter 5: Conclusions and future works	119
5.1 Conclusions on HDX-MS characterization of RLR activations.....	119
5.1.1 HDX-MS is an ideal tool to study the molecular mechanism of RLR signalling.....	119
5.1.2 The complete surveillance of virus RNA PAMPs is ensured by the different resting states of RIG-I and MDA5	120
5.1.3 RIG-I and MDA5 are activated by separate and distinct mechanisms .	120
5.2 Conclusions of proteomics studies on host factors incorporated into influenza A virions.....	123
5.3 Future directions	124
References.....	130
Appendix.....	146
Publications.....	165
Posters, awards, invited talks	165

List of figures

Figure1.1 Schematic representation of the domain structure of RIG-I like receptors.....	5
Figure1.2 Interferon signaling pathway.....	9
Figure1.3 The RIG-I/MDA5 signaling pathway.....	12
Figure1.4 Proposed model of the molecular basis of the RLR signaling pathway. .	13
Figure1.5 Backbone amide hydrogens for HDX-MS	15
Figure1.6 pH and temperature affect amide hydrogen exchange rate	17
Figure1.7 HDX sample preparation and on-line, off-line pepsin digestion HDX-MS.....	19
Figure1.8 House-built on-line pepsin digestion HDX-MS system.....	20
Figure1.9 The genome organization of influenza H1N1/PR/8/34.....	23
Figure1.10 Schematic representation of influenza A virus.....	23
Figure2.1 SDS-PAGE gels and size-exclusion chromatogram of the recombinant RIG-I protein.....	44
Figure2.2 RIG-I apoenzyme conformation.....	49
Figure2.3 Characteristic HDX profiles of RIG-I peptides induced by different RNA ligands	53
Figure2.4 HDX plots of RIG-I peptides associated with ATP effect	55
Figure2.5 Characterization of ATP analogue effects on RIG-I domain peptides...	57
Figure2.6 Structure model of 3p10L bound LGP2	58
Figure2.7 Perturbation view: differential deuterium exchange Map for LGP2 ± 3p10L	60
Figure2.8 A model of the RIG-I activation pathway	64
Figure3.1 SDS-PAGE gels and size-exclusion chromatogram of the recombinant MDA5 protein.....	71
Figure3.2 MDA5 apoenzyme conformation.....	74
Figure3.3 Structure basis of MDA5 cooperativity.....	80
Figure3.4 Structure model of PolyIC bound LGP2	83
Figure3.5 Perturbation view: differential deuterium exchange Map for LGP2 ± PolyIC	84
Figure3.6 Perturbation view: differential deuterium exchange Map for polyIC bound MDA5 ± LGP2	88
Figure3.7 MDA5 early activation mechanisms revealed by HDX-MS.....	91

Figure4.1 Subtilisin digestion of bovine serum albumin (BSA) and its digestion compatibility with 3 % Sodium dodecyl sulfate (SDS).....	101
Figure4.2 Expression level of IAV proteins and their sequence length	103
Figure4.3 Experimental flowchart of subtilisin-treated, subtilisin & 3% SDS-treated, and untreated IAV virions	106
Figure4.4 LC-MS/MS spectra of the triply charged peak at m/z 608.62 (M2 peptide ⁶² GPSTEGVPESMREEYR ⁷⁷)	129
Figure4.5 Validation of host proteins and viral proteins by western blot analysis.....	110
Figure4.6 Staining patterns of viral NP and host α -tubulin in IAV infected and non-infected DF-1 cells by fluorescence microscopy.....	112
Figure4.7 Staining patterns of viral NP and host annexin A2 in IAV infected and non-infected DF-1 cells by fluorescence microscopy	113
Figure4.8 Knockdown of annexin A2 in DF-1 cells.....	115
Figure4.9 Effect of annexin A2 knockdown on IAV infected or mock infected DF-1 cells.....	116

List of Tables

Table1.1 A summary of innate immune receptors.....	2
Table1.2 Reprehensive proteomics applications to selected viruses	30
Table2.1 Comparison view of the conformational dynamics of RIG-I and RIG-I Δ CARDs induced by RNA ligand (5' triphosphorylated RNA hairpin 3p10L and polyIC) and ATP hydrolysis	51
Table3.1 Comparison view of the conformational dynamics of MDA5 and MDA5 Δ CARDs induced by polyIC and ATP hydrolysis	76
Table3.2 Comparison view of the conformational dynamics of MDA5 (MDA5 Δ CARDs) and LGP2 induced by polyIC	85
Table4.1 Summary of current approaches on characterizing the proteome of enveloped virions.....	126
Table4.2 Identification and comparison of viral proteins in mock, subtilisin-treated, and subtilisin-3%SDS treated virions	127
Table4.3 Identification and comparison of host proteins in mock, subtilisin-treated, and subtilisin-3%SDS treated virions	128
Table4.4 siRNA designs for gallus annexin A2 knockdown.....	114
Appendix A HDX-MS sequence coverage of RIG-I	146

Appendix B HDX-MS sequence coverage of RIG-IΔCARDs	147
Appendix C HDX-MS sequence coverage of MDA5	148
Appendix D HDX-MS sequence coverage of MDA5ΔCARDs	149
Appendix E Host proteins uniquely identified in protease-treated virions by IP2150	
Appendix F A summerization of differential hydrogen/deuterium exchange data for RIG-I and RIG-IΔCARDs.....	152
Appendix G A summerization of differential hydrogen/deuterium exchange data for MDA5 and MDA5ΔCARDs	160

Abbreviations

ATP	Adenosine triphosphate
ADP	Adenosine diphosphate
AIM2	Absent in melanoma 2
ANX2	Annexin A2
BCA	Bicinchoninic acid assay
BIRs	Baculovirus inhibitor repeats
CARD	Caspase recruitment domain
CDS	Cytosolic DNA sensors
CLRs	C-type lectin receptors
CTD	C-terminal domain
DAI	DNA-dependent activator of interferon-regulatory factors
DC	Dendritic cell
DEME	Dulbecco's Modified Eagle Medium
DI	Defective interfering
DNA	Deoxyribonucleic acid
DRA	Double-stranded RNA stimulated ATPase
ECL	Enhanced chemiluminescence
<i>E.coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
ERK	extracellular signal-regulated kinase
ESI	Electrospray ionization
FBS	Foetal Bovine Serum
FDA	Food and Drug Administration
FPLC	Fast protein liquid chromatography
GFP	Green fluorescent protein
GPM	Global Proteome Machine
HA	Hemagglutinin
HDX	Hydrogen/Deuterium Exchange
HEL	Helicase domain

IAV	Influenza A virus
IFN	Interferon
IP	Immunoprecipitation
IP2	Integrated Proteomics Pipeline
IPTG	Isopropyl- β -D-thiogalactopyranoside
IRFs	Interferon regulatory factors
JNK	c-Jun N-terminal kinase
kDa	Kilo Dalton
LB	Lysogeny broth
LC-MS/MS	Liquid chromatography in couple with tandem mass spectrometry
LGP2	Laboratory of Genetics and Physiology 2
LRRs	Leucine rich repeats
Lrrfp1	Leucine-rich repeat interacting protein-1
LPS	Lipopolysaccharide
MALDI	Matrix assisted laser desorption/ionization
MAPKs	Mitogen-activated protein kinases
MAVS	Mitochondrial antiviral-signaling protein
MDA5	Melanoma Differentiation Associated gene 5
MOI	Multiplicity of infection
MS	Mass spectrometry
MW	Molecular weight (in Dalton)
M1	Matrix 1 protein
M2	Matrix 2 protein
NA	Neuraminidase
NEP	Nuclear export protein
NES	Nuclear export signal
NCBI	National Center for Biotechnology Information
NF- κ B	Nuclear factor- κ B
Ni ²⁺ -NTA	Nickel nitrilotriacetic acid
NLRs	Nucleotide-binding and oligomerization domain-like receptors
NLS	Nuclear localization signals
NMR	Nuclear Magnetic Resonance
NP	Nucleoprotein
NS1	Non-structural protein 1
°C	Degree Celsius
OD _{600nm}	Ocular density measured at 600 nm
ORF	Open reading frame
PA	Polymerase acidic protein
PAMPs	Pathogen-associated molecular patterns
PB1	Polymerase basic protein 1
PB1-F2	Polymerase basic protein one-F2
PB2	Polymerase basic protein 2
PCR	Polymerase chain reaction
PDB	Protein Data Bank

PI3K	Phosphatidylinositol 3-kinase
PMSF	Phenyl methyl sulfonyl fluoride
Pol III	RNA polymerase III
PolyIC	Polyinosinic:polycytidylic acid
PRRs	Pathogen recognition receptors
PYD	Pyrin domain
RanBP5	Ran binding protein 5
RIG-I	Retinoic acid-Inducible Gene 1
RLRs	RIG-I-like receptors
RNA	Ribonucleic acid
siRNA	Small interfering RNA
TAP	Tandem affinity purification
TBK1	TANK-binding kinase 1
TIR	Toll/interleukin-1 receptor
TLRs	Toll-like receptors
TMD	Transmembrane domain
TNG	Trans-Golgi network
rpm	Revolutions per minute
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
UV	Ultra-violet
vRNPs	Viral ribonucleoproteins

Abstract

In the first two projects, we applied hydrogen/deuterium exchange coupled with mass spectrometry (HDX-MS) to study the structural dynamics of the RIG-I like immune receptors, RIG-I, MDA5 and LGP2. We probed the *apo* states of RIG-I and MDA5 in solution and dissected the molecular details of their shared and distinct features with respect to RNA recognition and CARDs activation. The RNA recognition by LGP2 was compared with RIG-I and MDA5 using the same RNA ligand.

RIG-I and MDA5 are the major intracellular immune receptors that recognize viral RNA species and undergo a series of conformational transitions leading to the activation of the interferon-mediated antiviral response. However, to date, full-length RLRs have resisted crystallographic efforts and a molecular description of their activation pathways remains hypothetical. Here we employ hydrogen/deuterium exchange coupled with mass spectrometry (HDX-MS) to probe the *apo* states of RIG-I and MDA5 and to dissect the molecular details with respect to distinct RNA species recognition, ATP binding and hydrolysis, and CARDs activation. We show that human RIG-I maintains an auto-inhibited resting state owing to the intra-molecular HEL2i-CARD2 interactions while *apo* MDA5 lacks the analogous intra-molecular interactions and therefore adopts an extended conformation. Our work demonstrates that RIG-I binds and responds differently to short triphosphorylated RNA and long duplex RNA and that sequential addition of RNA and ATP triggers specific allosteric effects leading to RIG-I CARDs activation. We also present a high-resolution protein surface mapping technique that refines the cooperative oligomerization model of neighboring MDA5 molecules on long duplex RNA. Taken together, our data provide a high-resolution view of RLR activation in solution and offers new evidence for the molecular mechanism of RLR activation.

In the third project, we utilized proteomics to analyze protease-treated influenza A virus particles to identify viral and host proteins that are specifically incorporated into the virions. Proteomic analysis of exogenous protease treated influenza A viruses could largely increase the confidence to sensitively identify host factors incorporated into virus particles. In addition, western blot, immunofluorescence microscope, and siRNA-mediated knockdown were employed to validate the association of identified host proteins with viral NP protein.

CHAPTER 1

Introduction

1.1 RIG-I like receptors and innate immunity

1.1.1 Innate immunity receptors

Animal viruses are infectious and pathogenic. Viruses could evolve and mutate much more quickly than host organisms. This could be partially attributed to the fact that RNA virus genomes exhibit hyper-variations and error-prone mutations due to lack of proofreading RNA dependent RNA polymerases [1, 2]. In addition, there exist a diversity of virus families and a vast number of formidable viruses. The representatives are human immunodeficiency virus, influenza virus, respiratory syncytial virus, severe acute respirator syndrome coronavirus, and dengue virus. To date, the numbers of complete viral genomes and proteomes that cover 118 taxonomy groups have reached to 2853 and 1932 in the NCBI and Uniprot databases, respectively.

Confronting threats of animal viruses, host organisms could occupy a network of multiple defensive measures to combat pathogenic invasions [9]. In the vertebrate immune system, it can be divided into innate and acquired immune response. Acquired immunity, also known as adaptive immunity, is mediated via T and B lymphocyte that could generate antigen-binding receptors after pathogens invade [10]. In contrast to acquired immunity, innate immunity is characterized by several factors, such as fast immune responses to combat pathogen infection, occupying specific receptors inherited in the genome, and targeting broad classes of pathogens. And innate immunity contributes to the first line of defense against viruses. In general, innate immune response relies on recognition of conserved pathogenic signatures, which are usually unique to microorganisms and considered

as “non-self” products by the host organisms. Therefore, these conserved molecules are called pathogen-associated molecular patterns (PAMPs). Distinct pathogens express different PAMPs, which require a specific group of receptors to recognize and countermeasure. Correspondingly, these receptors are called pathogen recognition receptors (PRRs) and they play vital roles in antiviral innate immune responses [11].

Table 1.1 A summary of innate immune receptors

Innate immune receptors	Number of identified receptors	Localization	Main recognition ligands
TLRs	10 (human) 12 (mouse)	Cellular surface; endocytic compartments	Lipopolysaccharide; lipoprotein; viral or bacterial nucleic acid
NLRs	22 (human)	Cytoplasm	Bacterial products, e.g. peptidoglycan
RLRs	3 (human)	Cytoplasm	Intracellular viral RNAs (dsRNA)
CDS	6 (human)	Cytoplasm	Intracellular viral DNAs
CLRs	> 60 (human)	Cellular surface of dendritic cells	Carbohydrate structures

The main PRRs families of the innate immunity consist of Toll-like receptors (TLRs), NOD-like receptors (NLRs), RIG-I-like receptors (RLRs), cytosolic DNA sensors (CDS), and C-type lectin receptors (CLRs) (Table 1.1).

TLRs belong to type I membrane glycoproteins. A total of 10 and 12 TLRs

have been identified in human and mice respectively [11]. According to cellular distribution, TLR 1, 2, 4, 5, and 6 reside on cellular surface and they are mainly responsible for recognition of bacterial or viral components, such as lipopolysaccharide and lipoprotein. TLR3, 7, 8, 9 with their location at endocytic compartments, are primarily dedicated to recognize ssRNA or dsDNA viruses [12]. These different TLRs share similar or evolve different signaling pathways for activation, which lead to induction of various genes involved in host defense system, such as NF- κ B, JNK and MAPKs [11-13]. NLRs are conserved intracellular PRRs that sense danger signals derived from bacterial infections as well as a variety of human diseases [14]. In human cells, a total of 22 NLRs have been identified and these receptors could concertedly cooperate with TLRs to regulate inflammatory and apoptotic responses. In addition, NLRs have also been discovered in lymphocytes, macrophages, and dendritic cells. Among the NLRs, some share similar activation mechanism via formation of inflammasome, which is mediated via caspase-1 pathway to generate mature cytokines; whereas some other NLRs could regulate and activate pathways involved in NF- κ B, MAPKs and IRFs to trigger innate immunity [15]. In addition, a total of six cytosolic DNA sensors have been reported to be involved in innate immunity, including DAI, AIM2, Pol III, Lrrfip1, DHX9, DHX36, and IFI16 [16]. These DNA sensors display distinct pathways leading to activation of the transcription factor NF- κ B and TBK1-mediated phosphorylation of IRF3. For example, Lrrfip1 could recognize viral DNA to induce IFN- β via IRF3 pathway [17]. DAI has been reported to activate type I interferon via TBK1-IRF3-dependent pathway [18]. Cytosolic DNA has also been reported to activate caspase-1 dependent maturation of pro-inflammatory cytokines interleukin IL-1 β and IL-18. For instance, AIM2 has been identified to act in the upstream caspase-1 dependent pathway [19]. CLRs comprise a large family of receptors that are expressed in dendritic cells [20]. In human cells, more than 60 CLRs have been identified and most of CLR members examine various carbohydrate structures, e.g. mannose, fucose and glucan [21]. CLRs are known to induce different signaling pathways that are involved in NF- κ B activation, TLR

dependent IFN pathway (e.g. TLR8 and TLR9), and NLR inflammasome dependent pathway (e.g. NLRP3) [20].

1.1.2 RNA recognition features of RLRs

RIG-I like receptors play important roles in mediating antiviral IFN responses by recognizing cytoplasmic “non-self” viral RNAs [22-24]. Due to the absence of RNA-dependent RNA polymerase, duplex RNA is seldom found in mammalian cells and it is widely accepted as a viral PAMP. In particular, polyIC, a mixture of relatively long duplex RNA with blunt, 5’ or 3’ heterogeneous ends, have exhibited comparative activity to viral infection as IFN-inducing compounds [25]. It is known that polyIC are potent RNA ligands triggering both RIG-I and MDA5 mediated antiviral responses in various assays [26-28]. The 5’ triphosphate or diphosphate moiety of short duplex RNA is another viral PAMP that is often generated during the viral replication process [29, 30]. For instance, the 5’ triphosphate and 5’ diphosphate moiety can be found at the termini of the genomes from influenza virus and reovirus [28]. In contrast, “self” RNAs are absent of such viral PAMP or are associated with chemical modifications [31]. For instance, a 7-methyl-guanosine cap structure is found at the 5’ end of messenger RNA in host cell. Transfer RNA undergoes 5’ cleavage and a series of nucleotide base modifications. Thus duplex RNA and 5’ triphosphate moiety are considered as “non-self” RNAs by RLRs whereas “self” RNAs escape from detection due to chemical modifications.

There are three members of RLR genes: RIG-I (Retinoic acid-Inducible Gene 1), MDA5 (Melanoma Differentiation Associated gene 5) and LGP2 (Laboratory of Genetics and Physiology 2). In Figure 1.1, all RLRs contain conserved RNA sensing module consisting of a helicase domain (HEL) linked to the C terminal domain (CTD), but only RIG-I and MDA5 contain the signaling module at their N-terminus, tandem caspase activation and recruitment domains

(CARDs; CARD1 and CARD2) [3, 7, 32, 33]. The function of the CARDs domain is mainly responsible for triggering the downstream signaling events leading to IFN production.

Figure 1.1 Schematic representation of the domain structure of RIG-I like receptors. Various domains are color-coded for RIG-I, MDA5 and LGP2. Truncated RIG-I and MDA5 (deleting the CARDs and the linker to HEL) are also shown.

The RNA sensing modules of RLRs are homologous in sequence and structure. The HEL domain is made of the conserved Rec-A-like-helicase core (HEL1 and HEL2) and a novel insertion domain HEL2i. HEL1 and HEL2 contain various conserved features. Motif Q, I, II, and VI harbor the ATP binding and hydrolysis pocket. Motif Ia, Ib, Ic, IV, IVa, V, Vb and Vc are associated with RNA binding activities [34, 35]. Because of this unique domain architecture RLRs are classified as double-stranded RNA stimulated ATPases (DRAs), which also

include Dicer and Dicer-related helicases [36]. As the name implies, DRAs recognize preferentially duplex RNA and play important roles in antiviral defense, RNA interference, to regulation of gene activities and genomic integrity [36]. In addition, DRAs do not exhibit RNA unwinding activities due to the lack of a β -hairpin motif [35]. The CTD is a Zinc-ion binding domain and is also involved in binding to the duplex RNA. RIG-I CTD has been known to specifically recognize blunt-ended duplex RNA bearing 5' triphosphate-specific RNA species generated during virus replication process [29, 30]. Very recently, RIG-I has also been shown to mediate specific antiviral responses to RNAs bearing 5' diphosphates which can be found at the termini of the genomes from reoviruses (segmented, double-stranded RNA viruses) [28]. This is mediated by the highly conserved triphosphate-binding pocket and the RNA duplex end capping loop [4, 37, 38]. RIG-I has also been shown to be activated by long duplex RNA (polyIC) but the conformation of RIG-I CTD and HEL binding internally to long duplex RNA remains elusive [23, 26, 27, 39, 40]. Unlike RIG-I, MDA5 does not contain such functional features but binds to the internal regions of long duplex RNA with weak affinity [41, 42]. The crystal structure of CARDs-deleted MDA5 in complex with 12mer duplex RNA shows several different features in terms of RNA recognition compared with RIG-I [41]. The CTD of MDA5 binds the stem rather than the end of the duplex RNA, thus it has a different structural orientation compared with RIG-I CTD. In addition, the HEL2 insertion loop of MDA5 directly inserts to the major groove of the duplex RNA, further enhancing the internal binding feature of MDA5. In contrast, the RIG-I HEL2 insertion loop resides near the end of duplex RNA and is not directly involved in RNA binding [43]. Another distinct feature of RNA recognition of MDA5 is that it forms head-to-tail interactions between MDA5 monomers along long duplex RNA. One electron microscopy study has shown that both full length and CARDs-deleted MDA5 could form helical filaments on long duplex RNA, suggesting that MDA5 utilizes HEL-CTD to mediate RNA recognition and cooperative binding [44]. Therefore, RIG-I and MDA5 are known to have preferences towards the RNA substrates: RIG-I

recognizes short, duplex RNA with a 5'end triphosphate and MDA5 cooperatively binds long duplex RNA [27, 45-47]. As indicated by these molecular recognition features, RIG-I and MDA5 recognize different but overlapping groups of RNA viruses [48, 49]. For instance, RIG-I has been reported to sense influenza viruses, Sendai virus, and hepatitis C viruses, and defective interfering (DI) particles [50, 51]. In contrast, MDA5 has been known to exhibit distinct recognition specificity for long duplex RNA viruses and dsRNA replication intermediates of some positive sense ssRNA, such as picornavirus and encephalomyocarditis virus [26, 52, 53].

Compared to RIG-I and MDA5, LGP2 is less known in its function involved in signaling pathway and structure. Contradictory observations have been reported in terms of LGP2 role in virus infections. LGP2 has been reported as a positive regulator to RLR mediated signaling pathway, since LGP2 knockout cells and mutant that abrogated its ATPase activity showed impaired IFN production after picornavirus infection [54]. LGP2 has also been observed to negatively regulate viral recognition and activation of RLRs. The CTD structure of LGP2 has revealed higher affinity binding to dsRNA terminus compared to that of MDA5 CTD [55]. LGP2 CTD has also been shown to inhibit interaction between activated RIG-I and MAVS [56]. Unlike RIG-I and MDA5, high basal ATP hydrolysis, which is independent of dsRNA stimulated ATP hydrolysis, has been shown to enhance LGP2 recognition towards various RNA ligands [57]. Currently, little is known about the similarities and differences of LGP2 from RIG-I and MDA5 in RNA recognition. Thus understanding the conformations of LGP2 apoenzyme and RNA bound state would assist to shed light on the inherent properties of RLRs.

1.1.3 Interferon β antiviral pathways

There are three classes of IFN have been identified, known as type I to III. So far little is known towards type III IFNs and type II class of IFN mediates broad

immune responses to pathogens other than viruses [58]. The IFN- α/β response constitutes a critical element of the innate immune system that is particularly important in combating against virus infections. Secretion of IFN- α/β results in the induction of a cellular antiviral response involving the transcriptional up-regulation of more than 100 IFN-stimulated genes (ISGs), which have antiviral, antiproliferative, and immunomodulatory functions. Once IFN- α/β has been synthesized, it is secreted and binds to the IFN- α/β receptors, IFNAR1 and IFNAR2. As a result, the cytoplasmic kinases JAK1 and TYK2 become activated and phosphorylate the STAT-1 and STAT-2 molecules that signal to the nucleus [59]. This process promotes generation of the ISGF3 transcription factor, a complex of STAT-1, STAT-2, and IRF9 that accumulates in the nucleus. Nuclear ISGF3 binds to promoters that contain IFN-stimulated response elements (ISRE) and stimulates the transcriptional induction of antiviral genes, including Mx, protein Kinase R (PKR), 2'-5'-oligoadenylate synthetase (OAS), ISG20, promyelocytic leukemia protein (PML) (Figure 1.2).

IFN-induced enzyme systems represent major antiviral pathways and many have been studied, including PKR signaling pathway, the 2'-5' OAS/RNaseL system, and the Mx GTPase pathway. PKR is synthesized in an inactive form and it undergoes dimerization and activation in response to dsRNA produced during viral replication [60]. Eukaryotic translational initiation factor 2 (eIF2a) is a well-characterized substrate for PKR. The phosphorylation by PKR prevents recycling of eIF2a [61]. PKR also mediates virus clearance by phosphorylation of protein kinase IKK β . IKK β then phosphorylates the NF- κ B inhibitor (I κ B), which activates the NF- κ B transcription factor. Dissociation and subsequent degradation of the phosphorylated inhibitor allows NF- κ B to translocate to the nucleus where NF- κ B regulates the expression of pro-inflammatory cytokines such as tumour necrosis factor- α and interleukin-6 [62]. Like PKR, OAS is also synthesized in an inactive form, and activation of OAS by pathogen-associated molecular patterns allows it to oligomerize ATP through an unusual 2'-5' phosphodiester linkage. The 2'-5'-

oligoadenylates bind to and activates RNase L, which degrades cellular and viral RNAs [63]. The Mx GTPase pathway is another antiviral response mediated by IFN-stimulated response element (ISRE). After stimulation by type I IFNs, myxovirus-resistance A (MXA) gene expression is induced with an ISRE as the gene promoter. The activated MxA protein oligomerizes and accumulates on the intracellular membranes. After viral invasion, the accumulated MxA oligomers are released for trapping viral components prior to degradation [64].

Figure 1.2 Interferon signaling pathway. IFN-α/β binds to the type I IFN receptor (IFNAR) and activate the expression of numerous IFN responsive genes (ISGs) via the JAK/STAT pathway prior to activation and production of PKR, Mx or OAS antiviral activities.

1.1.4 RLR signaling pathways

Recent advances in structural biology have delivered several interesting structures of RIG-I, MDA5 and Mitochondrial antiviral-signaling protein (MAVS) and greatly increased our understandings of the molecular basis of RLR signaling pathway [3, 5, 65]. Activation of RLR signaling is a carefully regulated process [3, 32, 36, 66]. The auto-inhibitory conformation of RIG-I in resting state has been proposed based on the crystal structure of the *apo* duck RIG-I, in which the CARDs are trapped in an inactive conformation through the intra-molecular interactions between the CARD2 and HEL2i [67]. Leu109-bearing latch peptide in CARD2 and gate α -helical ridges containing Phe540 and Phe570 in HEL2i cooperatively contribute to this intra-molecular interface constructed via salt bridges and hydrophobic interactions. Reported crystal structures of 5'ppp duplex RNA bound mammalian RIG-I lacking tandem CARDs reveal consistent structural similarities that RIG-I encircles and collapses compactly on dsRNA backbone while keeps CTD 5'ppp moiety and ATP affinity pockets tightly sealed by 5'ppp moieties and ATP/ADP analogs respectively [43, 68, 69]. Recognition of RNA alone may not be sufficient to activate RIG-I and ATP binding has been suggested to be important for RIG-I to exit from the auto-inhibitory state by fully releasing the CARDs [8, 67, 68, 70-72]. Post-translational modifications including ubiquitination, phosphorylation and non-covalent K63-linked polyubiquitin binding to the CARDs have been reported to be essential for full activation of RIG-I [8, 71]. More recently, two new structures – a human RIG-I CARDs tetramer bound to poly-ubiquitin chains and a RIG-I CARDs-MAVS CARD complex - were determined, providing the molecular basis for MAVS fiber formation [65, 73, 74]. In contrast, MDA5 binds long RNA duplex cooperatively and form helical fiber-like mega-structures through its RNA sensing module, revealed by a combination of biochemical, biophysical, crystallographic, electron microscopic approaches [41, 44, 46, 47, 75]. Unlike the CARDs of RIG-I [5], the CARDs of MDA5 attached to the HEL-CTD: RNA complex through the long non-structured linker have not been

demonstrated to form high-order structure with functional implications. It has been shown that the non-covalent K63-linked polyubiquitin was also needed for MDA5 activation [6]. However, the activation mechanism of MDA5 is less known compared to RIG-I owing to limited crystal structures reported.

Upon sensing viral RNA, CARDS of RIG-I and MDA5 interact with the common adaptor protein MAVS (also known as IPS-1, VISA or CARDIF) located in the mitochondria [48] (Figure 1.3). MAVS is a membrane-bound, CARD-containing protein and its N-terminal CARD shares homology with RIG-I and MDA5 CARD1. Therefore MAVS is activated by homotypic CARD-CARD interaction. This process also involves ATP hydrolysis and binding of lysine-63 poly ubiquitin chains, which is functionally important to activate downstream transcription factors NF- κ B and IRF3 (interferon regulatory factor 3). Upon signal transduction, IRF3 is phosphorylated by kinases IKK ϵ or TBK1 and accumulates in the nuclei where it binds to target sequences to drive gene transcription. In contrast, NF- κ B activation requires the IKK complex-mediated phosphorylation of I κ B α , which is subjected to ubiquitin-dependent degradation by proteasomes. The activation of IRF3 and NF- κ B pathway could further lead to production of type I interferon (IFN) and inflammatory cytokines [7, 45, 49, 76-78].

Figure 1.3 The RIG-I/MDA5 signaling pathway. RIG-I and MDA5 recognize intracellular viral RNAs and undergo a series of conformational changes to interact with MAVS, a mitochondrial protein that functions downstream of RIG-I and MDA5. MAVS can further trigger NF- κ B and IRF3 signaling pathways by activating the IKK and TBK-1/IKK ϵ kinase complexes. Once activated, NF- κ B and IRF3 translocate into the nucleus and turn on the IFN- β gene promoter.

Figure 1.4 Proposed model of the molecular basis of the RLR signaling pathway. This model describes unanswered questions in RNA sensing and activation of RIG-I (A), MDA5 (B) and LGP2 (C). In the resting state, RIG-I adopts an auto-repressed conformation. Upon binding to RNA PAMP, RIG-I undergoes major conformational changes via its RNA sensing module CTD-HEL. RNA PAMP alone could not fully activate RIG-I and it requires ATP binding and hydrolysis to further release the signaling module CARDs. For MDA5, it has been shown to bind to the internal regions of long duplex RNA and oligomerizes cooperatively on long duplex RNA. Ubiquitination and the non-covalent binding of polyubiquitin chain to RIG-I CARDs are essential for RIG-I activation. MAVS resides on the outer membrane of mitochondria and is activated by the RIG-I and MDA5 CARDs via the formation of fiber-like oligomers prior to activation of IFN production.

Although tremendous efforts have been invested in better understanding the molecular mechanism of RLR activations, there are still several unanswered questions in the early events of RLR activation pathway that are listed as follows (Figure 1.4): (1) the hypothesis that human RIG-I apoenzyme adopts analogous CARD2-HEL2i intra-molecular interaction as its avian counterpart has never been demonstrated. (Q2 and Q4) The sequential activation of RIG-I CARDS following RNA and ATP remains unproven. (Q3) The conformation of polyIC bound RIG-I is unclear (Q5) The *apo* conformation of MDA5 remains elusive. (Q6) Recent structural studies using gel shift, chemical crosslinking, small angle x-ray scattering, x-ray crystallography, and electron microscopy methods [6, 41, 44, 46, 75, 79] have provided evidence to support the cooperative binding of MDA5 on long duplex RNA and, to a certain extent, elucidated the likely intermolecular interface between the HEL-CTD domains of MDA5. However, the detailed mapping of the interface is not available. (Q7) What is the function of ATP in MDA5 activation? (Q8, 9 and 10) The structural basis of LPG2 is little known in apoenzyme conformation as well as its RNA sensing events.

1.2 Hydrogen/Deuterium Exchange Mass spectrometry (HDX-MS)

There are several biophysical tools to probe single molecular protein conformational insights. Recent advances in the optical methods have greatly enhanced our understandings towards protein-nucleic acid interactions and their conformational changes involved in transcriptions and replications [80-82]. X-Ray crystallography and Nuclear Magnetic Resonance (NMR) spectroscopy remain as the main techniques to obtain high-resolution structures. However, flexible or highly dynamic regions of protein often elude crystal analysis. Higher order complex and large size proteins (> 50 kDa) always challenge and limit NMR rational designs. In addition to these available methodologies, hydrogen/deuterium exchange coupled with mass spectrometry (HDX-MS) has emerged as an effective

and sensitive way to study protein folding, ligand-protein interaction, protein intra-molecular and inter-molecular interactions in solution and it has been proved to gain complementary information accommodating static state structures [83-86]. The advantageous aspects of HDX-MS analysis in structural biology include compatibility with large size protein studies (i.e. > 50 kDa), low sample preparation requirement, and directly deciphering protein conformational dynamics in solutions.

1.2.1 Principle of HDX MS and its applications

There are three types of hydrogen atoms that are exchangeable to surrounding solvent, including O-H, N-H, and C-H groups (Figure 1.5). Among these different linkages, only backbone amide hydrogen (N-H) is exchangeable to deuterium with an exchange rate accessible for liquid chromatography mass spectrometry (LC-MS) analysis.

Figure 1.5 Backbone amide hydrogens for HDX-MS. Three types of hydrogen atoms that exist at in a peptide sequence, NDSCK. Hydrogens (red) at the backbone amide positions exchange at rates that can be measured by LC-MS; hydrogens (green) covalently bonded to carbon essentially do not exchange; hydrogens on the side chains exchange very fast and typically cannot be detected by LC-MS.

The kinetics of amide H/D in protein structure characterization is dependent on many factors. Solvent accessibility allows differentiating amide H/D exchange rate according to secondary structure of specific regions. For instance, regions containing α -helical and β -sheet exchanged to less extent of deuterated solvent compared to protein regions associated with terminus and loops. Also, regions involved in central core positions are inaccessible to deuterated solvent and thus expected to exhibit low H/D exchange kinetics whereas amide hydrogen atoms present in protein surface are readily exchangeable to deuterium at a fast rate. Amide hydrogen bonding activities are also governed by local fluctuations or protein stabilities [87]. Unstable regions of protein accompanying with local fluctuations exhibit higher deuterium uptake compared to stable regions or regions that fluctuate at a lower frequency. Therefore, the deuterium labeling induced mass shift and H/D exchange rate could reflect the protein conformational dynamics and hydrogen bonding activities of protein secondary structures [88, 89]. In recent decades, HDX-MS has been widely applied to important viral proteins. For instance, hexameric packaging motor of cystovirus enabled to bind viral RNAs through its RNA-binding channel coupled with ATPase activities. The conformational dynamics of P4 in the presence and absence of RNA have been examined by HDX coupled with mass spectrometry. The kinetics of HDX revealed distinctive states for different domains of P4 in response to nucleotide-binding, RNA loading and translocation as well as ATPase activities, and thus provided a comprehensive understanding to P4 molecular architecture in different biological states [90]. HDX also characterized subunit conformational dynamics of HIV-1 capsid protein from immature state to protein maturation [91]. Assembly of MS2 viral coat protein is initiated by binding with a RNA stem-loop, resulting in a conformational switch from a symmetric dimer to an asymmetric structure. In this circumstance, detailed structural information was characterized by HDX and mass spectrometry that some known RNA binding regions showed a more dynamic HDX kinetics [92]. Protein-protein interaction dynamics could also be studied by HDX and mass spectrometry. For instance, Kong et al. comparatively studied the

local conformational rearrangements within HIV-1 gp120 in the presence or absence of CD4 [93]. Monroe et al. also applied HDX combined with LTQ-FT mass spectrometry to study the immature, mature, and mutant Gag polyprotein to further unravel the capsid assembly [91]. As such, an intermolecular interaction was observed in the immature virions but not in the mature viral particles [91].

1.2.2 HDX experimental design

Amide hydrogen exchange is sensitive to temperature and pH conditions (Figure 1.6). Poly-D, L-alanine (PDLA) and poly-D, L-lysine (PDLL) have been used to examine HDX rate in respect to different temperature and pH conditions. Under pH 2.5 and 0 °C condition, this random polypeptide displayed minimum HDX kinetics [94].

Figure 1.6 pH and temperature affect amide hydrogen exchange rate (k_{ex}) [95]. (A) plots of the log of the k_{ex} versus pH. Both acid-catalyzed and base-catalyzed processes occur, and the rates of these two processes are minimal at approximately pH 2.5 (adapted from Smith et al., 1997). (B) Temperature affects k_{ex} and $t_{1/2}$ (the time when half the deuterium label is lost).

The general design of HDX-MS experiment is based on quench condition (0°C, pH 2.5). To initiate amide H/D exchange, purified protein samples are incubated with deuterated solvent for various time points at a designed temperature. After each HDX time point, on-exchange protein samples are quenched by fast decreasing the pH to 2.4 and temperature to 0 °C, which also denatures proteins at low pH conditions. The HDX of on-exchange protein samples for various time intervals are considered as “in-exchange” samples. Reversely, “back exchange” also occurs during sample analysis prior to detection by a mass spectrometer. To maximally limit back exchange, the whole HDX equipment system coupled with LC-MS is maintained at pH 2.4 and 0 °C. In addition, this set up also requires a low pH compatible protease to generate peptide fragments for MS identification and characterization. Pepsin and protease XIII are ideal acidic digestion enzymes capable for HDX studies [96]. These proteases are able to provide fast digestions within short time, yielding reproducible peptides for direct experimental comparisons under different conditions. Currently, there are two ways that incorporate acidic protease digestion into the HDX-MS system. Off-line pepsin digestion is one way to rapidly digest proteins by incubating protein with protease within short time prior to injecting into HDX-MS system (Figure 1.7 B). In addition, on-line immobilized pepsin digestion has been conveniently incorporated into reverse-phase LC system, which greatly enhances HDX analysis (Figure 1.7 C). In this way, yielded peptic peptides are readily identified by tandem MS (LC-MS/MS) generating a peptide pool with isotopic mass, retention time, and charge state recorded. For all the on-exchange and non-deuterated samples, raw MS spectra are recorded only in LC-MS mode. We have established house-built HDX on-line digestion system and applied it for our HDX studies (Figure 1.8).

Figure 1.7 HDX sample preparation and on-line, off-line pepsin digestion HDX-MS. (A) Amide hydrogen atoms of protein exchange to deuterium under native conditions prior to decreasing pH and temperature (pH 2.4, 0°C) (adapted from Hoofnagle et al., 2003) [97]. On-exchanged protein samples are proceeded to either off-line or on-line digestion by pepsin. (B) The design of off-line digestion system for HDX-MS. (C) The design of on-line system for HDX-MS.

Figure 1.8 House-built HDX-MS system incorporated with on-line digestion [98]. For capillary-flow LC, buffer A at pH 2.4 was H₂O containing 0.3% (v/v) formic acid. Buffer B at pH 2.4 was acetonitrile containing 0.3% (v/v) formic acid. Protein samples were then digested online by passing through an immobilized pepsin-coupled column (2.1 mm i.d. x 30 mm) (house-packed, immobilized pepsin purchased from Invitrogen) and were de-salted for 3 min on a house-packed C4 trap (0.75 mm i.d. x 10 mm, C4 beads purchased from Michrom). The mobile phase for on-line pepsin digestion was buffer A and the flow rate was 150 $\mu\text{l min}^{-1}$ driven by the LC loading pump (Dionex 3000 RSLC). A programmed gradient on a house-packed C4 column (0.3 mm i.d. x 50 mm, C4 beads purchased from Michrom) was used to separate and elute peptic peptides prior to MS analysis. All parts are connected by 1/16" OD x 50 μm ID PEEK tubing and the flow rate was 15 $\mu\text{l min}^{-1}$ driven by LC NC pumps (nano/capillary pumps). MS raw files were acquired in the range of m/z 300-2000 for 30 min in positive mode on a LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific). The injection valve and high-pressure switch valve are controlled by LC program via contact closure. All HDX system is buried on ice&water.

A number of software applications for HDX MS are growing recently to enhance HDX MS data analysis, such as HDX Workbench [99], HDEaminer (<http://www.massspec.com/HDEaminer.html>), and ExMS [100]. The efficiency and accuracy for determining large sets of amide hydrogen exchange MS data remains as a tedious process for the HDX community. The HD desktop (http://hdx.florida.scripps.edu/hdx_workbench/Home.html), developed by Scripps Florida, is fully automated software specifically designed to enhance HDX MS analysis with statistical significance.

1.3 Influenza A virus

Orthomyxoviridae family comprises the influenza A, B, and C viruses, *Thogotovirus* and *Isavirus*. One characteristic of this virus family is that they have negative-sense, single-stranded and segmented RNA genomes [101]. For instance, influenza A, B viruses have eight genomic RNA segments. Influenza C virus possesses seven RNA segments. Another characteristic of *Orthomyxoviruses* is that they rely on host cell nucleus rather than the cytoplasm for transcription and replication purposes [102].

The subtypes of influenza A virus are known to be distinguished by two viral glycoproteins present on the virus envelop, hemagglutinin (HA) and neuraminidase (NA). To date, there are sixteen types of HA and nine types of NA detected in wide aquatic birds and poultry [103]. Among these various subtypes, three HA and two NA are known to infect humans, such as H1N1, H2N2, H3N2, and H1N2 [104]. In addition, H5N1 and H7N9 are highly pathogenic to humans, causing pandemics in Asia [105, 106]. Influenza A viruses undergo continual antigenic drift and shift along their evolutionary lifespan. *Antigenic drift* is attributable to the accumulation of point mutations in the sequence of hemagglutinin and neuraminidase, resulting in changes of antigenic sites or epitopes that could escape the recognition of host immune system [107, 108]. One

of the reasons could be attributed to the intrinsically error-prone viral RNA dependant RNA polymerases, which could lead to spontaneous mutation rate at one nucleotide mutation per genome per replication [109]. The changes of these epitopes not only strengthen the virulent capabilities of the virus variants, but also desensitize the vaccine's protection that targets predominantly the variable regions of the hemagglutinin [110]. In addition, some influenza pandemics are associated with *antigenic shift*, which refers to the reassortments of influenza strains that co-infect the host cell therefore generating a novel virus strain [108]. The fact that the viral segmented genomes are independently imported into host nucleus for assembly could increase the possibilities to generate a genetically recombined virus. In this way, reassortments between hemagglutinin and neuraminidase subtypes are prone to occur. *Antigenic drift* and *shift* account for the diversity of repertoire of influenza A viruses.

1.3.1 Genome structure of influenza A virus

Influenza A virus has complex genome structures, which are comprised of eight distinctive genomic RNA segments (Figure 1.9). Viral genome segment 1 to 6 encode PB2, PB1, PA, HA, NP, NA, respectively. Segment 7 encodes M1 and M2 protein while segment 8 encodes NS1 and NS2 (NEP) due to the *gene splicing* mechanisms. PB1-F2 is encoded in the +1 reading frame of the PB1 gene segment. And each of these segments contains conserved non-coding regions at both the 3' and 5' ends [111].

Figure 1.9 The genome organization of influenza H1N1/PR/8/34 (adapted from Knipe, 2007).

Figure 1.10 Schematic representation of the structure of an influenza A virus particle (adapted from Nelson et al., 2007). And NS2 (NEP) present in the virus is not shown.

1.3.2 Influenza A viral proteins

Hemagglutinin (HA) and Neuraminidase (NA) are two internal membrane proteins of influenza A virus (Figure 1.10). The mature HA molecule is generated by proteolytic cleavage of precursor HA0, resulting in formation of two subunits HA1 and HA2 linked by a disulfide bond. HA2 region is more conserved than HA1 [112]. Unlike HA, the maturation of NA does not require proteolytic cleavage. Its N-terminus (amino acids 7-35) processes a hydrophobic sequence that serves as an anchor on virus surface [113]. Hemagglutinin is one of the most critical determinants for host adaptations. The non-permissive recognition of hemagglutinin to sialic acid receptors on different host cell membranes could restrict the host range specificity [114]. For instance, avian and human influenza virus could bind to the sialic acids NeuA α 2, 3Gal and NeuA α 2, 6Gal to initiate receptor binding for virus entry respectively [112]. In contrast, Neuraminidase could reversely catalyze sialic acids to facilitate the release and spread of progeny virions [115].

Matrix 2 protein (M2) is another internal membrane protein acting as a proton channel for equilibrating pH across the viral envelope. During virus entry, M2 has been reported to mediate an acidic condition to catalyze HA and endosomal membrane fusion [116]. In the late stage of posttranslational modifications of the internal proteins (HA, NA, M2), M2 has been shown to keep a suitable pH environment in trans-Golgi network (TNG) from being too low [117].

A Matrix 1 (M1) layer is formed beneath the viral envelope and is essential for virus membrane stability and integrity. Morphological evidences have revealed that M1 is implicated in the morphology of virus particles. It is favorable for the viral growth as well as generating different kinds of filamentous phenotypes conducive for virus assembly [118, 119]. In addition, M1 could mediate the bidirectional transport of vRNPs across the nuclear envelope via nuclear pore [120]. M1 plays a pivotal role in virus assembly and budding. It has been found to

directly interact with the cytoplasmic tails of two glycoproteins (HA, NA) and M2, organizing the assembly of complete virus particles [121-124].

Eight viral ribonucleoproteins (vRNPs) reside inside of influenza A virions, consisting of nucleoprotein (NP) and three polymerase proteins e.g. polymerase basic protein 1 (PB1), polymerase basic protein 2 (PB2) and polymerase acidic protein (PA). Nucleoprotein is a multifunctional viral protein. It contains conserved residues for RNA binding and keeps the formation of single stranded RNA [125]. Encapsulated by the nucleoprotein, viral RNA could be prevented from forming replicative intermediates (dsRNA). NP plays vital role in viral transcription and replication activities. Nuclear localization signals (NLS) on NP could assist vRNPs transportation to the nucleus, where transcription and translation occur. And serine-to-alanine substitution in the C-terminal of NP and its interaction with polymerase proteins suggest NP's role in regulating the switch from viral transcription to viral replication activities [126]. Three viral polymerase proteins concertedly play vital roles in initiating replication and transcription [127-129]. PB1 is known to be involved in the viral RNA chain elongation. PB2 acts as the major cap binding protein of the viral polymerase proteins via its conserved cap recognition pocket, which specifically target m7G of the cap [130, 131]. And the m7G bound structure of PB2 has been highlighted [132]. Also, PB2 is found to contribute to the replication and initial transcription activities of viral RNA [133]. PA is reported to bear endonuclease active sites and it is responsible for the cap snatching mechanisms, in which PA steals host mRNA cap to assemble it on viral RNA to initiate transcription [134]. Also, PA is required to associate with PB1 to transport from cytoplasm to the nucleus.

Besides vRNPs, nuclear export protein (NEP) is another viral protein located inside mature IAV particles. NEP bears a leucine-rich nuclear export signal (NES), which could mediate the nucleocytoplasmic export through the corporation with a cellular factor CRM1 [135]. The interaction of C-terminal of NEP with matrix 1 (M1) viral protein is also found to be involved in the nuclear export

process of vRNPs [136]. In addition, NEP is reported to directly inhibit transcription but facilitate translation activities during the virus infection [137].

Polymerase basic protein one-F2 (PB1-F2) is encoded by an alternative open reading frame of PB1 genome and it is known to be predominantly localized in the mitochondria of the some virus infected cells to initiate the apoptosis pathway. Also, PB1-F2 is a virulent factor that it could aggravate virus infection by enhancing the secondary bacterial inflammations [138]. Recently, N-terminal of PB1-F2 has been shown to interact with PB1 and increased expression levels of PB1 were observed, suggesting its role in regulating the vRNA polymerase activities [139]. However, the exact role of PB1-F2 is not well understood as it is found to be in strain-dependent and host-dependent manners. Like PB1-F2, non-structural protein 1 (NS1) is not associated with mature virus particles. It is mainly expressed in nucleus of virus infected cells. NS1 of influenza A virus bears nuclear localization signal (NLS), which could mediate its sub-cellular localization in host cell nucleus. Different influenza virus subtypes have distinctive nuclear localization signals, for instance, H1N1/WSN/33 has only one N-terminal NLS1 enriched of arginine or lysine while H3N2 virus subtypes have an additional NLS2 at the C-terminal [140, 141]. NS1 is a multifunctional viral protein, and it is known to accelerate the translation initiation of viral mRNA [142], blocking pre-mRNA splicing and transport [143], and inhibiting the host immune surveillances, such as type I interferon (IFN) system and dendritic cell (DC) activation [144, 145].

1.3.3 Influenza A virus infectious cycle

Virus attachment and entry is the initial step in which HA of influenza viruses recognizes and specific receptors present on the host cell surfaces. Sialic Acid (N-acetylneuraminic acid) is a α -keto acid with a nine-carbon backbone and it is ubiquitously distributed at the surface of all vertebrate cells [146, 147]. Carbon-2 of sialic Acid could associate with carbon-3 or carbon-6 of galactose to form SA

2,3 Gal linkages or SA 2,6 Gal linkages, which are predominantly regarded as two specific receptors for avian influenza virus and human influenza virus, respectively [148]. The receptor-bound virus enters the cell by receptor-mediated endocytosis. Majority of influenza viruses enter the infected cells via clathrin-mediated endocytosis whereas some evidences illustrated that influenza viruses could also infect cells in a clathrin-independent pathway [149]. During the early entry, virus infection could activate the extracellular signal-regulated kinase (ERK) and phosphatidylinositol 3-kinase (PI3K) pathways in an ATP-dependent manner, which results in the acidification of endosome [150]. Membrane fusion requires proteolytic cleavage of HA [151]. The low pH (pH 5~pH 6) inside of the endosome triggers the irreversible conformational changes of HA to favor viral and endosomal membrane fusion activities [112, 152, 153]. In addition, M2, an ion channel protein, assists accelerate acidification of the interior of the virions by conducting the protons across the virus envelope [117]. This acidification induces depolymerization of M1 layer on viral membrane and dissociation of M1 from vRNPs [154]. Influenza virus replication and transcription activities are nucleus-dependent. Released vRNPs could utilize the host cell nuclear import machinery to efficiently transport into the nucleus via nuclear pore complexes. For instance, importin α protein family has NLS binding sites with NP and viral polymerase subunit PB2 and they could mediate the nucleus import process [155, 156]; Hsp90 is found to interact with viral polymerase proteins PB1 and PB2, assisting the assembly and nuclear transport of viral polymerase proteins [157]; The association between Ran binding protein 5 (RanBP5) and PB1-PA dimer also contributes to the nuclear import of viral RNAs [158].

IAV transcription and replication occur in distinctive ways [101]. In transcription mechanisms, the binding between 5'- terminal of vRNA and viral polymerase could activate the cap-binding sequence on PB2; subsequently, the binding between the 3'- terminal of vRNA and PB1 activates the endonuclease activity [131]. These bindings could result in stealing the capped primer from the

host pre-mRNA molecules and initiating the mRNA synthesis. Viral mRNA is synthesized by viral RNA-dependent RNA polymerase using negative sense vRNA as template, the process of which requires the capped 5'- primer snatched from the host pre-mRNA molecules, elongation factor PB1 as well as polyadenylation at the 3'- end. In contrast to the transcriptional activities, viral genome replication is neither cap-dependent nor polyadenylated as mRNA. The positive cRNA is generated via direct copying of vRNA and it is used as templates for large production of vRNA. In addition, the host factors also participate and contribute to the procedures of the viral transcription and translation [159], such as α -Importins [160], Heat Shock Protein 70 [161], eIF-4G [162].

Viral mRNAs are expressed in the cytoplasm after the viral transcription, from where the newly synthesized viral proteins, e.g. NP, viral polymerase proteins, are imported to the nucleus to form new vRNPs. The newly formed M1 and NEP are also transported back to the nucleus for assisting the nuclear export of vRNPs. One model has been proposed that NEP could conduct this nuclear-cytoplasm transport by interacting with nucleoporins [163]. Shimizu et al. found that the C-terminal of NEP is essential to bind with M1 as well as guiding the export of vRNPs since the dissociation of M1 from vRNPs can inhibit this effect [136, 164]. Neumann et al. discovered that NEP could mediate export of vRNPs in a CRM1-dependent manner via forming a ternary export complex with CRM1 and Ran-GTP [165].

Virus assembly and budding are the last steps of viral infectious circle and influenza virus assembly and bud from the plasma membrane. HA and NA are known to utilize the exocytic pathway to be transferred from trans-Golgi network (TGN) to the lipid-raft associated plasma membrane. Specifically, ectodomain or transmembrane domain (TMD) of NA and HA contains determinants or signals for apical sorting, which could target to apical plasma membrane or lipid rafts [166, 167]. M2 undergoes the similar pathway but not involved in the raft-derived microdomains [168]. M1 plays a key role in virus assembly and it could

independently form the virus-like particles and then bud from the cells [121]. M1 has been observed to interact with HA cytoplasmic tail and transmembrane domain during the HA transportation to the plasma membrane [169]. However, the detailed mechanism of how M1-vRNP-NEP complex transfers to the assembly sites from the cytoplasm remains elusive. IAV also utilizes the host components to aid its budding and release. Lipid rafts are enriched in sphingolipids and cholesterol and are abundantly associated with the plasma membrane, late-Golgi compartments and endocytic membranes [170]. These specialized domains are known to be related with infectious cycle of many enveloped viruses, such as lentivirus [171], flavivirus [172], and respiratory syncytial virus [173]. For viral proteins of influenza A virus, HA but not M2 is organized into clusters into this micro-membrane domain, which also concentrates NA [174]. Also, lipid raft is reported to aid the transportation of NP specifically targeting to the apical plasma membrane [175].

1.4 Proteomics and applications on virology

The conjunction of mass spectrometry based proteomics with realms of virology exhibits a multidisciplinary effort to enhance our understanding into a wide range of viral diseases. Several approaches have been used together with liquid chromatography (LC) separation followed by either electrospray ionization (ESI) or matrix assisted laser desorption/ionization (MALDI) tandem mass spectrometry to study protein-protein interactions. These techniques, namely tandem affinity purification (TAP), immunoprecipitation (IP), and quantification by stable isotope labeling, have been increasingly advanced to put perspectives into the interactions between viruses and hosts at the protein level (Table 1.2).

Table 1.2 Reprehensive proteomics applications to selected viruses

<i>Virus family</i>	Other virus	Methods	Reference
<i>Arenaviridae</i>	pichindé virus	SELDI-TOF MS, tandem MS	[176]
<i>Baculoviridae</i>	<i>autographa californica</i> nucleopolyhedrovirus	MUDPIT, 2D-LC MS/MS	[177]
<i>Coronaviridae</i>	mouse hepatitis virus	SILAC, LC MS/MS	[178]
<i>Flaviviridae</i>	bovine viral diarrhea virus	2D-LC MS/MS	[179]
	west nile virus	2D DIGE, LC MS/MS	[180]
	swine fever virus	2D DIGE, MALDI-TOF MS/MS	[181]
<i>Hepadnaviridae</i>	hepatitis B virus	2D gel, MALDI TOF/TOF MS; iTRAQ, 2D-LC MS/MS	[182-184]
<i>Hepeviridae</i>	hepatitis E virus	CO-IP, MALDI-TOF MS	[185]
<i>Herpesviridae</i>	epstein-barr virus	LC MS/MS	[186]
	pseudorabies virus	SILAC, MALDI TOF/TOF MS	[187]
	human cytomegalovirus	2D-LC MS/MS	[188]
	Marek's disease virus	2D gel, LC MS/MS; 2D gel, MALDI- TOF/TOF	[176, 189]
	murine cytomegalovirus	LC MS/MS	[190]
<i>Iridoviridae</i>	iridovirus	iTRAQ, LC MS/MS	[191]
	chilo iridescent virus	LC MS/MS	[192]
	tiger frog virus	2D gel, MALDI-TOF MS	[193]

<i>Nimaviridae</i>	white spot syndrome virus	2D gel, MALDI-TOF/TOF MS	[194]
<i>Paramyxoviridae</i>	measles virus	HLA purification, nanoLC MS/MS	[195]
<i>Poxviridae</i>	myxoma virus	MALDI-TOF MS	[196]
	fowlpox virus	2D gel, MALDI-TOF MS	[197]
	vaccina virus	MALDI-TOF MS	[198]
<i>Retroviridae</i>	simian immunodeficiency virus	MALDI-TOF MS	[199]
	murine leukemia virus	MALDI MS and LC MS	[200, 201]
	reovirus	2D DIGE, MALDI TOF/TOF MS	[17]
<i>Rhabdoviridae</i>	rabies virus	2D gel, LC MS/MS	[202]
<i>Roniviridae</i>	yellow head virus	2D gel, MALDI-TOF MS, nanoLC MS/MS	[203]

1.4.1 Proteomic searching engines

With the advent of high throughput short-gun proteomics, liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) could generate numerous fragment ion spectra and thus is able to identify large number of peptide sequences. Presently, several proteomic search engines are available from distinctive searching algorithms and compatible with mass spectrometric data, including some traditional ones, e.g. Mascot, SEQUEST, X!TANDEM, as well as some new search engine, e.g. ProluCID, etc. Mascot incorporates peptide mass

fingerprint, sequence query, and MS/MS ions search. It is a possibility-based scoring engine by calculating the observed match between experimental data and theoretical sequence data [204]. Although Mascot provides high throughput protein identification dependent on probability rankings, there exist some limitations associated with searching non-independent dataset, and non-statistical validation of atypical sequence entries. SEQUEST could correlate ion fragmentations in the processed tandem mass spectrometric data with their corresponding amino acid sequences in the FASTA database files. There are generally four steps of this search algorithm, including tandem mass spectra reduction, matching spectra with amino acids, generating high-ranked sequences, and correlating with protein identification [205, 206]. SEQUEST also is able to search several covalent modification-bearing peptides by matching the nascent tandem mass spectra [207]. X!TANDEM is an open-source platform for proteomic researchers to efficiently process MS/MS data [208]. Its analysis on a mixture of peptides is based on one axiom: for each detectable protein in the original protein mixture, there will be at least one good tryptic peptide match within a designed scope. In the first step of the analysis, a smaller set of protein sequences is generated from the original protein database by thoroughly filtering with the designed scope that set as small as possible. From this step onwards, the subsequent searches are within this refined protein sequences, thus improving the efficiency and saving the overall search time. In the second step of analysis, a bigger scope is set to perform multiple comparisons of the spectra with those refined protein sequences in respect to the different peptide modifications, number of missed cleavages, and non-specific hydrolysis, etc [209]. And global proteome machine (gpm) is a well-established open source search engine based on TANDEM (<http://www.thegpm.org/>). Recently, Xu et al. developed a new search engine ProLuCID, which is based on binomial probability preliminary scoring scheme to filter candidate peptides for further isotopic distribution analysis [210, 211]. And high sensitivity and specificity of ProLuCID could be achieved compared to SEQUEST.

In addition to these searching algorithms, there emerge several multifunctional proteomic pipelines that could incorporate statistical analysis and guarantee high confident searching results. Since different search engines have distinctive algorithms and sensitivities, some important low abundant hits may be identified by only few software. To cope with this problem, one advantage of Scaffold proteomic pipeline is that peptides simultaneously identified by several different searching engines, e.g. Mascot, SEQUEST, TANDEM, etc, could be integrated into “a folder” by Peptide Prophet Algorithm, resulting in a list of combined peptide sequences. And a further statistical calculation and validation by Protein Prophet algorithm is processed with MS/MS data to generate protein identifications [212]. Similarly, Trans-Proteomic Pipeline (TPP) also supports the original data generated by Mascot, SEQUEST, TANDEM, etc. Peptide or protein identification is performed by peptide prophet or protein prophet algorithm, respectively. It also combines the advantages of statistical validations by iProphet tool and quantitative analysis by XPRESS, ASAPRatio, or Libra algorithm [213]. Moreover, the Scripps Research Institute developed an integrated Proteomics Pipeline (IP2), which provides a simple and efficient platform for rapid identifications and quantifications to proteomic researchers (<http://integratedproteomics.com/>). It is compatible with both SEQUEST and ProLuCID search engines for high resolution mass spectrometry spectra analysis [210]. Subsequently, DTASelect set spectrum filtering parameters to ensure low false positive rate and refine the confidence of protein output [214, 215]. Census is further incorporated into IP2 to enable large-scale quantitative analysis on isotope-labeled, e.g. N15, SILAC, and iTRAQ, or label-free samples [216]. Some well-established post translational modifications, especially phosphopeptides, could also be specifically searched during the IP2 ProLuCID analysis step, providing clues for further functional validations.

1.5 Objectives and organization of this dissertation

The main objective of this Ph.D project is to study and provide new evidences of the early events associated with the activation of RIG-I like receptors. And we could eventually depict a detailed step-by-step view of the early events during RLR activation. Towards this goal, this dissertation is divided into the following chapters.

Chapter 1 reviews the “non-self” RNA recognition features of RLRs and their roles in innate immune responses. In addition, the mainstream technique - Hydrogen/deuterium exchange coupled with mass spectrometry (HDX-MS) - is introduced in this chapter. Furthermore, as a side project of proteomics studies on influenza A virus, the current knowledge of influenza A virus is also reviewed within this chapter.

Chapter 2 presents the in depth analysis of RNA PAMP:RIG-I:ATP interactions by HDX-MS. Within this chapter, our data provided the first in-solution evidence that full-length human RIG-I adopts an auto-inhibitory conformation via CAR2-HEL2i intra-molecular interaction. We also recorded the unique allosteric effects upon sequential addition of RNA PAMPs (3p10L and polyIC) and ATP during the RIG-I CARDs activation process. In addition, the HDX-MS profiles of 3p10L bound LGP2 suggest that it resembles 3p10L bound RIG-I Δ CARDs.

Chapter 3 discusses the activation mechanism of MDA5 by HDX-MS. The *apo* state of MDA5 remains mysterious. And the immunology field has been puzzled by the activation process of MDA5 since its discovery. Our work provided the first direct structural evidence that the *apo* state of MDA5 is radically different from RIG-I and is absent of auto-inhibition. Then, we showed that MDA5 oligomerizes on polyIC in solution by mapping the cooperative binding surfaces on the MDA5 HEL-CTD and present a high-resolution model of a cooperative MDA5-RNA complex. Unlike MDA5, the HDX-MS study showed that LGP2

doesn't oligomerize on polyIC but maintains tight interactions with polyIC via its HEL-CTD domains.

Chapter 4 is devoted to a proteomics study on host factors incorporated into influenza A virions. Proteomic analysis of exogenous protease treated influenza A viruses identified five host factors that reside within influenza A viral particles. Western blot, immunofluorescences and knockdown studies were employed to validate the association of Annexin A2 with viral NP in the virus infected DF-1 cells.

Chapter 5 concludes all the above studies and suggests future works.

CHAPTER 2

HDX-MS characterization of RIG-I activation mechanism

2.1 Introduction

RIG-I belongs to RIG-I like receptors (RLRs) and plays a major role in innate immunity against virus infection. RIG-I undergoes conformational transitions upon encountering viral RNA and harnesses ATP hydrolysis to switch on mechanical rearrangements, mediating a downstream signaling relay leading to interferon- α/β (IFN- α/β) production [98]. RIG-I adopts tandem CARDs at its N-terminus, a C-terminal domain (CTD) and a core-helicase domain, which consists of ATP hydrolysis pocket and two rigid RecA-like domains (HEL1 and HEL2) with a specialized insertion domain within HEL2 (HEL2i) [32]. Recent endeavors in structural biology have greatly increased our understandings of RIG-I mediated signaling events. The crystal structure of *apo* duck RIG-I suggests that the CARDs are trapped in an inactive conformation through intra-molecular interactions between CARD2 and HEL2i [67]. Recognition of RNA alone may not be sufficient to activate RIG-I and thus ATP binding has been suggested to play a role for RIG-I to exit from the auto-inhibitory state by providing energy to fully release the CARDs [4, 6, 8, 43, 67, 68, 217]. Post-translational modifications including ubiquitination, phosphorylation and non-covalent K63-linked polyubiquitin binding to the CARDs have been reported to be essential for full activation of RIG-I [8, 71]. More recently, two new structures - a human RIG-I CARDs tetramer bound to poly-ubiquitin chains and a RIG-I CARDs-MAVS CARD complex - were determined, providing the molecular basis for MAVS fiber formation [65, 73, 74]. Nevertheless, full-length RIG-I has resisted crystallographic efforts and a molecular description of its activation pathways remains hypothetical.

LGP2 differs from RIG-I by the lack of CARDs signaling domain but shares HEL and CTD domains with RIG-I. The structure of full length LGP2 bound RNA has not been reported. The CTD domain of human LGP2 has been refined and it has been characterized to bind the termini of dsRNA, suggesting that the function of LGP2 CTD is closer with RIG-I CTD rather than MDA5 CTD domain, which binds to the internal part of dsRNA [218]. Due to a lack of the signaling domain, LGP2's precise role in antiviral immunity remains to be fully elucidated. Characterization of LGP2 in the presence of short dsRNA could greatly assist to refine its structural context and comparatively explain the mechanisms of RLR signaling regulations.

Hydrogen/deuterium exchange mass spectrometry (HDX-MS) provides a way to study the large size RLR *apo* enzymes, higher-order RNP complexes and the structure dynamics driven by ATP hydrolysis. First, we chose biologically relevant RNA ligands – short 5' triphosphorylated 10 base pair hairpin RNA and long duplex RNA polyIC – as RNA PAMPs to activate RIG-I and MDA5 and performed HDX-MS to visualize the activation process. Next, using our functional HDX experiment, we recorded the unique allosteric effects upon sequential addition of RNA PAMPs and ATP during the RIG-I CARDs activation process. Moreover, to reveal RNA binding state of 3p10L bound LGP2 and compare it with RIG-I or RIG-IΔCARDs, we performed HDX studies on *apo* LGP2 and 3p10&LGP2. We were able to collect high-resolution and high-coverage data across seven HDX time points for various conformational states of full-length and truncated RIG-I. Direct comparisons of these time-resolved HDX-MS snapshots across *apo*, RNA bound and ATP-stimulated enzymes enable us to dissect the early activation steps of RIG-I signaling cascade. Based on this work and in the context of previous biochemical and structural studies on RIG-I, we present a detailed side-by-side view of the early events during its activation.

2.2 Materials and methods

2.2.1 Materials

The plasmids that encode human RIG-I, human RIG-I Δ CARDs, mouse LGP2 were provided by Dr. Luo Dahai lab [4]. pET-SUMO, SUMO protease, TEV protease were purchased from Invitrogen. All restriction enzymes, ligases and T7 polymerase were purchased from New England Biolabs. All PCR reagents were purchased from Invitrogen. The competent Rosetta II (DE3) *Escherichia coli* cells were purchased from Novagen. Ni-NTA beads were purchased from QIAGEN. All the chemicals were purchased from Sigma-Aldrich. All FPLC system was purchased from GE Healthcare. Pepsin column was purchased from Invitrogen. Two external switch valves were ordered from VICI instruments. 1/16" OD x 50 μ M ID PEEK tubing was purchased from uppurge. C4 magic beads were purchased from Michrom. Liquid chromatography (Dionex 3000 RSLC) and LTQ-Orbitrap mass spectrometer were purchased from Thermo Fisher Scientific (Singapore).

2.2.2 Cloning, Expression, and Purification of RIG-I, RIG-I Δ CARDs, and LGP2

The full-length RIG-I and N-terminal CARDs (1-238) deletion constructs, hereafter named RIG-I (Δ CARDs 1-238), was cloned into the pET-SUMO vector (Invitrogen). Mouse LGP2 construct was cloned into the pNIC28-Bsa4 plasmid. The purification procedures of RIG-I constructs and LGP2 are similar except that SUMO protease is used to cleave the SUMO tag on MDA5 proteins and LGP2 is cleaved by TEV protease for tag removing. Transformed Rosetta II (DE3) *Escherichia coli* cells (Novagen) were grown at 37°C in Luria broth medium supplemented with 40 μ g ml⁻¹ kanamycin and 34 μ g ml⁻¹ chloramphenicol to an OD_{600nm} of 0.6–0.8. Protein expression was induced at 18°C by adding isopropyl- β -

D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. After 20 hr growth, cells were harvested by centrifugation at $8,000 \times g$ for 10 min at 4°C and stored at -20°C . Cells resuspended in buffer A (25 mM HEPES [pH 8.0], 0.5 M NaCl, 10 mM imidazole, 10% glycerol, 5 mM β -ME) were lysed by passing three times through a MicroFluidizer at 15,000 psi and the lysate was clarified by centrifugation at $15,000 \times g$ for 60 min at 4°C . The supernatant was purified by batch binding with QIAGEN Ni-NTA beads. The beads were collected in Biorad polyprep columns and the SUMO-tagged proteins were eluted with buffer B (25 mM HEPES [pH 8.0], 0.3 M NaCl, 10% glycerol, 5 mM β -ME, 200 mM imidazole). The fraction containing His₆-Sumo-RIG-I was then digested with ulp protease (Invitrogen), 4°C overnight. His-LGP2 was then digested with TEV protease overnight at 4°C . The cleavage mixture was loaded onto a HisTrap HP column to remove the His₆-Sumo protein and ulp protease from the mixture. The recombinant protein was then further purified by using a HiTrap Heparin HP column (GE Healthcare) by running buffer C with an additional 1 M NaCl gradient. Concentrated proteins were subjected to a final gel-filtration purification step through a HiPrep 16/60 Superdex 200 column (Amersham Bioscience) in buffer D (25 mM HEPES [pH 7.4], 150 mM NaCl, 2 mM MgCl_2 , 5% glycerol, 5 mM β -ME). Fractions containing monomeric RIG-I were pooled, concentrated, and stored at -80°C . Recombinant protein RIG-I (Δ CARDs: 1-238) was expressed and purified using the same method. The concentrations of the proteins were determined by measuring the absorbance at 280 nm by using extinction coefficients of $95,300 \text{ M}^{-1} \text{ cm}^{-1}$ for full-length RIG-I and $60,040 \text{ M}^{-1} \text{ cm}^{-1}$ for RIG-I (Δ CARDs: 1-238). The purity for each protein was $> 95\%$ as verified on SDS-PAGE gel.

2.2.3 RNA Work.

As described previously [27], 5' triphosphorylated RNA hairpin (3p8L: 5'pppGGCGCGGCUUCGCCGCGCC, 3p10L : 5'pppGGACGUACGUUUCGACGUACGUCC) were transcribed with T7 polymerase and purified on 20% denaturing polyacrylamide gels. LMW polyIC (Invivogen) was dissolved in buffer containing 25 mM Hepes (pH 7.4), 150 mM NaCl, 5% glycerol, 5 mM β -ME to a final concentration of 10 mg/ml. Concentrations were determined spectrophotometrically.

2.2.4 Hydrogen/Deuterium Exchange

The buffer for HDX on-exchange was the same composition (25 mM HEPES [pD 7.9], 150 mM NaCl, 2 mM MgCl_2 , 5 mM β -ME) except that H_2O was replaced with D_2O (99.99%) and glycerol was removed. Specifically, we added 1 mM ATP into D_2O buffer (ATP& D_2O buffer) for ATP hydrolysis studies associated with all proteins. All samples with and without RNA ligands were incubated on ice for 1 hr prior to HDX. 4 μls of 18 μM RIG-I or RIG-I Δ CARDs in the presence or absence of RNA ligand (100 μM 3p10L or 5 mg/ml polyIC) were mixed with 16 μls D_2O buffer (final D_2O concentration was 80%) or H_2O buffer for 0s samples, and incubated at 4 $^\circ\text{C}$ at various time intervals for on-exchange, e.g. 10s, 30s, 60s, 300s, 900s, 3600s. To study ATP hydrolysis, 4 μls 18 μM of RIG-I or RIG-I Δ CARDs with respective RNA ligands (100 μM 3p10L and 5 mg/ml polyIC) were incubated on ice for 1 hr before mixing the RNA&protein complex with 16 μl ATP& D_2O or ATP& H_2O buffer for 0s samples, and incubated at 4 $^\circ\text{C}$ at the same time intervals. Ice-cold 20 μls quench solution consisting of 1 M guanidine hydrochloride and 1.5 % (v/v) formic acid were added to each on-exchanged samples after specific time intervals prior to rapid freezing by liquid Nitrogen until HDX MS analysis. For fully deuterated control samples, 4 μls of 18 μM RIG-I

(± 200 μM 3p10L) were incubated with 16 μls D_2O solvent at 37 $^\circ\text{C}$ for 20 hr. The following steps were performed the same as other HDX time point samples. Three replicates were performed for each HDX time point.

For RNA ligand-RIG-I activation studies, 4 μls of 18 μM RIG-I were incubated with 100 μM 3p8L or 100 μM 3p10L or 5 mg/ml polyIC on ice for 1 hr prior to HDX experiment. 4 μls of each sample were then mixed with 16 μls of D_2O buffer for 1 hr, quenched, and analyzed by LC-MS. Ice-cold 20 μls quench solution consisting of 1 M guanidine hydrochloride and 1.5 % (v/v) formic acid were added to each on-exchanged sample after specific time intervals prior to rapid freezing by liquid Nitrogen until HDX MS analysis. For ATP analogue studies, 4 μls of 18 μM RIG-I were incubated with 100 μM 3p10L on ice for 1 hr. Stabilized RIG-I&3p10L complex was then incubated with or without ATP/ADP/ADP-AlF_x (final concentration 1 mM in RIG-I&3p10L complex) for another hr to reach an equilibrium conformation prior to the HDX experiment. 4 μls of RIG-I&3p10L&ATP/ADP/ADP-AlF_x (18 μM RIG-I, 100 μM 3p10L, and 1 mM ATP/ADP/ADP-AlF_x in final concentration) were then incubated with 16 μls D_2O buffer for 5 min, quenched, and analyzed by LC-MS. Ice-cold 20 μls quench solution consisting of 1 M guanidine hydrochloride and 1.5 % (v/v) formic acid were added to each on-exchanged sample after specific time intervals prior to rapid freezing by liquid Nitrogen until HDX MS analysis. The following steps were performed the same as other HDX time point samples. Three replicates were performed for each HDX time point.

2.2.5 LCMS under quench condition

For capillary-flow LC, buffer A was H_2O containing 0.3% (v/v) formic acid, and buffer B was acetonitrile containing 0.3% (v/v) formic acid. Protein samples were then digested online by passing through an immobilized pepsin-coupled column (2.1 mm i.d. x 30 mm) (Invitrogen) and were de-salted for 3 min on a

house-packed C4 trap (0.75 mm i.d. x 10 mm, C4 beads purchased from Michrom). The mobile phase for on-line pepsin digestion was buffer A and the flow rate was $150\ \mu\text{l min}^{-1}$ driven by the LC loading pump (Dionex 3000 RSLC). A 20 min gradient on a house-packed C4 column (0.3 mm i.d. x 50 mm, C4 beads purchased from Michrom) was used to separate and elute peptic peptides prior to MS analysis. All parts are connected by 1/16" OD x 50 μM ID PEEK tubing and the flow rate was $15\ \mu\text{l min}^{-1}$ driven by LC NC pumps. The gradient started from 5 % buffer B and increased to 35 % buffer B within 20 min, followed by washing with 90 % buffer B for 3 min and equilibration with 1 % buffer B for 5 min. MS raw files were acquired in the range of m/z 300-2000 (with mass resolution set to 100,000 at m/z 400) for 30 min in positive mode on a LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific) equipped with an ESI source (capillary temperature $275\ ^\circ\text{C}$ and spray voltage of 5 kV). All the HDX systems were strictly performed at $0\ ^\circ\text{C}$ (fully buried on ice & water) and on-line pepsin digestion was carried out at $16\ ^\circ\text{C}$. Blank injections were made between every two samples to remove carryover peptides. Data for each time point were repeated three times. All HDX data were normalized to 100 % D_2O content, corrected for an estimated average deuterium recovery of 70 %, and analyzed by the software HDX Workbench [219]. The HDX Workbench has built-in statistical software to evaluate the statistical significance using a one-way ANOVA and Tukey's multiple comparison tests. Initial peptic peptide identifications were performed with the same HDX set up as described above. Four microliters of protein sample (20 μM) were injected into the HDX MS system. Product ion (MS/MS) spectra were acquired in linear ion trap LTQ with eight most abundant ions selected in the precursor (MS) scan with a 7.5 sec exclusion time. MS and tandem MS files were extracted and searched by using Integrated Proteomic Pipeline (<http://integratedproteomics.com/products/ip2/>) for high-confident peptide identification (searching parameters: Precursor delta mass cutoff 20, Best peptide FP threshold 0.01, Best peptide delta mass threshold 10).

2.2.6 Structure modeling

The full length *apo* human RIG-I structure was modeled based on the *apo* duck RIG-I structure (PDB code: 4A2W) and the human RIG-I CTD structure (PDB code: 3LRR). The homology modeling program is HHPred (<http://toolkit.tuebingen.mpg.de/hhpred>) [220].

2.3 Results

2.3.1 Optimization of RIG-I protein expression and purification

Human RIG-I is a large-size protein with a molecular weight of 106 kDa. Several RIG-I constructs have been cloned and expressed to examine its soluble expression levels in *E.coli*, including a poly-histidine tag containing plasmid PNIC28-bsa4, a GST-tag containing plasmid PGEX-4T-1, a NUSA-tag containing plasmid PET50b, and a SUMO-tag containing plasmid PET-SUMO. Recombinant protein expression of RIG-I cloned in PNIC28-bsa4 was found to be highly insoluble when expressed in BL21 (DE3). The recombinant protein was mainly expressed as insoluble inclusion bodies and only modest amount was found in the soluble form. Although there was considerable amount of GST-RIG-I expressed as soluble form, poor binding efficiency of GST-RIG-I to glutathione magnetic beads resulted in large amount of soluble protein that could not be concentrated and further purified (Figure 2.1A). The above two RIG-I constructs failed to yield RIG-I monomeric protein, despite extensive efforts to clone these constructs and optimize the expression conditions. I also tried recombinant protein expression of NUSA-RIG-I in *E.coli*, which yielded high expression levels of soluble form NUSA-RIG-I with a double His-tag that binds well to Ni²⁺NTA beads. However, more than half of the proteins were lost during the overnight cleavage of NUSA tag (Figure 2.1B). Cleaved full-length RIG-I was prone to protein aggregation during the following FPLC purification steps.

Figure 2.1 SDS-PAGE gels and size-exclusion chromatogram of selected recombinant proteins. (A) SDS-PAGE of purification of GST-RIG-I (CL: Cell lysate S: supernatant P: inclusion bodies W1: 20mM glutathione in washing buffer W2 50mM glutathione in washing buffer E: elution buffer with 200mM glutathione M: protein marker) (B) SDS-PAGE of purification of NUSA-RIG-I (P: inclusion bodies CL: cell lysis FT: flowthrough after Ni^{2+} NTA beads W1: 20mM Imidazole in washing buffer W2: 50mM Imidazole in washing buffer W3:100mM Imidazole in washing buffer E: 500mM Imidazole in washing buffer M: protein marker). (C) Size-exclusion chromatogram of monomeric RIG-I (106 kDa).

SUMO-RIG-I construct is advantageous over above RIG-I constructs given to its high level of soluble expression, efficient protein affinity purification, and efficient protease digestion for removal of SUMO-tag to generate monomeric full-length RIG-I. The soluble fraction collected from Ni^{2+} NTA column was preceded with 3 hr protease digestion (SUMO protease). This protease digestion enables to

cleave 99% of Sumo-RIG-I. The digested products were loaded back to His-trap column (GE) to remove cleaved SUMO-tag and incompletely cleaved SUMO-RIG-I. The flowthrough was further purified by heparin column to remove bacterial RNA contaminants prior to a S200 size exclusion column for further purification (Figure 2.1C). The monomeric form of RIG-I was collected and concentrated.

2.3.2 Human RIG-I maintains the auto-inhibitory conformation through the intra-molecular CARD2:HEL2i interactions

To examine whether mammalian RIG-I adopts an auto-inhibitory conformation in solution similar to that observed in the static duck RIG-I structure, we performed HDX experiments with recombinant human RIG-I protein with and without ATP and RNA substrates. We also included truncated RIG-I (RIG-I Δ CARDs, residues 229-925) as a control, which lacks the N-terminal CARDs and the linker region to the helicase domain (Figure 2.2A). We used the 5' triphosphorylated 10mer hairpin RNA (3p10L) as the RNA ligand, which has recently been shown to bind and activate RIG-I *in vitro* and in cell-based IFN production assays [27]. In total, 221 peptides from RIG-I and 194 peptides from RIG-I Δ CARDs were identified and characterized across all HDX time points, representing a high sequence coverage of 89% for RIG-I and 86% for RIG-I Δ CARDs, respectively (Appendix A, B and F).

In the *apo* state, the latch peptide Y103-114 on RIG-I CARD2 exhibited the greatest magnitude of protection to HDX with less than 30% deuterium uptake after a 1 hr incubation in deuterated solvent. In contrast, in the presence of 3p10L, this protection was greatly compromised with deuterium incorporation of up to 70% during the same time interval (Figure 2.2B). Correspondingly, a 15% increase in average deuterium uptake was observed for this latch peptide at all HDX time

points (Table 2.1). This RNA dependent deuterium uptake was also clearly observed in three adjacent CARD2 peptides, such as K108-126 (9%), I139-155 (11%) and L160-185 (9%) (Table 2.1). These peptides reside next to the outermost helical latch peptide and face away from the RNA binding motifs in HEL-CTD. The CARDS are known to be rigid α -helical bundles with highly stable secondary structural compositions [5, 221]. Therefore, the significant differences in deuterium uptake in CARD2 peptides suggest that CARD2 in *apo* RIG-I contains a relatively solvent-inaccessible surface patch, which becomes more solvent-exposed in the presence of RNA ligand. Importantly, the sequences of the aforementioned CARD2 peptides also lie at the interface between CARD2 and HEL2i in duck RIG-I [67].

RNA-induced deuterium uptake was also observed in the human HEL2i α -helical peptide Y566-574 in the full length RIG-I protein. Specifically, a loss of protection (a 5 % increase in average deuterium uptake at all HDX time points) for the gate motif Y566-574 was observed upon 3p10L RNA binding. In contrast, the same surface patch in the RIG-I Δ CARDS showed statistically insignificant deuterium incorporations relative to the RNA bound state. Thus, RNA binding allosterically destabilizes the CARD2-HEL2i hydrophobic interface by releasing the RIG-I tandem CARDS. We also observed an intriguing HDX pattern in a second human HEL2i gate motif, V522-539. In *apo* RIG-I, peptide V522-539 exhibited a HDX kinetics pattern increasing from 15% to 55% over 1 hr of HDX (Figure 2.2B). In contrast, peptide V522-539 from *apo* RIG-I Δ CARDS exhibited a significantly higher deuterium uptake from 25% to 85% over the same time interval. Thus, the RIG-I Δ CARDS V522-539 peptide displayed a higher degree of HDX in the absence of the tandem CARDS, further supporting that human *apo* RIG-I adopts an auto-inhibitory conformation through intra-domain interactions between CARD2 and HEL2i. Interestingly, upon addition of 3p10L, we observed a four-fold higher protection against HDX for RIG-I Δ CARDS V522-539 (22 %) compared to that of RIG-I (5%) (Table 2.1). We therefore reason that the V522-

539 HEL2i peptide becomes protected from HDX by the CTD upon RNA binding following collaborative conformational rearrangements of the HEL and CTD domains. In the full length RIG-I protein, however, the V522-539 peptide was protected from HDX in both the apo and RNA bound state. This is likely because the CTD displaces the CARDs upon RNA binding and maintains the protection of the V522-539 peptide in the RNA bound state. These data are supported by structural alignments of different RIG-I states which clearly indicate that the CTD would displace CARD2 in the RNA and ATP bound states [4].

Among the CARD2 peptides that exhibited an RNA-dependent increase in uptake of deuterium (Table 2.1), peptides Y103-114 and L160-175 displayed unique partial local unfolding events following bimodal EX1 HDX kinetics at the 15 min and 1 hr time points (Figure 2.2C). In EX1 kinetics (also known as cooperative exchange/unfolding), the unfolding rate of the corresponding protein region is slower than the HDX rate. As a result, it gives rise to two distinct MS envelopes – the lower mass isotopic distribution represents molecules that have not unfolded/exchanged and the higher mass distribution depicts molecules that have undergone unfolding/exchange events. The majority of native proteins follow EX2 kinetics wherein only a single isotopic distribution increasing in mass is observed over time, oftentimes as a result of local structural fluctuations or perturbations [222]. Although the EX1 signature normally exists only under protein denaturing conditions (detergent or extreme pH), its occurrence under physiological conditions can reveal important insights into the nature of protein structural dynamics in solution such as large scale conformational changes [223-225]. Specifically, peptide Y103-114 is the CARD2 latch peptide that directly controls the opening of the RIG-I CARDs signaling module, while peptide L160-175 is the C-terminal peptide in CARD2 that is attached through a long 55-residue linker to HEL domain. The EX1 kinetic data for these two peptides suggest that binding to RNA PAMPs induces large scale conformational changes in RIG-I in which the tandem CARDs undergo partial unfolding at the region near the long linker (L160-

175) triggering disengagement from HEL2i. These two regions are therefore critical in controlling the gated activation of RIG-I.

In the absence of the RNA ligand, amide hydrogen within the RNA binding motifs of the HEL and CTD domains consistently exhibited greater exchange with the deuterium solvent (Table 2.1 and Appendix F1). We therefore conclude that the human RIG-I helicase and CTD domains adopt an open and extended conformation in the absence of RNA. This is in perfect agreement with the structural observations of the unliganded RIG-I proteins [67, 69, 226].

Figure 2.2 RIG-I apoenzyme conformation. (A) Schematic representation of full length and truncated form of RIG-I, displaying CTD (light orange), Pincer (light red), HEL1 (light green), HEL2 (light purple), HEL2i (light blue) and tandem CARDs (grey). RIG-I Δ CARDs is illustrated in the dashed box. (B) The auto-inhibitory conformation of *apo* RIG-I is maintained by the CARD2 and HEL2i domain (structure model based on duck RIG-I apo enzyme, PDB: 4a2w). Deuterium uptake plots of the CARD2 latch peptide (Y103-114) and the HEL2i gate peptides (Y566-574 and V522-539) from RIG-I and RIG-I Δ CARDs are shown. The data are plotted as percent deuterium uptake versus time on a logarithmic scale. Red, blue plots represent the RIG-I (RIG-I Δ CARDs) apoenzyme and the 3p10L bound state, respectively. (C) The RIG-I CARD2 peptides Y103-114 and L160-175 follow EX1 kinetics (in red) with two distinct mass isotopic distributions. In contrast, the RIG-I CARD2 peptides Y108-128 and the I139-155 follow EX2 kinetics (in black) in which only one mass isotopic distribution is present.

2.3.3 RIG-I binds and responds to different RNA species (3p8L, 3p10L, and polyIC) differently

In order to confirm the opening of the CARD2:HEL2i interface as the first step of RIG-I activation and to validate the HDX-MS method to monitor the activation process, three RNA species - 3p8L, 3p10L, and polyIC - were used in parallel to stimulate RIG-I in our HDX-MS assay. 3p8L and 3p10L are short RNA hairpins with 5' triphosphates but of different length (8 and 10 base pairs). 3p8L was shown to bind RIG-I but not elicit RIG-I-mediated interferon production in a cell-based reporter assay whereas 3p10L both bound and effectively stimulated a RIG-I mediated interferon response [4, 27]. In contrast to the short RNA hairpins, polyIC is mixture of long duplex RNAs bearing 5' diphosphates and has been well known to activate RIG-I in a variety of assays [26-28]. However, the molecular basis of how RIG-I reacts to these structurally different RNA species is still unknown.

Table 2.1 Comparison view of the conformational dynamics of RIG-I and RIG-IΔCARDs induced by RNA ligand (5' triphosphorylated RNA hairpin 3p10L and polyIC) and ATP hydrolysis

Domain / Peptides	RIG-I ±3p10L	RIG-I ±3p10L &ATP	RIG-I ΔCARDs ±3p10L	RIG-I ΔCARDs ±3p10L &ATP	RIG-I ±polyIC	RIG-I ±polyIC &ATP	RIG-I ΔCARDs ±polyIC
CARD1 F12-26	6 (4)*	8(2)**			5(3)*	5(2)*	
CARD2 latch Y103-114	15(4)**	18(2)**			12(2)**	9(2)**	
CARD2 K108-126	9(4)**	12(3)**			7(4)*	8(4)*	
CARD2 L139-155	11(3)**	14(2)**			7(3)**	7(2)**	
CARD2 L160-175	12(3)**	14(3)**			7(3)**	5(2)**	
CARD2 L160-185	8(4)*	14(3)**			10(4)**	13(5)**	
HEL1 motif Ia F296-306	-13(6)*	-10(3)*	-21(4)**	-23(3)**	-7(5)*	-9(5)**	-14(2)**
HEL1 motif Ic I343-363	-9(5)*	-6(3)*	-12(6)**	-14(6)**	0(5)	0(5)	-7(3)**
HEL1 motif IIa I370-387	-3(2)	-2(1)	-5(1)**	-6(2)**	0(2)	0(2)	-4(1)
HEL2i gate/CTD binding V522-539	-5(3)*	-2(2)	-22(4)**	-20(4)**	-5(3)*	-4(3)	-22(4)**
HEL2i gate Y566-574	5(2)**	7(1)**	-1(6)	-4(6)	8(3)**	6(3)**	-3(2)
HEL2 motif IV F633-645	-5(4)**	-4(4)	-3(2)	-4(2)	-1(3)	-1(3)	1(1)
CTD RNA capping loop V843-856	-11(6)*	-7(5)*	-11(6)*	-14(6)**	-2(5)	-2(5)	1(5)
CTD RNA binding F882-891	-6(1)**	-6(1)**	-6(1)**	-6(2)**	-1(2)	-1(2)	-1(1)
CTD RNA binding Y905-913	-8(5)**	-5(3)**	-12(4)**	-14(4)**	-6(3)*	-6(3)*	-7(2)**

ΔDeuterium uptake (%)	≥ 10	10 > ~ ≥ 5	N.S.	-5 ≥ ~ > -10	-10 ≥ ~ > -20	-20 ≥
-----------------------	------	------------	------	--------------	---------------	-------

Footnote: *, p ≤ 0.02; **, p ≤ 0.005.

The values ($\Delta D_2O\% \pm$ standard deviation) represent the average difference of deuterium uptake between RIG-I *apo* enzyme and RIG-I+RNA ligand state and between RIG-I *apo* enzyme and RIG-I + RNA ligand&ATP state across all HDX time points. Correspondingly, the same peptides (HEL-CTD region) from RIG-IΔCARDs are also shown. Regions with statistically significant differential deuterium uptake percentage were colored. The value in bracket means standard error. Asterisks indicate significant difference based on one-way ANOVA/Tukey (* = p < 0.02; ** = p < 0.005). Peptides exhibiting non-significant (N.S.) or undetectable changes are colored gray.

HDX profiles of polyIC bound RIG-I displayed significantly different characteristic patterns of RNA binding compared to that of 3p10L (Table 2.1). The most interesting HDX pattern was observed at the RIG-I CTD domain, in which the RNA capping loop V843-856 exhibited comparable HDX kinetics between apo and polyIC bound states, indicating this loop is not likely involved in binding to polyIC (Table 2.1). In contrast, the same peptide showed a high degree of protection against HDX (11% decrease in deuterium uptake) upon binding to the blunt-ended terminus of 3p10L (Table 2.1). Another RNA binding peptide F882-

891 in the CTD also showed statistically insignificant changes in the HDX profiles upon polyIC binding (Table 2.1). The diphosphates at the 5' ends of polyIC are likely to be the preferred binding sites of RIG-I especially during the earlier stages of the innate immune response where the physiological concentrations of RIG-I and RNA are low [28]. However, in our HDX experiments, there are more internally bound RIG-I molecules than long polyIC duplexes, and therefore the HDX signals were dominated by RIG-I molecules that bind internally (Table 2.1). Overall, the RIG-I CTD bound to polyIC weakly through a smaller interface as only peptide Y905-913 displayed 8% and 6% decreased deuterium uptakes in both the +3p10L and +polyIC states. Similarly, the overall binding strength of the RIG-I HEL domain to polyIC was weaker compared to 3p10L. For example, upon polyIC binding, HEL1 motif Ia (F296-306) showed a 7% decrease in deuterium uptake and several other conserved RNA binding motifs - Ic, IIa, and IV - displayed no significant differences in their HDX profiles (Table 2.1). In contrast, various degrees of deuterium uptake were observed in conserved RNA binding motifs of 3p10L bound RIG-I, including motif Ia (13%), Ic (9%) and IV (5%). Furthermore, unlike MDA5, we observed no evidence of cooperative binding of RIG-I on polyIC, as RIG-I lacked the inter-protein contacts observed in the MDA5 filament (Table 2.1, Table 3.1, Appendix F3 and G1). Finally, polyIC was also able to disrupt the RIG-I CARD2-HEL2i interface in our HDX experiment (Figure 2.3 and Table 2.1). At the 1 hr HDX timepoint, the CARD2 latch peptide Y103-114 exhibited greatly reduced protection from HDX (70% and 55%) upon binding to 3p10L and polyIC compared to the RIG-I apoenzyme (30%). The corresponding HEL2i gate peptide Y566-574 also displayed an increase in deuterium uptake in the 3p10L and polyIC group compared to apo RIG-I (Table 2.1 and Figure 2.3).

3p8L binds to the RIG-I HEL and CTD domains, as indicated by the HDX protection profiles of the HEL1 motif Ia F293-310 and CTD RNA binding peptide F882-891 (Figure 2.3). However similar HDX rates in the CARD2 and HEL2i interface peptides were observed between *apo* RIG-I and 3p8L bound RIG-I, indicating that 3p8L is unable to allosterically disrupt the CARD2-HEL2i interface,

thus explaining why 3p8L cannot elicit a RIG-I-mediated interferon response. This observation provides direct structure-based evidence in solution that the minimum length requirement of RNA PAMP is at least 10 base pairs.

Figure 2.3 Characteristic HDX profiles of RIG-I peptides induced by different RNA ligands. Deuterium uptake for peptides (Y103-114, Y566-574, F293-310, and F882-891) after a 1hr exposure to deuterated solvent. Asterisks indicate significant differences based on one-way ANOVA/Tukey between apo and +RNA states (* = $p < 0.04$; ** = $p < 0.005$). Isotopic distributions of selected peptides are shown for indicated states.

In RIG-IΔCARDs, conserved RNA binding motifs in the helicase domain consistently showed strong protection against HDX. This RNA binding effect is more obvious in RIG-IΔCARDs than RIG-I as the differences in deuterium uptake are larger for all RNA binding motifs in HEL (Table 2.1). For instance in RIG-IΔCARDs, peptide F296-310 on HEL1 motif Ia exhibited an increased protection (21 % and 14%) against HDX in the presence of 3p10L and polyIC whereas this

peptide in RIG-I showed almost half the extent of protection (13 % and 7%) upon binding to the same RNA ligand. Tighter interactions to polyIC were also shown in RIG-IΔCARDs compared to RIG-I that motif Ic exhibited enhanced HDX protection (7%) in the truncated RIG-I. These results are consistent with the fact that RIG-IΔCARDs has higher RNA binding affinity than RIG-I [227]. In contrast, CTD peptides showed similar HDX profiles between RIG-I and RIG-IΔCARDs in response to the same RNA ligand (3p10L or polyIC). The CTD capping loop in both full length and truncated RIG-I showed 11% decreases in deuterium uptake and statistical insignificance upon 3p10L and polyIC binding, respectively.

2.3.4 ATP allosterically modulates conformational rearrangements of RIG-I RNA complex

We next examined the role of ATP in driving the sequential conformational rearrangements of RIG-I subsequent to RNA recognition. HDX perturbations of 3p10L bound RIG-I were compared in the presence and absence of ATP (Table 2.1). RNA binding regions - motif Ia and Ic on HEL1, motif IV on HEL2 - consistently exhibited a discernible ATP dependent HDX kinetic profile, in which deuterium uptake plots in the presence of ATP fell between the *apo* and RNA bound states (Figure 2.4). This ATP-dependent HDX enhancement also applied to the CTD domain, as shown by the kinetic profiles of the RNA capping loop (V843-856) and RNA binding residues (Y905-913) (Figure 2.4). These results suggest that the RIG-I:RNA complex become less stable upon ATP binding and hydrolysis, which may drive the helicase and CTD to partially disengage from contacts with the RNA ligand and reach a conformational equilibrium in which thermo fluctuations occur more freely.

Figure 2.4 HDX plots of RIG-I peptides associated with ATP effect. Deuterium uptake plots of selected RIG-I peptides associated with ATP effects are shown. The data are plotted as percent deuterium uptake versus time on a logarithmic scale. Red, blue and green plots represent RIG-I apoenzyme, + 3p10L and + 3p10L&ATP state, respectively.

Intriguingly, ATP binding and hydrolysis also affected the HDX kinetics of the CARDS allosterically. When ATP was present in the samples of 3p10L bound RIG-I complex, RIG-I tandem CARDS consistently exhibited higher deuterium incorporations (Table 2.1). These ATP induced effects extended through the CARDS landscape – with CARD2 peptide L160-185 demonstrating the biggest increase in deuterium uptake (from 8% to 14%) (Table 2.1). The CARD2 latch peptide Y103-114 exhibited an additional increase in deuterium uptake (from 15% to 18%) in the ATP-activated state (Table 2.1). Synchronous increase in HDX kinetics was also observed in CARD1 F12-26 (from 6% to 8%), suggesting that ATP binding and hydrolysis increase the structural dynamics of the CARDS allosterically (Table 2.1). It is also worth noting that CARD1 is known to retain a relatively constant conformation with respect to CARD2 [5, 67]. Furthermore, the ATP enhanced HDX effects also took place on HEL2i gate motifs. Gate peptide Y566-574 exhibited a 5 % increase in deuterium uptake upon RNA binding. ATP further enhanced the HDX rate of the gate peptide to a 7% increase in deuterium

uptake (Table 2.1). The rate of amide hydrogen exchange with peripheral deuterium is also governed by local stability or fluctuations [87, 228]. Comparative HDX results suggest that RNA binding alone induces RIG-I to adopt a partially opened state, in which the CARD2-HEL2i intra-molecular interaction is loosened and RIG-I tandem CARDS form a conformational equilibrium where local fluctuation occurs. ATP binding and hydrolysis fully activate RIG-I tandem CARDS by completely dissociating them from HEL2i motif. Intriguingly, this is consistent with the observations that CARD2 latch peptide and its C-terminal peptide underwent partial unfolding events in solution upon RIG-I binding to RNA alone.

ATP binding and hydrolysis exhibited distinct HDX profiles between 3p10L bound RIG-I and RIG-I Δ CARDS. Upon introduction of ATP in the presence of RNA, slightly increased protections against HDX were observed across RNA binding motifs in all subunits, including HEL1, HEL2, HEL2i, and CTD (Table 2.1). These results are consistent with the crystallographic observations that ATP bound RIG-I Δ CARDS forms a more compact conformation by interacting with both strands of dsRNA [67, 68]. In the complex without ATP, RIG-I Δ CARDS is less compact as the helicase domain adopts a relatively open state wherein HEL2 is disordered [43].

To further dissect the role of ATP on RIG-I activation, we examined the allosteric effects of ATP, ADP and ADP-AlF_x (non-hydrolysable ATP analogue) on RNA binding and CARDS opening mechanics of RIG-I (Figure 2.5). 3p10L bound RIG-I was incubated with ATP, ADP and ADP-AlF_x respectively for 1hr before a 5min HDX in deuterated solvent. In CARD2, the deuterium uptake of latch peptide increased more than two folds in ADP-AlF_x group (~44%), compared to 3p10L alone, 3p10L+ADP and 3p10L+ATP (~21%) groups (Figure 2.5). In HEL domain, enhanced protections against HDX were observed for the conserved RNA binding motifs upon binding to ADP-AlF_x. HEL1 motif Ia, HEL2 motifs IV

and Vc exhibited approximately 40% decrease in deuterium uptakes in ADP-AIF_x group compared to 3p10L, 3p10L+ADP and 3p10L+ATP groups (Figure 2.5). However, in CTD, peptide F882-891 showed similar deuterium uptake events for all the 3p10L bound RIG-I in the presence or absence of ADP, ATP and ADP-AIF_x. These observations suggest that the helicase of RNA bound RIG-I adopts a much more compact conformation upon binding to ADP-AIF_x while the CTD RNA binding peptide is irresponsive to ATP analogue binding. RIG-I CARD2 became more open and accessible to the solvent upon ADP-AIF_x binding to the HEL domain.

Figure 2.5 Characterization of ATP analogue effects on RIG-I domain peptides. Deuterium uptake for peptides (Y103-114, Y566-574, F293-310, F633-645, Y715-724 and F882-891) in different ATP analogue binding states after a 5 min exposure to deuterated solvent. Asterisks indicate significant differences based on one-way ANOVA/Tukey (* = $p < 0.05$; ** = $p < 0.005$). Isotopic distributions of selected peptides are shown for indicated states.

2.3.5 The recognition of 3p10L by LGP2 is similar as RIG-I: HEL1 and CTD dominantly mediate hairpin RNA binding

Compared to *apo* state, HDX profiles of 3p10L bound LGP2 revealed various protections against deuterium uptake in solutions. These obvious HDX-resistant regions or peptides are mainly assembled in the HEL1 and CTD domains. In LGP2 HEL1, a peptide V54-66 that occupies the motif Ia α -helix displayed significantly increased protection against HDX (a 21% decrease in deuterium uptake) upon binding to 3p10L (Figure 2.6 and Figure 2.7). In the RIG-I- Δ CARDs crystal structure, this HEL1 counterpart holds firmly with the stem of 3p10L. A peptide F67-80 adjacent to motif Ia also showed a 7% decrease in deuterium uptake in the presence of RNA ligand, suggesting that this region is more stabilized or less dynamic in the RNA binding state. Motif IIa residues from I128-146 presented a 6% decrease in deuterium uptake (Figure 2.6 and Figure 2.7). In RIG-I, this peptide binds to one strand of the stem of dsRNA.

Figure 2.6 Structure model of 3p10L bound LGP2. The homology model of LGP2 was created based on reported RIG-I- Δ CARDs (4AY2) and MDA5- Δ CARDs (4GL2) using HHpred. Regions are colored based on HDX perturbation map of LGP2 (Figure 2.7).

The LGP2 HEL2 and HEL2i domain form a relatively opened conformation with 3p10L. Unlike RIG-I, the linker between HEL and CTD, residues V531-554, in LGP2 was greatly protected from HDX upon 3p10L recognition (a 20 % decrease in deuterium uptake) (Figure 2.6 and Figure 2.7). This linker in LGP2 (18 residues) is longer than that of RIG-I (13 residues) and its amide hydrogens consistently uptake less deuterium for all HDX time points in the presence of hairpin RNA, suggesting that this specific linker is involved in RNA binding and it enhances the CTD-HEL1 association in solutions. Like RIG-I, the RNA capping loop, residues Y588-615, showed strong protection against HDX (a 11% decrease in deuterium uptake) in the presence of 3p10L RNA (Figure 2.6 and Figure 2.7). This protection comes from the dsRNA terminal stabilization. Peptides - P631-639, L640-662 - also showed greatly less deuterium uptake with 20% and 15% decrease upon recognition of RNA (Figure 2.6 and Figure 2.7). In the reported crystal structure of LGP2 CTD domain, these two regions are tightly involved in RNA binding [229]. Most regions display similar HDX profiles between *apo* and RNA bound state. For instance, the HDX kinetics of the HEL2 insertion loop (peptide I404-426) in LGP2 adopted similar HDX patterns irrespective to 3p10L (Figure 2.7). In the reported RIG-I crystal structure, this insertion loop stays near the termini of duplex RNA and it does not interact with dsRNA [230].

Figure 2.7 Perturbation view: differential deuterium exchange Map for LGP2 ± 3p10L. The values ($\Delta D_2O\% \pm$ standard deviation) represent the average difference of deuterium uptake between LGP2 *apo* enzyme and LGP2+3p10L states across all HDX time points. Regions with statistically significant differential deuterium uptake percentage were colored according to the key. Peptides exhibiting non-significant or undetectable changes are colored gray based on statistics from t-test.

2.4 Discussion and conclusion

2.4.1 Differential behaviors in RNA recognition model HEL-CTD

The perturbation view of 3p10L bound RIG-I revealed that major protections from HDX in conserved RNA binding interfaces are in consistent with crystal structure work, including HEL1 motif Ia, Ic, HEL2 motif IV, as well as CTD capping loop (Table 2.1). In addition, protection from HDX was also observed in non RNA binding region, HEL2i peptide V522-539, which might be accounted to HEL2i-CTD intra-molecular interaction observed in the crystal structures of human RIG-I Δ CARDs:dsRNA [4, 43, 68]. While there is no crystal structure available for RIG-I with polyIC, insight into its binding mode can be inferred from its HDX profile. It is evident that long duplex RNA results in significantly distinctive HDX profiles compared with short duplex RNA (3p10L). Interestingly, in complex with polyIC, majority of RNA binding motifs - HEL1 Ic, HEL2 IV, and CTD capping loop – were not stabilized as compared with RIG-I apoenzyme (Table 2.1), a result could not have been predicted in short dsRNA:RIG-I Δ CARDs. These apo-like dynamics suggest that polyIC is deficient to stabilize these RNA binding regions. And only HEL1 motif Ia and CTD peptide Y905-913 revealed protections from HDX (Table 2.1). We therefore conclude that RIG-I binds polyIC transiently and weakly by adopting a relatively dynamic conformation involved in internal binding, which is different from 3p10L mediated terminal binding (Figure 2.8). In RIG-I and RIG-I Δ CARDs, however such cooperative binding events were clearly absent upon binding to long duplex RNA (Figure 2.8, Table 2.1, Appendix F3 and F7).

Compared to RIG-I, LGP2 is less known in its structure and optimal RNA ligands. The HDX kinetics profiles suggest that 3p10L bound LGP2 is more reminiscent of 3p10L bound RIG-I Δ CARDs, in which RNA recognition of LGP2 is mainly mediated via CTD and HEL1. Interestingly, we found that the HEL-CTD linker in LGP2 is involved in RNA interaction while this linker in RIG-I adopts a relatively opened state. The HDX profiles of LGP2 HEL2 and HEL2i domains

suggest that these two domains adopt relatively opened conformations in the presence of short dsRNA. Like RIG-I, the HDX plots of HEL2 insertion loop showed no statistical difference in deuterium uptake. It has been reported that LGP2 negatively regulates RIG-I mediated signaling activation in *vitro* and in *vivo* [22, 231]. The similarities in recognition of short dsRNA of LGP2 and RIG-I might compete each other in binding invading viral RNAs, thus alleviating RIG-I mediated signaling relay.

2.4.2 CARD2-HEL2i interface governs the key switch of RIG-I molecular machinery

It is hypothesized that the position of CARD2 latch peptide to HEL2i is the critical switch of RIG-I between inactive and active state. While HDX data could not provide accurate description on locations of CARD2 latch peptide to HEL2i upon different RNA binding, it could reveal the relative stability state in solutions. Upon 3p10L and polyIC binding, the HDX profiles specifically revealed obvious alleviation of protections from HDX in RIG-I peptides derived from CARD2-HEL2i interface (Table2.1), which is distal to the conserved RNA binding motifs within HEL-CTD. Such lack of stability would impair the hydrogen bonding within this interface and disrupt the gated conformation of RIG-I. Interestingly, in complex with 3p8L, this interface is still stabilized as compared with RIG-I apoenzyme (Figure 2.3). It is therefore highly possible that CARD2-HEL2i interface plays key role in surveillance on RNA PAMP from non-RNA PAMP. Furthermore, analysis of HDX isotopic distribution patterns of RIG-I and MDA5 CARD2 peptides revealed that only RIG-I CARD2 latch peptide Y103-114 and L160-175 underwent partial unfolding events (Figure 2.2C), indicating that these peptides become more dynamic or less stable in solutions. Given the locations on CARD2, Y103-114 latches the CARD2 to the HEL2i gate while L160-175 connects the most C-terminal of CARD2 with a 55-residue linker to HEL. It is highly possible that the stability state of these two CARD2 peptides are directly

correlated with RIG-I CARDs activation in solutions. Thus the change in receptor dynamics of CARD2 Y103-114 and L160-175 suggest that upon RNA PAMP recognition by RIG-I, the formation of RNP complex transforms RIG-I into a semi-closed conformation in which CARD2 is partially disengaged from HEL2i gate motif (Figure 2.8). In addition, ATP binding and hydrolysis could fully release RIG-I tandem CARDs (Figure 2.5 and 2.8). ATP binding could compress HEL-CTD-dsRNA complex and fully activates CARDs. ATP hydrolysis and release could result in disengagement of RNA from protein contact surfaces. In this way, the comparable HDX experiments provide important evidence to address the sequential activation of RIG-I CARDs following RNA PAMP recognition and ATP binding event (Figure 2.8).

Figure 2.8 A model of the RIG-I activation pathway. In the resting state, RIG-I adopts an auto-inhibitory conformation via CARD2-HEL2i interaction while the HEL and CTD domains maintain a flexible and extended conformation available for sensing viral RNA. Binding of 3p10L triggers a drastic structural transformation resulted in a ring-shape architecture in which CTD captures the terminus of the duplex RNA and HEL clamps the backbone. RIG-I then forms a semi-opened conformation in which tandem CARDs is partially released from the gate motif on HEL2i. ATP binding results in complete disengagement of CARDs away from HEL2i [3, 4]. Upon binding to polyIC, RIG-I adopts a dynamic conformation in which both CTD and HEL bind to polyIC and CARDs dissociate from HEL2i. There is no cooperativity between individual RIG-I proteins. K63-poly ubiquitin chains assist RIG-I CARDs tetramerization on mitochondrial outer membrane and MAVS oligomerization then occurs via MAVS-CARD: RIG-I CARDs interactions [5-8].

CHAPTER 3

HDX-MS characterization on MDA5 activation mechanism

3.1 Introduction

MDA5 is another RLR protein that defends viral infection by sensing viral RNAs. It is homologous with RIG-I in sequence and structure, consisting CARs, HEL, and CTD. MDA5 and RIG-I share a common adaptor protein, MAVS, for downstream signaling pathway. Unlike RIG-I, the CTD of MDA5 does not exhibit RNA recognition feature that binds to RNA terminus, but binds weakly to the internal regions of long duplex RNA [41, 42]. Unlike the CARs of RIG-I [5], the CARs of MDA5, which is attached to the HEL-CTD through a long, non-structured linker, has not been demonstrated to form higher-order structures with functional implications, although it has been shown that the association of non-covalent K63-linked polyubiquitin is also required for MDA5 activation [6]. MDA5 has a bias towards long duplex RNA and together they form helical fiber-like mega-structures through its RNA sensing module HEL-CTD [41, 44, 46, 47, 75]. MDA5 has been reported to play a functionally dominant role in polyIC induced IFN responses.

Precise RNA recognition by MDA5 plays a critical role in building a stable structural basis. The crystal structure of 12 bp dsRNA bound MDA5 Δ CARDs is clearly refined, which resembles liganded RIG-I Δ CARDs and forms a ring-like architecture with short dsRNA aligned perpendicular to MDA5 plane [41]. Distinctive from liganded RIG-I, this vertical dsRNA orientation, determined by precise positioning of CTD and Hel2 that specifically recognize the internal stem but not termini of dsRNA, is an important structural feature to form cooperative MDA5 filaments. Besides elaborate positions of helicase and CTD in terms of their orientations to dsRNA, another significant structural feature for cooperativity is

that dsRNA bound MDA5 occupy several α -helical frameworks within various subunits lying within a single plane at its two sides, therefore forming regular interfaces suitable for intermolecular interactions. To reveal apo MDA5 conformation, conformational changes of MDA5 induced by polyIC and ATP as well as the molecular basis for cooperativity, we performed HDX-MS on full length and CARDs-deleted MDA5 constructs.

LGP2 has been reported as a positive regulator of MDA5 mediated antiviral responses as the absence of LGP2 decreases IFN- β production induced by picornaviruses EMCV and poliovirus which are sensed by MDA5 [54]. It also has been shown that LGP2 is a concentration dependent regulator of MDA5 that low concentrations of LGP2 enhance only MDA5 signaling but not RIG-I signaling whereas high concentrations of LGP2 exhibit inhibitory effects towards RIG-I and MDA5 signaling activities. This suggests the differences of LGP2's role in regulating the other two RLRs [232]. PolyIC is the known ligand for activation of MDA5, and it is interesting to utilize PolyIC to trigger the conformational changes of LGP2 in order to reveal if there is any difference in the inherent properties of MDA5 and LGP2.

3.2 Material and methods

3.2.1 Materials

The plasmids that encode mouse MDA5, mouse MDA5 Δ CARDs were provided by Dr.Luo Dahai group [4]. pET-SUMO, SUMO protease, TEV protease were purchased from Invitrogen. All restriction enzymes, ligases and T7 polymerase were purchased from New England Biolabs. Low molecular weight PolyIC was purchased from Invivogen. All PCR reagents were purchased from Invitrogen. The competent Rosetta II (DE3) Escherichia coli cells were purchased from Novagen. Ni-NTA beads were purchased from QIAGEN. All the chemicals

were purchased from Sigma-Aldrich. All FPLC system was purchased from GE Healthcare. Pepsin column was purchased from Invitrogen. Two external switch valves were ordered from VICI instruments. 1/16" OD x 50 μ M ID PEEK tubing was purchased from uppurge. C4 magic beads were purchased from Michrom. Liquid chromatography (Dionex 3000 RSLC) and LTQ-Orbitrap mass spectrometer were purchased from Thermo Fisher Scientific.

3.2.2 Cloning, Expression, and Purification of MDA5, MDA5 Δ CARDs and LGP2

The MDA5 constructs were cloned into the pET-SUMO vector (Invitrogen) and transformed into Rosetta II (DE3) *E.coli* cells (Novagen). Mouse LGP2 construct is cloned into the pNIC28-Bsa4 plasmid. The purification procedures of MDA5 constructs and LGP2 are similar except that SUMO protease is used to cleave the SUMO tag on MDA5 proteins and LGP2 is cleaved by TEV protease for tag removing. The proteins were expressed in LB media upon the addition of 0.5 mM isopropyl- β -D-thiogalactopyranoside (IPTG) and grown at 18 °C overnight for 20 hr [27, 43]. The full-length MDA5 and N-terminal CARDs (1–306) deletion constructs, hereafter named MDA5 Δ CARDs, was cloned into the pET-SUMO vector (Invitrogen). Transformed Rosetta II (DE3) *Escherichia coli* cells (Novagen) were grown at 37°C in Luria broth medium supplemented with 40 μ g ml⁻¹ kanamycin and 34 μ g ml⁻¹ chloramphenicol to an OD_{600nm} of 0.6–0.8. Protein expression was induced at 18°C by adding isopropyl- β -D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. After 20 hr growth, cells were harvested by centrifugation at 8,000 $\times$ g for 10 min at 4°C and stored at -20°C. Cells resuspended in buffer A (25 mM HEPES [pH 8.0], 0.5 M NaCl, 10 mM imidazole, 10% glycerol, 5 mM β -ME) were lysed by passing three times through a MicroFluidizer at 15,000 psi and the lysate was clarified by centrifugation at 15,000 $\times$ g for 60 min at 4°C. The supernatant was purified by

batch binding with QIAGEN Ni-NTA beads. The beads were collected in Biorad polyprep columns and the SUMO-tagged proteins were eluted with buffer B (25 mM HEPES [pH 8.0], 0.3 M NaCl, 10% glycerol, 5 mM β -ME, 200 mM imidazole). The fraction containing His₆-Sumo-MDA5 (MDA5 Δ CARDs) was then digested with ulp protease (Invitrogen), 4°C overnight. His-tag bearing LGP2 was then digested overnight with TEV protease at 4°C. The cleavage mixture was loaded onto a HisTrap HP column to remove the His₆-Sumo protein and ulp protease from the mixture. The recombinant protein was then further purified by using a HiTrap Heparin HP column (GE Healthcare) by running buffer C with an additional 1 M NaCl gradient. Concentrated proteins were subjected to a final gel-filtration purification step through a HiPrep 16/60 Superdex 200 column (Amersham Bioscience) in buffer D (25 mM HEPES [pH 7.4], 150 mM NaCl, 2 mM MgCl₂, 5% glycerol, 5 mM β -ME). Fractions containing monomeric MDA5 or MDA5 Δ CARDs were pooled, concentrated, and stored at -80°C. The concentrations were determined spectrophotometrically using the extinction coefficients of ϵ = 115,971 M⁻¹ cm⁻¹ for the mouse MDA5, 82,992 M⁻¹ cm⁻¹ for the MDA5 Δ CARDs (1-306 deleted) proteins. The storage buffer contained 25 mM Hepes, pH 7.4, 150 mM NaCl, 5 % glycerol, 5 mM β -ME. The protein preparations were flash frozen with liquid nitrogen and stored at -80 °C. The purity for each protein was >95% as verified on SDS-PAGE gel.

3.2.3 RNA Work

LMW polyIC (Invivogen) was dissolved in buffer containing 25 mM Hepes (pH 7.4), 150 mM NaCl, 5% glycerol, 5 mM β -ME to a final concentration of 10 mg/ml. Concentrations were determined spectrophotometrically.

3.2.4 Hydrogen/Deuterium Exchange

The buffer for HDX on-exchange was the same composition except that H₂O was replaced with D₂O (99.99%) and glycerol was removed. Specifically, we added 1mM ATP into D₂O buffer (ATP&D₂O buffer) for ATP hydrolysis studies associated with all proteins. 4μl of MDA5 or MDA5 Δ CARDs or LGP2 (with or without polyIC, 18μM) was mixed with 16μl D₂O buffer (final D₂O concentration was 80%) or H₂O buffer for 0s samples, and incubated at 4 °C at various time intervals, e.g. 10s, 30s, 60s, 300s, 900s, 3600s. For ATP hydrolysis group, 4μl of MDA5 or MDA5 Δ CARDs (18μM) with polyIC was mixed with 16μl ATP&D₂O or ATP&H₂O buffer for 0s samples, and incubated at 4°C at the same time intervals. 4μl polyIC bound MDA5 (18μM) with and without LGP2 (36μM) was mixed with 16μl D₂O buffer for 0s samples, and incubated at 4°C at the same time intervals. Ice-cold 20μl quench solution consisting of 1M guanidine hydrochloride and 1.5% (v/v) formic acid was added to each on-exchanged samples after specific time intervals prior to rapid freezing by liquid nitrogen till HDX MS analysis.

3.2.5 LCMS Under Quench Condition

For capillary-flow LC, buffer A was H₂O containing 0.3% (v/v) formic acid, and buffer B was acetonitrile containing 0.3% (v/v) formic acid. Protein samples were then digested online by passing through an immobilized pepsin-coupled column (2.1 mm i.d. x 30 mm) (Invitrogen) and were de-salted for 3 min on a house-packed C4 trap (0.75 mm i.d. x 10 mm, C4 beads purchased from Michrom). The mobile phase for on-line pepsin digestion was buffer A and the flow rate was 150 μl min⁻¹ driven by the LC loading pump (Dionex 3000 RSLC). A 20 min gradient on a house-packed C4 column (0.3 mm i.d. x 50 mm, C4 beads purchased from Michrom) was used to separate and elute peptic peptides prior to MS analysis. All parts are connected by 1/16" OD x 50 μM ID PEEK tubing and the flow rate was 15 μl min⁻¹ driven by LC NC pumps. The gradient started from 5 % buffer B

and increased to 35 % buffer B within 20 min, followed by washing with 90 % buffer B for 3 min and equilibration with 1 % buffer B for 5 min. MS raw files were acquired in the range of m/z 300-2000 (with mass resolution set to 100,000 at m/z 400) for 30 min in positive mode on a LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific) equipped with an ESI source (capillary temperature 275 °C and spray voltage of 5 kV). All the HDX systems were strictly performed at 0 °C (fully buried on ice & water) and on-line pepsin digestion was carried out at 16 °C. Blank injections were made between every two samples to remove carryover peptides. Data for each time point were repeated three times. All HDX data were normalized to 100 % D₂O content, corrected for an estimated average deuterium recovery of 70 %, and analyzed by the software HDX Workbench [219]. The HDX Workbench has built-in statistical software to evaluate the statistical significance using a one-way ANOVA and Tukey's multiple comparison tests. Initial peptic peptide identifications were performed with the same HDX set up as described above. Four microliters of protein sample (20 µM) were injected into the HDX MS system. Product ion (MS/MS) spectra were acquired in linear ion trap LTQ with eight most abundant ions selected in the precursor (MS) scan with a 7.5 sec exclusion time. MS and tandem MS files were extracted and searched by using Integrated Proteomic Pipeline (<http://integratedproteomics.com/products/ip2/>) for high-confident peptide identification (searching parameters: Precursor delta mass cutoff 20, Best peptide FP threshold 0.01, Best peptide delta mass threshold 10).

3.2.6 Structure modeling

The full length *apo* mouse MDA5 structure was modeled based on the RIG-I CARDs (PDB code: 4NQK) and the human MDA5 HEL-CTD (PDB code: 4GL2). The homology modeling program is HHPred (<http://toolkit.tuebingen.mpg.de/hhpred>) [220].

3.3 Results

3.3.1 Protein expression and purification of MDA5

SUMO-MDA5 and its truncated construct were designed to generate monomeric MDA5. The soluble fraction collected from Ni²⁺NTA column was preceded with 3 hr protease digestion (SUMO protease). The digested products were loaded back to His-trap column (GE) to remove cleaved SUMO-tag and incompletely cleaved SUMO-MDA5. The flow-through was further purified by heparin column to remove bacterial RNA prior to a S200 size exclusion column for further purification (Figure 3.1). The monomeric form of MDA5 or MDA5 Δ CARDs were collected and concentrated.

Figure 3.1 SDS-PAGE gels and size-exclusion chromatogram of the recombinant MDA5 protein.

3.3.2 CARDs of *apo* MDA5 are exposed in solution

To determine the MDA5 conformation in its resting state and to compare it with RIG-I, we performed HDX experiments with recombinant mouse MDA5 and MDA5 Δ CARDs (residues 306-1025) in parallel (Figure 3.2A). To assess the

conformational changes that lead to MDA5 activation, we used polyIC as the RNA ligand and performed the HDX kinetic experiments in the presence of ATP. In total, 182 and 120 peptides were recorded and mapped to full length MDA5 and MDA5 Δ CARDs, representing 82 % and 84 % sequence coverage, respectively (Appendix C, D, G1 and G3). We observed no significant difference in the perturbation view of the tandem CARDs between *apo* and polyIC bound states (Figure 3.2B and Table 3.1). In particular, MDA5 CARD2 peptide T100-114, structurally corresponding to the RIG-I latch peptide Y103-114, exhibited high amide hydrogen exchange activities irrespective of polyIC binding, yielding 70% deuterium uptake after 1 hr incubation in deuterated solvent (Figure 3.2B). Its adjacent CARD2 peptides (L118-129, T145-167 and D183-193) and linker peptides (F194-222 and D271-293) consistently showed statistically insignificant HDX patterns (Table 3.1). Furthermore, unlike RIG-I, there were no partial unfolding events involved in MDA5 CARD2 peptides T100-115 and D183-193 (corresponding to RIG-I Y103-114 and L160-175) (Figure 3.2C). Therefore these uniform HDX rates associated with MDA5 CARDs suggest that it remains solvent-exposed in the presence and absence of polyIC.

To further highlight the radical differences between *apo* MDA5 and *apo* RIG-I, we observed that the peptide F630-642 on MDA5 HEL2i surface, corresponding to the RIG-I HEL2i gate motif, showed considerably higher deuterium uptake (a 15 % increase in average deuterium uptake across all HDX time points) in the absence of polyIC (Figure 3.2B and Table 3.1). Furthermore, HDX kinetics of this peptide in full length MDA5 was indistinguishable from that in MDA5 Δ CARDs, which also showed higher deuterium exchange activities (a 17 % increase in average deuterium uptake across all HDX time points) in the *apo* state (Figure 3.2B and Table 3.1). These results together indicate that the surfaces of MDA5 CARDs as well as HEL2i are not involved in the intra-molecular auto-inhibitory interactions as with RIG-I.

In the perturbation views of MDA5 and MDA5 Δ CARDs (Appendix G), MDA5 surface peptides that are neither proximal to RNA binding motifs nor involved in intermolecular interactions (see next section), showed negligible HDX difference in the presence and absence of PolyIC. These regions were distributed in various domain surfaces, including HEL1 (Y342-363), HEL2 (F517-555), HEL2i (L679-700) (Appendix G1). Conserved RNA binding motifs or regions involved in intermolecular interactions in helicase and CTD domains consistently sustained higher deuterium exchange activities in their *apo* state. It is thus evident that MDA5 tandem CARDs and CTD are flexible and connected to the central helicase domain which forms an extended and open conformation (Figure 3.2B). This result is consistent with previous MDA5 SAXS studies [46].

Figure 3.2 MDA5 apoenzyme conformation. (A) Schematic representation of full length and truncated form of MDA5, displaying CTD (light orange), Pincer (light red), HEL1 (light green), HEL2 (light purple), HEL2i (light blue) and tandem CARDs (grey). MDA5 Δ CARDs are illustrated in the dashed box. (B) The CARD2-HEL2i interface is absent in the MDA5 apoenzyme (structure model based on the duck RIG-I CARDs, PDB: 4a2w and human MDA5 Δ CARDs, PDB: 4gl2). Deuterium uptake plots of the CARD2 peptide (T100-114) and the HEL2i peptide (F630-642) from MDA5 and MDA5 Δ CARDs are shown. The data are plotted as percent deuterium uptake versus time on a logarithmic scale. Red and purple plots represent the MDA5 (MDA5 Δ CARDs) apoenzyme and + polyIC state, respectively. (C) MDA5 CARD2 peptides T100-115 and D183-193 follow EX2 kinetics (in black) in which only one mass isotopic distribution is present.

3.3.3 MDA5 binds polyIC through HEL and CTD

Since MDA5 forms cooperative filaments rather than monomeric complexes with RNA, obtaining high-resolution structural information of the function unit of this protein remains challenging. To map out the molecular interface within the co-operative MDA5 polymer on long duplex RNA, we recorded HDX kinetics for an MDA5:polyIC complex. We observed that MDA5 in complex with polyIC was easily digested when passing through an on-line pepsin column, yielding abundant peptic peptide signals acquired in mass spectrometric raw files. We also carried out HDX studies on MDA5 Δ CARDS in complex with polyIC, which were also thoroughly digested by pepsin protease. The association between MDA5 and polyIC was neither tight nor rigid, but did constitute an ordered complex. In contrast, we failed to obtain good MS signals for a RIG-I:polyIC complex.

MDA5 helicase and CTD formed extensive interactions with polyIC (Table 3.1). In HEL1, residues V410-422, covering the helicase motif Ic, showed an over 20% decrease in deuterium uptake in the presence of PolyIC (Table 3.1). Two HEL1 peptides (I441-459, T489-514), covering RNA binding motif IIa, motif III and the Pincer axel, also exhibited 13 % and 14 % less deuterium uptake, respectively. These data indicate HEL1 forms tight interactions with the RNA duplex, which is in agreement with the structural observations [41]. In HEL2, one peptide, A745-762, showed increased protection (11 %) against HDX compared to the unliganded state. This peptide forms a loop inserting into the major groove of the duplex RNA [41]. The CTD also established several direct interactions with polyIC. Peptide L913-938 showed a low extent of protection from deuterium exchange (a 3 % decrease in HDX) upon PolyIC binding, likely because this peptide contains an α -helix and two β -sheets and is close to the RNA binding surface [41]. Peptide K992-1006 on the CTD surface displayed a high degree of protection against HDX in the presence of RNA ligand, as shown by the marked decrease of deuterium uptake (11%). This peptide was observed to interact directly

with one strand of duplex RNA [41]. To further evaluate these RNA binding regions, we performed MDA5 Δ CARDS HDX performances in response to polyIC binding. As expected, all above mentioned RNA binding motifs or regions showed similar degrees of protection from deuterium exchange between MDA5 Δ CARDS and full length MDA5 (Table 3.1).

Table 3.1 Comparison view of the conformational dynamics of MDA5 and MDA5 Δ CARDS induced by polyIC and ATP hydrolysis

Domain / Peptide	MDA5 ±polyIC	MDA5 ±polyIC&ATP	MDA5ΔCARDs ±polyIC	MDA5ΔCARDs ±polyIC&ATP		
CARD1 Y25-37	1(3)	2(2)				
CARD2 T100-114	1(6)	0(4)				
CARD2 L118-129	2(3)	1(2)				
CARD2 T145-167	0(4)	-1(3)				
CARD2 D183-193	2(5)	2(3)				
Linker F194-222	2(5)	1(3)				
Linker F270-293	1(4)	-1(3)				
HEL1 front interface F374-391	-20(2)**	-21(2)**	-25(3)**	-24(3)**		
HEL1 front interface K398-409	-7(2)**	-6(2)**	-9(2)**	-8(3)**		
HEL1 motif Ic V410-422	-20(2)**	-18(2)**	-26(2)**	-25(2)**		
HEL1 motif IIa I441-459	-13(2)**	-12(2)**	-14(2)**	-12(3)**		
HEL1 motif III and Axel T489-512	-12(3)**	-13(3)**	-12(5)**	-10(4)**		
HEL2 back interface A560-581	-5(3)*	-3(2)	-8(4)**	-2(5)		
HEL2i front interface F630-642	-15(4)**	-15(6)*	-17(5)**	-12(5)**		
HEL2i surface loop N643-674	0(5)	0(3)	-4(6)	1(5)		
HEL2 insertion loop A745-762	-11(5)**	-11(2)**	-10(5)**	-9(5)**		
HEL2 front interface V763-785	-17(4)**	-11(5)**	-16(4)**	-14(6)**		
Pincer back interface R850-868	-14(4)**	-14(3)**	-14(5)**	-14(6)**		
Pincer back interface Q877-902	-12(5)**	-10(5)**	-17(4)**	-12(5)**		
CTD back interface L913-938	-3(2)*	-3(2)*	-5(2)*	-5(2)*		
CTD RNA capping loop I940-959	-14(4)**	-16(3)**	-18(5)**	-14(6)**		
CTD RNA binding K992-1006	-11(4)**	-13(3)**	-12(4)**	-12(6)**		
ΔDeuterium uptake (%)	≥ 10	10 > ~ ≥ 5	N.S.	-5 ≥ ~ > -10	-10 ≥ ~ > -20	-20 ≥

The values ($\Delta D_2O\% \pm$ standard deviation) represent the average difference of deuterium uptake between MDA5 *apo* enzyme and MDA5 + polyIC state and between MDA5 *apo* enzyme and MDA5 + polyIC&ATP state across all HDX time points. Correspondingly, same peptides (HEL-CTD region) from MDA5 Δ CARDS are also shown. Regions with statistically significant differential deuterium uptake percentage were colored. The value in bracket means standard error. Asterisks indicate significant difference based on one-way ANOVA/Tukey (* = $p < 0.02$; ** = $p < 0.005$). Peptides exhibiting non-significant (N.S.) or undetectable changes are colored gray.

3.3.4 Highly Cooperative binding of MDA5 to polyIC

Cooperative binding of MDA5 to dsRNA was clearly demonstrated by the polyIC-induced HDX resistant peptides (Table 3.1). We identified a total of eleven peptides mapped to the surface of HEL-CTD, the ring shape RNA recognition module of MDA5. Six were clustered on one face of the ring (hereafter termed as the front face) (Figure 3.3A) and five on the reverse face (back face) (Figure 3.3B). These peptides formed the intermolecular interface that mediated the cooperative binding of MDA5 to long RNA duplexes [44, 46, 47].

On the front face, six peptides played important roles in establishing an intermolecular interface between monomeric MDA5 monomers (Figure 3.3A and Table 3.1). In HEL1, peptide F374-391 showed significant protection against HDX with an apparent decrease of 20 % deuterium uptake upon binding to polyIC. This peptide adopts a position closer to the outer rim of the MDA5 ring-like architecture. Another short α -helical peptide, K398-409, recorded a drop of 7 % in deuterium uptake in the presence of long polyIC. In HEL2, peptide V763-785 was observed with significantly increased - 17 % - protection against HDX in the presence of polyIC. This peptide occupies an α -helical framework in conjunction with the HEL2 insertion loop [41]. In HEL2i, peptide F630-642 was identified as a HDX resistant peptide, which exhibited a 15 % decrease in deuterium uptake upon binding to polyIC. This peptide occupies the C-terminus of one α -helix on the surface of HEL2i in the reported crystal structure, with intra-molecular interaction between HEL2i and the CTD stabilizing this region from deuterium exchange [41]. In contrast, the adjacent peptide residues N643-674, comprising an unstructured surface loop, displayed almost identical HDX kinetics between *apo* MDA5 and polyIC bound MDA5, indicating this region is disordered and solvent-exposed even in the presence an RNA ligand (Figure 3.3A). The HDX kinetic profiles for the CTD peptide I940-959 are interesting. In *apo* MDA5, over 80 % deuterium uptake was observed after 1 hr HDX incubation, suggesting it was fully solvent-accessible in the absence of RNA. Upon polyIC binding, it only reached 60 %

deuterium uptake after incubating over the same time interval (Figure 3.3A). This CTD peptide is structurally equivalent to the RNA-capping loop in RIG-I, but is disordered in the crystal structure of the MDA5 dsRNA complex [41]. Therefore, this protection against HDX was not likely due to RNA binding but rather intermolecular contacts. Thus these six peptides from various subdomains of MDA5 provide an extensive contact surface for intermolecular interactions.

On the back face, five peptides were observed to form the other contact surface for the cooperative intermolecular interaction based on our HDX data (Figure 3.3B and Table 3.1). In HEL1, peptide T489-512 displayed increased protection against HDX (12 %) upon polyIC binding. This peptide includes RNA binding motif III and the Pincer axel, and is also partially involved in RNA binding activities. Peptide A560-581 on the HEL2i surface also displayed increased protection against HDX (4 %) in the presence of long dsRNA. Peptide R850-868 at the Pincer turn exhibited considerable protection against HDX (14%) in the presence of polyIC. Another peptide, Q877-902, adjacent to the Pincer also showed a decrease of deuterium uptake (12 %) compared to *apo* state. This peptide belong to the loop linking the HEL to CTD, and it was largely disordered in the crystal structure of MDA5-dsRNA complex [41]. Thus this region might be more critical for intermolecular interactions on long RNA duplex.

Previous X-linking experiments revealed that K777 on HEL2 of one MDA5 monomer is oriented in close proximity to mutated residues C556, C559 and C563 on HEL2i of a second MDA5 molecule [41]. Interestingly, K777 is within peptide V763-785 (front face) and residues 556-559-563 is on the same helix covering the peptide A560-581 (back face). With these eleven HDX resistant peptides and previous X-linking experiment, we are able to propose a more refined and high-resolution model of the MDA5 intermolecular interaction map. This model is in good agreement with the hypothesis that the front and back planes of MDA5 interface with each other at every ~ 74 degrees following the helical dsRNA backbone (Figure 3.3 C). In addition, these eleven peptides exhibited similar HDX

kinetic profiles between full length MDA5 and MDA5 Δ CARDs (Table 3.1). These results indicate that HEL and CTD domain concertedly mediate cooperative RNA binding of MDA5.

Figure 3.3 Structure basis of MDA5 cooperativity. (A) The front view: deuterium uptake plots of MDA5 peptides involved in intermolecular interactions (F374-391, K398-409, F630-642, A745-762, V763-785 and I940-959). The data are plotted as percentage deuterium uptake versus time on the logarithmic scale. Red plots indicate MDA5 *apo* enzyme and purple plots represent +polyIC state, showing that these surface peptides are more protected against HDX in the presence of polyIC. These peptides are colored in blue and mapped to MDA5 front surface. The HEL2i surface loop N643-674 (in gray), which is truncated in the reported crystal structure, is also modeled according to its HDX profile. (B) The back view: deuterium uptake plots of the peptides involved in intermolecular interactions (T489-512, F630642, R850-868, Q877-902 and L913-938). (C) A model displays cooperative binding of three neighboring MDA5 monomers to the long dsRNA. This model is constructed based on our HDX data taking the previous cross-linking experiments into consideration [41, 44, 47]. The front and back interface peptides of the central MDA5 monomer (n) are colored in blue and light blue according to Figure 3.3A and 3.3B. The interfaces of two neighboring MDA5 monomers (n-1 and n+1) are highlighted as well. See also the supplementary movie S1 on the cooperative binding of MDA5 to duplex RNA.

3.3.5 ATP weakens interactions between MDA5 and polyIC but does not disrupt cooperativity

Similarly, although to a lesser extent than RIG-I, the interactions between MDA5 and polyIC were slightly weakened in the presence of ATP (Table 3.1). RNA binding peptides from HEL1 motif Ic and IIa exhibited slight increases in deuterium uptake upon ATP binding and hydrolysis. These peptides, V410-422, I441-459, demonstrated less than 2 % increases in deuterium uptake in the presence of ATP. The HEL2 insertion loop also exhibited an increase in deuterium uptake (3 %) after external ATP stimulus. As with the full-length construct, this ATP dependent disruption of RNA binding was not prominent in MDA5 Δ CARDs (Table 3.1). Of note, the eleven peptides that were identified to be involved in intermolecular cooperative binding had almost no ATP dependent HDX perturbation. Furthermore, ATP binding and hydrolysis did not alter the HDX kinetics of the CARDs (Table 3.1). We conclude that ATP only slightly weakens

the interactions between MDA5 and duplex RNA but not the intermolecular interactions of MDA5.

3.3.6 Recognition of PolyIC by LGP2

Our high-resolution HDX profiles have clearly confirmed that MDA5 or MDA5 Δ CARDs could cooperatively form extended and filamentous structures along long dsRNA or PolyIC. To reveal the inherent properties and rational difference between MDA5 and LGP2, we performed HDX-MS kinetics studies on *apo* LGP2 and PolyIC bound LGP2. We were able to acquire and record high sequence coverage of LGP2 (>80%) in both states across seven HDX time points.

The recognition of PolyIC by LGP2 is mediated through HEL and CTD domain (Figure 3.4 and 3.5). In HEL1, motif Ia and Ib - residues V54-66, F67-100 - exhibited high extent of protection against HDX with a 13% and 16% decrease in deuterium uptake in the presence of PolyIC, respectively (Figure 3.5). Peptide I128-146 belonging to motif IIa also showed a 5% decrease in deuterium uptake upon binding to PolyIC (Figure 3.5). In HEL2i α -helix bundle, peptides - I233-246, F292-303, L328-341 - all exhibited little or no difference in deuterium uptake plots in the presence or absence of PolyIC. The HEL2 peptides showed no statistical significance or no distinction in deuterium uptake with or without PolyIC (Figure 3.5). Adjacent to this pincer2 peptide, residues V531-553, the linker peptide between HEL and CTD, consistently showed less deuterium incorporations with a 7% decrease in the presence of PolyIC (Figure 3.5). This inter-domain linker peptide also exhibited greatly less deuterium uptake (15%) in the short hairpin (3p10L) bound LGP2. The CTD peptides involved in RNA binding all exhibited obvious protections against HDX with the exception of one peptide (residues E573-606), which resides on RNA capping loop and showed comparable deuterium uptakes in the presence of PolyIC. This intriguing observation suggests that CTD RNA capping loop is not involved in association with PolyIC. Other

HDX resistant peptides in CTD - I624-639, L640-662 - exhibited a 9% and 12% decrease in deuterium uptake upon binding to PolyIC, respectively (Figure 3.5).

Figure 3.4 Structure model of PolyIC bound LGP2 (front view on the left, back view on the right, PolyIC not shown). The homology model of LGP2 was created based on reported RIG-I Δ CARDs (4AY2) and MDA5 Δ CARDs (4GL2) using HHpred. Regions are colored based on HDX perturbation map of LGP2 $\pm$ PolyIC (Figure3.5).

Figure 3.5 Perturbation view: differential deuterium exchange Map for LGP2 ± PolyIC. The values ($\Delta D_2O\% \pm$ standard deviation) represent the average difference of deuterium uptake between LGP2 *apo* enzyme and LGP2+PolyIC states across all HDX time points. Regions with statistically significant differential deuterium uptake percentage were colored according to the key. Peptides exhibiting non-significant or undetectable changes are colored gray based on statistics from t-test.

3.3.7 LGP2 forms different complex with polyIC compared with MDA5:polyIC filamentous structure

Our HDX profiles have confirmed the potential interface between MDA5 monomers along polyIC and this interface is also reflected in MDA5 Δ CARDs: polyIC complex, indicating that the CARDs domain is not involved in MDA5 helicase activities or affect MDA5 HEL conformations. Absent of CARDs domain, LGP2 consists of HEL and CTD domain that share with RLRs. To investigate the conformations of LGP2 HEL and CTD in the presence of PolyIC, we specifically compare the HDX profiles of LGP2 peptides in homology with MDA5 interface peptide, of which the HDX profiles are consistent in full length and CARDs-deleted MDA5 (Table 3.2).

Table 3.2 Comparison view of the conformational dynamics of MDA5 (MDA5 Δ CARDs) and LGP2 induced by polyIC

Domain / Peptide	MDA5 ±polyIC	MDA5ΔCARDs ±polyIC	LGP2 peptide in homology with MDA5	MDA5ΔCARDs ±polyIC&ATP		
HEL1 front interface F374-391	-20(2)**	-25(3)**	HEL1 F67-80	-8(3)**		
HEL1 front interface K398-409	-7(2)**	-9(2)**	HEL1 F67-100	-16(6)**		
HEL1 motif III and Axel T489-512	-12(3)**	-12(5)**	HEL1 T167-188	-3(3)		
HEL2 back interface A560-581	-5(3)*	-8(4)**	HEL2i I233-246	-4(2)		
HEL2i front interface F630-642	-15(4)**	-17(5)**	Non detection			
HEL2 insertion loop A745-762	-11(5)**	-10(5)**	HEL2 loop L403-426	0(0)		
HEL2 front interface V763-785	-17(4)**	-16(4)**	HEL2 F427-434	0(6)		
Pincer back interface R850-868	-14(4)**	-14(5)**	Pincer A509-530	0(2)		
Pincer back interface Q877-902	-12(5)**	-17(4)**	Pincer L530-553	-7(4)**		
CTD back interface L913-938	-3(2)*	-5(2)*	CTD L569-584	-2(1)		
CTD RNA capping loop I940-959	-14(4)**	-18(5)**	CTD capping loop S585-615	1(0)		
ΔDeuterium uptake (%)	≥ 10	10 >~ ≥ 5	N.S.	-5 >~ -10	-10 ≥ >~ -20	-20 ≥

The values ($\Delta D_2O\% \pm$ standard deviation) represent the average difference of deuterium uptake between MDA5 *apo* enzyme and MDA5 +polyIC state and between MDA5&PolyIC and MDA5&PolyIC+1mM ATP state across all HDX time points. Correspondingly, same peptides (HEL-CTD region) from MDA5 Δ CARDs are shown on the right half of the table. Regions with statistically significant differential deuterium uptake percentage were colored. The value in bracket means standard error. Asterisks indicate significant difference based on one-way ANOVA/Tukey (* = $p < 0.02$; ** = $p < 0.005$). Peptides exhibiting non-significant (N.S.) or undetectable changes are colored gray.

In HEL1, peptides F67-80 and F67-100, which reside on outer-rim of LGP2 ring-like structure, exhibited 8% and 16% drops in deuterium uptake in the presence of PolyIC (Table 3.2). In MDA5 and MDA5 Δ CARDs, these two corresponding peptides respectively showed around 20% and 7% decrease in deuterium uptake upon PolyIC binding. Peptide T167-188 belonging to RNA binding motif III and the “Axel” showed a 3% less deuterium incorporations in LGP2:PolyIC complex (Table 3.2). In contrast, this region in MDA5:PolyIC filaments exhibited a 12% decrease in deuterium uptake. In HEL2, the LGP2 and MDA5 insertion loop showed no change and 10% decrease in deuterium uptake upon binding to polyIC, respectively. Also, LGP2 HEL2 peptide F427-434 displayed no significant changes in deuterium uptake in *apo* state and PolyIC bound state (Table 3.2). In MDA5, this homology region belongs to the HEL2 front interface and exhibited a drastically 17% decrease in deuterium uptake in the presence of PolyIC. The LGP2 Pincer linker peptide between HEL and CTD showed a 7% decrease in deuterium uptake in the presence of PolyIC. In the monomeric LGP2:3p10L complex, this linker peptide was found as a HDX-resistant peptide (15% decrease in deuterium uptake) upon short dsRNA binding. This observation suggests that HEL-CTD linker is involved in RNA binding but not intermolecular interaction between LGP2 in the presence of polyIC. Another LGP2 pincer peptide A509-530 also showed no difference in deuterium uptake in the presence of polyIC (Table 3.2). In contrast, the corresponding pincer region (R850-868) in MDA5 constituted the protein-protein interface in MDA5 filaments. The HDX profiles of these two LGP2 Pincer peptides indicate that LGP2 Pincer is not involved in protein-protein interaction in the LGP2:PolyIC complex. The CTD capping loop in LGP2:PolyIC complex displayed comparable deuterium uptake events in the presence and absence of polyIC whereas the MDA5 counterpart showed a high extent of protection (14% decrease in deuterium uptake) against HDX in MDA5 fibers (Table 3.2). These intriguing HDX observations indicate that unlike MDA5, there is no cooperative binding present in the LGP2:PolyIC complex.

3.3.8 LGP2 stays no interaction with polyIC bound MDA5 in solution

It has been reported that different concentrations of LGP2 could have different effects on MDA5 signaling [233]. To explore whether LGP2 interacts with MDA5 filaments in solution, we incubated MDA5-polyIC complex in the presence and absence of LGP2. MDA5:polyIC complex showed non-significant or undetected changes in the presence and absence of LGP2, suggesting that LGP2 exhibits no protein-protein interactions with MDA5 in the presence of polyIC (Figure 3.6). Very recently, Annie et al. showed that LGP2 could assist MDA5 signaling events by enabling many shorter MDA5 filaments to assemble quickly [234]. Interestingly, we found that polyIC bound LGP2 is absent of cooperativity. This observation could indirectly explain why LGP2 could enable MDA5 to form shorter filaments. By maintaining no cooperative binding and no interacting with MDA5, LGP2 in low concentration could bind to long polyIC and partially disrupt the MDA5 long filaments into more potent shorter filaments for signaling. In contrast, LGP2 in high concentration could greatly compete with MDA5 for binding to polyIC and prevent formation of MDA5 filaments.

3.4 Discussion and conclusion

RNA recognition of MDA5 has an intrinsic bias towards long poly IC. Since MDA5 is expected to form cooperative filaments rather than monomeric complex with RNA, obtaining high-resolution structural information of this ordered and elongated filament remains increasingly challenging. To address RNA recognition and intermolecular cooperativity governed by MDA5 subunit movements, we further fractionated LMW poly IC and used the relatively high molecular weight or longer dsRNA portion for HDX studies. This experimental design enabled us to collect high-resolution as well as high coverage data for various conformational states (varying ligands and substrates). Indeed, we achieved high sequence coverage across seven HDX time points for MDA5 and its CARDs-deleted form. Direct comparisons of these time-resolved HDX-MS snapshots across *apo*, RNA bound and ATP stimulated enzymes enabled us to dissect the early activation steps of MDA5 signaling cascade.

Previous studies have defined the basic RNA requirements for RIG-I and less precisely for MDA5 [3, 45]. Here the HDX profiles of MDA5 suggest that MDA5 contains no intra-molecular contacts between CARD2 and HEL2i and all domains seem to be equally accessible to deuterated solvent (Figure 3.2B and Table 3.1). In addition, MDA5 has longer linker between CARDs and HEL (~100 disordered residues) than RIG-I (~ 60 disordered residues), so it is likely that *apo* MDA5 adopts a more open and flexible conformation (Figure 3.7). In our HDX experiments, MDA5 HEL and CTD formed complex with polyIC in agreement with the previous structural studies [41]. We also identified eleven peptides on the non-RNA-binding-surface of HEL-CTD that were shown to be protected against HDX. These peptides - six from one face of the HEL-CTD ring and five from the back face - formed large buried surfaces on both sides of HEL-CTD ring-like architecture to mediate the intermolecular interactions, providing direct evidence for the cooperative binding of MDA5 on long RNA duplexes (Figure 3.7). ATP had shown little effects to MDA5:polyIC super complexes and little or no

disruption of the cooperative interfaces was observed in either full length MDA5 or MDA5 Δ CARDS (Figure 3.7).

Distinct from MDA5 cooperative properties, the HDX profiles revealed that the LGP2 pincer and HEL2 domain as well as the CTD capping loop adopt similar conformations in the presence and absence of polyIC, indicating that these regions are not associated with inter-molecular interactions as MDA5. In this case, LGP2 might form disordered complex with PolyIC without cooperativity. Importantly, our data is also in consistent with the electron microscopy observations that LGP2 does not form long helical filaments like MDA5 or MDA5 Δ CARDS but become dispersed aggregates along long dsRNA [235]. Furthermore, our HDX data showed that LGP2 exhibits little or no physical interactions with MDA5 filaments. This observation could indirectly explain why LGP2 could enable MDA5 to form shorter filaments [234]. By maintaining no cooperative binding and no interacting with MDA5, LGP2 in low concentration could bind to long polyIC and partially disrupt the MDA5 long filaments into more potent shorter filaments for signaling. In contrast, LGP2 in high concentration could greatly compete with MDA5 for binding to polyIC and prevent formation of MDA5 filaments.

Figure 3.7 MDA5 early activation mechanisms revealed by HDX-MS. MDA5 apoenzyme adopts an open and flexible conformation without CARD2-Hel2i auto-inhibition. Upon polyIC recognition, MDA5 front and back planes were interfaced with each other at every $\sim 74^\circ$ following the helical dsRNA backbone. Five polyIC bound MDA5 formed a repeated unite. In this state, tandem CARDS are loosely connected to the core HEL via a long linker. In MDA5 HDX profiles, there is little conformational change upon ATP binding and hydrolysis except the disengagement of RNA from HEL1 and HEL2 RNA binding regions. K63-poly ubiquitin chains assisted RIG-I/MDA5 CARDS tetramerization on mitochondrial outer membrane, where MAVS aggregation occurred via CARD-CARDS interactions to initiate downstream activation pathway.

CHAPTER 4

Protein analysis of annexin A2 in influenza A virus particles

4.1 Introduction

Animal viruses, which contain nucleic acid encased in a protein coat, are parasitically relative with living organisms. Albeit their simple structures, mature virus particles could intriguingly recruit endogenous cellular proteins as subviral components; observations have been made on several enveloped viruses [236]. However, a precisely delineated repertoire of host proteins specifically incorporated into virions still remains to be addressed. Due to the modest efficacy of virus purification, host contaminants could coexist with “spiked” virions, e.g. influenza virus [112], and coronavirus [237].

Influenza A virus (IAV) is highly contagious and pathogenic, causing recurring epidemics and overwhelming pandemics [238, 239]. Influenza A virus belongs to the *Orthomyxoviridae* group of viruses, and it has eight segments of negative-sense, single-stranded RNA genomes packaged in a membrane shell [101]. The lipid envelope is derived from apical plasma membrane of infected host cells, where the mature IVA particles assemble and pinch off [170]. Influenza A virus has eleven viral proteins and nine of them are specifically incorporated into the virus particles [101, 240]. Hemagglutinin (HA) and Neuraminidase (NA) are abundantly anchored in viral envelopes at approximately four to one ratio, playing important roles in receptor recognition, membrane fusion and virus disaggregation, respectively [112, 115]. The viral membrane also contains a minor but important protein, Matrix protein 2 (M2), acting as a proton channel equilibrating pH during membrane fusion step [117]. Beneath the envelope peripherally lies matrix protein one (M1), an abundant protein tightly involved in virus assembly and morphology

[119, 169]. Internally, eight viral RNA genome segments are independently encapsidated into viral ribonucleoprotein complexes (vRNPs) associated with nucleoprotein (NP) and tripartite RNA polymerase proteins (PB1, PB2, and PA), which are dedicated for viral translation and transcription activities [241]. Small quantities of nuclear export protein (NEP) or NS2 are also localized inside the virions [242]. Thus, these nine viral proteins intrinsically and essentially constitute mature IAV particles. In contrast, two viral proteins, nonstructural protein one (NS1) and PB1-F2, are alternatively distributed in nucleus and mitochondria of infected cells, respectively. NS1 and PB1-F2 could strengthen pathogenicity by undermining innate host countermeasures and hijacking intrinsic cellular signaling pathways [138, 145]. What is more, a nonessential viral polypeptide, PB1-N40, is additionally identified as an N-terminally truncated peptide from PB1; nevertheless, it is not consistently encoded by all IAV strains [243].

Recent endeavors in proteomic techniques have greatly assisted virology researches in past decades [244]. Several dominant mass spectrometry based approaches have been utilized to uncover the host components associated with different virions (Table 4.1). Nevertheless, exogenous protease treatment e.g. subtilisin, protease-K, which could externally digest viral particles without accessing to subviral structures, was seldom induced prior to virion proteomic analysis. In these cases, host contaminants, whether co-attached to viral envelope spike proteins or existed in microvesicles, could be accidentally incorporated into mass spectrometric data. Only kaposi's sarcoma associated herpesvirus and rhesus monkey rhadinovirus were exogenously digested before further MS characterization [245, 246]. For influenza A virus, the sucrose gradient purified IAV viruses were comprehensively analyzed by SDS PAGE followed by 1D LC-MS/MS and multi-dimensional LC-MS/MS, resulting in identification of 9 viral proteins and 36 host components associated with virions [240]. In our deduction, the number of identified host proteins of these purified viruses could be larger than those actually within their viral proteomes. This is because that host contaminants

could either stick externally to viral spike proteins or exist in cellular microvehicles co-purified with virions rather than reside internally. In addition, the existing approaches frequently adopted gel electrophoresis to separate proteins before LC-MS/MS analysis (Table 4.1). Some low abundant viral or host proteins were difficult to visualize after gel electrophoresis, thus these proteins could not be selected for in gel digestion and MS analysis.

To settle these questions, a novel proteomic approach, which combined exogenous protease catalytic digestion and one dimensional nanoflow LC-MS/MS, was undertaken to characterize proteome of mature IAV particles and compared it with detergent & protease treated, and non-treated parallel groups. Our results revealed that a total of nine viral proteins and five host proteins were characterized as residents inside the IAV mature particles. Next, western blot analysis was further utilized to validate the coexistence of these proteins with IAV virions. Immunofluorescence microscope was used to visualize the association of identified host proteins with virus nucleoprotein. siRNA mediated knockdown of selected host protein was performed to further functionally examine the host protein that specifically incorporated into the virions.

4.2 Materials and methods

4.2.1 Materials

Purified H1N1/WSN/33 was provided by Dr. Richard Sugrue lab. BCA Protein Assay Reagent (bicinchoninic acid) was purchased from Thermo Scientific. Subtilisin protease was purchased from Sigma-Aldrich. Acrylamide/bis acrylamide (29:1) was purchased from Bio Rad Laboratories. Nitrocellulose membrane was purchased from Invitrogen. The following antibodies and secondary antibodies were indicated respectively, including ezrin (Santa Cruz, sc-6407), actin (sigma a2668), annexin A2 (abcam40943), α -tubulin (abcam4074) and IVA NP (Bei-

resources), IVA NA (Bei-resources), anti-rabbit IgG and anti-mouse IgG (Sigma-Aldrich), Alexa Fluor 488 goat anti-rabbit IgG (H+L) and Texas Red goat anti-mouse IgG (H+L) (Invitrogen). Enhanced chemiluminescence (ECL) ECL protein detection reagents were purchased from Perkin Elmer. Protein marker and Mounting solution were ordered from Bio-Rad and DAKO. For siRNAs, siGFP and three siRNAs against annexin A2 were purchased Sigma-Aldrich. Lipofectamine 2000 was purchased from Invitrogen. All the chemicals were purchased from Sigma-Aldrich.

4.2.2 Virus production and purification

Twelve days old embryonated eggs were inoculated with 10^4 plaque forming units (pfu) of the virus. After 2 days incubation at 37 °C, the allantoic fluid was harvested and clarified at 2,500 rpm for 20 min in a tabletop centrifuge. Virus particles were pelleted down from clarified allantoic fluid by spinning at 50,000 x g for 90 min, at 4 °C. The virus pellet was soaked overnight in phosphate buffered saline (PBS) at 4 °C. The pellet was then resuspended in PBS and loaded onto a 30-60% sucrose gradient with 15% overlay and centrifuged at 140,000 x g for 90 min at 4 °C. The band at 40% sucrose was collected and the virus was pelleted down by centrifuging at 140,000 x g for 90 min at 4 °C. The pellet was then resuspended in PBS and store at - 80 °C. The concentration of purified virus in PBS buffer was determined by BCA assay.

4.2.3 Protease treatment of virions

Purified virus equivalent to 6 µg of protein was equally divided into three groups: subtilisin-treated group, detergent and subtilisin treated group, and control group. For subtilisin-treated group, purified virus equivalent to 2 µg of protein was

incubated with 4 µg of subtilisin protease (Sigma) in 30 µl 20 mM Tris-Cl (pH 8) and 1 mM CaCl₂ for 18 hr at 37 °C. For detergent and subtilisin treated group, purified virus equivalent to 2 µg of protein was incubated with 4 µg of subtilisin protease (Sigma) in 30 µl 3 % SDS (w/v) in 20 mM Tris-Cl (pH 8) and 1 mM CaCl₂ for 18 hr at 37 °C. For the control group, purified virus equivalent to 2 µg of protein was incubated in 30 µl 20 mM Tris-Cl (pH 8) and 1 mM CaCl₂ for 18 hr at 37 °C. After protease treatment, the treated virus in each group was diluted to 50 µl with NTE buffer (100 mM NaCl, 10 mM Tris-Cl (pH 7.4), and 1 mM EDTA). Phenyl methyl sulfonyl fluoride (PMSF, sigma) was added to the final concentration of 2mM and incubated for 15 min at room temperature to inhibit the subtilisin protease. Then the virus in each group was pelleted through a 20% sucrose cushion by ultracentrifuge (222,030 x g, 2 hr, 4 °C in a TLA 120.1 rotor [Beckman]). The 50 µl solution containing the virus pellet was collected and virus was pelleted down by ultracentrifuge (222,030 x g, 2 hr, and 4 °C in a TLA 120.1 rotor [Beckman]) to reduce the sucrose concentration prior to in gel digestion. These three different fractions were subsequently subjected to proteomic characterizations or western blot analysis.

4.2.4 In gel digestion

The 14 µl solution containing virus pellet in each group was dissolved in 14 µl 25 mM ammonium bicarbonate and 2% (w/v) SDS. Solubilized virus proteins were incorporated into a polyacrylamide gel matrix in a 1.5-ml microfuge tube by adding 12 µl 40% acrylamide/bis acrylamide (29:1), 1.4 µl 1% APS and 0.6 µl 100% TEMED. After 30 min for reaction of polymerization at room temperature, the gel slice was cut into small pieces and washed three times in 25 mM ammonium bicarbonate for 15 min with vortexing for removing sucrose and two times in 25 mM ammonium bicarbonate in 50 % acetonitrile with vortexing for removing SDS, respectively. The gel pieces were vacuum dried by SpeedVac and the proteins were

then reduced by 10 mM DTT at 60 °C for 60 min and alkylated by 55 mM of iodoacetamide in ammonium bicarbonate at room temperature in dark for 60 min. The gel pieces were washed with 25 mM ammonium bicarbonate and then dehydrated with 100% acetonitrile and vacuum dried. Overnight proteolytic digestion was carried out with Trypsin in ammonium bicarbonate (Protein:Trypsin = 10:1, w/w) at 37 °C for 16 hr. Peptides were extracted from gel pieces using sequential extraction with 25 mM ammonium bicarbonate, 0.1% (v/v) TFA in 25mM ammonium bicarbonate, 0.1% (v/v) TFA in 25 mM ammonium bicarbonate containing 50% acetonitrile and 100% acetonitrile. The peptides extracted in the supernatant from the four steps were combined together, concentrated by vacuum drying and the three samples were subjected to LC-MS/MS analysis separately.

4.2.5 One-dimensional LC/MS/MS analysis

The peptide mixture in each group was reconstituted in 8 µl of solution (0.1% formic acid in LC-MS water), and injected for LC separation. An autosampler (Dionex) with a switching valve (Dionex) and a trap column (C18 PepMap100, 5 µm, 100µm × 2cm, 100 Å, Dionex) were used for loading and desalting of peptides. Peptide separation was carried out using a nano-flow LC system (Ultimate plus, Dionex) with a nanoLC column (C18 PepMap100, 2 µm, 75 µm × 250 mm, 100 Å, Dionex). A solvent gradient with buffer B (0.1% formic acid / 100% acetonitrile) increasing from 0 to 40% in 90 min was used to elute the peptides followed by a 20-min wash with 80% buffer B and a 20-min equilibration with 100% buffer A. The flow rate was kept at 500 nl/min. The column was coupled to a Linear Trap Quadrupole (LTQ) Orbitrap (Thermo Fisher Scientific) having a T-union and a nanospray tip (20 µm ID tubing, 10 µm ID tip). The spray voltage of 1.8 kV was applied through the 22 T-union. The mass spectrometer was operated in data-dependent mode to switch between MS and MS/MS. Full scan MS spectra were acquired from m/z 300 to 2000 Da in the Orbitrap mass spectrometer with a

resolution of 60,000 at m/z 400. The 6 most intense precursor ions were selected for MS/MS scans in parallel with collision induced dissociation (CID) in the linear ion trap LTQ.

4.2.6 Database searching

4.2.6.1 Integrated Proteomics Pipeline searching engine

For protein identifications, Integrated Proteomics Pipeline (IP2) software developed by Integrated Proteomics Applications, Inc (<http://www.integratedproteomics.com/>) was utilized to identify viral and host proteins [216]. The acquired raw spectra were processed with RawXtract1_9_9_2.exe into ms1 and ms2 files. These mass spectral data were uploaded into IP2 software. Files were searched against the forward and reverse Uniprot taxonomy_gallus_8578_entries database (UniProt release, 1 Sep 2010) and the forward and reverse Uniprot H1N1_Wilson_smith database (UniProt release, 1 Sep 2010). Minimum one end tryptic peptide is specified as subtilisin is non-specific protease. For static amino acid modification, carbamidomethylation (57.02146 Da) is selected. For variable modification, oxidation on methionine (15.9949 Da) and N-terminal modification (42.0106 Da) are selected. IP2 uses the ProLuCID search engine. Precursor mass tolerance was set at +/- 50 ppm in high resolution with 3 isotopic peaks. For subtilisin-treated group, data were filtered based on the following values (False positive rate 0.05, Precursor delta mass cutoff 40, Best peptide FP threshold 0.01, Best peptide delta mass threshold 20). These criteria provided a 3.57% protein false discovery rate. For detergent and subtilisin-treated group, data were filtered based on the following values (False positive rate 0.05, Precursor delta mass cutoff 30, Best peptide FP threshold 0.01, Best peptide delta mass threshold 10). These criteria provided a 0% protein false discovery rate. For control group, data were filtered based on the following values (False positive rate 0.01, Precursor delta mass cutoff 40, Best peptide FP threshold 0.004, Best

peptide delta mass threshold 34). These criteria provided a 1.28% protein false discovery rate.

4.2.6.2 Global Proteome Machine (GPM) database searching

For Influenza A Virus, in addition to Mascot search engine, the Global Proteome Machine (GPM) search engine was also used. For each of the virus strains, the peak list in the MGF format was used to search the same databases.

4.2.7 SDS-PAGE and Western blot analysis

Protein samples were mixed with 1× boiling mixture (1% SDS, 15% glycerol, 1% β-mercaptoethanol, 60 mM sodium phosphate, pH 6.8) and heated at 100 °C for 5 min. The protein samples were analyzed by SDS-PAGE using a Mini-PROTEAN 3 (Bio-Rad). In Western blotting, the proteins were transferred onto nitrocellulose membranes using a mini blotting apparatus (Bio-Rad) after which the membranes were washed with 1 X PBS and blocked for 1 hr at room temperature in 1 X PBS containing 5% non-fat milk powder and 0.05% Tween 20. The membrane was incubated with the specific primary antibody followed by the appropriate anti-mouse or anti-rabbit IgG-peroxidase conjugate (Sigma). The protein bands were visualized using the ECL protein detection reagent (Perkin Elmer), and the apparent molecular masses were estimated using protein standards (Bio-Rad).

4.2.8 Immunofluorescence microscopy

Briefly, DF1 chicken fibroblast cells grown on 13-mm glass cover slips were infected with H1N1 (A/WSN/1933) at M.O.I of 1 or 2 when cells were around 50 % confluent. The infected cells were incubated at 37 °C in DEME

medium containing 2 % FBS and 1 % penicillin/streptomycin for 16 hr. Virus infected DF1 cells were generally fixed with cold methanol (5 min) and then incubated in 2% BSA in 0.1% PBS-Tween for 1hr to permeabilise the cells and block non-specific protein-protein interactions. The infected cells were then incubated with the specific antibodies (annexin A2 ab40943, α -tubulin ab4074 and IVA NP Bei-resources) for 1 hr at room temperature in dark. The secondary antibodies were Alexa Fluor 488 goat anti-rabbit IgG (H+L) and Texas Red goat anti-mouse IgG (H+L) used at a 1/1000 dilution for 1hr. Samples were washed three times and the cover slips were mounted in DAKO and analyzed by LSM 510 confocal microscopy.

4.2.9 Transfection of siRNA Molecules

DF1 cells were transfected with siGFP, siAnnexin A2 (Sigma) at 100 nM using Lipofectamine. At 36 hr post transfection, the cells were infected with H1N1 (A/WSN/1933), and at 18 hr post infection, the cells were subjected to western blot analysis.

4.3 Results

4.3.1 Subtilisin digestion of bovine serum albumin and its compatibility with 3 % sodium dodecyl sulfate (SDS)

Subtilisin Carlsberg (Sigma) is a non specific protease discovered in *Bacillus licheniformis*. In order to examine the digestion efficiency of subtilisin, different ratios of subtilisin to bovine serum albumin (BSA) were tested (Figure 4.1). BSA was completely digested when subtilisin/BSA (w/w) reached to 1:1 at 37 °C for 1hr. Moreover, after treatment with 3% SDS as detergent, BSA was completely digested when subtilisin/BSA (w/w) was 1:10 (Figure 4.1B). It is because that BSA became more soluble by SDS treatment, resulting in easier accessibility to subtilisin digestion. PMSF (Sigma) was frequently used as protease

inhibitor to quench subtilisin digestion and 2 mM PMSF could inhibit subtilisin digestion as noted in Figure 4.1C. It also suggested that subtilisin could induce self-digestion if it was not inhibited by PMSF (Figure 4.1C lane 2, lane 6).

Figure 4.1 Subtilisin digestion of bovine serum albumin (BSA) and its digestion compatibility with 3 % Sodium dodecyl sulfate (SDS). (A) Different ratios of subtilisin to BSA substrate (w/w) in 20mM Tris-Cl (pH 8) and 1mM CaCl₂ were tested for digestion; reactions were quenched by 2mM PMSF after 1 hr digestions at 37 °C, prior to SDS-PAGE analysis and coomassie blue staining. (B) Different ratios of subtilisin to BSA substrate (w/w) in 3 % SDS 20mM Tris-Cl (pH 8) and 1mM CaCl₂ were tested for digestion; reactions were quenched by 2mM PMSF after 1hr digestions at 37 °C, prior to SDS-PAGE analysis and coomassie blue staining; quench condition (lane 1): add subtilisin 4μg and 2mM PMSF to quench before adding 2μg BSA for digestion 1 hr at 37 °C. (C) Different concentrations of PMSF were tested to quench 4μg subtilisin before adding 2μg BSA for 1hr digestion at 37 °C.

4.3.2 “Tube gel” in-gel digestion and LC-MS/MS analysis

“Tube gel” [247], an in-gel digestion approach for characterizing membrane proteins, was employed to study IAV particles. One distinguishing feature of “Tube gel” is that it could additionally incorporate high concentration of detergent (2% SDS) in the initial steps to dissolve membrane proteins, whose cleavage sites are less accessible to proteolytic enzymes due to their low solubility and hydrophobicity. And in the subsequent washing steps, detergents could be efficiently removed without affecting *1D nano* LC-MS/MS performance. IAV proteins, e.g. HA, NA, M1, M2, are structural membrane proteins associated with virus envelopes. After SDS treatment, IAV lipid envelopes would be destructed and viral membrane and structural proteins would become soluble, resulting in enhanced proteolytic digestion efficiency and more observed peptides in mass spectrometric data. Moreover, dissolved proteins, either abundant or low abundant proteins, could be trapped in a polyacrylamide gel matrix without electrophoresis. In contrast, almost all the conventional approaches in Table 4.1 employed SDS-PAGE gel and 2D-gel electrophoresis, which have preference for detecting abundant proteins since some low abundant ones remain invisible after staining. This distinguishing characteristic is capable to render improved sensitivity by characterizing low abundant viral proteins. In particular, the viral proteins associated within IAV virions display different expression levels and distinctive polypeptide lengths (Figure 4.2). Hemagglutinin, neuraminidase, nucleoprotein and matrix 1 are abundant proteins whereas matrix 2 protein, nuclear export protein and three polymerase proteins are minor ones. What is more, matrix 2 protein and nuclear export protein, which only consists of corresponding 97 and 121 amino acids, could generate fewer tryptic peptides compared to the other minor proteins whose sequences are longer, e.g. viral polymerase proteins. Therefore it demands a more sensitive method to characterize the abundant and especially the low-abundant viral proteins, as well as some host proteins incorporated within viral particles. Also, we used *1D nano* LC-MS/MS instead of *2D* LC-MS/MS to

investigate the proteome of IAV particles. In this way, it requires much less time in both in gel digestion and LC-MS/MS analysis, compared to SDS-PAGE electrophoresis and 2D LC-MS/MS. And Integrated Proteomic Pipeline (IP2) was utilized to statistically analyze the mass spectrometric data [210].

Figure 4.2 Expression level of IAV proteins and their sequence length

4.3.3 Proteomic analysis of untreated IAV particles, subtilisin-treated IAV particles, and subtilisin & 3 % SDS treated IAV particles

In our initial analysis, a small quantity of 2 μ g IAV particles purified by sucrose gradient ultracentrifuge was subjected to “Tube gel” in gel digestion followed by 1D LC-MS/MS and IP2 analysis. A total of 10 viral proteins and around 70 host proteins were identified to be co-purified with IAV particles. Nevertheless, among the 10 IAV proteins, we paradoxically identified non-structural protein 1 (NS1), which bears a nuclear localization signal and predominantly localizes in nucleus or cytoplasm of IAV infected cells [140, 145, 248]. This paradox allowed us to raise hypotheses and causes associated with the co-purification of NS1 and some host proteins. Firstly, IAV is a highly “spiked”

virus decorated with two abundant viral proteins hemagglutinin and neuraminidase, which may be prone to exteriorly adhere NS1 or host proteins derived from infected cells. Secondly, it has been observed that micro vesicles, which contain some cellular materials, could be released by cells and found in purified HIV-1 virion preparations [249-251]. Thus in our deduction, it is possible that NS1 as well as some identified host proteins are contaminants, which were either associated with IAV exterior or micro vesicles. In order to solve this paradox, a novel approach, which combined exogenous protease digestion and detergent & protease treatment, was undertaken to investigate the host proteins specifically recruited into IAV particles (Figure 4.3).

Several studies have been reported to employ exogenous protease treatment to investigate the host proteins residing inside the virions [240, 249, 252]. In our trial, subtilisin treatment of IAV particles was induced prior to *ID* LC-MS/MS analysis. Subtilisin could efficiently digest nonspecific proteins that exteriorly adhere to viral spike proteins or envelopes while the interior viral or host proteins remain intact. It is because that subtilisin could not digest lipid membranes and thus the sub-viral materials encased inside are protected from exogenous subtilisin digestion. In addition, subtilisin treatment has been demonstrated to effectively digest host proteins associated with microvesicles during HIV-1 purifications [249, 253]. In this way, we used 2:1 subtilisin/virions (w/w) for 18 hr to ensure efficient digestion of exterior proteins and microvesicles. After digestion, phenyl methyl sulfonyl fluoride (PMSF, Sigma) was added to inhibit and quench the remaining subtilisin, preventing its low-specific cleavage activities during the subsequent steps. Density ultracentrifuge through 20% sucrose cushion was undertaken to isolate digested virions as pellets; whereas digested protein contaminants would become peptides and digested microvesicles would have lighter densities, which could not pass through the 20% sucrose cushion. Two additional rounds of ultracentrifuge purifications may be required to dilute sucrose since sucrose was not compatible with LC column. The digested virion pellets were collected and

subjected to “Tube gel” in gel digestion and *ID* LC-MS/MS analysis described previously.

In parallel to subtilisin-treated IAV, we also introduced detergent in presence of exogenous protease to the IAV particles. Sodium dodecyl sulfate (SDS) is an anionic detergent and it is known that high concentration of SDS could induce dissociation of membranes and delipidation [254]. The protection of the interior of the virions was interrupted once viral envelope was solubilized by the detergent, resulting in exposure of sub-viral materials accessible to exogenous protease. Yet some detergents are potential protein denaturants incompatible with biological activities of protease. High concentration SDS was used as detergent and it was compatible with subtilisin digestion. A much less ratio of subtilisin/BSA (1:10, w/w) in 3% SDS was needed to ensure complete digestion of BSA for 1 hr, compared to the ratio of subtilisin/BSA (1:3, w/w) in SDS deficient conditions (Figure 4.1B). Therefore, virions are treated with subtilisin in 3% SDS for 18 hr, and PMSF is added after the digestion. The following steps are operated similar to protease treated group.

Figure 4.3 Experimental flowchart of subtilisin-treated, subtilisin & 3% SDS-treated, and untreated IAV virions. A total of 2 µg purified IAV virions (each group) derived from sucrose gradient ultracentrifuge were parallelly treated with subtilisin (4 µg) in 3 % SDS 20mM Tris-Cl and 1mM CaCl₂, subtilisin (4 µg) in 20mM Tris-Cl and 1mM CaCl₂, or mock-treated in 20mM Tris-Cl and 1mM CaCl₂ for incubating 18 hr at 37 °C. PMSF was added to quench the remaining subtilisin activities prior to 20 % sucrose cushion and ultracentrifuge. The virus pellet was collected and subjected to “Tube gel” in gel digestion, followed by 1D *nano*LC-MS/MS analysis and IP2 search identification.

In mock-treated IAV group (Table 4.2), a total of ten viral proteins were detected with accession number, sequence coverage, sequence count, spectrum count indicated. These were abundant proteins e.g. M1, NP, HA, NA, and low abundant proteins, e.g. PA, PB1, PB2, M2, NEP. And the sequence counts of each viral protein generally reflected the level of abundance of each protein in Figure 4.2. Detection of the minor protein M2 remains to be the most difficult with only two peptides of M2 mapped by MS. In addition, NS1 was identified in the mock-treated group. Although NS1 was an abundant protein expressed in the nucleus and cytoplasm, only two peptides of NS1 were mapped, suggesting that a limited amount of NS1 may adhere to the viral spike proteins and co-purified with IAV virions. In subtilisin-treated IAV virions, IP2 detected a total of seven viral proteins. These were abundant proteins, e.g. HA, NP, M1, and low abundant ones, e.g. PA, PB1, PB2, NEP. NS1 that found in the mock-treated group was not detected after subtilisin digestion, suggesting that NS1 was localized outside of virions and was digested by protease treatment. In addition, NA was also not detected after subtilisin treatment for 18 hr. And it was reported to be easily released by protease and thus could not pass the 20% sucrose cushion [255]. Even in the one hr subtilisin digestion trial (data not shown), only 2 peptides of NA were mapped by MS, supporting that subtilisin could efficiently remove this spike protein. Yet different from NA, HA was more resistant to the protease digestion as some peptides from external region are incompletely digested by subtilisin. However, M2 was not identified by IP2, and it might be derived from the fact that IAV particles after subtilisin treatment became more fragile and sensitive, resulting in fewer virions and less M2 after drastic movement by sucrose cushion ultracentrifuge. Another searching engine GPM weakly detected one triply charged peptide “GPSTEGVPESMREEYR” of M2, which was resulted from an incomplete tryptic digestion that there was a missed cleavage of arginine (R) in the peptide. The tandem MS spectrum of this peptide appeared credible as most of the doubly charged y ions matched accurately with the LC-MS/MS spectra at m/z 608.62 (Figure 4.4). In the subtilisin & 3% SDS treated IAV virions, none of the viral

proteins were detected as all these viral proteins lost the protection by the viral envelopes.

In respect to the host proteins, five host proteins were detected in both the mock-treated and subtilisin-treated IAV virions (Table 4.3). These were actin, tubulin, annexin A2, ezrin, and polyubiquitin-B. Actin was an abundant cellular protein and the spectra counts of actin in control and subtilisin-treated group were much higher than that of 3% SDS-subtilisin treated group, indicating that actin reached almost maximum digestion in 3% SDS-subtilisin treated group. A few actin may remain undigested due to its abundance. The spectrum counts of polyubiquitin-B did not differ much comparing with subtilisin-treated and 3% SDS-subtilisin treated group, and this may be attributed to the resistance of ubiquitin to enzymatic digestion [247]. Together with the nine viral proteins, five proteins from the host were identified as the specifically recruited protein by influenza A virions.

In addition, 53 host proteins are uniquely identified in control group (Appendix E), suggesting that they may adhere to the outside of the virions or present in sub-cellular compartments, e.g. microvesicles or exosomes, resulting in co-purifications with virions. These host proteins are also functionally important, as they would closely participate in different steps of virus infectious cycle via interactions with viral proteins and thus are more likely to be co-purified with the virions.

4.3.4 Western blot validation of host proteins incorporated into IAV particles

We next examined the proteomic result by immunoblot analysis to further verify the presence of identified viral or host proteins inside the virions (Figure 4.5). For consistency, the sample preparation and treatment are similar with

proteomic analysis, including mock, protease-treated and detergent&protease treated group. In addition, we also included mock-infected and IAV-infected DF-1 cells as control groups for the reactivity of the antibodies and size of the protein. DF-1 Cells were infected or mock infected with IAVs with MOI 1 upon reaching 60 % confluent and preceded for 24 hr infection prior to western blot analysis. Western blot analysis of selected viral proteins showed that NA, which resides on viral envelope, was detected in the protease untreated virions but not in the subtilisin-treated virions. In contrast, viral NP protein located inside of the virions was detected in both protease-treated and untreated virions. Expectedly, NP and NA were identified in virus infected DF1 cell lysate and none of these two viral proteins existed in the detergent-protease treated virions and mock-infected DF-1 cells. These observations suggest that viral NP protein is unaffected to protease treatment as it is located inside of the viral particles. In contrast, viral NA protein that resides on the surface of viral envelope was externally digested by subtilisin, which is consistent with the proteomic data that no peptides derived from NA was identified in the protease-treated virions. We also examined host factors - actin, tubulin, and annexin A2 - in the presence or absence of protease treatment. These host proteins were detected in protease treated and non-treated groups as well as IAV infected and non-infected DF-1 cells. The bands of actin and tubulin were slightly weaker in protease treated samples than those in the non-treated samples. This observation suggests that some amounts of these host proteins are associated externally with the virions while the protease treatment could largely remove them and result in weaker bands detected by immune blotting in the protease-treated IAV group. Unlike actin and tubulin, similar amounts of annexin A2 were detected in the protease-treated and mock-treated virions. These observations are also in agreement with proteomics data that actin, tubulin and annexin A2 are specifically incorporated into the virions.

Figure 4.5 Validation of host proteins and viral proteins by western blot analysis. 2 μ g purified IAV particles were either treated by subtilisin and SDS (lane 1), or digested by subtilisin for 18 hr and concentrated by 20 % sucrose cushion (lane 2), or mock-treated (lane 1) prior to western blot analysis. 15 μ g of IAV infected (lane 5) or mock infected chicken embryo fibroblasts (lane 5) were lysated and analyzed by western blot with respective antibodies against indicated proteins.

4.3.5 Viral NP is partially associated with α -tubulin and is co-enriched with annexin A2 in the virus infected cells visualized by fluorescence microscopy

DF-1 chicken fibroblast cells were infected with H1N1 (A/WSN/1933) and at 36 hr post infection, the cells were fixed and stained using appropriate antibodies to visualize the association between viral NP and host proteins - tubulin and annexin A2 - in virus infected host cells.

Microtubules are major cytoskeleton components. Tubulin has three distinct forms including α -, β -, and γ -tubulins, with α -, β -tubulin forming microtubule filaments. To examine whether IVA virus co-localizes with α -tubulin during virus infection, we employed immunofluorescence microscope to visualize the

association of α -tubulin with viral NP. In Figure 4.6, the microtubule filaments in the infected and non-infected DF-1 cells were well stained, and spread throughout the DF-1 cells as cytoskeleton components. Some of the stained NP molecules showed association with microtubules, suggesting that viral NP might partially interact with α -tubulin in the virus infected cells. In addition, long filament-like structures co-stained with NP and annexin A2 were observed in the surface of virus-infected DF-1 cells (Figure 4.7). In the non-infected cells, such filamentous structure was not observed. It has been known that IAV viruses bud and release from the cellular plasma membrane, forming elongated viral filaments during the assembly and budding step [256]. At cell surface where virus budding occurred, both annexin A2 and NP molecules were more enriched compared to other cellular regions, suggesting that new generations of IAV viruses might bud and release from the cellular membrane. Therefore, the immunofluorescence images indicate that viral NP or RNPs are partially associated with α -tubulin filaments in during virus infection and that annexin A2 and viral NP are more enriched at the cellular membrane where virus budding occurs.

Figure 4.6 Staining patterns of viral NP and host α -tubulin in IAV infected and non-infected DF-1 cells by fluorescence microscopy. (A) DF-1 Cells were infected with H1N1 (A/WSN/1933). At 36 hpi, the cells were fixed and stained using anti- α -tubulin (Green) and anti-NP (Red). (B) DF-1 Cells were mock infected with H1N1 (A/WSN/1933). At 36 hpi, the cells were fixed and stained using anti- α -tubulin (Green).

Figure 4.7 Staining patterns of viral NP and host annexin A2 in IAV infected and non-infected DF-1 cells by fluorescence microscopy. (A) DF-1 Cells were infected with H1N1 (A/WSN/1933). At 36 hpi, the cells were fixed and stained using anti- α -tubulin (Green) and anti-NP (Red). (B) DF-1 Cells were mock infected with H1N1 (A/WSN/1933). At 36 hpi, the cells were fixed and stained using anti- α -tubulin (Green).

4.3.6 Knockdown of annexin A2 in DF-1 cells does not affect IVA virus productions

Three different siRNAs against Annexin A2 were examined for their knockdown efficiencies in DF-1 cells. siGFP was introduced as control group following the same procedures. Cells were transfected with siRNAs upon reaching to about 30 % confluent. After 36 hr knockdown, cells were collected and lysised prior to western blot analysis. In Figure 4.8, all of these siRNAs against annexin A2 showed good efficiency in knockdown (> 80%) in DF-1 cells, compared to siGFP knockdown cells. In contrast, the expression levels of α -tubulin in these different siRNA-treated samples were simultaneously examined as control group. The expression levels of α -tubulin detected by western blot were similar for siAnnexin A2 knockdown and siGFP knockdown samples.

Table 4.4 siRNA designs for gallus Annexin A2 knockdown (Gene number 396297)

Target Name	Sense/Antisense	siRNA Design	Start	Target Sequence
siANX2-1	sense	GUGAUUGACUAUGA ACUGAdTdT	600	GTGATTGACT ATGAACTGA
siANX2-1	antisense	UCAGUUCAUAGUCA AUCACdTdT	600	TCAGTTCATA GTCAATCAC
siANX2-2	sense	GAAAUCAUUUGCUC ACGGAdTdT	435	GAAATCATT GCTCACGGA
siANX2-2	antisense	UCCGUGAGCAAUG AUUUCdTdT	435	TCCGTGAGCA AATGATTTC
siANX2-3	sense	CUGUGAUGUGGGAA AUGGAdTdT	1071	CTGTGATGTG GGAAATGGA
siANX2-3	antisense	UCCAUUUCCCAU CACAGdTdT	1071	TCCATTTC CATCACAG
siGFP	sense	GCUGACCCUGAAGU UCAUCdTdT	N.A	N.A
siGFP	antisense	GAUGAACUUCAGGG UCAGCdTdT	N.A	N.A

Figure 4.8 Knockdown of annexin A2 in DF-1 cells. Cells were either transfected with siGFP, siANX2-1, siANX2-2, siANX2-3 (Table 4.4) and subjected to western blot analysis probed with anti α -tubulin and annexin A2.

We next determined the effects of annexin A2 knockdown on IAV virus productions (Figure 4.9). DF-1 cells were transfected with siRNA directed against Annexin A2 or GFP. Cells were infected or mock infected with IAVs with MOI 10 at 36 hr posttransfection and preceded for 36 hr prior to western blot analysis.

For all the siRNA-treated samples that were infected with IAV, similar expression levels of NP were observed as abundantly expressed protein, suggesting that Annexin A2 knockdown plays little role in affecting IAV production. Importantly, the overall expression levels of Annexin A2 differed in IAV infected and mock infected samples that were transfected by siGFP RNA. The IAV infected samples exhibited lower expression levels of annexin A2 compared to the mock-infected samples. This observation was also reflected in IAV infected and mock-infected samples that were treated with same siRNA against Annexin A2. By contrast, the expression levels of α -tubulin are irresponsive to siRNA knockdown in all samples. These observations suggest that severe IAV virus infections (MOI 10) induce knockdown of Annexin 2 in DF-1 cells by incorporating cellular

annexin A2 into the mature virions, which is consistent with previous proteomic and western blot data.

Figure 4.8 Effect of annexin A2 knockdown on IAV infected or mock infected DF-1 cells. DF-1 cells were transfected with three siRNA against annexin A2 (Sigma) as well as a mock siRNA against GFP at 100 nM using Lipofectamine. At 36 hr post-transfection, the cells were infected with H1N1 (A/WSN/1933) at MOI 10, and at 24 h postinfection, the cells were subjected to western blot analysis probed with anti α -tubulin, annexin A2, and viral NP.

4.4 Discussion and conclusion

The propagation for influenza A virus in chicken embryos is well-established and cost-effective, yielding considerable amount of virions or vaccine strains. For instance, the FDA approved H1N1 pandemic vaccine of GlaxoSmithKline is egg-propagated (<http://www.gsk.com/media/pandemic-flu.htm>). It is of great biomedical importance to delineate gallus host components inside of H1N1 virions. In this study, we have optimized and developed a sensitive

method to rapidly detect host components and viral proteins that are specifically recruited into the IAVs. Subtilisin digestion was employed to remove contaminant proteins that attach exteriorly to the IAV particles. The protease-treated, mock-treated and detergent&protease treated virions were subjected to proteomics analysis. This method enables to differentiate viral and host proteins residing within virions from the contaminant proteins that are co-purified with IAV particles.

All viral proteins that locate on and within IAV particles were identified by proteomics in the protease-treated virions. One exception is viral protein NA, which was easily digested by subtilisin and thus was not detected. In addition, NS1, which was abundantly expressed and distributed in the nucleus of IAV infected cells, exhibited its two peptides in the mock-treated virions. In contrast, NS1 peptides were not detected after protease treatment, suggesting that NS1 is a contaminant viral protein that is co-purified with the virions. By applying this method, it also greatly reduces the list of potential host factors associated with virions that only five host proteins were consistently detected in protease-treated and non-treated groups. In addition, this method enables to enhance sample preparation analysis that we successfully identified the proteome of IAVs by analyzing 2 μ g purified virions as the starting amount. What is more, this method could also be further improved by incorporating quantitative proteomics analysis. By quantitatively recording the protein abundance, it can further improve the resolution to classify various contaminant proteins that co-exist with the virions. Five host proteins were identified in both protease-treated and mock-treated virions, including actin, tubulin, ubiquitin, ezrin, and annexin A2. Western blot analysis and immunofluorescence were employed to validate the proteomics data by examining the association between viral NP and host α -tubulin and annexin A2. Further, the effects of annexin A2 knockdown were examined with respect to virus production in IAV infected DF-1 cells. Viral NP was abundantly expressed

post 36 hr infection irresponsive to Annexin A2 knockdown, indicating that Annexin A2 plays little role in virus entry step as well as virus production.

Recent biochemical studies have revealed that Annexin A2 is a potential RNA binding protein that specifically recognizes the higher-order secondary structures of RNA complex [257, 258], which exists commonly in virus genomes. Several other researches have also revealed that annexin A2 is incorporated into other viruses, such as HCV [259], H1N1 [260], HIV [261], RSV [262] and IBV [263]. For instance, one study has specifically characterized the role of annexin A2 on interacting with IBC pseudoknot RNA for regulating the frame shifting [264]. The association of annexin A2 with these virions may be derived from the nucleic acid binding properties of annexin A2. In the case of HIV, it has been shown that annexin A2 played little role in virus assembly in a variety of cell types. And it is possible that the incorporation of host protein annexin A2 into the virions may be not necessarily an indication for function.

CHAPTER 5

Conclusions and future works

5.1 Conclusions on HDX-MS characterization of RLR activations

5.1.1 HDX-MS is an ideal tool to study the molecular mechanism of RLR signalling

Understanding the early activation process of the RLR signaling cascade promises to provide important insight into combatting viral infection and consequently will be of great biomedical interest. However, the complete delineations of the RLR activation steps require high-resolution structures of full length RIG-I and MDA5 captured in many different states. This is challenging if not impossible for the conventional methods such as crystallography and electron microscopy. To study the relatively large RLR *apo* enzymes, higher-order RNP complexes and the structural dynamics driven by ATP hydrolysis, we utilized Hydrogen/Deuterium Exchange coupled with high-resolution mass spectrometry to gain novel insights into the transitions of RIG-I and MDA5 from the resting state to the active state. In addition to full length RIG-I and MDA5, we also designed parallel HDX studies on RIG-I Δ CARDs and MDA5 Δ CARDs. This experimental design allowed us to collect high coverage data for various conformational states (varying ligands and substrates). Indeed, we achieved high sequence coverage across seven HDX time points for RIG-I, MDA5 and the corresponding CARDs-deleted constructs. Direct comparisons of the time-resolved HDX-MS snapshots across these *apo*, RNA bound and ATP stimulated enzymes enabled us to dissect the early activation steps of the RIG-I and MDA5 signaling cascades. Furthermore, our work also provided direct evidence for distinct RNA recognition and activation mechanics by using three RNA species of different lengths and secondary structure.

5.1.2 The complete surveillance of virus RNA PAMPs is ensured by the different resting states of RIG-I and MDA5

Our HDX data clearly show that human RIG-I adopts an auto-inhibitory conformation in solution. This conformation is maintained by the intra-molecular interactions between the CARD2 latch peptide Y103-114 and the HEL2i gate motifs Y566-574 and V522-539. In this *apo* state, the HEL and CTD domains are flexible and fully solvent exposed. This model based on our HDX data agrees with the crystallographic studies on duck *apo* RIG-I, so it is likely that all vertebrate RIG-I proteins share a common auto-inhibitory conformation in the resting state (Figure 2.8). In contrast, the HDX profiles of MDA5 reveal that MDA5 contains no intra-molecular contacts between the CARD2 and HEL2i domains as each domain is highly solvent-accessible in the *apo* state (Figure 3.2). In addition, MDA5 has a longer linker between the CARDS and HEL domains (~100 disordered residues) compared to that of RIG-I (~ 60 disordered residues). These intrinsic properties of RIG-I and MDA5 in their resting states ensure complete surveillance against virus infections by sensing all types of foreign RNAs which are heterogeneous both structurally and chemically.

5.1.3 RIG-I and MDA5 are activated by separate and distinct mechanisms

Our HDX data indicate that for RIG-I, RNA loosens the CARD2-HEL2i intra-molecular interactions and then ATP binding further liberates the CARDS (Table 2.1 and Figure 2.8). We demonstrate that binding of 3p10L to RIG-I triggers a dramatic structural transformation in which both the HEL and CTD domains are bound tightly around the terminus of the duplex RNA. This RNP complex formation transforms RIG-I into a partially open conformation in which the tandem CARDS are partially released and in conformational equilibrium with the gate motifs of HEL2i. We believe that ATP binding then alters the HEL-CTD-dsRNA architecture and further destabilizes the CARD2-HEL2i interaction (Table

2.1). We observe that ATP binding compresses the HEL-CTD-dsRNA complex, whereas ATP hydrolysis and ADP release result in disengagement of RNA from protein contact surfaces. Consequently, ATP drives the tandem CARDs completely away from HEL2i in a spatiotemporal manner as evidenced by the enhanced hydrogen bond exchange activities of the RIG-I tandem CARDs. Interestingly, following the addition of ATP we detected an increased deuterium uptake across RNA binding regions in the HEL and CTD domains that we attribute to thermo fluctuations from ATP-hydrolysis cycles. It is well established that the rate of amide hydrogen exchange with peripheral deuterium is governed by local stability or fluctuations [87, 228].

Despite exhibiting no cooperativity upon binding to polyIC, the terminal binding of polyIC clearly disrupts the CARD2-HEL2i interface in RIG-I (Table 2.1 and Figure 2.8). Based on our data, it is evident that RIG-I possesses inherent flexibility in recognizing distinct RNA structures and in activating CARDs (Figure 2.8 and Table 2.1), although it is also clear that the overall binding affinity of the RIG-I HEL-CTD domains to polyIC terminus is weaker compared to 3p10L with a triphosphorylated blunt end. The apo-like dynamics of RNA capping loop upon polyIC binding also suggests that RIG-I CTD specifically accommodates 5' triphosphates rather than 5' diphosphates. The CTD plays the key role in modulating the two states of RIG-I (Table 2.1). Although it is not apparent under which states RIG-I is more potently activated, the presence of multiple binding sites near the terminus of long duplex RNAs will increase the local concentration of activated RIG-I and may trigger a more potent signaling event.

In contrast, we found that the MDA5 tandem CARDs are not locked into an auto-inhibitory state and show no remarkable changes in conformation in the presence or absence of polyIC or ATP (Table 3.1 and Figure 3.2). From our HDX experiments, we confirmed that the HEL-CTD of both the full-length and truncated MDA5 form a higher order complex with polyIC in agreement with previous structural studies [41]. We identified eleven peptides on the non-RNA-binding

surface of the HEL-CTD that were resistant to HDX. These peptides – six from the front face and five from the back face – formed large buried surfaces on both sides of the HEL-CTD ring-shaped architecture to mediate the cooperative binding of MDA5 (Figure 3.3). Introduction of ATP induced little or no changes to the MDA5-polyIC complex as no disruption of the cooperative interfaces was observed in either full length MDA5 or MDA5 Δ CARDs (Table 3.1). Our data explain why RIG-I and MDA5 play non-redundant roles in sensing pathogenic RNA species; these proteins require different activation signals, which could be a combination of the following factors: chemical and structural features of the RNA substrates, active concentrations of the RNA substrates and RLRs, and the subcellular localizations of the RLR machinery [3, 45].

In addition, HDX-MS data revealed that the recognition of 3p10L by LGP2 is similar as CARDs deleted RIG-I in which HEL1 and CTD dominantly mediate hairpin RNA binding. We also observed that LGP2 forms tight interactions with polyIC mainly through HEL1 and CTD domain and that there is no cooperative binding present in LGP2 polyIC complex.

In summary, our data illustrate the early events in the RLR activation pathways, which occur in a carefully choreographed order. RIG-I and MDA5 undergo completely different molecular pathways of activation, despite sharing high sequence similarities and conserved domain arrangements. RIG-I and MDA5 exist in different resting states, however both bind preferentially to distinct RNA ligands through their conserved HEL-CTD modules. For RIG-I, RNA binding occurs preferentially at the ends of dsRNA and partially releases its autoinhibitory closed conformation; ATP then further liberates the CARDs signaling module. Alternatively, RIG-I is also able to bind terminally and becomes activated by polyIC. The structural flexibility of RIG-I enables the protein to recognize different RNA species distinctively and to activate its CARDs selectively, providing a molecular basis for the breadth of surveillance during infection by various viruses [45, 78, 265]. In contrast, MDA5 exists in a relatively open CARDs conformation

with a lower threshold of activation than RIG-I. Perhaps to compensate for this, MDA5 binds the stem of long dsRNA in a cooperative manner to form filamentous structures that place the CARDs physically close to each other. Taken together, we believe these different steps in the early activation of RIG-I and MDA5 serve the same purpose, which is to place the CARDs in a conformation and spatial arrangement to bind to their downstream adaptor protein MAVS and further activate the RLR signaling pathway by promoting the fiber-like oligomerization of MAVS [7, 65, 217].

5.2 Conclusions of proteomics studies on host factors incorporated into influenza A virions

In this chapter, I described the proteomic strategies used to specifically characterize host components that are incorporated into the mature influenza A viral particles after virus assembly and releasing. Five host proteins – actin, α -tubulin, ezrin, annexin A2 and ubiquitin – have been highlighted for further studies. Western blot, confocal microscope, and siRNA-mediated knockdown were further performed to functionally validate the proteomic data. Although the knockdown of annexin A2 was found not to compensate the productions of influenza A virus, the reasons why annexin A2 incorporates into the influenza A virions remains to be further understood. The characterization of the interaction between purified annexin A2 and viral RNA, e.g. short duplex hairpin RNA that resembles influenza A virus stem-loop RNA structures, will further help us to gain a better understanding of the role of annexin A2 involved in the highly dynamic and complex processes of virus maturation process.

5.3 Future directions

We have shown that HDX-MS is suitable to examine the activation of RLRs by RNA PAMPs as well as the distinctive inherited properties of RLRs. Recently, two new structures - a human RIG-I CARDs tetramer bound to poly-ubiquitin chains and a RIG-I CARDs-MAVS CARD complex - have been determined, providing the molecular basis for MAVS fiber formation [65, 73, 74]. However, how the CARDs become CARDs tetramer in full-length RIG-I remains to be understood. It would be interesting to construct RIG-I mutants to explore ubiquitin:CARDs:HEL2i interactions. In our HDX-MS studies on RIG-I and MDA5, we have not observed any CARDs-CARDs interactions upon binding to respective RNA PAMPs and ATP, suggesting the essential call for poly-ubiquitin chains to assist CARDs oligomerization. Thus, how poly-ubiquitin chains affect the sequential activation of RIG-I after RNA PAMP recognition and ATP binding and hydrolysis could be an interesting topic. More interestingly, we also have shown that MDA5 CARDs is drastically different from that of RIG-I. The minimal molecular basis of MDA5 CARDs oligomer that is capable for MDA5 activation along long duplex RNA remains elusive. Exploring the roles that poly-ubiquitin chains play in terms of MDA5 activation could further address its full activation mechanisms in the combination of RNA PAMPs, ATP and poly-ubiquitin chains. Also, it would be worthy to design and test the mutant constructs, which further examine the significantly changed peptides in deuterium uptake observed by HDX-MS. By answering these questions, it could help to provide valuable validation and new insights into the structural biology of RLRs and will add notably to our understanding of these proteins that play important roles in the immune response to virus infection.

As for the side project, annexin A2 has been identified to be specifically incorporated into influenza A virions during virus maturations. In the future endeavors, by using the immunofluorescence microscopy, it would be interesting to design a detailed time course experiment covering various time points in each

infection cycle to study the association of viral NP with interested host proteins. In addition, emerging evidences have revealed that annexin A2 is a potential RNA binding protein. Hydrogen/deuterium exchange mass spectrometry will be performed to investigate the conformational change and RNA binding sites of annexin A2 in response to RNA ligands with distinctive chemical and structural properties, e.g. complex secondary structures frequently present in viral RNAs. In conjugation with mutagenesis, binding assays will allow us to map out the potential binding motifs or important residues for RNA interaction. Lastly, by comparison of the sequences and structure models, we hope to provide a possible explanation for the function for annexin A2 in RNA recognition as well as influenza A virus maturation procedures.

Table 4.1 Summary of current approaches on characterizing the proteome of enveloped virions

Purified virus	Virion starting material amount	Exogenous protease treatment	Protein electrophoresis gel	Trypsin digestion	Detection method	Proteome coverage (protein numbers)	Reference
Influenza A virus	15 µg	No	SDS-PAGE Gel	In slice-gel digestion	1D LC/MS/MS	45	[240]
	100 µg	No	No	In solution digestion	2D LC/MS/MS		
Human immunodeficiency virus type-1 (HIV-1)	120 mg	Subtilisin	SDS-PAGE Gel	No	HPLC and Protein sequence analysis	8 host protein	[266]
	1 mg	Subtilisin	SDS-PAGE Gel	No	Western blot; HPLC and sequence analysis	4 host protein	[249]
Kaposi's sarcoma associated herpesvirus	Purified from 5 liters of infected cell supernatants	Trypsin	SDS-PAGE Gel	In slice-gel digestion	MALDI-MS	18	[245]
Murine cytomegalovirus	Not indicated	No	SDS-PAGE Gel	In slice-gel digestion	<i>Nano</i> LC-MS/MS	38	[200]
Human cytomegalovirus	250 µg	No	NuPAGE-Gel	In slice-gel digestion	2D LC-MS/MS	71	[188]
Vaccinia virus	50 µg	No	No	In solution digestion	2D LC-MS/MS	98	[267]
Rhesus monkey rhadinovirus	~50 µg	Protease-K	SDS-PAGE	In slice-gel digestion and in solution digestion	LC-MS/MS and 2D LC-MS/MS	43	[246]
Infectious bronchitis virus	250 µg	No	2D-Gel	In slice-gel digestion	MALDI-TOF/TOF	70	[252]

Table 4.2 Identification and comparison of viral proteins in mock, subtilisin-treated, and subtilisin-3%SDS treated virions

Viral Protein	Accession Number	Un-treated IAV particles			Subtilisin-treated IAV particles			Subtilisin&3%SDS-treated IAV particles		
		seq coverage	seq count	spec count	seq coverage	seq count	spec count	seq coverage	seq count	spec count
Matrix protein 1	sp P05777 ;tr Q20MH7 ;tr Q1K9G9	82.1	59	226	70.6	29	194	0	0	0
Nucleoprotein	sp P15682 ; tr Q20MH5	76.7	85	321	39.8	25	81	0	0	0
Hemagglutinin	sp P03454 ;tr Q1K9H3 ;tr Q20MH9	39.1	44	147	12.9	14	149	0	0	0
Polymerase basic protein 2	sp P03427 ;tr Q20MG9 ;tr Q1K9H6	28.7	22	27	9.2	6	14	0	0	0
Polymerase acidic protein	sp P15659 ; tr Q20MH2 ; tr Q1K9H4	28.8	22	27	2.2	2	3	0	0	0
Polymerase basic protein 1	sp P03430 ; tr Q20MH1 ; tr Q1K9H5	28.1	23	30	2.2	2	3	0	0	0
Matrix protein 2	sp P05780	30.9	2	4	16.5	1	1	0	0	0
Nuclear export protein	sp Q89733 ; tr Q20MH4	26.4	6	8	7.4	1	2	0	0	0
Neuraminidase	sp P03470 ; tr Q20MH6 ; tr Q1K9H1	38.9	25	69	0	0	0	0	0	0
Nonstructural protein 1	sp Q82506 ; tr Q20MH3	11.7	2	2	0	0	0	0	0	0

Table 4.3 Identification and comparison of host proteins in mock, subtilisin-treated, and subtilisin-3%SDS treated virions

Host Protein	Accession Number	Un-treated IAV particles			Subtilisin-treated IAV particles			Subtilisin&3%SDS-treated IAV particles		
		seq coverage	seq count	spec count	seq coverage	seq count	spec count	seq coverage	seq count	spec count
Actin, cytoplasmic type 2	sp Q5ZMQ2 ; sp P60706	61.9	36	63	28.2	8	20	0	0	0
Actin, cytoplasmic type 5	sp P53478	57.2	34	55	28.2	8	20	5.3	2	4
Tubulin alpha-1 chain	sp P02552	28.2	14	17	5.8	2	5	0	0	0
Tubulin beta-7 chain	sp P09244	27.7	13	21	5	2	5	0	0	0
Tubulin beta-1 chain	sp P09203	24.9	10	13	4.9	2	3	0	0	0
Ezrin	tr Q9YGW6	18.1	10	12	1.7	1	1	0	0	0
Annexin A2	sp P17785	57.5	30	38	8.3	3	8	0	0	0
Polyubiquitin-B	sp P0CG62	7.2	2	2	5.2	1	2	5.2	1	1
Guanine nucleotide-binding protein G (i) subunit	sp P50146	0	0	0	3.1	1	1	0	0	0
Fatty acid synthase	sp P12276	0	0	0	0.4	1	7	0	0	0

Figure 4.4 LC-MS/MS spectra of the triply charged peak at m/z 608.62 (M2 peptide $^{62}\text{GPSTEGVPESMREEYR}^{77}$). There is a missed cleavage of this peptide that R73 is not cleaved by trypsin, resulting in triply charged precursor peak. In the MS2 spectra of precursor peak at m/z 608.62, most y ions (y_9^{2+} , y_{10}^{2+} , y_{11}^{2+} , y_{12}^{2+} , y_{13}^{2+} , y_{14}^{2+}) match well to their theoretical score.

Reference:

1. Holland, J., et al., *Rapid evolution of RNA genomes*. Science, 1982. **215**(4540): p. 1577-85.
2. Steinhauer, D.A., E. Domingo, and J.J. Holland, *Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase*. Gene, 1992. **122**(2): p. 281-8.
3. Luo, D., *Toward a crystal-clear view of the viral RNA sensing and response by RIG-I-like receptors*. RNA Biol, 2014. **11**(1): p. 25-32.
4. Luo, D., et al., *Visualizing the determinants of viral RNA recognition by innate immune sensor RIG-I*. Structure, 2012. **20**(11): p. 1983-8.
5. Peisley, A., et al., *Structural basis for ubiquitin-mediated antiviral signal activation by RIG-I*. Nature, 2014.
6. Jiang, X., et al., *Ubiquitin-induced oligomerization of the RNA sensors RIG-I and MDA5 activates antiviral innate immune response*. Immunity, 2012. **36**(6): p. 959-73.
7. Jiang, Q.X. and Z.J. Chen, *Structural insights into the activation of RIG-I, a nanosensor for viral RNAs*. EMBO Rep, 2011. **13**(1): p. 7-8.
8. Zeng, W., et al., *Reconstitution of the RIG-I pathway reveals a signaling role of unanchored polyubiquitin chains in innate immunity*. Cell, 2010. **141**(2): p. 315-30.
9. Finlay, B.B. and G. McFadden, *Anti-immunology: evasion of the host immune system by bacterial and viral pathogens*. Cell, 2006. **124**(4): p. 767-82.
10. Litman, G.W., J.P. Rast, and S.D. Fugmann, *The origins of vertebrate adaptive immunity*. Nat Rev Immunol, 2010. **10**(8): p. 543-53.
11. Kumar, H., T. Kawai, and S. Akira, *Pathogen recognition by the innate immune system*. Int Rev Immunol, 2011. **30**(1): p. 16-34.
12. Medzhitov, R., *Toll-like receptors and innate immunity*. Nat Rev Immunol, 2001. **1**(2): p. 135-45.
13. Cui, J., et al., *JNK pathway: diseases and therapeutic potential*. Acta Pharmacol Sin, 2007. **28**(5): p. 601-8.
14. Chen, G., et al., *NOD-like receptors: role in innate immunity and inflammatory disease*. Annu Rev Pathol, 2009. **4**: p. 365-98.
15. Akira, S., S. Uematsu, and O. Takeuchi, *Pathogen recognition and innate immunity*. Cell, 2006. **124**(4): p. 783-801.
16. Sharma, S. and K.A. Fitzgerald, *Innate immune sensing of DNA*. PLoS Pathog, 2011. **7**(4): p. e1001310.
17. Yang, P., et al., *The cytosolic nucleic acid sensor LRRFIP1 mediates the production of type I interferon via a beta-catenin-dependent pathway*. Nat Immunol, 2010. **11**(6): p. 487-94.
18. Ferguson, B.J., et al., *DNA-PK is a DNA sensor for IRF-3-dependent innate immunity*. Elife, 2012. **1**: p. e00047.
19. Hornung, V., et al., *AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC*. Nature, 2009. **458**(7237): p. 514-8.

20. Geijtenbeek, T.B. and S.I. Gringhuis, *Signalling through C-type lectin receptors: shaping immune responses*. Nat Rev Immunol, 2009. **9**(7): p. 465-79.
21. van Vliet, S.J., J.J. Garcia-Vallejo, and Y. van Kooyk, *Dendritic cells and C-type lectin receptors: coupling innate to adaptive immune responses*. Immunol Cell Biol, 2008. **86**(7): p. 580-7.
22. Yoneyama, M., et al., *Shared and unique functions of the DExD/H-box helicases RIG-I, MDA5, and LGP2 in antiviral innate immunity*. J Immunol, 2005. **175**(5): p. 2851-8.
23. Yoneyama, M., et al., *The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses*. Nat Immunol, 2004. **5**(7): p. 730-7.
24. Barbalat, R., et al., *Nucleic acid recognition by the innate immune system*. Annu Rev Immunol, 2011. **29**: p. 185-214.
25. Samuel, C.E., *Antiviral actions of interferons*. Clin Microbiol Rev, 2001. **14**(4): p. 778-809, table of contents.
26. Kato, H., et al., *Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses*. Nature, 2006. **441**(7089): p. 101-5.
27. Kohlway, A., et al., *Defining the functional determinants for RNA surveillance by RIG-I*. EMBO Rep, 2013. **14**(9): p. 772-9.
28. Goubau, D., et al., *Antiviral immunity via RIG-I-mediated recognition of RNA bearing 5'-diphosphates*. Nature, 2014. **514**(7522): p. 372-5.
29. Pichlmair, A., et al., *RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates*. Science, 2006. **314**(5801): p. 997-1001.
30. Hornung, V., et al., *5'-Triphosphate RNA is the ligand for RIG-I*. Science, 2006. **314**(5801): p. 994-7.
31. Fujita, T., *Virology. Sensing viral RNA amid your own*. Science, 2006. **314**(5801): p. 935-6.
32. Kolakofsky, D., E. Kowalinski, and S. Cusack, *A structure-based model of RIG-I activation*. RNA, 2012. **18**(12): p. 2118-27.
33. Bruns, A.M. and C.M. Horvath, *Activation of RIG-I-like receptor signal transduction*. Crit Rev Biochem Mol Biol, 2012. **47**(2): p. 194-206.
34. Pyle, A.M., *Translocation and unwinding mechanisms of RNA and DNA helicases*. Annu Rev Biophys, 2008. **37**: p. 317-36.
35. Fairman-Williams, M.E., U.P. Guenther, and E. Jankowsky, *SF1 and SF2 helicases: family matters*. Curr Opin Struct Biol, 2010. **20**(3): p. 313-24.
36. Luo, D., A. Kohlway, and A.M. Pyle, *Duplex RNA activated ATPases (DRAs): platforms for RNA sensing, signaling and processing*. RNA Biol, 2013. **10**(1): p. 111-20.
37. Wang, Y., et al., *Structural and functional insights into 5'-ppp RNA pattern recognition by the innate immune receptor RIG-I*. Nat Struct Mol Biol, 2010. **17**(7): p. 781-7.
38. Lu, C., et al., *The structural basis of 5' triphosphate double-stranded RNA recognition by RIG-I C-terminal domain*. Structure, 2010. **18**(8): p. 1032-43.

39. Binder, M., et al., *Molecular mechanism of signal perception and integration by the innate immune sensor retinoic acid inducible gene-I (RIG-I)*. J Biol Chem, 2011.
40. Peisley, A., et al., *RIG-I forms signaling-competent filaments in an ATP-dependent, ubiquitin-independent manner*. Mol Cell, 2013. **51**(5): p. 573-83.
41. Wu, B., et al., *Structural Basis for dsRNA Recognition, Filament Formation, and Antiviral Signal Activation by MDA5*. Cell, 2013. **152**(1-2): p. 276-89.
42. Li, X., et al., *Structural basis of double-stranded RNA recognition by the RIG-I like receptor MDA5*. Arch Biochem Biophys, 2009. **488**(1): p. 23-33.
43. Luo, D., et al., *Structural insights into RNA recognition by RIG-I*. Cell, 2011. **147**(2): p. 409-22.
44. Berke, I.C., et al., *MDA5 assembles into a polar helical filament on dsRNA*. Proc Natl Acad Sci U S A, 2012. **109**(45): p. 18437-41.
45. Schlee, M., *Master sensors of pathogenic RNA - RIG-I like receptors*. Immunobiology, 2013.
46. Berke, I.C. and Y. Modis, *MDA5 cooperatively forms dimers and ATP-sensitive filaments upon binding double-stranded RNA*. EMBO J, 2012. **31**(7): p. 1714-26.
47. Peisley, A., et al., *Cooperative assembly and dynamic disassembly of MDA5 filaments for viral dsRNA recognition*. Proc Natl Acad Sci U S A, 2011. **108**(52): p. 21010-5.
48. Ramos, H.J. and M. Gale, Jr., *RIG-I Like Receptors and Their Signaling Crosstalk in the Regulation of Antiviral Immunity*. Curr Opin Virol, 2011. **1**(3): p. 167-176.
49. Kawai, T. and S. Akira, *Antiviral signaling through pattern recognition receptors*. J Biochem, 2007. **141**(2): p. 137-45.
50. Baum, A., R. Sachidanandam, and A. Garcia-Sastre, *Preference of RIG-I for short viral RNA molecules in infected cells revealed by next-generation sequencing*. Proc Natl Acad Sci U S A, 2010. **107**(37): p. 16303-8.
51. Loo, Y.M., et al., *Distinct RIG-I and MDA5 signaling by RNA viruses in innate immunity*. J Virol, 2008. **82**(1): p. 335-45.
52. Kato, H., K. Takahasi, and T. Fujita, *RIG-I-like receptors: cytoplasmic sensors for non-self RNA*. Immunol Rev, 2011. **243**(1): p. 91-8.
53. Triantafilou, K., et al., *Visualisation of direct interaction of MDA5 and the dsRNA replicative intermediate form of positive strand RNA viruses*. J Cell Sci, 2012. **125**(Pt 20): p. 4761-9.
54. Satoh, T., et al., *LGP2 is a positive regulator of RIG-I- and MDA5-mediated antiviral responses*. Proc Natl Acad Sci U S A, 2010. **107**(4): p. 1512-7.
55. Takahasi, K., et al., *Solution structures of cytosolic RNA sensor MDA5 and LGP2 C-terminal domains: identification of the RNA recognition loop in RIG-I-like receptors*. J Biol Chem, 2009. **284**(26): p. 17465-74.
56. Saito, T., et al., *Regulation of innate antiviral defenses through a shared repressor domain in RIG-I and LGP2*. Proc Natl Acad Sci U S A, 2007. **104**(2): p. 582-7.

57. Bruns, A.M., et al., *ATP hydrolysis enhances RNA recognition and antiviral signal transduction by the innate immune sensor, laboratory of genetics and physiology 2 (LGP2)*. J Biol Chem, 2013. **288**(2): p. 938-46.
58. Sadler, A.J. and B.R. Williams, *Interferon-inducible antiviral effectors*. Nat Rev Immunol, 2008. **8**(7): p. 559-68.
59. Levy, D.E. and J.E. Darnell, Jr., *Stats: transcriptional control and biological impact*. Nat Rev Mol Cell Biol, 2002. **3**(9): p. 651-62.
60. Garcia, M.A., et al., *Impact of protein kinase PKR in cell biology: from antiviral to antiproliferative action*. Microbiol Mol Biol Rev, 2006. **70**(4): p. 1032-60.
61. Clemens, M.J. and A. Elia, *The double-stranded RNA-dependent protein kinase PKR: structure and function*. J Interferon Cytokine Res, 1997. **17**(9): p. 503-24.
62. Kawai, T. and S. Akira, *Signaling to NF-kappaB by Toll-like receptors*. Trends Mol Med, 2007. **13**(11): p. 460-9.
63. Silverman, R.H., *Viral encounters with 2',5'-oligoadenylate synthetase and RNase L during the interferon antiviral response*. J Virol, 2007. **81**(23): p. 12720-9.
64. Haller, O. and G. Kochs, *Human MxA protein: an interferon-induced dynamin-like GTPase with broad antiviral activity*. J Interferon Cytokine Res, 2011. **31**(1): p. 79-87.
65. Xu, H., et al., *Structural basis for the prion-like MAVS filaments in antiviral innate immunity*. Elife, 2014. **3**: p. e01489.
66. Leung, D.W. and G.K. Amarasinghe, *Structural insights into RNA recognition and activation of RIG-I-like receptors*. Curr Opin Struct Biol, 2012. **22**(3): p. 297-303.
67. Kowalinski, E., et al., *Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA*. Cell, 2011. **147**(2): p. 423-35.
68. Jiang, F., et al., *Structural basis of RNA recognition and activation by innate immune receptor RIG-I*. Nature, 2011. **479**(7373): p. 423-7.
69. Civril, F., et al., *The RIG-I ATPase domain structure reveals insights into ATP-dependent antiviral signalling*. EMBO Rep, 2011. **12**(11): p. 1127-34.
70. Luo, D., et al., *Visualizing the Determinants of Viral RNA Recognition by Innate Immune Sensor RIG-I*. Structure, 2012.
71. Jiang, X., et al., *Ubiquitin-Induced Oligomerization of the RNA Sensors RIG-I and MDA5 Activates Antiviral Innate Immune Response*. Immunity, 2012.
72. Hou, F., et al., *MAVS Forms Functional Prion-like Aggregates to Activate and Propagate Antiviral Innate Immune Response Cell*, 2011. **In Press**: p. doi:10.1016/j.cell.2011.06.041.
73. Wu, B., et al., *Molecular imprinting as a signal-activation mechanism of the viral RNA sensor RIG-I*. Mol Cell, 2014. **55**(4): p. 511-23.
74. Peisley, A., et al., *Structural basis for ubiquitin-mediated antiviral signal activation by RIG-I*. Nature, 2014. **509**(7498): p. 110-4.
75. Peisley, A., et al., *Kinetic mechanism for viral dsRNA length discrimination by MDA5 filaments*. Proc Natl Acad Sci U S A, 2012.

76. Fujita, T., et al., *Triggering antiviral response by RIG-I-related RNA helicases*. Biochimie, 2007. **89**(6-7): p. 754-60.
77. Seth, R.B., L. Sun, and Z.J. Chen, *Antiviral innate immunity pathways*. Cell Res, 2006. **16**(2): p. 141-7.
78. Loo, Y.M. and M. Gale, Jr., *Immune signaling by RIG-I-like receptors*. Immunity, 2011. **34**(5): p. 680-92.
79. Beckham, S.A., et al., *Conformational rearrangements of RIG-I receptor on formation of a multiprotein:dsRNA assembly*. Nucleic Acids Res, 2013.
80. Soltani, M., et al., *Nanophotonic trapping for precise manipulation of biomolecular arrays*. Nat Nanotechnol, 2014.
81. Monico, C., et al., *Optical Methods to Study Protein-DNA Interactions in Vitro and in Living Cells at the Single-Molecule Level*. Int J Mol Sci, 2013. **14**(2): p. 3961-92.
82. Li, M. and M.D. Wang, *Unzipping single DNA molecules to study nucleosome structure and dynamics*. Methods Enzymol, 2012. **513**: p. 29-58.
83. Englander, S.W., *Hydrogen exchange and mass spectrometry: A historical perspective*. J Am Soc Mass Spectrom, 2006. **17**(11): p. 1481-9.
84. Zhang, J., et al., *Hydrogen/deuterium exchange reveals distinct agonist/partial agonist receptor dynamics within vitamin D receptor/retinoid X receptor heterodimer*. Structure, 2010. **18**(10): p. 1332-41.
85. West, G.M., et al., *Protein conformation ensembles monitored by HDX reveal a structural rationale for abscisic acid signaling protein affinities and activities*. Structure, 2013. **21**(2): p. 229-35.
86. Lisal, J., et al., *Functional visualization of viral molecular motor by hydrogen-deuterium exchange reveals transient states*. Nat Struct Mol Biol, 2005. **12**(5): p. 460-6.
87. Maity, H., et al., *Protein hydrogen exchange mechanism: local fluctuations*. Protein Sci, 2003. **12**(1): p. 153-60.
88. Hamuro, Y., et al., *Rapid analysis of protein structure and dynamics by hydrogen/deuterium exchange mass spectrometry*. Journal of biomolecular techniques : JBT, 2003. **14**(3): p. 171-82.
89. Engen, J.R., *Analysis of protein conformation and dynamics by hydrogen/deuterium exchange MS*. Analytical Chemistry, 2009. **81**(19): p. 7870-5.
90. Lisal, J., et al., *Functional visualization of viral molecular motor by hydrogen-deuterium exchange reveals transient states*. Nature structural & molecular biology, 2005. **12**(5): p. 460-6.
91. Monroe, E.B., et al., *Hydrogen/deuterium exchange analysis of HIV-1 capsid assembly and maturation*. Structure, 2010. **18**(11): p. 1483-91.
92. Morton, V.L., et al., *RNA-induced conformational changes in a viral coat protein studied by hydrogen/deuterium exchange mass spectrometry*. Physical chemistry chemical physics : PCCP, 2010. **12**(41): p. 13468-75.
93. Kong, L., et al., *Local Conformational Stability of HIV-1 gp120 in Unliganded and CD4-Bound States as Defined by Amide Hydrogen/Deuterium Exchange*. Journal of Virology, 2010. **84**(19): p. 10311-10321.

94. Englander, S.W. and J.J. Englander, *Hydrogen--tritium exchange*. Methods Enzymol, 1978. **49**: p. 24-39.
95. Smith, D.L., Y. Deng, and Z. Zhang, *Probing the non-covalent structure of proteins by amide hydrogen exchange and mass spectrometry*. J Mass Spectrom, 1997. **32**(2): p. 135-46.
96. Zhang, H.M., et al., *Enhanced digestion efficiency, peptide ionization efficiency, and sequence resolution for protein hydrogen/deuterium exchange monitored by Fourier transform ion cyclotron resonance mass spectrometry*. Anal Chem, 2008. **80**(23): p. 9034-41.
97. Hoofnagle, A.N., K.A. Resing, and N.G. Ahn, *Protein analysis by hydrogen exchange mass spectrometry*. Annu Rev Biophys Biomol Struct, 2003. **32**: p. 1-25.
98. Zheng, J., et al., *High-resolution HDX-MS reveals distinct mechanisms of RNA recognition and activation by RIG-I and MDA5*. Nucleic Acids Res, 2014.
99. Pascal, B.D., et al., *HD desktop: an integrated platform for the analysis and visualization of H/D exchange data*. J Am Soc Mass Spectrom, 2009. **20**(4): p. 601-10.
100. Kan, Z.Y., et al., *ExMS: data analysis for HX-MS experiments*. J Am Soc Mass Spectrom, 2011. **22**(11): p. 1906-15.
101. Knipe, D.M.H., Peter M., *Fields Virology, 5th Edition*. 2007.
102. Portela, A., et al., *Replication of orthomyxoviruses*. Advances in virus research, 1999. **54**: p. 319-48.
103. Rohm, C., et al., *Characterization of a novel influenza hemagglutinin, H15: criteria for determination of influenza A subtypes*. Virology, 1996. **217**(2): p. 508-16.
104. Horimoto, T. and Y. Kawaoka, *Pandemic threat posed by avian influenza A viruses*. Clinical microbiology reviews, 2001. **14**(1): p. 129-49.
105. Beigel, J.H., et al., *Avian influenza A (H5N1) infection in humans*. The New England journal of medicine, 2005. **353**(13): p. 1374-85.
106. Jian, S.W., D.P. Liu, and F.Y. Chang, *Risk assessment of human infection with avian influenza A (H7N9) virus in Taiwan*. J Formos Med Assoc, 2014.
107. Webster, R.G., et al., *Molecular mechanisms of variation in influenza viruses*. Nature, 1982. **296**(5853): p. 115-21.
108. Nelson, M.I. and E.C. Holmes, *The evolution of epidemic influenza*. Nature reviews. Genetics, 2007. **8**(3): p. 196-205.
109. Drake, J.W., *Rates of spontaneous mutation among RNA viruses*. Proceedings of the National Academy of Sciences of the United States of America, 1993. **90**(9): p. 4171-5.
110. McHardy, A.C. and B. Adams, *The role of genomics in tracking the evolution of influenza A virus*. PLoS pathogens, 2009. **5**(10): p. e1000566.
111. de Wit, E., et al., *Rapid sequencing of the non-coding regions of influenza A virus*. Journal of virological methods, 2007. **139**(1): p. 85-9.
112. Skehel, J.J. and D.C. Wiley, *Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin*. Annual review of biochemistry, 2000. **69**: p. 531-69.

113. Bos, T.J., A.R. Davis, and D.P. Nayak, *NH2-terminal hydrophobic region of influenza virus neuraminidase provides the signal function in translocation*. Proceedings of the National Academy of Sciences of the United States of America, 1984. **81**(8): p. 2327-31.
114. Ito, T., et al., *Molecular basis for the generation in pigs of influenza A viruses with pandemic potential*. Journal of Virology, 1998. **72**(9): p. 7367-7373.
115. Tisoncik, J.R., et al., *Identification of critical residues of influenza neuraminidase in viral particle release*. Virology journal, 2011. **8**: p. 14.
116. Acharya, R., et al., *Structure and mechanism of proton transport through the transmembrane tetrameric M2 protein bundle of the influenza A virus*. Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(34): p. 15075-80.
117. Pielak, R.M. and J.J. Chou, *Influenza M2 proton channels*. Biochimica et biophysica acta, 2011. **1808**(2): p. 522-9.
118. Elleman, C.J. and W.S. Barclay, *The M1 matrix protein controls the filamentous phenotype of influenza A virus*. Virology, 2004. **321**(1): p. 144-53.
119. Liu, T., J. Muller, and Z. Ye, *Association of influenza virus matrix protein with ribonucleoproteins may control viral growth and morphology*. Virology, 2002. **304**(1): p. 89-96.
120. Martin, K. and A. Helenius, *Nuclear transport of influenza virus ribonucleoproteins: the viral matrix protein (M1) promotes export and inhibits import*. Cell, 1991. **67**(1): p. 117-30.
121. Gomez-Puertas, P., et al., *Influenza virus matrix protein is the major driving force in virus budding*. Journal of Virology, 2000. **74**(24): p. 11538-47.
122. Chen, B.J., et al., *The influenza virus M2 protein cytoplasmic tail interacts with the M1 protein and influences virus assembly at the site of virus budding*. Journal of Virology, 2008. **82**(20): p. 10059-70.
123. Schmitt, A.P. and R.A. Lamb, *Influenza virus assembly and budding at the viral budzone*. Advances in virus research, 2005. **64**: p. 383-416.
124. Barman, S., et al., *Transport of viral proteins to the apical membranes and interaction of matrix protein with glycoproteins in the assembly of influenza viruses*. Virus research, 2001. **77**(1): p. 61-9.
125. Ye, Q., R.M. Krug, and Y.J. Tao, *The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA*. Nature, 2006. **444**(7122): p. 1078-82.
126. Biswas, S.K., P.L. Boutz, and D.P. Nayak, *Influenza virus nucleoprotein interacts with influenza virus polymerase proteins*. Journal of Virology, 1998. **72**(7): p. 5493-501.
127. Biswas, S.K. and D.P. Nayak, *Mutational analysis of the conserved motifs of influenza A virus polymerase basic protein 1*. Journal of Virology, 1994. **68**(3): p. 1819-26.
128. Perales, B. and J. Ortin, *The influenza A virus PB2 polymerase subunit is required for the replication of viral RNA*. J Virol, 1997. **71**(2): p. 1381-5.

129. Perez, D.R. and R.O. Donis, *Functional analysis of PA binding by influenza A virus PB1: effects on polymerase activity and viral infectivity*. J Virol, 2001. **75**(17): p. 8127-36.
130. Ulmanen, I., B.A. Broni, and R.M. Krug, *Role of two of the influenza virus core P proteins in recognizing cap 1 structures (m7GpppNm) on RNAs and in initiating viral RNA transcription*. Proceedings of the National Academy of Sciences of the United States of America, 1981. **78**(12): p. 7355-9.
131. Li, M.L., P. Rao, and R.M. Krug, *The active sites of the influenza cap-dependent endonuclease are on different polymerase subunits*. The EMBO journal, 2001. **20**(8): p. 2078-86.
132. Guilligay, D., et al., *The structural basis for cap binding by influenza virus polymerase subunit PB2*. Nature structural & molecular biology, 2008. **15**(5): p. 500-6.
133. Perales, B. and J. Ortin, *The influenza A virus PB2 polymerase subunit is required for the replication of viral RNA*. Journal of Virology, 1997. **71**(2): p. 1381-5.
134. Dias, A., et al., *The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit*. Nature, 2009. **458**(7240): p. 914-8.
135. Fukuda, M., et al., *CRM1 is responsible for intracellular transport mediated by the nuclear export signal*. Nature, 1997. **390**(6657): p. 308-311.
136. Shimizu, T., et al., *Crucial role of the influenza virus NS2 (NEP) C-terminal domain in M1 binding and nuclear export of vRNP*. FEBS letters, 2011. **585**(1): p. 41-6.
137. Robb, N.C., et al., *NS2/NEP protein regulates transcription and replication of the influenza virus RNA genome*. The Journal of general virology, 2009. **90**(Pt 6): p. 1398-407.
138. Krumbholz, A., et al., *Current knowledge on PB1-F2 of influenza A viruses*. Medical microbiology and immunology, 2011. **200**(2): p. 69-75.
139. Kosik, I., et al., *N-terminal region of the PB1-F2 protein is responsible for increased expression of influenza A viral protein PB1*. Acta virologica, 2011. **55**(1): p. 45-53.
140. Melen, K., et al., *Nuclear and nucleolar targeting of influenza A virus NS1 protein: striking differences between different virus subtypes*. Journal of Virology, 2007. **81**(11): p. 5995-6006.
141. Mukaigawa, J. and D.P. Nayak, *Two signals mediate nuclear localization of influenza virus (A/WSN/33) polymerase basic protein 2*. Journal of Virology, 1991. **65**(1): p. 245-53.
142. de la Luna, S., et al., *Influenza virus NS1 protein enhances the rate of translation initiation of viral mRNAs*. Journal of Virology, 1995. **69**(4): p. 2427-33.
143. Fortes, P., A. Beloso, and J. Ortin, *Influenza virus NS1 protein inhibits pre-mRNA splicing and blocks mRNA nucleocytoplasmic transport*. The EMBO journal, 1994. **13**(3): p. 704-12.
144. Fernandez-Sesma, A., *The influenza virus NS1 protein: inhibitor of innate and adaptive immunity*. Infectious disorders drug targets, 2007. **7**(4): p. 336-43.

145. Hale, B.G., et al., *The multifunctional NS1 protein of influenza A viruses*. The Journal of general virology, 2008. **89**(Pt 10): p. 2359-76.
146. Angata, T. and A. Varki, *Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective*. Chemical reviews, 2002. **102**(2): p. 439-69.
147. Varki, A., *Sialic acids in human health and disease*. Trends in molecular medicine, 2008. **14**(8): p. 351-60.
148. Beare, A.S. and R.G. Webster, *Replication of avian influenza viruses in humans*. Archives of virology, 1991. **119**(1-2): p. 37-42.
149. Whittaker, G.R. and S.B. Sieczkarski, *Influenza virus can enter and infect cells in the absence of clathrin-mediated endocytosis*. Journal of Virology, 2002. **76**(20): p. 10455-10464.
150. Marjuki, H., et al., *Influenza A virus-induced early activation of ERK and PI3K mediates V-ATPase-dependent intracellular pH change required for fusion*. Cellular microbiology, 2011. **13**(4): p. 587-601.
151. Wiley, D.C. and J.J. Skehel, *The structure and function of the hemagglutinin membrane glycoprotein of influenza virus*. Annual review of biochemistry, 1987. **56**: p. 365-94.
152. Thoennes, S., et al., *Analysis of residues near the fusion peptide in the influenza hemagglutinin structure for roles in triggering membrane fusion*. Virology, 2008. **370**(2): p. 403-14.
153. Li, Z.N., et al., *Length requirements for membrane fusion of influenza virus hemagglutinin peptide linkers to transmembrane or fusion peptide domains*. Journal of Virology, 2008. **82**(13): p. 6337-48.
154. Bui, M., G. Whittaker, and A. Helenius, *Effect of M1 protein and low pH on nuclear transport of influenza virus ribonucleoproteins*. Journal of Virology, 1996. **70**(12): p. 8391-401.
155. Melen, K., et al., *Importin alpha nuclear localization signal binding sites for STAT1, STAT2, and influenza A virus nucleoprotein*. The Journal of biological chemistry, 2003. **278**(30): p. 28193-200.
156. Tarendeau, F., et al., *Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit*. Nature structural & molecular biology, 2007. **14**(3): p. 229-33.
157. Naito, T., et al., *Involvement of Hsp90 in assembly and nuclear import of influenza virus RNA polymerase subunits*. Journal of Virology, 2007. **81**(3): p. 1339-49.
158. Deng, T., et al., *Role of ran binding protein 5 in nuclear import and assembly of the influenza virus RNA polymerase complex*. Journal of Virology, 2006. **80**(24): p. 11911-9.
159. Nagata, K., A. Kawaguchi, and T. Naito, *Host factors for replication and transcription of the influenza virus genome*. Reviews in medical virology, 2008. **18**(4): p. 247-60.
160. Resa-Infante, P., et al., *The host-dependent interaction of alpha-importins with influenza PB2 polymerase subunit is required for virus RNA replication*. PloS one, 2008. **3**(12): p. e3904.

161. Li, G., et al., *Heat shock protein 70 inhibits the activity of Influenza A virus ribonucleoprotein and blocks the replication of virus in vitro and in vivo*. PloS one, 2011. **6**(2): p. e16546.
162. Burgui, I., et al., *PABP1 and eIF4GI associate with influenza virus NS1 protein in viral mRNA translation initiation complexes*. The Journal of general virology, 2003. **84**(Pt 12): p. 3263-74.
163. O'Neill, R.E., J. Talon, and P. Palese, *The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins*. The EMBO journal, 1998. **17**(1): p. 288-96.
164. Kim, J., et al., *Nuclear export of influenza viral ribonucleoprotein is temperature-dependently inhibited by dissociation of viral matrix protein*. Virology, 2003. **306**(2): p. 244-253.
165. Neumann, G., M.T. Hughes, and Y. Kawaoka, *Influenza A virus NS2 protein mediates vRNP nuclear export through NES-independent interaction with hCRM1*. The EMBO journal, 2000. **19**(24): p. 6751-8.
166. Nayak, D.P., E.K. Hui, and S. Barman, *Assembly and budding of influenza virus*. Virus research, 2004. **106**(2): p. 147-65.
167. Barman, S., et al., *Role of transmembrane domain and cytoplasmic tail amino acid sequences of influenza a virus neuraminidase in raft association and virus budding*. Journal of Virology, 2004. **78**(10): p. 5258-69.
168. Leser, G.P. and R.A. Lamb, *Influenza virus assembly and budding in raft-derived microdomains: a quantitative analysis of the surface distribution of HA, NA and M2 proteins*. Virology, 2005. **342**(2): p. 215-27.
169. Ali, A., et al., *Influenza virus assembly: effect of influenza virus glycoproteins on the membrane association of M1 protein*. Journal of Virology, 2000. **74**(18): p. 8709-19.
170. Suomalainen, M., *Lipid rafts and assembly of enveloped viruses*. Traffic, 2002. **3**(10): p. 705-9.
171. Kolegraff, K., P. Bostik, and A.A. Ansari, *Characterization and role of lentivirus-associated host proteins*. Experimental biology and medicine, 2006. **231**(3): p. 252-63.
172. Lee, C.J., et al., *Cholesterol effectively blocks entry of flavivirus*. Journal of Virology, 2008. **82**(13): p. 6470-80.
173. McDonald, T.P., et al., *Evidence that the respiratory syncytial virus polymerase complex associates with lipid rafts in virus-infected cells: a proteomic analysis*. Virology, 2004. **330**(1): p. 147-157.
174. Leser, G.P. and R.A. Lamb, *Influenza virus assembly and budding in raft-derived microdomains: A quantitative analysis of the surface distribution of HA, NA and M2 proteins*. Virology, 2005. **342**(2): p. 215-227.
175. Carrasco, M., M.J. Amorim, and P. Digard, *Lipid raft-dependent targeting of the influenza A virus nucleoprotein to the apical plasma membrane*. Traffic, 2004. **5**(12): p. 979-992.
176. Bowick, G.C., et al., *Proteomic analysis of Pichinde virus infection identifies differential expression of prothymosin-alpha*. Journal of biomedicine & biotechnology, 2010. **2010**.

177. Braunagel, S.C., et al., *Determination of the protein composition of the occlusion-derived virus of Autographa californica nucleopolyhedrovirus*. Proc Natl Acad Sci U S A, 2003. **100**(17): p. 9797-802.
178. Vogels, M.W., et al., *Identification of host factors involved in coronavirus replication by quantitative proteomics analysis*. Proteomics, 2011. **11**(1): p. 64-80.
179. Lee, S.R., et al., *Bovine viral diarrhea virus infection affects the expression of proteins related to professional antigen presentation in bovine monocytes*. Biochim Biophys Acta, 2009. **1794**(1): p. 14-22.
180. Pastorino, B., et al., *Identification of cellular proteome modifications in response to West Nile virus infection*. Molecular & cellular proteomics : MCP, 2009. **8**(7): p. 1623-37.
181. Li, S., et al., *Proteomic analysis of primary porcine endothelial cells after infection by classical swine fever virus*. Biochim Biophys Acta, 2010. **1804**(9): p. 1882-8.
182. Zhao, C., et al., *Proteomic analysis of cell lines expressing small hepatitis B surface antigen revealed decreased glucose-regulated protein 78 kDa expression in association with higher susceptibility to apoptosis*. J Med Virol, 2010. **82**(1): p. 14-22.
183. Niu, D., H. Feng, and W.N. Chen, *Proteomic analysis of HBV-associated HCC: insights on mechanisms of disease onset and biomarker discovery*. J Proteomics, 2010. **73**(7): p. 1283-90.
184. Feng, H., et al., *Protein profile in HBx transfected cells: a comparative iTRAQ-coupled 2D LC-MS/MS analysis*. J Proteomics, 2010. **73**(8): p. 1421-32.
185. Zheng, Z.Z., et al., *Role of heat-shock protein 90 in hepatitis E virus capsid trafficking*. The Journal of general virology, 2010. **91**(Pt 7): p. 1728-36.
186. Johannsen, E., et al., *Proteins of purified Epstein-Barr virus*. Proc Natl Acad Sci U S A, 2004. **101**(46): p. 16286-91.
187. Skiba, M., et al., *Gene expression profiling of Pseudorabies virus (PrV) infected bovine cells by combination of transcript analysis and quantitative proteomic techniques*. Veterinary Microbiology, 2010. **143**(1): p. 14-20.
188. Varnum, S.M., et al., *Identification of proteins in human cytomegalovirus (HCMV) particles: the HCMV proteome*. Journal of Virology, 2004. **78**(20): p. 10960-6.
189. Thantrige-Don, N., et al., *Proteomic analysis of host responses to Marek's disease virus infection in spleens of genetically resistant and susceptible chickens*. Dev Comp Immunol, 2010. **34**(7): p. 699-704.
190. Franco, C.F., et al., *Monitoring virus-like particle and viral protein production by intact cell MALDI-TOF mass spectrometry*. Talanta, 2010. **80**(4): p. 1561-8.
191. Chen, L.M., et al., *iTRAQ analysis of Singapore grouper iridovirus infection in a grouper embryonic cell line*. Journal of General Virology, 2008. **89**(Pt 11): p. 2869-76.
192. Ince, I.A., et al., *Proteomic analysis of Chilo iridescent virus*. Virology, 2010. **405**(1): p. 253-8.

193. Luo, Y., et al., *Tiger frog virus can infect zebrafish cells for studying up- or down-regulated genes by proteomics approach*. Virus research, 2009. **144**(1-2): p. 171-9.
194. Xie, X. and F. Yang, *Interaction of white spot syndrome virus VP26 protein with actin*. Virology, 2005. **336**(1): p. 93-9.
195. Ovsyannikova, I.G., et al., *Identification of HLA-DRB1-bound self-peptides following measles virus infection*. J Immunol Methods, 2005. **297**(1-2): p. 153-67.
196. Zachertowska, A., D. Brewer, and D.H. Evans, *MALDI-TOF mass spectroscopy detects the capsid structural instabilities created by deleting the myxoma virus cupro-zinc SOD1 homolog M131R*. J Virol Methods, 2004. **122**(1): p. 63-72.
197. Puehler, F., et al., *An interferon-gamma-binding protein of novel structure encoded by the fowlpox virus*. The Journal of biological chemistry, 2003. **278**(9): p. 6905-11.
198. Chen, T.S., J.D. Yoder, and D.E. Hruby, *Preparation of a large hydrophobic protein for mass spectrometry analysis: vaccinia virus ATI protein*. Analytical biochemistry, 2003. **315**(2): p. 277-80.
199. Chackerian, B., L.M. Rudensey, and J. Overbaugh, *Specific N-linked and O-linked glycosylation modifications in the envelope VI domain of simian immunodeficiency virus variants that evolve in the host alter recognition by neutralizing antibodies*. Journal of Virology, 1997. **71**(10): p. 7719-27.
200. Kattenhorn, L.M., et al., *Identification of proteins associated with murine cytomegalovirus virions*. Journal of Virology, 2004. **78**(20): p. 11187-97.
201. Medvedkin, V., et al., *Mass spectrometry of murine leukemia virus core proteins*. J Virol Methods, 2010. **169**(2): p. 290-5.
202. Zandi, F., et al., *Proteomics analysis of BHK-21 cells infected with a fixed strain of rabies virus*. Proteomics, 2009. **9**(9): p. 2399-407.
203. Bourchookarn, A., et al., *Proteomic analysis of altered proteins in lymphoid organ of yellow head virus infected Penaeus monodon*. Biochim Biophys Acta, 2008. **1784**(3): p. 504-11.
204. Perkins, D.N., et al., *Probability-based protein identification by searching sequence databases using mass spectrometry data*. Electrophoresis, 1999. **20**(18): p. 3551-67.
205. Eng, J.K., A.L. McCormack, and J.R. Yates, *An Approach to Correlate Tandem Mass-Spectral Data of Peptides with Amino-Acid-Sequences in a Protein Database*. Journal of the American Society for Mass Spectrometry, 1994. **5**(11): p. 976-989.
206. Yates, J.R., *Database searching using mass spectrometry data*. Electrophoresis, 1998. **19**(6): p. 893-900.
207. Yates, J.R., et al., *Method to Correlate Tandem Mass-Spectra of Modified Peptides to Amino-Acid-Sequences in the Protein Database*. Analytical Chemistry, 1995. **67**(8): p. 1426-1436.
208. Craig, R. and R.C. Beavis, *TANDEM: matching proteins with tandem mass spectra*. Bioinformatics, 2004. **20**(9): p. 1466-7.

209. Craig, R. and R.C. Beavis, *A method for reducing the time required to match protein sequences with tandem mass spectra*. Rapid communications in mass spectrometry : RCM, 2003. **17**(20): p. 2310-6.
210. Xu, T., et al., *ProLuCID, a fast and sensitive tandem mass spectra-based protein identification program*. Molecular & Cellular Proteomics, 2006. **5**(10): p. S174-S174.
211. Tao Xu, J.V., Sung Kyu Park, Daniel Cociorva, Bingwen Lu, Lujian Liao, Johannes Hewel, Catherine C L Wong, Xuemei Han, James Wohlschlegel, John R. Yates III, *ProLuCID: using probability and statistical scores to improve sensitivity and specificity of CID and ETD database search results*.
212. Searle, B.C., *Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies*. Proteomics, 2010. **10**(6): p. 1265-9.
213. Deutsch, E.W., et al., *A guided tour of the Trans-Proteomic Pipeline*. Proteomics, 2010. **10**(6): p. 1150-1159.
214. Cociorva, D., L.T. D, and J.R. Yates, *Validation of tandem mass spectrometry database search results using DTASelect*. Current protocols in bioinformatics / editorial board, Andreas D. Baxevanis ... [et al.], 2007. **Chapter 13**: p. Unit 13 4.
215. Tabb, D.L., W.H. McDonald, and J.R. Yates, 3rd, *DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics*. Journal of proteome research, 2002. **1**(1): p. 21-6.
216. Park, S.K., et al., *A quantitative analysis software tool for mass spectrometry-based proteomics*. Nature Methods, 2008. **5**(4): p. 319-22.
217. Hou, F., et al., *MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune response*. Cell, 2011. **146**(3): p. 448-61.
218. Pippig, D.A., et al., *The regulatory domain of the RIG-I family ATPase LGP2 senses double-stranded RNA*. Nucleic Acids Res, 2009. **37**(6): p. 2014-25.
219. Pascal, B.D., et al., *HDX workbench: software for the analysis of H/D exchange MS data*. J Am Soc Mass Spectrom, 2012. **23**(9): p. 1512-21.
220. Soding, J., A. Biegert, and A.N. Lupas, *The HHpred interactive server for protein homology detection and structure prediction*. Nucleic Acids Res, 2005. **33**(Web Server issue): p. W244-8.
221. Ferrage, F., et al., *Structure and dynamics of the second CARD of human RIG-I provide mechanistic insights into regulation of RIG-I activation*. Structure, 2012. **20**(12): p. 2048-61.
222. Beemon, K.L., et al., *Rapid induction of B-cell lymphomas by avian leukosis virus*. Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, U.K, 1997. **11 Suppl 3**: p. 179-82.
223. Fang, J., J.R. Engen, and P.J. Beuning, *Escherichia coli processivity clamp beta from DNA polymerase III is dynamic in solution*. Biochemistry, 2011. **50**(26): p. 5958-68.
224. Deng, Y.H. and D.L. Smith, *Identification of unfolding domains in large proteins by their unfolding rates*. Biochemistry, 1998. **37**(18): p. 6256-6262.
225. Eyles, S.J., et al., *Protein conformational stability probed by Fourier transform ion cyclotron resonance mass spectrometry*. Journal of the American Chemical Society, 2000. **122**(3): p. 495-500.

226. Deimling, T., et al., *Crystal and solution structure of the human RIG-I SF2 domain*. Acta Crystallogr F Struct Biol Commun, 2014. **70**(Pt 8): p. 1027-31.
227. Vela, A., et al., *The thermodynamic basis for viral RNA detection by the RIG-I innate immune sensor*. J Biol Chem, 2012. **287**(51): p. 42564-73.
228. Englander, S.W. and N.R. Kallenbach, *Hydrogen exchange and structural dynamics of proteins and nucleic acids*. Q Rev Biophys, 1983. **16**(4): p. 521-655.
229. Li, X., et al., *The RIG-I-like receptor LGP2 recognizes the termini of double-stranded RNA*. J Biol Chem, 2009. **284**(20): p. 13881-91.
230. Luo, D.H., et al., *Structural Insights into RNA Recognition by RIG-I*. Cell, 2011. **147**(2): p. 409-422.
231. Komuro, A. and C.M. Horvath, *RNA- and virus-independent inhibition of antiviral signaling by RNA helicase LGP2*. J Virol, 2006. **80**(24): p. 12332-42.
232. Rodriguez, K.R., A.M. Bruns, and C.M. Horvath, *MDA5 and LGP2: Accomplices and Antagonists of Antiviral Signal Transduction*. J Virol, 2014. **88**(15): p. 8194-8200.
233. Rodriguez, K.R., A.M. Bruns, and C.M. Horvath, *MDA5 and LGP2: accomplices and antagonists of antiviral signal transduction*. J Virol, 2014. **88**(15): p. 8194-200.
234. Bruns, A.M., et al., *The innate immune sensor LGP2 activates antiviral signaling by regulating MDA5-RNA interaction and filament assembly*. Mol Cell, 2014. **55**(5): p. 771-81.
235. Bruns, A.M., et al., *The Innate Immune Sensor LGP2 Activates Antiviral Signaling by Regulating MDA5-RNA Interaction and Filament Assembly*. Mol Cell, 2014.
236. Cantin, R., S. Methot, and M.J. Tremblay, *Plunder and stowaways: incorporation of cellular proteins by enveloped viruses*. Journal of Virology, 2005. **79**(11): p. 6577-87.
237. Beniac, D.R., et al., *Conformational reorganization of the SARS coronavirus spike following receptor binding: implications for membrane fusion*. PLoS one, 2007. **2**(10): p. e1082.
238. Webster, R.G., et al., *Evolution and ecology of influenza A viruses*. Microbiological reviews, 1992. **56**(1): p. 152-79.
239. Taubenberger, J.K. and D.M. Morens, *1918 Influenza: the mother of all pandemics*. Emerging Infectious Diseases, 2006. **12**(1): p. 15-22.
240. Shaw, M.L., et al., *Cellular proteins in influenza virus particles*. PLoS pathogens, 2008. **4**(6): p. e1000085.
241. Klumpp, K., R.W. Ruigrok, and F. Baudin, *Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure*. The EMBO journal, 1997. **16**(6): p. 1248-57.
242. Richardson, J.C. and R.K. Akkina, *NS2 protein of influenza virus is found in purified virus and phosphorylated in infected cells*. Archives of virology, 1991. **116**(1-4): p. 69-80.
243. Wise, H.M., et al., *A complicated message: Identification of a novel PBI-related protein translated from influenza A virus segment 2 mRNA*. Journal of Virology, 2009. **83**(16): p. 8021-31.

244. Zheng, J., R.J. Sugrue, and K. Tang, *Mass spectrometry based proteomic studies on viruses and hosts--a review*. *Analytica chimica acta*, 2011. **702**(2): p. 149-59.
245. Bechtel, J.T., R.C. Winant, and D. Ganem, *Host and viral proteins in the virion of Kaposi's sarcoma-associated herpesvirus*. *Journal of Virology*, 2005. **79**(8): p. 4952-64.
246. O'Connor, C.M. and D.H. Kedes, *Mass spectrometric analyses of purified rhesus monkey rhadinovirus reveal 33 virion-associated proteins*. *Journal of Virology*, 2006. **80**(3): p. 1574-83.
247. Lu, X. and H. Zhu, *Tube-gel digestion: a novel proteomic approach for high throughput analysis of membrane proteins*. *Molecular & cellular proteomics : MCP*, 2005. **4**(12): p. 1948-58.
248. Han, H., et al., *New regulatory mechanisms for the intracellular localization and trafficking of influenza A virus NS1 protein revealed by comparative analysis of A/PR/8/34 and A/Sydney/5/97*. *The Journal of general virology*, 2010. **91**(Pt 12): p. 2907-17.
249. Ott, D.E., et al., *Cytoskeletal proteins inside human immunodeficiency virus type 1 virions*. *Journal of Virology*, 1996. **70**(11): p. 7734-43.
250. Trams, E.G., et al., *Exfoliation of membrane ecto-enzymes in the form of micro-vesicles*. *Biochimica et biophysica acta*, 1981. **645**(1): p. 63-70.
251. Allan, D., et al., *Release of diacylglycerol-enriched vesicles from erythrocytes with increased intracellular (Ca²⁺)*. *Nature*, 1976. **261**(5555): p. 58-60.
252. Kong, Q., et al., *Proteomic analysis of purified coronavirus infectious bronchitis virus particles*. *Proteome science*, 2010. **8**: p. 29.
253. Ott, D.E., et al., *Analysis and Localization of Cyclophilin-a Found in the Virions of Human-Immunodeficiency-Virus Type-1 Mn Strain*. *Aids Research and Human Retroviruses*, 1995. **11**(9): p. 1003-1006.
254. Becker, R., A. Helenius, and K. Simons, *Solubilization of the Semliki Forest virus membrane with sodium dodecyl sulfate*. *Biochemistry*, 1975. **14**(9): p. 1835-41.
255. Biddle, F., *The action of protease on influenza A2 virus*. *The Journal of general virology*, 1968. **2**(1): p. 19-28.
256. Vijayakrishnan, S., et al., *Cryotomography of budding influenza A virus reveals filaments with diverse morphologies that mostly do not bear a genome at their distal end*. *PLoS Pathog*, 2013. **9**(6): p. e1003413.
257. Hollas, H., et al., *Annexin A2 recognises a specific region in the 3'-UTR of its cognate messenger RNA*. *Biochim Biophys Acta*, 2006. **1763**(11): p. 1325-34.
258. Chabanon, H., et al., *An AU-rich stem-loop structure is a critical feature of the perinuclear localization signal of c-myc mRNA*. *Biochem J*, 2005. **392**(Pt 3): p. 475-83.
259. Saxena, V., et al., *Annexin A2 is involved in the formation of hepatitis C virus replication complex on the lipid raft*. *J Virol*, 2012. **86**(8): p. 4139-50.
260. Shaw, M.L., et al., *Cellular proteins in influenza virus particles*. *PLoS Pathog*, 2008. **4**(6): p. e1000085.

261. Rai, T., A. Mosoian, and M.D. Resh, *Annexin 2 is not required for human immunodeficiency virus type 1 particle production but plays a cell type-dependent role in regulating infectivity*. J Virol, 2010. **84**(19): p. 9783-92.
262. Radhakrishnan, A., et al., *Protein analysis of purified respiratory syncytial virus particles reveals an important role for heat shock protein 90 in virus particle assembly*. Mol Cell Proteomics, 2010. **9**(9): p. 1829-48.
263. Kong, Q., et al., *Proteomic analysis of purified coronavirus infectious bronchitis virus particles*. Proteome Sci, 2010. **8**: p. 29.
264. Kwak, H., M.W. Park, and S. Jeong, *Annexin A2 binds RNA and reduces the frameshifting efficiency of infectious bronchitis virus*. PLoS One, 2011. **6**(8): p. e24067.
265. Baum, A. and A. Garcia-Sastre, *Differential recognition of viral RNA by RIG-I*. Virulence, 2011. **2**(2): p. 166-9.
266. Ott, D.E., et al., *Actin-binding cellular proteins inside human immunodeficiency virus type 1*. Virology, 2000. **266**(1): p. 42-51.
267. Chung, C.S., et al., *Vaccinia virus proteome: identification of proteins in vaccinia virus intracellular mature virion particles*. Journal of Virology, 2006. **80**(5): p. 2127-40.

Appendix A. HDX-MS sequence coverage of RIG-I

Appendix B. HDX-MS sequence coverage of RIG-IΔ CARDS

Appendix C. HDX-MS sequence coverage of MDA5

Appendix D. HDX-MS sequence coverage of MDA5ΔCARDs

Appendix E. Host proteins uniquely identified in protease untreated virions by IP2

Host proteins	seq coverage ^a	seq count ^b	spec count ^c	Host proteins	seq coverage	seq count	spec count
Actin, cytoplasmic 1 2	53.9%,	23	48	Nonmuscle myosin heavy chain	3.7%,	6	6
Polyubiquitin-B	3.0%,	1	2	Pyruvate kinase muscle isozyme	2.1%,	1	1
Annexin A2	53.7%,	22	35	Isoform A2 of V-type proton ATPase catalytic subunit A	1.8%,	1	1
Tubulin beta-7 chain	28.6%,	12	24	Keratin, type I cytoskeletal 14	1.5%,	1	4
Tubulin beta-1 chain	25.2%,	10	16	Heat shock protein HSP 90-alpha	1.0%,	1	1
Tubulin alpha-1 chain	24.3%,	11	18	Putative uncharacterized protein	15.5%,	3	4
Ezrin	16.8%,	9	12	Actin, aortic smooth muscle	28.9%	14	29
Guanine nucleotide-binding protein G(i) subunit alpha-1	3.1%	1	1	Type I alpha-keratin 15	3.5%,	2	5
Elongation factor 1-alpha 1	18.0%,	8	10	Protein S100-A11	11.9%,	1	2
Peptidyl-prolyl cis-trans isomerase	15.8%,	3	4	Isoform MLC3nm of Myosin light polypeptide 6	5.3%	1	1
Fatty acid-binding protein	15.8%,	2	2	Cell division control protein 42 homolog	5.8%,	1	1
Tubulin beta-5 chain	14.8%,	6	10	14-3-3 protein gamma	3.2%,	1	1
Peroxiredoxin-1	14.6%,	3	3	Myosin regulatory light chain 2, smooth muscle minor isoform	6.4%,	1	1
Destrin	13.9%,	2	3	Isoform Smooth muscle of Tropomyosin alpha-1 chain	3.5%,	1	1

Thioredoxin	12.4%,	1	2	THO complex subunit 5 homolog 2	2.4%,	1	1
Keratin, type II cytoskeletal cochlear	12.0%,	7	12	Putative uncharacterized protein	3.5%,	1	1
Alpha-enolase	11.5%,	3	3	Protein S100-A6	14.1%,	1	1
Annexin I	10.8%,	3	3	Gallinacin-1	12.3%,	1	1
Keratin, type I cytoskeletal 19	8.5%,	4	7	Protein S100-A11	11.9%,	1	2
Heat shock cognate 71 kDa protein	8.0%,	5	6	ADP-ribosylation factor 5	11.7%,	2	3
Type II alpha-keratin IIC	7.9%,	7	15	Proteasome subunit alpha type	10.8%,	1	1
Type II alpha-keratin IIA	6.8%,	5	16	GTPase cRac1B	9.9%,	2	2
Ras-related protein Rab-11A	6.0%,	1	1	Glyceraldehyde -3-phosphate dehydrogenase	9.6%,	2	2
Claudin-1	5.7%,	1	1	Calmodulin	8.7%,	1	2
Claudin-3	5.6%,	1	2	Rho-related GTP-binding protein RhoC	5.7%,	1	1
Myosin-9	5.1%,	8	9	Apolipoprotein A-I	3.8%,	1	1
Galectin-3	4.1%,	1	2	Myosin-11	2.3%,	4	5
Fanconi anemia protein FANCD2	1.2%,	1	1	2',5'-oligo adenylyate synthetase A	3.3%,	1	1
Isoform 2 of Low-density lipoprotein receptor-related protein 1	0.5%,	1	1	Rab-GDP dissociation inhibitor	3.1%,	1	1
Spectrin alpha chain, brain	0.4%,	1	1	Myosin regulatory light chain 2, smooth muscle minor isoform	6.4%,	1	1
Isoform Beta-2 of F-actin-capping protein subunit beta isoforms 1 and 2	3.7%,	1	1				

Appendix F

Differential hydrogen/deuterium exchange data for RIG-I is summarized. Differential deuterium exchange data for RIG-I in the presence and absence of (1) 3p10L and (2) 3p10L&ATP and (3) polyIC and (4) polyIC&ATP are shown using a perturbation view by HDX Workbench. Correspondingly, differential deuterium exchange data for RIG-IΔCARDs in the presence and absence of (5) 3p10L and (6) 3p10L&ATP and (7) polyIC are shown using a perturbation view by HDX Workbench. Peptides are presented using rectangular strips below the respective portion of the protein sequence. Colors are used to characterize difference in average deuterium uptake for each peptide. The color key shows the colors assigned to deuterium differences for each peptide and the grey color represents no significant change.

Appendix F1: Differential HDX data for RIG-I ± 3p10L

Appendix F2: Differential HDX data for RIG-I±3p10L&ATP

Appendix F3: Differential HDX data for RIG-I± polyIC

[illegible]

Appendix F5: Differential HDX data for RIG-IΔCARDs±3p10L

Figure 1: Sequence logos for the 1000 Genomes Project. The figure displays sequence logos for the 1000 Genomes Project, showing the conservation of nucleotides across the genome. The logos are organized into 10 panels, each representing a different genomic region. The x-axis represents the position of the nucleotide in the genome, ranging from 1 to 1000. The y-axis represents the information content of the nucleotide, measured in bits. The logos are color-coded by nucleotide: A (blue), C (green), G (red), and T (yellow). The logos show that certain nucleotides are highly conserved in specific regions, while others are more variable. For example, in the first panel, the nucleotide 'A' is highly conserved at position 1, while in the second panel, 'C' is highly conserved at position 2. The logos also show that some regions are more conserved than others, with the first and second panels showing higher conservation than the third and fourth panels.

Appendix F7: Differential HDX data for RIG-IΔCARDs±polyIC

Appendix G

Differential hydrogen/deuterium exchange data for MDA5 is summarized. Differential deuterium exchange data for MDA5 in the presence and absence of (1) polyIC and (2) polyIC&ATP are shown using a perturbation view by HDX Workbench. Correspondingly, differential deuterium exchange data for MDA5 Δ CARs in the presence and absence of (3) polyIC and (4) polyIC&ATP are shown using a perturbation view by HDX Workbench. Peptides are presented using rectangular strips below the respective portion of the protein sequence. Colors are used to characterize difference in average deuterium uptake for each peptide. The color key shows the colors assigned to deuterium differences for each peptide and the grey color represents no significant change.

Appendix G1: Differential HDX data for MDA5±polyIC

[illegible]

Appendix G3: Differential HDX data for MDA5ACARDs±polyIC

Appendix G4: Differential HDX data for MDA5ACARDs±polyIC&ATP

Publications

- Zheng, J., et al., High resolution HDX-MS reveals distinct RNA recognition and activation by RIG-I and MDA5. *Nucleic Acid Research*, 2014.
- Zheng, J., et al., Mass spectrometry based proteomic studies on viruses and hosts. *Anal Chim Acta*, 2011. **702**(2): p. 149-59.
- Zheng, J., et al., Current approaches on viral infection: proteomics and functional validations. *Front Microbiol*, 2012. **3**: p. 393.

Posters

- Zheng, J. Protein analysis of host components within protease-treated influenza A virions. 61st ASMS Conference on Mass Spectrometry & Allied Topics. Minneapolis, USA. June 13, 2013.
- Zheng, J. HDX-MS depicts intrinsic structural rearrangements of RIG-I upon sensing viral RNA and ATP hydrolysis. 62nd ASMS Conference on Mass Spectrometry & Allied Topics. Baltimore, USA, June 19, 2014.
- Zheng, J. HDX-MS depicts intrinsic structural rearrangements of RIG-I upon sensing viral RNA and ATP hydrolysis. 5th Asia Oceania Mass Spectrometry Conference. Beijing, China, July 15, 2014

Oral presentation

- Zheng, J. HDX-MS reveals distinct RNA recognition and activation mechanism by RIG-I and MDA5. Young Scientist Forum of 5th Asia Oceania Mass Spectrometry Conference. Beijing, China, July 15, 2014