Georgia State University
ScholarWorks @ Georgia State University

Chemistry Dissertations | Department of Chemistry

5-14-2021

Structure And Function Of Viral RNA-Host Protein Interactions Among Flaviviruses

Jessica Siemer

Follow this and additional works at: https://scholarworks.gsu.edu/chemistry_diss

Recommended Citation
Siemer, Jessica, "Structure And Function Of Viral RNA-Host Protein Interactions Among Flaviviruses." Dissertation, Georgia State University, 2021.
doi: https://doi.org/10.57709/22671380

This Dissertation is brought to you for free and open access by the Department of Chemistry at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Chemistry Dissertations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

STRUCTURE AND FUNCTION OF VIRAL RNA-HOST PROTEIN INTERACTIONS AMONG FLAVIVIRUSES

by

JESSICA L SIEMER

Under the Direction of Markus W Germann, PhD

ABSTRACT

All viruses are dependent on interactions with host cellular proteins to perpetuate their lifecycles. The genus *Flavivirus*, one of the four genera that comprise the family *Flaviviridae*, contains members that cause significant morbidity and mortality worldwide. Among the mosquito-transmitted flaviviruses, West Nile Virus (WNV), dengue (DENV), and Zika virus (ZIKV) have emerged in the Western hemisphere, which now joins Southeastern Asia and Africa as an endemic region. There are no effective antiviral therapies for treating individuals with flavivirus infections, and only a few flavivirus vaccines are available for humans.

Current knowledge about the complex mechanisms that differentially regulate flavivirus positive and negative strand RNA synthesis during the various stages of the replication cycle is

incomplete. Previous work identified a multi-contact site interaction between the cell protein eEF1A and the conserved flavivirus 3′(+)SL RNA as well as an interaction between the cell protein TIAR and the viral 3′(-)SL RNA. Both of these interactions have been shown to affect viral replication.

The formation of TIAR-RNA complexes was studied using a number of biochemical and analytical chemistry techniques. Further studies were performed to examine TIAR complex formation in the context of the 3′(-) WNV SL. Finally, a structural analysis of EF1A1 was preformed using a deposited cryo-EM structure. Ultimately, elucidating the structures and functions of these host cellular protein-viral RNA complexes is essential for fully understanding flavivirus replication cycles and for the future rational design of broad spectrum anti-flavivirus therapies.

INDEX WORDS: Nucleic Acids, Flavivirus, West Nile Virus, Stem Loops, EF1A1, TIAR, RNA Binding Proteins, Virus-host Interactions, Host Factors

STRUCTURE AND FUNCTION OF VIRAL RNA-HOST PROTEIN INTERACTIONS AMONG FLAVIVIRUSES

by

JESSICA L SIEMER

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2021

Copyright by
Jessica Lynn Siemer
2021

STRUCTURE AND FUNCTION OF VIRAL RNA-HOST PROTEIN INTERACTIONS AMONG FLAVIVIRUSES

by

JESSICA L SIEMER

Committee Chair: Markus W Germann

Committee: Margo A Brinton

Gregory K Poon

W David Wilson

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2021

DEDICATION

I dedicate my dissertation to my mother, Terisita, and my grandmother, Henrietta.

ACKNOWLEDGEMENTS

I have been particularly blessed to have two wonderfully supportive mentors in Margo Brinton and Markus Germann. I am incredibly grateful for the time and effort that they have both spent guiding me in this endeavor. I am a better scientist because of them. Thank you to my committee members, Gregory Poon and David Wilson for all of your insights and guidance throughout my time at Georgia State University. The members of the Germann and Brinton labs, both past and present, have provided me with so much support, and I especially want to thank Dr. Alexander M. Spring-Connell, Dr. Marina Evich, and Dr. Joseph C. Madden for their mentorship as well as Qiushi (Shauna) Li and Thao Le for helping me to purify protein for these studies. Thank you to Dr. Matthew Brewer and all of the past and present members of the iGEM program for cultivating my teaching and mentoring skills. My experience working with this program has brought me so much happiness and many fond memories. To all of my family and friends, especially my father, John Siemer, thank you for all of your support and confidence in me over the years. I could not have done this without you. Finally, I am grateful to the Molecular Basis of Disease program at Georgia State University as well as the Provost's Dissertation Fellowship award for providing me with the financial support to complete my studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BHK: baby hamster kidney cells
CD: circular dichroism
cHP: capsid coding region hairpin
Cryo-EM: cryogenic electron microscopy
CYC: cyclization sequence
DAR: downstream AUG region
DENV: Dengue virus
DHF: dengue hemorrhagic fever
DMSO: dimethyl sulfoxide
DNA: deoxyribonucleic acid
DOSY: diffusion-ordered spectroscopy
DSS: dengue shock syndrome
DTT: dithiothreitol
eEF1A1: eukaryotic elongation factor 1 alpha 1
EMSA: electrophoretic mobility shift assay
ER: endoplasmic reticulum
FPLC: fast protein liquid chromatography
GST: glutathione-S-transferase
HCV: hepatitis C virus
HDT: host-directed therapy
HSQC: heteronuclear single-quantum correlation spectroscopy
IPTG: isopropyl β-D-thiogalactopyranoside
ITC: isothermal titration calorimetry
JEV: Japanese encephalitis virus
NCR: noncoding region
NMR: nuclear magnetic resonance
NOE: Nuclear Overhauser Effect
NOESY: Nuclear Overhauser Effect spectroscopy
ORF: open reading frame
PBS: phosphate buffered saline
PDB: protein data bank
PRD: prion-related domain
RBP: RNA binding protein
RdRp: RNA-dependent RNA polymerase
RNA: ribonucleic acid
RNP: ribonucleoprotein
RRM: RNA recognition motif
SDS PAGE: sodium dodecyl sulphate-polyacrylamide gel electrophoresis
sHP: small hairpin
SL: stem loop
TBEV: tick-borne encephalitis virus
TOCSY: total correlation spectroscopy
TIA-1: T-cell intracellular antigen 1
TIAR: T-cell intracellular antigen related protein

UAR: upstream AUG region
UV: ultraviolet light
WNV: West Nile virus
YFV: yellow fever virus
ZIKA: Zika virus

1 INTRODUCTION

1.1 The Mutability of RNA Viruses

The mutability of RNA viruses remains one of the greatest challenges in the development of antiviral therapies (1). The high mutation rate of viruses occurs for three reasons: 1) the higher error rate of viral polymerases, 2) only a few viruses with proofreading or repair mechanisms and 3) the activity of cellular RNA editing enzymes such as cytosine deaminases (2, 3). Given that host immune systems generate selective pressure on rapidly reproducing viral populations, new viral strains can emerge on a relatively short timescale. As a consequence of their smaller genomes, the majority of RNA viruses, which replicate in the cytoplasm, lack proofreading mechanisms and therefore exhibit a higher mutation rate – between 100 and 10,000-fold higher – than DNA viruses (Figure 1) (2-5). Therapeutic consequences of these high mutation rates include immune escape and the failure of both vaccines and small molecule therapeutics to produce robust protection (2, 3).

In the case of hepatitis C virus (HCV), combination therapy with NS3/4A inhibitors and NS5A or NS5B inhibitors, which target multiple components of the viral lifecycle, has been proven somewhat effective (6). However multi-drug approaches have an increased incidence of side-effects and the longevity of these types of therapies is still challenged by the rise of virus strains with multiple mutations. Additionally, while multiepitope vaccines are a promising emerging solution, development of these treatments is challenging (7).

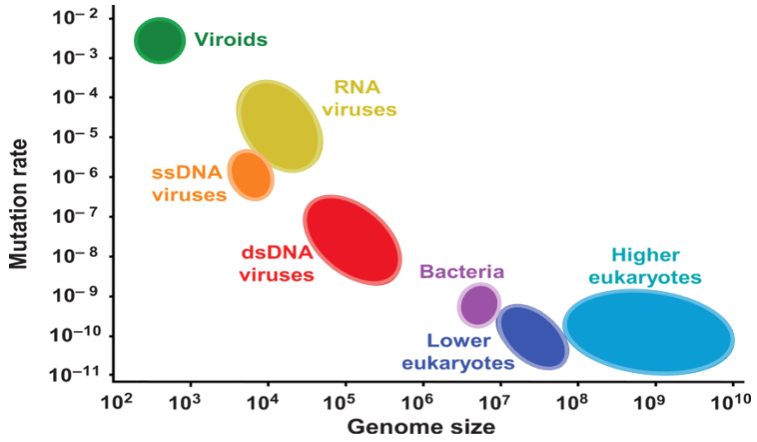


Figure 1. Mutation rate as a function of genome size. A general trend is observed that as genome size decreases, genomic mutation rate increases. This is a consequence of the loss of access to cellular or encoded proofreading and repair enzymes as genomes become smaller. RNA viruses have some of the highest observed mutation rates among viruses (5).

1.2 Host Dependency as an Antiviral Target

As a result of their limited genome size, all viruses are dependent on host-encoded proteins to support their lifecycles. While viruses can rapidly mutate, host cellular proteins are conserved; thereby, the necessity of maintaining essential viral-host cellular protein interactions constrains viral evolution (8). Therefore, host-directed therapies (HDTs) are less susceptible to viral resistance. Host cellular protein interactions exist with both viral proteins and viral nucleic acids. As the ability to study complex nucleic acid structures (pseudoknots, G-quadraplexes, riboswitches etc.) has expanded, viral nucleic acid-host proteins interactions have become more appealing as therapeutic targets. A smaller library of mutational options – 5 bases (A, C, T, G, and U) compared to 20 amino acids – combined with conserved structures increase the possibility of durable broad range therapeutics. Moreover, HDTs that specifically target viral nucleic acid-host protein interactions have less off-target effects and are therefore preferential to therapeutics with broader mechanisms such as synthetic guanidine analogs or receptor antagonists.

1.3 A General Introduction to Flaviviruses

Historically, viruses transmitted by insect vectors were divided into groups under the classification of arboviruses (9). The taxonomic families *Togaviridae* and *Flaviviridae* arose from arbovirus Groups A and B, respectively, with many of the remaining viruses becoming members of the families *Bunyaviridae* and *Reoviridae* (9, 10). Members of the family *Flaviviridae* cause significant morbidity and mortality worldwide among both humans and livestock (11-13). The family *Flaviviridae* is divided into four genera, *Flavivirus, Hepacivirus, Pestivirus* and *Pegivirus* (9-12). The over 70 viruses within the genus *Flavivirus,* have been separated into around 50 species, that are further divided into four groups: tick-borne, mosquito-borne, unknown vector, and unclassified (11).

Mosquito-borne flaviviruses are most commonly spread by the *Aedes* and *Culex* mosquito species, and viruses within this subset include dengue virus (DENV), Japanese encephalitis virus (JEV), West Nile virus (WNV), yellow fever virus (YFV) and Zika virus (ZIKV) (11). Disease caused by these viruses in humans can be characterized as encephalitic (JEV and WNV) or hemorrhagic (YFV and DENV) (13). A 2015 outbreak revealed that ZIKV can cause both encephalitic and hemorrhagic disease as well as congenital birth defects (14, 15). WNV, DENV and ZIKV have recently emerged in North and South America, adding the Western Hemisphere to Southeastern Asia and Africa as endemic regions (Figure 2) (13, 15). Over 3.5 billion people are at risk for DENV infection per year with an estimated 390 million infections occurring annually (16).

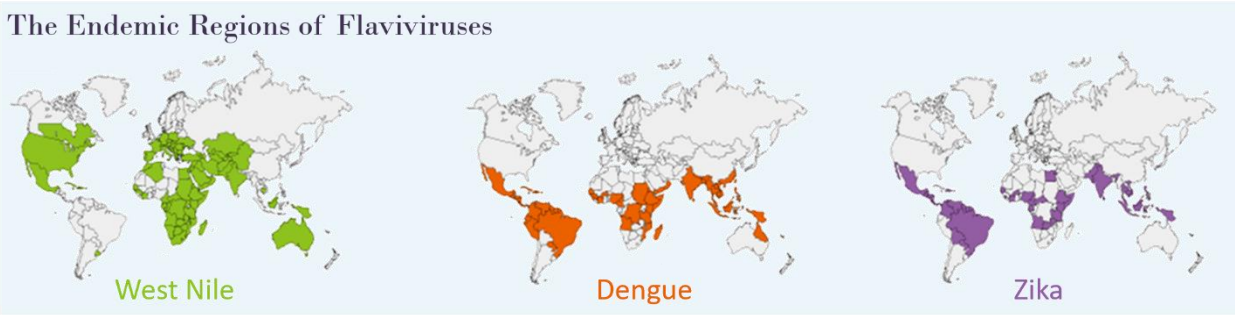


Figure 2. The Current Endemic Regions of the Flaviviruses Dengue, West Nile and Zika. Dengue, West Nile and Zika have recently spread to the Western hemisphere making all three globally pervasive. The 1999 WNV outbreak in New York demonstrated that flavivirus outbreaks are not limited to tropical regions (13, 15).

1.4 Current Antiviral Treatments for Flaviviruses

Although flaviviruses cause significant disease worldwide, anti-flavivirus treatments remain limited. In 1938, Max Theiler developed the first flavivirus vaccine, a live-attenuated vaccine for YFV, and subsequently received the 1951 Nobel Prize in Physiology or Medicine. Since then, vaccines have been developed for JEV and for Tick-borne Encephalitis Virus (TBEV), both inactivated and live-attenuated (11, 15, 17). There is an inactivated veterinary vaccine for WNV that is used in livestock; however, there is no WNV vaccine available for human use.

Antivirals for DENV are urgently needed. Each year, between 67-136 million DENV cases manifest clinically (16). The prevalence of severe disease as a result of DENV infection is increased due to immune response complications resulting from sequential infections with different serotypes. There are four serotypes of DENV (11), and all serotypes are commonly found in many of the endemic regions. It has been observed that dengue shock syndrome (DSS) and dengue hemorrhagic fever (DHF) are more prevalent after a secondary infection with another serotype of DENV (11, 18). Infections with a third and fourth serotype rarely cause severe disease (18). In 2019, a trial for a tetravalent vaccine containing 4 live-attenuated

recombinant viruses, one for each of the 4 DENV serotypes was approved by the FDA. The level of antibody protection in response to the vaccine was significantly higher in individuals that were already seropositive but reduced responses to some serotypes (DENV1 efficacy was ~58% for individuals over 9 years-old) was observed (18). Moreover, vaccination of seronegative individuals has been shown to increase the risk of DSS and DHF if these individuals do not produce adequate neutralizing antibody titers to all four serotypes and then, are subsequently naturally infected with DENV (18).

Overall, vector control strategies are still the most widely adopted and most effective methods for preventing flavivirus infection. However, the continued re-emergence of these viruses and the significant burden on global health from the diseases caused by them, highlight the continuing need for antiviral research. High throughput screening techniques have provided some promise for the discovery of flavivirus antivirals, but to date none are commercially available. Therefore, further research into the mechanisms regulating these virus' lifecycles is essential to the development of specific pan-flavivirus antivirals.

1.5 Terminal Flavivirus Genomic Structures

Flaviviruses are single-stranded, positive sense RNA viruses. The genomes are approximately 11 kb in length and have a 5′ type I cap with a 3′ terminal CU_{OH} instead of a poly-A tail (10, 11). There is a short 5′ non-coding region (NCR) (~96 nts), a single open reading frame (ORF), and a 3′ NCR (337-649 nts) (12, 19, 20). While only a few short sequences in these NCRs are conserved throughout the genus, there are genomic structures that are common across all flaviviruses (11, 19). Conserved structures at the 5′ end of the genome include: the terminal stem loop A (SLA), the SLB and the capsid coding region hairpin (cHP) (12). The 3′ end of the

genome contains two stem loops, the 3’ terminal SL and the sHP (also called the SSL) (12, 19, 21). Portions of each of these structures contain *cis*-acting elements and complementary regions such as the 3′ and 5′ cyclization (CYC) sequences, the 3′ and 5′ upstream AUG regions (UAR), and the 3′ and 5′ downstream AUG regions (DAR) that facilitate genome cyclization (11). At the 3′ terminus, the conserved sequence regions are the terminal CU_{OH}, a 5′-ACAC-3′ near the 3′ end, a five-residue motif in the top loop 5′-CACAG-3′, and the sHP (12, 19, 22, 23). Additionally, two base pairs in the stem flanked by a C/A bulge above and C/C bulge below (in the WNV genome) were identified as a metastable cis-acting element (12, 21, 22).

Deletion of either terminal stem loop, 5′ SLA or 3′ SL, is lethal to flavivirus infectious clones (12). Morpholino oligonucleotides targeting regions in the 5′ SLA were shown to negatively affect translation in both DENV and WNV infectious clones; while morpholinos targeting the 3′ SL negatively affected genome replication in addition to translation (24, 25). Mutations in terminal regions involved in long-range base-pairing, such as the 3′ and 5′ CYC, negatively affected viral replication but did not affect translation with substitutions that caused mismatches in long-range base pairs having the greatest negative effect (12, 20). Additionally, mutations in the highly conserved sequence at the top of the 3′ SL, the metastable element in the stem of the 3′SL, the top loop of the 3′ SL and the sHP negatively affected viral replication (12, 22, 23, 26).

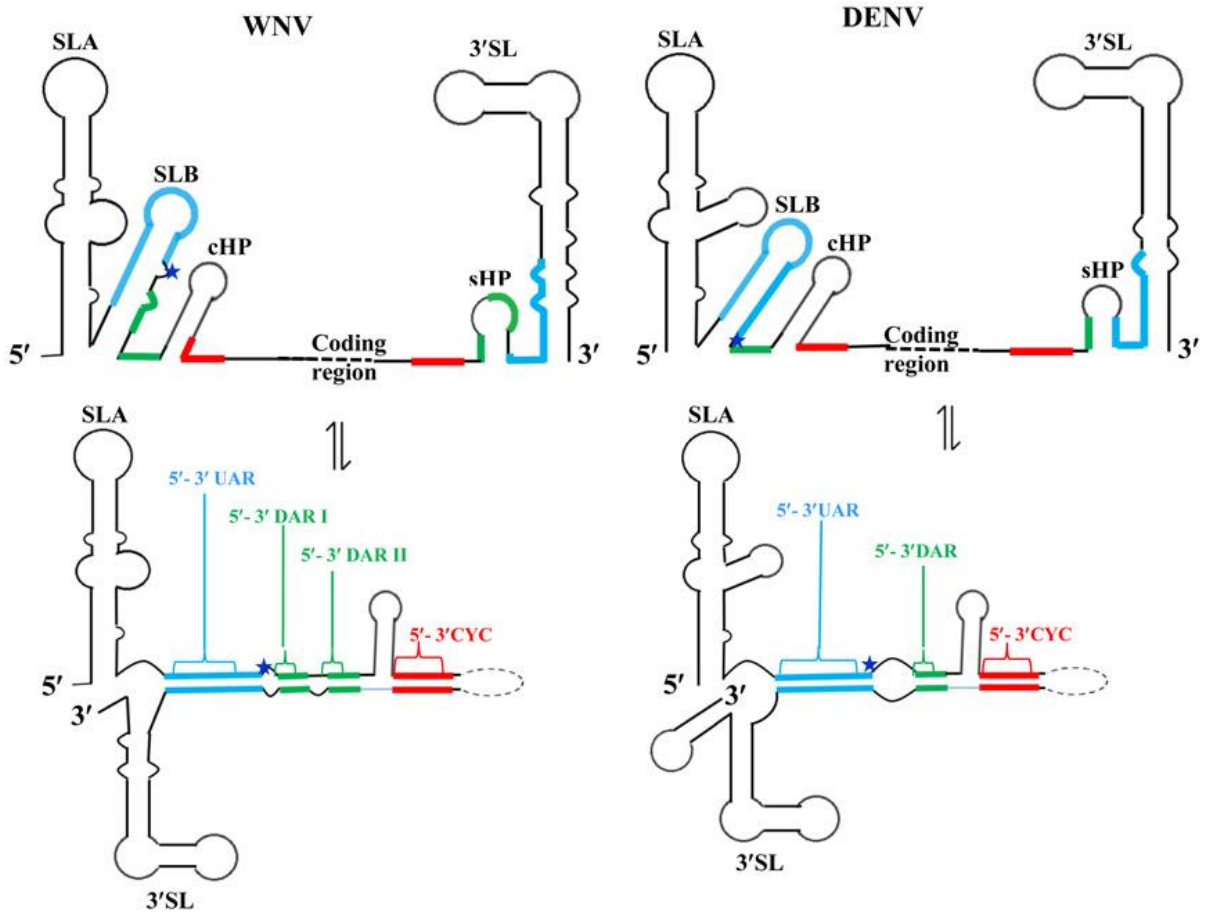


Figure 3. Terminal Structures and Cis-Acting Elements in the Flavivirus Genome. The 5′ and 3′ terminal structures of WNV (left) and DENV (right) are shown. The respective 5′ and 3′ CYC (red), DAR (green), and UAR (cyan) sequences are shown in the context of the linearized forms (top) and cyclized forms (bottom) of the genomes. Figure from Brinton, M. A., & Basu, M (12).

1.6 The Viral Replication Cycle

Flavivirus virions are small, spherical, enveloped particles that are ~50 nm in diameter (11, 19, 27). Virions achieve entry into a cell through receptor-mediated endocytosis via clathrin-coated pits (11). Several receptors have been implicated in cell entry including the glucose-regulating protein/binding immunoglobulin protein (GRP78/BiP) and C-type lectin receptors: dendritic-cell-specific ICAM-grabbing nonintegrin (DC-SIGN) and DC-SIGNR (11, 17, 19, 27).

Members of the T-cell immunoglobulin domain and muscin domain (TIM) and Tyro3, Axl and Mertk (TAM) families of phosphatidylserine receptors have also been shown to facilitate flavivirus entry (17, 19). CD14 as well as glycosaminoglycans, such as heparin, have been proposed as coreceptors (27).

Upon entry, acidification of the endosomal vesicle induces irreversible trimerization of the envelope (E) proteins facilitating the fusion of the viral endosomal membranes followed by release of the viral genome (11, 17, 27). Initially, the viral polyprotein is translated from the genome in the cytoplasm. Various signal sequences and anchor domains translocate regions of the polyprotein into the endoplasmic reticulum (ER) membrane and host and the viral NS3 protease processes process the polyprotein into the mature viral proteins (17, 27). Following processing, the nonstructural (NS) proteins form replication complexes and induce invaginations in the ER membrane (19).

For viral RNA replication to proceed, a negative-sense template strand must first be generated. An unknown mechanism regulates the long-range base-pairing of the genomic 3′ and 5′ ends to form a cyclized genome. From the cyclized form, the viral polymerase (NS5) synthesizes a single negative strand forming a double-stranded replicative intermediate (11, 17, 19). At the early-stage of infection, genome transcription is modest; however, during the late-stage of infection, genome replication is exponential (11, 19). One component of this switch in transcription efficiency is the formation of vesicles by ER invagination that contain the viral replication complexes (11, 17). These vesicles provide a protective environment for nascent RNA against cellular RNA degradation processes (11, 17, 19). However, the mechanism that mediates rapid re-initiation by NS5 of negative strand RNA synthesis remains unknown. Virion assembly occurs on the surface of the ER, with immature virions budding into the ER lumen and

moving to the Golgi where the pre-membrane (prM) protein is processed in the trans-Golgi compartment. Mature virions are then released from infected cells by exocytosis (11, 19).

1.7 Host Cellular Protein Interactions with the Viral (+) and (-) 3′ SL RNAs

As previously stated, it is common for viruses to involve host proteins in their replication cycles. A number of cellular proteins have been shown to interact with the 3′ (+) and 3′(-) SLs of flaviviruses. La and Mov34 have been shown to bind the DENV-4 and JEV 3′ (+) SL, and eukaryotic elongation factor1 alpha (eEF1A) has been shown to bind the 3′(+) SL of DENV, TBEV, WNV and YFV (28-31). The AU-rich element RNA-binding protein 1 (AUF1) has been reported to bind the 3′(+) SL of WNV, DENV, and ZIKV (32, 33). La, T-cellular antigen-1 (TIA-1), and T-cellular antigen related protein (TIAR) have been reported to bind to the 3′(-) SL (28, 34, 35). However, only EF1A1 and TIAR have been shown to specifically bind their respective viral SL binding partners and to have the ability to affect viral RNA replication efficiency (31, 35-37).

1.7.1 EF1A1

In studies with BHK cells, three proteins with molecular weights of 56, 84 and 105 kDa were identified as specifically binding to the 3′ (+) SL (30). The 50 kDa protein was later identified as eEF1A1 using Edman degradation chemical sequencing (31). EF1A is an abundant host protein that is involved in a variety of cellular processes including the transfer of aminoacylated tRNAs and GTP to the ribosomal A site during translation. EF1A1 has been shown to specifically bind to multiple flavivirus 3′ (+) SL RNAs. The binding sites on the viral RNA were subsequently mapped through foot-printing and the use of truncated or mutated RNAs resulting in the identification of one major binding site located above the metastable

region and two minor binding sites located in the top loop of the 3′ SL and the loop of the sHP (30, 31). Moreover, mutations in the viral RNA that decreased eEF1A1 binding to the 3′(+) SL RNA had a negative effect on WNV growth as well as on negative strand synthesis, and mutations that increased protein binding efficiency also increased negative strand RNA synthesis (37). It has been proposed that EF1A1 helps to regulate the switch from viral translation to negative strand synthesis by opening the bottom of the 3′SL thereby facilitating the long-range RNA interactions involved in genome cyclization (12, 37).

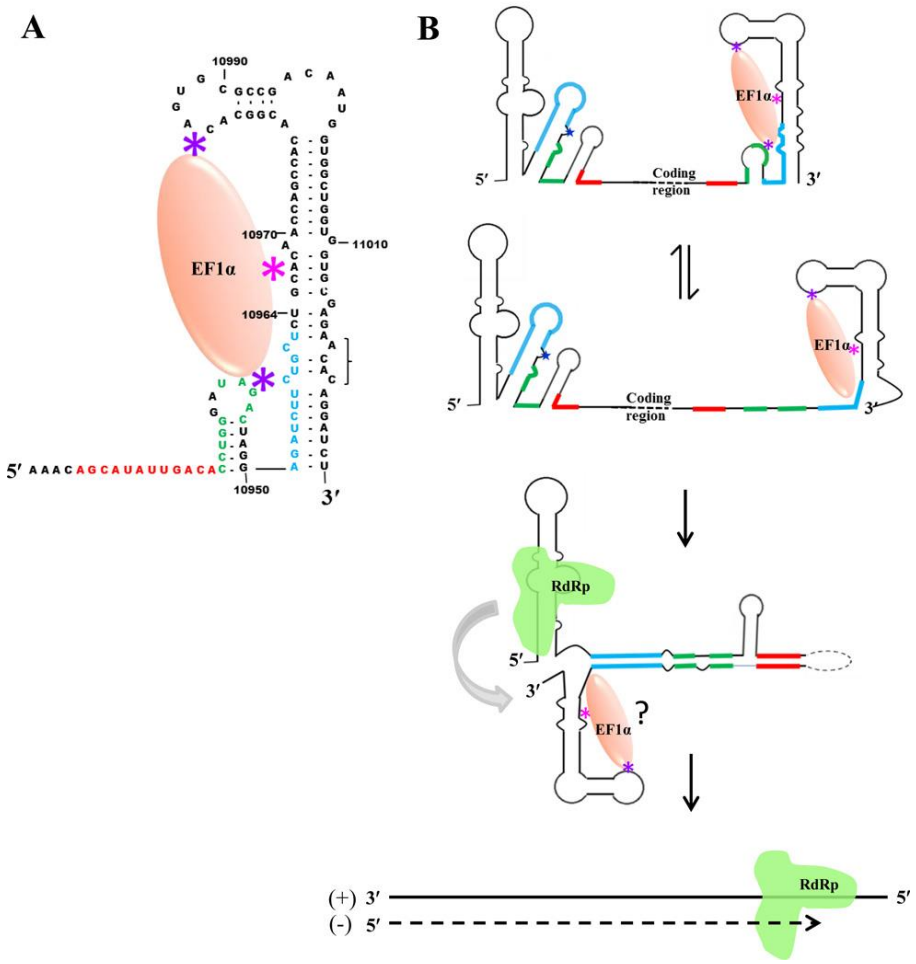


Figure 4. Proposed Mechanism for EF1A1 Regulation of Negative Strand Synthesis. A) The mapped binding sites of EF1A1 to WNV 3'(+) SL are indicated by (*). One major site accounting for 60% of the binding activity (pink) and two minor sites accounting for 20% of the binding activity each (purple) have been previously identified (29). B) Previously reported data indicate that EF1A1 binding has an effect on negative strand RNA synthesis. EF1A1 binding may facilitate the long-range RNA-RNA interaction necessary for genome cyclization by opening the base-pairs at the bottom of the 3′ (+) SL. Figure from Brinton, M. A., & Basu, M (12).

1.7.2 TIA-1/TIAR

Experiments were also performed using cytoplasmic extracts of BHK cells and in vitro synthesized WNV 3′ (-) SL as a probe (34). Four cellular proteins were detected with molecular masses of 42, 50, 60, and 108 kDa (34); the 42 kDa protein was identified as TIAR (35). In the cell, TIA-1/TIAR are involved in alternative splicing and stress granule formation. Both TIA-1 and the homolog TIAR specifically bind the 3′ (-) SL (35) in the AU-rich loop regions (35). Mutations in the 3′ (-) SL RNA that negatively affected TIA-1/TIAR binding also had a negative effect on WNV growth as well as on the efficiency of genome RNA amplification (35). TIA-1/TIAR were also shown to colocalize with viral replication complexes in WNV and DENV infected cells (38). Based on the previously reported data, TIA-1/TIAR binding to the 3′ (-) SL may facilitate rapid re-initiation at the 3′(-) end to enable the asymetric genome RNA amplification observed during late-stage infection.

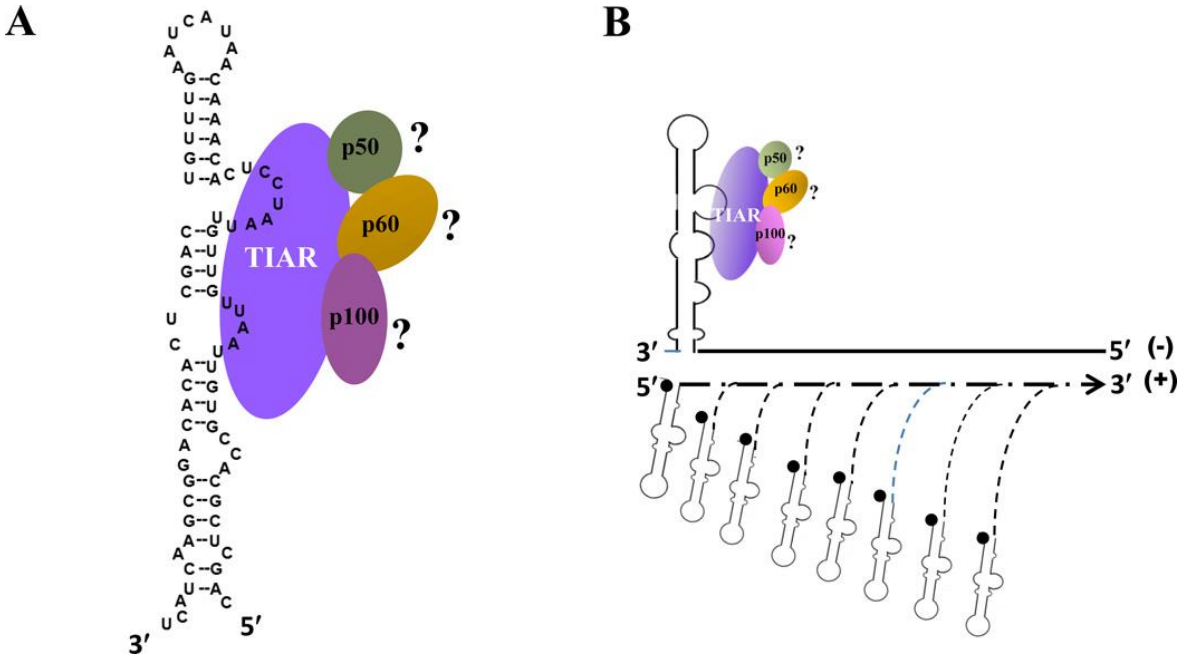


Figure 5. TIAR Binding to the 3′(-) SL. A) TIAR was identified with three other proteins to bind to the WNV 3′(-) SL. B) During genome amplification,, the 5′ (+) SL may form as the nascent strands are synthesized. 5′ caps synthesized on the nascent strands are indicated by black circles. TIAR may facilitate formation of the SLA on the negative strand template enabling rapid re-initiation of the template. Figure from Brinton, M. A., & Basu, M (12).

1.8 Goals of this Dissertation

The high binding specificities of eEF1A1 to the flavivirus 3′(+) SL and of TIA-1/TIAR to the WNV 3′(-) SL coupled with the impact of these interactions on virus replication efficiency and strand-specific viral RNA levels, makes these host cellular protein-viral RNA interactions promising candidates for further study. This dissertation aims to provide a structural and functional perspective of these interactions using analytical and biochemical techniques. Fundamental understanding of these protein-RNA interactions provides a basis for further elucidation of the complex regulatory mechanisms governing the flavivirus lifecycle and may provide novel targets for antiviral therapy development.

1.9 References

1. Lederberg, J., Hamburg, M. A., & Smolinski, M. S. (Eds.). (2003). *Microbial threats to health: emergence, detection, and response*. National Academies Press.
2. Domingo, E. (2007). Virus Evolution. In D. Knipe & P. Howly (Eds.), Fields Virology (5th ed.) (641-705). Lippincot Williams and Wilkins.
3. Sanjuán, R., & Domingo-Calap, P. (2016). Mechanisms of viral mutation. Cellular and molecular life sciences: CMLS, 73(23), 4433–4448. https://doi.org/10.1007/s00018-016-2299-6
4. Peck, K. M., & Lauring, A. S. (2018). Complexities of viral mutation rates. *Journal of virology*, *92*(14).
5. Drake, J. W. (1999). The distribution of rates of spontaneous mutation over viruses, prokaryotes, and eukaryotes. *ANNALS-NEW YORK ACADEMY OF SCIENCES*, *870*, 100-107.
6. Zając, M., Muszalska, I., Sobczak, A., Dadej, A., Tomczak, S., & Jelińska, A. (2019). Hepatitis C–New drugs and treatment prospects. European Journal of Medicinal Chemistry, 165, 225-249.
7. Zhang, L. (2018). Multi-epitope vaccines: a promising strategy against tumors and viral infections. Cellular & molecular immunology, 15(2), 182-184.
8. Kaufmann, S. H., Dorhoi, A., Hotchkiss, R. S., & Bartenschlager, R. (2018). Host-directed therapies for bacterial and viral infections. Nature Reviews Drug Discovery, 17(1), 35.
9. White, D. E., White, D. O., White, D. O., & Fenner, F. J. (1994). Medical virology. Gulf Professional Publishing.

10. Hernandez, R., Brown, D. T., & Paredes, A. (2014). Structural differences observed in arboviruses of the alphavirus and flavivirus genera. Advances in Virology, 2014.
11. Lindenbach, B.D., Thiel, H.J., & Rice, C.M. (2007). Flaviviridae: The Viruses and Their Replication. In D. Knipe & P. Howly (Eds.), Fields Virology (5th ed.) (1101-1152). Lippincot Williams and Wilkins.
12. Brinton, M. A., & Basu, M. (2015). Functions of the 3′ and 5′ genome RNA regions of members of the genus Flavivirus. Virus research, 206, 108-119.
13. Gubler, D., Kuno, D., & Markoff, L. (2007). Flaviviruses. In D. Knipe & P. Howly (Eds.), Fields Virology (5th ed.) (1101-1152). Lippincot Williams and Wilkins.
14. Mlakar, J., Korva, M., Tul, N., Popović, M., Poljšak-Prijatelj, M., Mraz, J., ... & Vizjak, A. (2016). Zika virus associated with microcephaly. New England Journal of Medicine, 374(10), 951-958.
15. Collins, M. H., & Metz, S. W. (2017). Progress and works in progress: update on flavivirus vaccine development. Clinical therapeutics, 39(8), 1519-1536.
16. Bhatt, S., Gething, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., ... & Myers, M. F. (2013). The global distribution and burden of dengue. Nature, 496(7446), 504-507.
17. Pierson, T. C., & Diamond, M. S. (2020). The continued threat of emerging flaviviruses. Nature Microbiology, 1-17.
18. World Health Organization. (2018). Dengue vaccine: WHO position paper–September 2018. Wkly Epidemiol Rec, 93(36), 457-476.
19. Brinton, M. A. (2014). Replication cycle and molecular biology of the West Nile virus. Viruses, 6(1), 13-53.

20. Basu, M., & Brinton, M. A. (2011). West Nile virus (WNV) genome RNAs with up to three adjacent mutations that disrupt long distance 5′–3′ cyclization sequence basepairs are viable. Virology, 412(1), 220-232.
21. Brinton, M. A., Fernandez, A. V., & Dispoto, J. H. (1986). The 3′-nucleotides of flavivirus genomic RNA form a conserved secondary structure. Virology, 153(1), 113-121.
22. Elghonemy, S., Davis, W. G., & Brinton, M. A. (2005). The majority of the nucleotides in the top loop of the genomic 3′ terminal stem loop structure are cis-acting in a West Nile virus infectious clone. Virology, 331(2), 238-246.
23. Tilgner, M., Deas, T. S., & Shi, P. Y. (2005). The flavivirus-conserved penta-nucleotide in the 3′ stem-loop of the West Nile virus genome requires a specific sequence and structure for RNA synthesis, but not for viral translation. Virology, 331(2), 375-386.
24. Deas, T. S., Binduga-Gajewska, I., Tilgner, M., Ren, P., Stein, D. A., Moulton, H. M., ... & Shi, P. Y. (2005). Inhibition of flavivirus infections by antisense oligomers specifically suppressing viral translation and RNA replication. Journal of virology, 79(8), 4599-4609.
25. Holden, K. L., Stein, D. A., Pierson, T. C., Ahmed, A. A., Clyde, K., Iversen, P. L., & Harris, E. (2006). Inhibition of dengue virus translation and RNA synthesis by a morpholino oligomer targeted to the top of the terminal 3′ stem–loop structure. Virology, 344(2), 439-452.
26. Davis, W. G., Basu, M., Elrod, E. J., Germann, M. W., & Brinton, M. A. (2013). Identification of cis-acting nucleotides and a structural feature in West Nile virus 3′-

terminus RNA that facilitate viral minus strand RNA synthesis. Journal of virology, 87(13), 7622-7636.

27. Mukhopadhyay, S., Kuhn, R. J., & Rossmann, M. G. (2005). A structural perspective of the flavivirus life cycle. Nature Reviews Microbiology, 3(1), 13-22.
28. Yocupicio-Monroy, M., Padmanabhan, R., Medina, F., & del Angel, R. M. (2007). Mosquito La protein binds to the 3′ untranslated region of the positive and negative polarity dengue virus RNAs and relocates to the cytoplasm of infected cells. Virology, 357(1), 29-40.
29. Ta, M., & Vrati, S. (2000). Mov34 protein from mouse brain interacts with the 3′ noncoding region of Japanese encephalitis virus. Journal of virology, 74(11), 5108-5115.
30. Blackwell, J. L., & Brinton, M. A. (1995). BHK cell proteins that bind to the 3'stem-loop structure of the West Nile virus genome RNA. Journal of Virology, 69(9), 5650-5658.
31. Blackwell, J. L., & Brinton, M. A. (1997). Translation elongation factor-1 alpha interacts with the 3'stem-loop region of West Nile virus genomic RNA. Journal of virology, 71(9), 6433-6444.
32. Friedrich, S., Schmidt, T., Geissler, R., Lilie, H., Chabierski, S., Ulbert, S., ... & Behrens, S. E. (2014). AUF1 p45 promotes West Nile virus replication by an RNA chaperone activity that supports cyclization of the viral genome. Journal of virology, 88(19), 11586-11599.
33. Friedrich, S., Engelmann, S., Schmidt, T., Szczepankiewicz, G., Bergs, S., Liebert, U. G., ... & Behrens, S. E. (2018). The host factor AUF1 p45 supports flavivirus

propagation by triggering the RNA switch required for viral genome cyclization. Journal of virology, 92(6).

34. Shi, P. Y., Li, W., & Brinton, M. A. (1996). Cell proteins bind specifically to West Nile virus minus-strand 3'stem-loop RNA. Journal of virology, 70(9), 6278-6287.
35. Li, W., Li, Y., Kedersha, N., Anderson, P., Emara, M., Swiderek, K. M., ... & Brinton, M. A. (2002). Cell proteins TIA-1 and TIAR interact with the 3′ stem-loop of the West Nile virus complementary minus-strand RNA and facilitate virus replication. Journal of virology, 76(23), 11989-12000.
36. Emara, M. M., Liu, H., Davis, W. G., & Brinton, M. A. (2008). Mutation of mapped TIA-1/TIAR binding sites in the 3′ terminal stem-loop of West Nile virus minus-strand RNA in an infectious clone negatively affects genomic RNA amplification. Journal of virology, 82(21), 10657-10670.
37. Davis, W. G., Blackwell, J. L., Shi, P. Y., & Brinton, M. A. (2007). Interaction between the cellular protein eEF1A and the 3′-terminal stem-loop of West Nile virus genomic RNA facilitates viral minus-strand RNA synthesis. Journal of virology, 81(18), 10172-10187.
38. Emara, M. M., & Brinton, M. A. (2007). Interaction of TIA-1/TIAR with West Nile and dengue virus products in infected cells interferes with stress granule formation and processing body assembly. Proceedings of the National Academy of Sciences, 104(21), 9041-9046.

2 CHARACTERIZING THE RNA-LENGTH DEPENDENT STOICHIOMETRY OF TIAR/TIA-1-RNA COMPLEXES

2.1 Preface

Model systems with reduced variables are commonly used to parse the individual contributions of elements in a complex interaction. This chapter examines the stoichiometry and formation of complexes between TIAR/TIA-1 and linear U-rich sequences of different lengths. It has become increasingly apparent through this work, that a simplified system is not always as simple as expected.

2.2 Introduction

2.2.1 RNA Recognition Motifs (RRMs)

The RNA recognition motif (RRM) is the most abundant and ubiquitously found binding domain among all of the RNA binding proteins (RBPs) in nature. Despite this, among the almost 3500 ribonucleoprotein (RNP) structures in the RCSB Protein Data Bank, less than 10% are of protein-RNA complexes containing RRMs. RRM proteins typically contain one to four RRM domains that are often connected to each other by flexible linkers and also have various auxiliary domains (1). Each RRM contains two conserved ribonucleoprotein motifs, the RNP 1 octamer ([K/R]-G-[F/Y]-[G/A]-[F/Y]-[I/L/V]-X-[F/Y]) and the RNP 2 hexamer ([I/L/V]-[F/Y]-[I/L/V]-X-N-L) (2, 3). All RRM domains are structurally similar with an antiparallel β-sheet that contain the two RNP motifs and two α-helices, located posterior to the RNA binding surface.

Despite this, structural similarity, data from binding and structural studies for different members of the RRM family of proteins have revealed a somewhat surprising diversity in single-stranded RNA binding modes and sequence recognition (2, 3) with some RRMs even

participating in protein-protein interactions rather than nucleic acid binding (4, 5). The "tunable" nature of these interactions is often attributed to the modularity of the protein structure: the spatial positioning of the RRM domains mediated by the flexible linkers, the variability in the number RRM subunits and the type of terminal domain. Binding studies involving tandem RRM domains, have also revealed different affinities and sequence preferences for individual RRM subunits within the same protein (1, 3). Some studies were able to elucidate complex interactions between proteins containing tandem RRMs and particular RNA structures – as in the case of the Prp24 protein-snRNA U6 complex (6). However, the majority of research on RRM family proteins has been focused on individual RRM subunits and short oligonucleotide sequences to reduce the complexity of both the protein and the RNA binding partner.

2.2.2 TIA-1 and TIAR

T-cell intracellular antigen (TIA-1) and T-cell intracellular antigen related protein (TIAR) are members of the RNA recognition motif (RRM) family (1). Both TIA-1 and TIAR are comprised of three RRM domains and a C-terminal glutamine-rich prion related domain (PRD) (7, 8). The two proteins have 80% amino acid sequence identity, and there are two isoforms of each protein that are generated by alternative splicing (9). The RRMs of TIA-1b and TIARb share the high sequence identity of ~90% (10). Both, TIA-1 and TIAR are able to shuttle between the cytoplasm and nucleus (11). In the nucleus, both proteins regulate alternative splicing of pre-mRNA (5) and TIA-1 has been shown to facilitate Fas-mediated apoptosis (7). In the cytoplasm, TIA-1/TIAR bind U-rich motifs in the 3′ untranslated region (UTR) of mRNAs to achieve translational regulation (12), and in response to cellular stresses, TIA-1/TIAR nucleate mRNA-protein aggregation during stress granule formation that is facilitated by their PRDs (1, 7).

Multiple studies on TIAR and TIA-1 have identified a preference for uridylate-rich RNA sequences with the RRM2 domain as the main contributor to RNA binding (1, 13, 14). Solution nuclear magnetic resonance (NMR) as well as x-ray crystallography data have shown that the RNA binding interaction occurs on the surface of the RRM2 β-sheet with additional contacts located in the flexible linker region between RRM2 and RRM3 (13-16). Although crystallization and binding characterization studies have only identified direct contacts between the RRM2 subunit and RNA sequences (10, 17), nucleic acid binding has also been observed for RRM1 as well as RRM3. Isolated RRM1 has been reported to bind DNA with a sequence preference for T and isolated RRM3 has been shown to bind C-rich RNA with high micromolar affinity (18). Based on these data, it was postulated that RRM1 and RRM3 contribute to RNA sequence recognition and stabilize the binding interaction through weaker contacts with RNAs (10). However, in the context of splicing, TIA-1 was shown to promote recruitment of the U1 snRNP through direct protein-protein interactions with U1-C and the PRD in conjunction with RRM1 (5).

Similar to studies on other members of the RRM family of proteins, the majority of the previous TIA-1/TIAR RNA binding studies only utilized short, single-stranded RNA oligos. However, multiple copies of TIA-1/TIAR were shown to bind to a 16-mer isolated from the *Fas* intron and to the 75 nt 3′SL of the antigenome minus strand RNA of West Nile virus (WNV 3′(-) SL) (19-21). These data prompted an investigation of the binding properties of TIARb with oligo U sequences of different lengths to determine the stoichiometry of the complexes and the involvement of the RRM1 and RRM3 domains in this interaction.

2.3 Materials and Methods

2.3.1 Protein expression and purification

Individual GST-tagged RRM cDNAs and an RRM123 cDNA for the hTIARb isoform (Table 1) were subcloned into the PGEX-6p-3 (GE Healthcare) vector and separately expressed in Escherichia coli, strain Rosetta 2(DE3) pLysS, cells (Novagen) as previously reported (19). The constructs of each RRM subunit include the RNA contact residues in the individual C-terminal flexible linker regions (14): RRM123 (residues 1-294), RRM1 (residues 1-91), RRM2 (residues 95-185), RRM3 (residues 187-283). Cells were incubated at 37 °C until an OD600 of 0.6-0.8 was reached. Protein expression was induced by addition of 1 mM isopropyl β-D-thiogalactopyranoside (IPTG). After an overnight post-induction incubation at 25°C, cells were harvested by centrifugation. Cell pellets were resuspended in phosphate buffered saline (PBS) buffer containing CelLytic Express (Sigma-Aldrich Co. LLC), 1 mM dithiothreitol (DTT) and 1 tablet of complete EDTA-free protease inhibitor cocktail (Roche). The cells were sonicated on ice until lysed, and the lysate was clarified by centrifugation. The lysate was then incubated with 0.2% polyethylenimine (PEI) at 4°C for half an hour. Ammonium sulfate precipitation was then performed at 75% saturation at 4°C for another half hour. The pellet was resuspended in PBS buffer and incubated with 1 mL of Glutathione-Sepharose 4B resin (GE Healthcare) overnight at 4°C. The GST tag was cleaved on the column using HRV-3C protease (Pierce). Free protein was collected from the column flow-through and several subsequent washes. For the individual RRM subunits, the collected fractions were concentrated and further purified using a Superdex-75 preparative column. RRM123 was purified using a centrifugal concentrator with a molecular weight cutoff of 50 kDa (Amicon).

Table 1. Properties of hTIARb RRM proteins. Expected molecular weights, isoelectric points (pIs), and extinction coefficients for RRM1, RRM2, RRM3 and RRM123 were calculated using ExPASy ProtParam.

Protein	MW (kDa)	pI	Extinction Coefficient @ 280 nm (M^{-1} cm^{-1})
RRM1	10.22	6.71	9970
RRM2	10.05	9.40	13980
RRM3	10.82	7.93	9970
RRM123	31.48	8.44	33920

2.3.2 RNA samples

The RNA oligos rC_5 (5'-CCCCC-3') and rA_5 (5'-AAAAA-3') used in this study were purchased from Integrated DNA Technologies, Inc. and the 2-hydroxyl protected rU_5 (5′-UUUUU-3′), rU_{20} (5′-UUUUUUUUUUUUUUUUUU-3′), N_{20} (5′-UUAAUUGUUGUUAAUCCUCUCA-3′), $rU_{15}T$ (5′-UUUUUUUUUUUUUUT-3′), fluorescein-labeled rU_{20} (FrU_{20}), rU_{15} (FrU_{15}) and rU_6 (FrU_6) were purchased from Dharmacon. The 2-hydroxyl protected RNA samples were deprotected in house following the manufacturer's protocol and desalted using a HiTrapTM desalting column (GE Healthcare). The RNA samples were then lyophilized.

2.3.3 CD spectroscopy

Circular dichroism experiments were performed on a Jasco J-1500 CD spectrometer in 1 cm quartz cuvette at 25 °C. For each titration point, four spectra were averaged from 250 to 190 nm wavelength with scan speed 20 nm/min and integrated over 0.5 s. A baseline measurement using buffer in the same cuvette was collected and subtracted from subsequent measurements.

2.3.4 Electrophoretic mobility shift assay

FrU_{20} (8-10 μM) was mixed with increasing amounts of RRM2 or RRM123 in 25 μl of storage buffer (10 mM HEPES, 1 mM DTT, 0.1 mM EDTA, 100 mM KCl, 2 mM $MgCl_2$, pH 7.3). Sucrose was added to achieve adequate density for loading into the gel. The sample mixture was incubated at 37°C for 30 minutes prior to electrophoresis. Samples were analyzed on nondenaturing polyacrylamide gels (10%-12%) made from a 40% acrylamide/Bis (19:1) solution. A voltage of 12 V/cm was applied for 3 hours. After electrophoresis, RNA on the gel was visualized using a UV transilluminator and camera to image the labeled RNA.

2.3.5 Steady-state fluorescence anisotropy measurement

Fluorescence anisotropy experiments were carried out on a Perkin Elmer LS55 fluorescence spectrometer (Waltham, MA) using an excitation and emission wavelengths of 495 nm and 521 nm and slit widths of 10 mm and 3 mm, respectively. FrU_{20} and FrU_6 were prepared in 500 μl of high (5 mM PBS, 120 mM KCl at pH 7.0), medium (5 mM PBS, 50 mM KCl at pH 7.0) or low (5 mM PBS, 15 mM KCl at a pH 7.0) salt buffer using a 0.5 cm fluorimeter quartz cuvette. Anisotropy of FrU_{20} was measured after each addition of a protein aliquot. The data were analyzed with GraphPad Prism.

2.3.6 Isothermal titration calorimetry (ITC) experiments

All the ITC experiments were carried out on a VP-ITC microcalorimeter (MicroCal, LLC, Northampton, MA) at 25 °C with a reference power of 18 μcal/sec and a stirring speed of 351 rpm. For the RNA into protein titrations, protein samples were first dialyzed overnight in D-Tube Dialyzers MWC 6-8 kD (EMD Millipore) against 20 mM sodium phosphate, pH 7.0 and 100 mM KCl and adjusted to the desired concentration. For the protein into RNA titrations, RNA samples were first dialyzed against the high (5 mM sodium phosphate buffer, 120 mM KCl at pH

7.0) or the low salt buffer (5 mM sodium phosphate, 15 mM KCl at pH 7.0). The lyophilized RNA was resuspended in appropriate protein dialysis buffer to ensure buffer matching. Following resuspension of the protein samples, ^{1}H NMR spectra were acquired to ensure correct protein folding (Figure 7). Samples were degassed prior to loading into the sample cell or the autopipette. The titration was performed by adding 8-10 μl aliquots to the sample cell every 300 s until the syringe volume was depleted.

2.3.7 NMR

Data were collected on a Bruker 500 MHz NMR equipped with a 5 mm TBI probe. Protein samples were diluted with NMR sample buffer containing 10 mM PBS, 50 mM NaCl, 0.1 mM EDTA and 10% D_2O to achieve a final concentration between 80-100 μM. Samples were adjusted to pH 6.8. ^{1}H spectra were acquired using a pulse program with watergate solvent suppression (zgprw5). The $rU_{15}T$ samples were resuspended in 10 mM PBS, 50 mM NaCl and 0.2 mM EDTA with 10% D_2O, pH 6.8. RRM123 was added at 1:1, 2:1 or 3:1 ratios and the pH was adjusted to match the free RNA sample (50 μM). TOCSY data was collected using a pulse program with watergate solvent suppression (mlevgpphw5) at 293 K with a relaxation time of 2 s and a mixing time of 50 ms.

2.4 Results and Discussion

2.4.1 Confirmation of the folding of hTIARb RRMs

Recombinant hTIARb RRMs were expressed, purified and visualized by SDS PAGE and Coomassie Blue staining. Spectroscopic methods were used to ensure the purity and folding of the recombinantly expressed RRM1, RRM2, RRM3 and RRM123. Circular dichroism (CD) spectra showed negative peaks consistent with β-sheet formation, 210 -220 nm, (22) (Figure 6).

Proton NMR spectra showed dispersed amide peaks (6.5-10.5 ppm), β-strand components (5-6 ppm), and shifted methyl peaks (-1-0 ppm) (Figure 7). All of these features are indicative of proper protein folding.

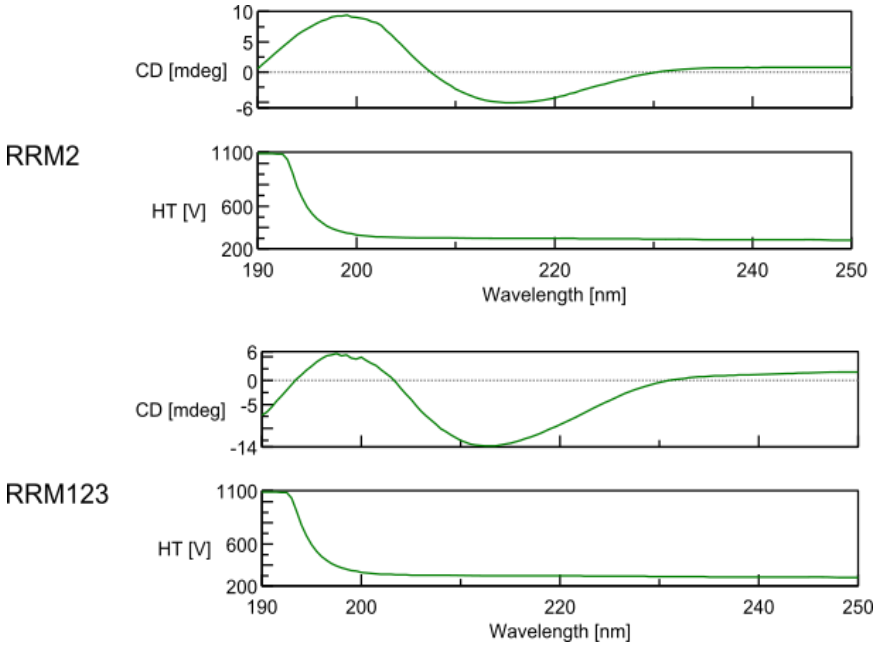


Figure 6. CD spectra of RRM2 and RRM123. Far-UV spectra indicates β-sheet characteristics. RRM1 and RRM3 had similar spectra (data not shown). Protein samples were in 10mM PBS with 15 mM NaCl.

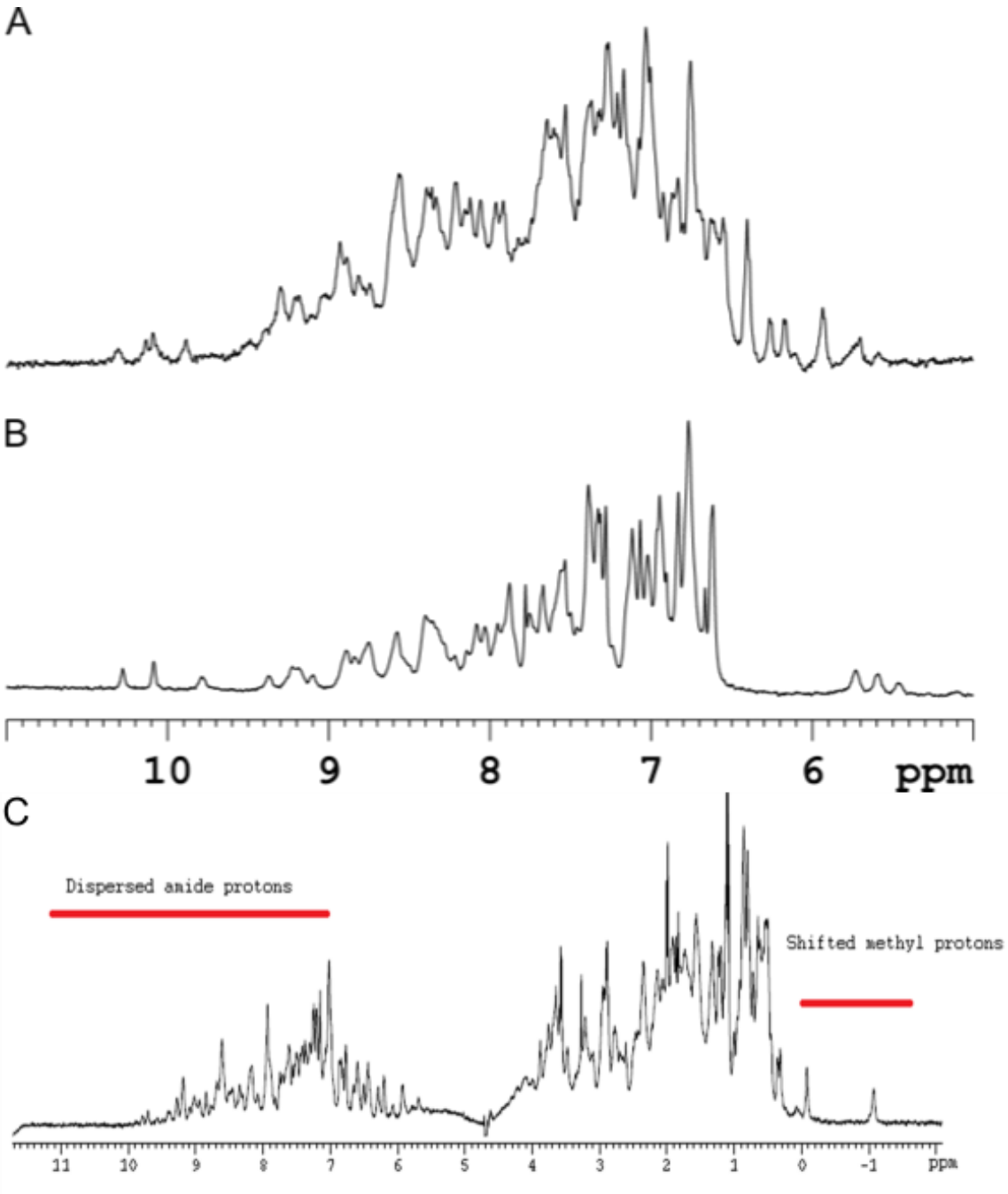


Figure 7. ^{1}H NMR spectral region of RRM2, RRM3 and RRM123. Downfield ^{1}H spectra of: A) RRM123 (80 µM), B) RRM2 (100 µM), and C) Full ^{1}H NMR spectrum of RRM3. Spectra of each protein were acquired in 10 mM PBS, 50 mM NaCl, 0.1 mM EDTA at pH 6.8 and 10% D_2O using a pulse program with watergate solvent suppression (zgprw5).

2.4.2 Multiple complexes observed in electrophoretic mobility shift assays (EMSA)

To probe the formation of multimeric complexes, a fluorescently labeled high-affinity substrate, FrU_{20}, was titrated with increasing concentrations of RRM2 or RRM123 (see Methods and Materials for protein constructs) and the complexes formed were analyzed by EMSA (Figure 8). RRM2 forms multiple complexes as evidenced by the increasing molecular weight of the top band as the protein to RNA ratio increased; however, the complexes are not stable in the gel as indicated by smearing. RRM123 not only has a higher affinity for the RNA substrate than RRM2, as evidenced by the rapid probe depletion with increasing amounts of protein but also by the more stable RNA-protein complexes formed as indicated by the reduced smearing. The data demonstrate that while RRM2 is sufficient to achieve RNA binding, as previously reported, the presence of RRM1 and RRM3 with their respective linker regions enhances the protein affinity for the RNA target and the stability of the overall complex. In addition, the detection of multiple bands signifies the formation of multiple protein/RNA complexes.

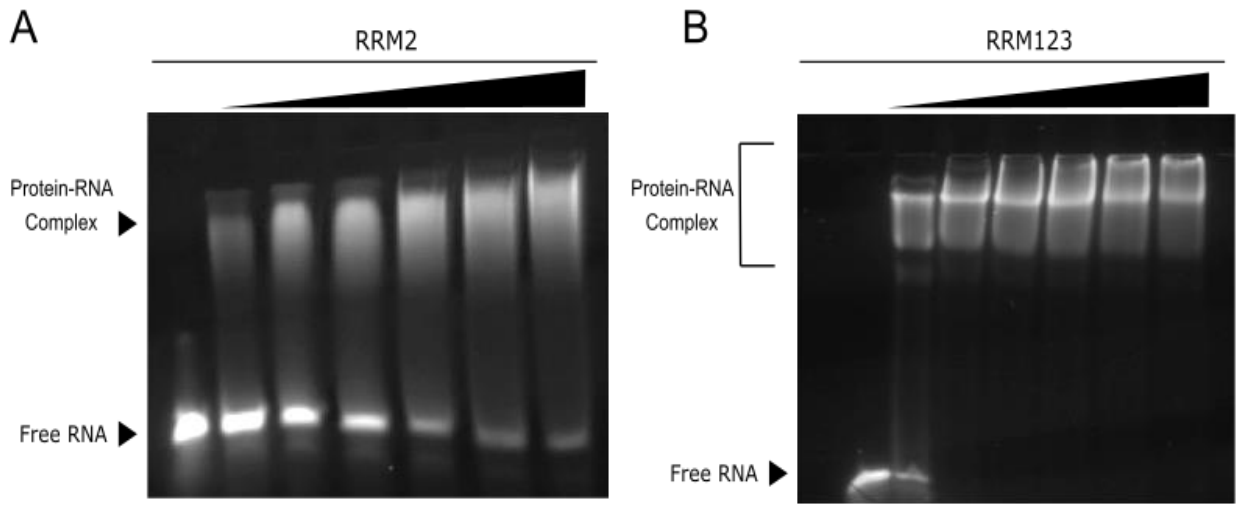


Figure 8. Electrophoretic mobility shift assays of RRM2 and RRM123 with FrU_{20}. A) Fluorescein labeled rU_{20} (10 µM) was titrated with varying concentrations of RRM2 to achieve ratios ranging from 1:1 to 6:1 protein to RNA (lanes 2-7). B) Fluorescein labeled rU_{20} (8 µM) was titrated with varying concentrations of RRM123 to achieve ratios ranging from 1:1 to 6:1 protein to RNA (lanes 2-7). Samples were incubated in 25 µL of 10 mM HEPES, 1 mM DTT, 0.1 mM EDTA, 100 mM KCl, 2 mM $MgCl_2$, pH 7.3 at 37°C for 1 hour prior to gel loading.

2.4.3 Determination of the binding contribution of individual RRM domains

To analyze the contributions of the RRM1 and RRM3 subunits to RNA binding affinity, isothermal titration calorimetry (ITC) for RRM1 and RRM3 was performed with rU_5, a well-characterized substrate for other RRMs. No significant change in enthalpy was recorded for either titration (Figure 9 A). These data suggest that the affinities of RRM1 and RRM3 for short rU rich sequences are very low. This conclusion agrees with data from previous studies and indicates that despite each of the RRMs having similar protein tertiary structures and conserved residues in their RNP motifs (Table 2), only RRM2 has significant affinity for rU_n sequences (1). The isotherms obtained for the titrations with RRM123 and the known low-affinity substrates rA_5 and rC_5, showed only very small changes in enthalpy (Figure 9 B). At higher concentrations, low-affinity binding was indicated by the isotherm for RRM2 and rC_5 (Figure 9 B). These data indicate that neither rA_5 nor rC_5 are high-affinity substrates for RRM1, RRM2, or RRM3.

Finally, a U-rich 20-mer (N_{20}) was used to assess the affinities of each individual RRM and of RRM123 in the context of a mixed RNA sequence. Again, no binding was observed for RRM1 and RRM3; however, the data for the RRM2 and RRM123 titrations showed similar binding affinities with K_d values in the 10^{-7} range (Figure 9 C). These data confirm that while RRM2 and RRM123 did not bind rA_5 and rC_5 with high affinity, both proteins are able to bind a mixed U-rich sequence at high affinity demonstrating that U residues are only needed at certain positions. These data also indicate that RRM1 and RRM3 do not directly bind to the RNA in the context of a U-rich mixed sequence (Figure 9 C). Additionally, the presence of RRM1 and RRM3 does not enhance binding by an amount consistent with independent multi-domain contacts as evidenced by the similar K_d values for RRM2 and RRM123.

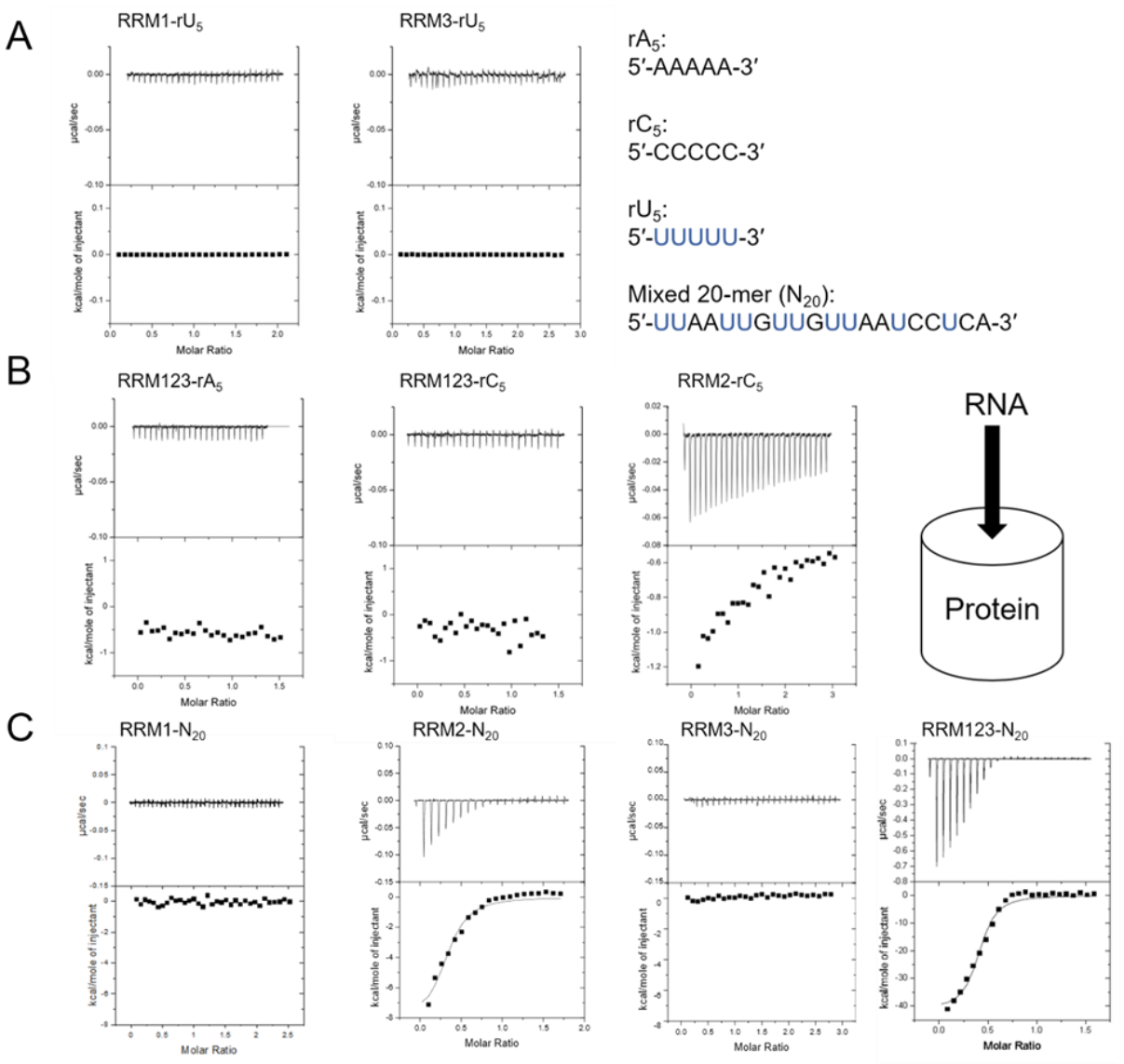


Figure 9. Isothermal titration calorimetry (ITC) of low affinity RNA substrates with individual RRMs and RRM123. A) RRM1 and RRM3 (5 µM) were titrated with rU_5 (50 µM). B) RRM123 (5 µM) was titrated with rA_5 and rC_5 (50 µM), respectively. RRM2 (12 µM) was titrated with rC_5 (172 µM). C) RRM1, RRM2, RRM3 and RRM123 (5 µM) were titrated with U-rich mixed 20-mer (50 µM). All titrations were performed in 20 mM sodium phosphate buffer with 100 mM NaCl.

2.4.4 Characterizing the differences between TIA-1/TIAR RRM domains

An analysis was done to identify possible differences contributing to the dissimilar binding preferences of each RRM using deposited structures in the Protein Data Bank (PDB). To visualize the structures in three-dimensional space, electrostatic and hydrophobic maps for RRM1, RRM2 and RRM3 were generated using UCSF Chimera (Figure 10) (23). A prominent difference observed between RRM2 and the other two RRMs was that the β-sheet surface of RRM2 that is involved in RNA binding (13, 14) is positively charged while the corresponding surfaces on RRM1 and RRM3 are neutral or even repulsive to nucleic acids. The surface of RRM2 anterior to the β-sheet is negatively charged, which under low salt conditions would be expected to negatively affect the binding of additional copies of RRM2 to the same RNA in close vicinity. However, each of the RRMs contain hydrophobic patches that could support protein-protein contacts between multiple copies of RRM123 bound to the same RNA.

To assess the importance of electrostatic contributions to RRM2-RNA binding, fluorescence anisotropy experiments were carried out under varying salt conditions (Figure 11). To avoid complications due to protein-protein contacts, a short fluorescently labeled RNA probe (FrU_6) was used to ensure that a 1:1 complex was formed. A slightly longer RNA, rU_6 rather than rU_5, was chosen to mitigate potential effects of the fluorophore on protein binding. Binding curves were recorded under low (15 mM KCl), medium (50 mM KCl) and high (120 mM KCl) salt conditions (Figure 11), and K_d values were derived from curve fitting as 90 ± 20 nM, 450 ± 80 nM, and 25,840 ± 14,540 nM respectively. A reliable fit was not obtained for the high salt condition due to low affinity and incomplete saturation that resulted in a high standard deviation for the calculated K_d. Despite this, a clear trend was still observed. Under high salt conditions, electrostatic interactions between RRM2 and the RNA were attenuated as expected, which

resulted in a weaker protein binding affinity compared to the low salt condition. However, with longer RNA targets, to which multiple RRMs can bind, high salt conditions could also attenuate repulsive electrostatic protein-protein interactions. This is relevant when multiple highly charged proteins bind to the same RNA and both hydrophobic and salt-dependent electrostatic interactions contribute to the binding interaction.

Table 2. Alignment of the RNP1 and RNP2 motifs of TIAR. All three RRMs have conserved RNP 1 and RNP 2 motifs containing aromatic residues at key positions. Shaded residues indicate amino acids conservation between the TIAR/TIA-1 RRMs.

	RNP1								RNP2					
RRM_1	D	P	Y	C	F	V	E	F	L	Y	V	G	N	L
RRM_2	K	G	Y	G	F	V	S	F	V	F	V	G	D	L
RRM_3	K	G	Y	S	F	V	R	F	V	Y	C	G	G	V

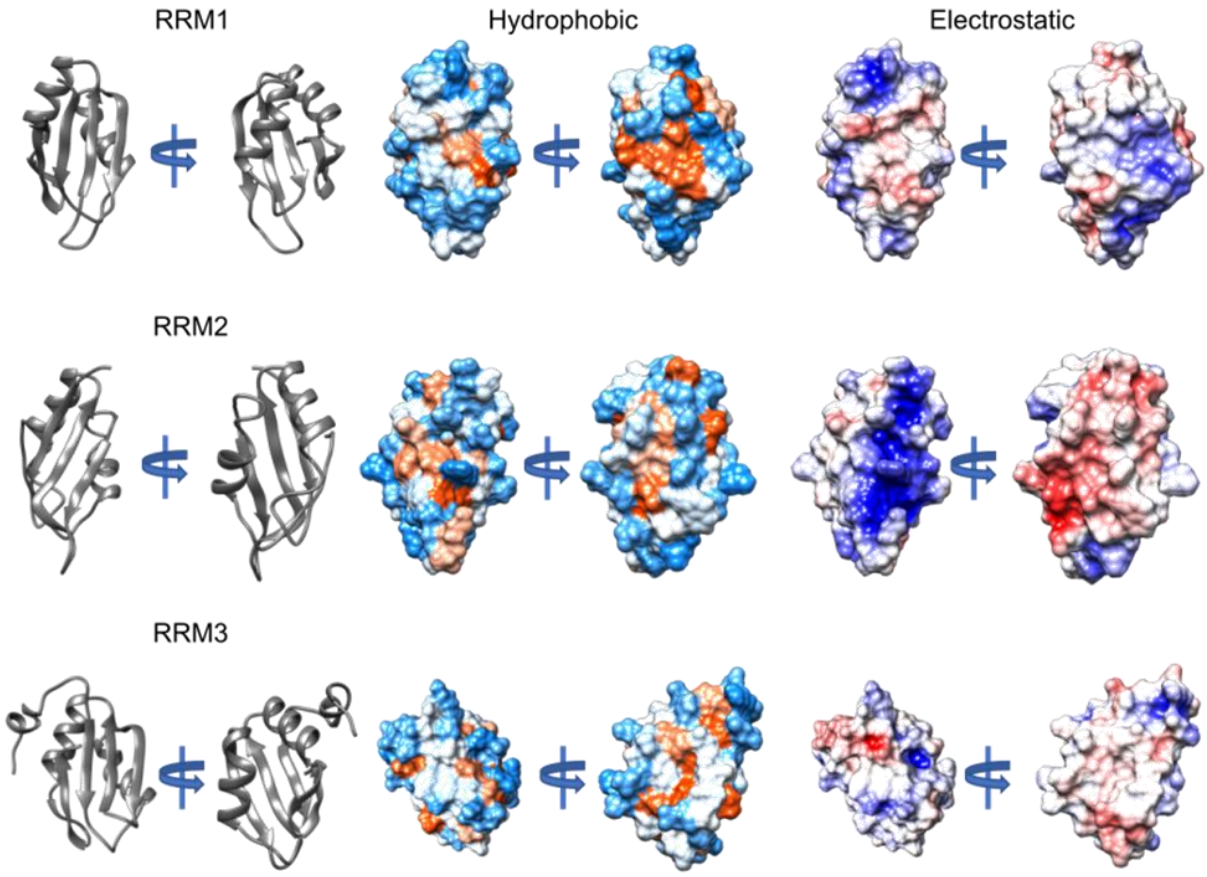


Figure 10. Electrostatic and hydrophobic maps of the TIAR RRM subunits. Ribbon representation, hydrophobic and electrostatic maps of each RRM subunit were generated using UCSF Chimera. For the hydrophobic maps, orange indicates hydrophobic and blue indicates polar patches. Electrostatic maps are from +5eV (blue) to -5eV (red). Left sides: bottom view, into the β-sheet; right sides: top view into the α helices. Only the RNA binding surface of RRM2 is highly charged. PDB 2CQI, 2DH7, 1X4G

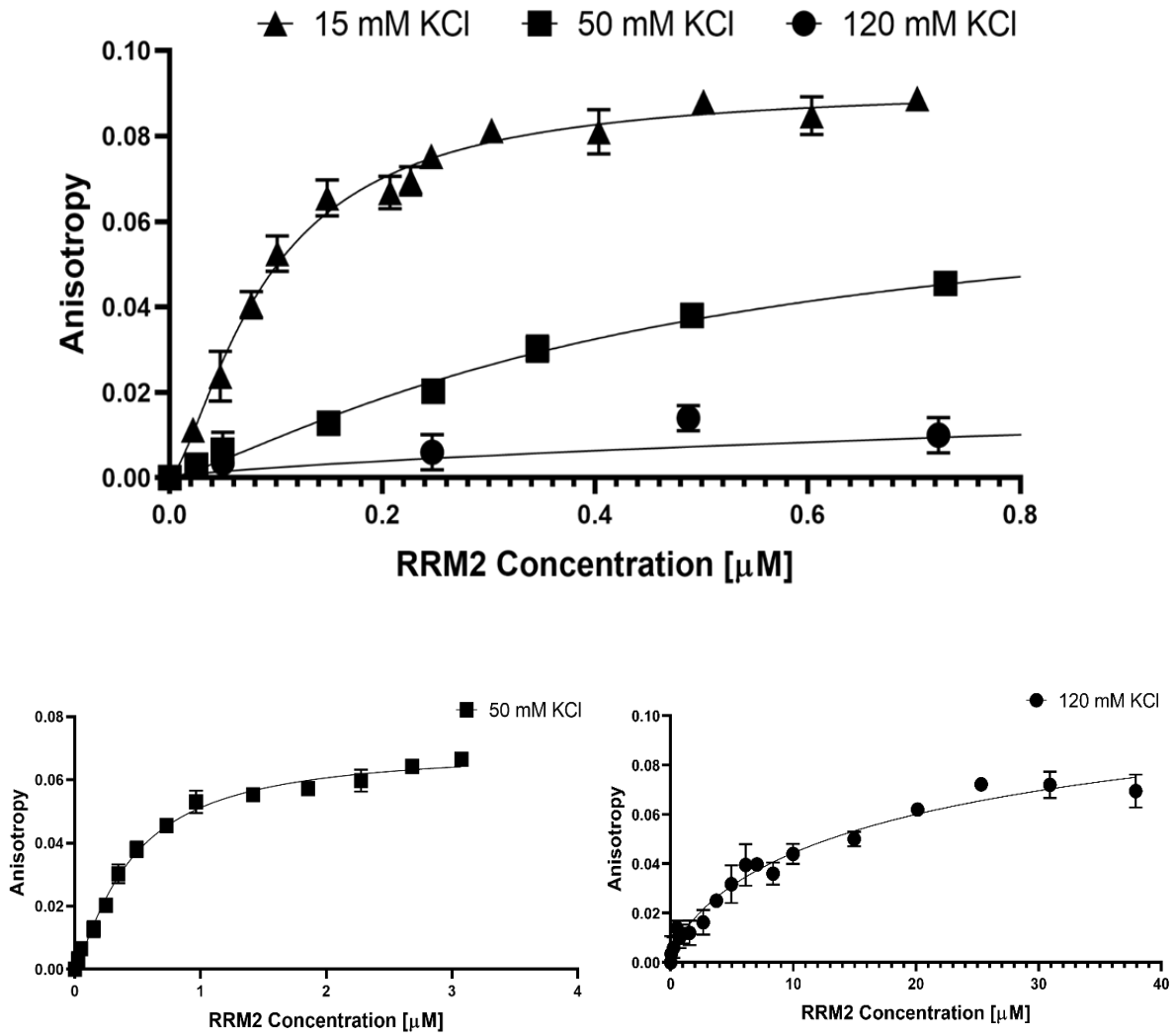


Figure 11. Salt-dependent binding of RRM2 to fluorescently labeled rU_6. FrU_6 (50 nM) was titrated with increasing concentrations of RRM2 in 5 mM sodium phosphate buffer with 15, 50 or 120 mM KCl. Binding was monitored using steady-state fluorescence anisotropy.(Top) Comparison of the binding curves at different salt conditions (Bottom) Full binding curves for the 50 mM KCl and 120 mM KCl conditions

2.4.5 ITC data confirms RNA length dependent stoichiometry

Additional ITC experiments were performed to quantitatively measure affinity (Table 3) as well as to determine the stoichiometry of the RNA-protein complexes (Figure 12). All binding isotherms were initially fit using an independent one-site model in order to determine stoichiometry and overall enthalpy. A one-site model in this context indicates multiple binding sites with equivalent dissociation constants and no interaction between sites. The n values increased as the length of the oligo U sequence increased (Figure 12 C). The number of RRM2 domains bound to a single RNA increases from 1:1 to 2:1 to 3:1 as the length of the RNA sequence increases (Figure 12 A). The ITC data for RRM123 showed that the RNA length also determined the number of these proteins that could bind (Figure 12 B). Both RRM2 and RRM123, have very similar enthalpies (ΔH) when binding to rU5 and both show a 1:1 stoichiometry, which supports the previous observation that RRM1 and RRM3 are not involved in direct RNA binding.

Based on the first iteration of fitting, it was determined that a more generalized binding polynomial would be appropriate to fit the cases where multiple proteins are bound to one RNA (See Appendix A). All isotherms were reprocessed to ensure consistent baselining and fitted using AFFINImeter software (24). Both datasets for rU_5 were refit using independent one-site models. The binding isotherms for the rU_{15} and rU_{20} were fitted using a two-step stoichiometric binding model (Table 3) (25, 26). In the case of rU_{15} a two-step model represents the formation of the singly-bound [PR] complex followed by the formation of the doubly-bound complex [P_2R]. In the case of rU_{20}, it is likely that a two-step model represents the formation of the doubly-bound complex [P_2R] followed by the formation of the triply-bond complex [P_3R]. Overall enthalpies were consistent between both sets of models (AFFINImeter vs MicroCal).

The affinity of RRM2 for rU_5 with 5 nucleotides (nts) available per RRM2, is relatively modest with a K_d of 7900 nM in 100 mM KCl. For the rU_{15} substrate, which supports the binding of two RRM2 proteins (7.5 nts per RRM2), a 60-fold stronger binding affinity was observed for the first copy of RRM2 than for rU_5. The second copy of RRM2 binds to rU_{15} at ~4.6-fold higher affinity than a single copy for rU_5. Increasing the RNA length to 20 nucleotides allows the binding of 3 RRM2 domains (6.5 nts per RRM2) but provides no marked additional increase in affinity for the first step compared to rU_{15}; however, there is an additional increase in affinity for the third copy of RRM2, which binds at ~5.3-fold higher affinity than to rU_5.

The same trend was observed for RRM123 with the different length rU substrates. Although a rU_{15} RNA accommodates two RRM2 or two RRM123 proteins, the second K_d value for RRM123 was 12 times lower than that for RRM2 (Table 3) indicating tighter binding to the RNA. The free energy of binding (ΔG) for each system is calculated using the equation:

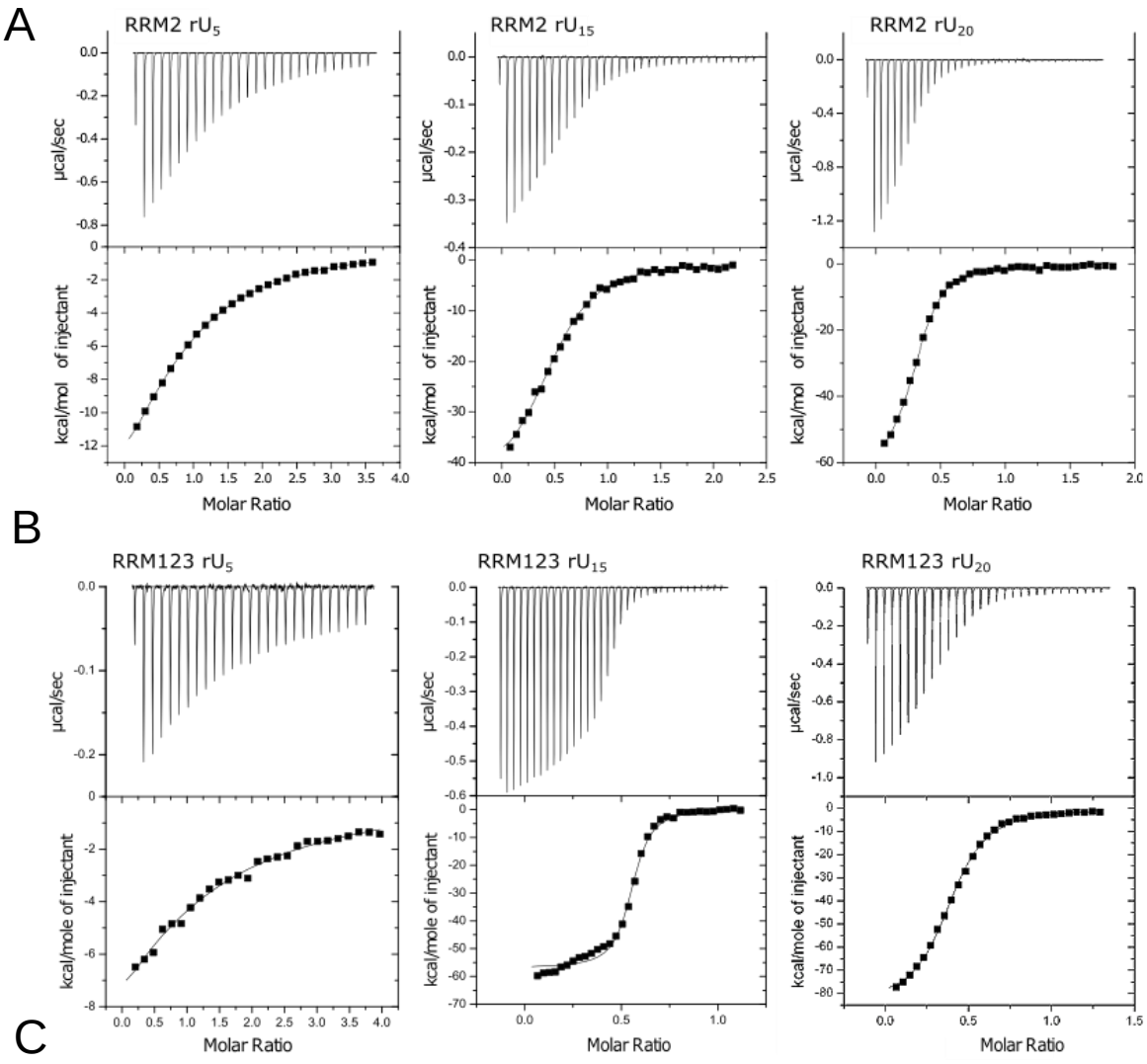
$$\Delta G = -RTlnK$$

where R is the ideal gas constant, T is the temperature, and K is the affinity constant.

Comparing the ΔG values for RRM2-rU_{15} and RRM123-rU_{15} the bound RRM123 dimer is stabilized by an additional ~1.4 kcal/mol compared to the bound RRM2 dimer. Although RRM1 and RRM3 do not directly bind to the RNA, they provide additional stabilization likely through protein interactions. There was a smaller increase in affinity between RRM2-rU_{20} and RRM123-rU_{20} (ΔK_d 3 x 10^{-7} M for K_2) than with rU_{15} systems (ΔK_d 16 x 10^{-7} M for K_2). This is likely an effect of the larger size of the RRM123 protein, which made it more difficult for three RRM123 molecules to bind optimally to rU_{15}.

Cooperativity was further assessed through steady-state fluorescence experiments for each RNA and protein combination. Data were collected using the low salt condition previously

described in order to best achieve binding saturation. Anisotropy data were fit using the specific binding with a Hill slope model in GraphPad Prism. The derived Hill coefficients showed cooperative binding for all of the rU_{15} and rU_{20} protein interactions measured (Table 3).



C

System	n	Ratio	ΔH (kcal/mol)
RRM2 rU_5	0.89 ± 0.01	1:1	-24.0 ± 0.3
RRM2 rU_{15}	0.52 ± 0.01	2:1	-45.0 ± 1.6
RRM2 rU_{20}	0.34 ± 0.00	3:1	-53.6 ± 0.7
RRM123 rU_5	0.85 ± 0.28	1:1	-25.8 ± 9.7
RRM123 rU_{15}	0.55 ± 0.00	2:1	-57.0 ± 0.5
RRM123 rU_{20}	0.40 ± 0.00	2.5:1	-84.4 ± 8.0

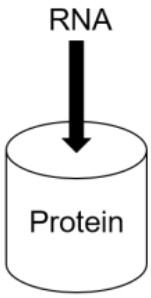


Figure 12. ITC of the RRM2 and RRM123 interactions with rU sequences of varying lengths. A) 3-5 µM RRM2 was titrated with 50-100 µM of a rU_n oligo (where n is the length) in 20 mM sodium phosphate buffer with 100 mM NaCl. B) RRM123 (3-7 µM) was titrated with 30-100 µM of rU_n oligo (where n is the length) under the same buffer conditions as in A. C) The isotherms obtained with the rU_5, rU_{15} and rU_{20} RNAs were fit with one-site independent binding models using the Origin MicroCal software.

Table 3. ITC and anisotropy data show high affinity cooperative binding. AFFINImeter was used to process the raw ITC data. Titrations with rU_5 were fit using a one-site independent, while the rU_{15} and rU_{20} titrations were fit using a two-step stoichiometric model. Hill coefficients were derived from fits of the steady-state fluorescence anisotropy data obtained at the low salt condition (5 mM PBS, 15 mM NaCl, pH 7.0).

	ITC		**Anisotropy**	
System	**K_d (M)**	**ΔH (kcal/mol)**	**System**	**Hill coefficient**
RRM2 rU_5	7.9×10^{-6}	-23.7 ± 0.2	**RRM2 rU_6**	1.1 ± 0.2
RRM2 rU_{15}	K_1: 1.3×10^{-7} K_2: 1.7×10^{-6}	H_1: -20.3 ± 0.1 H_2: -25.8 ± 0.5	**RRM2 rU_{15}**	1.5 ± 0.2
RRM2 rU_{20}	K_1: 6.3×10^{-7} K_2: 1.5×10^{-6}	H_1: -24.4 ± 0.1 H_2: -39.4 ± 0.2	**RRM2 rU_{20}**	1.4 ± 0.2
RRM123 rU_5	4.0×10^{-6}	-23.8 ± 0.2	**RRM123 rU_6**	1.1 ± 0.2
RRM123 rU_{15}	K_1: 5.0×10^{-8} K_2: 1.4×10^{-7}	H_1: -25.2 H_2: -33.7	**RRM123 rU_{15}**	2.4 ± 0.3
RRM123 rU_{20}	K_1: 1.2×10^{-7} K_2: 1.2×10^{-6}	H_1: -31.0 H_2: -46.7 ± 0.6	**RRM123 rU_{20}**	1.5 ± 0.1

2.4.6 The anisotropy data indicated that RRM2 only binds RNA

Data obtained from previous binding models of the TIAR RRM123-RNA interaction have suggested that while RRM2 is the main binding partner, RRM1 and RRM3 also directly bind to RNA. In these models, a single RRM123 would be oriented parallel to the RNA strand and all 3 RRMs would bind to the RNA forming a 1:1 complex (16). However, the ITC data described above indicate that although RRM1 and RRM3 stabilize the binding of multiple copies of RRM123 to RNA, this is not due to direct binding of RRM1 or RRM3 to the RNA. To further confirm multimeric binding of RRM123 to rU_{20}, the steady-state fluorescence anisotropy of FrU_{20} in complex with RRM2 or RRM123 was measured (Figure 13). If a single RRM123 binds to FrU_{20} forming a 1:1 complex with all 3 RRMs binding to the RNA, given the similar sizes of each RRM, it would be expected that the observed anisotropy would be similar to that when three RRM2 modules bind to the RNA forming a 3:1 complex. However, the measured fluorescence anisotropy of the two complexes was not similar. The anisotropy of the RRM123-U_{20} complex was markedly higher than the anisotropy of the RRM2-rU_{20} complex, supporting the binding of multiple RRM123 proteins to the rU_{20} target (Figure 13).

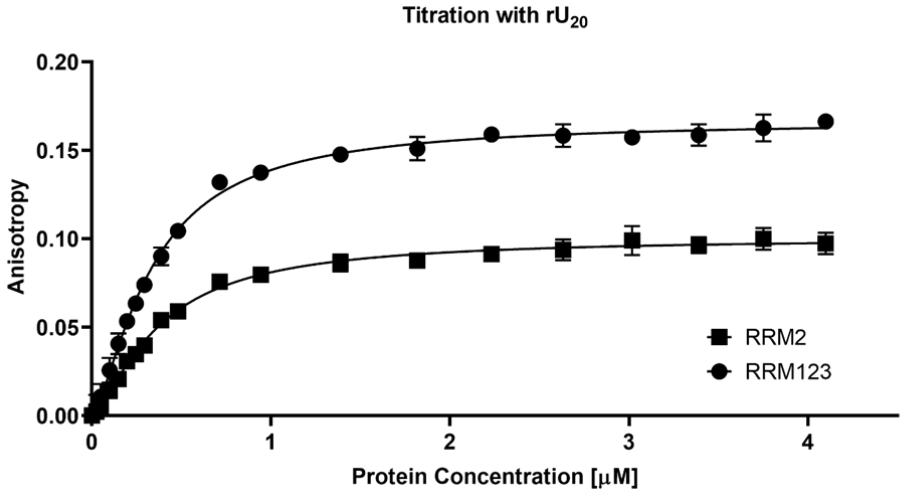


Figure 13. Binding of RRM2 or RRM123 to fluorescently labeled rU_{20}. FrU_{20} (250 nM) was titrated with increasing concentrations of RRM2 or RRM123 in 5 mM sodium phosphate buffer with 15 mM KCl. Binding was monitored using steady-state fluorescence anisotropy.

2.4.7 Reverse ITC titration and NMR data confirm cooperative RRM123 binding

Further studies were performed to examine the cooperativity of the rU_{15} and rU_{20} RRM interactions. Multiple published reviews have described the importance of using alternative methods such reverse ITC titrations and NMR to confirm cooperativity and stoichiometry in multivalent systems (25, 26). Reverse ITC measurements were performed for RRM123 binding to rU_{20} using the previously defined high and low salt conditions (Figure 14 A). For these experiments, the protein was titrated into the sample cell containing the RNA target. In this method, all protein is bound until complete saturation is achieved. As a result, multiple binding events can be observed when the binding sites are nonidentical.

For the titrations under both conditions, saturation was achieved at a molar ratio of ~3 confirming a stoichiometry greater than 1:1. Under low salt conditions, two distinct binding events with different affinities were detected with rU_{20}. As the protein was titrated into the RNA, a 2:1 complex formed before the binding of a third copy of RRM123 that bound at a decreased affinity due to the suboptimal binding site length that remains available on the RNA and the resulting repulsive electrostatic interactions. Under high salt conditions, because the affinity of the RRM123 binding to the rU_{20} RNA is reduced and the electrostatic repulsion between the bound proteins is also expected to be reduced, the different binding events were resolved into separate events.

A reverse titration was also done for the RRM123-rU_{15} under the low salt condition. Only one binding event was observed. In order to observe the interaction by NMR, an rU_{15} RNA was designed with a terminal thymidine residue ($rU_{15}T$). The inclusion of a T provided a unique signal from the methyl group that can be used to distinguish one end of the RNA sequence from the other. NMR spectra were collected using a TOCSY pulse program. A cross-peak was

observed between the T H6 and methyl protons (Figure 14 B). Upon titration of RRM123 into the RNA, an upfield chemical shift of the cross-peak was observed. At a 1:1 ratio two cross-peaks were observed, at a 2:1 ratio the majority of the volume shifted upfield, and at a 3:1 ratio only the upfield-shifted cross-peak was visible. In this method binding is monitored through the RNA peaks, and it is not possible to detect excess protein. Based on the ITC data discussed previously, it is expected that at the 3:1 titration ratio RRM123 is doubly bound to $rU_{15}T$ with excess free RRM123 present. The data demonstrate that only two major species exist at a 1:1 ratio. These species are most likely unbound RNA and the doubly bound RNA as the chemical shifts of the two cross-peaks correspond with the cross-peak that was observed for the unbound RNA and the cross-peak observed at the 2:1 and 3:1 ratio.

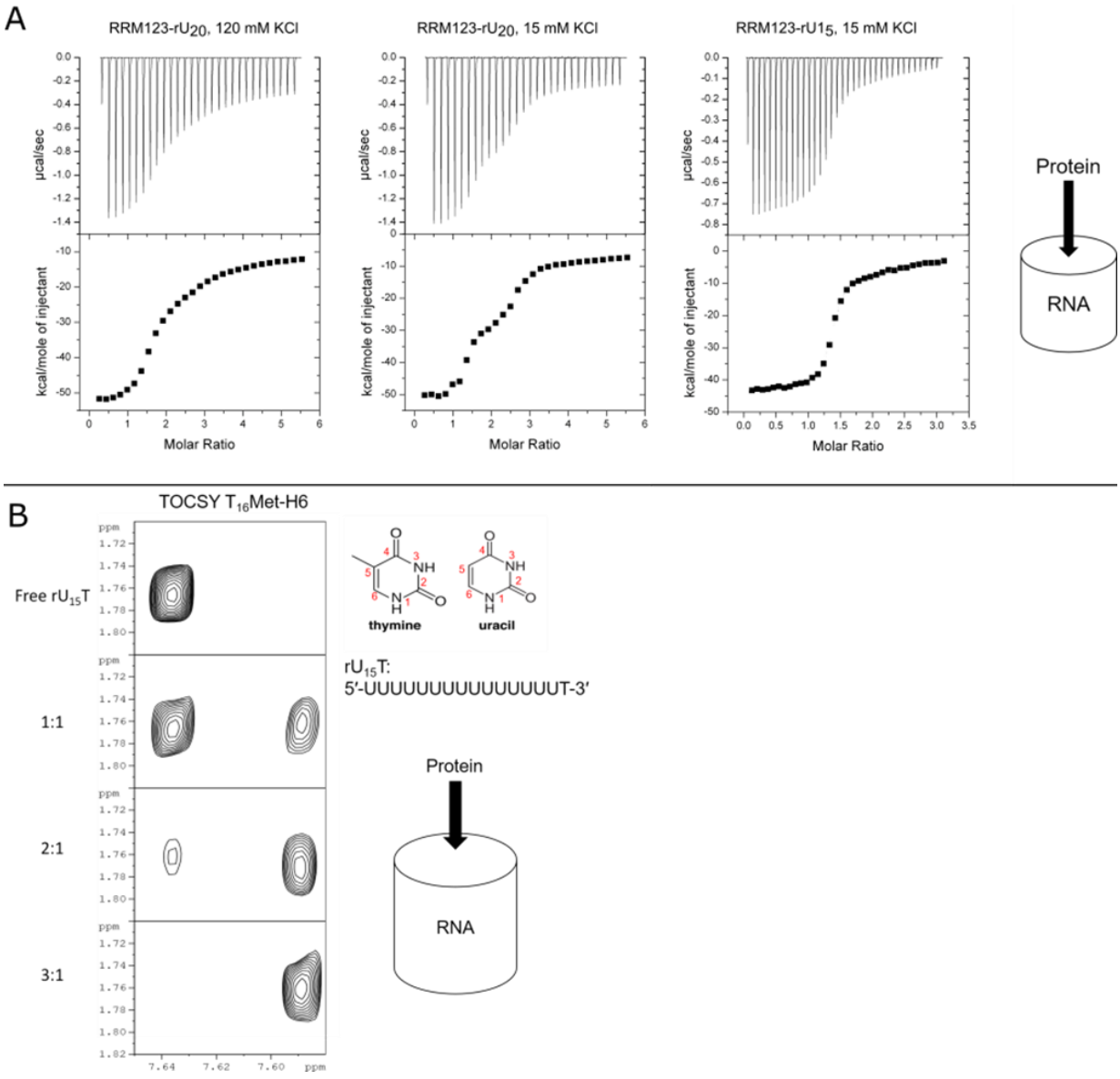


Figure 14. Reverse ITC titrations and NMR confirm multivalent stoichiometry and positive cooperativity. A) ITC experiments performed under high and low salt buffer conditions. High salt (left isotherm). Low salt (center and right isotherm). RRM123 (75 µM) was titrated into rU_{20} (5 µM) (left and center panel). RRM123 B) TOCSY of $rU_{15}T$ (50µM) titrated with increasing amounts of RRM123 to form 1:1, 2:1 and 3:1 complexes of protein to RNA. A chemical shift of the T_{16}Met-H6 cross-peak was observed upon protein binding.

2.5 Conclusions

The diversity and modular composition of RRM-containing proteins make it difficult to determine the function of individual RRMs. Some RRM-containing proteins form monomeric complexes with RNA as is the case for PTB (28). Others, like the Daz-1 homodimer (29) and the Nab3-Nrdl1 heterodimer (30) form multimeric complexes with RNA. Moreover, for proteins containing multiple RRMs, the RNA binding properties of individual RRM subunits may differ with only a single RRM or multiple RRMs binding to the RNA as reported for IMP3 (31) and Prp24, respectively (6).

Since multiple RRMs of the same protein or even multiple RRM proteins may bind to a single RNA, longer RNA sequences can exhibit a more complicated binding behavior that is more challenging to analyze. For simplicity, it is therefore instructive to evaluate the binding properties of each RRM in isolation. However, residues within the flexible linker region may contribute to RNA binding (2, 14) and consequently should be included with a single RRM. Moreover, the use of data generated after the removal of an RRM to infer the binding contribution of that component could result in interpretations that overlook the effects of interactions other than direct RNA contact. Furthermore, it is important to consider the effects of RNA sequence context as well as the experimental methods and pH and salt conditions used.

The data in this chapter provide further evidence that only the TIAR RRM2 binds directly to U-rich RNA sequences (Figure 15). For TIAR RRM1, we observed no binding to short RNA targets in agreement with data from previous surface plasmon resonance experiments with rU_{17} (14). Under our experimental conditions, we observed no binding of RRM3 to rU_5 RNA. However, other reports documented binding of RRM3 in the high micromolar range (140 μM) to rU_5 (18) and in the 30-50 nM range to rU_9 (20) in 50 mM NaCl solutions. It was also reported

that RRM3 (in the presence of RRM2) favors RNA sequences containing Cs (18) with enhanced binding at pH 5 due to protonation of histidines, which may be relevant to the cellular location of the protein (32). RRM3 was also reported to co-precipitate a population of cellular proteins but was not reported to interact with U-rich sequences precipitated by the full-length protein in a study by Dember et. al (1).

High-affinity binding of TIAR RRM2 to a rU_5 target in the low micromolar range (13) was confirmed and a nanomolar affinity for FrU_6 under low salt conditions (15 to 50 mM KCl) was observed. The RRM123 construct did not bind to C- or A-rich sequences with an observable affinity indicating that neither RRM1, RRM2, nor RRM3 bind to these sequences with high affinity. Given these results, it can be concluded that under saturating protein concentrations only RRM2 binds directly to the RNA and that Us are necessary for high-affinity protein-RNA interactions.

Although RRM modules are relatively small peptides, their size (~41x38x24Å) must be considered when studying their binding to RNA. A crystal structure of RRM2 with a T_3 ligand (10) shows that three nucleotides do not cover the entire binding surface of RRM2 (Figure 16 A). AutdockVina was previously used to model rU_5 binding to RRM2 (33, 34). Again, the five nucleotides covered only a portion of the β-sheet (Figure 16 B). The ITC data from the rU_{15} titrations with RRM2 and RRM123 both showed that an RNA of this size can support high-affinity binding of only two RRM2 or RRM123 proteins. These results suggest that each copy of RRM2, whether isolated or in the presence of RRM1 and RRM3, requires 7-8 nucleotides to optimally bind (Figure 15). The longer rU_{20} RNA (6.5 nts/RRM2) supports the binding of 3 RRM2 proteins albeit with reduced affinity compared to rU_{15} but the larger protein size of RRM123 makes it even more difficult for 3 proteins to bind optimally as shown by the ITC data

where the n indicates a ratio of 2.5 RRM123 for 1 rU_{20} (Figure 12). In summary, the results indicate that both RNA target length and protein size determine binding characteristics.

While RRMs are defined by their RNP motifs, which contain aromatic residues that enable RNA base-specific stacking interactions, electrostatic interactions also contribute significantly to protein binding affinity. It is apparent that single-stranded Us in an RNA are preferred by RRMs at key positions but are not necessary at all positions to achieve RNA binding (Figure 9 C) (12). Therefore, high binding affinity is a function of both "fit" and "span", where Us must be present at certain positions (likely involved in base stacking interactions) and the single-stranded region must be long enough to span the binding interface of RRM2 and engage in non-residue-specific and electrostatic interactions. The importance of electrostatic interactions is also further indicated by the strong salt dependence of the affinity of RRM2 for a FrU_6 substrate (Figure 11), and by the 27-fold increase in affinity observed by Kim et al. due to the inclusion of charged residues in the C-terminal flexible linker region of RRM2 (14).

The results of the present work show that longer rU RNA targets allow multiple TIAR RRM2 or RRM123 proteins to bind with higher affinity and also that multiple RRM123s bind with a higher affinity to the RNA target than multiple RRM2s. This was indicated by the EMSA data, which showed evidence for a higher affinity and binding complexity of RRM123 compared to RRM2. Also, a significant decrease was observed between the K_d values for the RRM123 ITC titrations compared to the equivalent RRM2 ITC titrations. This was especially evident when the global association constants (β_2) (See Appendix A) for the rU_{15} RNA substrate were compared. For RRM123 β_2 was ~10^{14}, and for RRM2 β_2 was ~10^{12}. This comparison showed a 100-fold higher affinity for the doubly bound RRM123 than for the doubly bound RRM2 in the context of rU_{15}. Since RRM1 and RRM3 have been shown not to bind to U-rich RNA under the

experimental conditions presented in this chapter, the enhancement of binding affinity and complex stability was attributed to protein-protein interactions (Figure 15). The effects of salt on these interactions and on the overall affinity of the protein for RNA were observed by the inverse ITC titrations (protein into RNA) and anisotropy experiments. A low salt environment promotes electrostatic attractive protein-RNA interactions and attractive electrostatic protein-protein interactions; while a high salt environment attenuates attractive electrostatic interactions and mitigates repulsive interactions.

The number and affinity of the proteins that can bind to a U-rich RNA depend on RNA length and are modulated by the salt concentration. Protein-protein contacts between multiple copies of protein support complex formation. However, binding of proteins also shortens the length of available space on the RNA which negatively affects complex formation. Therefore, it is important to consider RNA and protein size in addition to RNA sequence, as well as the stoichiometry and possible interactions of the components of the complex when considering binding models for RRM containing proteins and RNA partners.

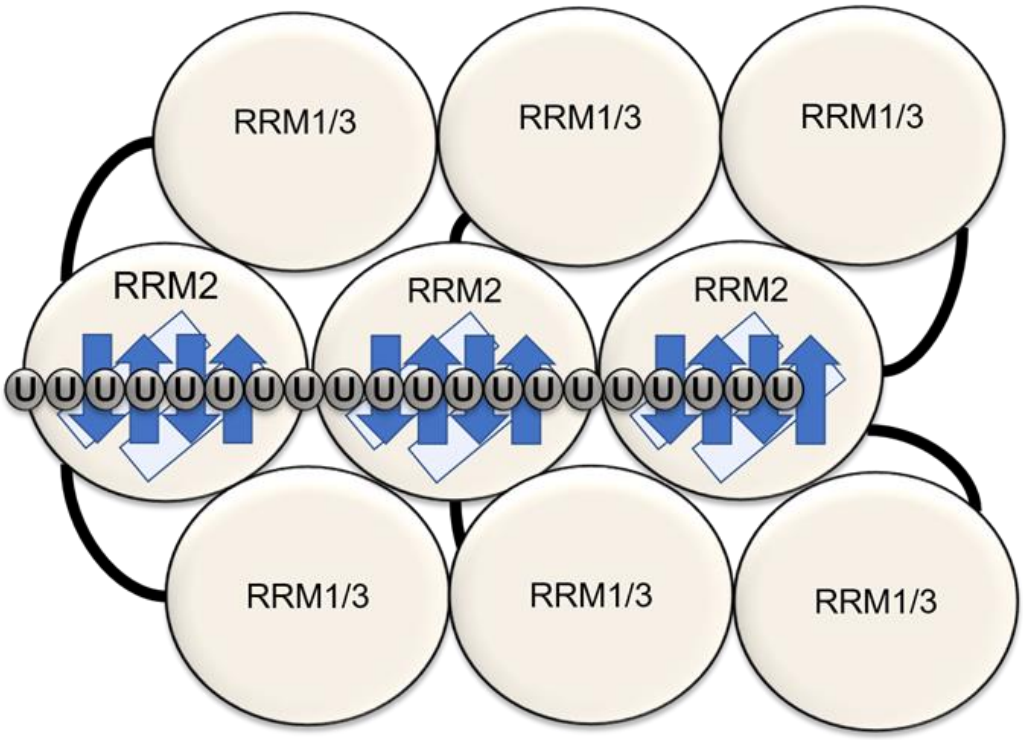


Figure 15. Proposed RNA binding model for TIA-1/TIAR. Steric constraints as well as the lack of evidence for RNA contacts between RRM1 or RRM3 and rU_{20} suggest that only RRM2 in the RRM123 protein directly contacts the RNA.

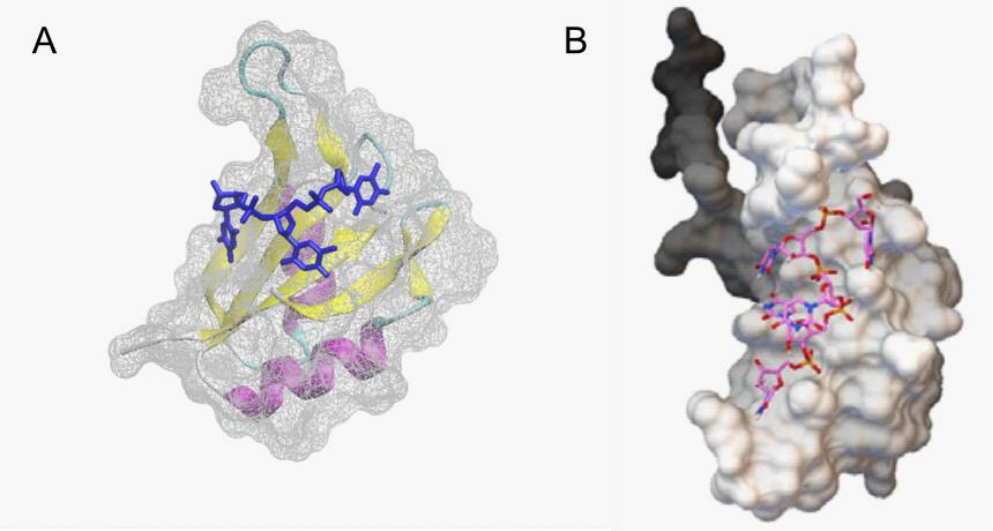


Figure 16. TIA-1/TIAR RRM2 in complex with short oligonucleotides. A) TIA-1 RRM2 (based on PDB 5ITH) is shown in surface representation. Secondary structure elements are colored: β-sheet, yellow; α-helix, purple; coil, cyan. The bound oligodeoxyribonucleotide, dT_3 (blue) is rendered in licorice mode. B) A molecular docking analysis was performed with TIAR RRM2 and rU_5 using AutodockVina (33, 34).

2.6 References

1. Dember, L. M., Kim, N. D., Liu, K. Q., & Anderson, P. (1996). Individual RNA recognition motifs of TIA-1 and TIAR have different RNA binding specificities. Journal of Biological Chemistry, 271(5), 2783-2788.
2. Afroz, T., Cienikova, Z., Cléry, A., & Allain, F. H. (2015). One, two, three, four! How multiple RRMs read the genome sequence. In Methods in enzymology (Vol. 558, pp. 235-278). Academic Press.
3. Muto, Y., & Yokoyama, S. (2012). Structural insight into RNA recognition motifs: versatile molecular Lego building blocks for biological systems. Wiley Interdisciplinary Reviews: RNA, 3(2), 229-246.
4. Kielkopf, C. L., Rodionova, N. A., Green, M. R., & Burley, S. K. (2001). A novel peptide recognition mode revealed by the X-ray structure of a core U2AF35/U2AF65 heterodimer. Cell, 106(5), 595-605.
5. Förch, P., Puig, O., Martínez, C., Séraphin, B., & Valcárcel, J. (2002). The splicing regulator TIA-1 interacts with U1-C to promote U1 snRNP recruitment to 5′ splice sites. The EMBO journal, 21(24), 6882-6892.
6. Montemayor, E. J., Curran, E. C., Liao, H. H., Andrews, K. L., Treba, C. N., Butcher, S. E., & Brow, D. A. (2014). Core structure of the U6 small nuclear ribonucleoprotein at 1.7-Å resolution. Nature structural & molecular biology, 21(6), 544-551.
7. Tian, Q., Streuli, M., Saito, H., Schlossman, S. F., & Anderson, P. (1991). A polyadenylate binding protein localized to the granules of cytolytic lymphocytes induces DNA fragmentation in target cells. Cell, 67(3), 629-639.

8. Kawakami, A., Tian, Q., Duan, X., Streuli, M., Schlossman, S. F., & Anderson, P. (1992). Identification and functional characterization of a TIA-1-related nucleolysin. Proceedings of the National Academy of Sciences, 89(18), 8681-8685.
9. Beck, A. R., Medley, Q. G., O'Brien, S., Anderson, P., & Streuli, M. (1996). Structure, tissue distribution and genomic organization of the murine RRM-type RNA binding proteins TIA-1 and TIAR. Nucleic acids research, 24(19), 3829-3835.
10. Waris, S., García-Mauriño, S. M., Sivakumaran, A., Beckham, S. A., Loughlin, F. E., Gorospe, M., ... & Wilce, J. A. (2017). TIA-1 RRM23 binding and recognition of target oligonucleotides. Nucleic acids research, 45(8), 4944-4957.
11. Zhang, T., Delestienne, N., Huez, G., Kruys, V., & Gueydan, C. (2005). Identification of the sequence determinants mediating the nucleo-cytoplasmic shuttling of TIAR and TIA-1 RNA-binding proteins. Journal of cell science, 118(23), 5453-5463.
12. Kim, H. S., Kuwano, Y., Zhan, M., Pullmann, R., Mazan-Mamczarz, K., Li, H., ... & Wilce, J. A. (2007). Elucidation of a C-rich signature motif in target mRNAs of RNA-binding protein TIAR. Molecular and cellular biology, 27(19), 6806-6817.
13. Kuwasako, K., Takahashi, M., Tochio, N., Abe, C., Tsuda, K., Inoue, M., ... & Taguchi, S. (2008). Solution structure of the second RNA recognition motif (RRM) domain of murine T cell intracellular antigen-1 (TIA-1) and its RNA recognition mode. Biochemistry, 47(24), 6437-6450.
14. Kim, H. S., Headey, S. J., Yoga, Y. M., Scanlon, M. J., Gorospe, M., Wilce, M. C., & Wilce, J. A. (2013). Distinct binding properties of TIAR RRMs and linker region. RNA biology, 10(4), 579-589.

15. Kim, H. S., Wilce, M. C., Yoga, Y. M., Pendini, N. R., Gunzburg, M. J., Cowieson, N. P., ... & Wilce, J. A. (2011). Different modes of interaction by TIAR and HuR with target RNA and DNA. *Nucleic acids research*, *39*(3), 1117-1130.
16. Bauer, W. J., Heath, J., Jenkins, J. L., & Kielkopf, C. L. (2012). Three RNA recognition motifs participate in RNA recognition and structural organization by the pro-apoptotic factor TIA-1. Journal of molecular biology, 415(4), 727-740.
17. Kumar, A. O., Swenson, M. C., Benning, M. M., & Kielkopf, C. L. (2008). Structure of the central RNA recognition motif of human TIA-1 at 1.95 Å resolution. Biochemical and biophysical research communications, 367(4), 813-819.
18. Cruz-Gallardo, I., Aroca, Á., Gunzburg, M. J., Sivakumaran, A., Yoon, J. H., Angulo, J., ... & Díaz-Moreno, I. (2014). The binding of TIA-1 to RNA C-rich sequences is driven by its C-terminal RRM domain. RNA biology, 11(6), 766-776.
19. Li, W., Li, Y., Kedersha, N., Anderson, P., Emara, M., Swiderek, K. M., ... & Brinton, M. A. (2002). Cell proteins TIA-1 and TIAR interact with the 3′ stem-loop of the West Nile virus complementary minus-strand RNA and facilitate virus replication. Journal of virology, 76(23), 11989-12000.
20. Wang, I., Hennig, J., Jagtap, P. K. A., Sonntag, M., Valcárcel, J., & Sattler, M. (2014). Structure, dynamics and RNA binding of the multi-domain splicing factor TIA-1. Nucleic acids research, 42(9), 5949-5966.
21. Zhang, J., Pearson, J. Z., Gorbet, G. E., Cölfen, H., Germann, M. W., Brinton, M. A., & Demeler, B. (2017). Spectral and Hydrodynamic Analysis of West Nile Virus RNA–Protein Interactions by Multiwavelength Sedimentation Velocity in the Analytical Ultracentrifuge. Analytical chemistry, 89(1), 862-870.

22. Corrêa, D. H., & Ramos, C. H. (2009). The use of circular dichroism spectroscopy to study protein folding, form and function. African Journal of Biochemistry Research, 3(5), 164-173.

23. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. Journal of computational chemistry, 25(13), 1605-1612.

24. Piñeiro, Á., Muñoz, E., Sabín, J., Costas, M., Bastos, M., Velázquez-Campoy, A., ... & Rial, J. (2019). AFFINImeter: A software to analyze molecular recognition processes from experimental data. Analytical biochemistry, 577, 117-134.

25. Freire, E., Schön, A., & Velazquez-Campoy, A. (2009). Isothermal titration calorimetry: general formalism using binding polynomials. Methods in enzymology, 455, 127-155.

26. Brown, A. (2009). Analysis of cooperativity by isothermal titration calorimetry. International journal of molecular sciences, 10(8), 3457-3477.

27. Martinez, J. C., Murciano-Calles, J., Cobos, E. S., Iglesias-Bexiga, M., Luque, I., & Ruiz-Sanz, J. (2013). Isothermal titration calorimetry: thermodynamic analysis of the binding thermograms of molecular recognition events by using equilibrium models (Vol. 4). chapter.

28. Monie, T. P., Hernandez, H., Robinson, C. V., Simpson, P., Matthews, S., & Curry, S. (2005). The polypyrimidine tract binding protein is a monomer. Rna, 11(12), 1803-1808.

29. Ruggiu, M., & Cooke, H. J. (2000). In vivo and in vitro analysis of homodimerisation activity of the mouse Dazl1 protein. Gene, 252(1-2), 119-126.

30. Carroll, K. L., Ghirlando, R., Ames, J. M., & Corden, J. L. (2007). Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements. Rna, 13(3), 361-373.

31. Jia, M., Gut, H., & Chao, J. A. (2018). Structural basis of IMP3 RRM12 recognition of RNA. RNA, 24(12), 1659-1666.

32. Cruz-Gallardo, I., Aroca, Á., Persson, C., Karlsson, B. G., & Díaz-Moreno, I. (2013). RNA binding of T-cell intracellular antigen-1 (TIA-1) C-terminal RNA recognition motif is modified by pH conditions. Journal of Biological Chemistry, 288(36), 25986-25994.

33. Zhang, Jin, "Macromolecular Interactions in West Nile Virus RNA-TIAR Protein Complexes and of Membrane Associated Kv Channel Peptides." Dissertation, Georgia State University, 2013. https://scholarworks.gsu.edu/chemistry_diss/81

34. Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computational chemistry, 31(2), 455-461.

3 STRUCTURAL AND FUNCTIONAL ANALYSES OF THE TIAR-3′(-) SL RNA INTERACTION

3.1 Preface

Nucleic acids are able to form a diverse range of 3D structures much like proteins. The 2002 discovery of riboswitches, RNA elements that undergo ligand-dependent conformational changes to regulate transcription, highlighted the specificity and dynamic nature of RNA structure (1-3). The work in this chapter explores the structure of the 3′(-) SL of WNV in the context of its interaction with TIAR. Determining the structure of these RNA-protein complexes provides a foundation for understanding their function in the regulation of viral genome replication and potentially provides a basis for the design of small molecule inhibitors.

3.2 Introduction

Crosslinking studies with cellular extracts revealed proteins with molecular weights of 42, 50, 60, and 108 kDa interacting with the 3′ SL of the negative (-) strand RNA of flaviviruses (4). The 42 kDa protein was identified as TIAR (5). TIAR was shown to bind with a K_d of about 10^{-8} to the 3′(-) SL RNA (5). The related protein, TIA-1 was also shown to bind the WNV 3′(-) SL at an affinity 10 times lower than TIAR (5). The binding sites within the stem loop had previously been located within two adjacent loops (L1 and L2) (6). The L1 sequence is 5′-UAAUU-3′ and the L2 sequence is 5′-UUAAUCCUC-3′ with these sequences separated by a 4 base pair stem (6). Deletion or mutation to Cs of the nts in L1 and L2 in a viral infectious clone resulted in reduced viral replication (6), and lower virus replication was also observed in a TIAR knockout cell line (5). It was hypothesized that TIAR/TIA-1 binding may help to disrupt the duplex between the viral genome and the full length negative strand RNA by facilitating the

formation of the 3′(-) SL on the negative template strand by binding to “breathing” A-U base-pairs and so enhancing recognition of this template by the polymerase (7).

3.3 Materials and Methods

3.3.1 Protein Purification

GST tagged RRM2 C DNAs and a RRM123 cDNA from hTIARb were subcloned into the PGEX-6p-3 (GE Healthcare) vector and separately expressed in *Escherichia coli*, strain Rosetta 2(DE3) pLysS, cells (Novagen) as previously reported (5). Constructs of each RRM subunit were created from hTIARb to include the contact residues in the flexible linker regions: RRM123 (residues1-294) and RRM2 (residues 94-196). Cells were incubated at 37 °C until an OD_{600} of 0.6-0.8 was reached. Protein expression was induced by addition of 1 mM IPTG. After an overnight, post-induction incubation at 25°C, cells were harvested by centrifugation. Cell pellets were resuspended in phosphate buffered saline (PBS) buffer containing CelLytic Express (Sigma-Aldrich Co. LLC), 1 mM dithiothreitol (DTT) and 1 tablet of complete EDTA-free protease inhibitor cocktail (Roche). The cells were sonicated on ice until lysed, and the lysate was clarified by centrifugation. The lysate was then incubated with 0.2% PEI at 4°C for half an hour. Ammonium sulfate precipitation was then performed at 75% saturation at 4°C for another half hour. The pellet was resuspended in PBS buffer and incubated with 1 mL of Glutathione-Sepharose 4B resin (GE Healthcare) overnight at 4°C. The GST tag was cleaved on the column using HRV-3C protease (Pierce). Free protein was collected from the column flow through and from several subsequent washes. RRM_{123} was purified using a centrifugal concentrator with a molecular weight cutoff of 50 kDa (Amicon).

3.3.2 RNA Preparation

The L3, WNV45 and WNV75 RNA oligos were purchased from Dharmacon as 2-hydroxyl protected RNAs. The samples were deprotected in-house following the manufacturer's protocol and desalted using a HiTrapTM desalting column (GE Healthcare). The RNA samples were then lyophilized. Proper folding was monitored by NMR.

3.3.3 Isothermal Titration Calorimetry

All the ITC experiments were carried out on a VP-ITC microcalorimeter (MicroCal, LLC, Northampton, MA) at 25 °C with reference power of 18 μcal/sec and a stirring speed of 351 rpm. Protein samples were dialyzed overnight in D-Tube Dialyzers MWC 6-8 kD (EMD Millipore) against 20 mM PBS, pH 7.0 and 100 mM KCl and adjusted to the desired concentration. Lyophilized RNA was resuspended in protein dialysis buffer to ensure buffer matching. Following resuspension of protein samples, ^{1}H NMR spectra were acquired to ensure correct protein folding. Samples were degassed prior to loading into the sample cell or the autopipette. The titration was performed by adding 8-10 μl aliquots to the sample cell every 300 s until the syringe volume was depleted.

3.3.4 NMR Spectroscopy

RNA samples were resuspended in NMR sample buffer containing 10 mM PBS, 50 mM NaCl 2 mM EDTA and 10% D_2O. Final RNA concentration was ~80 μM. Protein was added to achieve 1:1, 2:1 and 3:1 protein to RNA ratios for RRM2 and RRM123, respectively. Each sample was adjusted pH ~6.8. Imino proton spectra were collected on a Bruker Avance 600 MHz

spectrometer equipped with a 5mm QXI probe (Bruker) using Jump and Return solvent suppression (mwgp11) at 298 K.

^{31}P spectra were collected on a Bruker Avance 600 MHz spectrometer equipped with a 5mm QXI probe. RNA samples were prepared as stated above. Spectra were acquired at 298 K with proton decoupling (zgdc). The peaks were referenced using in-house chemical shift vales for 10mM PBS, 50 mM NaCl in D_2O at pH 6.8.

NOESY data was collected on a Bruker Avance 600 MHz spectrometer equipped with a 5mm QXI probe (Bruker using Jump and Return solvent suppression (mwgnoesyph11) with a mixing time of 100 ms at 298 K.

Diffusion measurements were collected on a Bruker Avance IIIHD 700 MHz spectrometer equipped with a 5mm indirect cryoprobe. RNA samples prepared in NMR sample buffer (as described above), lyophilized, resuspended in 98% D_2O and pH* ~6.8. RRM2 and RRM123 samples remained PBS in H_2O. Spectra were collected at 298 K using a pseudo-2D diffusion-ordered spectroscopy experiment (DOSY) pulse program using stimulated echo with bipolar gradient pulses (stebpgp1s). Data was acquired with 16k x 16 data points with Δ of 100 ms and δ of 2.5 ms. Data was collected for gradient strengths between 4% and 96%. Data was processed using Bruker Topspin T1/T2 software using manual peak picking. The derived diffusion coefficients from the fitted data were adjusted to account for the viscosity differences due to the changing percentage of D_2O in the sample. All diffusion coefficients were reported as the value in 100% H_2O. An example of this calculation is shown below for the 3:1 titration:

Volume of H_2O = 60 µL

Volume of D_2O = 270 µL

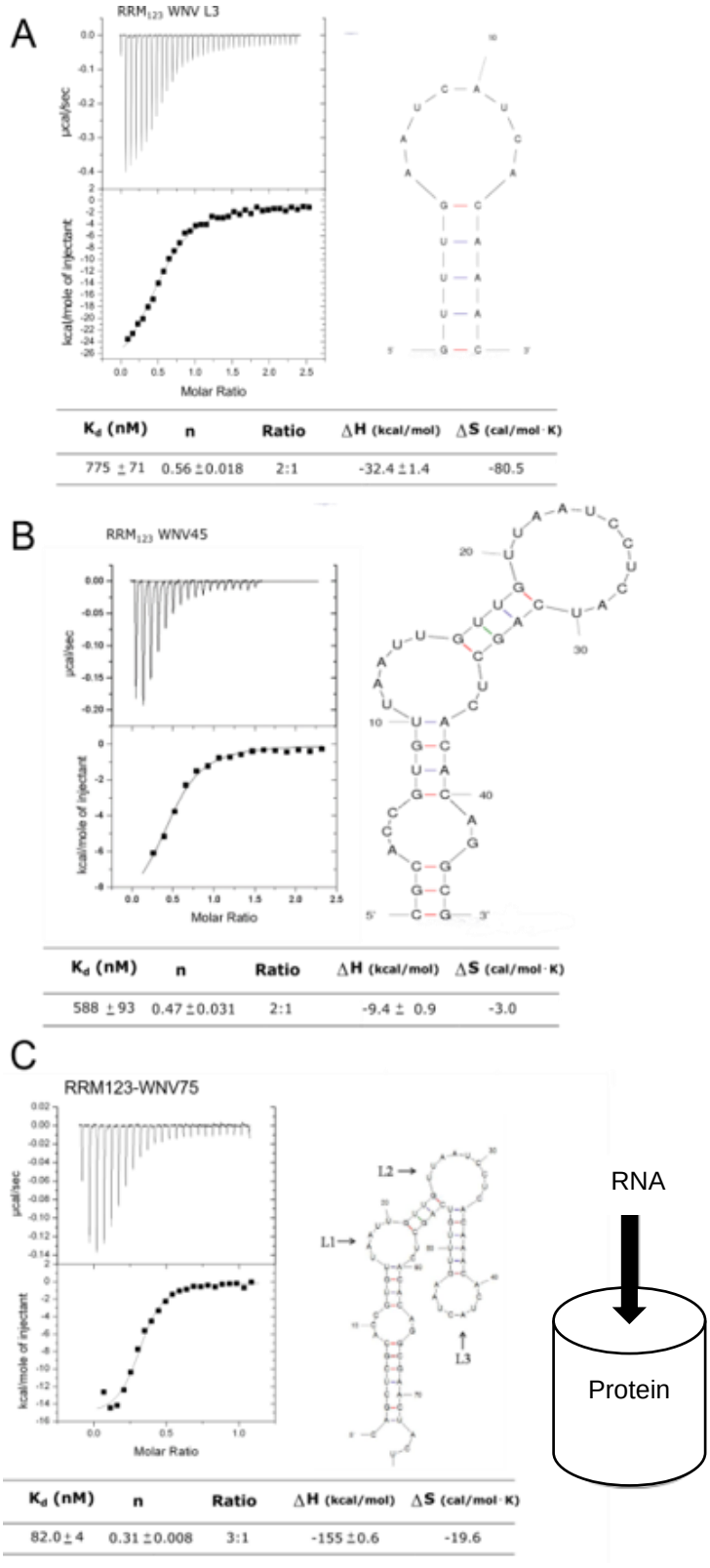
Total Volume = 330 µL

$$Corrected\ viscosity = \frac{60}{330}\eta_{H_2O} + \frac{270}{330}\eta_{D_2O}$$

3.4 Results and Discussion

3.4.1 ITC confirms multiple high affinity TIAR RRM123 binding sites on the WNV75 SL

Previous studies identified L1 and L2 of the WNV 3′ (-) SL as binding sites for TIAR. Further studies were conducted using RNAs consisting of loop 3 (L3) (Figure 17 A), a 45-mer consisting of L1 and L2 (Figure 17 B), and a 75-mer that encompasses the entire terminal SL (WNV75) (Figure 17 C). TIAR RRM123 was shown to bind to L3 in addition to the previously reported L1 and L2. The ITC data was fit using both independent site models and stoichiometric models. All three binding isotherms were best fit with one-site independent binding models. The stoichiometry values (n) from the ITC fits indicated that multiple copies of RRM123 were able to bind all three RNA sequences. RRM123 bound all three RNAs with K_d values in the 10^{-7} to 10^{-8} range, which is consistent with previously reported data (6).



K_d (nM)	n	Ratio	ΔH (kcal/mol)	ΔS (cal/mol·K)
775 ± 71	0.56 ± 0.018	2:1	-32.4 ± 1.4	-80.5

K_d (nM)	n	Ratio	ΔH (kcal/mol)	ΔS (cal/mol·K)
588 ± 93	0.47 ± 0.031	2:1	-9.4 ± 0.9	-3.0

K_d (nM)	n	Ratio	ΔH (kcal/mol)	ΔS (cal/mol·K)
82.0 ± 4	0.31 ± 0.008	3:1	-155 ± 0.6	-19.6

Figure 17. ITC titrations of RRM123 with WNV75 3′ (-) SL RNA. RNA was titrated into protein for all three experiments. A) WNV L3, B) WNV45, and C) WNV75 with RRM123. All titrations were performed in 20 mM PBS, 100 mM KCl, pH 7.0.

3.4.2 ^{1}H NMR Titrations

As discussed in the previous chapter, ITC data are not sufficient to characterize interactions in multivalent systems. Additionally, these interactions involve a highly structured RNA in contrast to the single-stranded sequences used in the previous chapter. RNA structure adds an additional level of complexity to the interaction as binding-induced conformational changes to the RNA and changes in RNA conformation have been shown to regulate processes such as transcription (1-3).

To further examine the interaction, complex formation was monitored by ^{1}H NMR. RRM2 and RRM123 were sequentially added to WNV75 to obtain 1:1, 2:1 and 3:1 ratios of protein to RNA. Upfield chemical shifts were observed for protein methyl peaks in both titrations (Figure 18). These data indicate that bound protein is the majority species at all ratios of the titration, which confirms that multiple copies of RRM2 and RRM123 bind to WNV75.

The imino protons of WNV75 were also monitored by NMR (Figure 19). Again, RRM2 and RRM123 were added in increasing concentrations to a constant amount of WNV75. Surprisingly, only minimal changes to the imino proton spectrum were observed, which indicated minor perturbation of base-pairing (Figure 16). The broad signal at ~11 ppm for the free WNV75 is consistent with solvent protected Us in the loop regions. The disappearance of the signal is consistent with protein binding to the loop region. These data suggest that both RRM2 and RRM123 bind to the loop regions (L1, L2 and L3) and the double-stranded regions of the SL remain conserved.

To further confirm this observation a 2D NOESY was performed using free WNV75 and a sample containing a 1:1 ratio of WNV75 and RRM123 (Figure 20). A strong nuclear Overhauser effect (NOE) cross-peak between the imino protons of U22 and G56 indicate the presence of a G-U wobble base-pair. A NOE cross-peak can still be observed after the addition of RRM123. These data show that the 4 base-pair stem located in between L1 and L2 was not perturbed by protein binding.

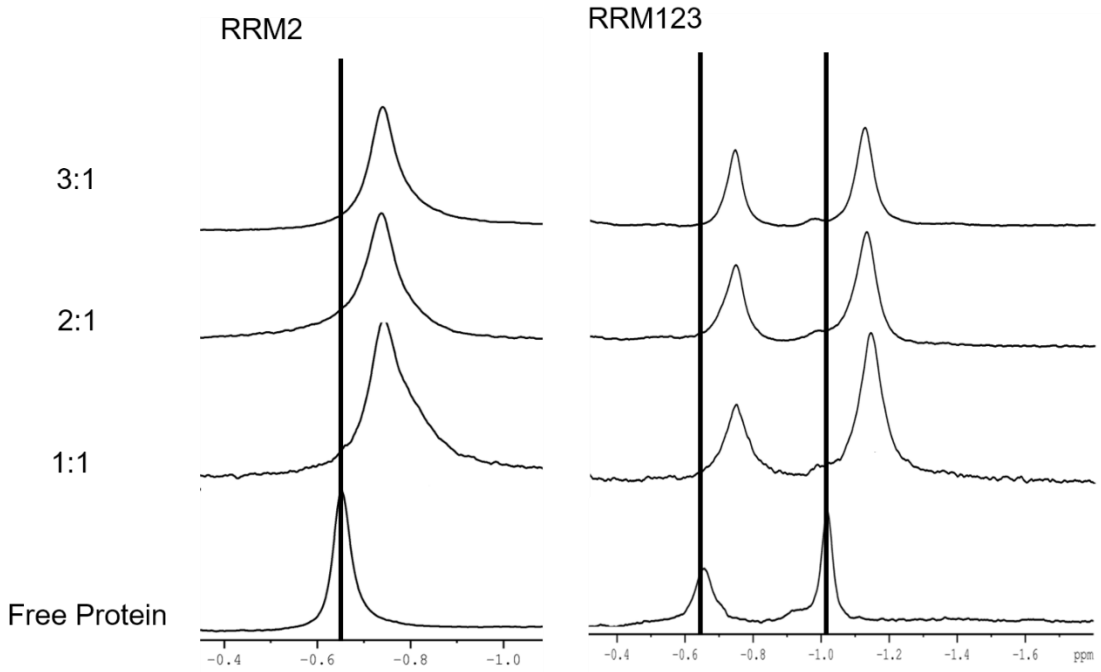


Figure 18. Complex formation monitored by NMR. Chemical shifts are observed for methyl peaks of both RRM2 and RRM123 when combined with WNV75 (80 μM). The observed chemical shifts indicated that RRM2 and RRM123 remain in complex when combined in a ratio of 3:1 with WNV75. These data support multiple proteins binding to one copy of RNA.

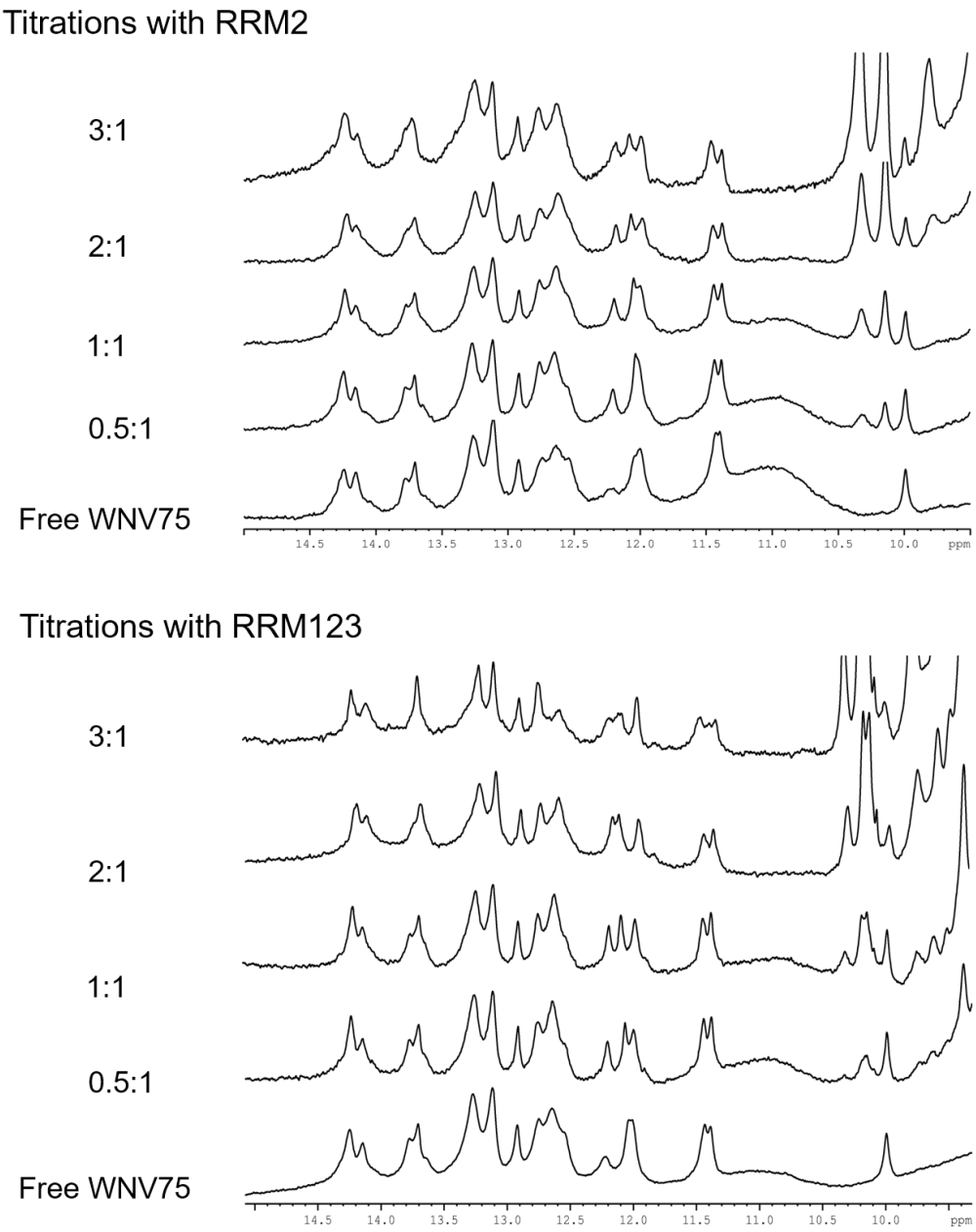


Figure 19. Imino proton spectra of WNV75 titrated with RRM2 and RRM123. RRM2 (top panel) and RRM123 (bottom panel) were titrated into WNV75 (80 μM) at ratios of 0.5:1, 1:1, 2:1 and 3:1 (protein to RNA).

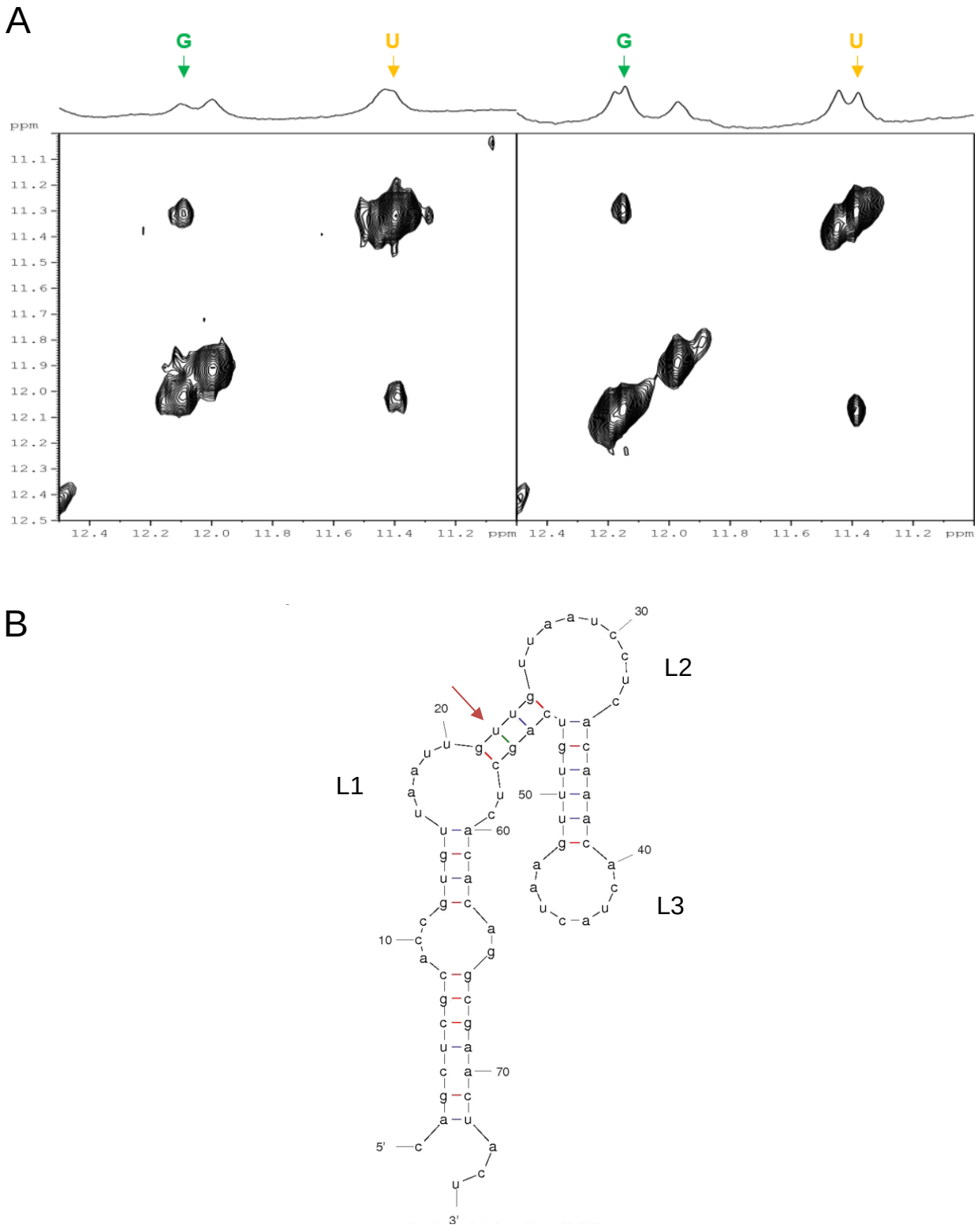


Figure 20. NOESY of free WNV75 and RRM123-WNV75 confirms conservation of a G-U basepair. A) NOESY spectra of free WNV75 (left) and a sample containing a 1:1 ratio of RRM123 and WNV75 (right) B) Predicted secondary structure of the 75nts comprising the 3′ (-) SL of WNV (WNV75). Red arrow indicates a G-U basepair in the structure.

3.4.3 ^{31}P NMR of RRM123 Titrations

Base-pairing is only one aspect of nucleic acid structure. The RNA backbone structure can be monitored by phosphate NMR. It is relatively unusual to collect ^{31}P spectra for long unlabeled nucleic acid sequences because the overlap of the signals typically makes it difficult to obtain useful structural information. However, some information can be determined by looking at the envelope of the backbone signals.

Proton decoupled 1D ^{31}P spectra were collected for the free WNV75 and samples containing 1:1, 2:1 and 3:1 ratios of RRM123 to the SL RNA as previously described in other experiments (Figure 21). As expected, the free WNV75 had a broad envelope indicative of many overlapping signals. However, as the concentration of added RRM123 increased, the width of signal envelope decreased. These data indicate that RRM123 binding to WNV75 decreases the structural variability of the phosphate backbone making the structure more uniform with less flexibility.

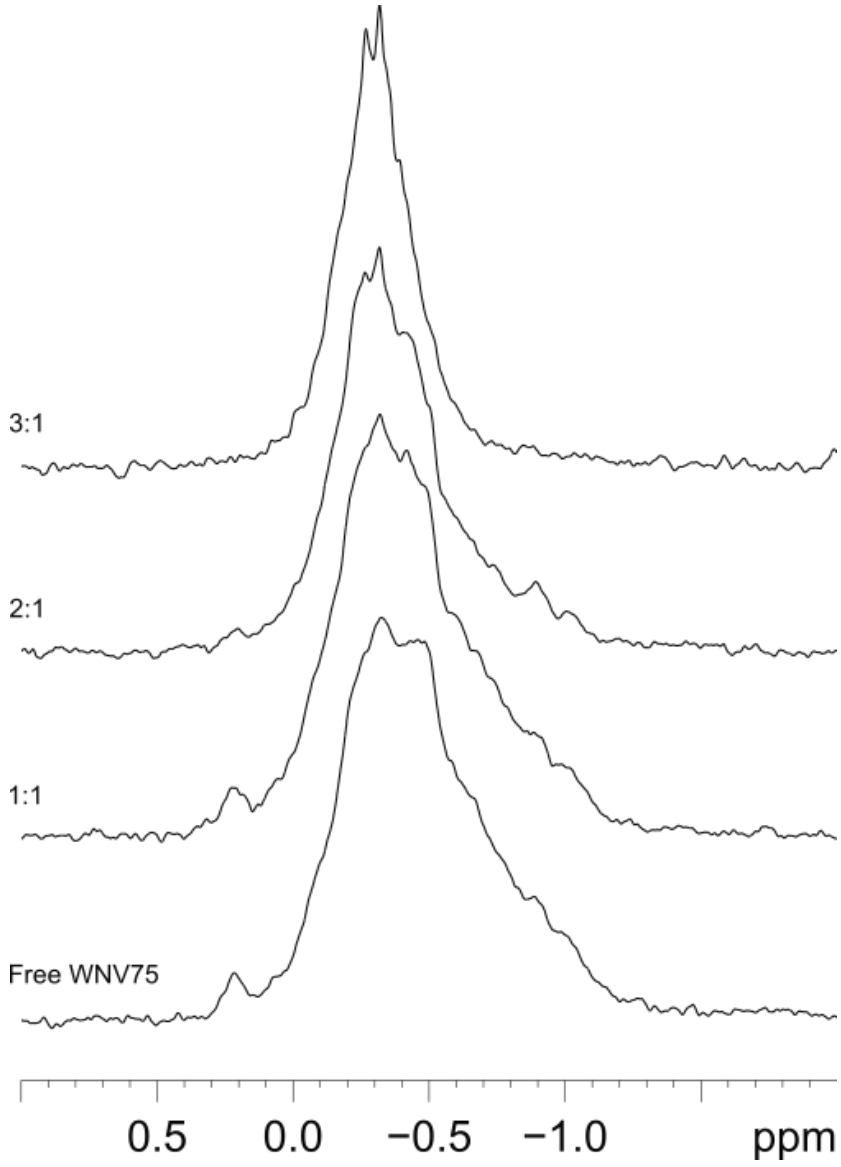


Figure 21. Monitoring RRM123-WNV75 complex formation by ^{31}P NMR. Spectra were collected for WNV75(60 μM) and samples containing increasing ratios of RRM123.

3.4.4 NMR Diffusion Experiments

The size of the RNA-protein complexes was also investigated by NMR spectroscopy. NMR diffusion experiments were initially proposed as a method for confirming multiprotein binding to WNV75. NMR diffusion experiments employ gradients to dephase and then refocus the signal. The greater the distance a molecule diffuses in the tube, the less signal will be measured upon refocusing. The diffusion coefficient of a spherical particle is inversely related to the radius of the particle, meaning that particles with smaller radii will diffuse faster than particles with larger radii. This relationship is described by the Stokes-Einstein equation (8):

$$D = \frac{k_b T}{6\pi\eta r} \quad (1)$$

where D is the diffusion coefficient, k_b is the Boltzmann constant 1.382 x 10^{-23} m^2 kgs^{-2} K^{-1}, T is the temperature, η is the viscosity, and r is the radius of the sphere. It was hypothesized that the hydrodynamic radius of the RRM123-WNV75 complex would increase as multiple copies of protein bound to the RNA and therefore the diffusion coefficient would decrease.

The experimental data yielded a different trend than what was predicted. A DOSY experiment was done with a sample of free WNV75 and samples containing a 1:1, 2:1, or 3:1 ratio of RRM2 or RRM123 and WNV75. Since it was determined that the majority of the protein is bound at all three ratios (as previously described), the intensity of a protein peak was used to calculate the diffusion coefficients. Observed intensity is related to the diffusion coefficient by the following equation:

$$I = I_0 * \exp\left(-D(2\pi G\gamma\delta)^2 * (\Delta - \frac{\delta}{3})*1000\right) \quad (2)$$

where I is the observed intensity, I_0 is the initial intensity, D is the diffusion coefficient, G is the gradient strength, γ is the gyromagnetic ratio for the observed nuclei, δ is the gradient length (ms) and Δ is the delay between gradients (ms).

For the RRM2 titrations, diffusion coefficients increased as the amount of protein increased (Table 4). Additionally, the RRM2 complexes diffused only marginally slower than the free RNA. For the RRM123 titrations, the diffusion coefficients for the complexes were larger than that of the free RNA (Table 4). In both cases, the diffusion coefficients did not significantly change between the 1:1, 2:1 and 3:1 ratios which indicates that 3 copies of RRM2 and RRM123 may bind to WNV75 at the 1:1 ratio. Since there is an inverse relationship between the diffusion coefficient and the hydrodynamic radius of the Stokes particle, the data suggest binding of RRM2/RRM123 to the WNV75 caused compaction of the complex. RRM123-WNV75 complexes diffused much faster than RRM2-WNV75 complexes indicating a higher degree of compaction which may be attributed to protein-protein interactions facilitated by RRM1 and RRM3.

As a validation for the values obtained by the DOSY experiments, the hydrodynamic radius for RRM2 was calculated using Equation 1 and compared to the distance measured from the deposited crystal structure of RRM2 (PDB 5ITH) (9). This calculation approximates the shape of the complex as a sphere. The calculated r value was 16.2 Å. Chimera was used to estimate the longest distance across RRM2. The measured distance was 38.8 Å corresponding to an r of 19.4 Å. While this is a rough estimate, it does confirm that the calculated r value is reasonable. The calculated radius for the free WNV75 was 24.8 Å; while the free RRM123 was 26.1 Å. The smallest RRM123-WNV75 complex, which was obtained at a 2:1 ratio of protein to RNA, had a radius of 17.7 Å demonstrating that there is significant compaction of the protein-RNA complex.

Table 4. Diffusion coefficients for RRM2 and RRM123 complexes. Coefficients for the complexes were calculated using the intensities from protein peaks.

RRM2 with WNV75	**m/s^2**
Free Protein	13.5×10^{-11}
Free RNA	8.8×10^{-11}
1:1	7.7×10^{-11}
2:1	8.4×10^{-11}
3:1	8.6×10^{-11}
RRM123 with WNV75	
Free Protein	8.4×10^{-11}
Free RNA	8.7×10^{-11}
1:1	11.1×10^{-11}
2:1	12.3×10^{-11}
3:1	11.8×10^{-11}

3.5 Conclusions and future directions

Much like the data obtained with the longer ssRNA sequences presented in Chapter 2, the ITC and NMR titration data for the 3′(-) SL and RRM2/RRM123 indicated the formation of multimeric complexes. These data are further supported by the previously reported analytical ultracentrifugation (AUC) data in a 2017 study by Zhang et. al. showing TIAR binding to WNV75 at a ratio of 4:1 when samples were mixed at a 6:1 ratio of protein to RNA (10).

Upon addition of RRM2 and RRM123 in increasing concentrations to a constant amount of WNV75, minimal changes to the imino proton spectrum indicated minor perturbation of RNA base-pairing. However, significant changes to the phosphorus backbone were observed by NMR and suggested that the RNA backbone became more uniform. Additionally, diffusion experiments showed faster diffusion of the complexes and these data were also consistent with the AUC data. Sedimentation is proportional to the molecular mass and inversely proportional to the friction coefficient. The friction coefficient is directly related to the hydrodynamic radius. Therefore, as molecular mass increases and hydrodynamic radius decreases and the sedimentation coefficient increases.

RRM2 was found to be the only TIAR RRM to bind RNA at a nM affinity. The smaller binding footprint (RRM2 vs RRM123) would account for minimal perturbation of base-pairing. The formation of protein-protein interactions could also be a possible explanation for the observed complex compaction.

Further studies are needed to fully understand, the structure of this cellular protein-viral RNA interaction and the functional role of this interaction on enhancing initiation of genome RNA synthesis. At the time of the writing of this dissertation, samples have already been sent to the Stanford-SLAC Cryo-EM Center (S^2C^2) for analysis (See Appendix B). Recent advances in

cryo-electron microscopy (cryo-EM) allow structure determination of nucleic acids as small as 40 kDa (10). Structure determination of this complex would represent only the second structure to be obtained for a viral SL RNA in complex with a full-length cellular protein. The only structure of this nature previously resolved is the truncated SLIV of the poliovirus IRES in complex with poly-C binding protein 2 (PCBP2) (EMD-21133) (12).

However, some conclusions can still be drawn about the structure and function of the WNV 3′(-) SL-TIA-1/TIAR complex in the absence of a cryoEM structural analysis. All of the presented data suggest that TIA-1/TIAR has a stabilizing effect on the terminal SL. The data also suggest that a conformational change resulting in compaction occurs upon protein binding. The conclusions from the data about the composition and structure of these complexes coupled with previous data demonstrating the effects of TIA-1/TIAR on viral genome amplification, support the hypothesis that SL stabilization and compaction facilitate rapid re-initiation by the viral RdRp at the 3′ end of the negative strand template during the later stages of the flavivirus replication cycle.

3.6 References

1. Mironov, A. S., Gusarov, I., Rafikov, R., Lopez, L. E., Shatalin, K., Kreneva, R. A., ... & Nudler, E. (2002). Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. Cell, 111(5), 747-756.
2. Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L., & Breaker, R. R. (2002). Genetic control by a metabolite binding mRNA. Chemistry & biology, 9(9), 1043-1049.
3. Winkler, W. C., Nahvi, A., Sudarsan, N., Barrick, J. E., & Breaker, R. R. (2003). An mRNA structure that controls gene expression by binding S-adenosylmethionine. Nature Structural & Molecular Biology, 10(9), 701-707.
4. Shi, P. Y., Li, W., & Brinton, M. A. (1996). Cell proteins bind specifically to West Nile virus minus-strand 3'stem-loop RNA. Journal of virology, 70(9), 6278-6287.
5. Li, W., Li, Y., Kedersha, N., Anderson, P., Emara, M., Swiderek, K. M., ... & Brinton, M. A. (2002). Cell proteins TIA-1 and TIAR interact with the 3′ stem-loop of the West Nile virus complementary minus-strand RNA and facilitate virus replication. Journal of virology, 76(23), 11989-12000.
6. Emara, M. M., Liu, H., Davis, W. G., & Brinton, M. A. (2008). Mutation of mapped TIA-1/TIAR binding sites in the 3′ terminal stem-loop of West Nile virus minus-strand RNA in an infectious clone negatively affects genomic RNA amplification. Journal of virology, 82(21), 10657-10670.
7. Panavas, T., Stork, J., & Nagy, P. D. (2006). Use of double-stranded RNA templates by the tombusvirus replicase in vitro: Implications for the mechanism of plus-strand initiation. Virology, 352(1), 110-120.

8. Einstein, A. (1905). On the electrodynamics of moving bodies. Annalen der physik, 17(10), 891-921.
9. Waris, S., García-Mauriño, S. M., Sivakumaran, A., Beckham, S. A., Loughlin, F. E., Gorospe, M., ... & Wilce, J. A. (2017). TIA-1 RRM23 binding and recognition of target oligonucleotides. Nucleic acids research, 45(8), 4944-4957.
10. Zhang, J., Pearson, J. Z., Gorbet, G. E., Cölfen, H., Germann, M. W., Brinton, M. A., & Demeler, B. (2017). Spectral and Hydrodynamic Analysis of West Nile Virus RNA–Protein Interactions by Multiwavelength Sedimentation Velocity in the Analytical Ultracentrifuge. Analytical chemistry, 89(1), 862-870.
11. Zhang, K., Li, S., Kappel, K., Pintilie, G., Su, Z., Mou, T. C., ... & Chiu, W. (2019). Cryo-EM structure of a 40 kDa SAM-IV riboswitch RNA at 3.7 Å resolution. Nature communications, 10(1), 1-6.
12. Beckham, S. A., Matak, M. Y., Belousoff, M. J., Venugopal, H., Shah, N., Vankadari, N., ... & Wilce, J. A. (2020). Structure of the PCBP2/stem–loop IV complex underlying translation initiation mediated by the poliovirus type I IRES. Nucleic acids research, 48(14), 8006-8021.

4 STRUCTURAL ANALYSES OF EF1A1 AND THE FLAVIVIRAL 3′(+) STEM LOOP (SL) RNA

4.1 Preface

The structural analysis of EF1A1 presented in this chapter was done in collaboration with Dr. Graeme Conn of Emory University. The ribosomal cryo-EM structure containing EF1A1 was identified as having possibly useful structural information for EF1A1. The PDB file was sent to Dr. Conn and crystallography data analysis software was used to visualize the interactions of EF1A1 with ribosomal RNA and also protein post-translational modifications.

4.2 Introduction

As previously described three cellular proteins, 52, 84, and 105 kDa were found to bind to a positive strand SL probe through UV-crosslinking assays with cellular extracts (1). The 52 kDa protein was identified as EF1A1 (2). Similar eEF1A binding activities were detected with WNV, YFV, DENV2 and TBEV 3′(+) SL RNA probes demonstrating the conservation of this interaction (1). The dissociation constant (K_d) for EF1A in complex with the viral 3′SL RNAs was reported to be in the low nanomolar range (10^{-9}) (2). A major binding site responsible for 60% of the EF1A1 binding activity and two minor sites, each responsible for 20% of the binding activity, were identified (2, 3). The major site is located in the upper stem of the terminal 3′(+)SL, one of the minor sites is located in the top loop of the 3′SL (5′-AGUGC-3′) and the other in the loop of the penultimate sHP (3). Both EF1A binding to the 3′ RNA *in vitro* and viral negative strand synthesis in infected cells are greatly decreased by mutations in the critical nts of each of the binding sites supporting the functional relevance of this cellular protein-viral RNA interaction. However, EF1A binding to the 3′ RNA *in vitro* and viral negative strand synthesis were enhanced by disruption of the base pairing of three of the four residues in the major binding

site, a CACA on the 5′ side of the upper stem of the 3′(+) SL (3,4). This mutation increased the levels of negative strand RNA but did not increase levels of viral genome indicating that a specific ratio of negative to positive RNAs is required for optimal viral replication in infected cells (3).

Flavivirus genomes are translated after release from infecting virions in infected cells. The viral polyprotein produced is processed into the mature viral proteins including the viral replication complex proteins. A genome RNA must switch from translation to replication to become a template for negative strand RNA synthesis (5). For this to occur, the genome must become cyclized through formation of 3′-5′ terminal base pairs. How the disruption of the 3′ SL base pairs and the formation of the new 3′-5′ base pairs is triggered and regulated is not known. EF1A binding to the 3′(+) SL of WNV may promote the disruption of base pairing at the lower part of the 3′ SL stem and in the sHP to facilitate genome cyclization (3, 4). Based on these data, it is proposed that EF1A may regulate the switch of the genome RNA from translation to replication.

EF1A was also previously shown to bind RNA structures in the UTRs of a number of plant viruses including turnip yellow mosaic virus (TYMV) and tomato bushy stunt virus (TBSV) (7). EF1A1 binds to TBSV 3′ SL structures in the positive-sense RNA genome and regulates the synthesis of the negative strand (8). Previous studies also showed that phosphorylation of EF1A1 was necessary for binding to the 3′(+) SL; however, the specific phosphorylation site(s) that are essential for this interaction remain unknown (2).

4.3 Materials and Methods

4.3.1 Cloning of EF1A1

A cDNA sequence derived from EF1A extracted from BHK cells was cloned into a pCRct TOPO (Novagen) plasmid encoding a C-terminal 6x His tag. A codon optimized cDNA of EF1A1 from *Sus Scrofa* was cloned into the PGEX-6p-3 (GE Healthcare) encoding a C-terminal GST tag.

4.3.2 Phosphomimetic Mutants

Site-directed mutagenesis primers were designed to introduce phosphomimetic mutations into the EF1A1 cDNA (Table 5). A QuikChange II XL Kit (Agilent) was used to introduce these mutations into the EF1A1 expression plasmids.

Table 5. Mutagenesis primers for phosphomimetic mutations.

Mutants	**Primers**
S163E	Forward: 5′-**GAT TTC CTC GTA TCT CTT CTG CTC GTA TGG TGG CTC GGT GGA ATC**-3′ Reverse: 5′-**GAT TCC ACC GAG CCA TAC GAG CAG AAG AGA TAC GAG GAA ATC**-3′
S300E	Forward: 5′-**ATT GTC CCC AGG AAG AGC TTC CTC CAA AGC TTC ATG GTG CAT TTC**-3′ Reverse: 5′-**GAA ATG CAC CAT GAA GCT TTG GAG GAA GCT CTT CCT GGG GAC AAT**-3′
T432E	Forward: 5′- **TTG ATG ACA CCC ACA GCA ACC TCC TGC CTC ATG TCA CGA ACA G**-3′ Reverse: 5′- **CTG TTC GTG ACA TGA GGC AGG AGG TTG CTG TGG GTG TCA TCA A**-3′

4.3.3 RNA samples

The 2-hydroxyl WNV32 and ZIKV35 RNA oligos were purchased from Dharmacon. The 2-hydroxyl protected RNA samples were deprotected in-house following the manufacturer's protocol and desalted using a HiTrapTM desalting column (GE Healthcare). The RNA samples were then lyophilized and resuspended in NMR buffer consisting of a sample of purified RNA in 10 mM PBS, 50 mM NaCl, 1 mM EDTA, pH 6.8 and 10% D_2O.

4.3.4 ^{1}H NMR

Spectra were collected on a Bruker Avance 500 MHz spectrometer equipped with a 5 mm TBI probe. Imino proton spectra were collected for WNV32 (38 μM) and ZIKV35 (158 μM) at 298 K and 278 K, respectively. Jump and return solvent suppression was used (mwgp11).

4.4 Results and Discussion

4.4.1 ^{1}H NMR

A pseudoknot was initially proposed to exist between the 3′(+) SL and sHP; however, subsequent NMR mapping did not find evidence of this interaction (4). NMR was previous used to map a 67 nt WNV 3′(+) SL that consisted of the middle and lower parts of the 3′ SL and the adjacent sHP (6). However, the top portion of the 3′(+) SL RNA was not analyzed by NMR in the previous study. A highly conserved pentanucleotide sequence (5′-CACAG-3′) is present in the top loop of the 3′ SL and has been shown to be essential for genome replication (9). ^{1}H NMR was used to confirm the predicted structure of the top portion of the 3′(+) SLs for WNV and ZIKV. The imino proton spectra for both ZIK35 and WNV32 were consistent with their predicted secondary structures (Figure 22). Signals from nine imino protons were observed for ZIKV35 (Figure 22 A). Signals from eight base pairs were observed for WNV32 (Figure 22 B). One base pair was missing in each spectrum, which is commonly due to base pair opening at the end of the SL.

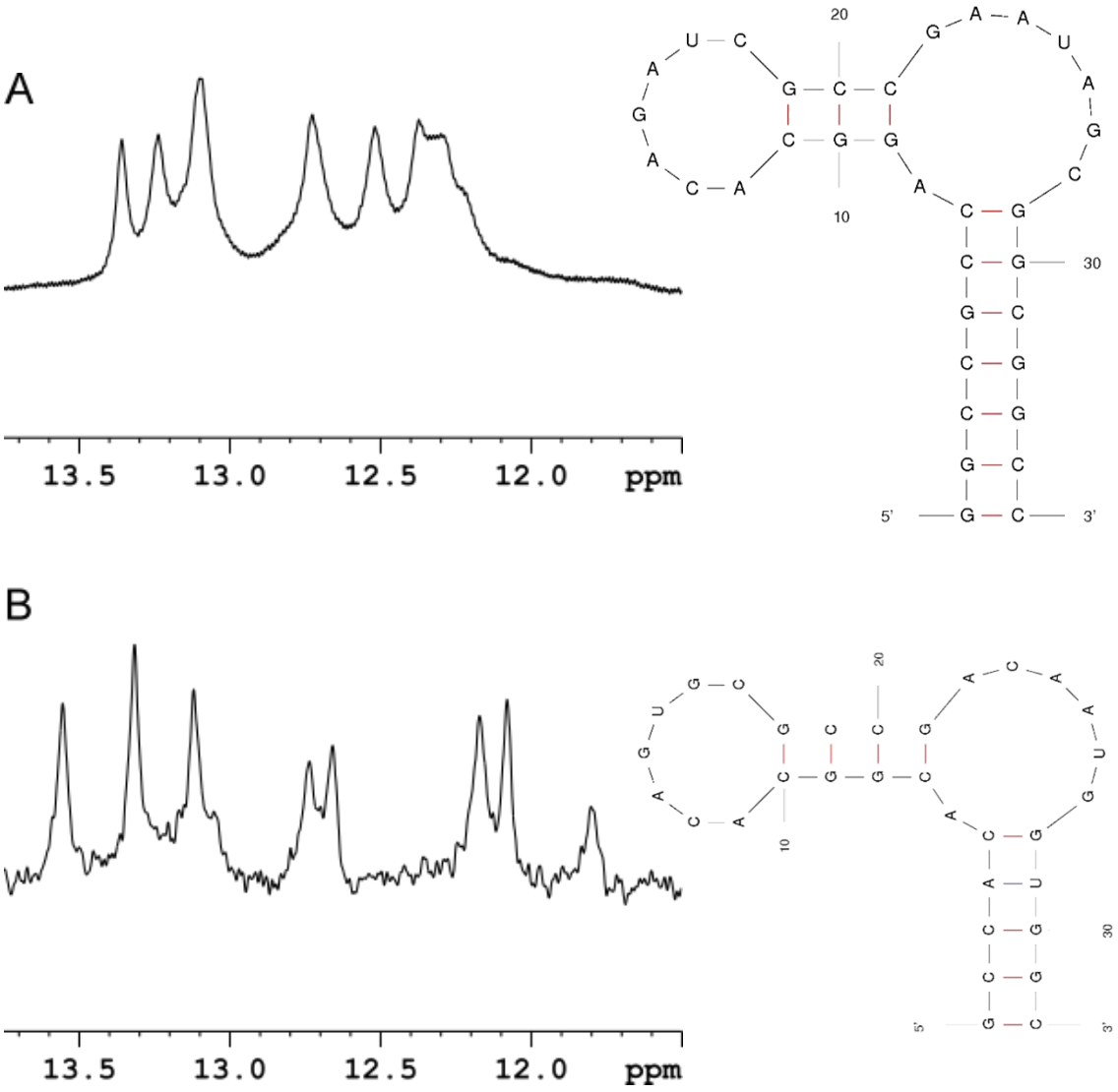


Figure 22. ^{1}H imino proton spectra of ZIKV35 and WNV32. A) Imino proton spectra and the predicted RNA structure for A) ZIKV35 and B)WNV32.

4.4.2 Protein Expression

Recombinant protein expression was attempted for EF1A1 in *E. coli* using a 6x His tag expression system. Two induction temperatures were tested, 15 °C and 30 °C. Protein expression and purification yielded a low amount of purified protein for both expression conditions. Purified protein could not be detected via Coomassie blue staining. A Western blot using anti-EF1A1 primary antibody showed that a large amount of protein remained in the cell pellet and flowthrough for both protein preparations. Multiple bands were detected, which may be due to a combination of protein degradation and background signal caused by nonspecific antibody binding. The primary antibody, obtained from Thermo Fisher, was a rabbit polyclonal antibody against an EF1A1 fusion protein. The bacterial gene EF-Tu was (incorrectly) listed as a gene alias on the product page. It is possible that there was some cross-reactivity between the primary antibody and the bacterial analog.

In an attempt to improve protein solubility, the expression vector was switched to a GST-tag system and the cDNA sequence was codon optimized. A protein expression trial yielded no purified protein visible on a Coomassie blue stained gel. Additionally, no protein was visible on a Western Blot when an anti-GST primary antibody was used.

The difficulty expressing EF1A1 in an *E. coli* system may be due to a number of factors. Protein misfolding may have occurred due to missing post-translational modifications that cannot be achieved in a bacterial system. Additionally, overexpression of mammalian proteins is sometimes toxic to bacteria. Since there is a bacterial analog to EF1A1, it is possible that overexpression of EF1A1 was not well tolerated. This is consistent with what was observed. Post-induction, the *E. coli* cultures grew more slowly than anticipated and a fair amount of protein was detected in the cellular pellets following cell lysis.

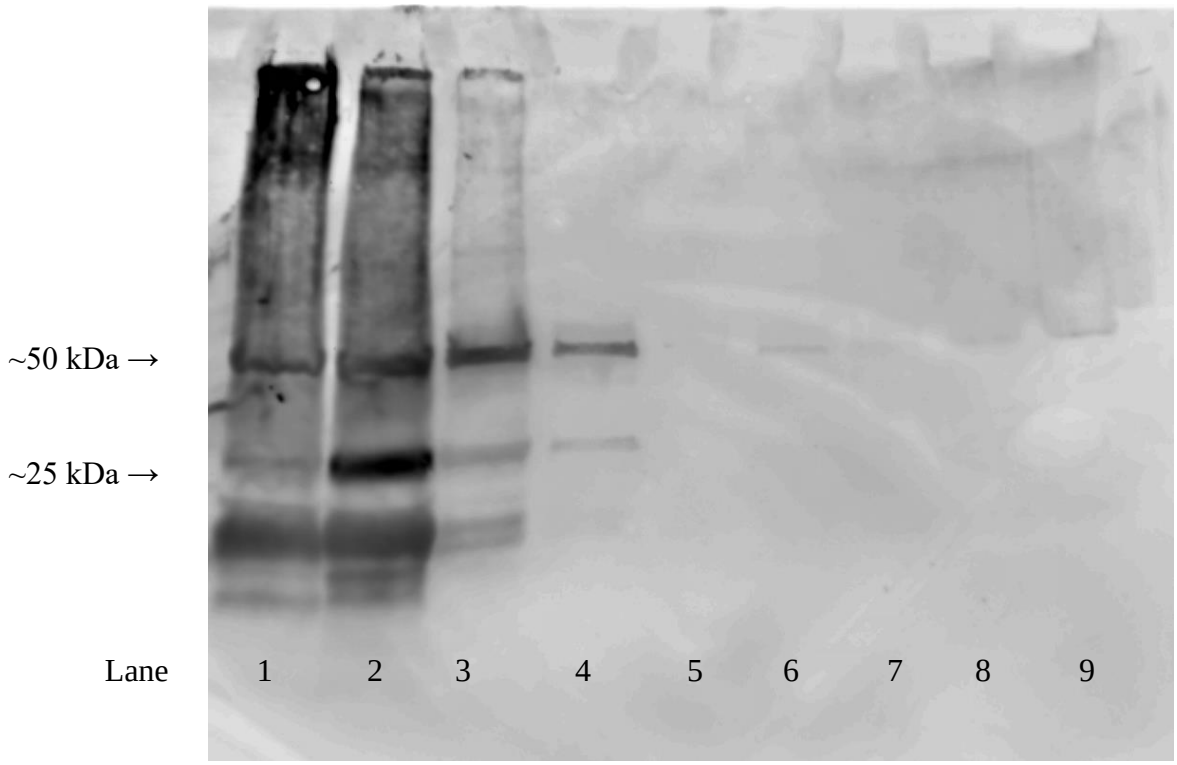


Figure 23. Anti-EF1A1 Western Blot. Following expression and purification of His-tagged EF1A1, a Western Blot was performed on different fractions of the purification process using an anti-EF1A1 primary antibody. Lane 1: cell pellet from 15°C induction, Lane 2: : cell pellet 30°C induction, Lanes 3&4: washes from 15°C induction, Lanes 6&7: elutions from 15°C induction Lanes 8&9: elutions from 30°C induction

4.4.3 Structural Analysis of EF1A1

Prior to 2016 the only structure of EF1A in the PDB was a crystal structure of the less common isoform EF1A2 (PDB 1JNY) which is encoded by a separate gene (10). However, in 2016, a cryo-EM structure of a mammalian translating ribosomal was deposited in the PDB (PDB 5LZS) (11). EF1A1 was spatially isolated from this structure ensemble and used to identify the RNA binding sites and gain other important information. EF1A1 was shown to bind to three different RNAs, charged t-RNA, the sarcin-ricin loop of the 28S rRNA, and a loop of the 18S rRNA (Figure 21). All three EF1A1 domains contact RNA in this structure.

Phosphorylation sites on EF1A1 were previously reported for Ser300 (12) and Thr432 (13). A third phosphorylation site, Ser163, was identified in the cryo-EM structure (Figure 21). Interestingly, the two phosphorylation sites reported in literature were not observed in the cryo-EM structure. Previous work reported that treatment of EF1A1 with calf intestinal alkaline phosphatase inhibited binding to the 3′(+) SL, but the individual sites that are essential for the interaction have not been identified (2). To examine the roles of the two reported and one newly identified phosphorylation site, mutants with phosphomimetic substitutions were designed (See Materials and Methods for mutagenesis primers). Glutamate was selected to replace Ser300, Thr432, and Ser163 because of its negative charge.

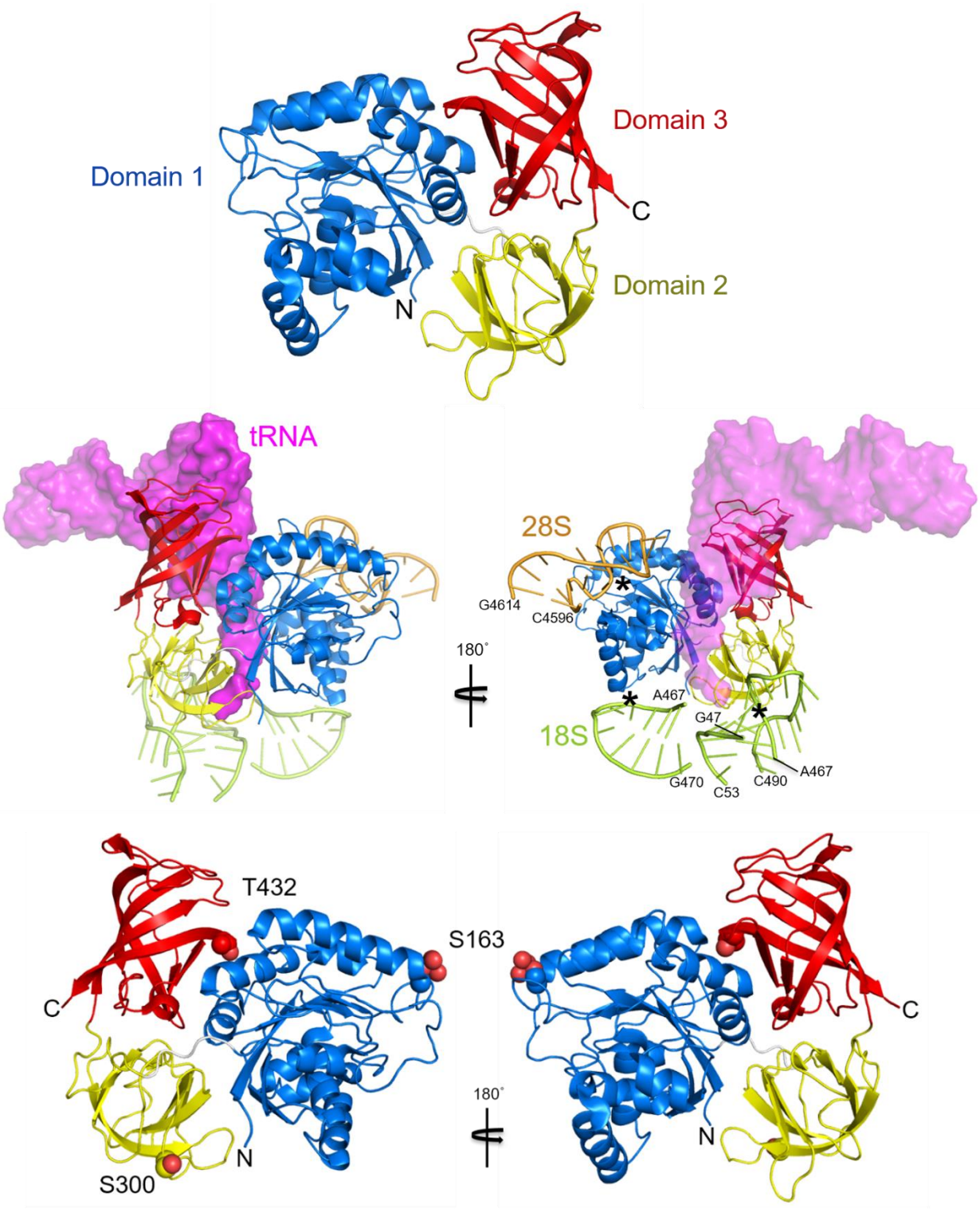


*Figure 24.Structural analysis of EF1A1 from a cryo-EM structure of a translating ribosome. All three domains of EF1A1 have contacts with RNA. An unreported phosphorylation site, Ser163, was identified in Domain 1 of EF1A1. Two previously reported phosphorylation sites Ser300 and T432 were mapped to the structure but were not observed in the deposited structure data. Interaction sites (labeled with *) were identified with the 28S and 18S of the ribosome. (PDB 5LZS)*

4.5 Conclusions and future directions

Imino proton spectra of the tops of the 3′(+) SL loop RNAs of WNV and ZIKV confirmed their predicted secondary structures. Completion of the mapping of the WNV 3′(+) SL structure by NMR provides the ability to monitor RNA folding for the entire SL for future structural work. Analysis of the cryo-EM structure of EF1A1, informed the design of phosphomimic mutants that could be recombinantly expressed. However, recombinant expression of EF1A1 proved challenging, for purifying the protein in amounts adequate for NMR structure work and crystallography.

As previously mentioned, advances in the field of cryo-EM have made it possible for high resolution structures to be obtained for complexes as small as 40 kDa. A major benefit of cryo-EM is the small amount of material needed to obtain a structure. As a result, purification of endogenous EF1A1 from mammalian cells has become a viable option for EF1A structural studies. An added benefit of this method is that this EF1A1 should have all the essential post-translational modifications for optimal binding to the WNV 3′(+) SL if purified in phosphate buffer.

4.6 References

1. Blackwell, J. L., & Brinton, M. A. (1995). BHK cell proteins that bind to the 3'stem-loop structure of the West Nile virus genome RNA. Journal of Virology, 69(9), 5650-5658.
2. Blackwell, J. L., & Brinton, M. A. (1997). Translation elongation factor-1 alpha interacts with the 3'stem-loop region of West Nile virus genomic RNA. Journal of virology, 71(9), 6433-6444.
3. Beckham, S. A., Matak, M. Y., Belousoff, M. J., Venugopal, H., Shah, N., Vankadari, N., ... & Wilce, J. A. (2020). Structure of the PCBP2/stem–loop IV complex underlying translation initiation mediated by the poliovirus type I IRES. Nucleic acids research, 48(14), 8006-8021.
4. Davis, W. G., Blackwell, J. L., Shi, P. Y., & Brinton, M. A. (2007). Interaction between the cellular protein eEF1A and the 3′-terminal stem-loop of West Nile virus genomic RNA facilitates viral minus-strand RNA synthesis. Journal of virology, 81(18), 10172-10187.
5. Lindenbach, B.D., Thiel, H.J., & Rice, C.M. (2007). Flaviviridae: The Viruses and Their Replication. In D. Knipe & P. Howly (Eds.), Fields Virology (5th ed.) (1101-1152). Lippincot Williams and Wilkins
6. Davis, W. G., Basu, M., Elrod, E. J., Germann, M. W., & Brinton, M. A. (2013). Identification of cis-acting nucleotides and a structural feature in West Nile virus 3′-terminus RNA that facilitate viral minus strand RNA synthesis. Journal of virology, 87(13), 7622-7636.

7. Matsuda, D., Yoshinari, S., & Dreher, T. W. (2004). eEF1A binding to aminoacylated viral RNA represses minus strand synthesis by TYMV RNA-dependent RNA polymerase. Virology, 321(1), 47-56.
8. Li, Z., Pogany, J., Tupman, S., Esposito, A. M., Kinzy, T. G., & Nagy, P. D. (2010). Translation elongation factor 1A facilitates the assembly of the tombusvirus replicase and stimulates minus-strand synthesis. PLoS Pathog, 6(11), e1001175.
9. Tilgner, M., Deas, T. S., & Shi, P. Y. (2005). The flavivirus-conserved penta-nucleotide in the 3′ stem-loop of the West Nile virus genome requires a specific sequence and structure for RNA synthesis, but not for viral translation. Virology, 331(2), 375-386.
10. Vitagliano, L. (2001). Crystal structure of sulfolobus solfataricus elongation factor 1 α in complex with GDP. FEBS Journal, 268(S1), 256-190.
11. Shao, S., Murray, J., Brown, A., Taunton, J., Ramakrishnan, V., & Hegde, R. S. (2016). Decoding mammalian ribosome-mRNA states by translational GTPase complexes. Cell, 167(5), 1229-1240.
12. Lin, K. W., Yakymovych, I., Jia, M., Yakymovych, M., & Souchelnytskyi, S. (2010). Phosphorylation of eEF1A1 at Ser300 by TβR-I results in inhibition of mRNA translation. Current Biology, 20(18), 1615-1625.
13. Eckhardt, K., Tröger, J., Reissmann, J., Katschinski, D. M., Wagner, K. F., Stengel, P., ... & Schläfli, P. (2007). Male germ cell expression of the PAS domain kinase PASKIN and its novel target eukaryotic translation elongation factor eEF1A1. Cellular Physiology and Biochemistry, 20(1-4), 227-240.

5 CONCLUSIONS

The work presented in the previous chapters used a number of analytical chemistry techniques to probe the interactions between host cellular proteins and viral RNA. Chapter 2 explored the stoichiometry of TIAR-RNA complexes using a linear single stranded oligos as the model system. In Chapter 3, multiple titration experiments were used to explore the structure of TIA-1/TIAR-3′(-) SL complexes. Chapter 4 provided initial structural information for the 3′(+) SL RNA and EF1A1 that can be used to inform future studies. Further development of structural analysis techniques, such as cryo-EM, and data processing methods, such as global ITC fitting, provides promise for obtaining further information relevant to elucidation of the roles of these cellular protein-viral RNA interactions in the regulation of flavivirus RNA replication.

REFERENCES

Afroz, T., Cienikova, Z., Cléry, A., & Allain, F. H. (2015). One, two, three, four! How multiple RRMs read the genome sequence. In Methods in enzymology (Vol. 558, pp. 235-278). Academic Press.

Basu, M., & Brinton, M. A. (2011). West Nile virus (WNV) genome RNAs with up to three adjacent mutations that disrupt long distance 5′–3′ cyclization sequence basepairs are viable. Virology, 412(1), 220-232.

Bauer, W. J., Heath, J., Jenkins, J. L., & Kielkopf, C. L. (2012). Three RNA recognition motifs participate in RNA recognition and structural organization by the pro-apoptotic factor TIA-1. Journal of molecular biology, 415(4), 727-740.

Beck, A. R., Medley, Q. G., O'Brien, S., Anderson, P., & Streuli, M. (1996). Structure, tissue distribution and genomic organization of the murine RRM-type RNA binding proteins TIA-1 and TIAR. Nucleic acids research, 24(19), 3829-3835.

Beckham, S. A., Matak, M. Y., Belousoff, M. J., Venugopal, H., Shah, N., Vankadari, N., ... & Wilce, J. A. (2020). Structure of the PCBP2/stem–loop IV complex underlying translation initiation mediated by the poliovirus type I IRES. Nucleic acids research, 48(14), 8006-8021.

Bhatt, S., Gething, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., ... & Myers, M. F. (2013). The global distribution and burden of dengue. Nature, 496(7446), 504-507.

Blackwell, J. L., & Brinton, M. A. (1995). BHK cell proteins that bind to the 3'stem-loop structure of the West Nile virus genome RNA. Journal of Virology, 69(9), 5650-5658.

Blackwell, J. L., & Brinton, M. A. (1997). Translation elongation factor-1 alpha interacts with the 3'stem-loop region of West Nile virus genomic RNA. Journal of virology, 71(9), 6433-6444.

Brinton, M. A. (2014). Replication cycle and molecular biology of the West Nile virus. Viruses, 6(1), 13-53.

Brinton, M. A., & Basu, M. (2015). Functions of the 3′ and 5′ genome RNA regions of members of the genus Flavivirus. Virus research, 206, 108-119.

Brinton, M. A., Fernandez, A. V., & Dispoto, J. H. (1986). The 3′-nucleotides of flavivirus genomic RNA form a conserved secondary structure. Virology, 153(1), 113-121.

Brown, A. (2009). Analysis of cooperativity by isothermal titration calorimetry. International journal of molecular sciences, 10(8), 3457-3477.

Collins, M. H., & Metz, S. W. (2017). Progress and works in progress: update on flavivirus vaccine development. Clinical therapeutics, 39(8), 1519-1536.

Corrêa, D. H., & Ramos, C. H. (2009). The use of circular dichroism spectroscopy to study protein folding, form and function. African Journal of Biochemistry Research, 3(5), 164-173.

Cruz-Gallardo, I., Aroca, Á., Gunzburg, M. J., Sivakumaran, A., Yoon, J. H., Angulo, J., ... & Díaz-Moreno, I. (2014). The binding of TIA-1 to RNA C-rich sequences is driven by its C-ter

Cruz-Gallardo, I., Aroca, Á., Persson, C., Karlsson, B. G., & Díaz-Moreno, I. (2013). RNA binding of T-cell intracellular antigen-1 (TIA-1) C-terminal RNA recognition motif is modified by pH conditions. Journal of Biological Chemistry, 288(36), 25986-25994.minal RRM domain. RNA biology, 11(6), 766-776.

Davis, W. G., Basu, M., Elrod, E. J., Germann, M. W., & Brinton, M. A. (2013). Identification of cis-acting nucleotides and a structural feature in West Nile virus 3′-terminus RNA that facilitate viral minus strand RNA synthesis. Journal of virology, 87(13), 7622-7636.

Davis, W. G., Blackwell, J. L., Shi, P. Y., & Brinton, M. A. (2007). Interaction between the cellular protein eEF1A and the 3′-terminal stem-loop of West Nile virus genomic RNA facilitates viral minus-strand RNA synthesis. Journal of virology, 81(18), 10172-10187.

Deas, T. S., Binduga-Gajewska, I., Tilgner, M., Ren, P., Stein, D. A., Moulton, H. M., ... & Shi, P. Y. (2005). Inhibition of flavivirus infections by antisense oligomers specifically suppressing viral translation and RNA replication. Journal of virology, 79(8), 4599-4609.

Dember, L. M., Kim, N. D., Liu, K. Q., & Anderson, P. (1996). Individual RNA recognition motifs of TIA-1 and TIAR have different RNA binding specificities. Journal of Biological Chemistry, 271(5), 2783-2788.

Domingo, E. (2007). Virus Evolution. In D. Knipe & P. Howly (Eds.), Fields Virology (5th ed.) (641-705). Lippincot Williams and Wilkins.

Drake, J. W. (1999). The distribution of rates of spontaneous mutation over viruses, prokaryotes, and eukaryotes. ANNALS-NEW YORK ACADEMY OF SCIENCES, 870, 100-107.

Eckhardt, K., Tröger, J., Reissmann, J., Katschinski, D. M., Wagner, K. F., Stengel, P., ... & Schläfli, P. (2007). Male germ cell expression of the PAS domain kinase PASKIN and its novel target eukaryotic translation elongation factor eEF1A1. Cellular Physiology and Biochemistry, 20(1-4), 227-240.

Einstein, A. (1905). On the electrodynamics of moving bodies. Annalen der physik, 17(10), 891-921.

Elghonemy, S., Davis, W. G., & Brinton, M. A. (2005). The majority of the nucleotides in the top loop of the genomic 3′ terminal stem loop structure are cis-acting in a West Nile virus infectious clone. Virology, 331(2), 238-246.

Emara, M. M., & Brinton, M. A. (2007). Interaction of TIA-1/TIAR with West Nile and dengue virus products in infected cells interferes with stress granule formation and processing body assembly. Proceedings of the National Academy of Sciences, 104(21), 9041-9046

Emara, M. M., Liu, H., Davis, W. G., & Brinton, M. A. (2008). Mutation of mapped TIA-1/TIAR binding sites in the 3′ terminal stem-loop of West Nile virus minus-strand RNA in an infectious clone negatively affects genomic RNA amplification. Journal of virology, 82(21), 10657-10670.

Freire, E., Schön, A., & Velazquez-Campoy, A. (2009). Isothermal titration calorimetry: general formalism using binding polynomials. Methods in enzymology, 455, 127-155.

Förch, P., Puig, O., Martínez, C., Séraphin, B., & Valcárcel, J. (2002). The splicing regulator TIA-1 interacts with U1-C to promote U1 snRNP recruitment to 5′ splice sites. The EMBO journal, 21(24), 6882-6892.

Friedrich, S., Engelmann, S., Schmidt, T., Szczepankiewicz, G., Bergs, S., Liebert, U. G., ... & Behrens, S. E. (2018). The host factor AUF1 p45 supports flavivirus propagation by triggering the RNA switch required for viral genome cyclization. Journal of virology, 92(6).

Friedrich, S., Schmidt, T., Geissler, R., Lilie, H., Chabierski, S., Ulbert, S., ... & Behrens, S. E. (2014). AUF1 p45 promotes West Nile virus replication by an RNA chaperone activity that supports cyclization of the viral genome. Journal of virology, 88(19), 11586-11599.

Gubler, D., Kuno, D., & Markoff, L. (2007). Flaviviruses. In D. Knipe & P. Howly (Eds.), Fields Virology (5th ed.) (1101-1152). Lippincot Williams and Wilkins.

Hernandez, R., Brown, D. T., & Paredes, A. (2014). Structural differences observed in arboviruses of the alphavirus and flavivirus genera. Advances in Virology, 2014.

Holden, K. L., Stein, D. A., Pierson, T. C., Ahmed, A. A., Clyde, K., Iversen, P. L., & Harris, E. (2006). Inhibition of dengue virus translation and RNA synthesis by a morpholino oligomer targeted to the top of the terminal 3′ stem–loop structure. Virology, 344(2), 439-452.

Jia, M., Gut, H., & Chao, J. A. (2018). Structural basis of IMP3 RRM12 recognition of RNA. RNA, 24(12), 1659-1666.

Kaufmann, S. H., Dorhoi, A., Hotchkiss, R. S., & Bartenschlager, R. (2018). Host-directed therapies for bacterial and viral infections. Nature Reviews Drug Discovery, 17(1), 35.

Kawakami, A., Tian, Q., Duan, X., Streuli, M., Schlossman, S. F., & Anderson, P. (1992). Identification and functional characterization of a TIA-1-related nucleolysin. Proceedings of the National Academy of Sciences, 89(18), 8681-8685.

Kielkopf, C. L., Rodionova, N. A., Green, M. R., & Burley, S. K. (2001). A novel peptide recognition mode revealed by the X-ray structure of a core U2AF35/U2AF65 heterodimer. Cell, 106(5), 595-605.

Kim, H. S., Headey, S. J., Yoga, Y. M., Scanlon, M. J., Gorospe, M., Wilce, M. C., & Wilce, J. A. (2013). Distinct binding properties of TIAR RRMs and linker region. RNA biology, 10(4), 579-589.

Kim, H. S., Kuwano, Y., Zhan, M., Pullmann, R., Mazan-Mamczarz, K., Li, H., ... & Wilce, J. A. (2007). Elucidation of a C-rich signature motif in target mRNAs of RNA-binding protein TIAR. Molecular and cellular biology, 27(19), 6806-6817.

Kim, H. S., Wilce, M. C., Yoga, Y. M., Pendini, N. R., Gunzburg, M. J., Cowieson, N. P., ... & Wilce, J. A. (2011). Different modes of interaction by TIAR and HuR with target RNA and DNA. Nucleic acids research, 39(3), 1117-1130.

Kumar, A. O., Swenson, M. C., Benning, M. M., & Kielkopf, C. L. (2008). Structure of the central RNA recognition motif of human TIA-1 at 1.95 Å resolution. Biochemical and biophysical research communications, 367(4), 813-819.

Kuwasako, K., Takahashi, M., Tochio, N., Abe, C., Tsuda, K., Inoue, M., ... & Taguchi, S. (2008). Solution structure of the second RNA recognition motif (RRM) domain of murine T cell intracellular antigen-1 (TIA-1) and its RNA recognition mode. Biochemistry, 47(24), 6437-6450.

Lederberg, J., Hamburg, M. A., & Smolinski, M. S. (Eds.). (2003). Microbial threats to health: emergence, detection, and response. National Academies Press.

Li, W., Li, Y., Kedersha, N., Anderson, P., Emara, M., Swiderek, K. M., ... & Brinton, M. A. (2002). Cell proteins TIA-1 and TIAR interact with the 3′ stem-loop of the West Nile virus complementary minus-strand RNA and facilitate virus replication. Journal of virology, 76(23), 11989-12000.

Li, Z., Pogany, J., Tupman, S., Esposito, A. M., Kinzy, T. G., & Nagy, P. D. (2010). Translation elongation factor 1A facilitates the assembly of the tombusvirus replicase and stimulates minus-strand synthesis. PLoS Pathog, 6(11), e1001175.

Lindenbach, B.D., Thiel, H.J., & Rice, C.M. (2007). Flaviviridae: The Viruses and Their Replication. In D. Knipe & P. Howly (Eds.), Fields Virology (5th ed.) (1101-1152). Lippincot Williams and Wilkins.

Lin, K. W., Yakymovych, I., Jia, M., Yakymovych, M., & Souchelnytskyi, S. (2010). Phosphorylation of eEF1A1 at Ser300 by TβR-I results in inhibition of mRNA translation. Current Biology, 20(18), 1615-1625.

Martinez, J. C., Murciano-Calles, J., Cobos, E. S., Iglesias-Bexiga, M., Luque, I., & Ruiz-Sanz, J. (2013). Isothermal titration calorimetry: thermodynamic analysis of the binding thermograms of molecular recognition events by using equilibrium models (Vol. 4). chapter.

Matsuda, D., Yoshinari, S., & Dreher, T. W. (2004). eEF1A binding to aminoacylated viral RNA represses minus strand synthesis by TYMV RNA-dependent RNA polymerase. Virology, 321(1), 47-56.

McGhee, J. D., & von Hippel, P. H. (1974). Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice. Journal of molecular biology, 86(2), 469-489.

Mironov, A. S., Gusarov, I., Rafikov, R., Lopez, L. E., Shatalin, K., Kreneva, R. A., ... & Nudler, E. (2002). Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. Cell, 111(5), 747-756.

Mlakar, J., Korva, M., Tul, N., Popović, M., Poljšak-Prijatelj, M., Mraz, J., ... & Vizjak, A. (2016). Zika virus associated with microcephaly. New England Journal of Medicine, 374(10), 951-958.

Monie, T. P., Hernandez, H., Robinson, C. V., Simpson, P., Matthews, S., & Curry, S. (2005). The polypyrimidine tract binding protein is a monomer. Rna, 11(12), 1803-1808.

Montemayor, E. J., Curran, E. C., Liao, H. H., Andrews, K. L., Treba, C. N., Butcher, S. E., & Brow, D. A. (2014). Core structure of the U6 small nuclear ribonucleoprotein at 1.7-Å resolution. Nature structural & molecular biology, 21(6), 544-551.

Mukhopadhyay, S., Kuhn, R. J., & Rossmann, M. G. (2005). A structural perspective of the flavivirus life cycle. Nature Reviews Microbiology, 3(1), 13-22.

Muto, Y., & Yokoyama, S. (2012). Structural insight into RNA recognition motifs: versatile molecular Lego building blocks for biological systems. Wiley Interdisciplinary Reviews: RNA, 3(2), 229-246.

Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L., & Breaker, R. R. (2002). Genetic control by a metabolite binding mRNA. Chemistry & biology, 9(9), 1043-1049.

Panavas, T., Stork, J., & Nagy, P. D. (2006). Use of double-stranded RNA templates by the tombusvirus replicase in vitro: Implications for the mechanism of plus-strand initiation. Virology, 352(1), 110-120.

Peck, K. M., & Lauring, A. S. (2018). Complexities of viral mutation rates. Journal of virology, 92(14).

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. Journal of computational chemistry, 25(13), 1605-1612.

Pierson, T. C., & Diamond, M. S. (2020). The continued threat of emerging flaviviruses. Nature Microbiology, 1-17.

Piñeiro, Á., Muñoz, E., Sabín, J., Costas, M., Bastos, M., Velázquez-Campoy, A., ... & Rial, J. (2019). AFFINImeter: A software to analyze molecular recognition processes from experimental data. Analytical biochemistry, 577, 117-134.

Ruggiu, M., & Cooke, H. J. (2000). In vivo and in vitro analysis of homodimerisation activity of the mouse Dazl1 protein. Gene, 252(1-2), 119-126.

Sanjuán, R., & Domingo-Calap, P. (2016). Mechanisms of viral mutation. Cellular and molecular life sciences: CMLS, 73(23), 4433–4448. https://doi.org/10.1007/s00018-016-2299-6.

Shao, S., Murray, J., Brown, A., Taunton, J., Ramakrishnan, V., & Hegde, R. S. (2016). Decoding mammalian ribosome-mRNA states by translational GTPase complexes. Cell, 167(5), 1229-1240.

Shi, P. Y., Li, W., & Brinton, M. A. (1996). Cell proteins bind specifically to West Nile virus minus-strand 3'stem-loop RNA. Journal of virology, 70(9), 6278-6287.

Ta, M., & Vrati, S. (2000). Mov34 protein from mouse brain interacts with the 3′ noncoding region of Japanese encephalitis virus. Journal of virology, 74(11), 5108-5115.

Tian, Q., Streuli, M., Saito, H., Schlossman, S. F., & Anderson, P. (1991). A polyadenylate binding protein localized to the granules of cytolytic lymphocytes induces DNA fragmentation in target cells. Cell, 67(3), 629-639.

Tilgner, M., Deas, T. S., & Shi, P. Y. (2005). The flavivirus-conserved penta-nucleotide in the 3′ stem-loop of the West Nile virus genome requires a specific sequence and structure for RNA synthesis, but not for viral translation. Virology, 331(2), 375-386.

Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computational chemistry, 31(2), 455-461.

Velázquez-Campoy, A. (2006). Ligand binding to one-dimensional lattice-like macromolecules: analysis of the McGhee–von Hippel theory implemented in isothermal titration calorimetry. Analytical biochemistry, 348(1), 94-104.

Vitagliano, L. (2001). Crystal structure of sulfolobus solfataricus elongation factor 1 α in complex with GDP. FEBS Journal, 268(S1), 256-190.

Wang, I., Hennig, J., Jagtap, P. K. A., Sonntag, M., Valcárcel, J., & Sattler, M. (2014). Structure, dynamics and RNA binding of the multi-domain splicing factor TIA-1. Nucleic acids research, 42(9), 5949-5966.

Waris, S., García-Mauriño, S. M., Sivakumaran, A., Beckham, S. A., Loughlin, F. E., Gorospe, M., ... & Wilce, J. A. (2017). TIA-1 RRM23 binding and recognition of target oligonucleotides. Nucleic acids research, 45(8), 4944-4957.`

White, D. E., White, D. O., White, D. O., & Fenner, F. J. (1994). Medical virology. Gulf Professional Publishing.

Winkler, W. C., Nahvi, A., Sudarsan, N., Barrick, J. E., & Breaker, R. R. (2003). An mRNA structure that controls gene expression by binding S-adenosylmethionine. Nature Structural & Molecular Biology, 10(9), 701-707.

World Health Organization. (2018). Dengue vaccine: WHO position paper–September 2018. Wkly Epidemiol Rec, 93(36), 457-476.

Yocupicio-Monroy, M., Padmanabhan, R., Medina, F., & del Angel, R. M. (2007). Mosquito La protein binds to the 3′ untranslated region of the positive and negative polarity dengue virus RNAs and relocates to the cytoplasm of infected cells. Virology, 357(1), 29-40.

Zając, M., Muszalska, I., Sobczak, A., Dadej, A., Tomczak, S., & Jelińska, A. (2019). Hepatitis C–New drugs and treatment prospects. European Journal of Medicinal Chemistry, 165, 225-249.

Zhang, J., Pearson, J. Z., Gorbet, G. E., Cölfen, H., Germann, M. W., Brinton, M. A., & Demeler, B. (2017). Spectral and Hydrodynamic Analysis of West Nile Virus RNA–Protein Interactions by Multiwavelength Sedimentation Velocity in the Analytical Ultracentrifuge. Analytical chemistry, 89(1), 862-870.

Zhang, K., Li, S., Kappel, K., Pintilie, G., Su, Z., Mou, T. C., ... & Chiu, W. (2019). Cryo-EM structure of a 40 kDa SAM-IV riboswitch RNA at 3.7 Å resolution. Nature communications, 10(1), 1-6.

Zhang, L. (2018). Multi-epitope vaccines: a promising strategy against tumors and viral infections. Cellular & molecular immunology, 15(2), 182-184.

Zhang, T., Delestienne, N., Huez, G., Kruys, V., & Gueydan, C. (2005). Identification of the sequence determinants mediating the nucleo-cytoplasmic shuttling of TIAR and TIA-1 RNA-binding proteins. Journal of cell science, 118(23), 5453-5463.

Zhao, H., Piszczek, G., & Schuck, P. (2015). SEDPHAT–a platform for global ITC analysis and global multi-method analysis of molecular interactions. Methods, 76, 137-148.

APPENDICES

Appendix A: ITC Fitting

Appendix A.1 Independent and Stoichiometric Models

There are two main types of models for fitting ITC data, independent site models and stoichiometric models. Both types of models are available in the software packages included with calorimetry instrumentation. For example, MicroCal Origin ITC Software, offers one-site independent, two-site independent and sequential (stoichiometric) models as preloaded fitting options. Global fitting software such as the open-source SEDPHAT (1) and the commercially available AFFINImeter (2) provide a wider range of models for fitting. The choice between an independent one-site model and a stoichiometric model for a 1:1 ratio is rather inconsequential; however, as the number of sites and binding partners increases, the selection of an appropriate model becomes more challenging and more essential.

Appendix A.1.1 Mathematical Basis of the Independent Two-Site Model

The most important assumption in an independent two-site model is that there is no interaction between sites. This means that there is no positive or negative cooperativity in the system. As a result, the association constants are only dependent on the fraction of sites occupied by a ligand (Θ). For an independent two-site model the site specific, also called microscopic association constants, are calculated using the equation:

$$k_i = \frac{\Theta_i}{(1-\Theta_i)[L]} \quad (1)$$

where i is the number of sites, k_i is the microscopic association constant for site "i", Θ_i is the fractional occupancy, and [L] is the concentration of free ligand.

The microscopic association constants for a two-site system can also be described using the following equations:

$$k_1 = \frac{[M^L]}{[M][L]},\ k_2 = \frac{[M_L]}{[M][L]} \quad (2, 3)$$

Appendix A.1.2 Mathematical Basis of the Stoichiometric Two-Site Model

Stoichiometric models are based on the macroscopic "step-wise" association constants. In these types of models there is no differentiation between binding sites, meaning that the complex $[M^L]$ (singly-bound ligand occupying site 1) is the treated as equivalent to $[M_L]$ (singly-bound ligand occupying site 2) and represented as a single species [ML]. The stepwise association constants, K_1 and K_2 for a two-site model are given by:

$$K_1 = \frac{[ML]}{[M][L]},\ K_2 = \frac{[ML_2]}{[ML][L]} \quad (4, 5)$$

The microscopic association constants can be related to the macroscopic association constants.

$$[M^L] + [M_L] = [ML] \quad (6)$$

Therefore,

$$K_1 = k_1 + k_2 \quad (7)$$

$$k_{1,2} = \frac{[ML_2]}{[M^L][L]},\ k_{2,1} = \frac{[ML_2]}{[M_L][L]} \quad (8)$$

$$K_2 = \frac{k_{1,2}k_{2,1}}{k_{1,2}+k_{2,1}} \quad (9)$$

Global association constants can also be calculated for a two-site model using the equation:

$$\beta_1 = \frac{[ML]}{[M][L]}, \beta_2 = \frac{[ML_2]}{[M][L]} \quad (10, 11)$$

The model employs a general binding polynomial, which does not assume independence between sites. For a two-site model the binding polynomial is:

$$Z = 1 + \beta_1[L] + \beta_2[L]^2 \quad (12)$$

or in terms of the stepwise association constants:

$$Z = 1 + K_1[L] + K_1K_2[L]^2 \quad (13)$$

or in terms of the microscopic association constants:

$$Z = 1 + (k_1 + k_2)[L] + \alpha k_1 k_2[L]^2 \quad (14)$$

where α is the cooperativity factor.

From Equations 6, 7 and 8 the relationship between the association constants, are derived as:

$$\beta_1 = K_1 = (k_1 + k_2), \beta_2 = K_1K_2 = \alpha k_1 k_2 \quad (15, 16)$$

Appendix A.2 Calculating the Interaction Parameter and Understanding Cooperativity

In a two-site model there is a single interaction factor ρ. The interaction factor ρ, or the deviation from independent site binding is given by the relationship:

$$\rho = \frac{4K_2}{K_1} = \frac{4\beta_2}{{\beta_1}^2} \quad (17)$$

This relationship is derived from the case where there are two independent identical sites. In this situation $k_1 = k_2$ and $\alpha = 1$. Therefore, K_1, which is given by the sum of k_1 and k_2, is equal to 2k. In this case, K_2 which is given by $\frac{k_1k_2}{(k_1+k_2)}$, simplifies to $\frac{k}{2}$. Thus, $K_1 = 4K_2$. The interaction factor, ρ, is a measure of how much the system deviates from the identical and independent case.

The interaction factor ρ can be used as the cooperativity factor α in a set number of cases:

1) Two identical independent sites.
2) Two identical sites with negative cooperativity.
3) Two identical sites with positive cooperativity.

Three more cases exist where ρ is not equivalent to α:

4) Two nonidentical independent sites. ($\rho < 1$, $\alpha = 1$, $k_1 \neq k_2$)

$K_1 = (k_1 + k_2)$, $K_2 = \alpha k_1 k_2/(k_1 + k_2)$

In this case, nonidentical sites are defined by $k_2 < k_1$. Independent sites have no cooperativity and therefore $\alpha = 1$. As a result, $4K_2 < K_1$ and $\rho < 1$.

5) Two nonidentical sites with negative cooperativity. ($\rho < 1$, $\alpha < 1$, $k_1 \neq k_2$)

   In this case, nonidentical sites are defined by $k_2 < k_1$. Negative cooperativity occurs when $\alpha < 1$. Therefore, $4K_2 < K_1$ and $\rho < 1$.

6) Two nonidentical sites with positive cooperativity. ($\alpha > 1$, $k_1 \neq k_2$)

   In this case, nonidentical sites are defined by $k_2 < k_1$. Positive cooperativity occurs when $\alpha > 1$. As a result, $4K_2$ may be greater than, less than or equal to K_1. Therefore, ρ may be any value including 1.

The Hill coefficient of the system is related to alpha by the equation:

$$nH = \frac{2}{1+\sqrt{\frac{1}{\alpha}}} \tag{18}$$

Appendix A.4 Conclusions

The systems presented in Chapter 2 are particularly challenging to fit because the effects of RNA length and cooperativity on the macroscopic association constants. The McGhee-von Hippel model for protein-nucleic acid binding seeks to compensate for length effects by treating linear nucleic acid sequence as lattice like repeats of a set binding footprint (4). Models using this method that account for cooperativity have been developed (5). However, these types of models are not included in widely available data analysis software (ie. MicroCal Origin, Sedphat, AFFINImeter), and are challenging to apply correctly to experimental data. Therefore, in terms of practical application of binding models in the analysis of ITC data, the best practice is to pick

the most generalized form of the binding model based on known information about the system and employ multiple experimental methods to collect information to refine the model.

Appendix A.4 References

1. Zhao, H., Piszczek, G., & Schuck, P. (2015). SEDPHAT–a platform for global ITC analysis and global multi-method analysis of molecular interactions. Methods, 76, 137-148.
2. Piñeiro, Á., Muñoz, E., Sabín, J., Costas, M., Bastos, M., Velázquez-Campoy, A., ... & Rial, J. (2019). AFFINImeter: A software to analyze molecular recognition processes from experimental data. Analytical biochemistry, 577, 117-134.
3. Freire, E., Schön, A., & Velazquez-Campoy, A. (2009). Isothermal titration calorimetry: general formalism using binding polynomials. Methods in enzymology, 455, 127-155.
4. McGhee, J. D., & von Hippel, P. H. (1974). Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice. Journal of molecular biology, 86(2), 469-489.
5. Velázquez-Campoy, A. (2006). Ligand binding to one-dimensional lattice-like macromolecules: analysis of the McGhee–von Hippel theory implemented in isothermal titration calorimetry. Analytical biochemistry, 348(1), 94-104.

Appendix B: Cryo-EM Sample Preparation

In order to optimize conditions for sufficient particle formation, a number of samples were prepared and sent for screening at the Stanford-SLAC Cryo-EM Facility using a Talos Arctica cryo-electron microscope (Thermo Fisher Scientific) operated at 200 kV. At the time of writing this dissertation, an optimal condition has not been confirmed; however, prescreening by native PAGE, size exclusion chromatography, and NMR data indicate that the most recently prepared samples showed enhanced complex enrichment and appear suitable for future structural work.

Appendix B.1 Screening Trial 1

Three samples were prepared using a 1:1, 2:1 and 3:1 ratio of RRM123 and WNV75. For the 1:1 sample, 20 μM RRM123 was added to 20 μM WNV75, yielding a sample concentration of ~1 mg/mL. For the 1:1 sample, 40 μM RRM123 was added to 20 μM WNV75, yielding a sample concentration of ~1.5 mg/mL. For the 3:1 sample, 60 μM RRM123 was added to 20 μM WNV75, yielding a sample concentration of ~2.0 mg/mL. Each sample was prepared in 10 mM PBS, 50 mM NaCl, 1 mM EDTA, pH 6.8. Samples were analyzed by NMR to confirm complex formation through the chemical shift of the RRM123 methyl groups. Each sample was lyophilized overnight and shipped dry. Screening results indicate that there is a high background of unbound RNA (Figure 26). These samples were not suitable for further structure work.

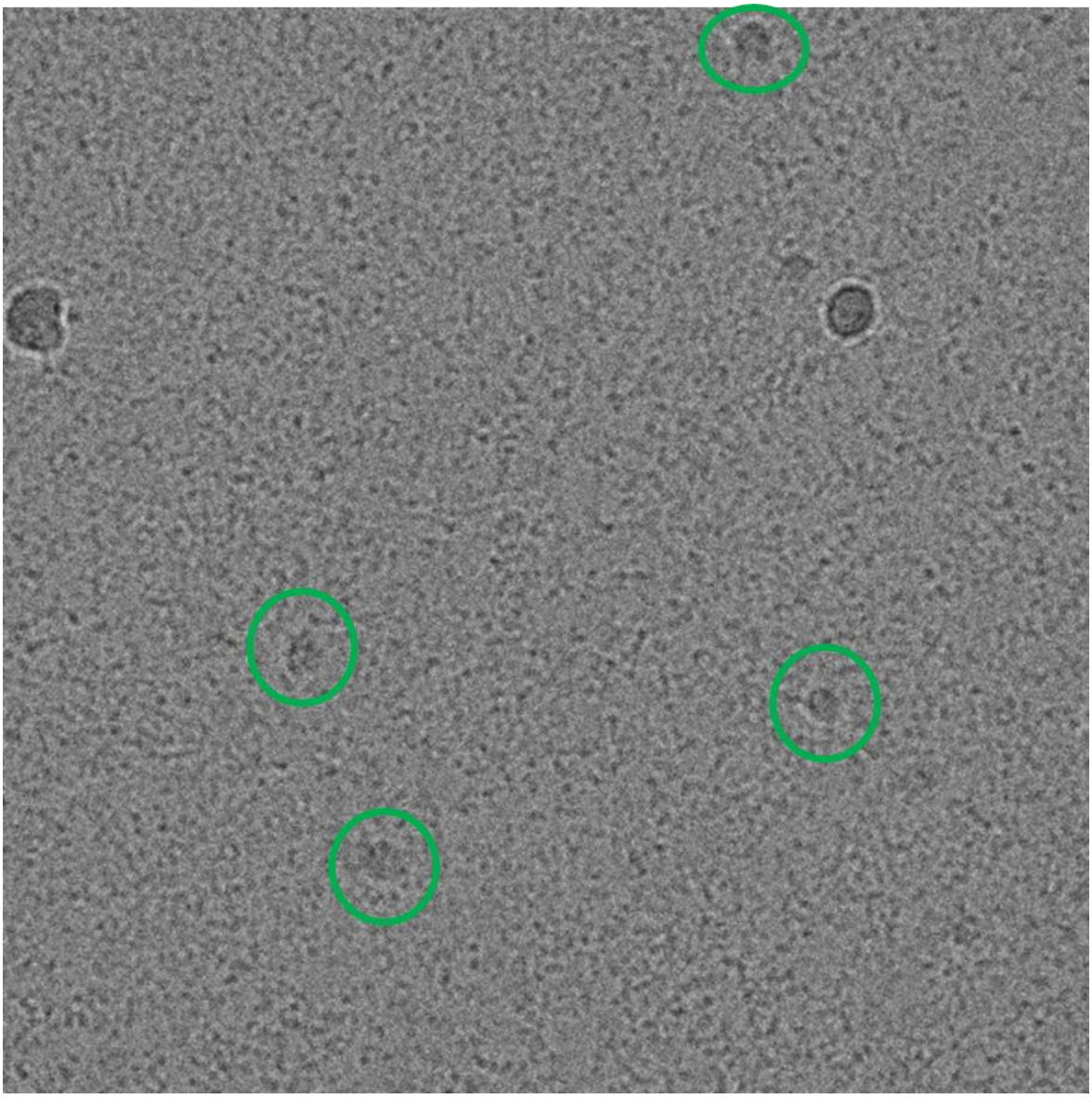


Figure 265. Screening results from first trial. Samples were flash frozen on an appropriate grid upon arrival and imaged using a Talos Arctica cryo-electron microscope (Thermo Fisher Scientific) operated at 200 kV. Larger particles (circled in green) were observed as well as a high background of smaller particles.

Appendix B.2 Screening Trial 2

Three samples were prepared using 1:1, 2:1 and 3:1 ratios of RRM123 and WNV75 with the concentrations and buffer conditions previously described. An additional two samples were prepared at the 3:1 ratio and applied to a centrifugal concentrator (EMD Millipore) with a molecular weight cutoff of approximately 50 kDa. A measurable absorbance at 260 nM was observed for both the concentrator flowthrough and sample retained in the upper chamber indicating the filtration of free RNA from the sample. Following concentration, one sample was diluted with sample buffer to 2.0 mg/mL. The remaining sample was left at a concentration of 2.5 mg/mL. All samples were analyzed by NMR to confirm complex formation through the chemical shift of the RRM123 methyl groups. Liquid samples were shipped overnight on ice. Screening results indicate the presence of 1-3 complex particles per grid at the 3:1 ratio. However, a high background of free RNA remained but at a reduced level from the previous screening results was observed (Figure 27). These conditions are not adequate for further structure work.

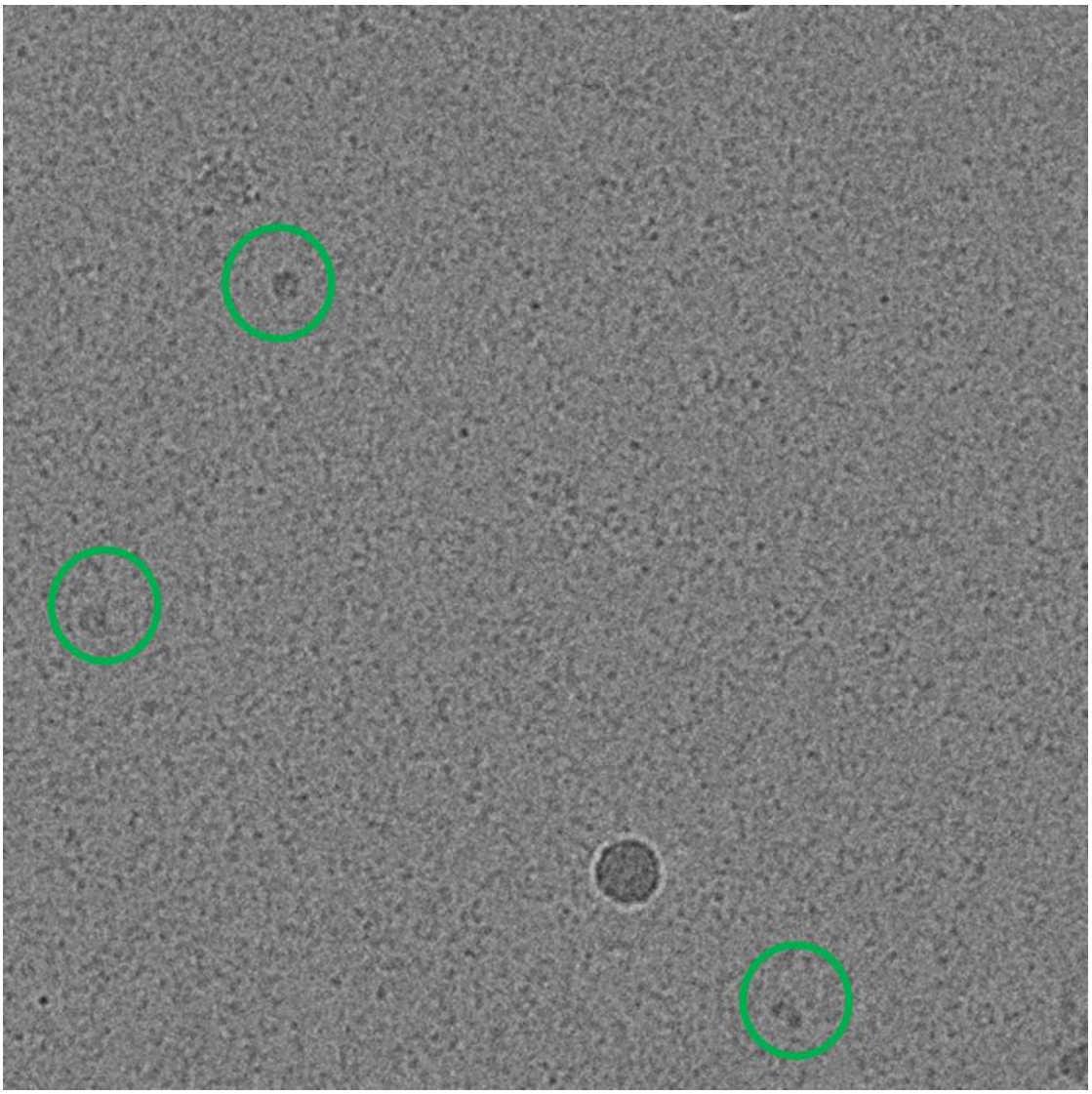


Figure 276. Screening results from second trial. Samples were flash frozen on an appropriate grid upon arrival and imaged using a Talos Arctica cryo-electron microscope (Thermo Fisher Scientific) operated at 200 kV. Larger particles (circled in green) were observed again; however a high background of smaller particles persisted.

Appendix B.3 Screening Trial 3

A 3:1 ratio of RRM123 and WNV75 was prepared using 180 μM RRM123 and 60 μM WNV75 in 10 mM PBS, 50 mM NaCl, 1 mM EDTA, 1mM DTT, pH 6.8. The sample was analyzed by NMR to confirm complex formation through the chemical shift of the RRM123 methyl groups. A portion of the sample was applied to a Superdex 200 (GE Healthcare) size exclusion column and 0.5 mL fractions were collected using an AKTA Pure (GE Healthcare) FPLC. The remaining 3:1 was aliquoted into one 300 μL sample and three 100 μL samples and UV crosslinked using a Stratagene UV Crosslinker 1800 by applying a flux of 0.25 J/cm^2. A portion of the 300 μL aliquot and a portion of a 100 μL aliquot were reserved as cryo-EM samples. The remaining crosslinked samples were applied to a Superdex 200 size exclusion column and 0.5 mL fractions were collected. An enrichment of complex was observable by the chromatogram graph of absorbance at 260 nM for the UV-crosslinked samples in comparison to the native sample (Figure 28). In total, five cryo-EM samples were shipped representing treatment with size exclusion only (1 sample), UV-crosslinking only (2 samples), and UV-crosslinking combined with size exclusion (2 samples). All samples were visualized by native PAGE. Staining with Sybr Gold (Invitrogen) and subsequently Spyro-Ruby Protein Gel Stain (Invitrogen) showed a significant comigration of protein and RNA components in the UV-crosslinked samples (Figure 28 B). Liquid samples were shipped overnight on ice.

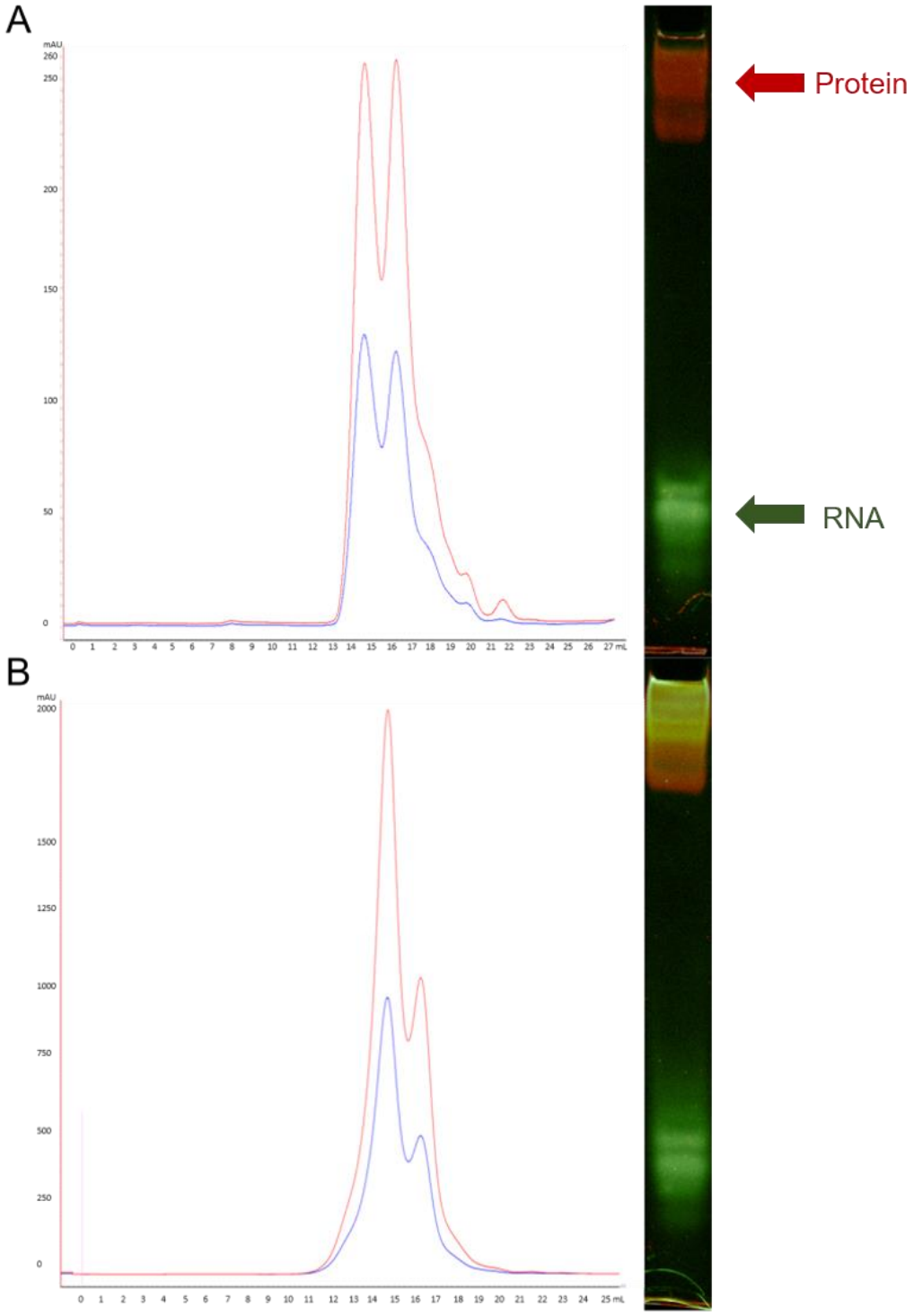


Figure 287. Size exclusion chromatography and native PAGE of samples. The resulting chromatograms and native PAGE for A) a 3:1 RRM123-WNV75 sample and B) a UV-crosslinked 3:1 RRM123-WNV75 sample. UV absorbance was monitored at 260 nm (red trace) and 280 nm (blue trace). The collected fractions from 13.5-14.5 mL post-injection were combined and loaded on an 8% native gel. After running the gel was stained with Sybr Gold (Invitrogen) (green channel) and SYPRO Ruby (Invitrogen) (red channel).