SOLUTION STRUCTURE OF A BENZIMIDAZOLE INHIBITOR IN COMPLEX WITH DOMAIN IIA OF THE HEPATITIS C VIRUS INTERNAL RIBOSOME ENTRY SITE

by

Ryan B. Paulsen

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Medicinal Chemistry

The University of Utah

December 2010

Copyright © Ryan B. Paulsen 2010

All Rights Reserved

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Ryan B. Paulsen**

has been approved by the following supervisory committee members:

Darrell R. Davis	, Chair	**05/26/2010** Date Approved
Thomas E. Cheatham, III	, Member	**05/26/2010** Date Approved
Peter F. Flynn	, Member	**05/26/2010** Date Approved
Eric W. Schmidt	, Member	**05/26/2010** Date Approved
Wesley I. Sundquist	, Member	**09/03/2010** Date Approved

and by **Darrell R. Davis** , Chair of the Department of **Medicinal Chemistry**

and by Charles A. Wight, Dean of The Graduate School.

ABSTRACT

The solution structure of domain IIa of the hepatitis C virus internal ribosome entry site in complex with a racemic benzimidazole inhibitor was determined by NMR spectroscopy with corroborating fluorescence data. A 38 base RNA construct representing the inhibitor-binding region of domain IIa was synthesized by T7 RNA polymerase. Fully and selectively ^{13}C and ^{15}N labeled and isotopically unlabeled RNA samples were produced and studied in complex with the inhibitor.

The inhibitor was previously shown to have inhibitory activity in an HCV replicon assay. It was also previously found to bind in the bulge region of domain IIa. In the free RNA, this five base bulge region introduces a bend in the extended domain II that situates the terminus of the domain over the mRNA cleft in the ribosomal E site. This hinge-like bulge region is not a known binding site for any host or viral translational cofactors, but domain II has been shown to be critical for IRES function, and the bulge-induced bend in domain IIa has been shown to be important for IRES function in mutagenesis assays.

Molecular dynamics refinement in explicit solvent and subsequent free energetic analysis indicate that the inhibitor enantiomers bind with comparable affinity and equivalent binding modes. The structure of this inhibitor/RNA complex suggests that the small molecule rearranges the base stacking in the bulge and introduces a significant conformational change that eliminates the bent RNA helical trajectory. This suggests a

possible mechanism of inhibition involving the displacement of the domain II terminus from the mRNA cleft on the ribosomal E site. Perhaps most importantly, this structure may serve as a guide in the development of second-generation higher affinity inhibitors of the hepatitis C IRES as well as provide general insights into small molecule inhibitor interactions with RNA.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF ABBREVIATIONS

A	Adenosine
C	Cytidine
COSY	Correlation spectroscopy
CSFV	Classical swine fever virus
DNA	Deoxyribonucleic acid
eIF	Eukaryotic initiation factor
EM	Electron microscopy
F1	Indirectly detected dimension of an NMR experiment
F2	Directly detected dimension of an NMR experiment
F1eF2e	Direct dimension edit indirect dimension edit
F1fF2e	Direct dimension edit indirect dimension filter
F1fF2f	Direct dimension filter indirect dimension filter
G	Guanosine
HCV	Hepatitis C virus
HMBC	Heteronuclear multiple bond correlation (spectroscopy)
HSQC	Heteronuclear single quantum correlation (spectroscopy)
IPAP	In phase antiphase NMR
IRES	Internal ribosome entry site
LTR	Long terminal repeat

MM-PBSA	Molecular mechanics Poisson-Boltzmann surface area
mRNA	Messenger RNA
MWCO	Molecular weigh cut off
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
rATP	Ribonucleotide adenosine triphosphate
rCTP	Ribonucleotide cytidine triphosphate
RDC	Residual dipolar coupling
rGTP	Ribonucleotide guanosine triphosphate
RNA	Ribonucleic acid
rNMP	Ribonucleotide monophosphate
rNTP	Ribonucleotide triphosphate
ROESY	Rotational nuclear Overhauser effect spectroscopy
rRNA	Ribosomal ribonucleic acid
TOCSY	Total correlation spectroscopy
tRNA	Transfer ribonucleic acid
U	Uridine
UTP	Uridine triphosphate
UTR	Untranslated region

ACKNOWLEDGMENTS

I would like to acknowledge Dr. Darrell R. Davis and his invaluable assistance in optimizing RNA NMR experiments and developing new ones to help me to find detours around the roadblocks I encountered in structure determination. Darrell R. Davis also performed the XPLOR structure refinement work. I would like to thank Dr. Jack Skalicky and Jay Olsen for teaching me about NMR and how to run the instruments. And thanks to Dr. Tom Cheatham for the AMBER structure refinement work.

The people at Isis Pharmaceuticals who did the early work with this compound and provided us with the opportunity to study it were instrumental in this effort. Specifically, Punit Seth, Rich Griffey, and Eric Swayze deserve special thanks.

This work was supported in part by a grant from the US-Israel Binational Science Foundation to Darrell R. Davis. The National Institutes of Health supported Thomas E. Cheatham (R01-GM081411) and NMR instrumentation. Computer time for the MD simulations came from a NSF TRAC award to Thomas E. Cheatham (MCA01S027) on resources at PSC, TACC, and UTK and also local computer time at the Center for High Performance Computing with computers funded in part by the NIH (S10-RR-17214).

I also acknowledge Elsevier Limited for the use of one figure from previously published work by T. Suzuki et al, and the American Association for the Advancement of Science for the use of one figure from previously published work by J. M. Ogle et al.

CHAPTER 1

INTRODUCTION

Hepatitis C

The World Health Organization estimates that 3% of the world's population is afflicted by the hepatitis C virus (HCV). Roughly 55-85% of infected patients develop chronic infections. Of those, 5-20% will develop cirrhosis and 1-2% will develop liver cancer. The virus is primarily spread through exposure to infected blood, but has been known to spread through sexual contact as well, especially in immune compromised individuals.[1]

Current treatment of hepatitis C consists of injection of interferon or pegylated interferon along with oral administration of ribavirin. There are serious side effects. This course of treatment can take several months and has mixed results depending on such factors as viral genotype and patient ethnicity. Average success rates are slightly greater than 50% as measured by sustained virological response (absence of detectable HCV RNA in serum by the end of treatment cycle).[1] The inefficacy of current therapies in treating some patients coupled with the general lack of broadly effective drugs for HCV highlight the need for new and better therapeutic alternatives.[2, 3]

Hepatitis C Virus

Hepatitis C is a positive-stranded RNA virus and a member of the hepacivirus genus of the Flaviviridiae family. The mode of viral entry into the cell is still not fully understood, but it is widely accepted that the host cell receptor proteins CD81 and SR-BI, and the tight junction protein claudin-1 are involved.[4-6] Having gained entry into the hepatocyte, the positive sense RNA genome can be released and translated into an approximately 3000 amino acid polyprotein which is subsequently cleaved and processed by host signal peptidases into four known structural (Core, E1, E2, p7) and six known non-structural (NS2, NS3, NS4A, NS4B, NS5A, and NS5B) proteins.[7] HCV employs an internal ribosome entry site (IRES) in the 5' untranslated region (UTR) of its genome to initiate ribosomal translation in a manner distinct from that of endogenous mRNA. HCV utilizes the viral RNA dependent RNA polymerase NS5B in conjunction with host factors to produce negative sense copies of its genome. These RNAs then serve as templates to produce a greater number of positive sense genomic copies.[7]

Viral assembly is still poorly understood, but is known to ultimately result in an immature virion consisting of encapsidated positive sense RNA in the cytoplasm. The immature virion acquires an envelope via budding through an intracellular membrane and is believed to be released from the cell through the secretory pathway to initiate another round of infection.[8] An overview of the HCV lifecycle can be seen in Figure 1.

Internal Ribosome Entry Site

The internal ribosome entry site is the RNA structure that facilitates synthesis of the HCV polyprotein. The IRES consists of 330 nucleotides in the 5' UTR of the

approximately 9600-nucleotide viral genome, and it folds into a complex three-dimensional structure with four characterized domains. With a minimal subset of endogenous transcription cofactors, the IRES utilizes the host ribosomal machinery to produce the viral polyprotein. In conventional eukaryotic translation, the 5' ^{7}meG-cap structure of the host mRNA assembles with eukaryotic initiation factor 4F (eIF4F). This eIF4F complex is a heterotrimer consisting of eIF4E (the cap-binding subunit), eIF4A (an RNA-dependent ATPase/ATP-dependent RNA helicase), and eIF4G (a high molecular weight scaffold for eIF4E and eIF4A binding). The eIF4F complex then serves as a scaffold for the recruitment of the 40S subunit, eIF1, eIF1a, the ternary complex (eIF2-Met-tRNA$_i^{Met}$-GTP), eIF3 and eIF4B. The eIF1 and eIF1A factors are thought to synergistically facilitate the proper formation of the 48S complex at the initiation codon, but they are not critical for simple complex formation.[9] The eIF4B subunit is thought to facilitate the association of mRNA with the 40S subunit.[10-12] This assembly then utilizes ATP hydrolysis to scan the 5' UTR of the mRNA for the AUG start codon to form the transient 48S complex. GTP hydrolysis in the ternary complex then triggers the release of initiator tRNA and the dissociation of eIF2 from the complex. This is followed by another GTP hydrolysis event involving eIF5B that recruits the 60S subunit while also releasing eIF3 to form the 80S complex, which is then followed by peptide bond formation.[13, 14]

In contrast, IRES mediated translation initiation involves the binding of the IRES to the 40S subunit in the absence of several canonical eukaryotic initiation factors. This complex then recruits eIF3 and the ternary complex to form a 48S-like complex that, like the normal 48S subunit, positions the AUG start codon of the IRES at the proper position

on the ribosome for initiation of translation, but without ATP hydrolysis dependent scanning. Interestingly, the IRES is capable of specifically binding eIF3 independently of the other initiation factors.[15] After the assembly of the 48S-like complex, translation occurs more or less conventionally. As in host cap-dependent translation initiation, GTP hydrolysis facilitates the release of the initiator tRNA and the dissociation of eIF2 from the complex. Also as in conventional translation, a second GTP hydrolysis event facilitated by eIF5B[16] then releases eIF3 and concurrently recruits the 60S subunit, and viral polyprotein translation commences. The eIF4B appears to play a similar role in both translational paths.[17] Thus, structural motifs in the IRES and/or interactions between the IRES and the ribosome seemingly obviate the requirement for eIF1, eIF1a, eIF4F and 5'-^{7}meG capping in ribosomal translation.[18-20] The differences between conventional eukaryotic translation and IRES translation are summarized in Table 1.

Four domains (I-IV) of the IRES have been identified[21] and characterized for activity.[15] The complete secondary structure of the HCV IRES is shown in Figure 2. Domains II-IV have been shown to interact with the 40S subunit, and have been shown to be necessary for translation.[15] Domains III and IV drive binding to the 40S subunit, and are independently capable of binding to the 40S subunit with a K_D approaching that of the complete IRES.[22]

Mutations of the apical loops of domain III negatively impact recruitment of eIF3 and, subsequently, the ternary complex. This is consistent with cyro-EM studies that suggest extensive interactions between domain III and eIF3.[23] Mutational studies also suggest that the basal region of domain III is responsible for tight binding to the 40S subunit.[24, 25]

Domain II is an independently folded extended duplex. The terminal half of the extended duplex is often referred to as domain IIb while the lower half is known as domain IIa. Domain IIa possesses a 5 base bulge that introduces a significant bend in the structure.[26-28] Mutational studies have shown than the deletion of this bulge region destroys IRES activity in vivo.[29-31] A number of potential interactions between domain II and the ribosome have been reported. The apical loop of domain II sits over the E site and closes off the mRNA cleft.[31] The apical loop is believed to induce changes in ribosome conformation that are critical for translation,[19] and it has been suggested that the loop may interact with the ribosomal protein S5 (rpS5) and thus impact positioning of the IRES on the 40S subunit during translation initiation.[32] Other cross-linking studies have suggested interactions of the domain II apical loop with proteins S14 and S16 on the 40S ribosomal subunit.[18] The secondary structure and three-dimensional solution structure of the lower domain IIa region has been determined by Lukavsky et al[27](Figure 3).

Deletion of domain II does not significantly impact 40S subunit binding affinity [20, 22, 25] or recruitment of the ternary complex and eIF3[20]; however, deletion does reduce translation by five-fold by impeding 80S formation.[15, 20, 24, 33] The Lukavsky group has suggested that this phenomenon is due to the fact that this domain (specifically the apical loop and the bulge region in domain IIa) is critical for eIF5 induced GTP hydrolysis and the subsequent release of eIF2 and GDP during the joining of the 60S ribosomal subunit to form the full 80S ribosome. Deletion of the entire domain arrested ribosome assembly at the 48S stage as determined by sucrose gradient experiments,[20] and the deletion also appeared to diminish 48S formation. Three mutants involving deletion of the five base

bulge in domain IIa, the apical loop of domain II, and the loop E motif (a G bulge and proximal mismatches in domain IIb that result in an S shaped bend in the backbone, a common RNA structural motif) all inhibited 80S formation like that of the full domain II deletion mutant.[30]

Spahn et al performed cyro-EM studies of the IRES RNA in complex with the ribosome. They found that domain IIa was bent in complex with the ribosome in much the same manner as has been described for the free IRES.[31] See Figure 4. The HCV IRES IIa bend is conserved in IRES structures from different viruses even though the sequences involved in generating the "bends" are quite different. The classical swine fever virus (CSFV) is known to have an IRES that is similar in structure and function to the IRES of HCV. Interestingly, the apical loop, loop E motif, and a similar global bend are all present in the IRES of the classical swine fever virus. The CSFV domain II structure has been shown to have a bend of 60.4± 9.1˚ compared to the HCV domain II with a 84.5 ± 9.6˚ bend.[20, 27] See Figure 5. This region in both viral genomes was shown to contain specific binding sites for divalent metals that provide stabilization of this motif.[26] Pestova et al also concluded that, like the HCV IRES domain II, the CSFV IRES domain II is critical for inducing the conformation changes in the 48S complex that are necessary for eIF5-induced hydrolysis of the GTP in the ternary complex. This step is necessary for disassembly of the 48S complex prior to 60S subunit recruitment.[34]

Targeting RNA with Small Molecules

RNA has the unique ability to hold both genetic information in its base sequence and to fold into complex three-dimensional structures capable of catalysis or involvement in regulatory binding interactions with proteins, small molecules or other RNAs. Thus,

RNA can be a therapeutic target in two different ways. One approach is to attack RNA as a bearer of protein sequence information. Antisense and siRNA drugs target RNA sequence as the basis for attack, and the therapeutic goal is the regulation of gene expression.

The folding of RNA into complex three-dimensional structures with catalytic or regulatory functions presents the possibility of treating RNA targets in much the same way as protein targets. That is, a key binding or catalytic domain in the RNA can be targeted with inhibitory small molecules. The many unique structural motifs presented by the complex and diverse folding patterns of RNA can potentially afford both specificity and tight binding for small molecules. See Figure 6. The various bulges, loops, and junctions seen in RNA enable potential drugs to access both the minor groove and the typically less accessible major groove. The therapeutic goal of this approach is to inhibit catalysis of a reaction or prevent some recognition or binding event from occurring.

In practice, the binding of many RNA targeted drugs is thermodynamically driven by electrostatic interactions.[35] Intercalating drugs also target nucleic acids (typically DNA) by interacting extensively with the bases and thus have a strong dipole-induced dipole packing component to their target-binding affinity. In contrast, the binding of most protein-targeted drugs is thermodynamically driven by hydrophobic interactions between rigid, largely hydrophobic drugs and well defined hydrophobic binding sites on the protein.

Typically, proteins can offer a broader variety of chemical environments due to the much larger list of conventionally encoded amino acids (20) compared to the four

nucleotides that constitute RNA. The extensive post-translational modifications seen in proteins provide still further chemical variation. Post-transcriptional modifications in RNA provide some additional binding site diversity, but these are neither widespread nor especially diverse outside of tRNA. Protein binding sites are also considered to be more static and well defined than the more flexible RNA targets.[36] The study of RNA binding drugs is a developing field, and higher affinities and specificities may be achieved in the near future.

Established RNA Targets

The Ribosome

The bacterial ribosome is the target of many RNA binding small molecule drugs. The large size of the ribosome, with its numerous functional sites and areas of interaction with tRNA, mRNA, and nascent peptides gives rise to a diversity of binding environments for small molecule inhibitors. It is not surprising that ribosomal peptide synthesis is susceptible to such a broad variety of inhibitors operating by diverse mechanisms. Many of the most successful, well-known, and widely used RNA binding therapeutics fall into this category.

Aminoglycosides represent an important subcategory of ribosome-targeting drugs that interact with the 16S rRNA of the 30S subunit in the bacterial ribosome. This class of drugs has been widely studied and is characterized by a 2-deoxystreptamine (2-DOS) core that drives binding.[37] See Figure 7. They effectively target prokaryotic ribosomal activity, but they often have low target affinity and specificity. Due to their many basic amine groups, they are cationic at physiological pH, which is in some measure

responsible for their antimicrobial activity,[38] but it can also make them very promiscuous binders to negatively charged RNA and other negatively charged cellular targets.

The mechanism of action of paromomycin**,** a representative aminoglycoside, has been studied extensively. Early NMR studies of paromomycin in complex with a model A site of the 16S rRNA suggested that the compound binds in the major groove of helix 44.[39, 40] These studies showed that Ring 1 of paromomycin inserts into the RNA helix and facilitates the flipping out of bases A1492 and A1493. It also stacks on G1491 and hydrogen bonds to A1408. Later crystal structures by the Ramakrishnan group corroborated the placement of the drug and yielded further insights into the specific mechanism of inhibition.[41, 42]

The flipping out of bases A1492 and 1493 from the interior of helix 44 is key to the mechanism of inhibition. Mutagenesis studies have revealed that ribosomal bases A1492 and A1493 are critical for A-site tRNA binding.[43] Analogously to the paromomycin bound 30S structures, the crystal structure of the 30S subunit in the presence of a codon and cognate tRNA anticodon stem loop similarly shows A1492 and A1493 flipped out of helix 44. However, in this context the two bases are involved in direct hydrogen bond contacts with the first two codon positions of the codon/anticodon interaction via the minor groove. The placement of the two adenines in the minor groove is dependent on Watson-Crick base pairing of the first two codon bases. Non-Watson-Crick base pairs in these positions disrupt the geometry and hydrogen bonding of the base triple while still keeping the bases in close proximity. This close proximity disallows the solvation of the exposed polar groups on the bases resulting in a large desolvation penalty

for near-cognate tRNA binding. In contrast, cognate tRNA binding is thermodynamically favorable and results in a domain closure in the 30S subunit.[42, 44]

Paromomycin fixes the flexible A1492 and A1493 nucleotides in a conformation like that induced by cognate tRNA, and thus decreases their entropy and subtly alters other structural aspects of the A site. The result is that the energetic difference between cognate tRNA binding and near-cognate tRNA binding becomes much smaller. This in turn facilitates closure of the 30S subunit in the presence of a near-cognate tRNA, which is typically disallowed in the absence of paromomycin. This results in error prone translation.[44] See Figure 8.

Later crystal structures involving paromomycin, streptomycin and spectinomycin revealed further insights into general binding characteristics of this drug group as well as variations between different aminoglycosides and similar compound classes. All three were found to interact with the 30S subunit. Spectinomycin, an aminocyclitol antibiotic, inhibits translocation of the peptidyl-tRNA from the A-site to the P-site, while streptomycin, like paromomycin, disrupts proofreading and leads to error prone protein synthesis. All three bind specifically through largely polar and electrostatic interactions.[41]

Macrolides, such as erythromycin, are another category of antibiotics that target the ribosome. This category of compounds typically binds in the polypeptide exit tunnel next to the peptidyl transferase center in the 50S subunit. They are thought to inhibit translation by impeding the exit of nascent polypeptides from the ribosome.[45, 46]

Chloramphenicol is a nonaminoglycosidic, nonmacrolide antibiotic known to disrupt peptidyl transferase activity. The crystal structure of the chloramphenicol bound to the large ribosomal subunit reveals that it binds in a hydrophobic crevice at the entrance to

the peptide exit tunnel, perhaps acting analogously to the macrolide antibiotics.[46, 47] Puromycin is a 3'-aminoacyl tRNA mimetic that halts peptide chain elongation as it becomes incorporated into the nascent chain. [47]

Hansen et al have shown that many non-aminoglycosidic, non-macrolide antibiotics (specifically Anisomycin, Virginiamycin M, Sparsomycin, and Blasticidin S) bind to the 50S subunit in and around the peptidyl-tRNA or aminoacyl-tRNA binding sites and thus competitively inhibit peptide bond formation. [47] This leads to amino acid misincorporation.

Tetracycline, a nonaminoglycosidic antibiotic, also attacks translation by promiscuously binding in the A site of the 30S subunit to disrupt aminoacylated tRNA binding.[48, 49] Pactamycin, another nonaminoglycosidic RNA binding drug, also inhibits ribosomal translation through binding in the E site of the 30S subunit. It is thought to adopt a form on the ribosome in which two aromatic rings in the compound fold over one another to mimic a dinucleotide. In this form, it displaces mRNA from its position in the E site.[49] Viomycin, another nonaminoglycosidic drug, is known to inhibit translation by keeping the peptidyl-tRNA stuck in the A-site. A survey of representative ribosome binding drugs is shown in Figure 9.

Tat/TAR

In addition to the ribosome, there are many unique RNA targets in pathogens that can possibly be exploited with specificity due to the absence or extreme dissimilarity of analogous targets in humans. RNA viruses and retroviruses offer especially compelling possibilities for such a targeting approach due to the unique and critical roles that RNA plays in their lifecycles.

The TAR element of HIV-1 is a highly promising and extensively studied RNA therapeutic target. It is known to interact with the HIV-1 Tat protein, the transactivator of the HIV-1 LTR. While there is currently no available therapeutic targeting this RNA, this is an area of active research.[50, 51]

Internal Ribosome Entry Site

Another interesting pathogenic RNA structural drug target is the IRES of picornaviruses and flaviviruses. The IRES of the hepatitis C virus in particular has been studied extensively. As described previously, this structure facilitates the utilization of host ribosomal machinery to translate viral proteins. The unique role of this structure in the viral lifecycle makes it an appealing therapeutic target. This complex three-dimensional structure with its multiple points of interaction with the human ribosome offers a number of potential sites for inhibitory attack.[2, 52]

There are currently no IRES targeted drugs in clinical use. There have been studies on a number of potential HCV therapeutics including siRNAs[53], and small molecules.[54] The mechanism of action of these small molecule inhibitors in the work of Seth et al.[54] will be discussed extensively in this dissertation.

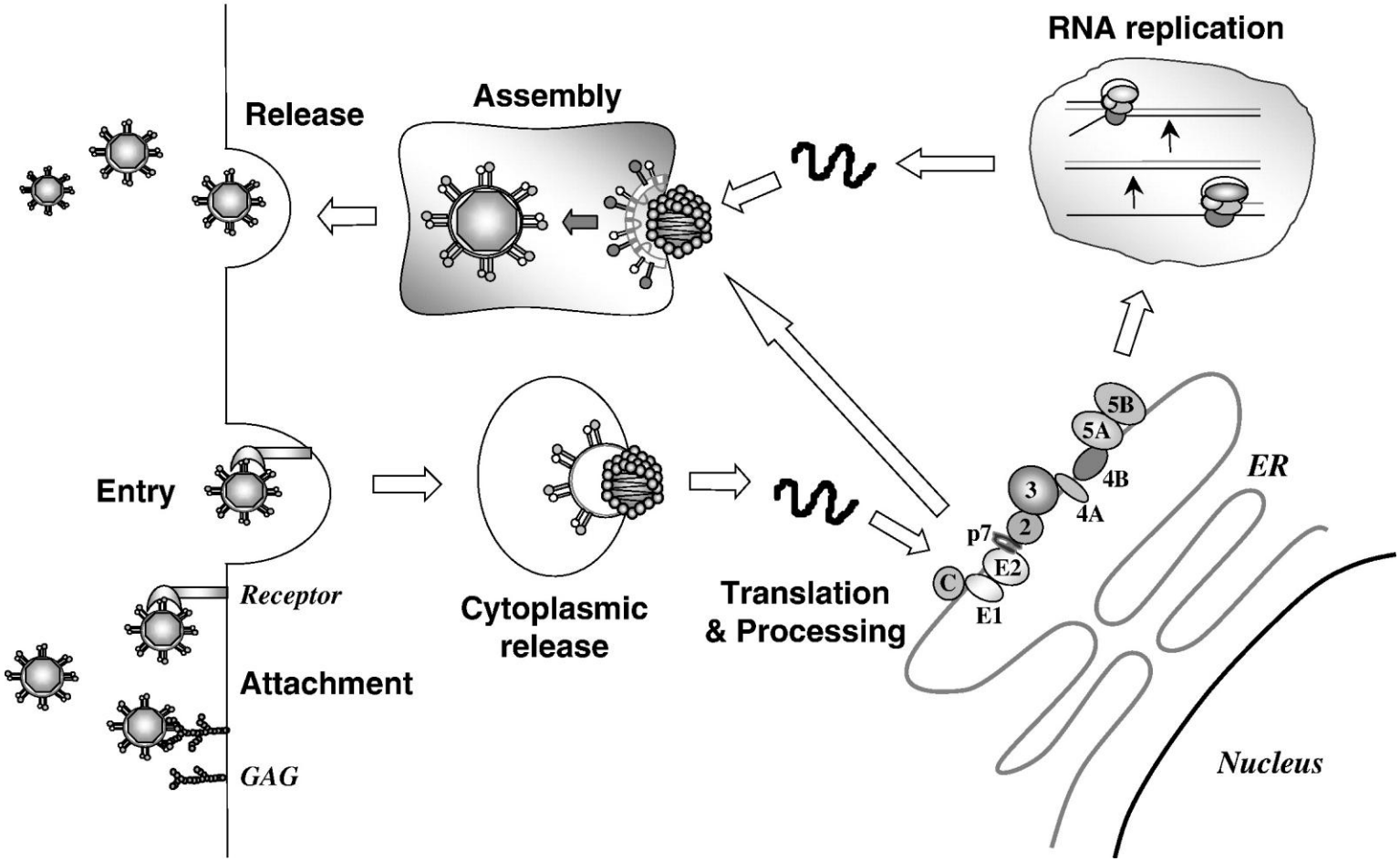

Figure 1: HCV life cycle. Reprinted with permission from Suzuki et al (2007, Elsevier B.V.)

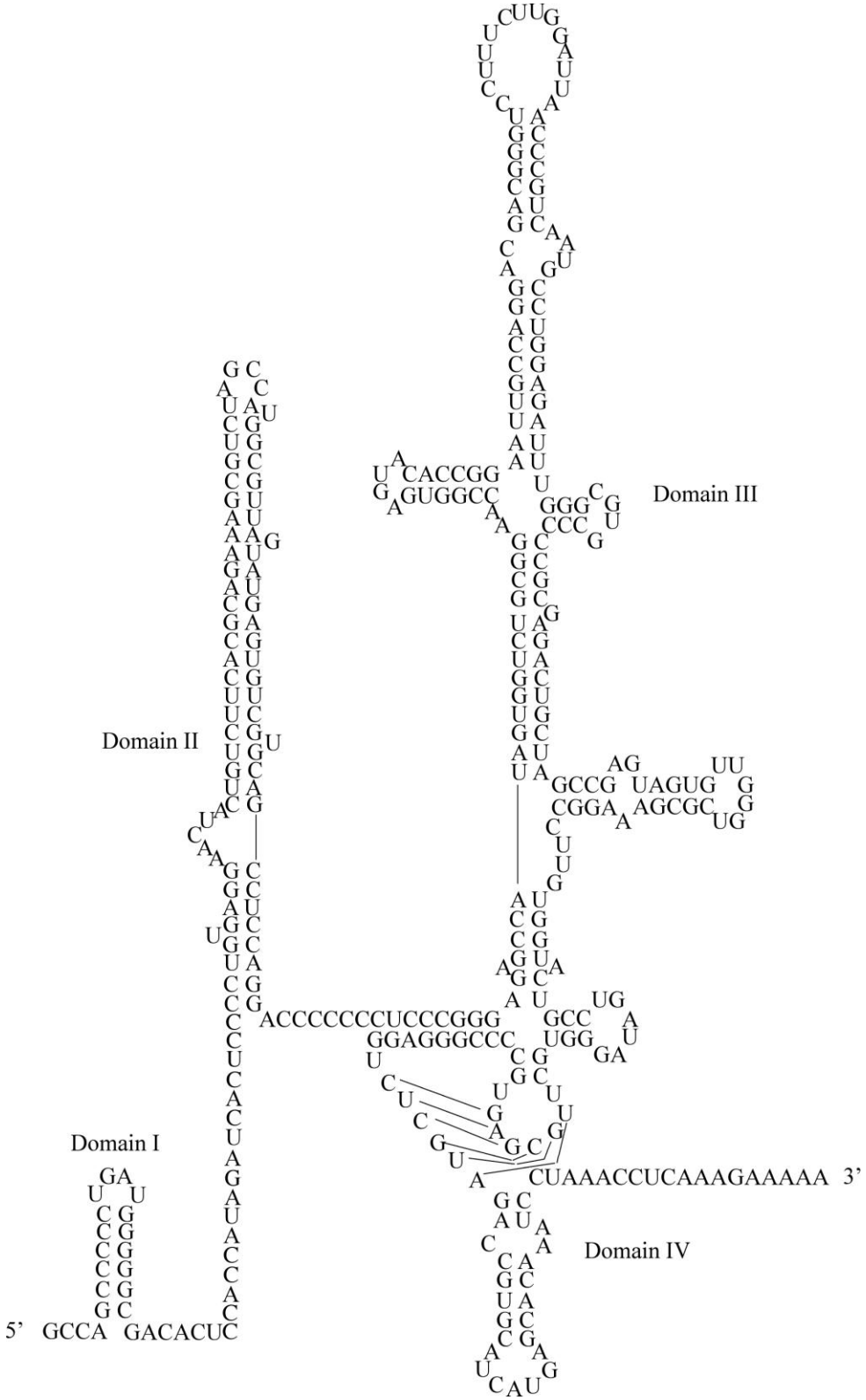

Figure 2: Full HCV IRES secondary structure. Domains I-IV are labeled. This figure is adapted from Fraser et al. (2007).

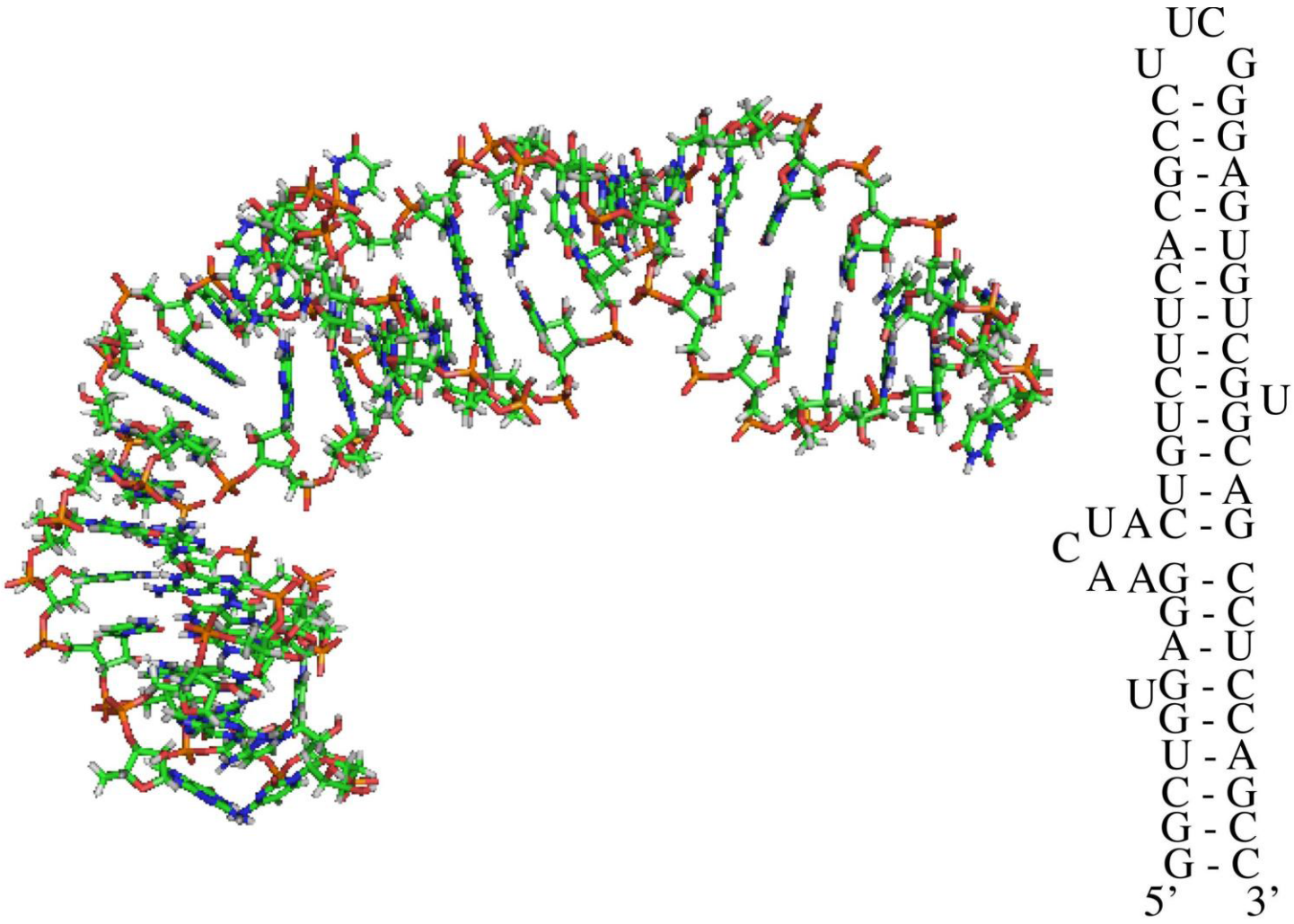

Figure 3: The three-dimensional and secondary structures of domain IIa. The 84.5º bend in the structure occurs at a five base bulge that can be easily seen in the secondary structure.

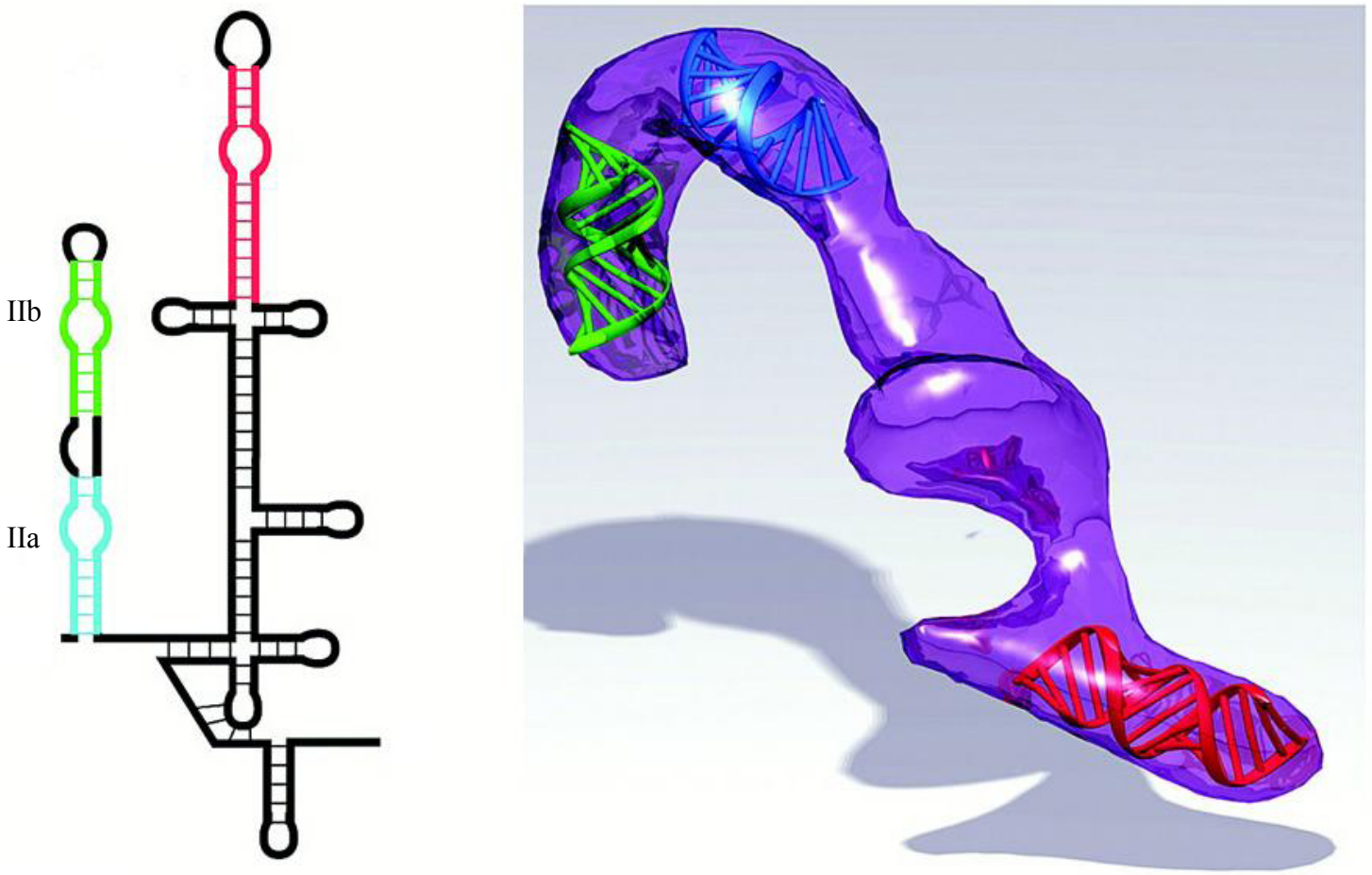

Figure 4: Cryo-EM structure of the HCV IRES bound to the ribosome. The ribosome is not shown. Adapted from Spahn et al (2001). Representation of a subset of A-form helices from the HCV IRES bound to the 40S subunit fit into the IRES surface generated through cryo-EM studies. The green helix corresponds roughly to domain IIb while the blue helix corresponds roughly to domain IIa. ~The 84.5° bend in domain II is clearly visible.

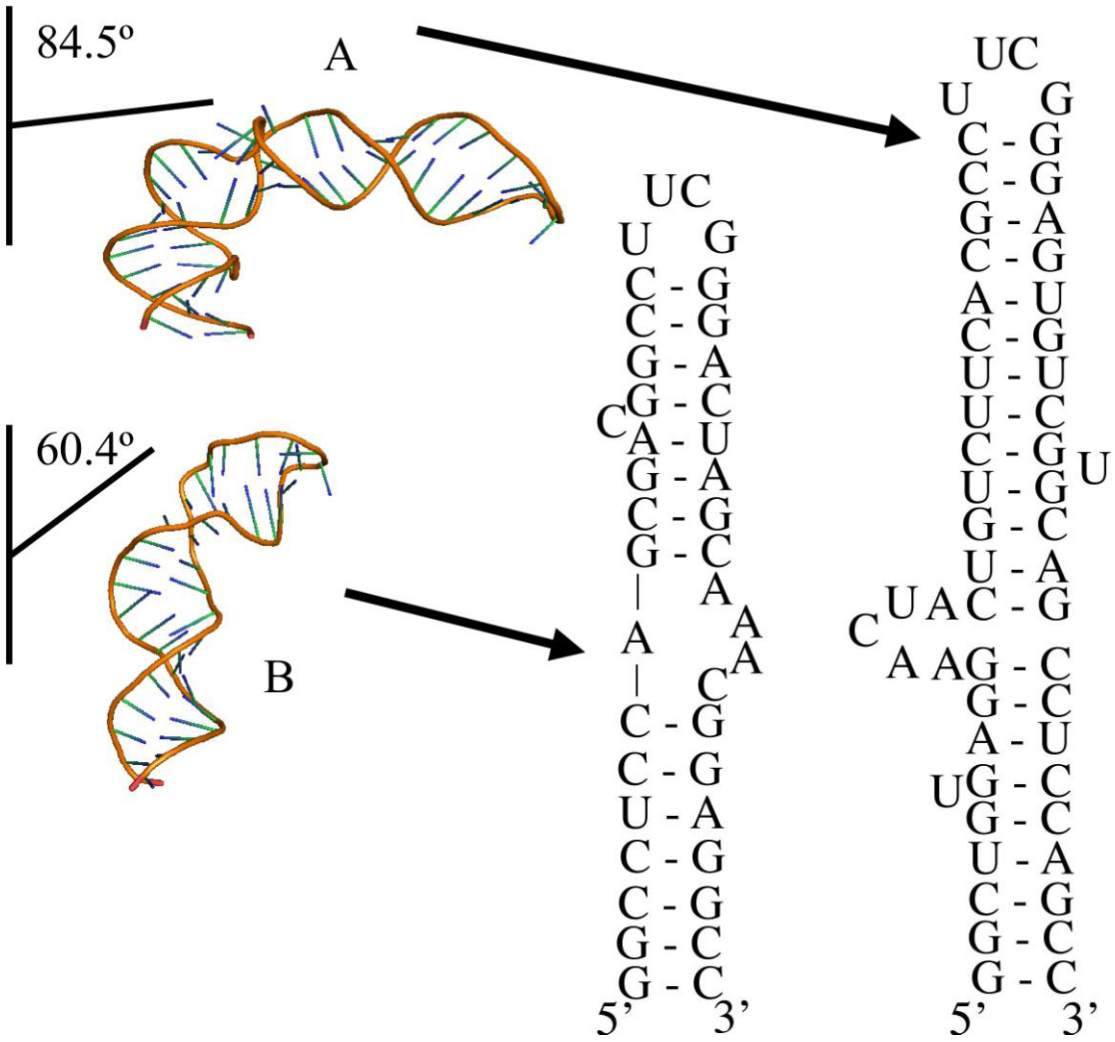

Figure 5: Comparison of HCV and CSFV IRES structures. Representations of the three dimensional and secondary structures of the IRES domain IIa constructs from the hepatitis C virus (A) and the classical swine fever virus (B). Both structures are similarly bent in spite of having divergent sequences in the bulge region.

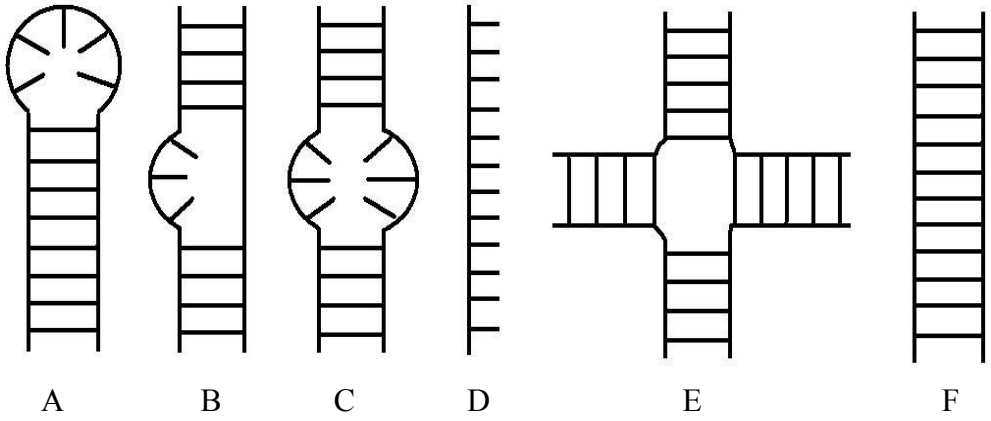

Figure 6: Sample of RNA secondary structural motifs: (A) a stem loop; (B) a bulge; (C) an internal loop; (D) a single strand; (E) a four-way junction; (F) a duplex.

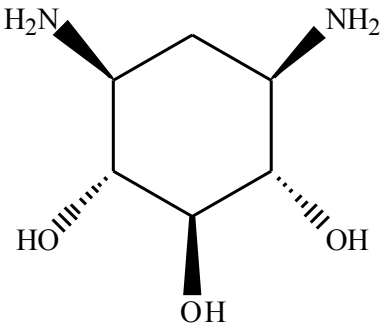

Figure 7: 2-DOS structure. This 2-deoxystreptamine core is common to all true aminoglycoside drugs.

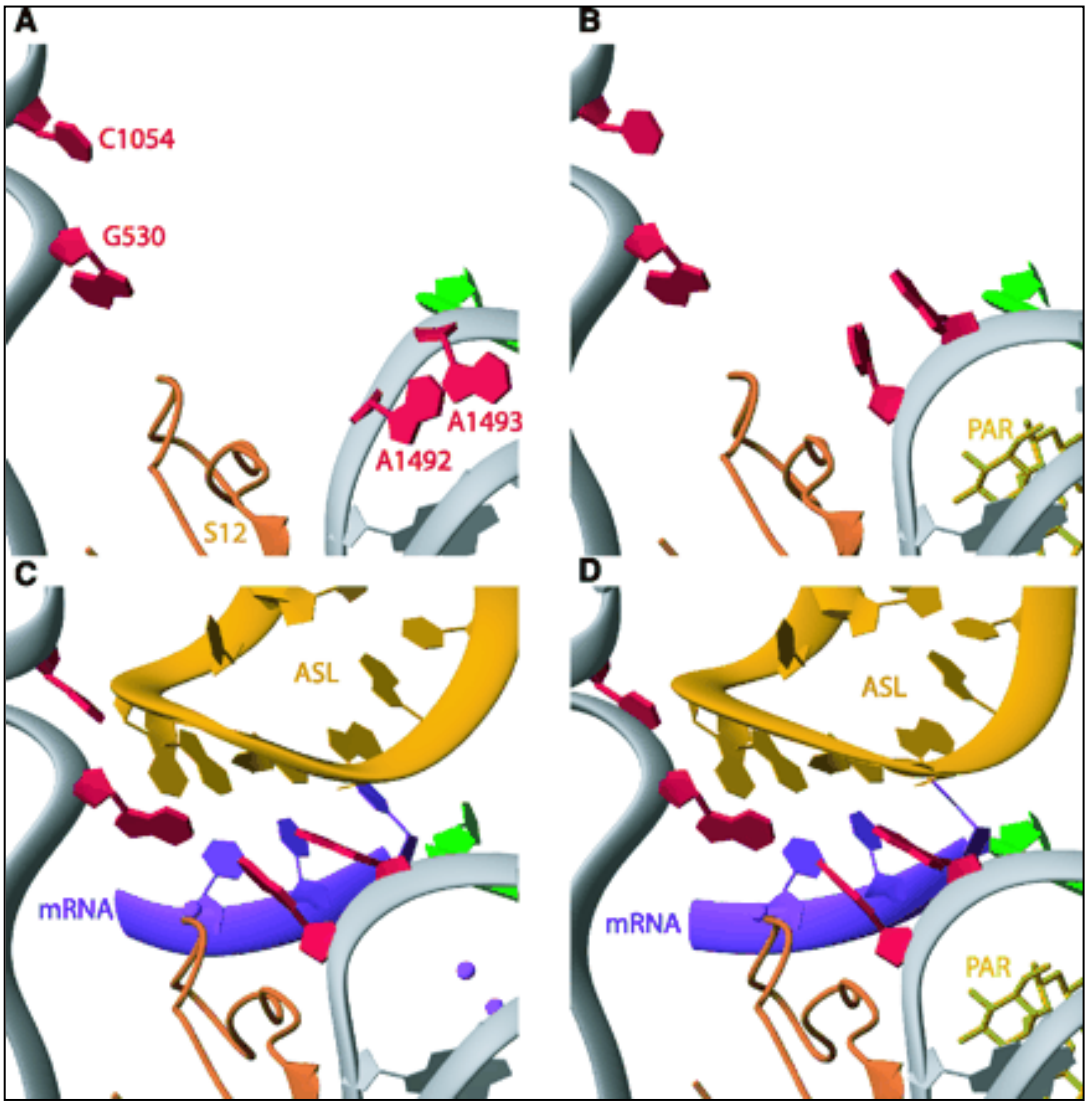

Figure 8: Influence of paromomycin on the A site of the 30S subunit. A) The native 30S subunit with A1492 and A1493 stacked in the interior of helix 44. B) Paromomycin binds in the A site and flips A1492 and A1493 out of helix 44 in a conformation like that shown in C. C) A cognate tRNA ASL paired to an mRNA codon in the A site results in the flipping out of A1492 and A1493 to form base triples with the first two codon positions of the codon/anticodon interaction. The stability and energetic favorability of this interaction is dependent on Watson Crick base pairing of the first two codon positions and thus serves to screen out near-cognate tRNA in the A site. D) The binding of paromomycin to the 30S subunit in the presence of the tRNA-ASL and the codon changes the structure very little compared to that shown in C. Paromomycin stabilizes the interaction of the 30S subunit with near-cognate tRNA. Thus, the energetic difference between the binding of cognate and near-cognate tRNA is minimized, and amino acid misincorporation is facilitated. Reprinted with permission from Ogle et al (2001, AAAS).

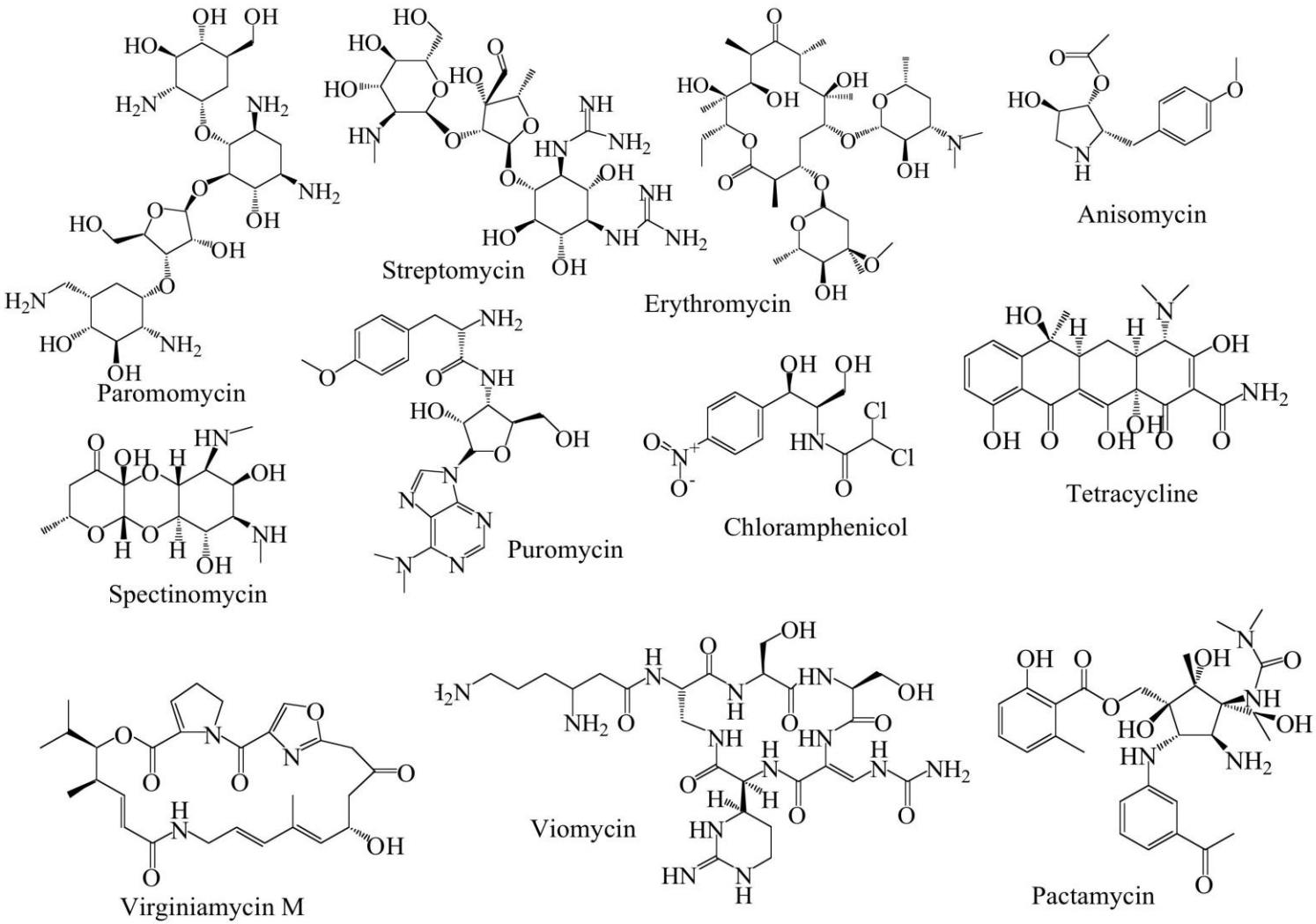

Figure 9: Sample of RNA targeted drugs known to target the ribosome. Paromomycin and streptomycin are true aminoglycosides since they contain the 2-deoxystreptamine core.

Table 1

Required cofactors for conventional ribosomal and IRES translation

Conventional Translation	IRES Translation
5' ^{7}meG cap	-
eIF4F	-
eIF1	-
eIF1a	-
Ternary complex	Ternary Complex
eIF3	eIF3
eIF4B	eIF4B
eIF5B	eIF5B

CHAPTER 2

ISIS-11 AND IRES5

Isis-11

Isis Pharmaceuticals, Inc. graciously provided us with milligram quantities of a lead compound described as compound 11 in the published account, hereafter described as Isis-11 (Figure 10).[54] This compound was developed from a K_D = 100 uM benzimidazole “hit” from a 180,000 member library screened through a mass spectrometry based method. A "hit" was required to have a K_D of l00 µM or less with a 29-mer HCV IRES domain IIa RNA construct, and that K_D had to be three fold lower than that of the "hit" with a 33-mer structured RNA control. Thus, a "hit" had to bind both tightly and specifically.

Employing a mass spectrometry-based method, Isis-11 was determined by Isis to bind to the IRES construct with a K_D of approximately 0.86 µM. Fluorescence data collected in our lab suggested a K_D of approximately 2 µM. The slight difference in K_D numbers could be attributable to the differing experimental approaches used, or to experimental uncertainty.

In a cellular HCV replicon assay, Isis-11 was found to inhibit the production of viral RNA with an EC_{50} of 3.9 µM. The comparable RNA binding constants and *in vivo*

activities for this molecule and others in the same chemical class suggest that the binding of the specific RNA target is the mechanism of viral replication inhibition.

The potential exists for up to three positive charges on the compound. The two tertiary amines are both likely to be charged at the pH 6-7 range, as the pK_as would be 9-10. The primary amine on the ring, by analogy to other known structures, is assumed to have a pK_a of approximately 6-7. Thus, this amine may also be charged under the solution conditions we employed. Modeling of the molecule in the binding site was done under the assumption that all three of these amines are positively charged at the pH of the experimental samples (pH 7).

The positive charges on these amines could lead to non-specific ionic or hydrogen bonding interactions with the negatively charged RNA; however, synthetic variations on Isis-3 (another benzimidazole compound from Isis that also contained the 2-amino group and two tertiary amines on methylene tethers) suggest that the charged groups interact in a specific manner at the binding site.[54] See Figure 11. A synthetic variant of Isis-3 that lacked the 2-amino group (Isis-7) showed a greater than 10-fold affinity decrease from Isis-3. A variant that lacked the methylene arm and tertiary amine at the N1 position (Isis-8) showed a similar decrease. Isis-9, which displaces the other arm without eliminating it, similarly had lower affinity.

Our collaborators at Isis used mass spectrometry based footprinting assays to localize binding of Isis-11 and analogous compounds in specific IRES constructs. Inhibitors were found to interact with bases in and around the adenosine-rich five base bulge which has been shown to be critical to the overall shape of domain II.[54] This compound was shown to bind to constructs that favored both the published IRES domain

II structure [27] and an alternate secondary structure previously proposed by the Doudna group.[55]

It is interesting to note that the compound could bind with similar affinity to different constructs with differing secondary structures. It is still unclear what this means as there is no published structure of any constructs biased to adopt the Doudna form. It is possible that the structure at the locus of binding is similar in either case. The portion of the secondary structure previously proposed by the Doudna group pertinent to the binding site of Isis-11 is approximated by the IRES2 construct shown in Figure 12.

Isis-11 NMR Assignments

In order to determine the structure of Isis-11 in complex with the IRES, it was first necessary to establish intermolecular distance restraints between Isis-11 protons and IRES protons as determined by Nuclear Overhauser Effect interactions. Thus, it was necessary to determine the ^{1}H chemical shifts of the Isis-11 protons. In order to do that, it was also necessary to determine the ^{13}C chemical shifts. The inhibitor proton assignments were established using TOCSY, ROESY, and HMQC/HMBC NMR experiments for a 1 mM Isis-11 sample dissolved in D_2O. The lone chiral proton provided an unambiguous starting point to connect protons on the furan chain by TOCSY connectivity. The methylene TOCSY connectivities for the long chain established those assignments, and the crucial distinction of the two dimethylamino methyl groups was made by ROE connectivity to protons on the relevant methylene chains. The two protons on the aromatic ring were assigned by HMBC connectivity to ring carbons. The Isis-11 chemical shifts are unchanged at 7-fold excess to the IRES target, as expected for a system in fast-dominated exchange. NOESY and ROESY NMR experiments established

that Isis-11 binds in an extended conformation; intramolecular NOEs were negative, consistent with cross-relaxation by the bound form.

The ^{1}H and ^{13}C NMR resonance assignments are shown in Figure 13. All Isis-11 assignment experiments were conducted on a Varian MERCURY 400 MHz and Varian INOVA 500 instruments at the University of Utah, which are well suited to small molecule NMR work.

Isis-11 Quantification

A gravimetric approach was used to quantify Isis-11. The synthesis product as provided to us contained an excess of acetic acid and acetate salts. The amount of acetic acid relative to Isis-11 was estimated based on NMR peak intensities, and the concentration of our solutions adjusted accordingly. There was some variability in compound concentration from sample to sample, but all of the NMR samples possessed Isis-11 at roughly a 7- to 9-fold excess to the IRES construct based on peak intensity comparison to the resolved peaks in the spectra from added spermine. The RNA chemical shifts were very consistent between the different complex samples with excess Isis-11 notwithstanding the slightly varying relative concentrations.

IRES5

The IRES5 RNA construct was developed in order to satisfy several requirements presented by this system. Firstly, it was necessary to incorporate all of what was believed to be the relevant binding region of the IRES domain IIa. Secondly, it was necessary to include a relatively long stem and a stable hairpin tetraloop at each end of the construct in order to limit possible destabilization and unfolding of the RNA upon compound binding

in the bulge. However, the construct should be as small as possible while meeting the other requirements in order to limit spectral complexity and the line broadening that accompanies high molecular weight.

To these ends, smaller constructs (IRES2 and IRES3) were used in early titration experiments. IRES2 appeared to bind well to Isis-11, but it did not assume the secondary structure of the published solution[27] and crystal structures[26] but rather adopted a secondary structure postulated by the Doudna group prior to the publication of the NMR and crystal structures.[55] IRES3 assumed a secondary structure like that of the published account, [27] but predicted free energy calculations in mfold showed it was also capable of assuming the Doudna structure with a low thermodynamic penalty. In practice, IRES3 was destabilized upon compound binding. The IRES4 construct showed promise, but it was abandoned for the slightly smaller IRES5 when it was shown that most of the long upper stem was left unperturbed by Isis-11 binding and thus was probably irrelevant to our study. While it is possible that IRES5 could be further minimized and still form an adequately stable complex with the compound, no further construct minimization was pursued for the determination of this structure. Figure 14 shows the IRES5 construct juxtaposed with the complete domain II construct from the previously published structure.[27]

Sample Preparation

Complex Conditions

Initial work was conducted on a 400 µM IRES5 sample with 25 mM KCl, 25 mM NaCl, and 10 mM sodium phosphate at pH 7.0. These conditions were suitable for the free IRES5 at a range of temperatures (10-30º C). Addition of low, sub-stoichiometric

amounts of Isis-11 (~1/10 RNA concentration) led to line broadening of a small subset of peaks; however, addition of higher amounts of Isis-11 led to heavy, global line broadening of all NMR peaks even at higher temperatures (>30º C). For example, ^{13}CHSQC spectra were uninterpretable due to the lack of any distinct, resolved cross peaks.

We hypothesized that the positive charges on the Isis-11 could lead to non-specific interactions with the negatively charged phosphate backbone of the RNA, especially in low salt conditions. I attempted to correct this problem with spermine. As a positively charged polyamine, spermine could compete out these nonspecific interactions. This approach showed some promise in titrations that were followed by monitoring changes in the imino region of the ^{1}H spectrum; however, spermine addition was not sufficient to correct the line-broadening problem in the ^{13}C HSQC experiments. Ultimately, all of the relevant data were collected in samples with 2 mM spermine based on the observation that spermine appeared to partly compete out weak, non-specific interactions between Isis-11 and the RNA as indicated by the improvement in the imino NMR spectrum. The addition of spermine alone did not cause any significant line broadening or changes in the spectra.

Increasing concentrations of KCl and NaCl were found to significantly diminish the line broadening caused by nonspecific Isis-11 binding. While the ^{13}C HSQC cross peaks were still broader in the complex than in the free RNA, the catastrophic line-broadening problem seen under low salt conditions was mitigated. Ultimately, the data pertinent to the complex structure determination were collected from samples containing both 150 mM NaCl and 150 mM KCl. Under these conditions, the non-specific line

broadening was minimized even at high molar excesses of Isis-11. The efficacy of monovalent salts in competing out the nonspecific compound binding suggests that the electrostatic interaction of the positively charged amine groups of Isis-11 with the negatively charged phosphate backbone of the RNA is the driving force behind the non-specific binding.

The high salt, high spermine conditions employed in these experiments represent a very potent means of competing out nonspecific, electrostatically driven interactions. These conditions could be generally useful in the screening of charged drug candidates against RNA. Indiscriminate binding of positively charged compounds to RNA can create false leads in drug discovery. High salt concentrations can effectively shield the negative charge on the backbone and thus eliminate or greatly reduce nonspecific binding. These conditions also represent a closer approximation of physiological conditions than is the case in a low salt NMR buffer.

Magnesium is known to interact specifically with domain II of the IRES. Free IRES5 had a slightly different imino spectrum in the presence and absence of magnesium. Addition of magnesium caused U12 and G33 to move downfield and upfield respectively. The free IRES5 also showed some differences in the aromatic ^{13}C HSQC spectrum in the presence and absence of Mg^{2+}.

Magnesium was not included in our complex study experiments because the Mg^{2+} appeared to slightly exacerbate the already extant line broadening problems. This may have been due to exchange phenomena caused by magnesium ions competing for the binding site of the compound as both the binding site of the molecule and at least two specific Mg^{2+} binding sites overlap. Alternatively, it could also be due to some degree of

RNA aggregation facilitated by divalent ion binding, or possibly due to intermediate exchange of non-specifically bound Mg^{2+} ions.[56] Regardless of the mechanism of the line broadening, Mg^{2+} was omitted to improve line widths.

While the presence of Mg^{2+} changes the imino spectrum of the free RNA, the imino spectrum of the complex looked essentially the same in the presence and absence of Mg^{2+}. See Figure 15. The fact that the imino spectra look the same in the presence and absence of magnesium lends credence to a model in which the compound interacts with portions of the IRES domain that are key magnesium binding domains and thus disrupts or out-competes magnesium binding. Also, RDC experiments on the free IRES5 under our conditions suggested that our construct adopted essentially the same fold and bend as the published structure that included magnesium. Thus, the changes in the structure of IRES5 in the complex are attributable to Isis-11 and not to the absence of magnesium.

Proper folding of the free IRES5 was sensitive to pH. Three 100 µM IRES5 samples were made in 25 mM KCl, 25 mM NaCl, 10% D_2O and 10 mM Sodium phosphate at pH 7, 6.4, and 5.5. The samples were then annealed in 5 mM Shigemi NMR tubes by placing them in a 150 mL beaker of boiling water which was subsequently allowed to come to room temperature. The samples at pH 7 and 6.4 annealed properly as evidenced by the imino spectra while the pH 5.5 sample misfolded. It is not clear if this is due to duplex formation or some other misfolding or aggregation problem. Nevertheless, due to this manifest pH sensitivity, the data used for structure determination were collected on samples buffered at a pH of 7.0.

The IRES5 construct was poorly behaved at concentrations higher than ~450 μM. At slightly higher concentrations, the imino spectrum peaks begin to broaden and run together. At very high concentrations (>600 μM), the spectrum becomes similar to the dilute sample spectrum, but the upfield tetraloop G peak is missing. This concentration dependent sample behavior is consistent with duplex formation. The NMR experiments conducted in this study were performed on samples with RNA concentrations of 400 μM or less. Of these, some of the more highly concentrated samples had artifact peaks that were absent or weaker in less concentrated samples. These artifact peaks were not shifted or broadened by compound binding. These peaks can be summarized as follows: At higher concentrations, there were two artifact C6-H6 cytosine peaks seen in the aromatic ^{13}C HSQC. There was one strong adenosine C8-H8 and one strong adenosine C2-H2 peak at higher concentrations. There was another very weak adenosine C8-H8 artifact peak as well. There was also one guanosine C8-H8 artifact peak. There were no substantive uridine artifact peaks.

At higher concentrations (800 uM), the G23 C8-H8 cross peak from the ^{13}C HSQC spectrum disappears completely from its usual place location while the rest of the spectrum changed little. If the high concentration sample is truly an extended duplex, then the domain IIa five base bulge region should be identical to that seen in the hairpin, but multiplied by two for each duplex dimer. Thus, the duplex dimer should behave like the monomer hairpin with respect to compound binding, which occurs far away from the tetraloop. Obviously, this situation is not acceptable for some experiments; however, some of this data is usable for looking at changes around the binding site, which should be intact. Some early ^{13}C HSQC titrations experiments on a 800 uM sample will be

referred to in this dissertation as they relate to line broadening at low stoichiometric equivalents of the compound, but all of the other data used for this study were generated from samples of lower concentration to avoid duplex formation.

RNA Production

All RNA constructs were produced by T7 transcription employing ^{13}C and ^{15}N labeled nucleotide triphosphates purchased from either Isotec or Cambridge Isotope Laboratories, or with ^{12}C and ^{14}N labeled nucleotide triphosphates purchased from Sigma Aldrich, Acros Organics, and MP Biomedicals. Single stranded DNA oligonucleotides were purchased from the peptide/oligonucleotide synthesis facility at the University of Utah. These were made with 2' O-Me at the two terminal 5' bases to facilitate the synthesis of homogeneous RNA transcipts.[57, 58] An optimized T7 RNA polymerase promoter sequence was placed upstream of the RNA coding sequence and a complementary DNA primer was annealed to the template DNA in 10 mM Tris-HCl pH 7.5 and 10 mM $MgCl_2$ buffer at 95º C for 1-3 minutes, depending on volume. This solution was then diluted 8 fold with the same annealing buffer listed above for a final concentration of about 1.8 µM template and 2.7 µM primer.

A solution was made with rNTPs at 6.5 mM each, 43 mM $MgCl_2$, 7 % RNAsecure (v/v) from Ambion, and then incubated at 65º C for 10 minutes to activate the RNAsecure, and then cooled for 20 minutes at room temperature as per the manufacturer's instructions. Subsequently, the solutions were incubated at 37º C for 5 minutes before adding the annealed DNA and the T7 polymerase enzyme. The annealed DNA is highly diluted and is added in a large aqueous volume effectively diluting the above listed concentrations by approximately 1.5 fold. The T7 polymerase was added in

excess compared to typical published or supplier recommended concentrations because this marginally improved reaction yields. Typically, the enzyme was added to a final concentration of 0.25 mg/mL by adding 50 μL of 5 mg/mL T7 polymerase per 1 mL reaction. Reactions were allowed to proceed for a minimum of 4 hours. Once substantial precipitation was visible in the tube, the reaction had usually gone to completion with high yields.

The RNA was purified by ion exchange chromatography employing a column packed with Pharmacia (Amersham) Source Q support, eluting with a gradient of $LiClO_4$. Buffer A consisted of 20 mM $LiClO_4$ and 20 mM ammonium acetate with 5% acetonitrile. Buffer B consisted of 600 mM $LiClO_4$ and 20 mM ammonium acetate with 5% acetonitrile. The gradient employed went from 0 to 50% B over 20 minutes. The column was heated to 50º C. The latest eluting large peak is usually the full length RNA, as the ion exchange column binds RNA oligonucleotides of longer length with higher affinity. The full length RNA is usually well resolved from the unincorporated rNTPs and short early termination transcripts. The appropriate fractions were pooled and then dialyzed (Spectrum 1000 MWCO) against water (1 x 2L), 0.5 mM EDTA (1 x 1L), 0.5 M NaCl (1 x 1L), and finally 3 x 2 L water. The purified samples were lyophilized and then dissolved in the appropriate buffer.

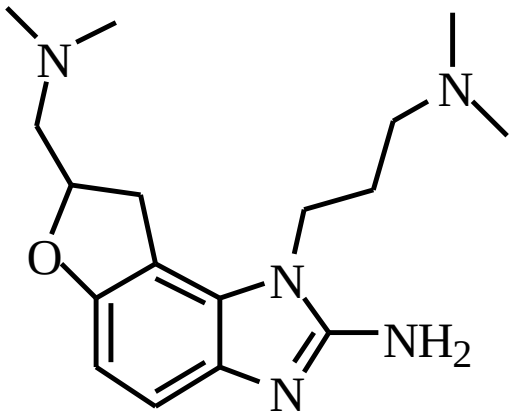

Figure 10: Structure of Isis-11 from Isis Pharmaceuticals.

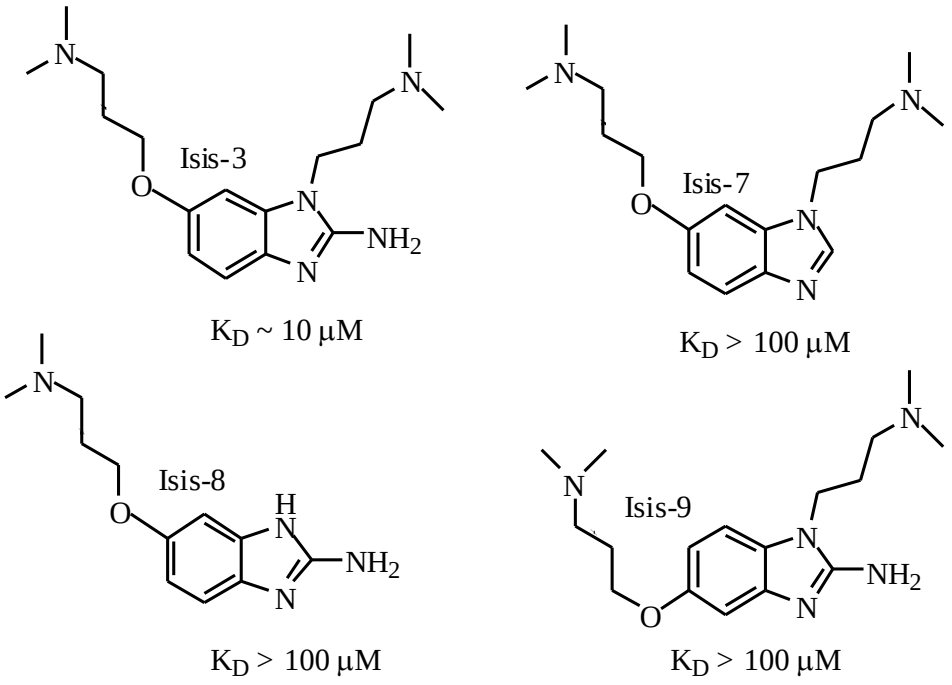

Figure 11: Specificity of benzimidazole binding. Synthetic variations that eliminate or displace key amine groups have a deleterious effect on binding affinity. Adapted from Seth et al (2005).

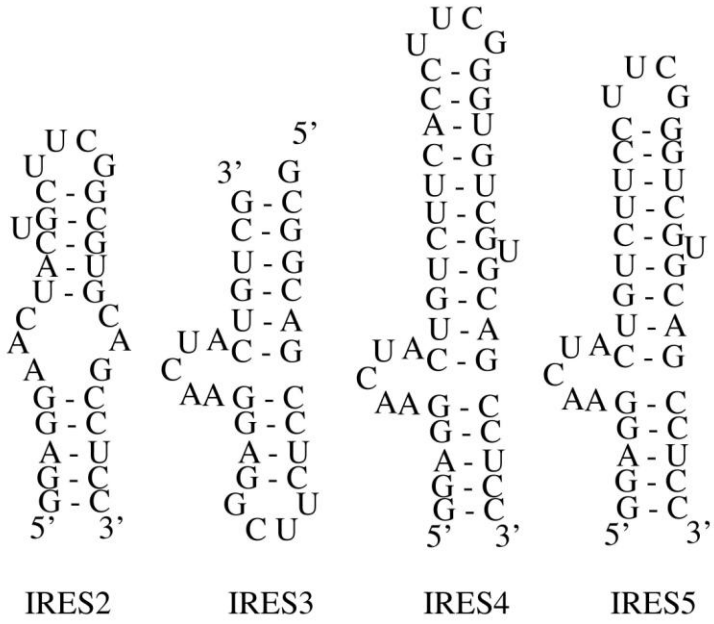

Figure 12: IRES constructs. Various IRES domain IIa constructs used in NMR experiments in the course of these studies.

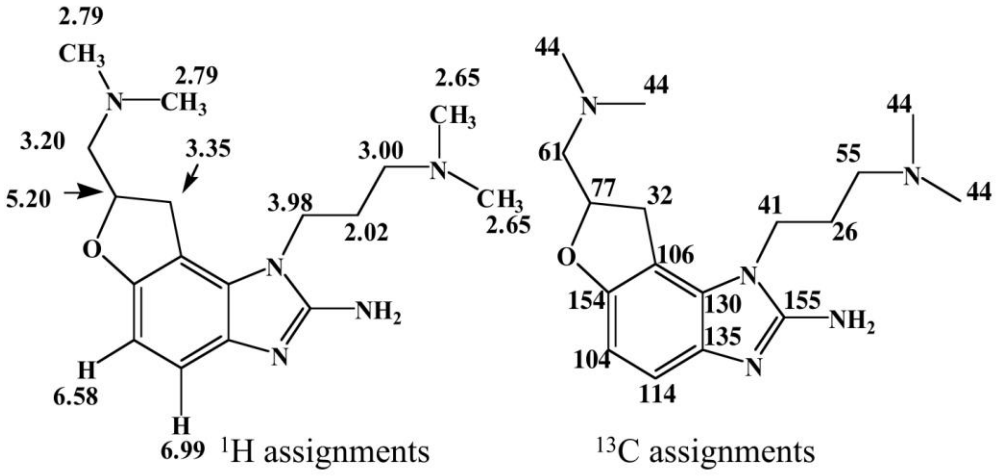

Figure 13: ^{1}H and ^{13}C assignments for the free Isis-11

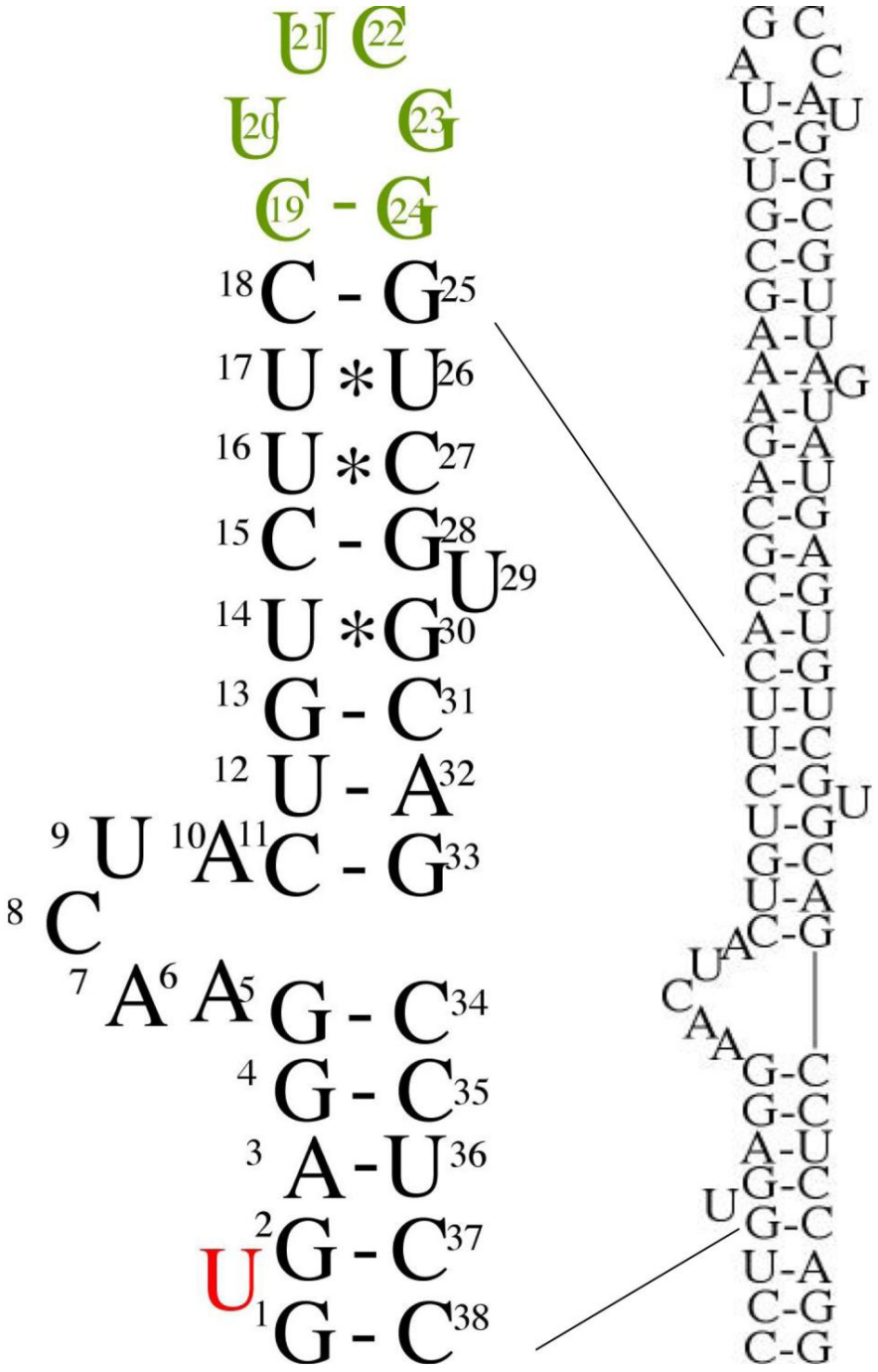

Figure 14: Secondary structure of the IRES5 construct and the full HCV IRES domain II. The base numbering in the figure will be referenced throughout this dissertation. The red uridine was omitted from our construct for stability and so does not have a number. The green bases were added to stabilize the construct.

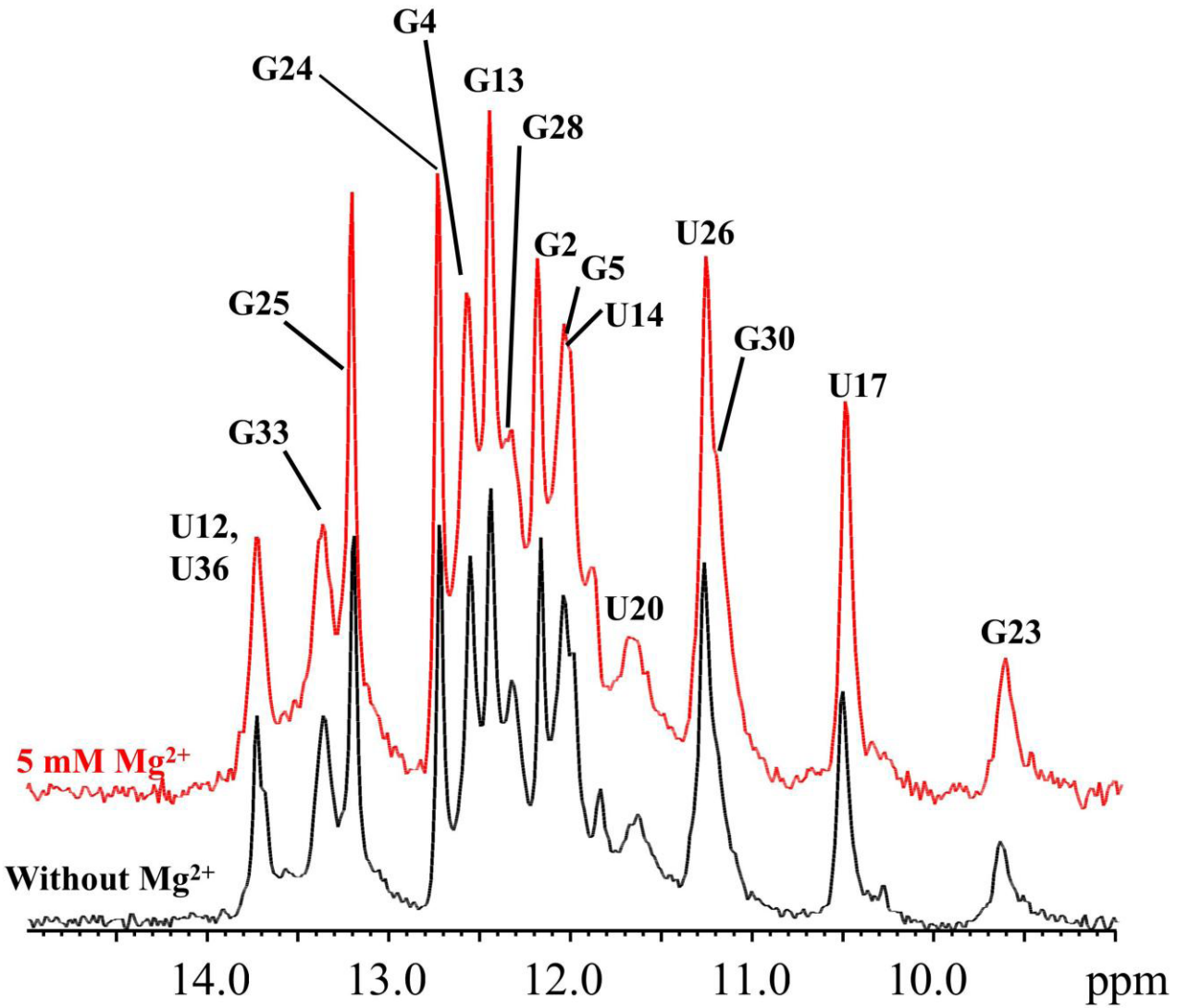

Figure 15: ^{1}H Imino spectra with an excess of Isis-11 in the presence and absence of magnesium. The common sample conditions were as follows: 230 µM IRES5, 150 mM NaCl, 150 mM KCl, 2 mM spermine, 10 mM sodium phosphate, pH 7.0, 20º C. The differences between the imino spectra of the complex in the presence and absence of Mg^{2+} are extremely subtle.

CHAPTER 3

COMPLEX STRUCTURE DETERMINATION

Methods

NMR Methodology

All NMR data used to determine the complex structure were produced on Varian INOVA 500 and 600 MHz instruments at the University of Utah. Some additional experiments were conducted on a Varian INOVA 800 MHz instrument at the University of Colorado at Boulder, but the line broadening in the complex was extreme at high field.

The assignment of resonances for this structure was greatly facilitated by the previously published work of the Puglisi lab in the determination of the complete structure of the HCV IRES domain II.[27] Our smaller domain IIa construct (38 residues vs. 55 residues) and differing NMR solution conditions (Ours: 150 mM KCl, NaCl; 2 mM spermine; 10 mM sodium phosphate, pH 7.0. Theirs: 100 mM NaCl, 5 mM $MgCl_2$, 10 mM sodium phosphate, pH 6.4) nevertheless yielded free RNA chemical shifts very similar to the published data.

Imino Proton Data

The imino protons on uracil and guanine provide simple data that are extremely powerful in the determination of RNA structure. In a stable base pair, these protons are shielded from solvent exchange to the extent that strong signals can be seen in a ^1H

spectrum. The chemical shift range of the imino protons is wide (~14-9 ppm). AU and GC base pairs tend to occupy distinct regions of this range with some overlap while non-canonical base pair interactions often have highly distinctive imino chemical shifts separate from the Watson-Crick pairings. Consecutive base pair imino protons fall within 5Å and produce an NOE interaction. Thus, a 2D NOESY can be run in H_2O to provide data that can then be used to assign the imino spectrum. This process is greatly facilitated by foreknowledge of the likely secondary structure of the RNA which can be predicted reliably for small, simple constructs with programs such as mfold.[59]

The free IRES5 provided narrow lines with high signal to noise in the imino spectrum, and we were able to assign this region with a 2D NOESY in H_2O.[60] In the complex, however, the line broadening reduced the signal to noise such that a 2D NOESY on the imino region was impractical. Instead, 1D difference NOE experiments were run on an unlabeled complex sample to establish connectivity between adjacent imino protons.[61] Imino peaks for uridine and guanosine were distinguished based on the ^{15}N shifts measured from ^{15}N HSQC on a ^{13}C ^{15}N labeled IRES5 complex sample.[62]

Carbon Bound Protons and NOESY Data.

In order to determine a high resolution structure of the complex, it was necessary to choose a suite of NMR experiments that would enable us to detect and assign intermolecular NOE cross peaks within the context of the limitations of our system. We were limited in the concentrations of RNA that could be properly folded. We were also limited in the temperature range. Lower temperatures, which under ideal circumstances might have enabled us to drive certain exchanges from the intermediate to the slow exchange regime, proved difficult to work with due to line broadening. This may have

been due to increased non-specific interactions, or it may have driven certain exchange rates farther into the intermediate range. Thus, all NMR experiments, with the exception of the aligned RDC experiments discussed later, were performed at 35º C.

Line broadening in the complex, coupled with the high level of resonance overlap due to the high molecular weight of the RNA, made it necessary to employ NMR experiments that could provide us with adequate resolution. One possibility was the use of 3D ^{13}C HSQC filtered NOESY experiments. This is the standard approach with proteins. It is often used for RNA, but there are disadvantages. Dispersion in both the sugar and aromatic regions are much poorer in RNA than in proteins.[63] Chemical shift anisotropy also leads to faster T2 relaxation times and consequently broader lines than are found in proteins of similar molecular weight.[64] Additionally, the inter-proton distances in RNA are typically longer on average than in proteins due to the lower density of protons per residue, especially outside of the sugar region. Thus we have less signal and broader lines in a less dispersed spectrum. In addition to these general difficulties, the concentration limitations and the exchange broadening in the IRES5/Isis-11 complex made most 3D experiments impractical.

Filter-Edit Experiments

The challenges presented by the IRES5/Isis-11 complex were overcome through the use of a suite of filtered/edited NOESY experiments described by Feigon et al[65] with selective isotopic labeling. Selective labeling is the use of specific ^{13}C labeled and unlabeled nucleotide triphosphates in one T7 transcription reaction in order to produce a sample that contains, for instance, ^{13}C labeled guanosines and cytosines and unlabeled adenosines and uridines. Such a sample can be used to produce 2D filtered/edited

NOESY experiments that are considerably easier to interpret than conventional 2D NOESY data produced on an unlabeled sample. These experiments are also more sensitive than a 3D ^{13}C HSQC filtered NOESY.

The term "filter" is used to describe a purge element that abolishes the signal from protons bound to an isotopically labeled nucleus (labeled protons) while retaining signal from protons bound to an unlabeled nucleus (unlabeled protons). The term "edit" refers to a purge element that removes the signal from unlabeled protons while keeping signal from labeled protons. Filtered/edited experiments work by applying either a filter or an edit before either the t_1 or t_2 evolution periods.

Filter/edit NOESY experiments are often used with complexes involving a labeled and unlabeled component. Such experiments can greatly facilitate the distinction of intramolecular NOE cross peaks from intermolecular NOE cross peaks while also sorting out the intramolecular NOE cross peaks from one complex component from the intramolecular NOE cross peaks from another complex component. With a selectively labeled RNA, these experiments can be used to greatly simplify intramolecular NOESY data. They are especially useful in assigning the highly congested aromatic to H1' and aromatic to H2'-H5" walk regions in RNA. Three powerful filter/edit experiments were employed that resulted in relatively simple NMR spectra, facilitating resonance assignments.

The F1 edit F2 edit (F1eF2e) experiment applies an edit before the t_1 and the t_2 evolution periods, thus removing all NOE interactions from ^{12}C bound protons as well as all NOE cross peaks to ^{12}C bound protons. Thus, the resulting NOESY spectrum contains only NOE cross peaks from ^{13}C bound protons to ^{13}C bound protons. The H1' to

aromatic and H2'-H5" to aromatic walk regions of such a spectrum thus contains only the intranucleotide NOE cross peaks of the ^{13}C labeled bases in addition to the internucleotide NOE cross peaks between labeled bases. The 4D-^{13}C-edited NOESY described by Vuister et al was run as a 2D experiment to generate these spectra.[66]

The F1 filter F2 filter (F1fF2f) experiment applies a filter before the t_1 and t_2 evolution periods and thus filters out all NOE interactions from ^{13}C bound protons as well as all NOE cross peaks to ^{13}C bound protons. The resulting NOESY spectrum contains only NOE cross peaks from ^{12}C bound protons to ^{12}C bound protons. The spectrum contains only the intranucleotide NOE interactions of the unlabeled nucleotides as well as the internucleotide NOE cross peaks between unlabeled nucleotides. An additional advantage of this experiment compared to the other filter/edit experiments is that there is less line broadening and signal loss from ^{12}C bound protons than that from bound ^{13}C bound protons. Proton relaxation and the associated signal decay are thus slower in this experiment. In practice, this proved to be a highly sensitive and informative experiment. The pulse sequence used in this study parallels that described by Zwahlen et al.[67]

Additionally, the F1fF2f data sets facilitated intermolecular NOE assignments. Intermolecular NOE cross peaks seen in the F1fF2f spectrum from the AU labeled sample indicated interactions between G or C nucleosides with Isis-11. Conversely, cross peaks seen in this experiment with the GC labeled sample were indicative of interactions between Isis-11 and A or U nucleosides.

The third experiment that completes this picture is the F1 filter F2 edit, or F1fF2e experiment. This experiment employs a filter prior to the t_1 evolution period and an edit prior to the t_2 evolution period. Thus, the f1 dimension contains only cross peaks from

unlabeled protons while the f2 dimension contains only cross peaks to labeled protons. This experiment filters out NOE cross peaks from ^{13}C bound protons and then edits out all NOE cross peaks to ^{12}C bound protons. Thus, this spectrum contains only cross peaks from ^{12}C bound protons to ^{13}C bound protons (i.e., cross peaks only between unlabeled and labeled nucleotides). The two halves of the spectrum on opposite sides of the diagonal offer unique information as well. For instance, in the aromatic to H1' NOESY walk region, one side will contain cross peaks between labeled sugar protons and unlabeled base protons while the other side will contain cross peaks between labeled base protons and unlabeled sugar protons. Thus, the spectral overlap is cut in half over and above the decrease in overlap from the removal of all labeled to labeled and unlabeled to unlabeled cross peaks. Another advantage of this experiment is that it removes all of the very intense and comparatively uninformative pyrimidine H5-H6 cross peaks from the H6/H8 to H1' aromatic walk portion of the spectrum. These spectra should be the least congested and the easiest to interpret. The pulse sequence used for these experiments parallels that described by Stuart et al.[68]

The F1fF2e can also be used with either a selectively ^{13}C labeled or fully labeled RNA construct to look for NOE cross peaks between the unlabeled Isis-11 and the RNA. With a fully ^{13}C labeled RNA construct, all of the NOE cross peaks in the spectrum will be intermolecular. As this interface is the most interesting part of this structure, the advantages of such a data set speak for themselves. All strong NOE cross peaks between the compound and the RNA were corroborated with this data set; however, in practice, the experiment was not sufficiently sensitive to see all intermolecular NOE cross peaks. Nevertheless, it was very helpful in locating medium to strong intermolecular NOE cross

peaks that were located in areas of the spectrum that were highly congested with intramolecular RNA cross peaks in the other NOESY experiments.

Combined, the three NOESY experiments described above contain all of the same NOE cross peaks as a conventional NOESY. See figure 16. In our experience, the signal to noise on the F1eF2e was very poor in comparison to the F1fF2e and the F1fF2f. A more fruitful approach was to conduct the F1fF2f experiment on both a selectively AU labeled and GC labeled RNA sample. An F1fF2f experiment on an AU labeled sample provides the same spectrum as an F1eF2e experiment on a GC labeled sample but with higher sensitivity. The F1fF2f experiment was also conducted on GU, GA and GUA labeled samples in order to clear up some remaining ambiguities from the other data sets.

Assignments

Titrations of IRES5 with Isis-11 enabled me to follow changes in the ^{13}C HSQC spectrum upon compound binding. Some resonances were broadened at very low stoichiometric equivalents of compound. In high salt, the overall quality of spectra is at its poorest at ~1 equivalent of Isis-11. As the Isis-11 compound is titrated to excess, the peak shapes begin to improve. Many of the broadened resonance peaks began to narrow at higher equivalents of compound. Some resonances moved gradually as more compound was added. Some resonances did not move at all. The resonances that did not move served as a convenient starting point for the assignment of the bound RNA. Many of the intrastrand NOE connectivities were retained. The resonances that were heavily broadened by exchange and then narrowed again at new chemical shifts were assigned based on connectivities to other known resonances. These titration experiments were conducted on a selectively GC labeled and a selectively AU labeled RNA sample in

addition to a fully labeled sample. The selectively labeled samples greatly facilitated the assignment and monitoring of changes in the pyrimidines as these protons have less dispersion in general and there are more pyrimidines than purines in the IRES5 construct.

Two-dimensional filter/edit NOESY spectra were used in conjunction with two-dimensional unlabeled NOESY data and ^{13}C HSQC data to make definitive H1' to H8/H6 walk assignments. The intraresidue H1' to H8/H6 correlations were confirmed as much as possible with HCN experiments, although this data was of poor quality with our complex samples. The NOESY data and ^{13}C HSQC also facilitated determination of the H5 and H2 assignments on the pyrimidines and adenines respectively. Further sugar assignments were determined using the two-dimensional unlabeled NOESY data in conjunction with HCCH COSY and HCCH Relay-COSY experiments.[69] This enabled the assignment of most of the H2' and H3' sugar protons. Short mixing time NOESY experiments also facilitated H2' and H3' assignments due to the fact that A-form helical RNA usually results in very short distances between n H2' protons and n+1 H6/H8 protons in addition to intraresidue H3' to H8/H6 protons in A-form or nearly A-form regions.

Due to line broadening, it was difficult to assign many NOE cross peaks in certain congested regions of the NOESY spectrum. For this reason, and the limited sensitivity of the HCCH-RELAY COSY, the H5', H5", H4', and H3' assignments are incomplete. This left us with NOE cross peaks heavily weighted toward H1', H2', and H3' to aromatic, and aromatic-to-aromatic interactions.

A handful of H4' and H5' and H5" complex assignments were made on the basis of the Puglisi structure assignments.[27] If the other resonance assignments for a given

residue did not change much upon compound addition, and if there was adequate resolution in the sugar region, then assignments were made for H4', H5', and H5" assuming that they very closely matched the Puglisi assignments. This was corroborated by filter/edit NOESY data.

Residual Dipolar Couplings

^{1}J C-H and N-H dipolar couplings are usually averaged out to zero due to isotropic tumbling of a molecule in solution. In a weakly aligned molecule, the dipolar couplings can be large enough to measure. A number of these can be used to calculate an order tensor frame. Individual bond vectors are then fit into this order tensor frame to minimize the difference between the theoretical and the measured couplings in conjunction with other restraints.[70]

The measurement of RDCs has become an important NMR method for determining macromolecular global structures and is especially powerful for positioning distal domains relative to each other.[27, 71-73] As all of the other typical NMR restraints (i.e. NOE restraints, dihedral angle restraints, base pairing restraints) are short range in nature, it is necessary to support these restraints with global structural information in order to attain good global structural convergence.

RDCs can be measured in a number of ways. Experimental methods fall into one of two categories: intensity based methods and frequency based methods. Intensity based methods are more prone to systematic errors than frequency based methods.[70] The primary drawback to frequency based methods is the greater spectral congestion in coupled spectra, but this can be neutralized through the right choice of experiments.[70] The best approach for the measurements of RDCs in this system was through the use of

in-phase anti-phase (IPAP) experiments.[74-76] The IPAP produces spin-state separated spectra within a single experiment. In this experiment, two spectra are collected in which the coupling evolves either in phase or anti-phase. These spectra are then added or subtracted to produce two spectra, each containing one half of the doublet.[74] This minimizes spectral congestion.

The alignment media consisted of a C12E5/hexanol mix that has been shown to create a dilute nonionic liquid crystalline media ideal for weak alignment.[77] The C12E5 term describes an alkyl-poly(ethylene glycol) molecule with 12 carbons in the alkyl moiety and 5 glycol units comprising the poly(ethylene glycol). RDC restraints for the aromatic protons were determined through the use of this alignment media at 3-3.2%. The optimal temperature for our alignment regime was found to be 31º C. This is lower than the optimal temperature of 35º C used for all of our other nonexchangeable proton NMR experiments, but there was no significant peak movement when decreasing the temperature from 35º C to 31º C. The cross peaks were broader for the aligned samples, probably due to a combination of additional exchange broadening at the lower temperature and the line broadening induced by the partial alignment.

The alignment and lower temperature-induced line broadening, coupled with the already congested spectra under ideal conditions necessitated the use of AU and GC selectively labeled samples. The AU and GC labeled samples were analyzed by ^{13}C HSQC-IPAP with either C6 or C5 band-selective WURST decoupling in order to determine the aromatic C-H couplings. HCCH-COSY-IPAP experiments with band selective WURST decoupling of the C2' carbons were used to measure the couplings for the C1'-H1' bond vectors. These values were corroborated with couplings measured from

F2 coupled ^{13}C HSQC spectra. Some additional couplings were obtained exclusively from these coupled experiments due to superior peak shape and intensity in some cases. Additionally, imino N-H couplings were measured from coupled ^{15}N HSQC spectra.

Extensive overlap and line broadening prevented the complete determination of the C6-H6 couplings of the cytosines; however, the C8-H8, C6-H6, and C2-H2 couplings of the adenosines, uridines, and guanosines were sufficiently resolved. Only four C5-H5 couplings were measured due to poor spectral quality. The error in our measurements was greater than the typical variability of the C-H coupling for unaligned RNA (<1 Hz). The unaligned values for the C2-H2, C6-H6, C5-H5, and C8-H8 bond vectors used to subtract from the aligned values were based on the measurements of Boisbouvier et al[78] since it has been shown that the unaligned coupling values do not vary significantly for the aromatic bond vectors of a particular base type.

The HCCH-COSY-IPAP experiments performed on these samples to determine the couplings on the C1'-H1' bond vectors proved to be the most challenging. It was necessary to determine the aligned and unaligned RDCs for these bonds due to the fact the unaligned values vary greatly depending on the torsional angles of the sugars. This compounded the sensitivity problem in that the peaks had to be strong, well shaped, and well resolved enough in four different spectra (aligned vs. unaligned, in phase vs. antiphase) in order to measure the RDCs for these bond vectors. Due to these complications, only nine C1'-H1' RDCs were measured.

In all cases, measurements were not included in the structure calculation if the peak was broadened to the point that it was not possible to locate the peak center with high confidence.

Sugar Puckers

While there are many degrees of freedom in a free ribose ring, the classical A form helix typical of an RNA duplex forces the sugar into a very narrow range of possible conformations. The sugar typically assumes a C_3'-endo conformation under these circumstances. The steric restraints placed on the ring by the hydroxyl groups and the base limit the possible ring conformations outside of the A form regions as well. Typically, most of the riboses will be in either the C_2'-endo or C_3'-endo conformation.

In the C_3' endo conformation, the H3' will have a stronger NOESY cross peak with the intraresidue aromatic H6 or H8 proton than other intraresidue sugar protons due to its close proximity to the base. This is especially evident in a NOESY experiment with a short mix time (<100 ms). Another strong indicator of the sugar pucker of a base is the presence of a H1' to H2' cross peak in a TOCSY experiment. If the cross peak is present, then the sugar must be in the C_2' endo conformation. The C_3' endo conformation places the H1' and H2' sugar protons at a dihedral angle approaching 90º which leads to minimal coupling between the two hydrogens according the Karplus relationship and, thus, the absence of a peak in the TOCSY. See Figure 17.

Fluorescence Experiments

2-amino purine nucleosides were used to probe the structural changes upon Isis-11 binding. The structure of 2-amino purine is very similar to adenine, as seen in Figure 18, so it should not significantly disrupt RNA structure or complex formation. The fluorescence of this nucleobase is sensitive to the structural environment, particularly the local stacking environment. A well-stacked 2-amino purine should be somewhat

quenched by the neighboring aromatic n+1 and n-l bases. An unstacked base, or a base in a region where the stacking is disrupted or in flux, should fluoresce more strongly.[79]

Structure Refinement

As these studies were conducted with a racemic mixture of compound, it was necessary to perform structure refinement with both isomers. Both isomers were capable of accommodating most of the restraints with very limited changes in the RNA structure or orientation of Isis-11. Not surprisingly, the relative strengths of some intermolecular NOE cross peaks were more consistent with one stereoisomer than with the other. This was especially true for NOEs to H11, the only proton on a chiral center. The intermolecular NOE cross peak list thus represented an ensemble from both stereoisomers.

Short mixing time (100 millisecond) NOE cross peaks arise from protons in close proximity and such interactions were given a restraint range of 1.8 to 3.0Å. The peaks seen in the long mixing time data that could not be found in the short mixing time data were divided into two groups. The stronger set was given a range of 1.8 to 4.5 Å while the weaker set was given a range of 1.8 to 5.0 or 5.5Å. The upper bound on these weak restraints were sometimes restrained at 5.0 Å if we deemed the actual distance to have a high probability of being within 5 Å and not due to spin diffusion based on common RNA distances. These restraint ranges were loosely based on the NOE cross peak intensities of hydrogens of known or relatively fixed distances, such as the H5-H6 cross peaks in the pyrimidines. Relative intensities of intermolecular NOEs were also used to fine-tune some restraints in the final iterations of structure refinement.

Some intermolecular lower bound restraints were used in the structure calculation on the basis of the unambiguous absence of an NOE cross peak where one should be visible in the spectrum if two hydrogens were in close proximity. Many of the intermolecular NOE cross peaks were very strong compared to the intramolecular RNA cross peaks due to the large excess of compound in the NMR samples. Many of them also had very unique chemical shifts with little overlap with intramolecular RNA cross peaks. Thus, it was easy to discern the absence of the peaks used as the basis for the intermolecular lower bound restraints. The absence of a peak was based on the NOESY spectrum of the unlabeled complex and the F1fF2f spectra on the selectively labeled RNA complexes, as these experiments were more sensitive and thus less likely to indicate peak absence on the basis of inadequate sensitivity.

Structures were calculated using a combination of Xplor-NIH 2.18 for in vacuo refinement by Darrell R. Davis and AMBER 10 for refinement with explicit solvation by Thomas E. Cheatham. The refinement protocol for simulated annealing in torsion angle space described by Clore and Kuszewski for the RNA aptamer/theophylline complex and provided in the Xplor-NIH distribution was used with only minor changes.[80] A notable exception to the distributed simulated annealing protocol was that the harmonic RDC potential was scaled to 0.4 to reflect the measurement uncertainty of the dipolar coupling measurements. Alignment tensor calculations were done with either the calcTensor python script provided with Xplor-NIH, or with PALES.[81] Electrostatic potential surface calculations were done using the AMBER/APBS module.[82] Structures were visualized and figures prepared using PyMOL and Chimera.[83] RMS differences and average structures were calculated using Clusterpose.[84]

The specific Xplor-NIH protocol is as follows: the published free NMR structure[27] was used as a model for an initial conformation of the free IRES5. An S isomer of Isis-11 was docked to this using low temperature molecular dynamics with intra and intermolecular NOE restraints. The starting structure was refined by restrained molecular dynamics *in vacuo* with NOE, base-pairing, and dihedral angle restraints (referred to hereafter as "NOE-only" restraints) to generate 20 structures. Of the 20, approximately half belonged to a general consensus family with a low number of NOE violations and an RMSD of ~4Å. This family of structures possessed a poor global fit due to a relatively low number of sequential NOEs and a lack of RDC restraints. Several representatives from this family were then used independently as starting structures for refinement with the NOE-only plus RDC restraints. After confirmation that each of these starting structures provided a similar family of 10 structures from the molecular dynamics refinement, one of the NOE-only refined RNA structures was used to generate a family of 30 complex structures for each of the Isis-11 enantiomers using NOE-only plus RDC restraints. From these families of structures, 16 (R-isomer) and 19 (S-isomer) lowest energy model structures were refined by molecular dynamics at 300K with explicit solvation in AMBER applying the NOE and RDC restraints. For each model, the > 10 ns length MD trajectories were clustered at 1 ps intervals into 5 groups based on the pair wise RMSD of all atoms of the RNA and Isis-11, and the average structure from the dominant cluster at the end of the simulations was minimized with implicit solvent. Although the entire family of structures fit the NMR data well, further energetic analysis from both minimization of the average structures and also MM-PBSA of the chosen

cluster from the MD trajectory pruned the entire set of 35 model structures to generate families of 10 final structures for each of the R and S isomers.

Results

Resonance Changes

Imino Protons

Under our standard conditions, the free RNA gave imino proton signals for the base pairs between G2-C37, A3-U36, G4-C35, G5-C34, C11-G33, U12-A32, G13-C31, U14-G30, C15-G28, U17-U26, C18-G25, C19-G24, and the tetraloop imino protons from G23 at and U21 at a range of temperatures between 10º and 30º C. Upon compound binding, some specific imino protons resonances shift, but all base pairs remain intact indicating limited stem disruption. See Figure 14. This is consistent with the placement of the compound in the bulge region. At very high excesses of the compound, the U14-G30 and C15-G28 imino protons become broader, suggesting some secondary interaction between the compound and this region of the IRES.

RNA Aromatic Protons

The TOCSY spectrum changes are easy to follow and useful for characterizing Isis-11 binding. In Figure 18, the changes of the TOCSY spectrum between the free and the bound RNA are shown. Most of the H5-H6 pyrimidine cross peaks with the largest chemical shift changes are in and around the bulge region of IRES5. The U9, C8, and C11 H5-H6 cross peaks show especially large chemical shift changes. The C34 and C35 H5-H6 cross peak chemical shifts also change. The U36, C37, and C38 cross peak chemical shifts change to a lesser degree, and the change is directly proportional to bulge

proximity. The U29 cross peak chemical shift also changes. The line widths of U29 remain narrow in the complex with a strong TOCSY cross peak, suggesting that U29 remains extruded from the helix. The cross peak chemical shifts in and very near the tetraloop do not appear to change at all, particularly C19-C22. As we proceed down the stem and get closer to the bulge region, the cross peak chemical shifts change slightly. The peaks with the weakest intensities are C8 and C11. This exchange broadening is consistent with binding of Isis-11 in this region. See Figure 19.

^{13}C HSQC aromatic proton spectra provide a more complete picture than the TOCSY because they contain peaks from both purines and pyrimidines, but these spectra suffer from greater spectral congestion. The overlap for the pyrimidines was particularly problematic. The spectral changes were independently monitored with titrations on AU and GC labeled samples in addition to the fully labeled RNA in order to decongest the spectra.

The titration of Isis-11 into the free fully labeled RNA at low stoichiometric equivalents provides some insight into the specificity of the binding site. G33 C8-H8, C8 C6-H6, A6 C8-H8 are well resolved in the free ^{13}C HSQC, and they are severely broadened at low equivalents of compound. Other peaks that show early broadening are A7-H8, A10-H8, A7-H2, A32-H2, and A6-H2. See Figure 20.

The ^{13}C HSQC on the GC labeled sample showed peak shifting mostly in and around the proposed binding site. The C6-H6 cross peaks for C8 and C11 move significantly, as does G33. The peaks in the tetraloop are the least affected. G5 moved very little, which is somewhat counter-intuitive, but not impossible assuming that the degree of shielding for the residue is similar in both the free RNA and in the complex.

There was a surprising amount of movement for G28. This is possibly related to the peak mobility seen for the U29 H6-H5 cross peak from the TOCSY. Figure 21 shows the changes in the ^{13}C HSQC for the GC sample upon compound binding.

The aromatic ^{13}C HSQC spectrum of the AU labeled sample showed a number of large changes, especially for the adenosines (H2 and H8 protons). This is to be expected given the fact that all of the adenosines are relatively close to the binding site. Specifically, A6, A7 and A10 are all in the bulge. A32 is very close to the bulge. A3 is the most isolated from the bulge, and does not move. The only uridine in close proximity to the binding site is U9, whose C6-H6 cross peak moves significantly. The U29 C6-H6 also changes. In the free form, the U29 is in proximity to the bulge, so it is possible that compound binding alters the environment of this base sufficiently to explain this chemical shift change.

There is a great deal of exchange broadening seen in the case of the adenosines. The A32 and A3 H2 peaks overlap based on the corresponding cross peaks in the NOESY spectra. The A7 H2 peak is particularly broad and sometimes looks like two very slightly resolved peaks. This might be due to slightly different binding environments for the two Isis-11 isomers. It is possibly due to intermediate exchange. The adenosine H8 peaks are also broad in the complex. The free and bound spectra for the H2 region of the ^{13}C HSQC labeled RNA are shown in Figure 22. The assigned RNA chemical shifts are summarized in Tables 2 and 3.

Nuclear Overhauser Effect Changes

The labeling scheme for the Isis-11 hydrogens used in the Amber structure refinement is summarized in Figure 23. This same scheme will be used to describe the intermolecular NOE cross peaks.

Although some geminal protons were resolved from one another (H10A-B, HA2A-B), it was impossible to distinguish them given the racemic mixture we were using. All other geminal sets were not resolved. Free rotation of the dimethyl groups resulted in redundant chemicals shifts, and thus it was not possible to distinguish methyl group HD1A-C from HE1A-C, and HG2A-C from HG3A-C. In these cases, the distance restraints were directed to the carbon for the methylenes and to the nitrogen for the methyl protons of the dimethylamino groups.

The F1fF2f and F1fF2e NOESY spectra were invaluable in distinguishing intermolecular NOEs from intramolecular NOEs. A region of the unlabeled NOESY spectrum with overlays from F1fF2f and F1fF2e spectra is shown in Figure 24. The F1fF2e experiment was particularly important for the determination of intermolecular NOEs between RNA protons and Isis-11 protons whose chemical shifts fell in or near the ranges typically occupied by RNA protons.

In one case, this approach facilitated the assignment of some critical intermolecular NOE cross peaks. There were potential NOE cross peaks from both the H2 from A6 and H6 from U9 to several protons on Isis-11. Unfortunately, these two RNA protons had nearly identical chemical shifts in the proton dimension. As both of these bases are in proximity to other bases known to interact with the molecule, either base could be in close proximity to Isis-11. Fortunately, the F1fF2f data from the

selectively labeled complex samples cleared up this ambiguity, and it became clear the intermolecular NOE cross peaks were to A6-H2.

Many of the strong intermolecular NOE cross peaks occur between Isis-11 and G5 and A6. A strong NOE occurs between proton H1 of the compound and the H1' of G5. There were also NOEs between G5 H1' and Isis-11 H2 and the HG3A-C and HG2A-C methyl proton set. There were also other strong NOEs between Isis-11 and other G5 sugar protons. Other pronounced and easily distinguished NOE cross peaks occur between H8 of this G5 and several compound protons (H11, H1, HA2A, HA2B, HG3A-C, HG2A-C). The large number of NOEs between Isis-11 and G5 clearly indicate that the binding site involves G5. There are also a large number of NOE peaks to H2 of A6 involving many of the same Isis-11 protons (H10A-B, H11, HG2A-C, HG3A-C, H1, H2, HA2A-B). This suggests that A6 is also located in the binding site.

There are a few weaker NOEs to G4 H1' and H8 from the HG3A-C and HG2A-C methyl proton set. Some of the NOEs to this methyl set are shown in Figure 25. Figure 25 demonstrates the utility of the F1fF2e experiment. The data shown clearly indicate that the F1fF2e experiment does an excellent job of filtering out intramolecular NOEs from Isis-11, which are typically very strong given the excess of Isis-11 relative to IRES5. The Isis-11 H1 to HG3A-C and HG2A-C methyl proton set peak is screened out completely in the F1fF2e experiment. Another fortunate outcome of this experiment is that it enabled us to see the intermolecular G5-H8 to HG3A-C and HG2A-C methyl proton set cross peak that would otherwise be mistaken for the intramolecular Isis-11 H2 to HG3A-C and HG2A-C methyl proton set cross peak. The common chemical shift of

G5-H8 and Isis-11 H2 would otherwise have made any intermolecular NOE assignments involving G5-H8 very doubtful.

There are few NOE cross peaks to the hydrogens of the dimethylamino group on the long methylene arm (protons HD1A-C and HE1A-C) and the methylene arm itself. There is, however, one strong and unambiguous NOE cross peak between the H5 on U12 and the methyl protons. There is also an NOE cross peak from H5 of C11 to these methyl protons (HD1A-C, and HE1A-C). This anchors that end of the molecule in the structure and proved to be critical for convergence in the structure refinement. See Figure 26.

In contrast to the paucity of NOEs for the long arm, there were a number of strong NOE cross peaks to the protons of the dimethylamino group of the short methylene arm (protons HG3A-C and HG2A-C). The cross peaks to these protons were easily distinguished due to their chemical shifts being upfield of the ribose sugar region. This was true for all the methyl protons and most of the methylenes in Isis-11.

There were also a number of NOE cross peaks to the H11 proton, which sits on the lone chiral center of Isis-11. Of course, as this is a racemic mixture of Isis-11, it must be taken into account that the NOE cross peaks to Isis-11 could be based on the binding of either the R or S isomer.

The 24 intermolecular NOEs are summarized in Table 4.

Sugar puckers

The ribose sugar conformations for the complex were predominantly 3' endo, as expected. The few sugar pucker conformations that could not be determined

unambiguously were left unrestrained in the structure calculations, as were the 2' endo sugars. The sugar conformations in the complex are summarized in Table 5.

Residual Dipolar Couplings

The measurements of the C6-H6 and C5-H5 from the pyrimidines and the C8-H8, and C2-H2 from the purines were of generally good quality. The fitting of the RDC measurements in the calculation of the structure yielded good alignment with very few violations. The list of RDC constraints used in the structure calculation is shown in Table 6. Figure 27 shows a portion of the aligned IPAP ^{13}C HSQC data from the G and C labeled sample.

Isis-11 Chemical Shift Changes

The ^{1}H chemical shifts of Isis-11 were dominated by the unbound form due to the high excess of compound relative to the RNA. Most of the protons in the complex were in fast to intermediate exchange; nevertheless, the aromatic (H1 and H2) and chiral (H11) compound protons were in slow exchange as a result of significant chemical shift changes between the bound and free forms. These assignments and this exchange behavior were confirmed with ROESY NMR experiments in which the phase of the cross peak between the free and bound forms of the same proton was the opposite of the phase of the other NOE cross peaks. The significant upfield shift is consistent with the placement of these protons near the center of the pi electron cloud of an aromatic group. While this information was too vague to admit as a specific restraint in the structure refinement, it did inform our analysis of the results. Given that these protons sit in a

rigid, relatively flat, mostly aromatic portion of Isis-11, this might mean that these protons are situated in a stacked environment in the complex.

Melting Experiments

Melting experiments done by the Hermann group have demonstrated that the presence of divalent metals not only raises the T_m of domain IIa, but also increases the cooperativity of melting. [26] As this molecule is known to bind near the critical Mg^{2+} binding site of the RNA, we were curious as to how the molecule would affect the T_m in the presence and absence of Mg^{2+}.

At a pH of 7 (10 mM sodium phosphate), in the absence of Mg^{2+} and in the presence of 50 mM NaCl and KCl, the IRES5 construct showed a small drop in the T_m in the presence of an excess of Isis-11. The free RNA yielded a T_m of 62.5±0.5°C while the complex yielded a T_m of 61.6±0.2°C. This is consistent with the model in that it indicates that there is not significant secondary structure disruption in the presence of the compound. While Isis-11 may bind near Mg^{2+} binding sites on the RNA, it does not have the same stabilizing effect since the T_m does not go up.

Given the apparent importance of Mg^{2+} in the complex, melting was studied in the presence of 1 mM $MgCl_2$. The salt concentration was lowered to 10 mM KCl and NaCl in order to lower the T_m and thus avoid the Mg^{2+} facilitated RNA hydrolysis at high temperatures. It was also necessary to keep the monovalent salt concentration low in order to avoid raising the T_m above the point at which it could be accurately measured in water. The sample was buffered at pH 7 with 10 mM sodium phosphate. In the absence of compound, the T_m was 68.3±0.4°C. In the presence of an excess of Isis-11 the T_m was slightly reduced (67.5±0.8°C). It is interesting that the compound lowers the T_m by about

1° C in both the Mg^{2+} and Mg^{2+}-free samples.

For the sake of comparison, the Mg^{2+} free melting study on IRES5 in the absence of compound was conducted a second time with lower salt concentrations as in the Mg^{2+} experiment described above. At 10 mM KCl and NaCl and 10 mM sodium phosphate pH 7, the T_m was determined to be 55.1±0.5°C. This is significantly lower than the T_m of the free IRES5 with Mg^{2+} (68.3±0.4°C). This is consistent with the general phenomenon of divalent metal nucleic acid stabilization, and is not necessarily indicative of any special stabilization.[85]

In summary, the compound does not appear to significantly facilitate melting of the RNA, although it does consistently drop the T_m by approximately 1° C in the presence and absence of Mg^{2+}. It also does not appear to stabilize the helix significantly. This is consistent with the placement of the compound in the bulge with limited perturbation of the upper and lower stems.

Fluorescence

The fluorescence of 2-aminopurine substituted RNAs at positions A6 and A7 increases upon compound binding. This is consistent with the unstacking of these bases. The fluorescence of the A10 2-aminopurine substituted RNA is more quenched upon compound addition, suggesting a higher degree of stacking in the complex.

The fluorescence intensity at 370 nm for the A6 2-aminopurine was also measured as a function of compound concentration in order to estimate the K_D. The fluorescence change upon compound binding was consistent with a simple two-state binding model. This curve fit indicated a K_D of approximately 2.4 μM, which is

consistent with the value reported previously.[54] The change in fluorescence for the A6 2-aminopurine titration can be seen in Figure 28.

Structures and Statistics

The structure was refined for both the S and R isomers. Given the limited number of both intra- and intermolecular NOE restraints, the convergence is good. The RMSDs of the R and S isomers were 1.7 and 2.2A, respectively. The R-isomer structures tend to have fewer bond and torsion violations, lower free energies for the RNA plus Isis-11 complex, and more favorable RNA free energies. Greater convergence to a common structure was also seen with the R isomers. The free energy of the binding of Isis-11, however, is comparable between the two isomers. See Table 7.

The convergence is worst in the area with the fewest restraints, namely the A7-C8 region of the bulge. There is also a high degree of mobility for U29, which is probably flipped out most of the time, but appears to sample a number of conformations. The convergence of the Isis-11 placement in the bulge is remarkable, as is the high degree of similarity between the R and S stereoisomer structures. The superpositions of the 10 lowest energy structures for the two structures are shown in Figure 29.

The fluorescence changes described previously are consistent with the structural changes indicated in NMR experiments. The NMR structure suggests that U9 flips in and becomes stacked with the A10. The structure also shows A6 unstacks from G5 and A7 and interacts with the molecule. A7 is unstacked from A6. It is not entirely clear what C8 is doing, but there are some NOE cross peaks to A7 that are consistent with, if not unambiguously indicative of, some stacking between C8 and A7.

Isomers

Interestingly, both isomers could be reasonably well accommodated in very similar RNA structures with limited violations. At first glance, the structures appear identical. Both isomers occupy essentially identical binding sites. Both are in largely extended conformations. This is consistent with the absence of NOEs between the dimethylamino side chains.

Given the additional fact that the two aromatic protons (H1 and H2) and the chiral proton (H11) of Isis-11 are significantly shifted upfield, the two families of structures must also be judged on the basis of the placement of these protons near the face of an aromatic group. In the S Isomer, the tricyclic body of Isis-11 is largely stacked between G5 and A6 with the two aromatic protons in fairly close proximity to the aromatic face of G5. The H11 proton is positioned under the face of A6. See Figure 30.

In the R isomer, the body is also similarly stacked, and the two aromatic protons are also consistently placed near the face of G5, though not as directly as was the case with the S isomer. The placement of H11 is less ideally situated over an aromatic face, as it is facing away from A6 and is not situated over G5. The Amber refinement suggests that the two isomers bind with similar affinity. The visual similarity of the refined structures is consistent with this conclusion. See Figure 31.

While all titrations and NMR experiments were conducted with a racemic mixture of Isis-11, this mixture was later separated via chiral HPLC. While the individual isomers could not independently be characterized as S or R, independent fluorescence titrations were performed to measure the K_Ds. The two isomers were found to bind with very similar affinities (9.0 μM and 2.8 μM) in low salt conditions.

Comparison to Published Structure

The Puglisi structure indicates a significant bend in domain IIa.[27] This has been corroborated by crystallographic studies performed by the Hermann group.[26] Domain IIa is known to contain several specific magnesium binding sites, and the magnesium is known to stabilize the single stranded region involved in the domain IIa bulge. Residues 6, 7 and 8 of this single-stranded region are stacked on the G5-C34 base pair, resulting in a continuous helical arrangement of these three bases over the lower stem. The accommodation of this motif produces an 84.5 ± 9.6˚ bend. Our construct, in the absence of the Isis-11 compound, was found to adopt essentially the same bent structure both in the presence and absence of magnesium, based on RDC measurements. Our structure suggests that compound binding induces significant straightening of this domain IIa bulge. Several independently acquired RDC sets were only consistent with this straight structure. See Figure 32.

Compound binding impacts the stacking and arrangement of the bases in the bulge. The free RNA shows continuous stacking from the 5' terminus (G1 in our construct) through C8 with U9 largely flipped out. Compound binding appears to swing out the A7 through A10 portion of the bulge from its previous position stacked over A6 through the terminus in such a way that A7 is very distant from G5. Concomitantly, the previously flipped out U9 becomes stacked with A10 on the C11-G33 base pair. See Figure 33. The NMR RDC fits indicate that the helical regions on either side of the bulge become coaxial and the previously L-shaped free RNA becomes straight.

Discussion

It is beyond the scope of this study to definitively suggest a mechanism of inhibition for Isis-11 on HCV domain II; however, it is tantalizing to suggest a possible mechanism based on what is known about domain II and its involvement in IRES mediated translation. The straightening of the structure upon compound binding is possibly the mode of action for the inhibition of IRES function. Cryo-EM studies by Siridechadilok et al indicate some proximity between domain IIa and elements of eIF3, but mutational studies have indicated that this domain has a negligible impact on eIF3 recruitment.[23] The only portion of domain II that is known to have a specific role in IRES function is the apical loop of domain IIb. While the compound does not directly interact with the apical loop portion of domain II, the conformational change in the RNA would move the domain terminus. As the lower portion of domain II is constrained by its attachment to the tightly bound domain III, it is immobilized in the translation initiation complex. The straightening of domain II should displace the apical loop from its position over the mRNA cleft at the ribosomal E site. This displacement could keep the apical loop from interacting with the ribosome and facilitating GTP hydrolysis and eIF2 release prior to 60S recruitment. See Figure 34. In this sense, Isis-11 could be considered an allosteric drug if we think of the domain IIb terminus as the "active site" of domain II. By binding in domain IIa, which is distal to the "active site", a conformational change is induced in the "active site" that prevents the "active site" from performing its catalytic function, which is to induce a structural change in the 40S subunit to facilitate a GTP hydrolysis event.

The only alternative mechanism of action that I can conceive is that the compound impedes the binding of initiation factors or binding to the 40S subunit. However, this seems unlikely given that domain II has been shown to be unnecessary for 40S binding, and it is not known to bind to any initiation factors.

Multiple sets of RDC restraints were shown to be only consistent with the straight structure. Recently published results from the Hermann lab have also suggested a conformational change in domain IIa upon inhibitor binding.[86] These FRET assays were performed with a similar tight binding benzimidazole compound first described by Isis Pharmaceuticals (compound 13 in the published account, Isis-13 hereafter)[54]. The conformational widening of the inter-helical angle by 23º C indicated by their data corroborates our general findings. The Hermann group also performed titrations of Isis-13 into an A7 2-aminopurine substituted IRES construct sample. In agreement with our findings with Isis-11 and a similar 2-aminopurine substituted IRES construct, they found that fluorescence, and thus destacking at the A7 position, increased upon Isis-13 addition.[86]

The binding of Isis-11 appears to be driven by stacking interactions as well as by cation-pi interactions. The tricyclic, largely aromatic body of Isis-11 is stacked between A6 and G5. The primary amine may be involved in a cation-pi interaction with G33. The tertiary amine on the short methylene arm may also be involved in a cation-pi interaction with A6. Although the architecture in and around the binding site has changed significantly from the free RNA, the placement of these amines in the complex may loosely approximate the placement of specific Mg^{2+} ions seen in the free structure.[26] See Figure 35.

It is interesting to note that electrostatic considerations are an important part of the binding interaction, but not in the obvious way we might have anticipated. The negative charge on the phosphate backbone is largely canceled out by the high concentration of monovalent counter cations. This leads to the predominance of more subtle negative charges elsewhere on the RNA that can interact with the positively charged amines on Isis-11. Figure 36 shows the electrostatic potential surface calculated for the RNA alone using the structure of the complex. As shown in the figure, the two tertiary amine groups of Isis-11 nestle into a region of predominantly negative potential, even though they are not interacting with the phosphate backbone.

It has been suggested that by Begley and Varani in describing the results of Parsons et al. that small molecules, such as the one in this study, "that binds the hepatitis C viral genome by 'locking' in a particular RNA conformation to inhibit viral protein production suggests a new paradigm for drug design."[87] They point out several advantages to the approach of directly targeting viral RNA, specifically the HCV IRES, but analogies could be made to other viral RNA systems. Firstly, the 5'-UTR of HCV contains some of the most highly conserved genome sequences of an otherwise highly mutation prone viral family. Secondly, the unique conformational flexibility of RNA, so crucial to its interactions with various cellular targets, becomes the point of attack. The abolition of this flexibility represents a heretofore-unexploited arena for viral inhibition. Thirdly, by targeting the RNA itself rather than attempting to disrupt the interaction between the RNA and a protein, the small molecule does not have to compete energetically or sterically with the protein.

In summary, this structure suggests the possibility of a novel mode of inhibition of hepatitis C virus. More broadly, this structure suggests the possibility of a new approach to drug design. This structure may also provide the basis for the future improvement of the binding affinity of this inhibitor class as later compounds are designed to better fit the binding site and take better advantage of potential energetically favorable binding interactions.

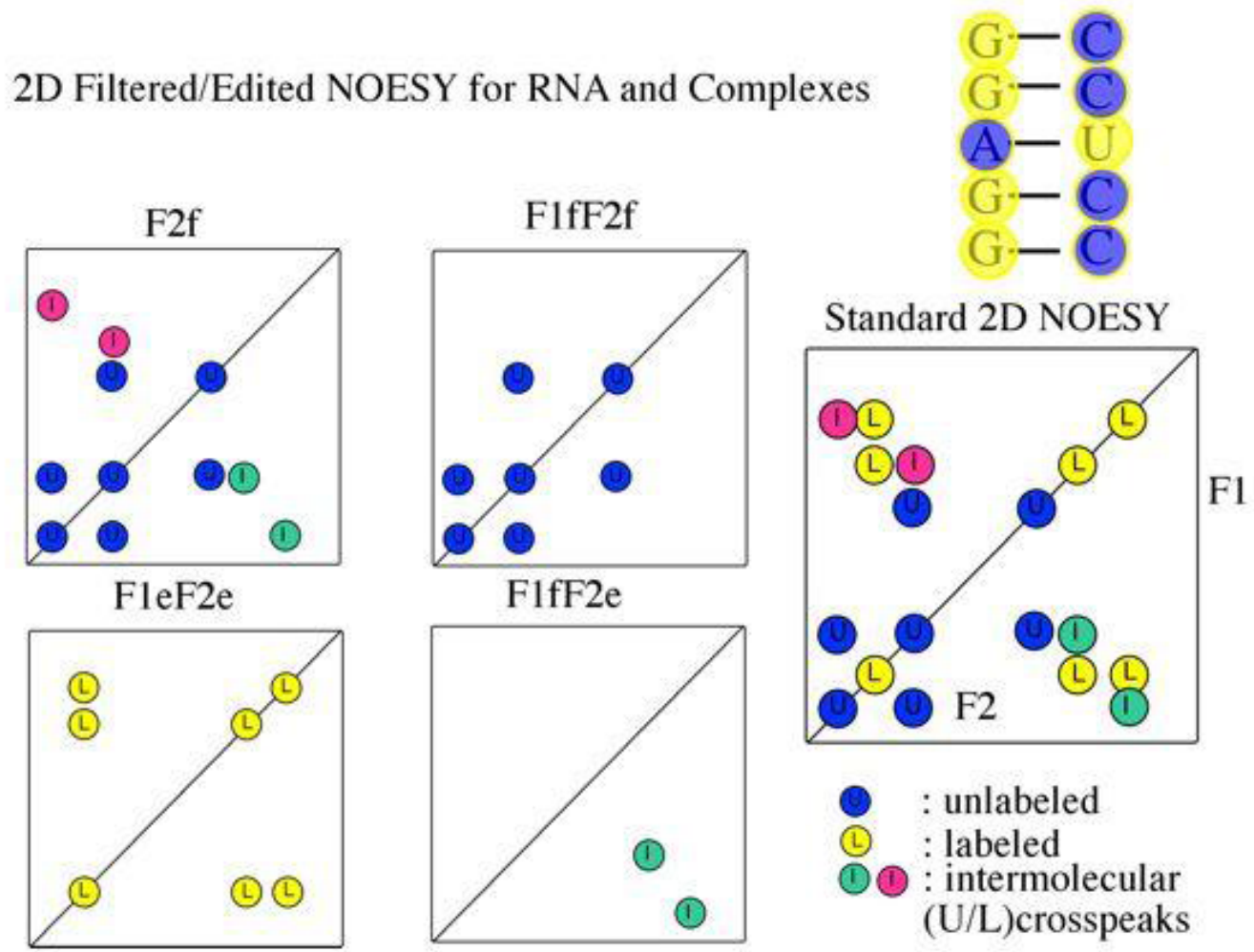

Figure 16: Filter/edit NOESY suite of experiments. Adapted from Peterson, Theimer, Wu and Feigon (2004).

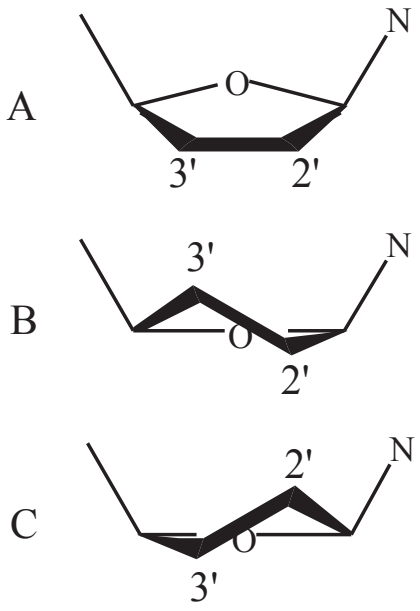

Figure 17: Common sugar puckers in RNA. (A) Planar sugar ring, never observed. (B) C_3'-endo sugar pucker. This conformation is the norm in A form helical RNA. (C) C_2'-endo sugar pucker. This conformation is also found in RNA, but usually only where the A-form helicity has been disrupted.

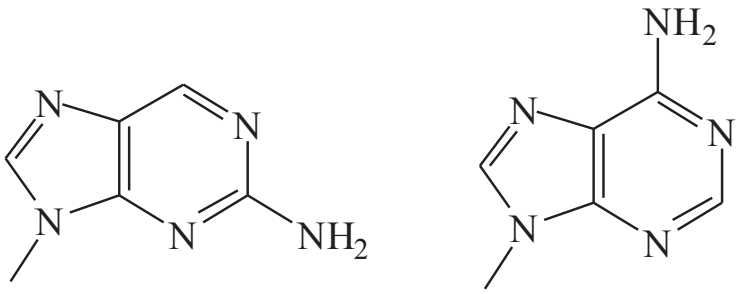

Figure 18: 2-aminopurine and adenine structures. 2-aminopurine nucleosides were used in lieu of selected adenosines in order to probe changes in the stacking of the IRES5 construct upon Isis-11 binding through monitoring of changes in fluorescence.

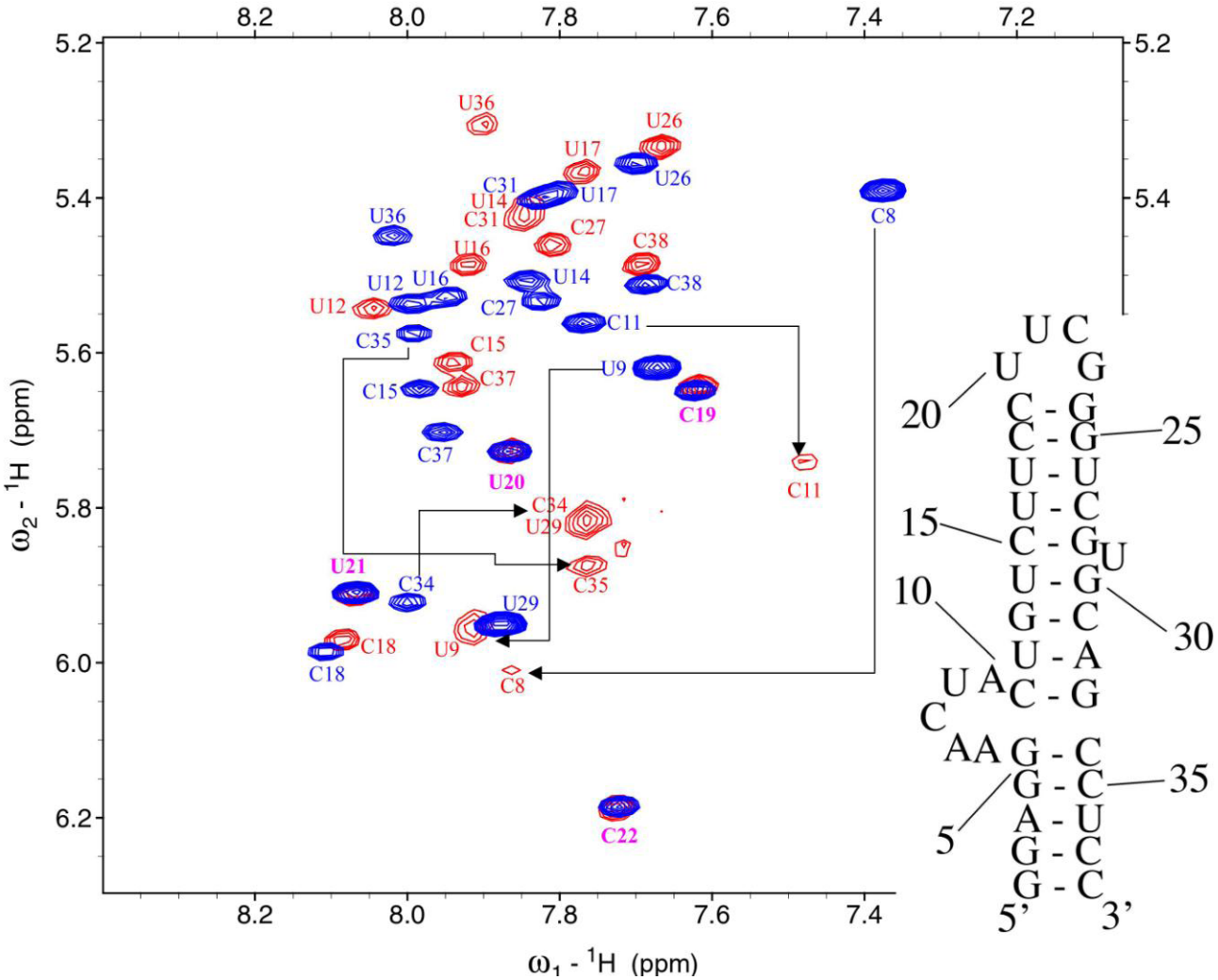

Figure 19: TOCSY spectrum of free and bound IRES5. The cross peaks with negligible movement are labeled in violet. The free spectrum is shown and labeled in blue while the bound spectrum is shown and labeled in red. Some of the peaks that show the most movement are highlighted with black arrows. Adapted from Paulsen et al (2010).

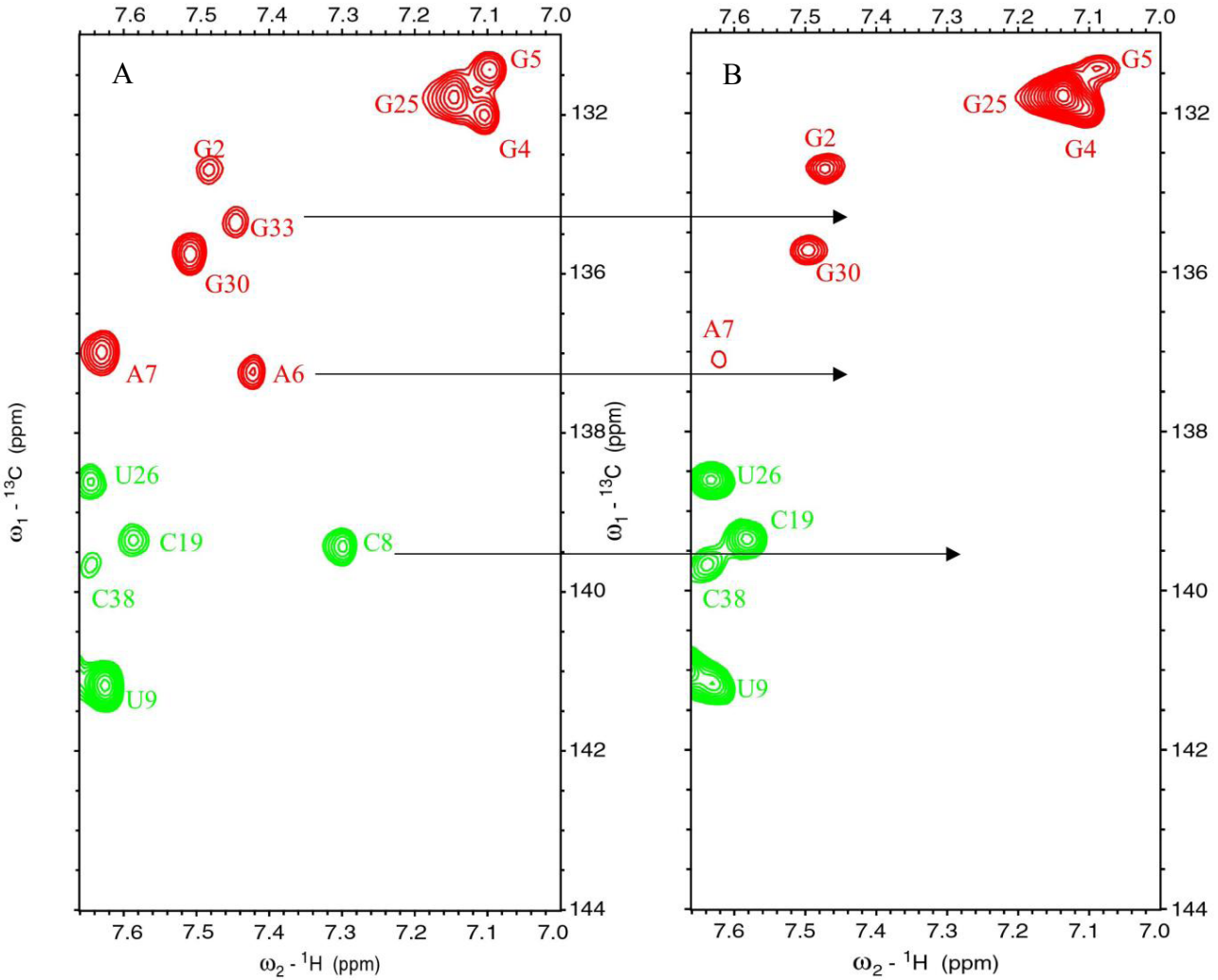

Figure 20: Specific line broadening in the complex. Low equivalents of Isis-11 lead to specific line broadening. A) Portion of the ^{13}C HSQC spectrum in the absence of Isis-11. B) Portion of the ^{13}C HSQC spectrum in the presence of ~0.5 equivalents of Isis-11.

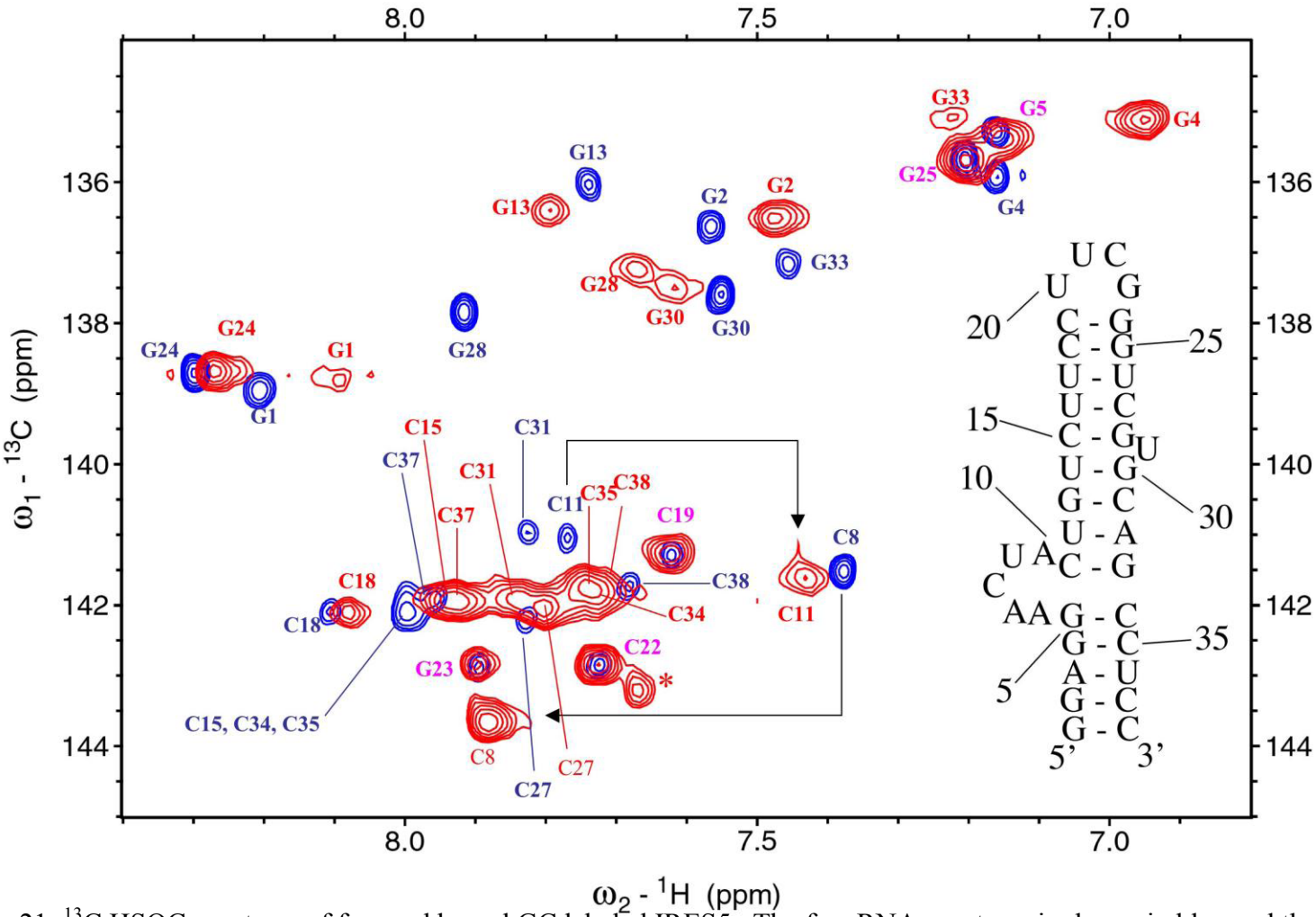

Figure 21: ^{13}C HSQC spectrum of free and bound GC labeled IRES5. The free RNA spectrum is shown in blue and the bound RNA spectrum is shown in red. Peaks that overlap in the free and bound forms are labeled in violet. The asterisk indicates an artifact peak (a peak that is absent at low RNA concentrations). Adapted from Paulsen et al (2010).

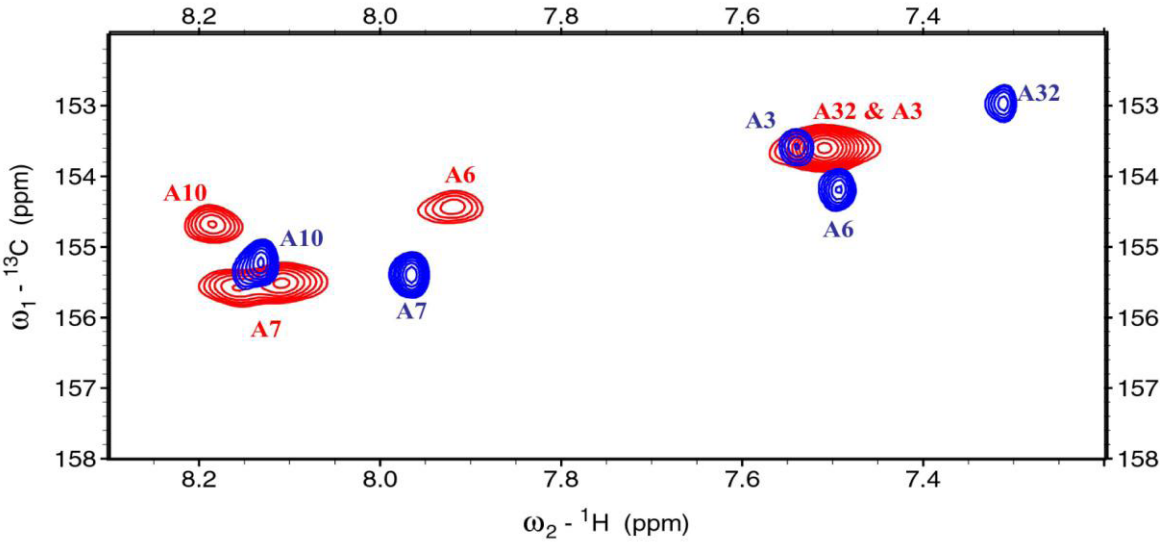

Figure 22: The H2 region of the ^{13}C HSQC spectrum of free and bound IRES5. The free RNA spectrum is shown and labeled in blue and the bound RNA spectrum is shown and labeled in red.

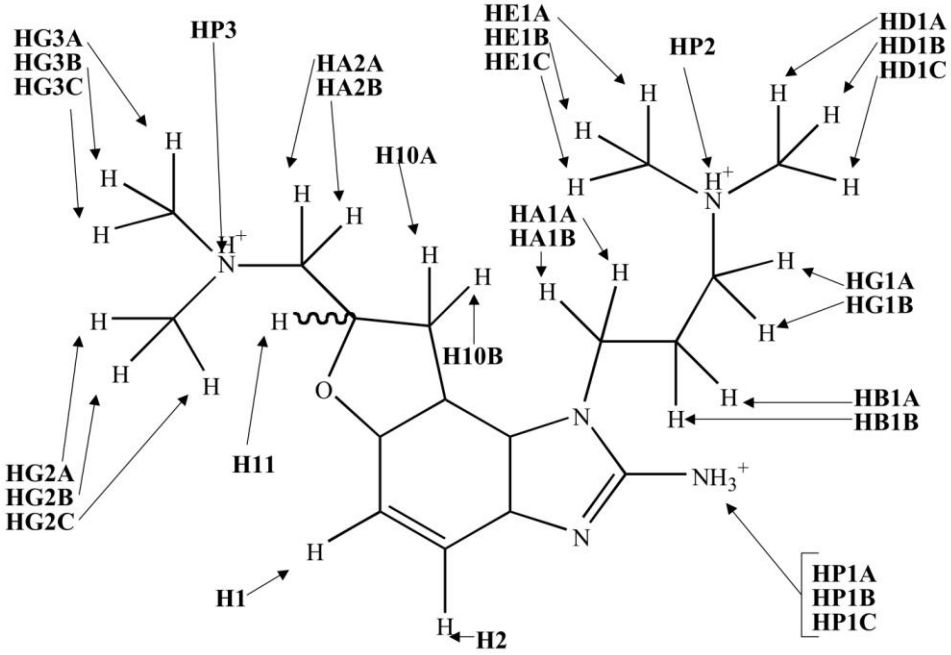

Figure 23: Proton labeling scheme for Isis-11.

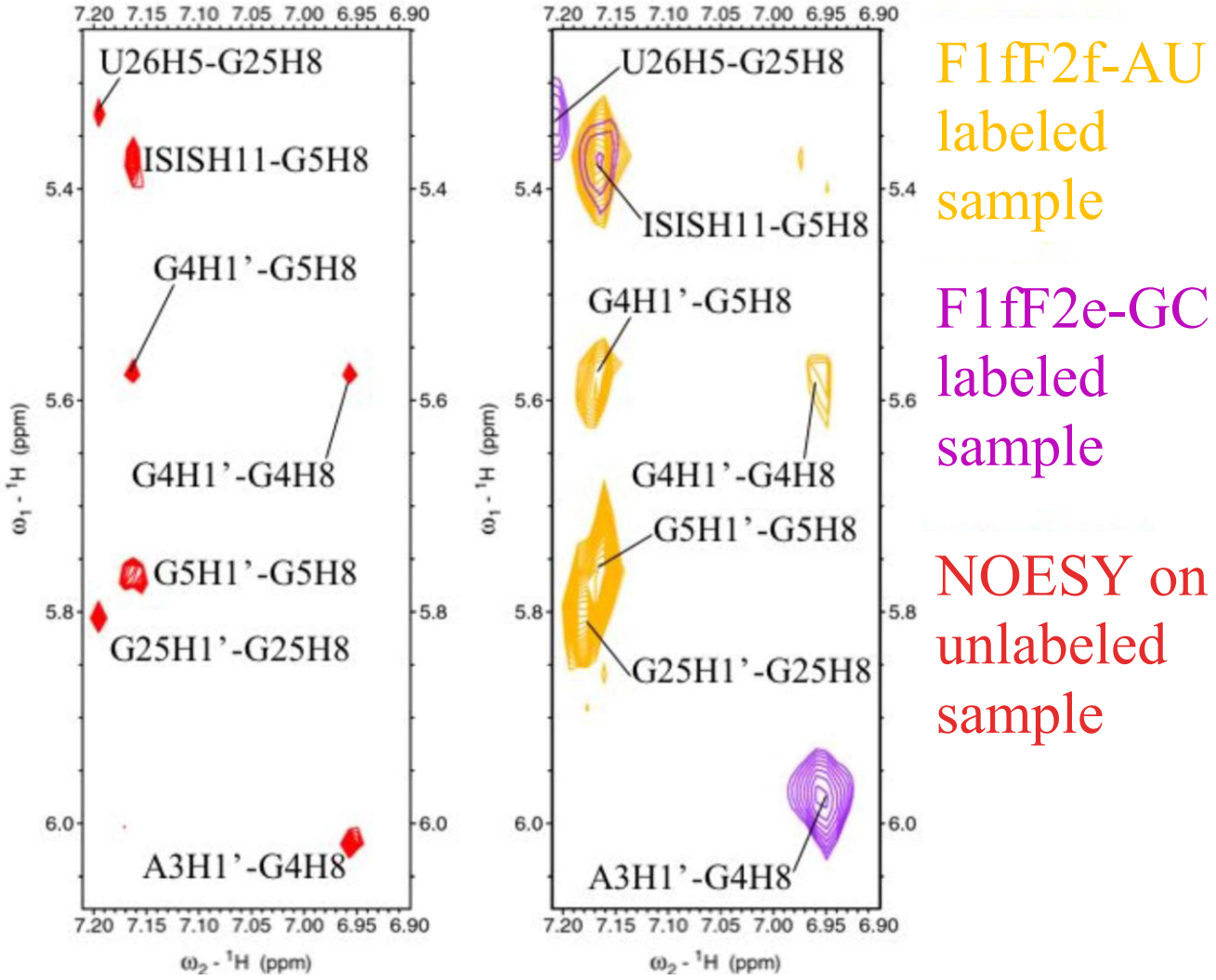

Figure 24: Application of filter/edit suite of experiments to IRES5. The different data sets provide a clear distinction between the unlabeled to unlabeled (F1fF2f) and unlabeled to labeled (F1fF2e) NOE cross peaks. Adapted from Paulsen et al (2010).

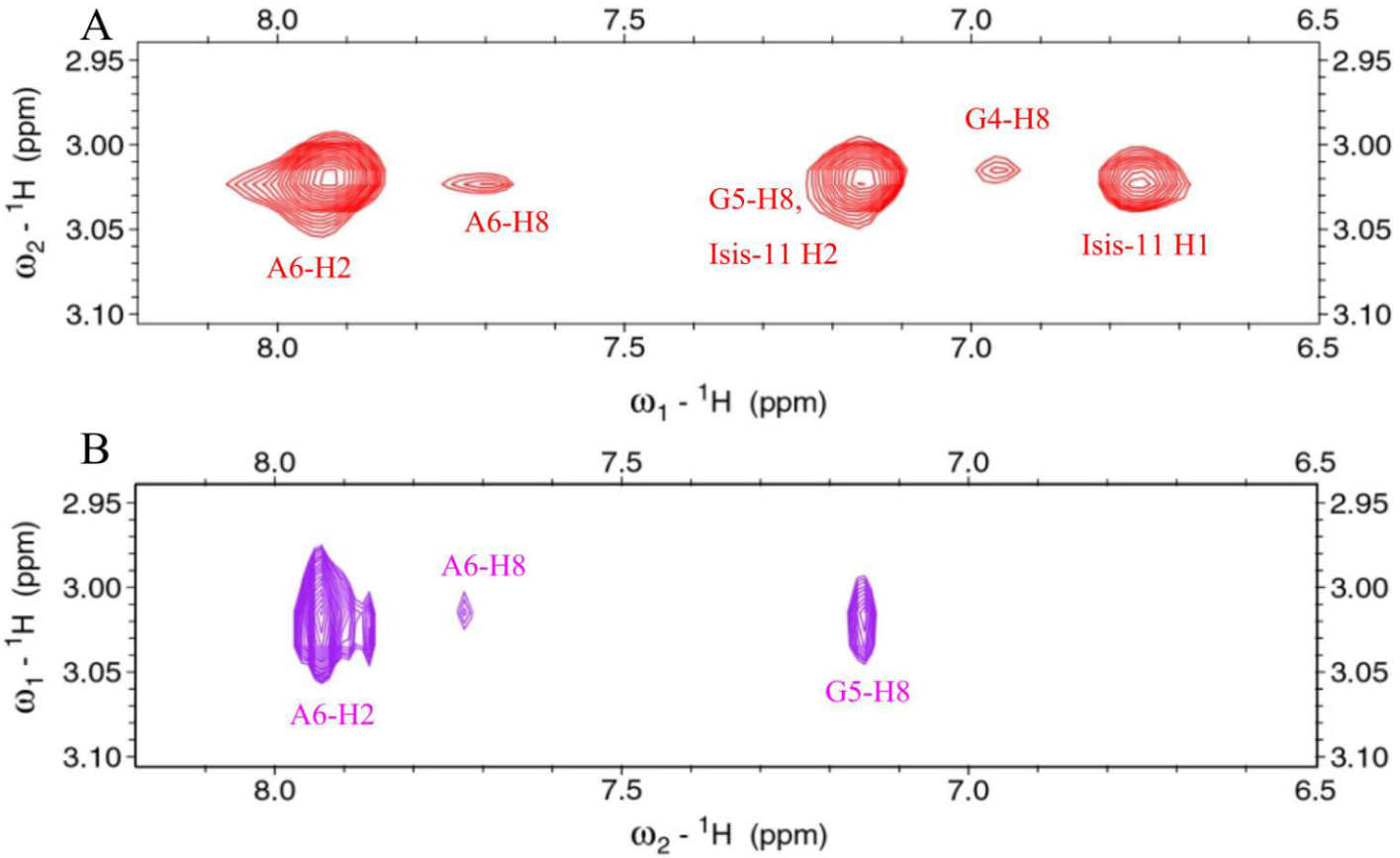

Figure 25: NOEs to the HG3A-C and HG2A-C methyl proton set. A) Selected region of the NOESY spectrum of the unlabeled IRES5-Isis-11 complex. B) Selected region of the F1fF2e NOESY spectrum of the fully labeled IRES5-Isis-11 complex. The F1 'filter' purge of the F1fF2e experiment yields NOEs only where the ω_1 frequency corresponds to the unlabeled Isis-11 proton. One intermolecular NOE (to G5-H8) and one intramolecular NOE (to Isis-H2) overlap. The F1fF2e experiment is less sensitive than the conventional NOESY, and weaker intermolecular NOEs do not appear in the F1fF2e spectrum. Adapted from Paulsen et al (2010).

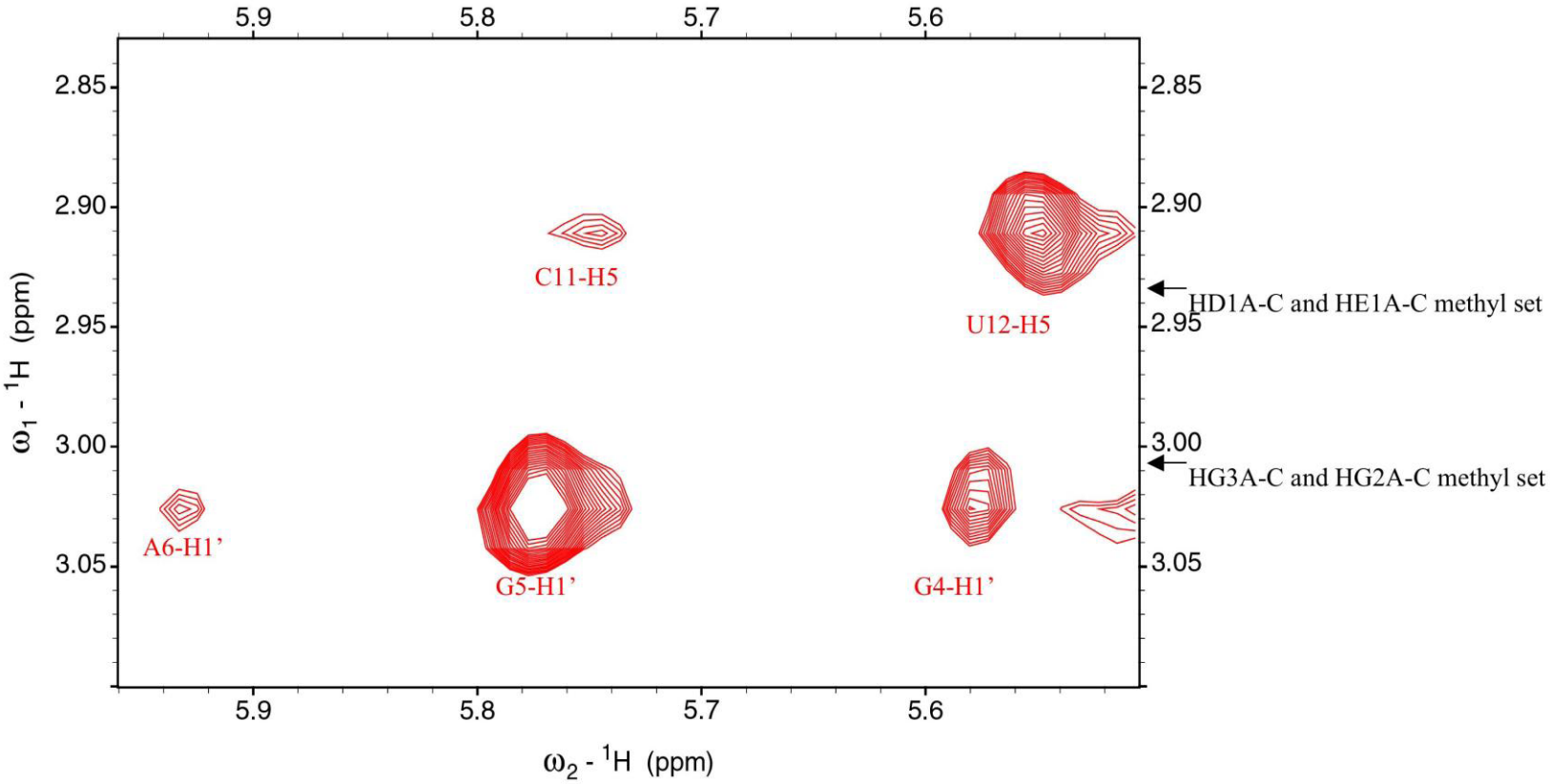

Figure 26: Intermolecular NOEs from the HD1A-C and HE1A-C methyl set and the HG3A-C and HG2A-C methyl set to IRES5. The data shown are from a NOESY spectrum on an unlabeled complex. The resolution of these two Isis-11 methyl proton sets can clearly be seen in this figure. Adapted from Paulsen et al (2010).

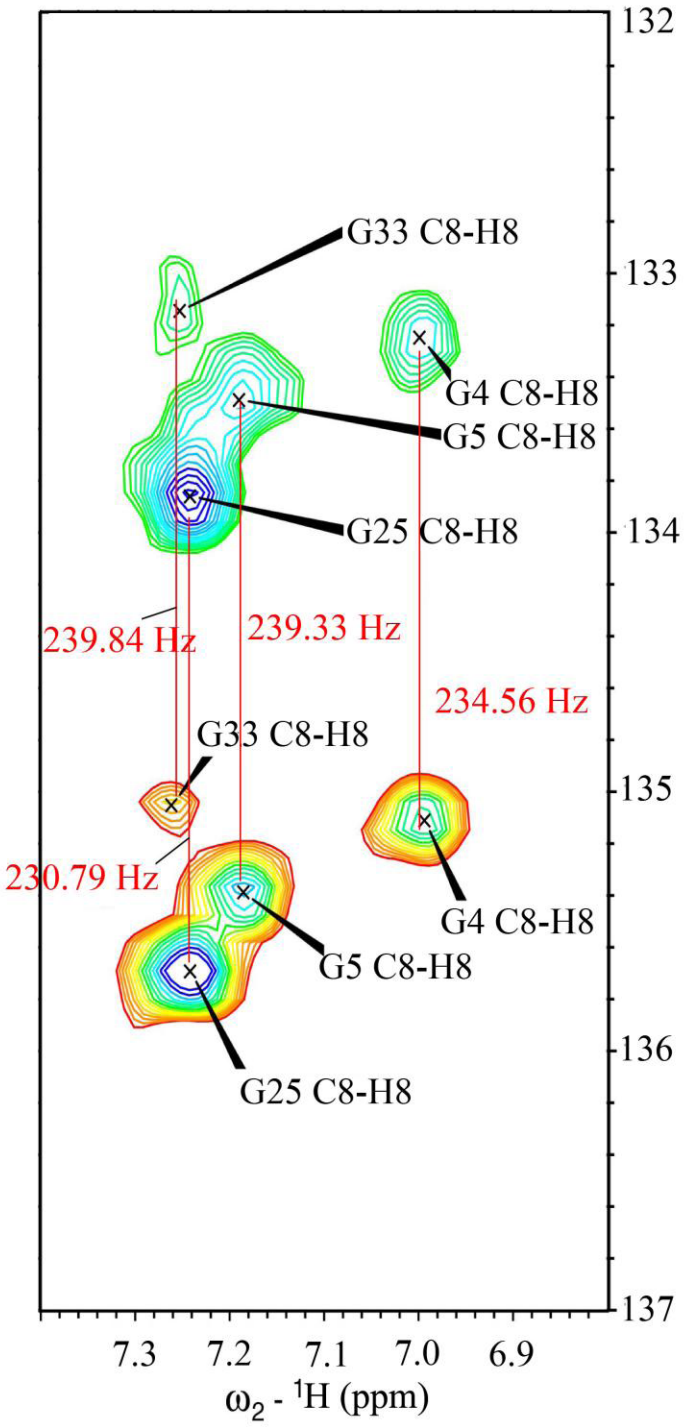

Figure 27: Residual dipolar couplings of the bound IRES5. Residual dipolar coupling measurements for a selection of guanosine C8-H8 bond vectors in the aligned complex sample. The unaligned $^{1}J_{CH}$ value for the C8-H8 vector in guanosine is 214.9 Hz.

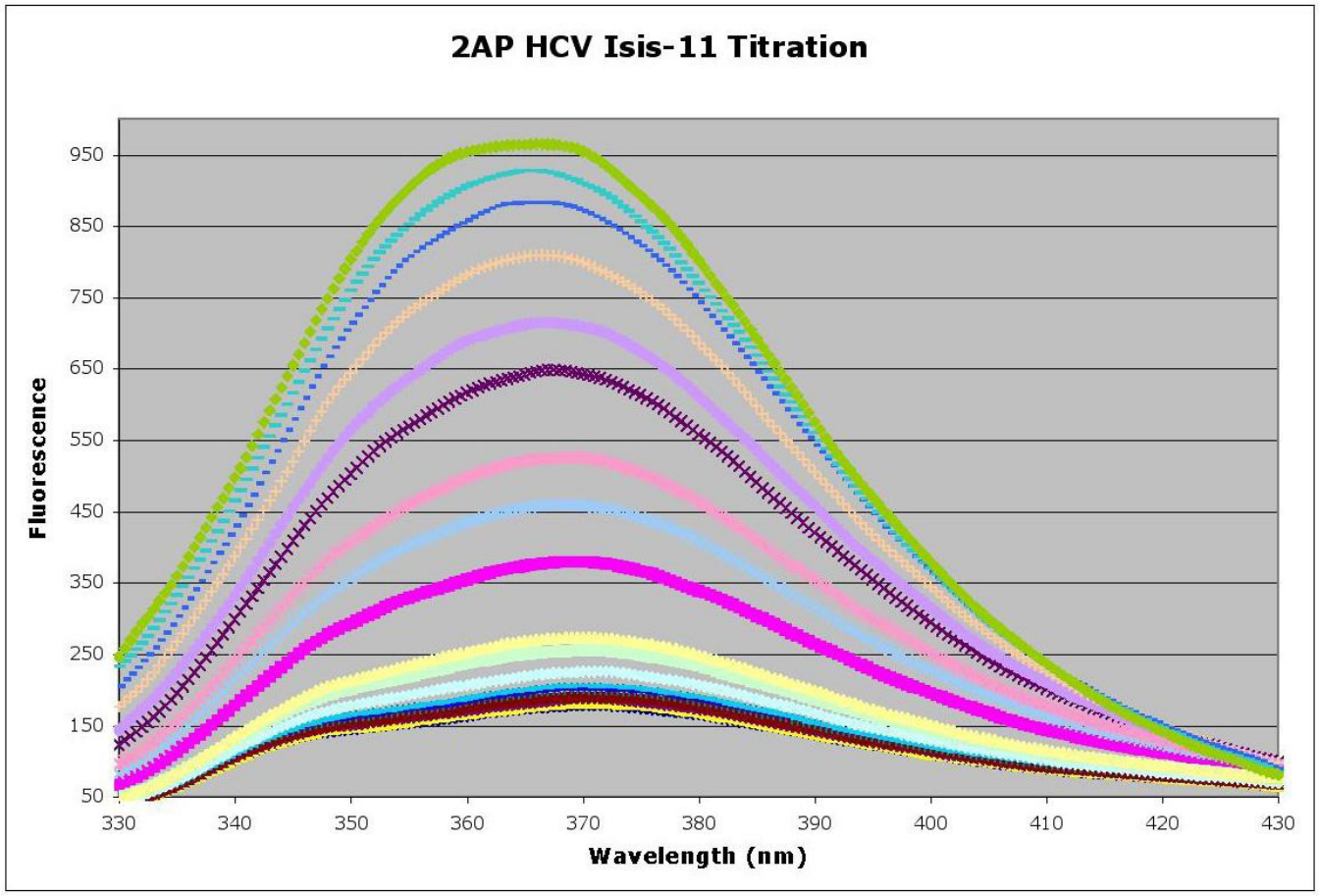

Figure 28: Fluorescence spectroscopy of IRES RNA with 2-aminopurine at position A6. The RNA concentration is 1 μM in 10 mM sodium cacodylate, pH 7.0 and 50 mM NaCl. Isis-11 was titrated at 0.6, 12, 24, 36, 66, 126, 186, 246, 546, 846, 1146, 1746, 2346, 2946, 2546, 4146 and 4746 nM. Fluorescence response increased with inhibitor concentration, consistent with decreased stacking upon compound binding. The fluorescence intensity at 370 nm was fit to a simple two-state binding model to yield a K_D of 2.4 μM. Adapted from Paulsen et al (2010).

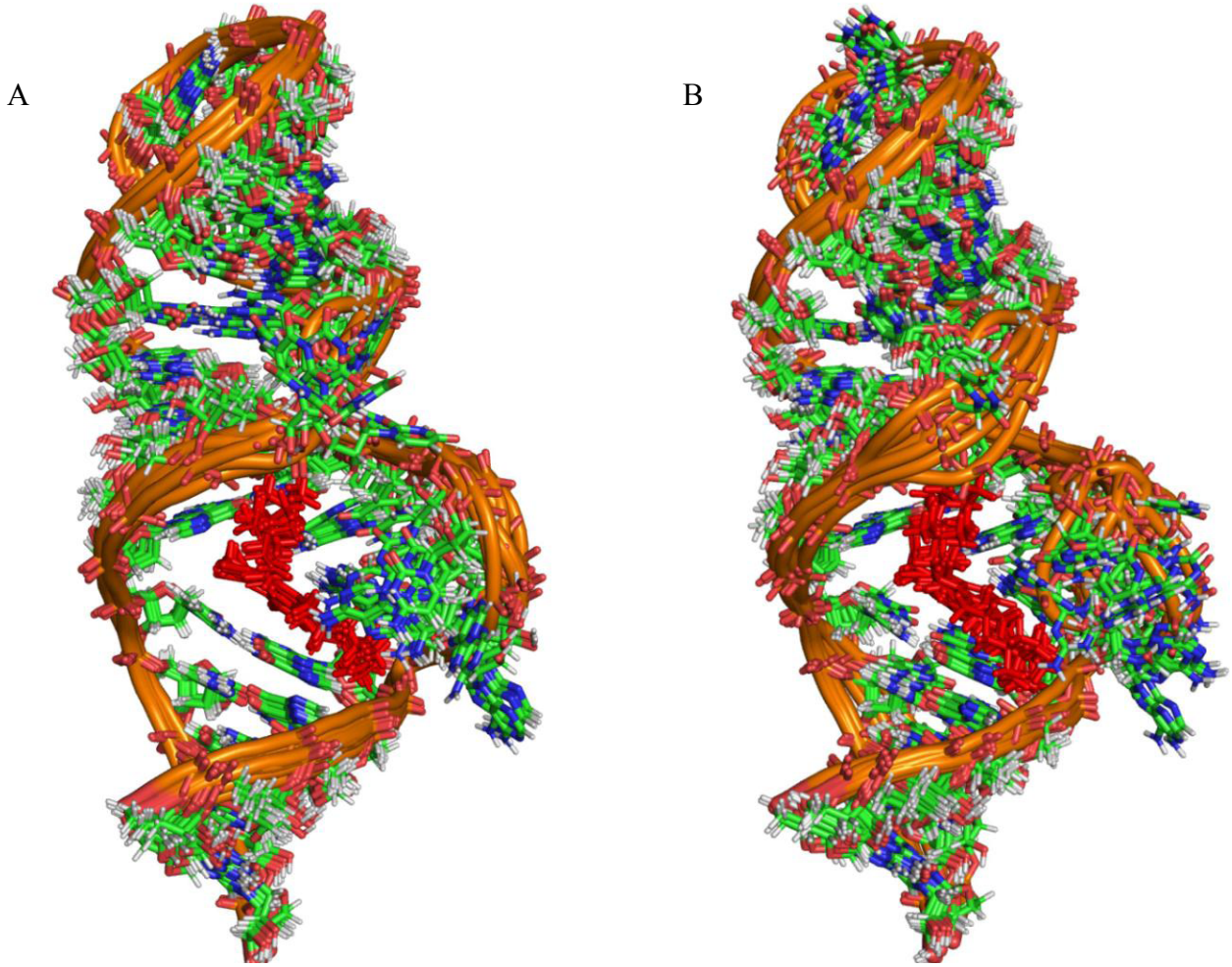

Figure 29: Superpositions of IRES5-Isis-ll complex structures. A) Superposition of 10 lowest energy structures of the R isomer of Isis-11. B) Superposition of 10 lowest energy structures of the S isomer of Isis-11.

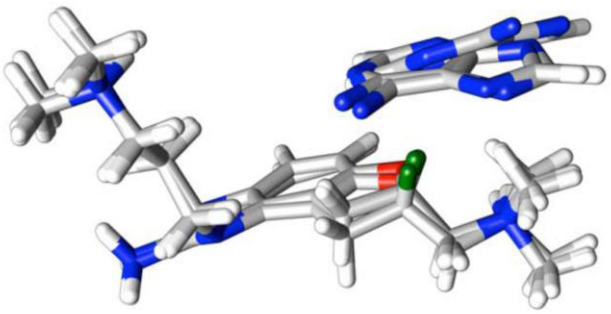

Figure 30: A6 - Isis-11 interactions consistent with upfield shift. Molecular graphics highlight the favorable orientation of A6 to the chiral proton of the S-isomer. This orientation is consistent with the upfield shift upon compound binding. The averaging of the NOEs due to both isomers being present appears to promote structures where A6 stacks over the benzimidazole rather than structures where A6 is involved in a cation-pi interaction with the dimethylamino group that would shield the H11 proton. Adapted from Paulsen et al (2010).

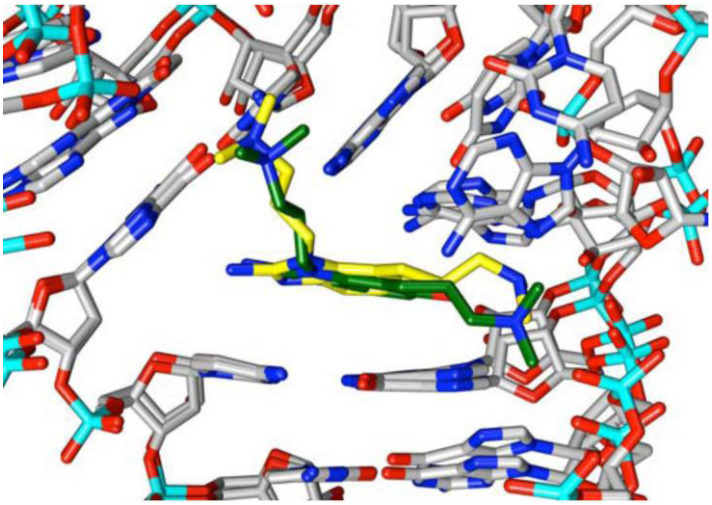

Figure 31: R and S isomer binding modes. Molecular graphics (local fit) of the S-isomer and R-isomer for Isis-11 after MD refinement in explicit solvent with the ff99SB+parmbsc0 force field. Shown are the lowest energy structures from a local fit over residues 5, 10, 11, 33, and 34 for each isomer (colored yellow and green). Adapted from Paulsen et al (2010).

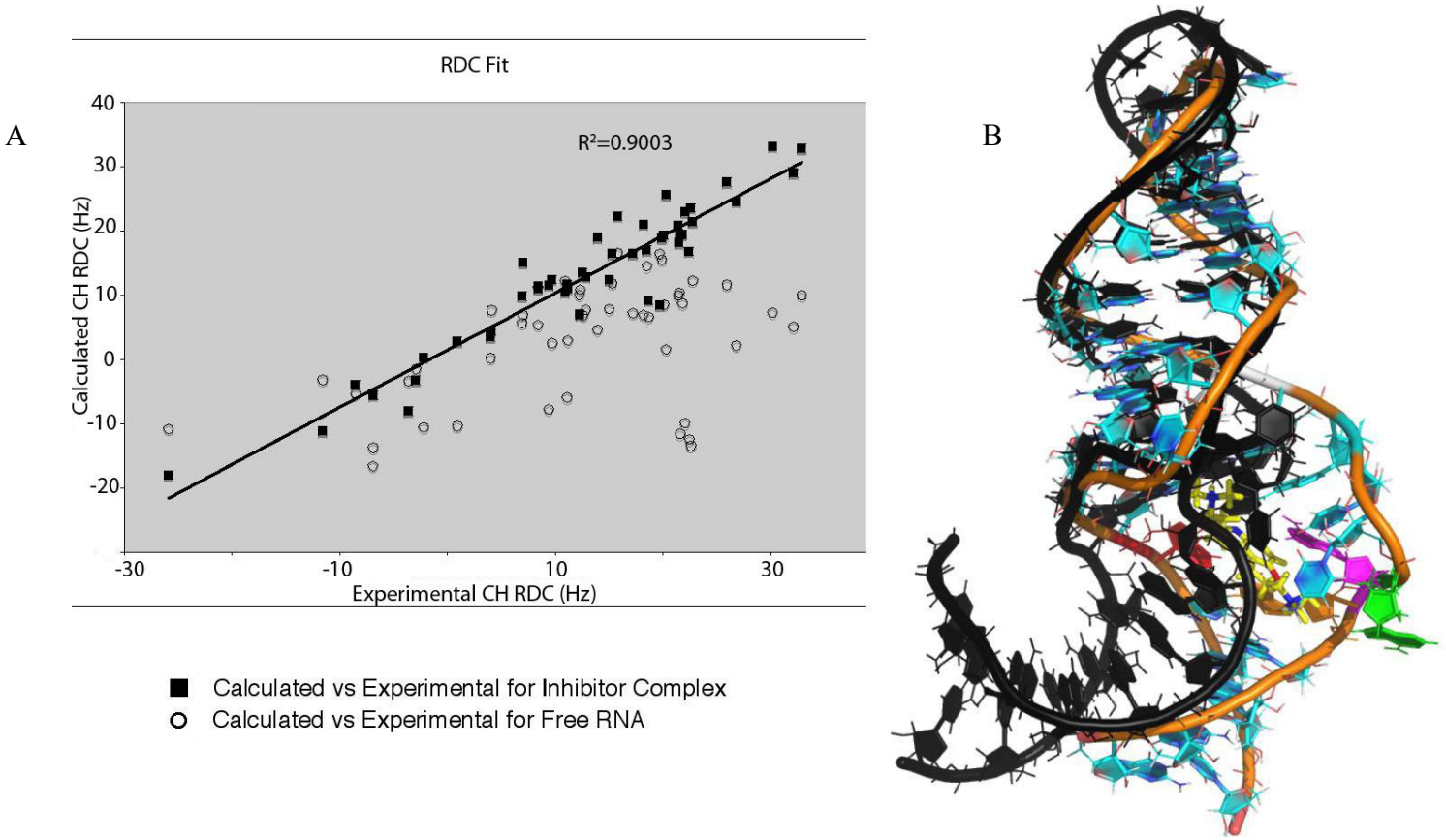

Figure 32: NMR structures and residual dipolar coupling data. (A) RDC fit. CH RDCs were calculated from a representative final complex structure and from the free RNA starting structure; each set of calculated RDCs is plotted against the experimental RDC set for the complex. Filled squares are for the complex, and open circles for the free RNA. (B) The free RNA model (black) is shown superimposed (residues 12:19,24:32) on the complex structure (orange backbone) illustrating the change in the helical axis. The inhibitor is shown in yellow. Adapted from Paulsen et al (2010).

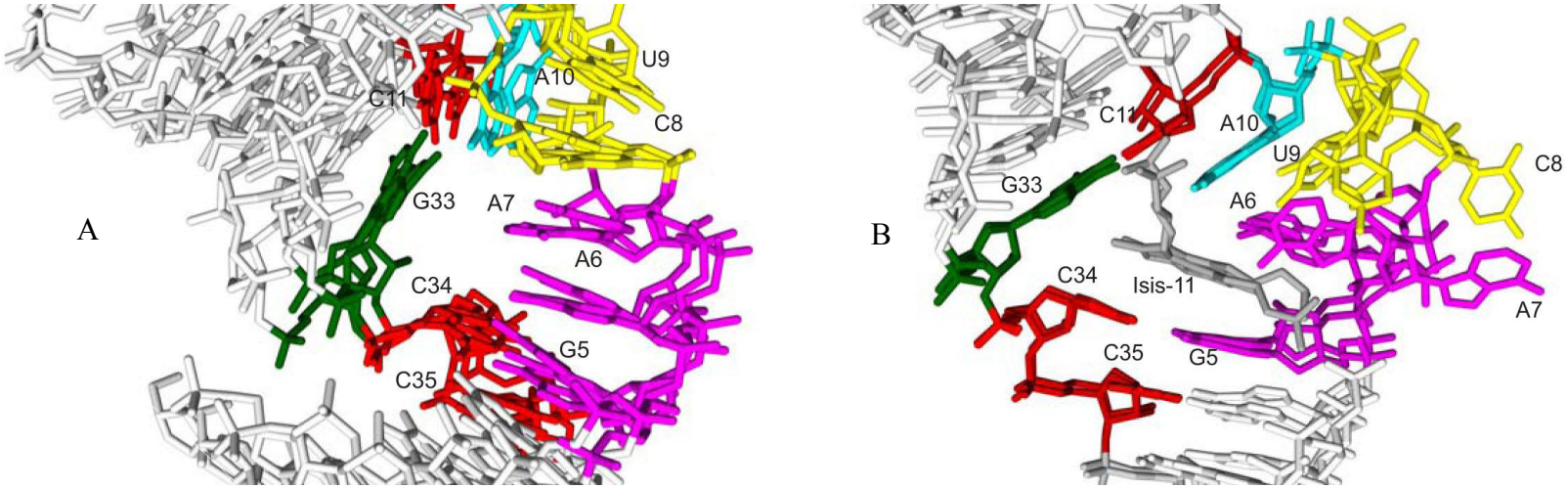

Figure 33: Bulge base stacking adjustments in complex. (A) Molecular graphics of the heavy atoms of the free first three models from PDB 1P5M and (B) Isis-11 bound RNA structures (the two lowest energy water-refined structures from each stereoisomer). The residues whose environment changes upon binding are labeled and colored (C11, C34, C35: red; C8, U9: yellow; G5, A6, A7: magenta; A10: cyan; G33:green; Isis-11: gray). The structures were fit using residues 5-11, 33-34 for the free structure and residues 5, 10, 11, 33, 34 for the bound structure. Adapted from Paulsen et al (2010).

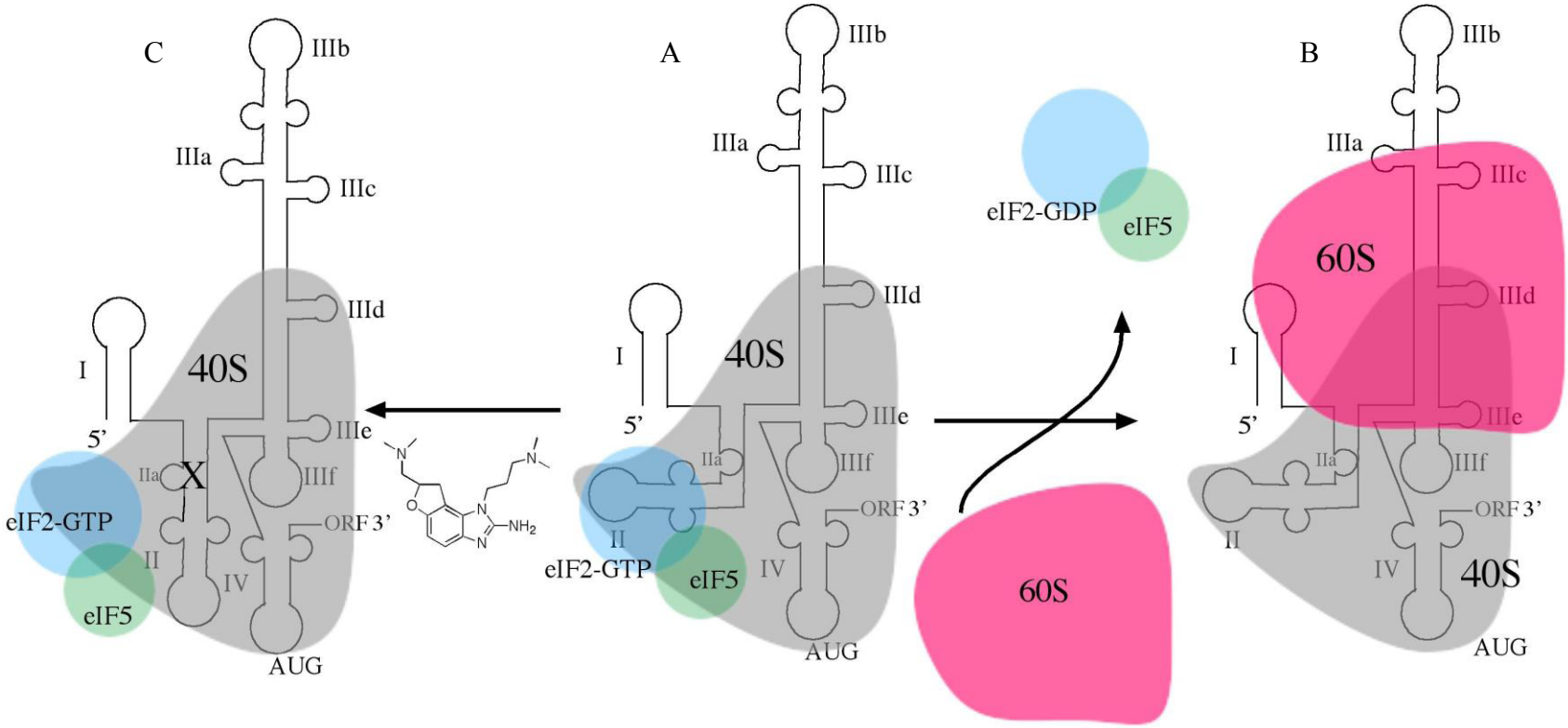

Figure 34: Proposed mechanism of inhibition. (A) Domain IIa of the HCV IRES has a 90° bend that positions the apical loop of domain II in contact with the 40S subunit near the E-site, and along with eIF5, promotes hydrolysis of eIF2-GTP and subsequent release of eIF2-GDP. (B) The release of eIF2-GDP allows recruitment of the 60S subunit and eIF3 release (not shown) to assemble a translation competent 80S ribosome. (C) The inhibitor (X) would straighten IRES domain II, reduce the eIF2-GTP hydrolysis rate, and slow 80S assembly. Adapted from Paulsen et al (2010).

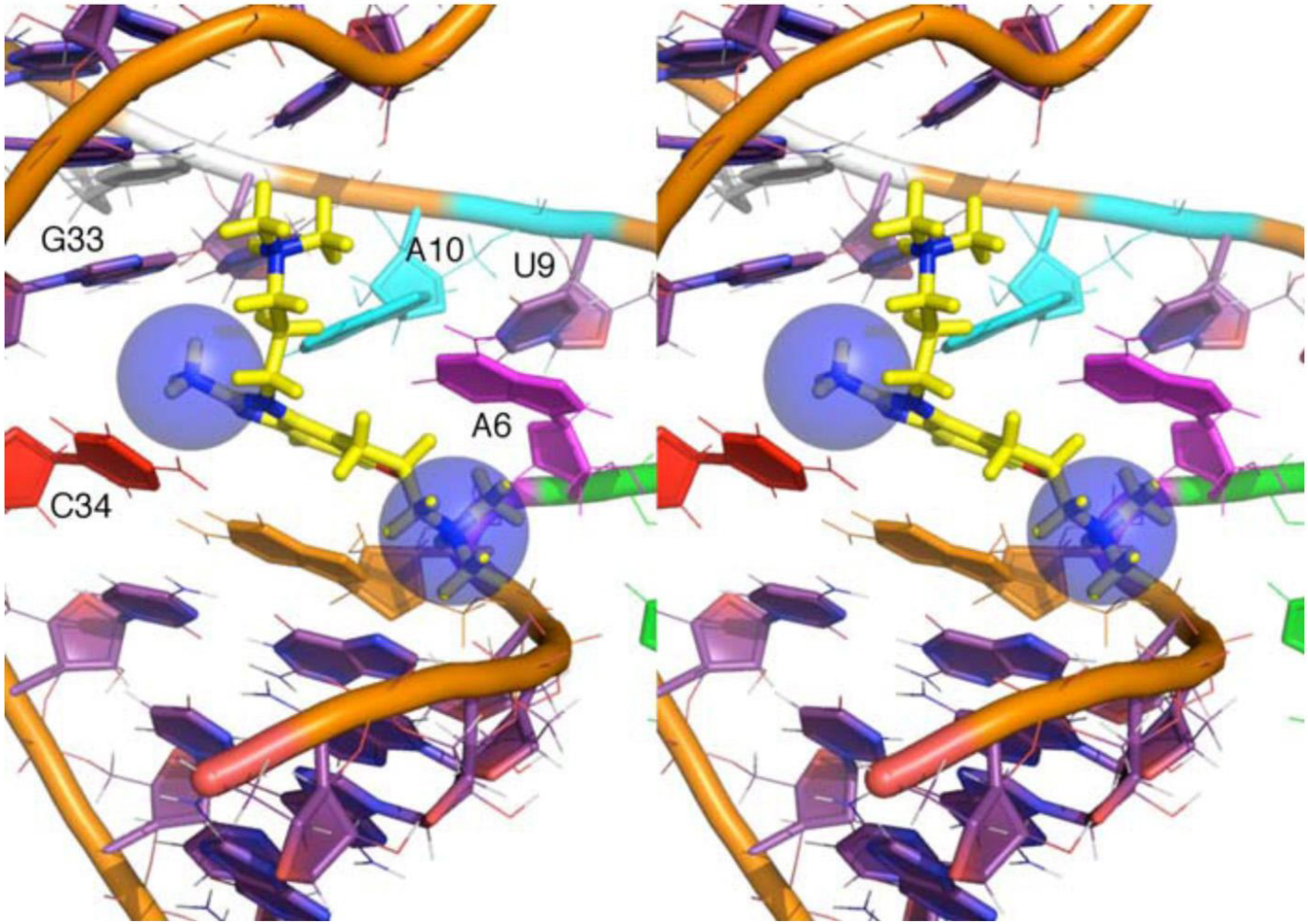

Figure 35: Stereo image of potential cation-pi interactions between IRES5 and Isis-11. Possible cation-pi interactions are highlighted with blue spheres. Adapted from Paulsen et al (2010).

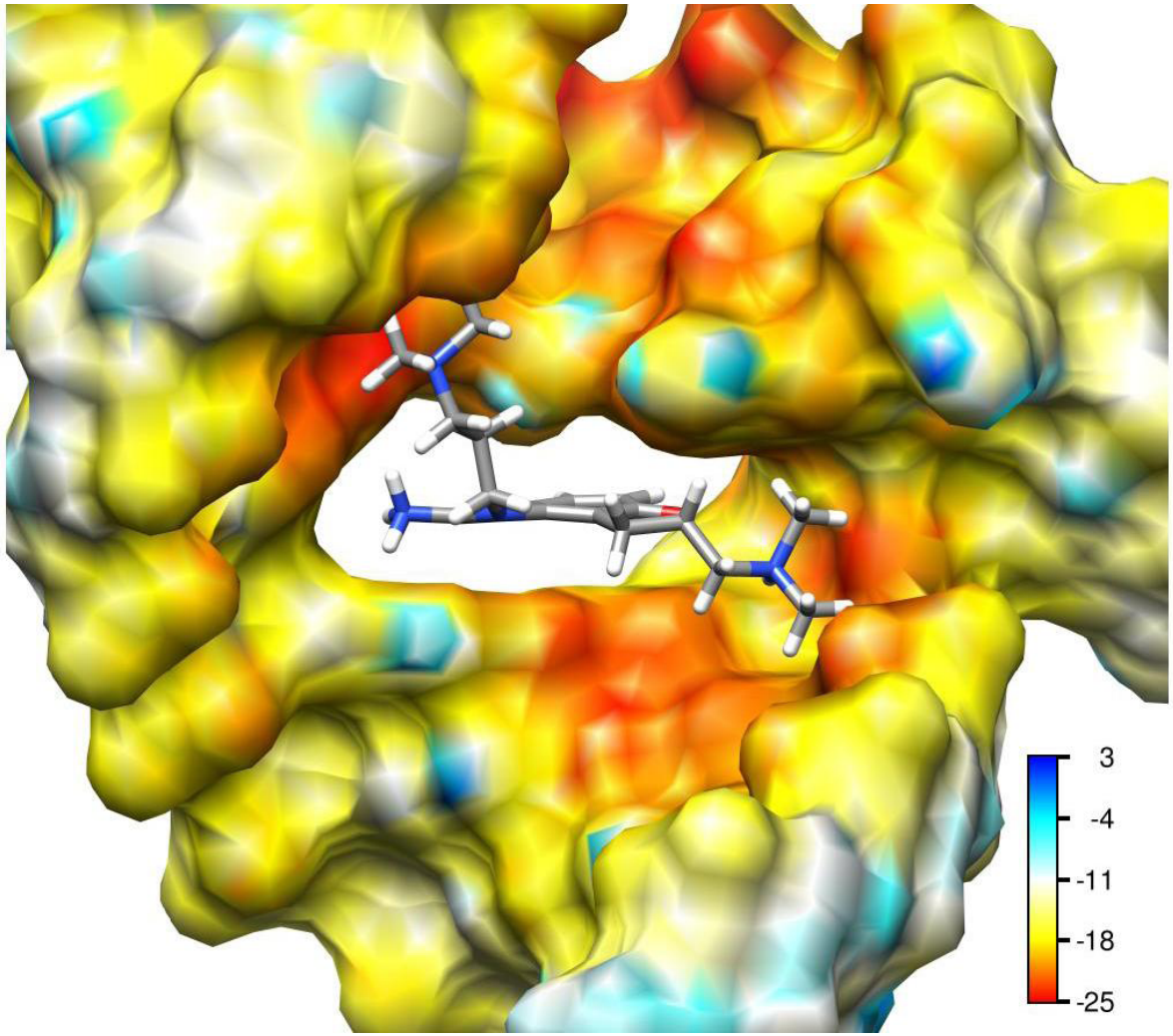

Figure 36: Electrostatic surface potential of the complex. The electrostatic potential surface was calculated for the RNA only using the structure of the complex. Areas of highest negative charge are colored red and are localized near the dimethylamino groups. Adapted from Paulsen et al (2010).

Table 2

Complexed RNA ^{1}H chemical shifts (ppm)

Residue	H8/H6	H5/H2	H1'	H2'	H3'	H4'	H5'	H5''	imino
G1	8.15	-	5.75	4.89	4.68	4.57	n.d.	4.30	n.d.
G2	7.50	-	5.89	4.64	4.90	n.d.	n.d.	n.d.	12.25
A3	7.81	7.51	6.01	4.70	4.69	n.d.	n.d.	n.d.	-
G4	6.96	-	5.58	4.47	4.31	n.d.	4.41	4.05	12.64
G5	7.16	-	5.77	4.15	4.47	n.d.	n.d.	n.d.	12.12
A6	7.75	7.92	5.94	4.11	4.31	n.d.	n.d.	n.d.	-
A7	8.43	8.12	6.08	4.86	4.91	n.d.	n.d.	n.d.	-
C8	7.87	6.02	5.75	4.05	4.84	n.d.	n.d.	n.d.	-
U9	7.92	5.99	6.01	4.51	4.80	n.d.	n.d.	n.d.	n.d.
A10	8.57	8.18	6.14	4.86	4.77	n.d.	n.d.	n.d.	-
C11	7.45	5.75	5.09	4.33	4.86	n.d.	n.d.	n.d.	-
U12	8.04	5.55	5.79	4.95	4.55	n.d.	n.d.	n.d.	13.84
G13	7.78	-	5.89	4.76	n.d	n.d.	n.d.	n.d.	12.53
U14	7.85	5.41	5.51	4.29	4.47	n.d.	n.d.	n.d.	12.09
C15	7.93	5.61	5.54	4.39	4.44	n.d.	n.d.	n.d.	-
U16	7.92	5.49	5.52	4.01	n.d	n.d.	n.d.	n.d.	n.d.
U17	7.77	5.36	5.49	4.29	4.51	n.d.	n.d.	n.d.	10.62
C18	8.08	5.97	5.72	4.52	n.d	n.d.	n.d.	n.d.	-
c19	7.62	5.64	5.57	4.62	4.19	n.d.	n.d.	n.d.	-
u20	7.87	5.73	5.66	3.84	4.57	n.d.	n.d.	n.d.	11.77
u21	8.07	5.91	6.11	4.69	4.08	n.d.	4.28	n.d.	n.d.
c22	7.73	6.19	5.98	4.15	4.55	n.d.	n.d.	n.d.	-
g23	7.90	-	6.00	4.92	5.62	n.d.	n.d.	n.d.	9.73
g24	8.28	-	4.61	n.d.	4.35	n.d.	n.d.	n.d.	12.81
G25	7.20	-	5.81	4.71	4.32	n.d.	n.d.	4.05	13.30
U26	7.66	5.34	5.44	4.13	4.42	n.d.	n.d.	n.d.	11.36
C27	7.81	5.45	5.52	3.79	4.59	n.d.	n.d.	n.d.	-
G28	7.73	-	5.85	4.67	4.81	n.d.	n.d.	n.d.	12.46
U29	7.76	5.79	5.92	4.42	4.68	n.d.	n.d.	n.d.	n.d.
G30	7.57	-	5.72	4.52	4.57	n.d.	n.d.	n.d.	11.39
C31	7.85	5.41	5.56	4.52	n.d	n.d.	n.d.	n.d.	-
A32	8.02	7.50	6.08	4.87	4.76	n.d.	n.d.	n.d.	-
G33	7.23	-	5.85	4.46	4.82	n.d.	n.d.	n.d.	13.44
C34	7.76	5.80	6.10	4.49	4.75	n.d.	n.d.	n.d.	-

Table 2 continued

Residue	H8/H6	H5/H2	H1'	H2'	H3'	H4'	H5'	H5''	imino
C35	7.76	5.87	5.61	4.17	n.d.	n.d.	n.d.	n.d.	-
U36	7.90	5.30	5.50	4.47	n.d.	n.d.	n.d.	n.d.	13.84
C37	7.93	5.64	5.59	4.22	n.d.	n.d.	n.d.	n.d.	-
C38	7.70	5.49	5.74	4.05	4.21	n.d.	n.d.	n.d.	-

Imino ^{1}H assignments determined at 10° C. Non-exchangeable ^{1}H assignments determined at 35° C. All assignments determined in 10 mM sodium phosphate buffer, pH 7.0, 150 mM NaCl, 150 mM KCl, 2mM spermine. Residues added for construct stability are shown in lowercase letters.

Table 3

Complexed RNA ^{1}H chemical shifts (ppm)

Residue	C8/C6	C5/C2	N1/N3
G1	138.9	-	n.d
G2	136.6	-	143
A3	139.2	153.6	n.d
G4	135.2	-	143.4
G5	135.5	-	144
A6	139.2	154.4	n.d
A7	141.8	155.5	n.d
C8	143.7	99.3	n.d
U9	143.9	105.9	n.d
A10	141.7	154.7	n.d
C11	141.7	98.5	n.d
U12	142.4	103.5	n.d
G13	136.5	-	143.5
U14	140.8	103.8	n.d
C15	141.9	97.6	n.d
U16	141.6	102.9	n.d
U17	142.4	102.8	153.8
C18	142.2	98.1	n.d
c19	141.3	98.5	n.d
u20	141.1	104.9	n.d
u21	144.7	105.4	n.d
c22	142.9	98.7	n.d
g23	142.9	-	n.d
g24	138.7	-	143.7
G25	135.8	-	144.8
U26	140.6	103.7	155.1
C27	142.1	96.5	n.d
G28	137.3	-	n.d
U29	144.1	105.0	n.d
G30	137.6	-	n.d
C31	142.0	96.8	n.d
A32	138.5	153.6	n.d
G33	135.2	-	n.d
C34	141.9	98.6	n.d
C35	141.8	98.5	n.d

Table 3 continued

Residue	C8/C6	C5/C2	N1/N3
U36	142.4	102.8	158.8
C37	142.0	97.2	n.d
C38	141.9	98.0	n.d

^{15}N assignments determined at 10° C. ^{13}C assignments determined at 35° C. All assignments determined in 10 mM sodium phosphate buffer, pH 7.0, 150 mM NaCl, 150 mM KCl, 2mM spermine. Residues added for construct stability are shown in lowercase letters.

Table 4

Intermolecular NOEs

IRES5 proton	Isis-11 proton or proton group
G4 H1'	HG3A-C, HG2A-C
G4 H8	HG3A-C, HG2A-C
G5 H1'	H1
G5 H1'	H2
G5 H1'	HG3A-C, HG2A-C
G5 H2'	H1
G5 H3'	H11
G5 H8	HA2A-B
G5 H8	H11
G5 H8	HG3A-C, HG2A-C
G5 H8	H1
A6 H1'	HG3A-C, HG2A-C
A6 H2	H11
A6 H2	H10A
A6 H2	H10B
A6 H2	HA2A-B
A6 H2	HG3A-C, HG2A-C
A6 H2	H1
A6 H2	H2
A6 H8	HG3A-C, HG2A-C
A7 H8	HG3A-C, HG2A-C
C11 H5	HD1A-C, HE1A-C
U12 H5	HD1A-C, HE1A-C

Table 5

Summary of ribose sugar puckers in the complex

3' endo	2' endo	Unknown
G1, G2, A3, G5, A7, A10, C11, U12, G13, U14, C15, U16, U17, C18, C19, U20, G23, G25, U26, C27, G28, G30, C31, A32, U36, C37, C38	G4, A6, U9, U21, C22, U29, G33,	C8, G24, C34, C35

Table 6

^{13}C and ^{15}N RDCs

Residue	Bond vector	RDC (Hz)	Residue	Bond vector	RDC (Hz)
A3	C2-H2	22.05	G5	C8-H8	21.8
A6	C2-H2	20.1	C11	C6-H6	9.4
A7	C2-H2	4	G13	C8-H8	12.84
A10	C2-H2	18.19	C18	C6-H6	18.51
A7	C8-H8	8.4	C18	C5-H5	32.88
U9	C6-H6	11.17	C19	C5-H5	15.78
A10	C8-H8	22.45	C19	C6-H6	12.56
U12	C6-H6	17.21	C22	C6-H6	12.24
U16	C6-H6	30.16	C22	C5-H5	-25.9
U17	C6-H6	25.93	G23	C8-H8	19.9
U21	C6-H6	19.7	G23	C1'-H1'	21.4
U29	C6-H6	4.1	G24	C8-H8	9.7
A32	C8-H8	8.4	G25	C8-H8	15
U36	C6-H6	12.32	C27	C5-H5	21.52
U21	C5-H5	6.99	G28	C8-H8	32.1
U26	C5-H5	6.92	G30	C8-H8	26.8
U29	C5-H5	0.92	G30	C1'-H1'	-6.9
U36	C5-H5	13.92	C31	C5-H5	20.28
A3	C1'-H1'	-6.92	G33	C8-H8	11.1
A6	C1'-H1'	12.62	C37	C5-H5	22.78
A7	C1'-H1'	-2.92	C38	C5-H5	15.3
U9	C1'-H1'	-2.22	G2	N1-H1	12.67
A10	C1'-H1'	-8.52	G4	N1-H1	16
U20	C1'-H1'	20.62	G5	N1-H1	24.75
U21	C1'-H1'	18.67	G13	N1-H1	14.4
G1	C8-H8	22.6	U17	N3-H3	22.4
G1	C1'-H1'	-11.6	G24	N1-H1	16.1
G2	C8-H8	21.6	G25	N1-H1	22.4
G4	C8-H8	10.9	U26	N3-H3	31.28
G4	C1'-H1'	-3.6	U36	N3-H3	9.8

Alignment Tensor D: -23.26 Rhombicity: 0.24

Table 7

NMR restraints and structural statistics

Experimental Restraints		
RNA intra-molecular NOE distance restraints	295	
Inter-molecular NOE upper bounds	23	
Inter-molecular NOE lower bounds	21	
Hydrogen bond distance and planarity restraints	40	
Dihedral restraints (chi, delta, gamma)	94	
RDC restraints	60	
Structure statistics[1]	S-isomer	R-isomer
Violations		
Distance violations –mean #, (s.d.)	108.1 (6.33)	102.9 (2.32)
Distance violation size - mean distance in Å, (s.d.)	0.028 (0.099)	0.021 (0.064)
Dihedral angle constraints - mean #, (s.d. in °)	1.75 (9.00)	2.26 (12.24)
RDC– mean value in Hz, (s.d)	1.33 (1.16)	1.05 (0.90)
Deviations from idealized geometry		
Bond lengths (Å) – mean value	0.01	0.01
Bond angles (º) – mean value	2.54 (0.03)	2.52 (0.03)
RMS difference for structures[2] (Å)		
All RNA atoms + Inhibitor	2.3 (1.5)	1.7 (1.1)

[1]Mean and standard deviations (SD) violations/deviations from experimental restraints over 10 structures for each isomer.

[2]RMS difference calculated between structures, and rms difference from the mean structure, in parenthesis. Ten structures were used in the calculation for each inhibitor isomer.

APPENDIX A

T7 PRODUCTION OF LONG RNAS WITH HOMOGENEOUS ENDS

The standard methodology for the production of long RNAs with homogenous 3' ends has been the use of ribozymes at the 3' end. These self-cleaving ribozymes leave clean, homogenous 3' ends, which can subsequently be used for ligation reactions requiring such clean ends. The problem with this methodology, at least in our experience, was that the yields were poor. The low yield could be ascribed to a number of possible reasons. For one, the addition of the ribozyme necessitated the use of long (>100 bases) DNA templates. These templates were too long, in most cases, to be synthesized in high yields by the synthesis core at the University of Utah. Thus, the DNA templates were incorporated into plasmids. Plasmid purification and subsequent endonuclease cleavage and purification did not lead to very clean product compared to the synthetic product provided by Dr. Schackmann and coworkers at the DNA synthesis core facility. It is possible that the DNA impurity created an additional variable that made T7 reaction optimization and subsequent ribozyme self-cleavage reactions more challenging than was anticipated. Ribozyme cleavage yields were poor, further limiting yield, and complicating RNA preparation.

Good yields were obtained with a T7 system that utilized a single stranded DNA template and a complementary 17 base single strand "top" oligonucleotide, forming a double stranded recognition sequence for the T7 enzyme at the 3' end of the DNA

template. The problem of 3' heterogeneity in the RNA could be assuaged by the use of 2' O-methyl substituted nucleotides at the 5' end of the template strand. These bulky O-Me groups tend to force the sugar into the 3' endo sugar conformation, which promotes A form helix formation and stability. This promotes stability at the 5' terminus of the single stranded template between the DNA and the 3' end of the nascent RNA, which leads to homogeneous 3' ends of the RNA product. A limitation of this method is the maximum length of the DNA template is restricted by the efficiency of chemical DNA synthesis; as the DNA must be chemically synthesized to incorporate the 2' O-Me groups, it is subject to the usual yield limitations mentioned above.

A new method was developed to incorporate the ribozyme-obviating 2' O-Me groups into the plasmid-based approach for making long RNA oligonucleotides. PCR could be performed on the plasmid of interest employing primers incorporating the 2' O-Me. We found that these altered primers worked as well as conventional primers in the colony PCR reaction. In this way, we were able to produce high yields of double stranded DNA ready for T7 synthesis. It was found that this template produced very clean, homogeneous ends as determined by mass spectrometric analysis.

While it seems as though someone else should have thought of this already, I am unaware of any published record describing this methodology. While we were unable to follow up on this to a sufficient degree to attempt to publish it independently, it proved very successful in producing long homogenous RNAs in good yields for other projects not mentioned in this dissertation.

APPENDIX B

PRODUCTION OF ISOTOPICALLY LABELED RNTPS FROM LABELED MINIMAL MEDIA

With the availability of inexpensive commercially available isotopically labeled rNTPs, it has become less necessary to produce these rNTPs for production of isotopically labeled RNA. Nevertheless, these rNTPs can be produced from ^{13}C labeled glucose and ^{15}N labeled ammonium chloride in reasonable yields and at reasonable cost. Our protocols parallel previously published work by the Pardi group with some modifications. [88]

E. coli is grown in minimal media and the ribosomes are harvested, fully hydrolyzed and phosphorylated to rNTPs. Given that the ribosome is found in similar amounts in all E. coli strains, any hardy strain is probably acceptable. This cell culture may be simultaneously used to produce isotopically labeled protein or tRNA. Isotopically labeled glucose is added at five grams per liter and isotopically labeled ammonium chloride is added at one gram per liter. Typically, M9 minimal media supplemented with trace metals was used, although other minimal media should work as well.

Cells in the trace metal minimal media can grow to A_{600} of four. The cells are resuspended in Pardi buffer C and then lysed with lysozyme and the use of a French press. The lysate is then centrifuged at 5000 rpm to remove the cell debris. The

supernatant is centrifuged overnight at 100,000 G, thus forming a pellet of the ribosomal RNA. This supernatant will contain most of the soluble protein and tRNA and may be harvested independently for the macromolecules of interest. The ribosomal pellet is then resuspended in STE buffer (approximately 10 grams of cells in 40 mLs of buffer) and phenol/chloroform extracted to remove the ribosomal protein material. The rRNA is then quantified by A_{260} measurement. Nuclease P1 provided by Sigma or other sources is then uses to hydrolyze the rRNA to rNMPs. This reaction is monitored by HPLC on an ion exchange column.

The rNMPs can be converted to rNTPs through an enzymatic reaction involving pyruvate kinase, nucleoside monophosphate kinase, myokinase or adenylate kinase, and guanylate kinase as described by Niconowicz et al.[88] This reaction is monitored by HPLC employing a Vydac ion exchange column. The initial buffer (A) consists of 0.045 M sodium formate (pH 4.5). The secondary buffer (B) consists of 0.5 M NaH_2PO_4 (pH 2.5) This method calls for a gradient of 0 to 100 % B over 10 minutes continuing with B buffer for 10 minutes. This method can readily resolve all four rNTPs from one another and from rNMPs. The reaction can reach 90% completion, but 80% is typical. The reaction mixture may be injected directly onto the HPLC column. The rNDPs are produced transiently in the reaction process, but are never present in large enough concentrations to see in the chromatograms.

The rNTPs have to be cleaned up prior to use in a T7 reaction. This is done with a boronate affinity column. The column is prepared by hydrating about 5 grams of Affi-gel 601 (boronate) in TE buffer (Maniatis) and packing the resin into a 20 x 2.5 cm glass

column. This large column gives a good flow rate and allows the resin to expand and contract as pH varies. This column is stored at 4˚C.

At 4˚C, a 1 M TEABC (triethylammonium biocarbonate) buffer is prepared at pH 9.3-9.6 by bubbling CO_2 from dry ice through 141 mL triethylamine in 700 mL of water with mixing until the desired pH is reached. Water is then added to 1 L. The column is equilibrated with 250-300 mL of 1 M TEABC buffer. The phosphorylation reaction mix (~25-50 mLs, typically), having been previously lyophilized, is dissolved in 10-25 mL of TEABC and added to the column. The column is then washed with TEABC until A_{260} is <0.1. The rNTPs are then eluted with bicarbonate buffer until A_{260} is again < 0.1. This bicarbonate buffer is prepared by vigorously bubbling CO_2 from dry ice through deionized water at 4º C with mixing for approximately one hour. The pH of the bicarbonate buffer varies and is hard to measure accurately with a pH meter due to the dearth of counter ions in solution. Complete elution may take 300-500 mL. The early fractions contain a high concentration of TEA and are not pooled with the rest of the eluent.

The clean eluent is pooled and lyophilized. Ethanol precipitation may be used for further purification, but is usually unnecessary. The labeled rNTPs are resuspended in deionized water to about 700 OD_{260} for T7 reactions. The relative concentrations of the four rNTPs in this mix vary somewhat, but the mix can be used as is for RNA production with consistent success.

REFERENCES

1. Strader, D. B.; Wright, T.; Thomas, D. L.; Seeff, L. B., Diagnosis, management, and treatment of hepatitis C. *Hepatology* **2004,** 39, (4), 1147-1171.

2. Tan, S.; Pause, A.; Shi, Y.; Sonenberg, N., Hepatitic C therapeutics: current status and emerging strategies. *Nat. Rev. Drug Discovery* **2002,** 1, 867-881.

3. Molla, A.; Kohlbrenner, W., Evolving therapeutic paradigms for HIV and HCV. *Curr. Opinion in Biotech.* **2003,** 14, 634-640.

4. Evans, M. J.; von Hahn, T.; Tscherne, D. M.; Syder, A. J.; Panis, M.; Wolk, B.; Hatziioannou, T.; McKeating, J. A.; Bieniasz, P. D.; Rice, C. M., Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature* **2007,** 446, (7137), 801-5.

5. Pileri, P.; Uematsu, Y.; Campagnoli, S.; Galli, G.; Falugi, F.; Petracca, R.; Weiner, A. J.; Houghton, M.; Rosa, D.; Grandi, G.; Abrignani, S., Binding of hepatitis C virus to CD81. *Science* **1998,** 282, (5390), 938-41.

6. Scarselli, E.; Ansuini, H.; Cerino, R.; Roccasecca, R. M.; Acali, S.; Filocamo, G.; Traboni, C.; Nicosia, A.; Cortese, R.; Vitelli, A., The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *Embo J* **2002,** 21, (19), 5017-25.

7. Suzuki, T.; Ishii, K.; Aizaki, H.; Wakita, T., Hepatitis C viral life cycle. *Adv Drug Deliv Rev* **2007,** 59, (12), 1200-12.

8. Murakami, K.; Ishii, K.; Ishihara, Y.; Yoshizaki, S.; Tanaka, K.; Gotoh, Y.; Aizaki, H.; Kohara, M.; Yoshioka, H.; Mori, Y.; Manabe, N.; Shoji, I.; Sata, T.; Bartenschlager, R.; Matsuura, Y.; Miyamura, T.; Suzuki, T., Production of infectious hepatitis C virus particles in three-dimensional cultures of the cell line carrying the genome-length dicistronic viral RNA of genotype 1b. *Virology* **2006,** 351, (2), 381-92.

9. Pestova, T. V.; Kolupaeva, V. G.; Lomakin, I. B.; Pilipenko, E. V.; Shatsky, I. N.; Agol, V. I.; Hellen, C. U., Molecular mechanisms of translation initiation in eukaryotes. *Proc Natl Acad Sci U S A* **2001,** 98, (13), 7029-36.

10. Fleming, K.; Ghuman, J.; Yuan, X.; Simpson, P.; Szendroi, A.; Matthews, S.; Curry, S., Solution structure and RNA interactions of the RNA recognition motif from eukaryotic translation initiation factor 4B. *Biochemistry* **2003,** 42, (30), 8966-75.

11. Trachsel, H.; Erni, B.; Schreier, M. H.; Staehelin, T., Initiation of mammalian protein synthesis. II. The assembly of the initiation complex with purified initiation factors. *J Mol Biol* **1977,** 116, (4), 755-67.

12. Benne, R.; Hershey, J. W., The mechanism of action of protein synthesis initiation factors from rabbit reticulocytes. *J Biol Chem* **1978,** 253, (9), 3078-87.

13. Sachs, A. B.; Sarnow, P.; Hentze, M. W., Starting at the beginning, middle, and end:translation initiation in eukaryotes. *Cell* **1997,** 89, 831-838.

14. Svitkin, Y. V.; Herdy, B.; Costa-Mattioli, M.; Gingras, A. C.; Raught, B.; Sonenberg, N., Eukaryotic translation initiation factor 4E availability controls the switch between cap-dependent and internal ribosomal entry site-mediated translation. *Mol Cell Biol* **2005,** 25, (23), 10556-65.

15. Pestova, T. V.; Shatsky, I. N.; Fletcher, S. P.; Jackson, R. J.; Hellen, C. U. T., A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes and Development* **1998,** 12, 67-83.

16. Pestova, T. V.; Lomakin, I. B.; Lee, J. H.; Choi, S. K.; Dever, T. E.; Hellen, C. U., The joining of ribosomal subunits in eukaryotes requires eIF5B. *Nature* **2000,** 403, (6767), 332-335.

17. Pacheco, A.; Reigadas, S.; Martinez-Salas, E., Riboproteomic analysis of polypeptides interacting with the internal ribosome-entry site element of foot-and-mouth disease viral RNA. *Proteomics* **2008,** 8, (22), 4782-90.

18. Babaylova, E.; Graifer, D.; Malygin, A.; Stahl, J.; Shatsky, I.; Karpova, G., Positioning of subdomain IIId and apical loop of domain II of the hepatitis C IRES on the human 40S ribosome. *Nucleic Acids Res* **2009,** 37, (4), 1141-51.

19. Kieft, J. S.; Zhou, K.; Grech, A.; Jubin, R.; Doudna, J. A., Crystal structure of an RNA tertiary domain essential to HCV IRES-mediated translation initiation. *Nat. Struct. Biol.* **2002,** 9, (5), 370-373.

20. Otto, G. A.; Puglisi, J. D., The pathway of HCV IRES-mediated translation initiation. *Cell* **2004,** 119, 369-380.

21. Brown, E. A.; Zhang, H.; Ping, L.; Lemon, S. M., Secondary structure of the 5' nontranslated regions of hepatitis C virus and pestivirus genomic RNAs. *Nucleic Acids Res* **1992,** 20, (19), 5041-5045.

22. Kieft, J. S.; Zhou, K.; Jubin, R.; Doudna, J. A., Mechanism of ribosome recruitment by hepatitis C IRES RNA. *RNA* **2001,** 7, 194-206.

23. Siridechadilok, B.; Fraser, C. S.; Hall, R. J.; Doudna, J. A.; Nogales, E., Structural roles for human translation factor eIF3 in initiation of protein synthesis. *Science* **2005,** 310, 1513-1515.

24. Ji, H.; Fraser, C. S.; Yu, Y.; Leary, J.; Doudna, J. A., Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA. *Proc. Natl. Acad. Sci.* **2004,** 101, (49), 16990-16995.

25. Otto, G. A.; Lukavsky, P. J.; Lancaster, A. M.; Sarnow, P.; Puglisi, J. D., Ribosomal proteins mediate the hepatitis C virus IRES-HeLa 40S interaction. *RNA* **2002,** 8, 913-923.

26. Dibrov, S. M.; Johnston-Cox, H.; Weng, Y.; Hermann, T., Functional architecture of HCV IRES domain II stabilized by divalent metal ions in the crystal and in solution. *Angew. Chem. Int. Ed.* **2007,** 46, 226-229.

27. Lukavsky, P. J.; Kim, I.; Puglisi, J. D., Structure of the HCV IRES domain II determined by NMR. *Nat. Struct. Biol.* **2003,** 10, (12), 1033-1038.

28. Zhao, Q.; Han, Q.; Kissinger, C. R.; Hermann, T.; Thompson, P. A., Structure of hepatitis C virus IRES subdomain IIa. *Acta Crystallographica Section D* **2008,** D64, 436-443.

29. Odreman-Macchioli, F.; Baralle, F. E.; Buratti, E., Mutational analysis of the different bulge regions of hepatitis C virus domain II and their influence on internal ribosome entry site translational ability. *J. Biol. Chem.* **2001,** 276, 41648-41655.

30. Locker, N.; Easton, L. E.; Lukavsky, P. J., HCV and CSFV domain II mediate eIF2 release during 80S ribosome assembly. *The EMBO Journal* **2007,** 26, 795-805.

31. Spahn, C. M. T.; Kieft, J. S.; Grassucci, R. A.; Penczek, P. A.; Zhou, K.; Doudna, J. A.; Frank, J., Hepatitis C virus IRES RNA-induced changes in the conformation of the 40S ribosomal subunit. *Science* **2001,** 291, 1959-1962.

32. Fukushi, S.; Okada, M.; Stahl, J.; Kageyama, T.; Hoshino, F. B.; Katayama, K., Ribosomal protein S5 interacts with the internal ribosomal entry site of hepatitis C virus. *J Biol Chem* **2001,** 276, (24), 20824-6.

33. Kolupaeva, V. G.; Pestova, T. V.; Hellen, C. U. T., An enzymatic footprinting analysis of the interaction of 40S ribosomal subunits with the internal ribosomal entry site of hepatitis C virus. *J. Virol.* **2000,** 74, (14), 6242-6250.

34. Pestova, T. V.; de Breyne, S.; Pisarev, A. V.; Abaeva, I. S.; Hellen, C. U., eIF2-dependent and eIF2-independent modes of initiation on the CSFV IRES: a common role of domain II. *Embo J* **2008,** 27, (7), 1060-72.

35. Wang, H.; Tor, Y., Electrostatic interactions in RNA aminoglycosides. *J. Am. Chem. Soc.* **1997,** 119, 8734-8735.

36. Gallego, J.; Varani, G., Targeting RNA with small-molecule drugs: therapeutic promise and chemical challenges. *Acc. Chem. Res.* **2001,** 34, (10), 836-843.

37. Shakil, S.; Khan, R.; Zarrilli, R.; Khan, A. U., Aminoglycosides versus bacteria--a description of the action, resistance mechanism, and nosocomial battleground. *J Biomed Sci* **2008,** 15, (1), 5-14.

38. Silva, J. G.; Carvalho, I., New insights into aminoclycoside antibiotics and derivatives. *Curr. Med. Chem.* **2007,** 14, 1101-1119.

39. Fourmy, D.; Recht, M. I.; Blanchard, S. C.; Puglisi, J. D., Structure of the A site of Escherichia coli 16S ribosomal RNA complexed with an aminoglycoside antibiotic. *Science* **1996,** 274, (5291), 1367-71.

40. Fourmy, D.; Yoshizawa, S.; Puglisi, J. D., Paromomycin binding induces a local conformational change in the A-site of 16 S rRNA. *J Mol Biol* **1998,** 277, (2), 333-45.

41. Carter, A. P.; Clemons, W. M.; Brodersen, D. E.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V., Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* **2000,** 407, (6802), 340-8.

42. Ogle, J. M.; Brodersen, D. E.; Clemons, W. M.; Tarry, M. J.; Carter, A. P.; Ramakrishnan, V., Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* **2001,** 292, (5518), 897-902.

43. Yoshizawa, S.; Fourmy, D.; Puglisi, J. D., Recognition of the codon-anticodon helix by ribosomal RNA. *Science* **1999,** 285, (5434), 1722-5.

44. Ogle, J. M.; Carter, A. P.; Ramakrishnan, V., Insights into the decoding mechanism from recent ribosome structures. *Trends Biochem. Sci.* **2003,** 28, (5), 259-266.

45. Hansen, J. L.; Ippolito, J. A.; Ban, N.; Nissen, P.; Moore, P. B.; Steitz, T. A., The structures of four macrolide antibiotics bound to the large ribosomal subunit. *Mol Cell* **2002,** 10, (1), 117-28.

46. Schlunzen, F.; Zarivach, R.; Harms, J.; Bashan, A.; Tocilj, A.; Albrecht, R.; Yonath, A.; Franceschi, F., Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature* **2001,** 413, (6858), 814-21.

47. Hansen, J. L.; Moore, P. B.; Steitz, T. A., Structures of five antibiotics bound at the peptidyl transferase center of the large ribosomal subunit. *J Mol Biol* **2003,** 330, (5), 1061-75.

48. Pioletti, M.; Schlunzen, F.; Harms, J.; Zarivach, R.; Gluhmann, M.; Avila, H.; Bashan, A.; Bartels, H.; Auerbach, T.; Jacobi, C.; Hartsch, T.; Yonath, A.; Franceschi, F., Crystal structures of complexes of the small ribosomal subunit with tetracycline, edeine and IF3. *Embo J* **2001,** 20, (8), 1829-39.

49. Brodersen, D. E.; Clemons, W. M., Jr.; Carter, A. P.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V., The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. *Cell* **2000,** 103, (7), 1143-54.

50. Bannwarth, S.; Gatignol, A., HIV-1 TAR RNA: the target of molecular interactions between the virus and its host. *Curr HIV Res* **2005,** 3, (1), 61-71.

51. Murchie, A. I. H.; Davis, B.; Isel, C.; Afshar, M.; Drysdale, M. J.; Bower, J.; Potter, A. J.; Starkey, I. D.; Swarbrick, T. M.; Mirza, S.; Prescott, C. D.; Vaglio, P.; Aboul-ela, F.; Karn, J., Structure-based drug design targeting an inactive RNA conformation: expoiting the flexibility of HIV-1 TAR RNA. *J. Mol. Biol.* **2004,** 336, 625-638.

52. Dasgupta, A.; Das, S.; Izumi, R.; Venkatesan, A.; Barat, B., Targeting internal ribosome entry site (IRES)-mediated translation to block hepatitis C and other RNA viruses. *FEMS Microbiol. Lett.* **2004,** 234, 189-199.

53. Chevalier, C.; Saulnier, A.; Benureau, Y.; Flechet, D.; Delgrange, D.; Colbere-Garapin, F.; Wychowski, C.; Martin, A., Inhibition of hepatitis C virus infection in cell culture by small interfering RNAs. *Mol. Ther.* **2007,** 15, (8), 1452-1462.

54. Seth, P. P.; Miyaji, A.; Jefferson, E. A.; Sannes-Lowery, K. A.; Osgood, S. A.; Propp, S. S.; Ranken, R.; Massire, C.; Sampath, R.; Ecker, D. J.; Swayze, E. E.; Griffey, R. H., SAR by MS: Discovery of a new class of RNA-binding small molecules of the hepatitis C virus: internal ribosome entry site IIA subdomain. *J. Med. Chem.* **2005,** 48, 7099-7102.

55. Kieft, J. S.; Zhou, K.; Jubin, R.; Murray, M. G.; Lau, Y. N. L.; Doudna, J. A., The hepatitis C virus internal ribosome entry site adopts an ion-dependent tertiary fold. *J. Mol. Biol.* **1999,** 292, 513-529.

56. Davis, J. H.; Tonelli, M.; Scott, L. G.; Jaeger, L.; Williamson, J. R.; Butcher, S. E., RNA helical packing in solution: NMR structure of a 30 kDa GAAA tetraloop-receptor complex. *J Mol Biol* **2005,** 351, (2), 371-82.

57. Kao, C.; Zheng, M.; Rudisser, S., A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. *Rna* **1999,** 5, (9), 1268-72.

58. Sherlin, L. D.; Bullock, T. L.; Nissan, T. A.; Perona, J. J.; Lariviere, F. J.; Uhlenbeck, O. C.; Scaringe, S. A., Chemical and enzymatic synthesis of tRNAs for high-throughput crystallization. *Rna* **2001,** 7, (11), 1671-8.

59. Zuker, M., Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **2003,** 31, (13), 3406-3415.

60. Varani, G.; Aboul-ela, F.; Allain, F. H.-T., NMR investigation of RNA structure. *Prog. Nucl. Magn. Reson. Spectrosc.* **1996,** 29, 51-127.

61. Neuhaus, D.; Williamson, M. P., *The Nuclear Overhauser Effect in Structural and Conformational Analysis*. VCH Publishers, Inc.: New York, 1989.

62. Kay, L. E.; Keifer, P.; Saarinen, T., Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **1992,** 114, 10663-10665.

63. Furtig, B.; Richter, C.; Wohnert, J.; Schwalbe, H., NMR spectroscopy of RNA. *Chembiochem* **2003,** 4, 936-962.

64. Shajani, Z.; Varani, G., NMR studies of dynamics in RNA and DNA by 13C relaxation. *Biopolymers* **2007,** 86, (5-6), 348-359.

65. Peterson, R. D.; Theimer, C. A.; Wu, H.; Feigon, J., New applications of 2D filtered/edited NOESY for assignment and structure elucidation of RNA and RNA-protein complexes. *J Biomol NMR* **2004,** 28, (1), 59-67.

66. Vuister, G. W.; Clore, M.; Gronenborn, A. M.; Powers, R.; Garet, D. S.; Tschudin, R.; Bax, A., Increased resolution and improved spectral quality in four-dimensional 13C/13C-separated HMQC-NOESY-HMQC spectra using pulsed field gradients. *J. Magn. Reson. B.* **1993,** 101, 210-213.

67. Zwahlen, C.; Legault, P.; Vincent, J. F. V.; Greenblatt, J.; Konrat, R.; Kay, L. E., Methods for Measurement of Intermolecular NOEs by Multinuclear NMR

Spectroscopy: Application to a Bacteriophage λ N-Peptide/boxB RNA Complex. *J. Am. Chem. Soc.* **1997,** 119, (29), 6711-6721.

68. Stuart, A. C.; Borzilleri, K. A.; Withka, J. M.; Palmer III, A. G., Compensating for Variations in 1H−13C Scalar Coupling Constants in Isotope-Filtered NMR Experiments. *J Am Chem Soc* **1999,** 121, (22), 5346-5347.

69. Pardi, A.; Nikonowicz, E. P., Simple procedure for resonance assignment of the sugar protons in ^{13}C-labeled RNAs. *J. Am. Chem. Soc.* **1992,** 114, 9202-9203.

70. Prestegard, J. H.; Al-Hashimi, H. M.; Tolman, J. R., NMR structures of biomolecules using field oriented media and residual dipolar couplings. *Q. Rev. Biophys.* **2000,** 33, (4), 371-424.

71. Sibille, N.; Pardi, A.; Simorre, J. P.; Blackledge, M., Refinement of local and long-range structural order in the theophylline-binding RNA using (13)C-(1)H residual dipolar couplings and restrained molecular dynamics. *J Am Chem Soc* **2001,** 123, (49), 12135-12146.

72. Tjandra, N.; Bax, A., Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* **1997,** 278, (5340), 1111-1114.

73. Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H., Nuclear magnetic dipole interactions in field-oriented proteins: information for structural determination in solution. *Proc Natl Acad Sci U S A* **1995,** 92, (20), 9279-9283.

74. Ottiger, M.; Delaglio, G.; Bax, A., Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.* **1998,** 131, (2), 373-378.

75. Tjandra, N.; Bax, A., Measurement of dipolar contributions to 1JCH splittings from magnetic-field dependence of J modulation in two-dimensional NMR spectra. *J. Magn. Reson.* **1997,** 124, 512-515.

76. Tzakos, A. G.; Grace, C. R. R.; Lukavsky, P. J.; Rick, R., NMR techniques for very large proteins and RNAs in solution. In *Nuclear Magnetic Resonance of Biological Macromolecules*, Abelson, J. N.; Simon, M. I., Eds. Elsevier: Amsterdam, 2005; Vol. 394, pp 399-416.

77. Ruckert, M.; Otting, G., Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. *J. Am. Chem. Soc.* **2000,** 122, (32), 7793-7797.

78. Boisbouvier, J.; Bryce, D. L.; O'Neil-Cabello, E.; Nikonowicz, E. P.; Bax, A., Resolution-optimized NMR measurement of $^{1}D_{CH}$, $^{1}D_{CC}$ and $^{2}D_{CH}$ residual dipolar couplings in nucleic acid bases. *J. Biomol. NMR* **2004,** 30, 287-301.

79. Ward, D. C.; Reich, E.; Stryer, L., Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives. *J Biol Chem* **1969,** 244, (5), 1228-37.

80. Clore, G. M.; Kuszewski, J., Improving the accuracy of NMR structures of RNA by means of conformational database potentials of mean force as assessed by complete dipolar coupling cross-validation. *J Am Chem Soc* **2003,** 125, 1518-1525.

81. Zweckstetter, M.; Hummber, G.; Bax, A., Prediction of sterically induced alignment in a dilute liquid crystalline phase: aid to protein structure determination by NMR. *J Am Chem Soc* **2000,** 122, (15), 3791-3792.

82. Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A., Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* **2001,** 98, (18), 10037-10041.

83. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF chimera-a visualization system for exploratory research and analysis. *J Comput Chem* **2004,** 25, (13), 1605-12.

84. Diamond, R., Coordinate-based cluster analysis. *Acta Crystallogr D Biol Crystallogr* **1995,** 1, (51(2)), 127-135.

85. Dove, W. F.; Davidson, N., Cation effects on the denaturization of DNA. *J. Mol. Biol.* **1962,** 5, 467-478.

86. Parsons, J.; Castaldi, M. P.; Dutta, S.; Dibrov, S. M.; Wyles, D. L.; Hermann, T., Conformational inhibition of the hepatitis C virus internal ribosome entry site RNA. *Nat Chem Biol* **2009,** 5, 823-825.

87. Begley, D. W.; Varani, G., Locking out viral replication. *Nat Chem Biol* **2009,** 5, 782-783.

88. Nikonowicz, E. P.; Sirr, A.; Legault, P.; Jucker, F. M.; Baer, L. M.; Pardi, A., Preparation of 13C and 15N labelled RNAS for heteronuclear multi-dimensional NMR studies. *Nucleic Acids Research* **1992,** 20, (17), 4507-4513.