1-1-2011

Nmr Solution Structures Of Two Hairpins Of E. Coli 16s Rrna: The Effects Of Mutations And Chemical Modifications On Structure And Function Of Rrna

Yu Liu
Wayne State University

NMR SOLUTION STRUCTURES OF TWO HAIRPINS OF *E. coli* 16S rRNA: THE EFFECTS OF MUTATIONS AND CHEMICAL MODIFICATIONS ON STRUCTURE AND FUNCTION OF rRNA

by

YU LIU

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2011

MAJOR: CHEMISTRY (Analytical Chemistry)

Approved by:

_____

Advisor Date

_____

_____

_____

DEDICATION

To my beloved husband, Jiang Du.

ACKNOWLEDGMENTS

First I would like to thank my Ph.D. advisor, Prof. John SantaLucia, Jr. for his support and encouragements in the last five years. Prof. SantaLucia's influence and mentoring is the most valuable treasure that will benefit me in the long run.

I am deeply thankful to Prof. Christine Chow and Prof. Philip Cunningham for giving me many useful suggestions and help in the collaboration projects. I also thank the graduate students in Prof. Chow and Prof. Cunningham's labs, especially Dinuka Abeydeera and Yogo Sakakibara for synthesizing the 7mer-peptide used in the 970 project.

I appreciate the helpful comments for my research and the thesis from my Ph. D committee members: Dr. Philip Cunnnigham, Dr. Tiffany Mathews and Dr. David Rueda. And also thanks Dr. Andrew Feig for his useful suggestions in the RNA club and in my oral exam. My special thanks go to my course work instructors Prof. Bhagwat, Prof. Chow, Prof. Poole, Prof. Spence, and Prof. Spaller.

I am also grateful to all my lab members over these years, especially Fred for teaching me NMR experiments when I joined the lab, Ravi for helping me to fix the structural restraints for calculating NMR structures, and Marcus for the valuable advice for the thesis. I want to acknowledge Larry, Jun and Laura for their help in many endeavors.

I also would like to thank the Department of Chemistry, Wayne State University and the National Institutes of Health (NIH) for their financial support. I

want to thank Dr. David Coleman for his support for the 700 MHz NMR facility and Dr. Bashar Ksebati for his help. I also would like to thank the Chemistry Department staff Melissa Barton, Nestor Ocampo, Mary Wood, Deborah McCreless, Diane Klimas, Diane Kudla, Beverly Jacobs, Bernadette Miesik, Sharon Kelley and all others who helped me.

I want to thank my parents for their unselfish and unconditional love to motivate me to be a better person. Finally, thank my husband, Jiang. With his love, I feel blissful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction of Ribonucleic Acids

1.1. Diversity and Functions of RNA

The flow of genetic information in organisms is from DNA to proteins (usually through transcription of DNA into RNA, and then translation into proteins, which is called the “central dogma” of molecular biology (Crick 1970)). The most common types of RNA are messenger RNA and transfer RNA, and both mRNA and tRNA are involved directly in the genetic information transfer processes. mRNA, whose sequences contain codons with genetic information for protein, are short-lived species that work as information-carrying intermediates that transfer genetic information from sections of DNA to proteins. The coding sequence of mRNA determines the sequence of the synthesized protein according to the genetic code (Caskey 1968). tRNA is relatively small RNA with about 80 nucleotides that can attach amino acids at its 3’-end, and with the ribosome transfers the amino acids to growing protein chains during translation. tRNA play an important role as molecular adapters to correlate the mRNA codes with the amino acid code (anti-codons) with the help of an RNA-protein complex machine named ribosome. The RNA in ribosome is called ribosomal RNA (rRNA), and the bacterial ribosomal RNA can be divided into 5S rRNA, 16S rRNA, and 23S rRNA. Ribosome is discussed with more details later in this chapter.

RNA is unique in biology in that it acts both as genetic information carrier and as a catalyst. The group I intron is the first RNA has been identified as catalytic (Kruger 1982). Ribozymes are defined as the enzymes with RNA as

catalytic centers. The second ribozyme discovered is ribonuclease P (RNase P). The crystal structures of RNase P from *Thermotoga maritime* and *Bacillus steraothermophilus* show that the catalytic moiety of RNase P in bacteria is only RNA (Guerrier-Takada 1983; Kazantsev 2005; Torres-Larios 2005).

RNA with diverse functions has been identified, recently discovered new types of RNA include small nuclear RNA (snRNA), RNA interference (RNAi), microRNA (miRNA), small interfering RNA (siRNA), piwi-interfering RNA (piRNA), and riboswitches. One of the main functions of these RNA is to involve in gene regulation and contributes to the diversity of gene expression in cells. For example, snRNAs (usually 60-300 nt) play an important role in the biological process called RNA splicing and the splicing can create alternatively spliced messages that result in different protein products (Wang 2008). RNAi was discovered based on its function to recognize and silence the similar genes of petunia (Napoli 1990). miRNA (usually 21 or 22 nt) can recognize some mRNAs and silence their gene expression by degrading the mRNA through the DICER protein complex and inhibiting the related translation processes (Lee 1993; Fire 1998). siRNA (usually 20-25 nt) is able to target nascent transcripts to specify local chromatin modifications and silencing gene expression (Herr 2005; Lida 2008). Some miRNA and siRNA can interact with some genes and cause the genes to be methylated, and hence affect gene transcriptions (Doran 2007; Pushparaj 2008), (Sontheimer 2005). piRNA (usually 29 or 30 nt) is likely to control the gene expression by directing genome silencing (Brennecke 2007).

Riboswitches are part of mRNA in many bacteria that control gene expression by binding with certain metabolites (Breaker 2008).

The diverse functions of RNA make “RNA world” hypothesis reasonable. Actually, as early as in 1967, Carl Woese already predicted that RNA could play catalytic roles in the cell similar to protein enzymes, and he also proposed that the earliest forms of life originated from a common ancestor in which RNA had multiple functions as both genetic information carrier, metabolite, and catalyst, thereby supporting what was later called the “RNA world” (Szathmáry 1999). The versatile roles of RNA lead to the “RNA world” hypothesis (Gilbert 1986): a world where RNA works as genetic information carrier instead of DNA, and RNA also works as enzyme instead of proteins in many basic biological processes, such as decoding the sequence of RNA for protein synthesis. Some research results support the “RNA world” hypothesis. An RNA ligase activity is a prerequisite to the hypothesis of the “RNA world” with self-replicating RNA (Robertson 2007). And the existence of the ribozyme with RNA ligase activity in the *in vitro* evolution system provide experimental support for the basic requirement of the “RNA world” hypothesis (Robertson 2007). Lincoln *et al.* also designed an RNA system to demonstrate the hypothesis, in which the first RNA enzyme catalyzes reactions that produce the second enzyme from two RNA molecules, and the second enzyme turn out to be the catalyst for producing the first enzyme from the original two RNA molecules, which maintains the system to replicate continuously (Lincoln 2009). Both Robertson and Lincoln data support the hypothesis that the functions of RNA are diverse and RNA is capable of originating life.

1.2. Structures of RNA

The structure of RNA is related closely to its function. Different structures of RNA can cause different interaction formats between RNA and its target, which makes RNA able to play different roles. And therefore, RNA structure is valuable for exploring RNA function. RNA structure can be described by primary, secondary, tertiary structure and quaternary structure.

1.2.1. Chemical Structures of RNA

RNA is a polymer that consists of linked nucleotide units. For most natural RNA, each nucleotide is composed of a nitrogenous base, a ribose sugar, and a phosphate (as shown in Figure1.1).

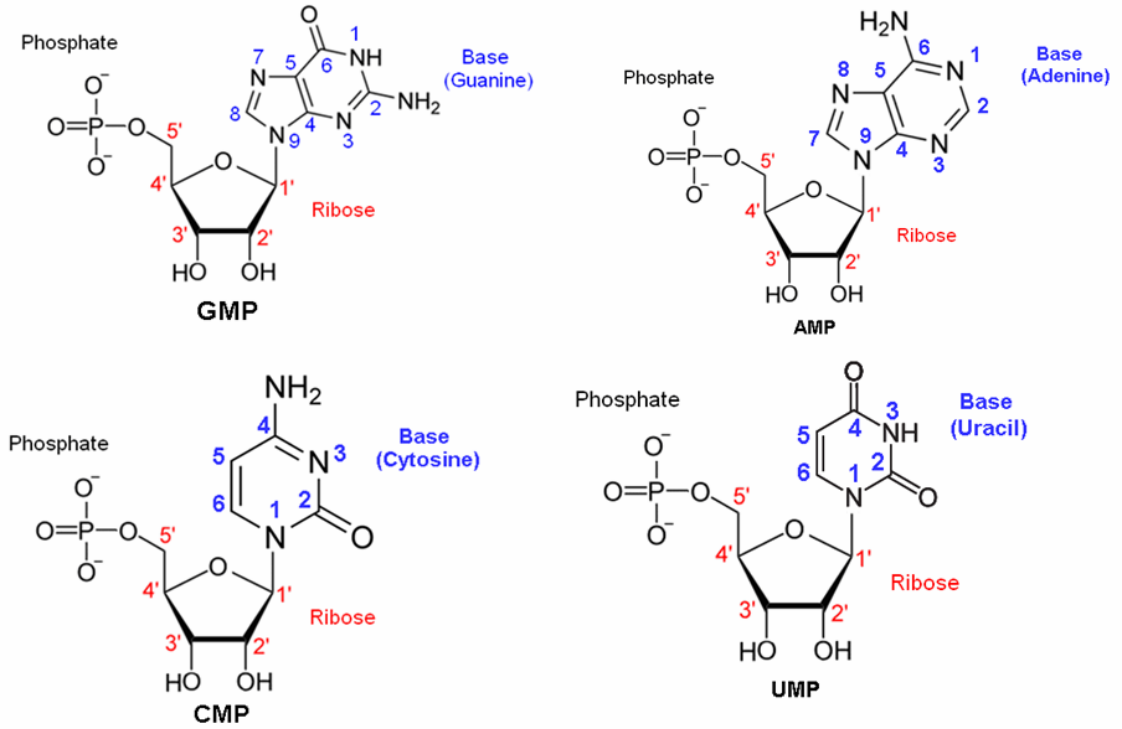


Figure 1.1. Chemical structure of the RNA nucleotide. The atom numbers of guanosine 5'-monophosphate (abbreviated GMP), adenosine 5'-monophosphate (abbreviated AMP), cytidine 5'-monophosphate (abbreviated CMP), and uridine 5'-monophosphate (abbreviated UMP) are labeled.

Bases in RNA are linked with the C1’ of ribose, and generally four common types of bases exist in RNA molecules, and they are adenine (A), guanine (G), cytosine (C) and uracil (U). Among the four bases, the first two bases are purines, and the latter two bases are pyrimidines. The phosphate group is connected to the O3’ and O5’ of two adjacent riboses. RNA is negatively charged because the phosphate groups are fully deprotonated at the near neutral pH in the cell.

1.2.2. Primary Structures of RNA

Each nucleoside comprises a ribose and a base. In an RNA polymer, the nucleotides are connected together by phosphodiester linkages. The primary structure of RNA is the sequential order of the bases and is commonly referred to as the “sequence”. An example of a four nucleotide RNA is given in Figure 1.2. The sequence of RNA is by convention specified from 5’ to 3’ direction.

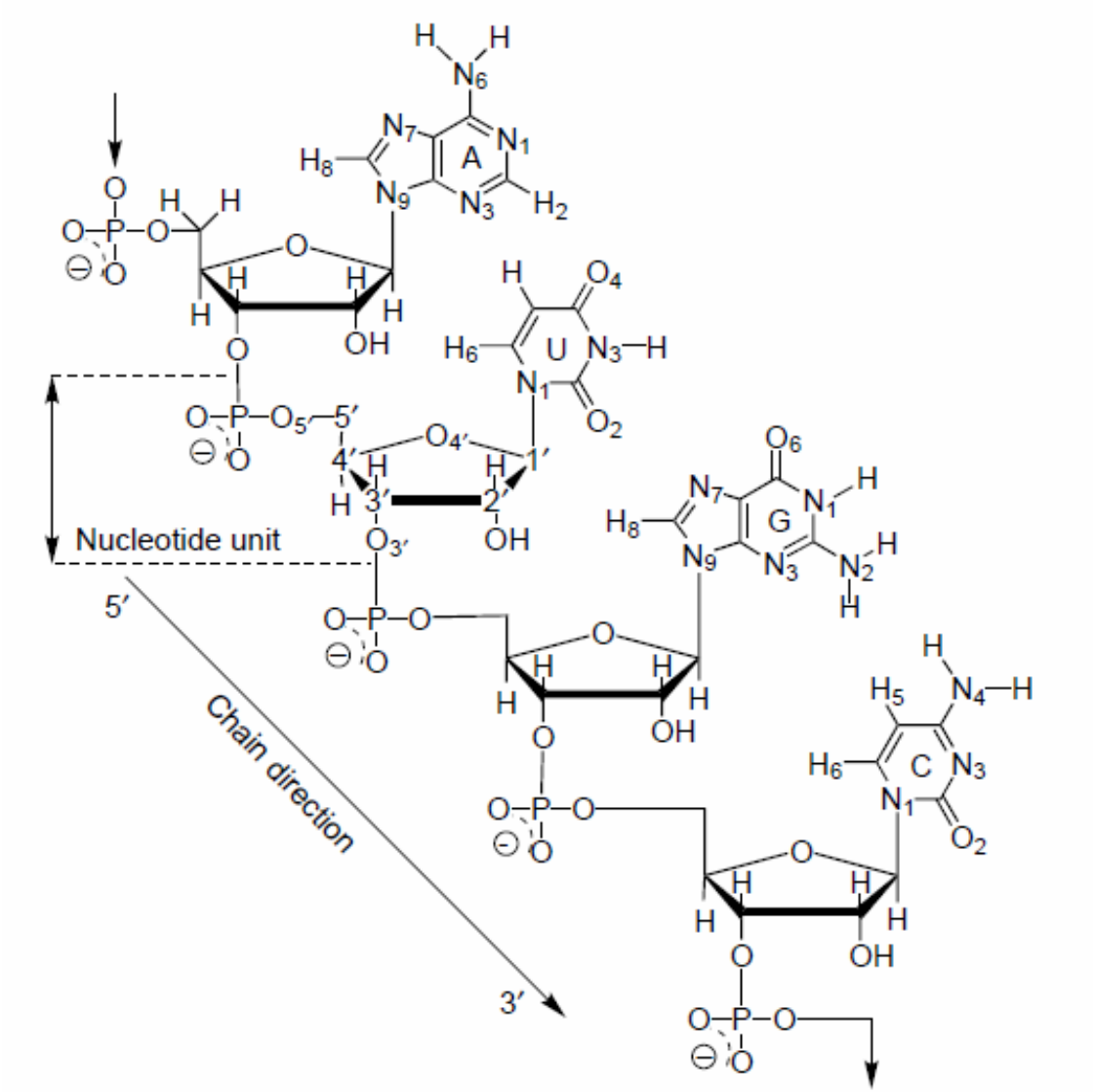


Figure 1.2. The primary structure of RNA with the sequence 5'-AUGC-3'. (Westhof 2000).

1.2.3. Secondary Structures of RNA

The secondary structure of RNA shows a 2D representation of the base-pairs that stabilize the folding of RNA. The double-stranded helix is one of the most common structural elements of RNA secondary structures. Double-stranded RNA duplexes are stabilized by Watson-Crick base pairs (Batey 1999; Tinoco 1999). The hydrogen bonds formed in the two base-pairs, guanine and cytosine, adenine and uracil are shown in Figure 1.3. The helical regions of RNA are comprised of Watson-Crick base-pairs (Figure 1.3).

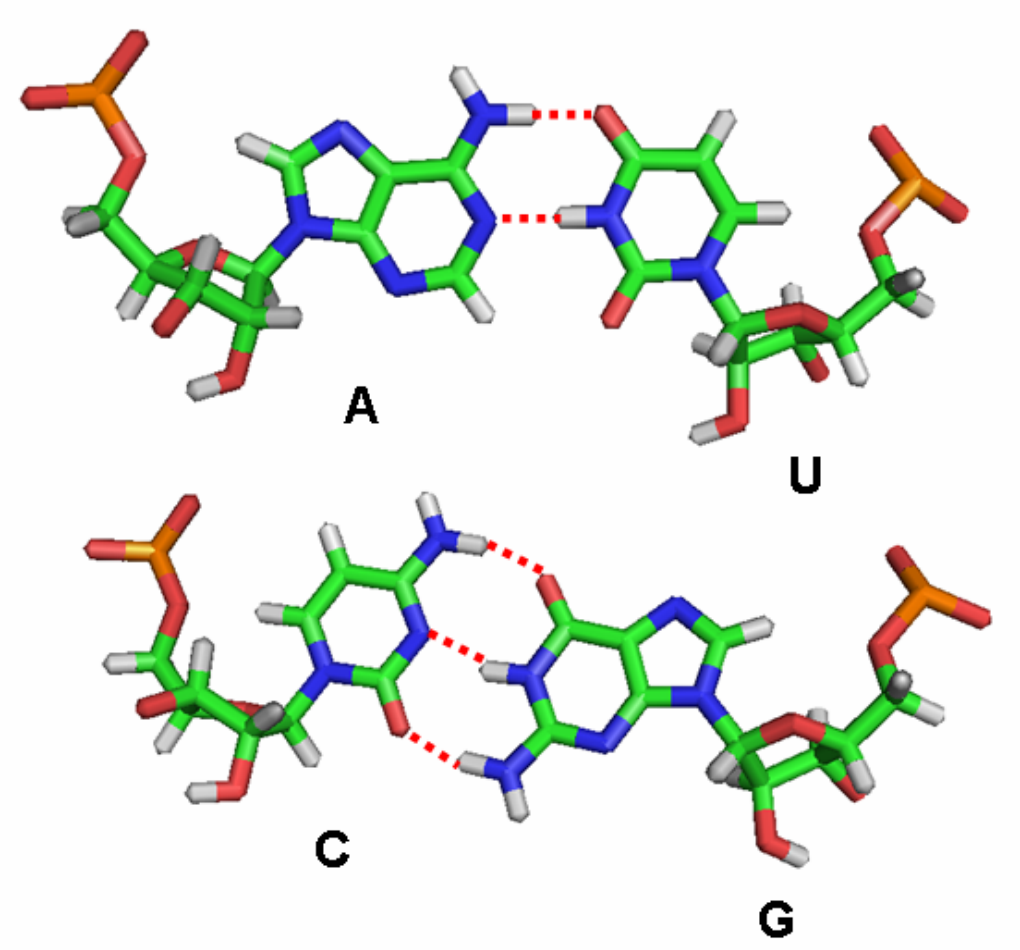


Figure 1.3. The standard Waston-Crick AU and GC base-pairs. The carbon, nitrogen, oxygen, phosphorus and hydrogen atoms in the base pairs are shown in green, blue, red, orange and white, respectively.

In addition to Watson-Crick base pairs, there are a wide variety of non-Watson-Crick pairs (often called mismatches) (Leontis 2001). Non-Watson-Crick base-base interactions frequently exist in RNA structures, which contribute to the specific binding by ligands, ions, proteins, or nucleic acids (Leontis 2001). Figure 1.4 shows a variety of the possible non-Watson-Crick pairs.

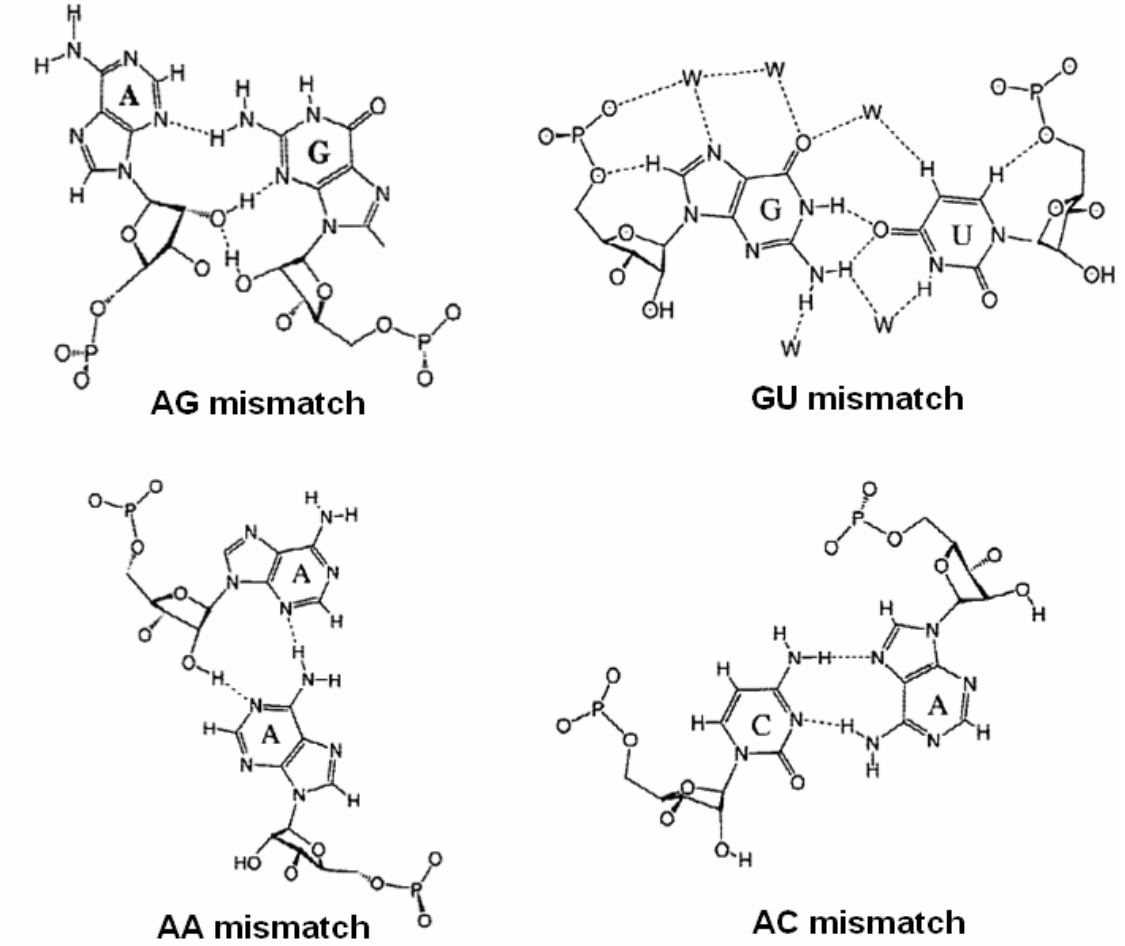


Figure 1.4. The mismatches, including AG, GU, AA, and AC mismatches are all non-Watson-Crick pairs (Leontis 2001).

Besides double-stranded helices, other single-stranded loop motifs are also found frequently in the secondary structures of RNA, including hairpins, single-stranded loops, bulges, internal loops and multi-helix junctions (Batey 1999) (Turner 1988). Hairpin loops are formed when RNA folds back upon itself to change the backbone direction. The unique characteristic of an internal loop is that one or more unpaired nucleotides are found in both strands between two helices (Figure 1.5). Bulges are a special case of an internal loop in which only one strand has unpaired residues between two helices (Figure 1.5). Junctions are formed when more than one double-stranded RNA regions are connected in a circular arrangement, sometimes with connectivity single stranded residues (Figure 1.5). And all the motifs of secondary structure of RNA can be observed in the 16S ribosomal RNA in *E. coli* (Figure 1.5). There are four domains in the secondary structures of 16S rRNA. Usually, the tertiary structure of 16S rRNA can be assigned into three domains as head, body and platform. As labeled by different color in the structure, the secondary structure of 16S rRNA is divided into four domains: the 5', central, 3'-major and 3'-minor domains (Figure 1.5) (Culver 2003).

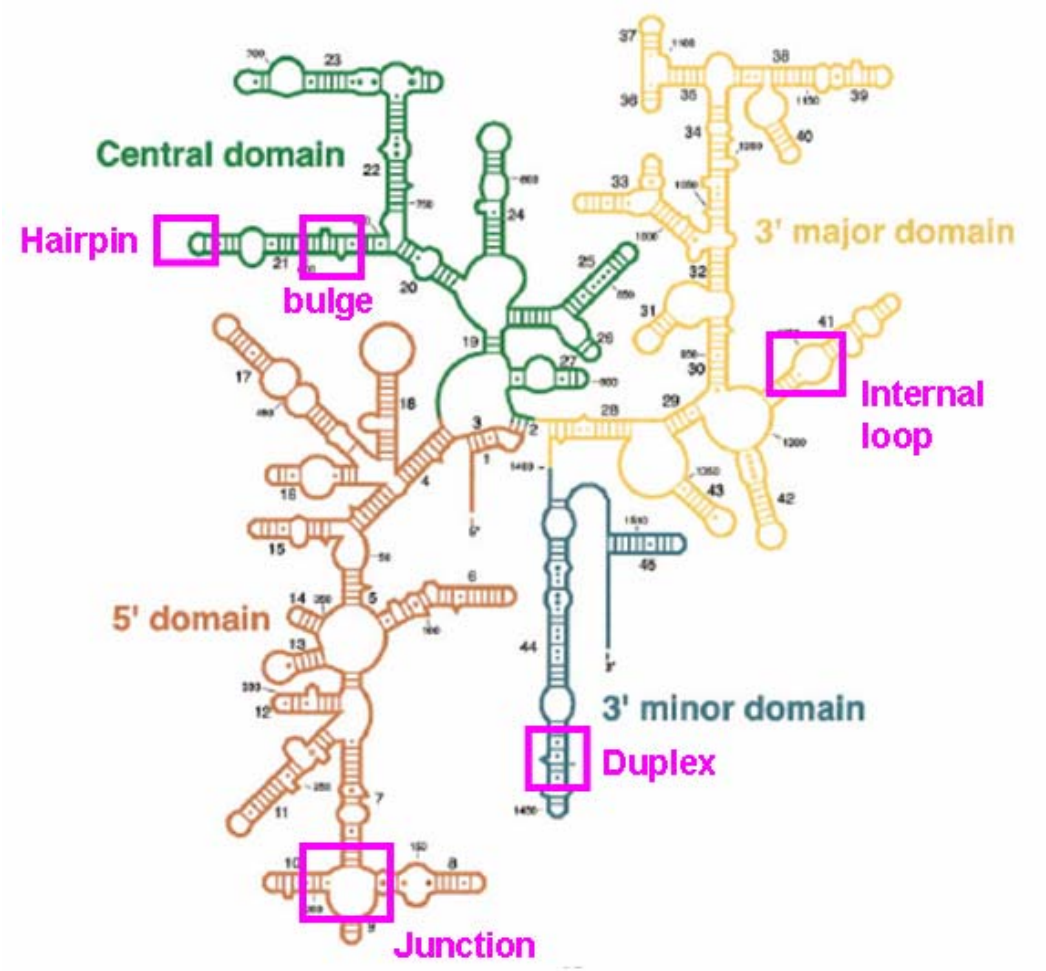


Figure 1.5. Secondary structure of the 16S rRNA from *E. coli* (Culver 2003). The 16S rRNA is made up of four domains: the 5'-major domain, the central domain, the 3'-major domain, and the 3'-minor domain. The secondary structure of the 16S rRNA is made up of many RNA structural motifs, including duplex, hairpin, bulge, internal loop and junction.

For RNA molecules, especially the RNA with large number of nucleotides in the sequences, more than one secondary structure is possible. Many factors are contributed for secondary structure formation, and thermodynamic stability is the most important factor for secondary structure formation. The thermodynamic stability of the secondary structures of RNA were predicted based on the hypothesis that more Watson-Crick base pairs (GC versus AU base pairs) will cause more stable structures due to the lower energy of the hydrogen bonds in these base pairs (Jaeger 1993). Later it was found that many other factors, such as coaxial stacking, stability of hairpin loops, and mismatches also affect how RNA folds from linear nucleotide chains into stable secondary structures. More recent studies suggest that building up native RNA structures based on the kinetics of the folding processes (Geis 2008). The major interactions among secondary structures include hydrogen bonds, London dispersion, and electrostatic interactions. Base stacking is a common interaction in RNA, and this is entropy driven and based upon London dispersion power (Petersheim 1983). The coaxial stacks of helical domains in RNA structures are dominated by base stacking, and the interactions that are most common in the stacks are Watson-Crick base pairs or non-Watson-Crick pairs (Lescoute 2006). The dominant role of base stacking is evident in that more than half of the bases in the structure of tRNA are involved in stacking (Kim 1976).

Commonly, chemical of enzymatic probing is used to determine secondary structures (Jaeger 1993) (Ehresmann 1987). These kinds of experiments are based on the fact that reactive chemicals of certain enzymes do not easily

access the inside of helical regions. Therefore, the affected reactive sites of RNA are located at easily accessible regions, are typically single-stranded, hairpins, bulges, *etc*. The modified sites of RNA are identified either by sequencing ^{32}P-labeled RNA or by reverse transcription of the modified sites of RNA with a ^{32}P-labled or fluorophore-labeled DNA primer. Most recently, the SHAPE method has been used in conjunction with thermodynamic energy minimization to accurately predict the secondary structure of RNAs as large as the HIV genome (Watts 2009). Comparative sequence analysis is more reliable method for secondary structure determination. Fox and Woese first used this method to predict the secondary structures of 5S rRNA, and the secondary structures of 16S rRNA and 23S rRNA predicted from comparative analysis closely matched the experimental data from later X-ray crystallographic analysis (Fox 1975; Woese 1980; Noller 1981; Woese 1993). Comparative sequence analysis is based on the hypothesis that the global structures of RNA, including the secondary structure, are conserved throughout evolutionary history despite variation in sequence. Computer programs such as MFOLD (Zuker 2003) and SFOLD (Ding 2004) are also used to predict RNA secondary structures. These programs build up the secondary structures based on calculating and comparing the free energy of various alternative secondary structures from fragments of the same primary sequence. And the structures with the lowest free energy are considered to be the most stable structures and to thus have the largest population at equilibrium. Although secondary structures of RNA can be calculated by computer programs

quickly and reasonably effectively, some calculated structures are not reliable as experimental data.

1.2.4. Tertiary Structures of RNA

Under appropriate conditions, secondary structural elements can assemble to form various three-dimensional structures with higher complexity, including side-by-side interaction of two double-stranded helices, loop-helix interactions, loop-loop interactions, bases-hairpin loop interactions, *etc*. For example, “kissing loops” are based on base pairs between complementary sequences of different hairpin loops that contribute to the formation of distinct structural domains in tertiary structures (Marino 1995; Lee 1998). To stabilize RNA tertiary structures, more hydrogen-bond motifs are found among the bases, such as triplex shown in Figure 1.6. The adenine in Figure 1.6, A interacts with one uracil by the regular Waston-Crick H-bonds, while the second uracil interacts with the Höogsteen edge of A (Batey 1999). The protonated cytosine in Figure 1.6, B forms a Höogsteen interaction with the G, and a Waston-Crick G-C pair is observed in the C^+-G-C (Figure 1.6, B).

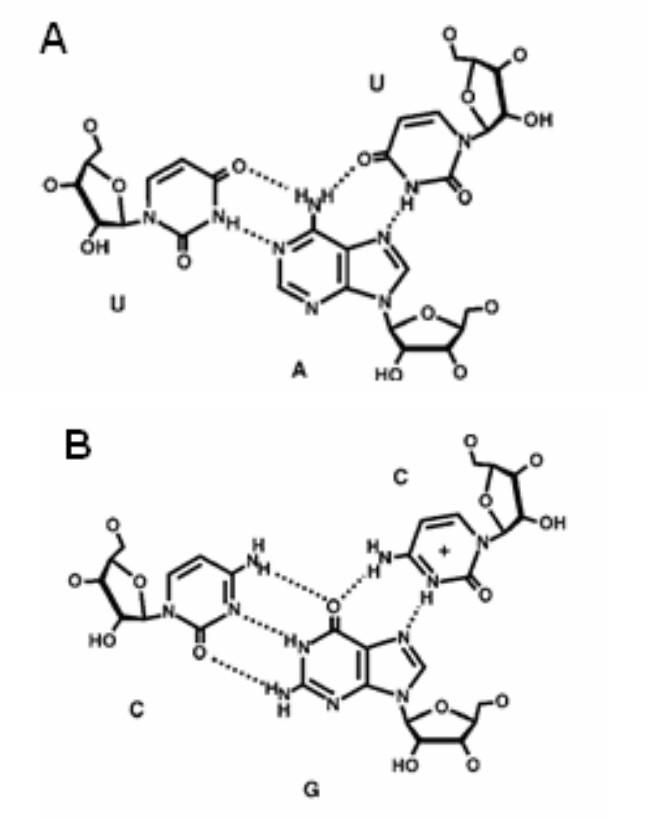


Figure 1.6. Examples of triple base pairs in RNA: U•A-U triple base (A) and C^+-G-C (B) (Batey 1999). In the triple base U•A-U, the adenine form standard Waston-Crick base pairs with the uracil on the left side, while the second uracil interacts with the Höogsteen edge of A through N7 and H62 (Figure 1.6, A). The protonated cytosine in the triple base C^+-G-C forms a Höogsteen interaction with the G, and another G-C pair is regular Watson-Crick base pair (Figure 1.6, B).

For each nucleotide in RNA, the atomic coordinates are defined by seven torsion angles. Six of these torsion angles are related to the sugar-phosphate backbone, and one torsion angle can be used to determine the rotation of the base relative to the sugar (Figure 1.7) (Saenger 1984).

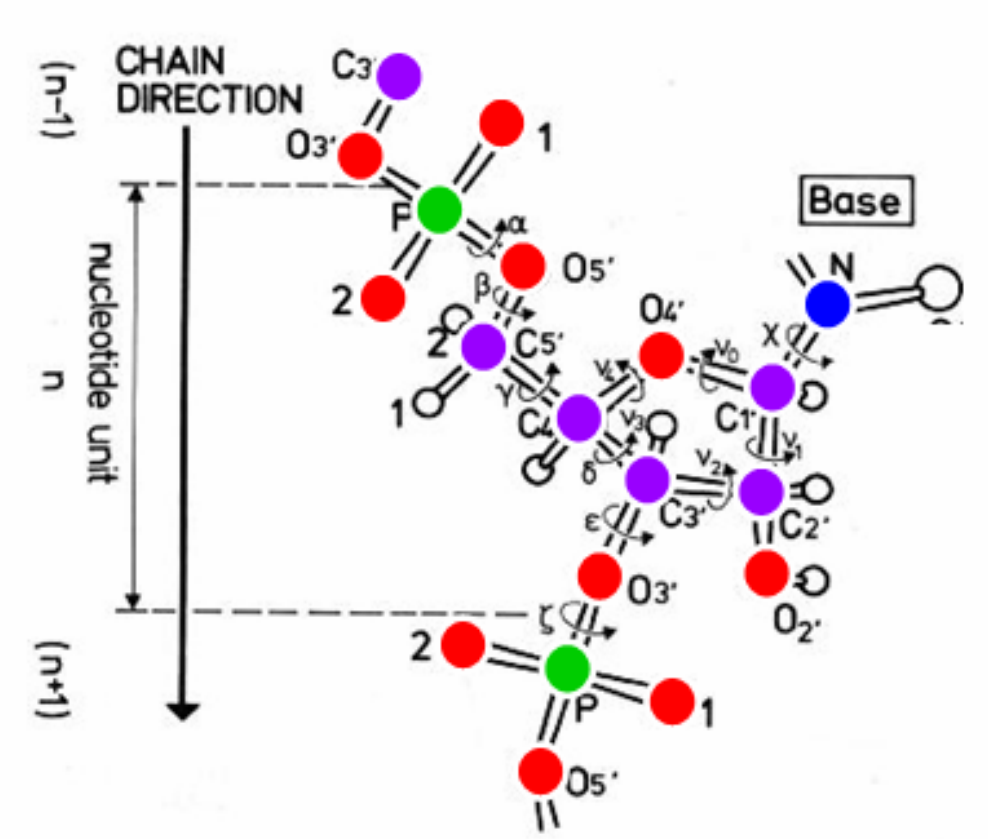


Figure 1.7. The dihedral angles that describe the conformation of one residue in RNA include α, β, γ, δ, ε, ζ, and χ. Different types of atoms in RNA are shown in different color balls, and the purple, blue, red, and green balls represent carbon, nitrogen, oxygen and phosphorus atoms, respectively. (Saenger 1984).

The ribose ring in RNA has two low energy “sugar pucker” conformations, called C2’-endo and C3’-endo (Figure 1.8) (Olson 1982; Olson 1982). The C3’-endo conformation is more common in A-form RNA helices. For the C2’-endo (B-form, found in DNA helices), all the carbon atoms are planar except carbon numbered 2’ (C2’), which is pucked above the plane of the ring, and the H1’-C1’-C2’-H2’ dihedral angle is approximately180º (the numbering of ribose carbons in RNA is labeled in Figure1.1). For C3’-endo (A-form, found in RNA helices). The C3’ is pucked about the plane of the other carbon atoms, and the dihedral angle is approximately 90º.

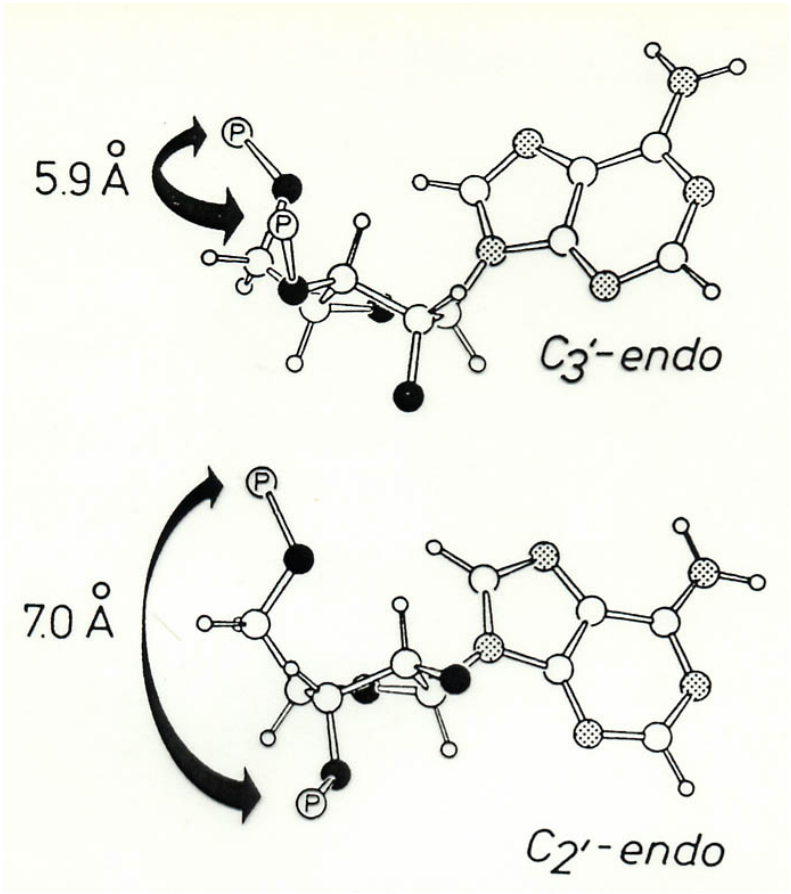


Figure 1.8. The common low energy sugar pucker conformations in nucleic acids. The upper conformation is C3'-endo and is common in RNA, also called A-form helices. The bottom conformation is C2'-endo and is more common in DNA, also called B-form helices (Saenger 1984). Note the difference in P-P distances in the two conformations.

The tertiary structures are difficult to predict due to the huge space of possible backbone conformations and because classical forcefields for RNA require further development, not many computer programs are effective to calculate tertiary structures for certain sequences. The homology modeling method has been used widely for 3D structure predictions of proteins, but only a few examles have been done for RNA (Wallner 2005). De novo prediction of RNA structure remains challenging (Das 2007). Tertiary interactions for a few rRNA and tRNA and group I intron were identified by comparative sequence analysis (Levitt 1969; Gutell 1985; Gutell 1986). However, most tertiary structures of RNA with high resolution were solved using X-ray crystallography and nuclear magnetic resonance (NMR). Although not able to obtain the structures directly, time-resolved fluorescence resonance energy transfer (FRET), hydroxyl radical footprinting, and single-molecule optical traps have been used to identify conformational changes during structural folding and assembly of RNA. More details about using X-ray crystallography, NMR and homology modeling methods will be discussed in Chapter 2.

1.3. Introduction of Ribosome Structure and Function

1.3.1. Ribosome Composition and Function

Ribosome is large ribonucleoprotein particles (RNP) made up of ribosomal RNA (rRNA) and proteins. Ribosome plays an important role as translation machinery in all living cells, and the peptidyl transferase reactions occur inside

the ribosome. Another important function of ribosome in translation is to maintain the accuracy of translation by allowing the mRNA and tRNA interact properly. High accuracy of translation of proteins is essential for all living organisms. Mutations in the RNA or protein components of ribosomes can cause the synthesis of defective proteins, which in turn triggers cell death (Zaher 2009). Thus, there is strong evolutionary selective pressure to retain ribosome sequences that are highly accurate (Zaher 2009). And the ribosome components are conserved in bacteria. The molecular mass for bacterial ribosomes is about 2.5 MDa with more than 4000 rRNA residues. Ribosomes consist of two asymmetric components: the small subunit and the large subunit. The unequally sized subunits have different compositions for bacteria and eukaryotes. The two subunits in bacterial ribosomes are termed 30S and 50S due to their sedimentation coefficients. The “S” refers to the Svedberg unit, and the Svedberg unit is used to measure sedimentation in centrifuge. Each subunit is made up of ribosomal RNA and proteins. For example, the 30S subunit in *E. coli*, with the molecular mass of 0.85 MDa, contains 16S rRNA (1542 nt) and 21 ribosomal proteins (denoted from S1 to S21, proteins in the smaller subunit can be labeled as Sn, where n=1~22). The 50S subunit in bacteria contains 5S rRNA, 23S rRNA and 34 proteins (proteins in the large subunit are labeled as Ln, where n=1~34). In eukaryotes, the subunits sediment with 40S and 60S sedimentation coefficients and the fully assembled ribosomes is 80S. The eukaryotic 40S subunit consists of 18S rRNA and more than 30 proteins, while the 60S consists

of 5S, 5.8S and 28S rRNA and 50 proteins. The tertiary structure of a bacterial ribosome is shown in Figure 1.9 (Yusupov 2001).

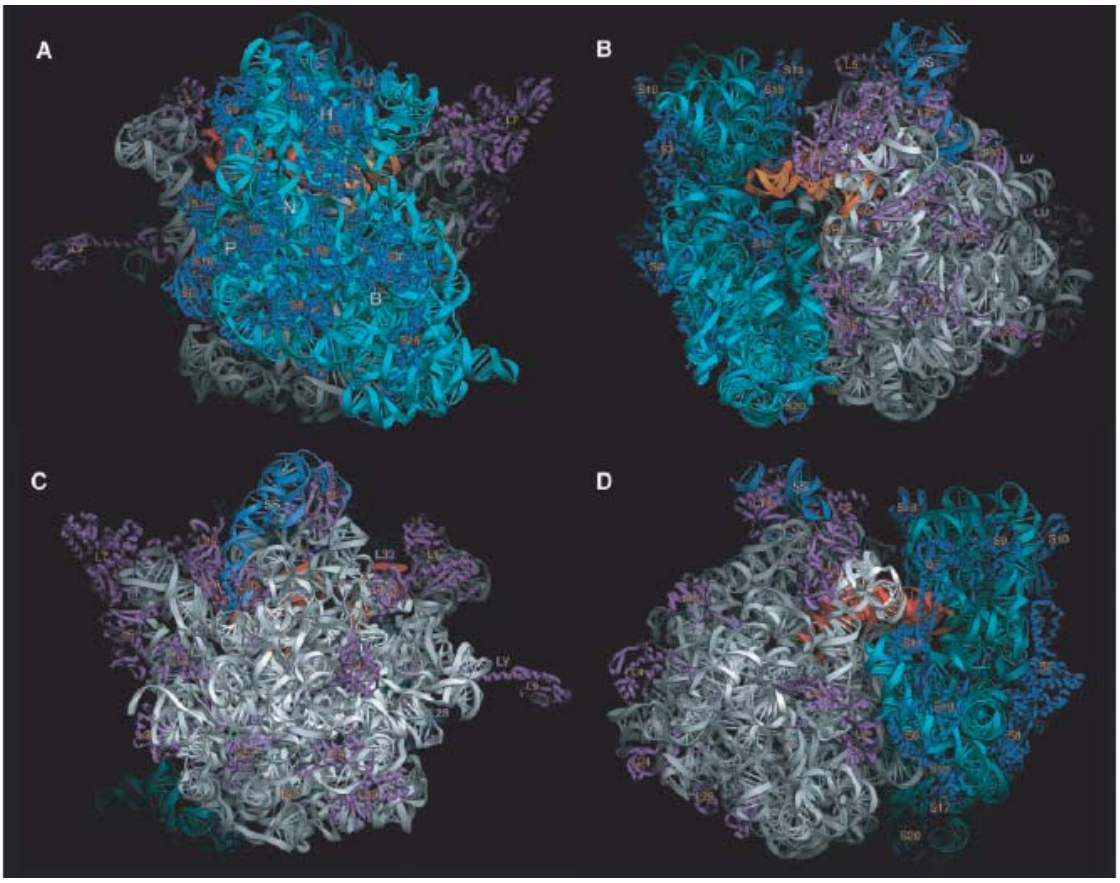


Figure 1.9. The structure of the *T. thermophilus* 70S ribosome. The 30S subunit is on the left and the 50S is on the right. The anticodon arm of the A-site tRNA (gold) is visible in the interface cavity. 5S rRNA, 16S rRNA, 23S rRNA, 30S proteins and 50S proteins are in light blue, cyan, gray, dark blue, and magenta (Yusupov 2001).

Both two subunits of ribosome, with many other cellular components, including mRNA, tRNA, initiation factors, elongation factors and release factors, are involved in translation processes. Translation can be divided into initiation, elongation, termination and ribosome recycling processes. There are three binding sites in the ribosome, the amino-acyl site (A-site), the peptidyl site (P-site) and the exit site (E-site). Both the 30S and 50S subunits contribute to the A-site and P-site, and E-site is mostly made up of the residues in the 50S subunit (Nelson 2004). In the initiation process (in Figure 1.10), the 30S subunit of ribosome binds two initiation factors, IF-1 and IF-3. IF-1 binding at the A-site of ribosome prevents tRNA bind the site, and IF-3 binding can prevent the 30S binds with the 50S prematurely. Then, mRNA binds to the small subunit of ribosome through interactions between the Shine-Dalgarno (SD) sequence on the mRNA and the anti-SD sequence near the 3'-end of the 16S rRNA. The interactions between mRNA and tRNA locate the start codon of mRNA (AUG) on the P-site of the 30S subunit. In the second step of the Figure 1.10, initiation complex is completed after GTP-bound IF-2 and initiator tRNA ($tRNA^{fmet}$) bind to P-site of ribosome and join the complex formed by the 30S subunit of ribosome, IF-3 and mRNA. $tRNA^{fmet}$ is the first tRNA binds with mRNA at the P-site of ribosome. In the third step of the translation initiation process, the large subunit of the ribosome associates with the 30S subunit to form the 70S subunit of the ribosome, simultaneously, GTP-bound IF-2 is hydrolyzed and released from the complex.

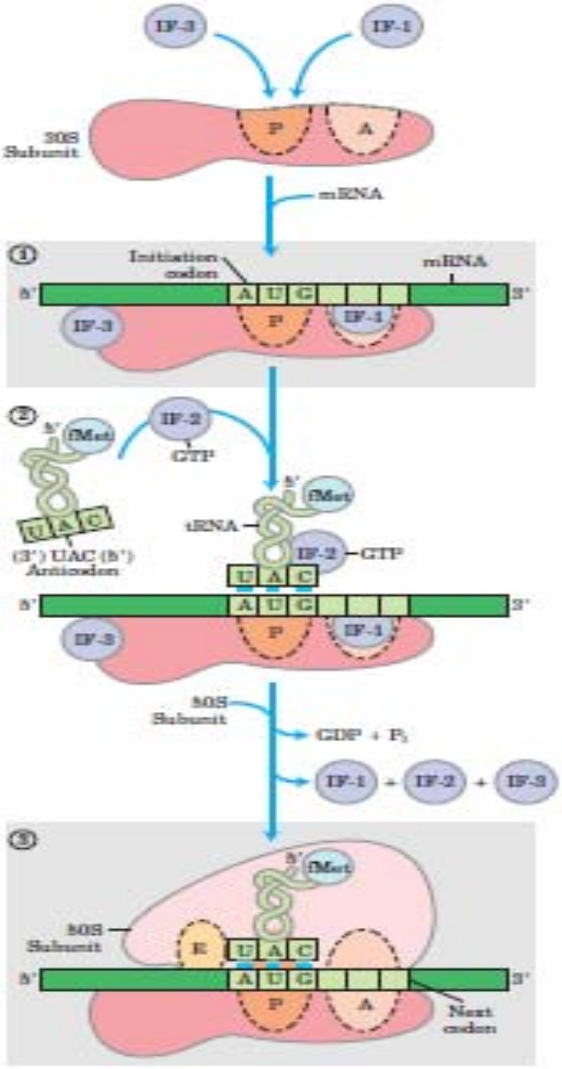


Figure 1.10. Formation of the initiation complex in bacteria (Nelson 2004).

The elongation process is shown in Figure 1.11 (Steitz 2008). When a charged tRNA arrives at the A-site of the ribosome, the anticodon loop of the tRNA is oriented to interact with the next codon of the bound mRNA. The carrier molecule of tRNA, the elongation factor (EF-Tu) is then apart from tRNA after the tRNA binds with mRNA. After the charged tRNA is released in the A-site by the EF-Tu, the new synthesized nascent polypeptide covalently bind with the amino acid on the tRNA in the A-site, extending the linear chain of polypeptide by one residue, the peptidyl transferase reaction. The elongation factor G (EF-G) then moves all tRNAs to another site of ribosome, the tRNA with no amino acid at the 3'-end move to the E-site of ribosome, the tRNA with polypeptide move to P-site, leaving A-site available for binding next charged tRNA. Because the anti-codons of tRNA bind with the codons of mRNA, therefore, the movement of the tRNA also causes the movement of mRNA in the ribosome. The elongation process proceeds until a stop codon in the sequence of mRNA is present, which is not complementary with any anti-codons of tRNA. The stop codon is recognized by a release factor, and new produced polypeptide move to E-site and then is released from the ribosome. The process is called translation termination process. Simply speaking, Ribosome is a nano-machine for protein synthesis. In translation, the obvious functions of ribosome are to select the suitable amino-acyl-tRNA (aa-tRNA) that contains anti-codon complementary with codon of mRNA and to catalyze the formation of polypeptide chains. Although the translation mechanism in bacteria is not the same as eukaryotes, ribosome is the essential nano-machine for protein synthesis in different organisms. That's to say,

the main role of ribosome is to provide appropriate environments for mRNA, tRNA and the proteins used in translation and to catalyze peptide bond formation. Ribosome may have another role, which stabilizes the association between the tRNA and mRNA, because the interactions between codon-anticodon are not very stable in solution (Lipsett 1960; McLaughlin 1966; Steitz 2008).

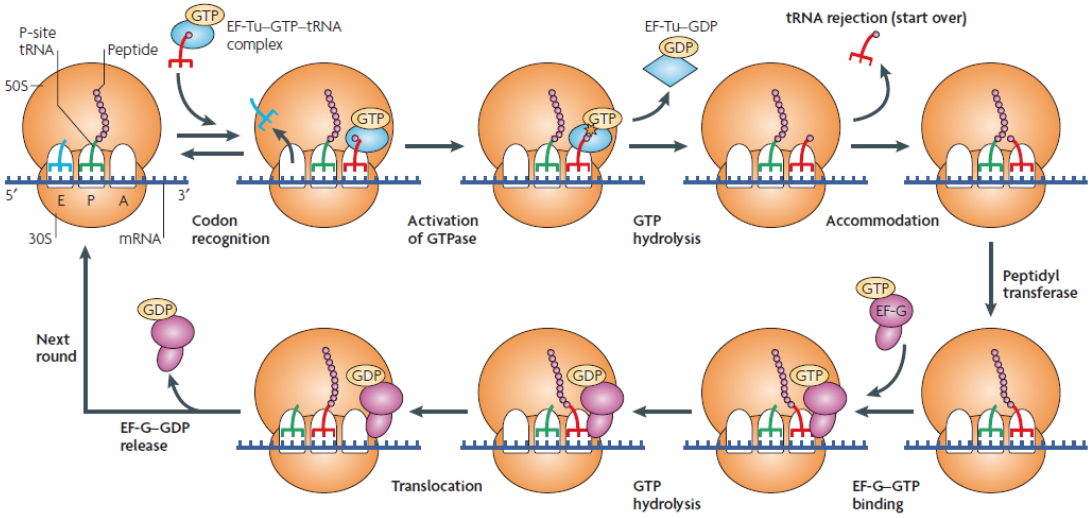


Figure 1.11. An overview of ribosome structure and elongation process of translation in bacteria (Steitz 2008).

The two subunits have different roles in the translation. Several crystal structures of the small subunit, large subunit, and the whole ribosome have been solved by X-ray crystallography (Wimberly 2000; Yusupov 2001; Selmer 2006) as well as cryoelectron microscopy (Gabashvili 2000; Frank 2003). The structures support the hypothesis that the large subunit actually contains the site and critical functional groups to act as a peptidyl-transferase ribozyme (Noller 2005; Strobel 2007) with the help of proteins to make RNA fold correctly and to improve the translation efficiency and accuracy (Ban 2000; Harms 2001; Spahn 2001; Yusupov 2001). The atomic-resolution crystal structures also reveal the positions and interactions of tRNA, mRNA and ribosome components for frozen states of translation. The small subunit contains the decoding site where the mRNA codon and tRNA anticodon sequences interact. In bacteria, the anti-Shine-Dalgarno sequence in 16S rRNA (i.e. CCUCC) is complementary to the 10-12nt Shine-Dalgarno sequence (i.e. GGAGG) in the mRNA. The Shine-Dalgarno sequence is located upstream of the start codon, and the base-paired Shine-Dalgarno interactions between mRNA and 16S rRNA, properly positions the AUG start codon in the P-site so that $tRNA^{fmet}$ can bind and then subunit association occurs (Ban 2000; Nissen 2000; Hansen 2002). Together, these processes are called initiation. And the large ribosomal subunit comprises the active site where peptidyl transfer is catalyzed (Noller 2005). Many ribosomal proteins and the ribosomal RNA are evolutionarily conserved in all three domains of life (Bacteria, Archaea and Eukarya). Some less-conserved proteins may also be used in some organisms for stabilizing the ribosome structures packed by ribosomal RNA and

proteins (Klein 2004). The ribosomal proteins work as accessory factors to improve the accuracy and efficiency of translation, while the proteins are not essential for the translation processes (Klein 2004). Strong evidence lies in the fact that peptidyl transferase activity of ribosomal RNA after removal of a large number of ribosomal proteins still exists although not as good as with the presence of ribosomal proteins (Khaitovich 1999). Several small proteins, such as S4, S5 and S12 are involved in the decoding process (Stöffler 1971; Piepersberg 1975; Andersson 1986; Stern 1989). Nomura's lab obtained the assembly map of how proteins bind to 16S rRNA to form the small subunit *in vitro* (Traub 1969; Traub 1969; Mizushima 1970; Held 1974; Nomura 1974; Culver 2003).

1.3.2. Ribosomal RNA Structural Elements

Watson-Crick base pairs are the main structural elements in rRNA, and about half of all of the nucleotides of rRNAs have Watson-Crick base pairing interactions. Like other RNA, the secondary structures of rRNA contain hairpins, bulges, internal loops and mutli-branched loops. Over 60% of the residues that are not involved in Watson-Crick base pairs are adenine. This suggests that loop residues may contribute specifically to different structures and functions of rRNA (Noller 2005). A-minor interactions were first found in the hammerhead ribozyme and in the group I ribozyme (Pley 1974; Scott 1995; Cate 1996), and A-minor interactions have been also found in rRNA (Fig, 1.12, A and B) (Ogle 2001; Noller 2005). A-minor interactions are usually minor-groove interactions between

an adenosine and the minor groove of Waston-Crick base pairs (typically GC base pairs) (Figure 1.12, A and B).

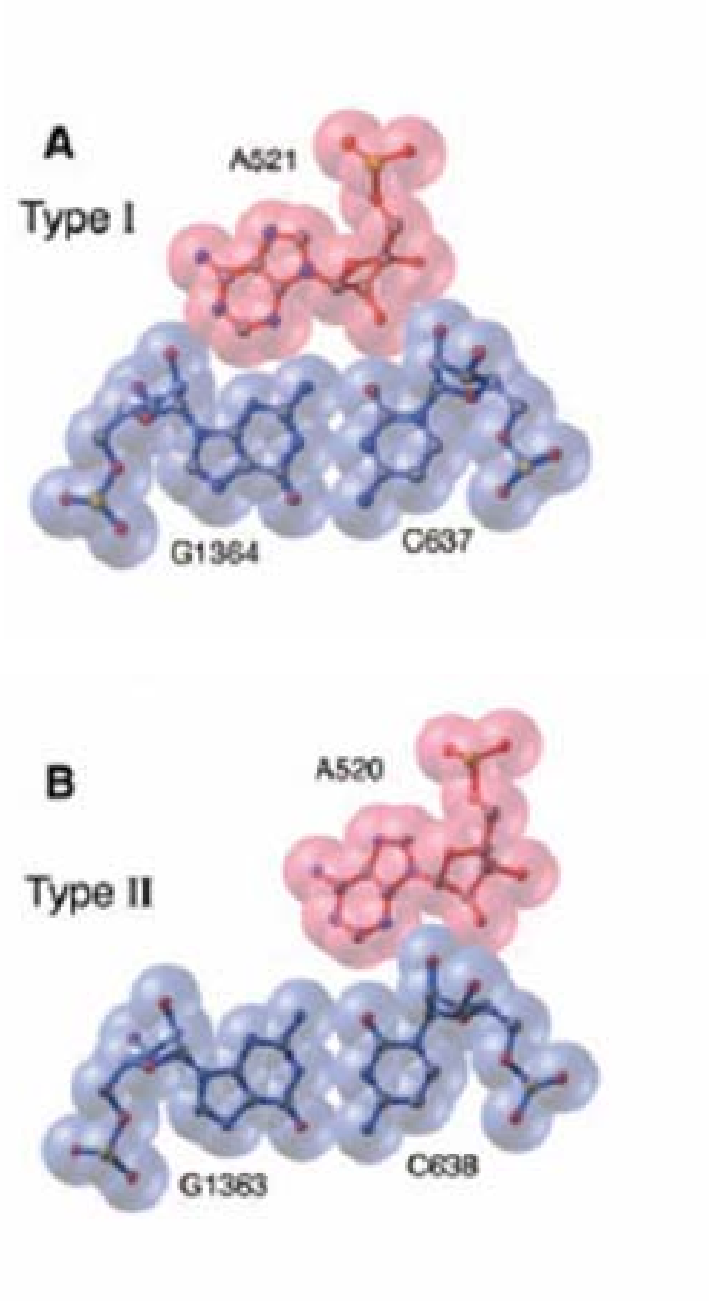


Figure 1.12. Two types of A-minor interactions extensively found in 16S and 23S rRNA (Noller 2005).

Numerous A-minor interactions exist within folded rRNAs. The crystal structure of the 16S rRNA of *T. thermophilus* revealed more than 50 A-minor motifs (Wimberly 2000; Noller 2005). Some A-minor interactions are long-range interactions involving residues that are located distantly in the secondary structure. Importantly, the A-minor interactions allow dynamic yet precise interactions between ribosome and tRNA, which occur due to the conformational changes of ribosome that occur during the different stages of translation and translocation (Nissen 2000; Ogle 2001; Noller 2005). A-minor interactions of 16S rRNA also help to improve the stereochemical fit of codon-anticodon pairing between mRNA and tRNA. The 3'-terminal adenosines of A- and P-site bound tRNAs interact with the peptidyl transferase center of 23S rRNA by A-minor interactions (Ogle 2001; Noller 2005).

Besides A-minor interactions, many other local folding motifs are found in rRNA, including the U-turn (Stallings 1997). U-turn changes the direction of the backbone of RNA, and there are many examples of the U-turn in rRNA (Stallings 1997). A U-turn is formed with a consensus sequences if UNR and is U (n)-R (n+3), stabilized a hydrogen bond between the N3 of U (n) and phosphate group of n+3 residue, and by a second H-bond from the N7 of the R residue (i.e. A or G) to the 2'-OH of the n residue (shown in Figure 1.13).

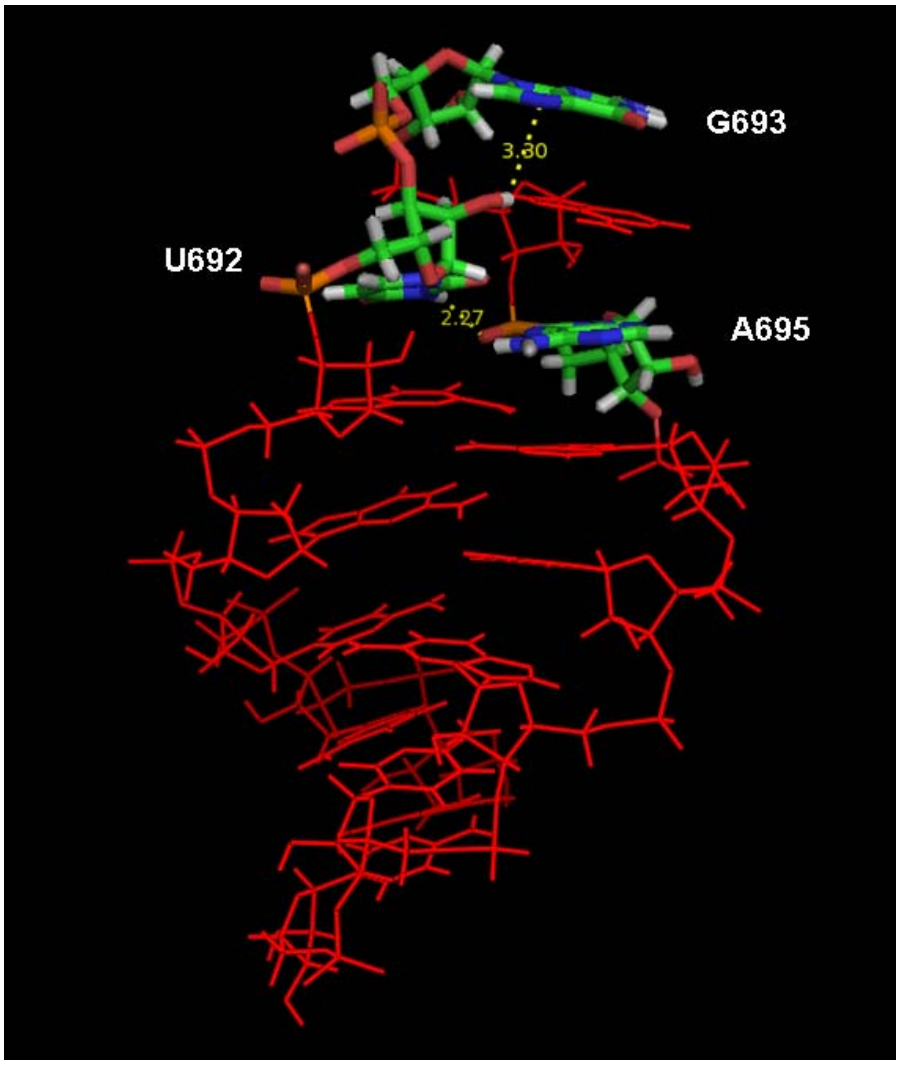


Figure 1.13. The “U-turn” is formed in the structures of the wild-type of the 690 loop of *E. coli* 16S rRNA (1FHK). Residues U692, G693 and A 695 are highlighted in color.

1.3.3. Ribosome Dynamics

The ribosome undergoes numerous dynamic conformation changes during assembly and as factors and tRNAs are bound and released and chemical reactions occur (peptidyl transfer and hydrolysis). The position of tRNA changes from the A-site where it enters the ribosome as aminoacyl-tRNA (except the first tRNA that enters the P-site of the ribosome to initiate the translation) to translocate the P-site of the peptide bond formation and finally translocates to the E-site where tRNA deacylated. Translocation of ribosome is necessary for the accuracy and efficiency of translation by facilitating the movements of tRNA during translation. Firstly, the acceptor arms of tRNA move to P- and E-sites of the large subunit while the anticodon ends of tRNA remain in the A- and P-sties of the smaller subunit, which is called the “hybrid state model” (Moazed 1989). Second, the anticodon ends of tRNA shift to the P-and E-sites of the smaller subunit (Moazed 1989). The movement of the ribosome is supported by the relative distance changes of residues in the crystal structures of the ribosome in different functional states (Cate 1999; Cater 2001; Harms 2001; Ogle 2001; Yusupov 2001; Noller 2002). The cryo-EM results the ribosome in different function states (Frank 2000; Frank 2003; Valle 2003). In addition, the L1 arm of the 50S subunit has been found to be hindrance for the release of deacylated-tRNA from E-site of ribosome, while the deacylated-tRNA is able be released if L1 arm of the 50S subunit move downward about 20 Å (Moazed 1989; Yusupov 2001; Noller 2002). The positions of the L1 arm of the 50S subunit are indeed obviously different in the *Dinococcus radiodurans* crystal compared with the L1

arm positions in the *T. thermophilus* crystal, and the difference is around 20 Å (Harms 2001; Noller 2005). The relative positions of 30S and 50S change during translocation: the rotational movement is about 6º and the maximum distance change is about 20 Å (Frank 2000; Valle 2003). Crystal structures of the ribosome in *E. coli*, also support multiple conformation changes such as the L1 stalk conformations in the 50S subunit in the empty ribosome and in the ribosome bound with E-site tRNA are different (Selmer 2006).

1.4. rRNA for Antibiotics Developments

Natural antibiotics tend to bind the highly conserved regions of RNA, DNA or proteins in cells. The ribosome catalyzes protein synthesis in all cells, and there are several sites in the ribosome whose sequences are conserved within bacteria, but not in humans and those sites are excellent targets for development of new antibiotics. Interactions between ribosome and specific antibiotics affect translation processes in various ways. Some specific interactions among tRNA, mRNA and rRNA are interfered by binding of an antibiotic to the highly conserved regions of the 30S rRNA, thereby hindering the translation. For example, the aminoglycoside antibiotics bind to the 30S ribosome are able to inhibit the translation of initiation step, and to change the conformation of the ribosome such that it is more tolerant for binding non-cognate tRNAs in the A site, and thus increase mRNA misreading errors (Moazed 1987; Woodcock 1991; Karimi 1994; Pape 2000). Ramakrishnan *et al.* obtained crystal structures of the 30S subunit bound with paromomycin, streptomycin and spectinomycin (Carter 2000). These

crystal structures support the hypothesized mechanism of miscoding in translation due to the interactions between rRNA and the antibiotics (Carter 2000). Puglisi *et al.* determined the NMR structures of A-site rRNA with and without aminoglycosides paromomycin and gentamycin C1a to elucidate the structural changes of A-site rRNA after binding with these antibiotics (Lynch 2003). Antibiotics, such as chloramphenicol, lincomycin, clindamycin and macrolides bind to the 50S ribosome subunit and hinder translation by different mechnisms. Chloramphenicol targets the peptidyl transferase cavity. Clindamycin or other macrolides physically hinders the peptide elongation by sterically blocking the exit channel (Schlünzen 2001). Tetracycline class antibiotics inhibit the transfer of activated amino acids to the ribosome by affecting interactions between A-site tRNA and the acceptor site of 30S rRNA. Edeine inhibits the initiation of translation by binding to 16S rRNA in the 30S subunit (Pioletti 2001).

Many antibotics have been found to target ribosome specifically and effectively, and several convincing several structures of ribosome bound with different antibiotics are shedding, and this knowledge is being used to develop next generation anti-infectives (Böttger 2001; Pioletti 2001; Schlünzen 2001; Moore 2002). The interaction modes between antibiotics and ribosomal RNA and can provide valuable information for antibiotics selectivity. Studies of antibiotics targeting ribosome are also an important tool for exploring ribosome function and the interaction modes between ribosome and others, which can be useful for designing new antibiotics with improved properties, such as lower susceptibility to

drug resistance. Resistance to antibiotics is a challenging problem in developing antibiotics, which requires design new antibiotics.

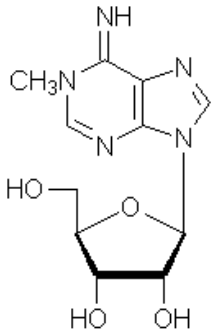
1.5. RNA Modifications

Chemical modifications are introduced into RNA post-transcriptionally and catalyzed by specific enzymes. In some cases, small nucleolar RNA (snoRNA) are involved in RNA modifications (Ye 2007). The chemical modifications are distributed widely in the bases and the sugars of RNA sequences with 107 unique modifications identified to date (Crain 1997; Agris 2004; Aduri 2007). More than 3000 modified nucleotides in RNA have been found in natural organisms (Czerwoniec 2009). The regular experimental methods that are used to identify modifications include modification-induced inhibition of the enzyme reverse transcriptase combining with HPLC analysis, electrospray ionization mass spectroscopy (ESI-MS) and digestion by RNase combining Matrix Assisted Laser Desorption Ionization mass spectroscopy (MALDI) analysis (Bakin 1993; Kowalak 1993; Andersen 2004). More than one hundred types of modified nucleotides exist in natural RNA (Crain 1997; Agris 2004; Aduri 2007). The main modifications include the methylation of bases, methylation of 2'-hydroxyl groups of ribose (2'-O-methylribose), and a variety of unusual bases, including pseudouridine (Ψ), ribothymidine (T), Inosine (I) and dihyrouridine (D). Both pseudouridine (Ψ) and ribothymidine (T) are frequently found in the TΨC loop of tRNA. In pseudouridine, unlike the common glycoside linkage between ribose and base in RNA, the pseudouridine linked with the ribose by C-C bond instead

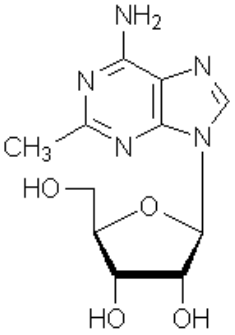
of N-C as the connection between uracil and ribose (Desaulniers 2008). The website of our lab [ozone3.chem.wayne.edu] provides AMBER parameters for 103 modifications that have been found in RNA (Aduri 2007). The website provides the AMBER parameter files, the input and output files for developing the parameters are available. The modifications appear frequently in tRNA and rRNA, for example, 16S and 23S rRNA of *E.coli* have 11 and 23 modified nucleotides, respectively (Rozenski 2000). Table 1.1 lists the modifications found in ribosomal RNA [http://rna-mdb.cas.albany.edu/RNAmods/cgi-bin/rnafind.cgi]. The natural modified nucleotides are likely to be related to some important functions of RNA based on the fact that most modified nucleotides in rRNA are located around the functional centers of the ribosome from the ribosome models built by Decatur *et al.* and some modified nucleotides are highly conserved in or very close to the anticodon loop of tRNA (Bruce 1982; Decatur 2002; Agris 2007).

Table 1.1. The modifications exist in ribosomal RNA.

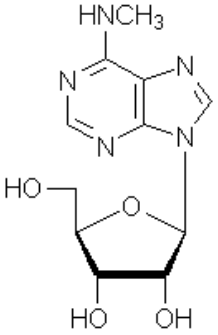
Modifications of adenosine



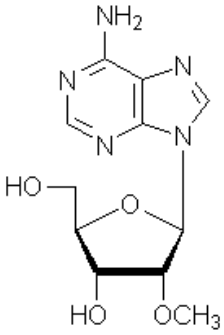
1-methyladenosine
(m^1A)



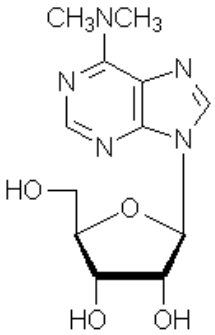
2-methyladenosine
(m^2A)



N^6-methyladenosine
(m^6A)

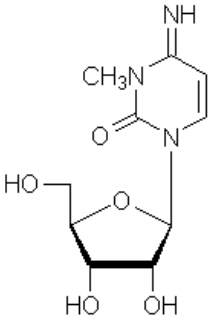


2'-O-methyladenosine
(Am)

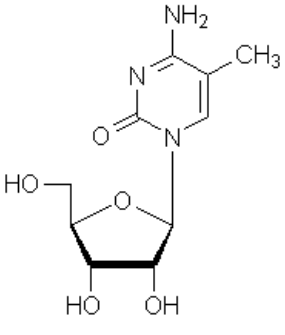


N6,N6-
dimethyladenosine
(m62A)

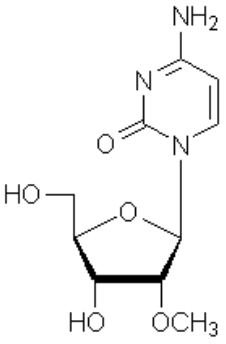
Modifications of cytidine



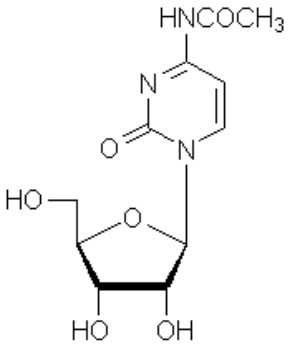
3-methylcytidine
(m^3C)



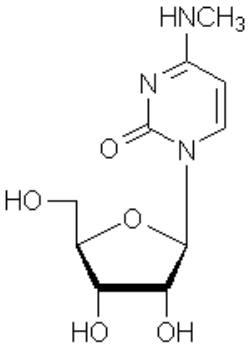
5-methylcytidine
(m^5C)



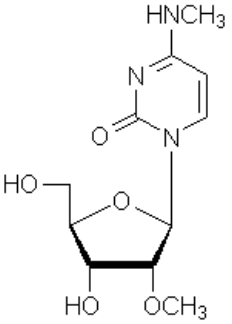
2'-*O*-methylcytidine (Cm)



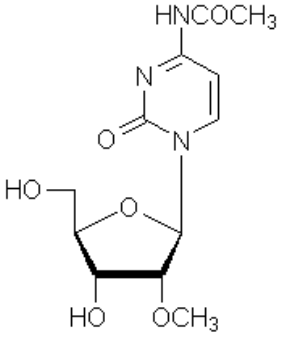
N^4-acetylcytidine
(ac^4C)



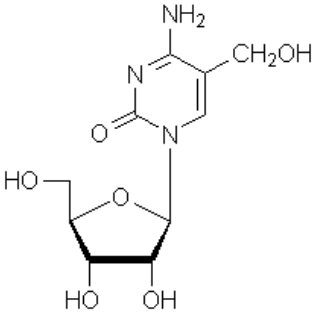
N^4-methylcytidine
(m^4C)



N^4,2'-O-dimethylcytidine
(m^4Cm)

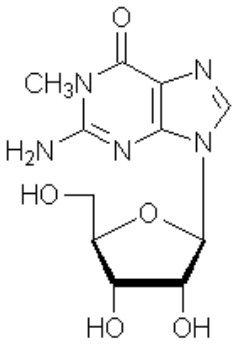


N^4-acetyl-2'-*O*-methyl-cytidine (ac^4Cm)

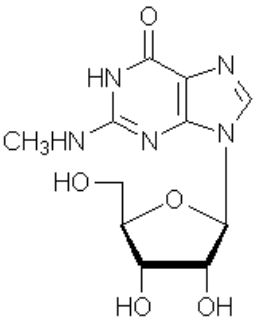


5-hydroxymethylcytidine
(hm^5C)

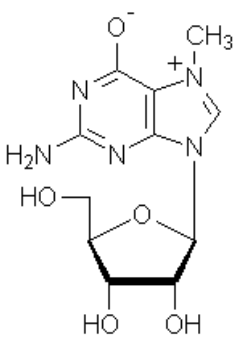
Modifications of guanosine



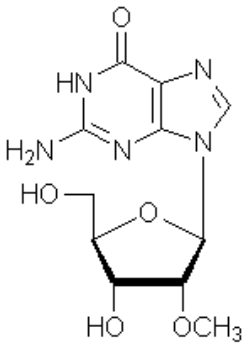
1-methylguanosine
(m^1G)



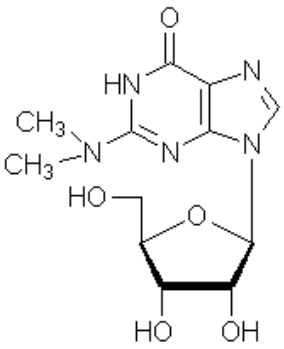
N^2-methylguanosine
(m^2G)



7-methylguanosine
(m^7G)

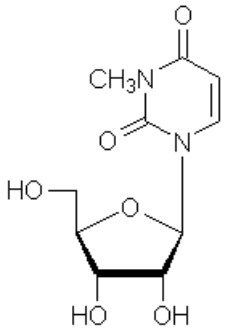


2'-O-methylguanosine (Gm)

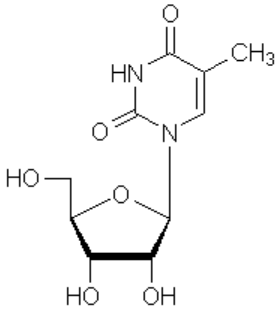


N^2,N^2-dimethylguanosine (m^2_2G)

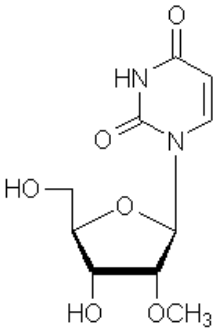
Modifications of uridine



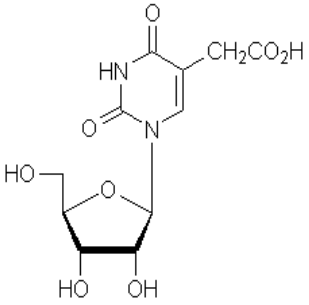
3-methyluridine (m^3U)



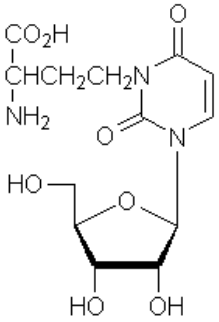
5-methyluridine (m^5U)



2'-*O*-methyluridine (Um)

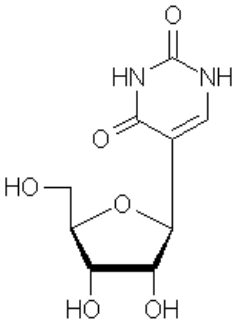


5-carboxymethyluridine (cm^5U)

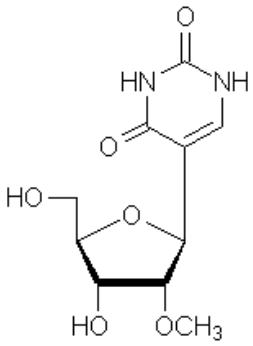


3-(3-amino-3-carboxypropyl)uridine (acp^3U)

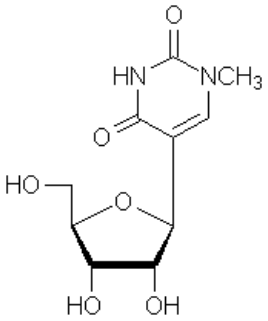
Pseudouridine and the modifications



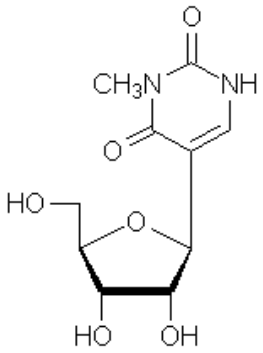
Pseudouridine (Ψ)



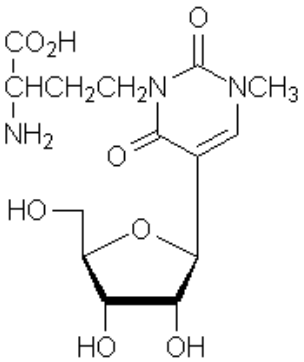
2'-*O*-methylpseudouridine (Ψm)



1-methyl pseudouridine ($m^1\Psi$)

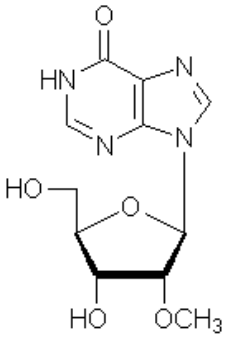


3-methylpseudouridine ($m^3\Psi$)

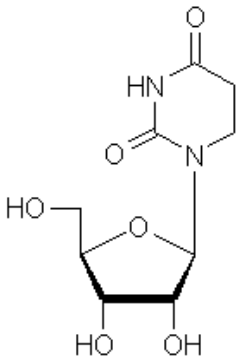


1-methyl-3-(3-amino-3-carboxypropyl) pseudouridine ($m^1acp^3\Psi$)

Other modifications



2'-*O*-methylinosine (Im)



Dihydrouridine (D)

CHAPTER 2

Methods for Solving Atomic-resolution Structures of Biomolecules

2.1. Introduction of Structure Determination Methods

RNA participates in several important biological functions: passing genetic information from DNA to proteins and catalysis in RNA processing and protein synthesis (Gesteland 1993). The diverse functions of RNA require various structural and conformational changes. Determining structures of different RNAs can greatly widen our understanding of biological roles of RNA. Relationships between structures and functions of RNAs can be used to design biomolecules with desired functions. In addition, the functional and evolutionary hints provided by 3D structures are valuable for drug design development [refer to Chapter 1]. The 3D structures of biomolecules can be obtained by two methods, either by experimental information, mostly from X-ray crystallography, cryoelectron microscopy and NMR spectroscopy, or purely theoretical methods such as homology modeling.

2.2. Comparison of X-ray Crystallography and NMR Methods

X-ray crystallography and NMR Spectroscopy are methods that have been used widely to provide atomic-resolution structures of large-sized biological molecules, including proteins, lipids, nucleic acids, and their complexes. To date, about 99% of the protein and nucleic acid structures in the Protein Data Bank

(PDB) have been determined by X-ray crystallography and NMR methods. Electron microscopy and IR spectroscopy make up most of the remaining structures (http://www.pdb.org/pdb/static.do?p=general_information/pdb_statistics/index.html).

The first crystal structures were solved in 1958 by Max Perutz for myoglobin and Sir John Cowdery Kendrew (they won a Nobel Prize for this achievement) by using the heavy atom method to solve the phase problem of X-ray crystallography (Hauptman 1991). X-ray crystallography is used to obtain structural information of macromolecules based on the diffraction pattern produced by bombarding a single crystal with X-rays. In some cases, powder diffraction and thin film X-ray diffraction coupled with lattice refinement algorithms have been used to obtain information about crystals when single crystals are not available (Moore 1997).

The first NMR experiment for biological molecules was done about 60 years ago (Saunders 1957), and the structure determination of a globular protein in solution was first solved in 1984 (Willionaon 1985). The development of NMR methods applied to the structures of proteins was pioneered by Kurt Wüthrich, who shared the Nobel Prize in 2002 (Wüthrich 1976; Wüthrich 1986). Nuclear magnetic resonance involves the interaction between the magnetic moments of atomic nuclei (i. e. nuclei with $I \neq 0$) and in the presence of a static homogeneous magnetic field, Bo along the Z-axis. The magnitude of the observed NMR signal is dependent on the tiny population difference between the two spin states of I=1/2 nuclei depending on the magnetic field strength. For an 11.7T magnetic

field, the population difference of the two I=1/2 spin states, α and β, is approximately one out of 10,000 at equilibrium. A resonant radio frequency pulse that satisfies the Bohr condition ($\Delta E = h\gamma$) applied to a sample can be absorbed by the molecular nuclei, which causes the population distribution between two spin states to change including the creation of coherence (i.e. creation of X-Y magnetization). Relaxation processes then result in photon emission that can be detected by a radiofrequency tuned Helmholtz coil, thereby creating an AC electrical current that is amplified and digitized. The time domain signal is then converted to frequency domain by Fourier Transform methods (Ernst 1966). Relaxation is the phenomenon in which the nuclear spins of perturbed population return to equilibrium distribution by the interactions of spins with the magnetic moments of neighboring spins (T_2 relaxation) and the environment (T_1 relaxation). Due to the low population difference of α and β spin states, NMR methods are insensitive compared to other spectroscopic techniques, such as UV and IR spectroscopy. Most biological NMR focuses on nuclei with spin 1/2, because they have no quadrupole moments and thus give rise to sharp spectral lines. Fortunately, the ^{1}H, ^{13}C, ^{15}N, and ^{31}P, isotopes found in RNA have spin 1/2. The proton, the most abundant nucleus in organic compounds, has the largest nuclear magnetic moment other than tritium and hence is the most sensitive nuclei measured by NMR methods. Below is a description of how NMR is used to extract distance, dihedral, and bond vector restraints for structural modeling.

X-ray crystallography and NMR spectroscopy have their distinctive advantages and disadvantages in solving 3D structures of bio-molecules at atomic-resolution. X-ray crystallography has been used widely to solve the structures of large biomolecules, such as the subunits of ribosome (Cate 1999; Wimberly 2000; Yusupov 2001; Selmer 2006). X-ray structures are, to date, the most reliable source of 3D structures when a diffraction quality crystal is available. Usually, tedious experiments sampling a wide variety of solution conditions are required to obtain protein or nucleic acid crystals that diffract to high resolution. Different buffer solutions, pH, precipitating agent, *etc*, need to be screened to obtain high-quality crystals and still, some bio-molecules are still not able to be crystallized after huge effort. Even if crystals can be obtained, some molecules have different structural and dynamic properties in crystals than in solution. For example, some RNA sequences prefer to crystallize in less biologically relevant duplex structures, even though their sequences form very stable hairpin structures in solution (Holbrook 1991; Wahl 1996). Besides, many factors during crystallization may cause crystal imperfections, such as, temperature variations when forming crystal, impurities precipitated with the sample crystals, twinning, or multiple stable conformations. These factors may prevent researchers from obtaining atomic-resolution protein or RNA structures at physiological conditions. It is also often difficult to obtain isomorphous heavy atom derivatives for solving the crystallographic phase problem. In some case, protein residues on the surface have been found to have different structure and in crystals than in solution (Snyder 2005). For RNA structure studies, X-ray crystallography is

sometimes not suitable because crystal packing interactions are severe and this is particularly true for small RNA due to the high surface area of RNA and hence crystal packing interacations are likely to be more troublesome in RNA crystallography. NMR has many distinctive advantages in RNA study although the molecular weight limit of molecules that can be measured by NMR is smaller than X-ray crystallography. Recent developments of Transverse Relaxation Spectroscopy (TROSY) (Fiaux 2002; Fernández 2003) and residual dipolar couplings (RDC) (Prestegard 2000) and advanced assignment methods have pushed the molecular size limit and improved the quality of NMR structures using isotope labeling methods (Nikonowicz 1992; Hines 1994; Batey 1995). The use of Small angle X-ray scattering (SAXS) provides low-resolution electron density for RNA in solution and is advancing the molecular weight limits of NMR (Zuo 2008; Zuo 2010), which are discussed in later part of this chapter.

Comparing with x-ray crystallography, one of the strengths of NMR is that it can provide dynamic structures of biomolecules under physiological solution conditions. NMR is limited to small biomolecules (currently fewer than 200nt for nucleic acids), because peaks are overlapped intensely in large-sized molecules, and is also limited to stable, soluble protein and/or RNA solutions that do not aggregate or degrade easily. Another problem in NMR is that the z-magnetization relaxes slower in larger molecules and thus less time can be used to detect the larger proteins or nucleic acids (de Alba 2000). In addition, larger molecules have x-y magnetization that relaxes faster, which causes broader linewidths (this is called the spin diffusion limit), thereby degrading both resolution and sensitivity.

Some of these problems have been alleviated by the improvements in NMR hardware, including higher magnetic field strength, superconducting magnets and development of cryoprobes. The Bruker Avance-700MHz at Wayne State University is equipped with a ^{1}H, ^{13}C, ^{15}N triple resonance cryoprobe with Z-axis pulsed field gradient coil has been used to collect most of the NMR data in this thesis. The use of the cryoprobe improves signal-to-noise ratio to be around 6000:1 for 0.1% ethylbenzene at room temperature. The development of modern NMR methodology has also had a large impact on advancing applications of NMR, such as Fourier transform spectroscopy, pulse field gradients and multi-dimensional experiments (Deltour 1965; Aue 1976). In the last twenty years, the introduction of Isotope-labeled samples has widened the applications of NMR studies of bio-molecules by alleviating the overlapping problem of large-sized molecules (Batey 1992; Nikonowicz 1992; Santoro 1992; Hines 1993; Hines 1994; Batey 1995). Isotopic labeling with ^{13}C and ^{15}N or selective observation of partial regions of the samples provides the possibility of resolving overlapped resonances and unambiguous resonance assignments in multidimensional spectra (Nikonowicz 1992; Hines 1994; Batey 1995). The improvements in NMR are designed to overcome the two main problems with NMR for determining the structure of bio-macromolecules: signal-to-noise ratio and spectral overlap.

Based on the previous introduction of the NMR and x-ray crystallography, the two methods are complementary in some ways. For NMR, sample preparation is relatively straightforward and expeditious, but data collection and resonance assignments can take a long time. For X-ray crystallography, on the

other hand, data collection and processing is much easier but a lot of crystallization conditions need to be screened and then optimized for getting diffraction-quality crystals. NMR spectroscopy provides data that are complementary to data from X-ray crystallography in many ways. The complementary information for residues on the surface of a protein from the two methods is of special interest, because protein functions depend largely on the nature of molecular surface areas. Therefore, combining NMR spectroscopy and X-ray crystallography can minimize the limits of solving structure problems, obtain more bio-molecular structures and may give a clearer insight into the relation between structure and function of protein molecules (Wüthrich 1990). Slow conversion between different structures results in multiple resonances observed by NMR. Thus, NMR can be used as a tool for screening of the best conditions for formation of pure monomeric or dimeric form of certain RNA molecules, which in turn can be used for crystallographic screening (Page 2003; Page 2005). Some researchers, such as Kurt Wüthrich, Cheryl Arrowsmith, Gaetano Montelione, have done beautiful work on the combination of NMR and X-ray crystallography for protein structural genomics (Page 2005; Snyder 2005; Yee 2005).

In my research, the structures of two hairpins in *E.coli* ribosomal RNA have been determined by using NMR to identify the structural effects of mutations, chemical modifications, and peptide binding. I will introduce more detailed information of NMR methodology for better understanding the experimental data and the approach used to determine RNA structure.

2.3. Introduction to NMR for RNA Study

The conserved sarcin-ricin loop (SRL) of 28S rRNA in the larger subunit interacts with lethal ribotoxins α-sarcin and ricin is one of the first rRNA structures obtained by NMR method (Heus 1991; Szewczak 1993; Allain 1995). NMR spectroscopy is a powerful method for studying the tertiary structures and dynamics of nucleic acids in solution. In addition, interactions of RNA with ligands, small proteins, and ions are amenable to NMR methods (Fürtig 2003). The Watson-Crick and non-standard base-pairing pattern can be used to determine the base-pair dynamics and to verify of the secondary structure of a RNA sample (Fürtig 2003). The imino protons of guanines and uracils can be observed in solvent water (H_2O) and are usually located between 10 and 15 ppm, and the H-bonded imino protons usually have downfield chemical shifts due to in-plane aromatic ring currents and electronic redistribution due to H-bonding. In contrast, the non H-bonded imino protons often show up in the region of 9.5-12.5 ppm due to stacking ring current effects and lack of H-bonding. The conversion between different structures in solution can be observed in NMR experiments and can be used to identify the conformational equilibrium among multiple structures. For example, the percentages between duplex and bulge at different sample concentrations, salt concentrations, and temperatures (Grüne 1996). Pulse-field-gradient NMR experiments have been used to measure diffusion constants to differentiate duplex RNA from hairpin RNA (Lapham 1997). Site-specific information about metal ion, protein, or nucleic acids binding to RNA can also be detected by NMR. Deshielding effects or a structural transition of RNA after

binding with a ligand may cause chemical shift changes, line broadening, or new NOE peaks to appear in NMR spectra (Rüdisser 2000; Schmitz 2000). Therefore, NMR can be used to map interaction sites based on the spectral change between the free and bound states of RNA samples.

2.3.1. NMR Study for Determining RNA Structures

NMR is a powerful method capable of providing tertiary structural information of RNA in solution by indirectly extracting the structural restraints from NMR spectra. The numbers of 3D structures of biomolecues solved by in the last 10 years has increased exponentially. To date, about 8,000 structures of RNA/protein complexes have been determined by NMR and deposited in the Protein Data Bank [http://www.pdb.org/]. Figure 2.1 shows the increasing tendency in solving bio-molecular structures by NMR.

The routine NMR procedures used to determine the 3D structures of RNA in our lab are shown in Figure 2.2.

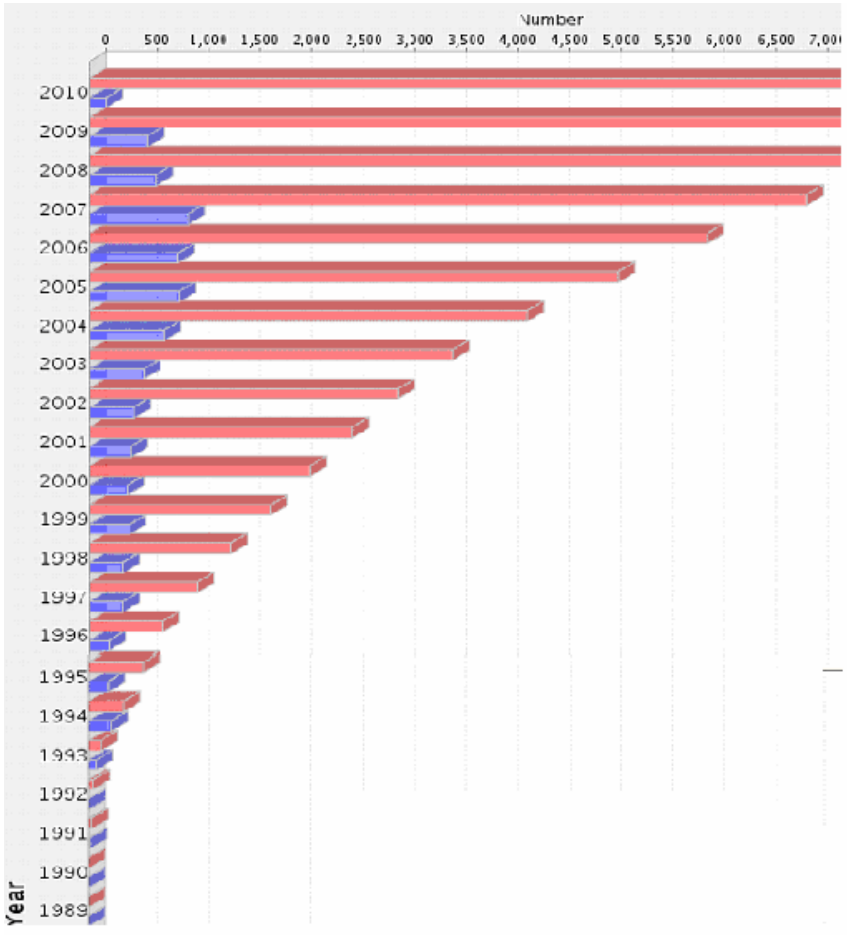


Figure 2.1. Yearly growth of structures solved by NMR from Protein Data Bank. The blue bars and red bars represent for the NMR structure numbers solved in yearly and the totally [http://www.pdb.org/].

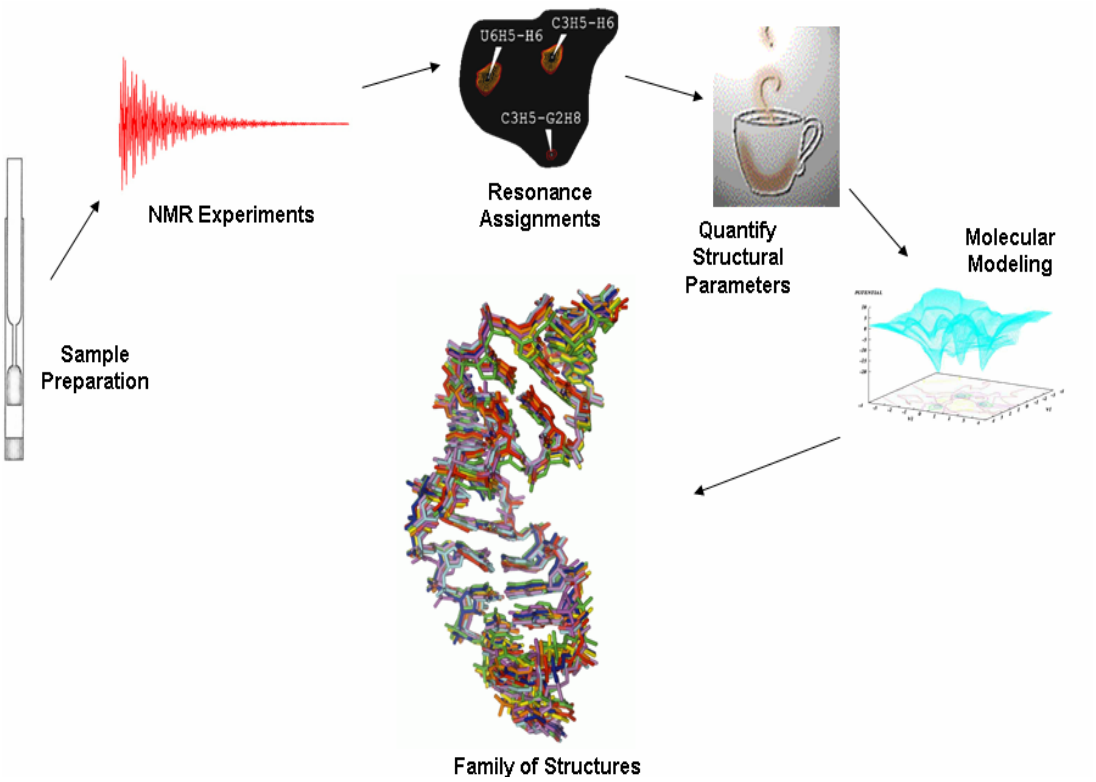


Figure 2.2. The routine procedures for solving RNA solution structures by NMR in our lab (Clos 2007). The first step is to prepare samples that are suitable for NMR study, and then, and then collect NMR data. The most time-consuming step is to analyze NMR spectra to assign peaks and to obtain structural restraints that are input in CNS software to run calculation and to obtain NMR structures of samples in later steps.

Protons are distributed in 1D-^{1}H spectra based on their chemical shifts. The exact resonance frequency of different atoms depends on the chemical environment of each spin, and these differences are called chemical shifts. The assignment process is divided into the identification of individual spin systems, residue type and the connection of neighboring nucleotides to obtain sequential assignments (Wüthrich 1976). The identification of individual spin systems can be obtained based on the chemical shift ranges of carbons that are bonded to different kinds of protons. The first step of a structure determination using NMR is to assign the chemical shifts of all the NMR-active atoms in biomolecule. The protons in RNA can be divided into two types: exchangeable protons, including imino and amino protons, and non-exchangeable protons, which include base protons (H5 or H6 of pyrimidines, H8 of purines, H2 of Adenine and sugar protons: H1', H2', H3', H4', H5', and H5'') . Both exchangeable and non-exchangeable protons can be observed in the NMR spectra if the samples are dissolved in H_2O, only non-exchangeable protons are observed in spectra if solvent is D_2O. Once enough NMR signals are assigned, the experimental restraints can be quantified and then used in structure calculations. However, in 1D NMR, peaks of many protons are overlapped because the protons in RNA samples are highly overlapped and resonate in a narrow range of 10 ppm (Figure 2.3 is the 690UC mutant in H_2O). The non-exchangeable protons are usually observed in the region 3.5~8 ppm, which makes peaks hard to distinguish from each other. The most important use of 1D-^{1}H in H_2O is that hydrogen bond information can be obtained based on imino peaks. Imino peaks are resolved

well in the region 10~14 ppm. The hydrogen-bonded imino peaks tend to be downfield shifted in the spectra, and imino protons of uridine H3 usually have larger chemical shifts than guanosine H1, due to the larger ring current from adenine (i.e. in a U-A base pair), then that from C (i.e. in a G-C base pair). The assignments of imino peaks can be completed with the help of 1D NOE difference spectrum (i.e. with vs. without resonance saturation) or 2D NOESY experiments. The H-bonding pattern can be specifically determined using through space J-coupling with HNN-COSY (Dingley 1998; Wöhnert 1999).

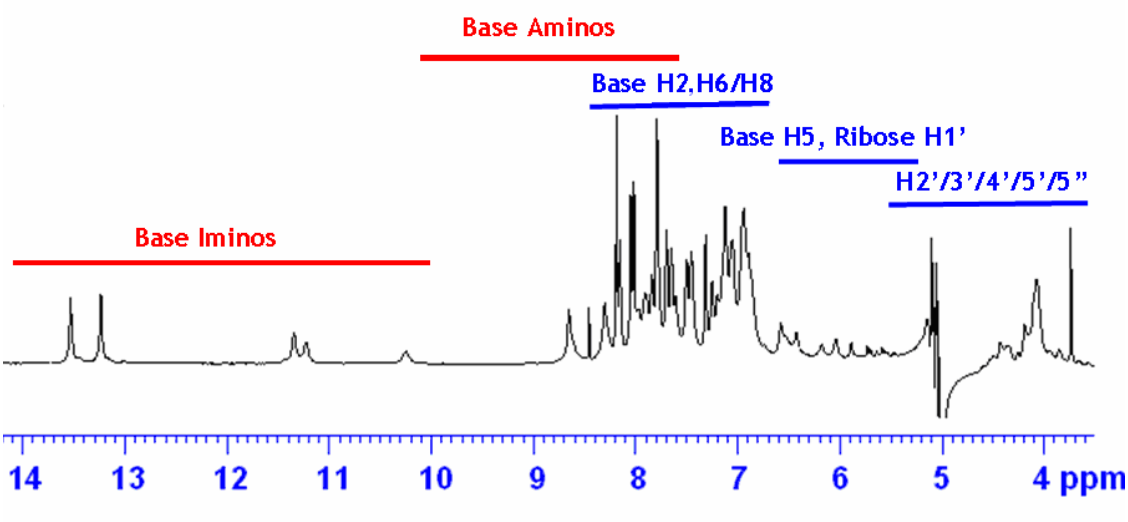


Figure 2.3. 1D-^{1}H NMR spectra of the 690UC mutant in H_2O at 25ºC. The distorted peak at 4.8ppm is the residual suppressed H_2O signal. The peaks between 10-14 ppm are imino proton peaks, and the hydrogen-bonded imino protons are located at the regions with higher chemical shifts than the non-hydrogen-bonded imino protons. Mostly, base amino protons are located between 7-9 ppm. Both imino and amino protons are exchangeable protons. The protons in the base include H2 in adenine, H6 of pyrimidine, and H8 of purine, which are usually located in the region of 6.5-8.5 ppm. And sugar proton H1' are regularly located in the region of 5-6 ppm, and other sugar protons are located in the region 3.5-4.5 ppm in most cases.

Structural studies of RNA using 1D-NMR methods are severely limited by the spectral overlap of the crowded sugar proton region. The sugar protons from H2' to H5'' are located between 4 and 5 ppm, which makes combining different spectra crucial. 2D and 3D NMR experiments are efficient methods to resolve peaks that are overlapped in 1D NMR. Besides improving peak resolution, the most important use of multi-dimensional NMR spectra is to obtain different types of structural restraints. Structural restraints can be classified as distance restraints, dihedral angle restraints and bond orientation restraints from RDCs (Prestegard 2000). The Nuclear Overhauser Enhancement Spectroscopy (NOESY) experiment is currently the most extensively used method for obtaining structure information (Neuhaus 2000). The NOE is a due to spin-spin cross-relaxation via through-space dipole-dipole interactions. A measurable NOE can be measured for two protons that are less than 5 Å. The "NOE-walk" region is used to obtain sequential assignments of the base-paired regions of the secondary structure (Wüthrich 1986). NOE peak volume are inversely proportional to the sixth power of the distance, therefore, distance information can be obtained by measuring NOESY cross-peak intensities when the distance between two atoms is less than 5.0 Å. NOE buildup experiments with short mixing times can be used to minimize the effects from spin diffusion, and measure distance between atoms in biological molecules more accurately (SantaLucia 1993). The structure calculation for large biomolecules requires a large number of NOE restraints. Non-observed NOEs can be added as restraints (i.e. distance > 4.5 Å) to avoid two hydrogen atoms from getting too close in the

calculated structures if no matching peaks can be observed in NOESY spectra (Diener 1998). NOESY can also be used to obtain the *chi* dihedral angels based on the base protons to H1' cross-peak strength. The *chi* dihedral angles are restrained to the *anti* conformation if the intra residue H1'-H6/H8 cross-peak is weak, and *chi* is restrained to *syn* if the intra residue cross-peak H1'-H6/H8 is strong.

Another class of structural restraints is dihedral angle restraints. J-coupling constants are affected by the dihedral angels of related chemical bonds based on appropriate Karplus equations (Karplus 1959). Therefore, J-coupling constants can be used to define the dihedral angle ranges of biomolecules. The experiments used in my research based on J-couplings are Heteronuclear Multiple Quantum Coherence (HMQC), DQF-COSY, TOCSY and HETCOR. HMQC is used to detect one-bond N-H or C-H cross-peaks. HMQC is used to confirm the assignments of NOESY. The ^{13}C chemical shift ranges in HMQC are much wider (~ 130 ppm) than the ^{1}H chemical shift ranges in NOESY in D_2O (~ 5 ppm). As a result of the heteronuclear chemical shift dispersion, the peak overlap in HMQC is much less intense than that in NOESY. Furthermore, the peaks in HMQC are grouped separately based on the different carbon containing functional groups (Varani 1991). The carbon chemical shift range from high to low frequency (i.e. 200 to 70 ppm) is: C2 of adenosine, C6/C8, C5 of uridine and cytosine, C1', C4', C2'/C3' and C5'. With the help of HMQC, the differentiation of C2 of adenosine from C6/C8 can be used to avoid the confusion of including H2 in "NOE-walk" (because the chemical shifts of H2 and H6 or H8 are similar). In

addition, although NOESY can divide purines from pyrimidines based on correlation between H5 and H6 aromatic resonances, but NOESY can not differentiate H5-H6 peaks of uridine from cytosine. In HMQC, the chemical shifts of C5 of uridine are down-field compared with C5 of cytosine, which aids the sequential assignments of NOESY (Varani 1991). Double Quantum Filtered Correlation spectroscopy (DQF-COSY), Total Correlation spectroscopy (TOCSY) and HETCOR provide information about through-bond correlations. The coherence transfer in DQF-COSY is restricted to directly spin-coupled protons that are 2-4 bonds apart and results in cross-peaks that are anti-phase. DQF-COSY is useful for determining basic connectivity via J-couplings (through bonds) (Varani 1996), and DQF-COSY has been used to obtain the sugar pucker conformations in my experiments. TOCSY transfers magnetization by the Hartman-Hahn mechanism (Davis 1985), which is different than DQF-COSY. As a result of Hartman-Hahn mechanism, TOCSY is able to generate cross-peaks via relaying of coherence among all the spins in a coupled system during the spin-locked mixing period, so that magnetization transfer can proceed from one spin to another even without direct coupling. The cross-peaks in TOCSY spectra are in-phase (Bax 1985). The ^{31}P-^{1}H HETCOR experiment can be used to quantify the β and ε dihedral angles and to limit the range of the α and ζ dihedrals using the ^{31}P chemical shift (Gorenstein 1984). The used experiments and their uses are summarized in Table 2.1.

Table 2.1. Routine NMR experiments used in our lab for unlabeled RNA samples.

Experiment	Use
1D-^{1}H (in H_2O)	Assignment of imino peaks Experimental optimization (e.g. temperature, pH, *etc*)
1D-^{1}H (in D_2O)	Experimental optimization (e.g. temperature, pH, *etc*)
NOESY (in H_2O)	Assignments of imino or amino protons Imino sequential walk
NOESY (in D_2O)	Sequential walk to confirm secondary structure Assignments of non-exchangeable protons
HMQC (in D_2O)	Differentiate H6/H8 from H2 Differentiate uridine H5 from cytosine H5 Confirmation assignments of non-exchangeable protons in NOESY
COSY (in D_2O)	Confirmation of proton assignments Sugar pucker conformations
HETCOR (in D_2O)	Dihedral Angle Restraints
TOCSY-NOESY (in D_2O)	Confirmation of proton assignments, especially sugar protons Dihedral Angle Restraints

By applying a third-dimension, 3D NMR experiments often resolve cross-peaks that are overlapped in 2D NMR based on chemical shift editing in the third-dimension. Therefore, 3D NMR experiments are very valuable for studying large nucleic acids (Wijmenga 1994). One distinct advantage of 3D TOCSY-NOESY is that the experiment doesn't require isotope labeled samples. Many 3D experiments, like 3D HSQC-NOESY, decrease resonance overlap by applying the more resolved carbon or nitrogen regions as the third dimension, but these experiments need isotope enriched samples. For RNA studies, one of most important roles of 3D TOCSY-NOESY is the ability to resolve the crowded sugar peaks between 4 and 5ppm, which can allow assignments of H5' and H5'' (Wijmenga 1994). 3D TOCSY-NOESY can also be used to confirm resonance assignments by combining NOESY and TOCSY cross-peaks. Figure 2.4 shows one of f1f3 planes of a 3D TOCSY-NOESY spectrum, and the NOE peaks between sugar protons and H6/8 protons are spread in the f1 dimension due to coherence transfer among the J-coupled sugar protons (Wijmenga 1994).

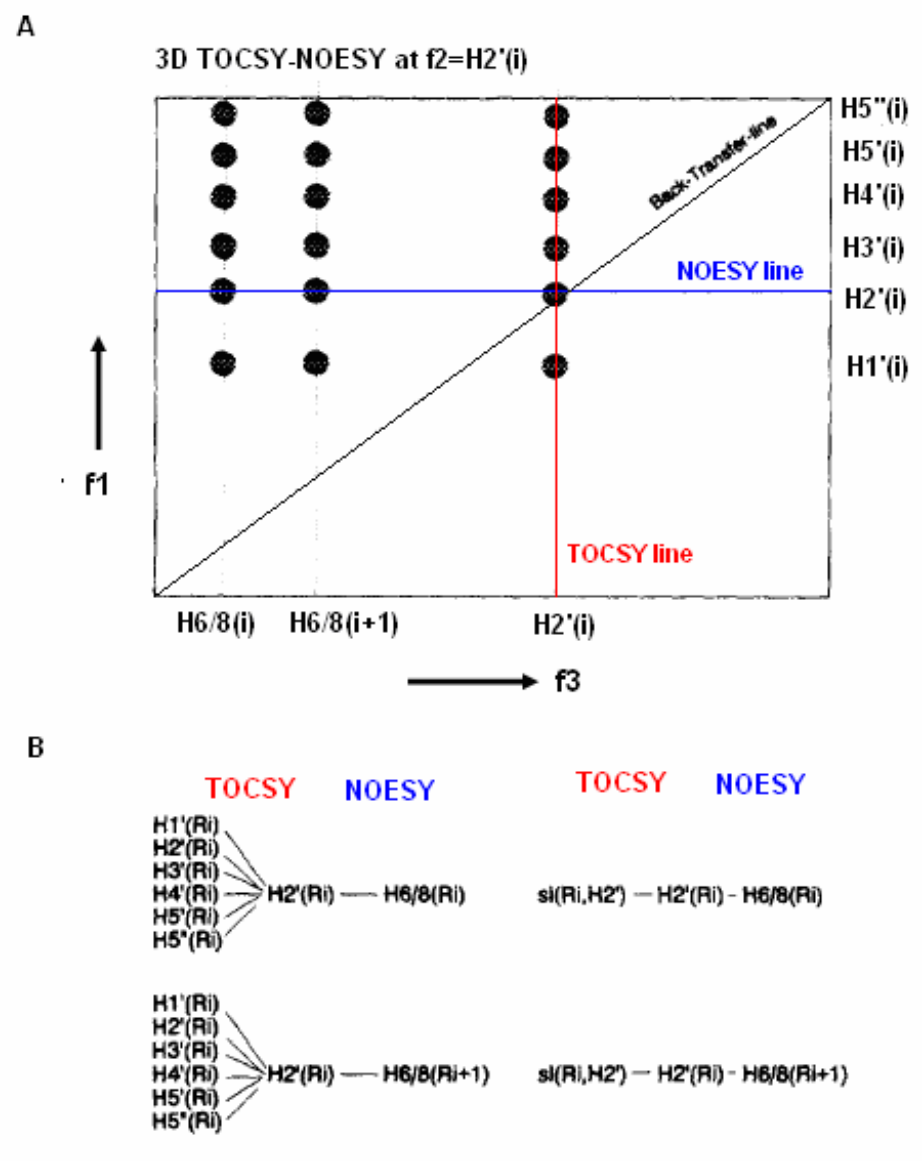


Figure 2.4. (A) One f1f3 plane of a TOCSY-NOESY spectrum at f2=H2'(i). The NOE peaks H6/8(i+1)— H2'(i), H6/8(i) — H2'(i) and H2' (i) —H2 (i) are spread out into a spin-lock pattern during to the presence of coherence transfer in TOCSY mixing time. (B) Magnetization transfer routes in TOCSY-NOESY (Wijmenga 1994).

Another valuable application of the 3D TOCSY-NOESY is to determine the sugar pucker and the γ backbone torsion. Intra-residue H1'-H2' peaks of most residues in RNA are usually very weak or not observable due to the small J-coupling between H1' and H2' in N-type sugar puckers. If strong cross-peaks appear at the H2' position in the H1' spin-lock lines, the sugar pucker conformation of that residue is predominantly S-type puckered (i.e. C2'-endo). The conformation of S-type sugar puckers in RNA is also indicated by the relatively weak inter-residue H2' (i) — H6/H8 (i+1) cross-peak along with a strong intra-residue H2' (i) — H6/8 (i). The γ torsion angles can be divided into the gauche (+) domain (γ^+), the trans domain (γ^t) and gauche (-) domain γ^-. TOCSY coherence transfer intensities among sugar protons can be used to determine γ torsion angles. For example, under the condition that the TOCSY mixing time is about 78 ms, for N-type sugars, coherence transfer to H5'/H5'' on the H2' and H3' spin-lock lines are very weak or not observable, which indicates the sugars are predominantly in γ^+, and the sugars are considered as γ^t or γ^- if intensive peaks at H5'/H5'' position can be observed on the H3' spin-lock lines (Wijmenga 1994). Analysis of the NOE intensities of H3' to H5'/H5'' and H4' to H5'/H5'' can be used to simultaneously determine stereo specific H5' and H5'' assignments and also to determine γ (Wijmenga 1994).

The structural restraints were input into CNS 1.1 software (Brünger 1998) and a family of global-folded structures with minimum energy was calculated by using simulated annealing and molecular dynamic methods to calculate a family of structures with minimum energy. The detailed protocol for structure

calculations is discussed in Chapter 3. The precision of output structures are correlated closely with the quantity and precision of the structural restraints. Typically, dihedral angle constraints or planarity restraints for forming A-form base-pairs derived from known structures are usually introduced in the calculations. All the structural restraints (distance restraints and dihedral angle restraints) derived from NMR spectra are used in restrained molecular dynamics/simulated annealing (MD/SA) structure calculation. MD is a force-field based simulation method that has been used to describe the physical properties of protein/ nucleic acid systems by using an analytical potential energy function and NMR derived restraints (Hashem 2009). The quality of output structures from the MD simulations are often correlated with the precision with the initial conditions. The initial conditions include the starting structures and inter-residue interactions that are taken into account by the force-field used in MD calculations. The starting structures are mainly derived from model-built systems with randomized geometry from X-ray or NMR data. The inter-residue interactions are derived from NMR data in my research. The result of an NMR structure calculation is a family of structures, which are consistent with the NMR data. In the output structures, the respective positions of the bases are usually better determined than the positions of the sugar-phosphate backbone because fewer structural restraints per atom in sugar-phosphate backbone are obtained from NMR spectra and therefore, more degrees of freedom for the sugar-phosphate backbone positions. The calculated structures with bad contacts, such as overlapped atoms, or distorted non-planar bases, are usually discarded. If

sufficient NMR restraints are provided, then the calculated structures in the family will converge to the same or very close fold. The agreement among the structures in the family is quantified by root-mean-square-deviation (RMSD).

2.3.2. NMR for Interaction Site Mapping

Molecular recognition plays an important role in biological function. NMR structural studies can provide insights into the factors involved in the recognition of RNA by proteins and small molecule ligandds. NMR methods are sensitive to molecular interactions due to the fact that changes in the chemical environment resulting from molecular interactions cause changes in the chemical shifts of nuclei and change the dynamics of the target, particularly in proximity to the ligand binding site. Many regions of RNA contain a significant degree of conformational flexibility, and NMR methods are very useful to provide structural information about conformational dynamics and mobility of biomolecules in solution, conformational or chemical exchange processes of biomolecules. Chemical shift mapping can be used to localize binding sites with ligand and to determine K_D (dissociation constants) and shed light on the mechanism for chemical reactions. NMR can also detect conformational changes of nucleic acid or protein residues that are involved in the formation of protein/RNA complexes (Williamson 2000; Leulliot 2001). This makes NMR a tool for functional screen of drugs (Williamson 2000; Leulliot 2001). Additionally, dynamics as well as the binding affinity of RNA/Protein complexes can also be extracted from NMR spectra. NMR titrations can also be used to determine the equilibrium binding

constants for bio-molecular complexes (Clos 2007). An advantage of using NMR method to determine the interaction geometry of nucleic acids with proteins and small molecules into the to mM dissociation constant range, which is not well covered by other biochemical binding methods (Clos 2007).

For NMR studies of RNA-ligand complexes, distinguishing the bound RNA-ligand complex ([RL]) from the free ligand ([L]) or free RNA ([R]) is a prerequisite for quantifying the amounts of the [L] or [R] and [RL]. Usually, chemical shift or linewidth changes of RNA can be used to monitor the interaction between the ligands and bio-molecules when ligands are titrated into the solution (Clos 2007). The complexes of protein/nucleic acid-ligand are dynamic, continuously binding and unbinding and the exchange time regime between [R] or [L] and [RL] are crucial for NMR determination. If the exchange rate is slow on the chemical shift time scale, resolved signals may be observed for the free and bound states, respectively. The concentrations of free and bound states are detectable by integration of separate resolved peaks. While in the other case, if the exchange rate is fast compared with the chemical shift time scale, then only one peak may be measured at the population weighted average frequency corresponding to the ratios of the two conformations (Clos 2007). The observed chemical shift corresponds to population weighted average of the free and bound states: $Mobs=X_{[L]}M_{[L]} + X_{[RL]}M_{[RL]}$. The $X_{[L]}$ and $X_{[RL]}$ in the equation is the mole fractions of free and bound ligand, $M_{[L]}$ and $M_{[RL]}$ are the NMR parameters for free and bound states, respectively. The NMR parameters (i.e. chemical shift) are chemical shifts in ppm or Hz, a linewidth, relaxation rates, translational diffusion

coefficients, or other NMR features. The frequency differences in Hz between the two peaks in fast exchange are proportional to the magnetic field strength B_o (i.e. $v=\gamma B_o$). Large line-widths indicate intermediate exchange. Therefore, the exchange rates and populations of the conformations between free proteins or nucleic acids and the bound molecules can be measured based on the peak numbers and line-widths. Usually, the measurable range of K_D that can be measured by NMR is around 10-100 μM and NMR can not be used to directly measure K_D smaller than the limit detection.

Part of this research to map the interaction sites of the 970 hairpin of *E. coli* ribosomal RNA after mixing with 7mer-peptide by NMR methods is in Chapter 4. Such study will be used to identify the functional sites of 970 hairpin that are accessible for small ligands and to provide useful structural and functional information of 970 hairpin for developing new classes of anti-bacterials. The 7mer-peptide is one of the ligands has been found in Dr. Chow and Dr. Cunningham's labs by phage display method and synthesized by solid-phase synthesis method that may bind with the 970 hairpin of *E. coli* ribosomal RNA.

2.3.3. Recent Progress in NMR methods

NMR methods have steadily improved to allow the determination of the structure of even larger biomolecules. This section describe some of the most important recent advances in the field.

2.3.3.1. NMR Methods for Isotope-labeled Samples

One of the main problems for NMR methods in determining the 3D structures of large-sized bio-molecules is chemical shift overlap. With the use of multi-dimensional NMR experiments, the overlapped ^{1}H peaks in the NMR experiments may be able to be separated in multiple dimensional NMR experiments based on the different chemical shifts of ^{13}C and ^{15}N. However, the isotopic natural abundance for ^{13}C and ^{15}N are only 1.1% and 0.37%, respectively. The natural abundance ^{13}C and ^{15}N isotopes are not suitable for multiple through bond magnetization transfers that are required for multi-dimensional NMR experiments. The efficient way to solve this problem is to prepare NMR samples that are enriched with NMR sensitive isotopes (Nikonowicz 1992; Hines 1994; Batey 1995). The common enriched isotopes in nucleic acids are ^{13}C, ^{15}N and ^{2}H. Isotope labeling with ^{17}O is not suitable for high-resolution NMR research since its nuclear spin is -5/2, which results in broad linewidths due to quadrupole relaxation. The related information of the common atoms in nucleic acids is shown in Table 2.2.

Table 2.2. The NMR property parameters for common nuclei in RNA.

Isotope	natural abundance	gyromagnetic ratio (10^7 $T^{-1}s^{-1}$)	Spin number
^{1}H	99.98%	26.75	½
^{2}H	0.02%	4.12	1
^{12}C	98.9%	NMR inactive	NMR inactive
^{13}C	1.1%	6.73	½
^{14}N	99.63%	1.93	1
^{15}N	0.37%	-2.71	½
^{16}O	99.9%	NMR inactive	NMR inactive
^{17}O	0.04%	-3.63	-5/2
^{31}P	100%	10.83	1/2

The NMR samples can be either selectively or uniformaly isotope labeled (Nikonowicz 1992; Hines 1994; Batey 1995). Selective isotope labeling of nucleic acids can reduce the complexity of some spectra and is useful for peak assignments, especially for some ambiguous assignments, which makes NMR potential to study larger bio-molecules that are isotope labeled (Batey 1992; Hines 1994; SantaLucia 1995). Another benefit of using isotope labeled samples is that the data collection for hetero-nuclear multidimensional NMR experiments, such as HMQC, HSQC, 3D HMQC-NOESY, *etc.*, can be completed in much less time with similar sensitivity as the NMR experiments for the un-labeled samples. What's more, partially selective isotope labeling facilitates the study of dynamics in NMR.

Although using ^{13}C and ^{15}N isotope labeled samples can improve peak separation in some multi-dimensional NMR experiments, the application of NMR is limited by magnetization transfer efficiency. For protonated bio-molecules, the desired magnetization transfers are reduced by homo-nuclear dipolar spin relaxation (^{1}H-^{1}H), and hetero-nuclear dipolar spin relaxation (^{1}H-^{13}C or ^{1}H-^{15}N). As a result, the sensitivity of hetero-nuclear multidimensional NMR experiments are decreased. By partially replacing some ^{1}H to ^{2}D, many magnetization transfer relaxation pathways are eliminated, which improve efficiency of the desired magnetization transfer pathways, sensitivity and resolution of peaks due to narrower linewidth. Also for NOESY, partially ^{2}D isotope labeled samples improve the accuracy of atomic distances that are calculated on the cross-peak intensities because the spin-diffusion pathways decrease (LeMaster 1988;

Torchia 1988; Grzesiek S 1993; Yamazaki 1994). Some NMR structures of large bio-molecules have been obtained by using selectively isotope labeled ^{2}D samples. Yu *et al.* used the deuterium substitution method to assign the ^{1}HN-^{1}HN peaks in the NOESY spectra of a 29 kDa methyl transferase (Yu 1997).

2.3.3.2. Transverse Relaxation Optimized Spectroscopy (TROSY)

Molecules in solution are surrounded by solvent molecules and undergo random “Brownian motion” caused by constant impact from surroundings. The larger the molecules are, the more slowly they respond to the impact and they reorient themselves in solution, and the slow tumbling deteriorates NMR spectra. The line widths in the NMR spectra are inversely proportional to the relaxation rates ($\Delta\gamma_{1/2}=1/(\pi T_2)$, $\Delta\gamma_{1/2}$: the line width at half-height, T_2: transverse relaxation rate). Therefore, fast relaxation will cause fast decay of signal and poor signal-to-ratio (s/n). Due to fast relaxation, traditional NMR methods have difficulty solving linewidth and the structures of large-sized molecules, such as proteins, nucleic acids, or their complexes with molecular weights higher than 50 KDa (Da is equivalent to the mass of one hydrogen atom). Two approaches have been used to overcome the poor signal-to-noise ratio problem: ^{2}D-labeling to eliminate ^{1}H mediated relaxation pathways and TROSY. TROSY was first developed by Kurt Wüthrich (Pervushin 1997). For two coupled I=1/2 atoms, there are four energy transitions, and routine NMR methods use decoupling which effectively detects the average all four transitions. However, such decoupling deteriorates the signal lineshape when the molecules are large-sized. TROSY belongs to single-

transition spectroscopy, and only one of the four transitions that are insensitive to “Brownian motion” are detected in TROSY, which can suppress the transverse nuclear spin relaxation and make measurement large-sized molecules possible for NMR (Pervushin 1997; Fiaux 2002). With TROSY, the size limit of NMR structures is expanded to much large sizes, and NMR method can be used to determine the proteins with molecular weights higher than 150 KDa with the help of TROSY. The distinctive improvement of TROSY used in large proteins is shown in Figure 2.5. TROSY is inefficient for proteins larger than 200 KDa due to very fast transverse relaxation rates of the proteins, while Cross Relaxation Enhanced Polarization Transfer (CRINEPT) becomes more effective method to extract structural information for very large bio-molecules (Riek 1999).

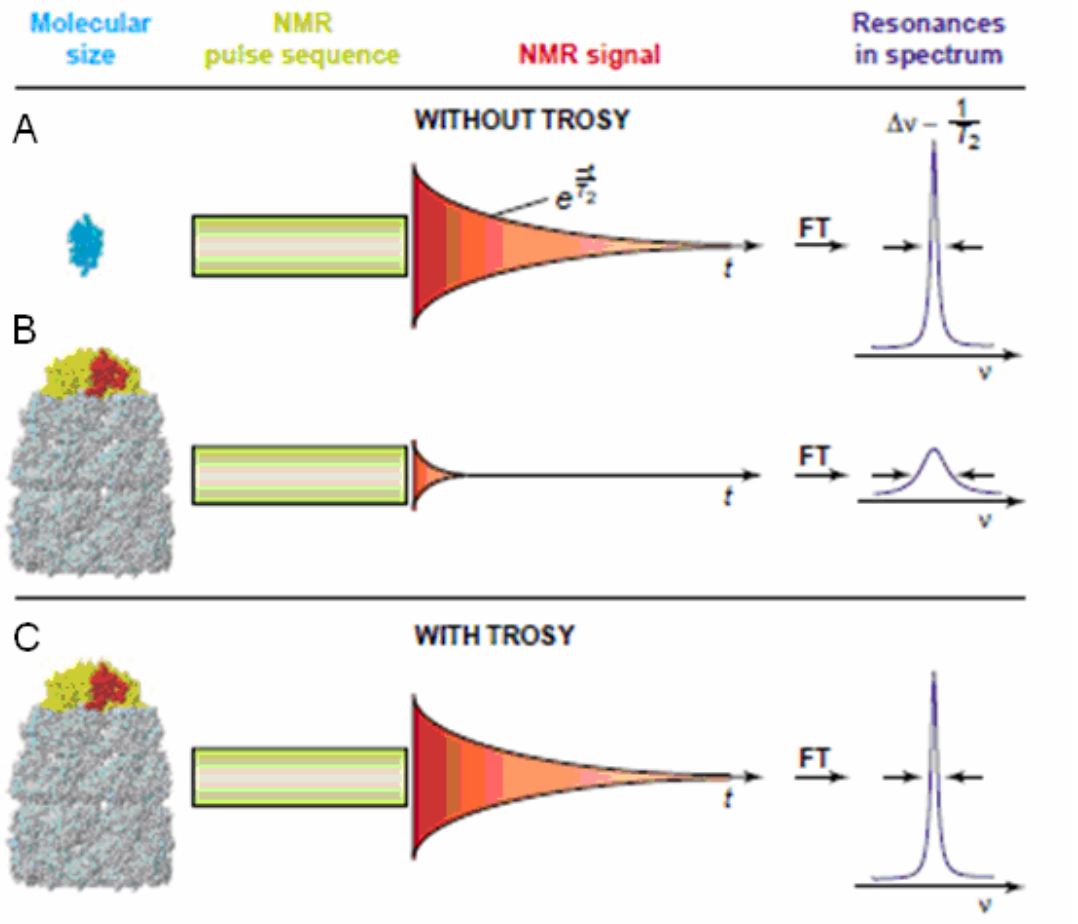


Figure 2.5. The NMR signal obtained from small molecules (A); The NMR signal obtained from large proteins (B); Using TROSY, much narrower peak obtained, which results improved spectral resolution and sensitivity for large molecules (C) (Fernández 2003).

2.3.3.3. Residual Dipolar Coupling (RDC)

Chemical shifts and J coupling constants give local structure information. J couplings provide information for atoms separated by 1-4 (or more in special circumstances) chemical bonds, and NOE data can relate atoms that may be far apart by many chemical bonds, but still, are close in 3D space. NOESY is very useful for obtaining distance restraints for the atoms that are closer than 5.0 Å. While the short-range nature of the NOESY is limited, it is difficult to extract information of long-range distance constraints or the relative spatial orientation of the molecule. In addition, the measurement of insufficient distances and cumulative errors in local distances can result in ill-defined relative positions of distant regions of bio-molecules. A weakness of structural restraints obtained traditional NMR methods is their short-range, for example, the distance restraints from NOESY are limited to be shorter than 5.0 Å, and dihedral angle restraints can only be used to restrict atoms that are apart by three chemical bonds. Uncertainties of such short-range restraints may lead to poor overall structure of the whole NMR structure even though individual local regions are well-defined.

Varani's lab firstly applied RDCs into refining RNA structures (Beyer 1999). In their research about determining the global structure of the U1A protein-RNA complex, the poorly defined global structure of the complex were improved by adding more than one hundred RDCs, which helped to determine the orientations of the two double-helical regions in the complex (Beyer 1999). RDCs (Residual Dipolar Couplings) can be used to provide the orientations of individual bond vectors with regard to the axes of the alignment tensor of the molecules and

contain global structural information. Such bond vector information from RDCs can provide the relative orientation of two domains of a biomolecule. RDCs in solid state are in the range of KHz, but in liquid states, dipolar couplings vary with the changes of distance and orientation of two spins due to molecular tumblings, however, the average of dipolar couplings is zero due to fast tumblings of samples in solution, which makes splittings in resonance lines unobservable. However, if NMR samples are prepared in slightly anisotropic solutions by adding some reagents which form bicelles, cellulose crystallites or filamentous phages, the dipolar coupling averages are no longer zero because bicelles, cellulose crystallites, or phages weakly align in the static magnetic, thereby providing weak dipolar couplings (Tjandra 1997). If only a fraction of one out of one thousand is aligned, the dipolar couplings are around a few Hz. The use of dilute liquid crystals creates anisotropy based on preferred alignments of a small fraction of the molecules and the produced residual dipolar couplings are sufficient to be detected in NMR. RDCs can be used to provide bond vector information in anisotropic solution if special crystalline media is used during the measurement. The distance and the angle dependence of dipolar couplings can be used to obtain orientational restraints based on the equation: $D=\text{const}/r^3\times\langle(3\cos^2\theta-1)/2\rangle$ (r: the distance between the two spins which are dipolar coupled; θ: the angle of the distance vector with respect to the principle axis of the alignment tensor). If we are considering a RDC between one-bond couplings, such as N-H, or C-H, the inter-atomic distances among different N-H or C-H bonds are very similar and the RDC depends only on the orientational components of different

bonds. This would also be correct for measuring two-bond RDCs, but three-bond RDCs may be affected by dihedral angles. RDCs provide crucial long range interactions, which are complementary to the short distance restraints obtained from NOEs. But two opposite orientations give the same RDC value, that's to say, RDC can not be used to differentiate the direction of an inter-nuclear vector from its inverse direction. This makes calculate a *de novo* structure based on residual dipolar couplings without NOEs very difficult and unreliable, and thus, RDCs have been used for refining structure models.

2.3.3.4. Small Angle X-ray Scattering (SAXS)

Conventional NMR and X-ray crystallography experiments have been used widely to obtain the atomic-resolution structures, and many structures have been solved successfully by these two methods. However, NMR can have size limitations and can lack global precision if insufficient RDCs are available. In addition, X-ray crystallography is limited by the need for diffracting crystals. Small-angle X-ray scattering (SAXS) or small-angle neutron scattering (SANS) have been used recently to avoid the ambiguous structures solved by sole NMR or X-ray crystallography methods. The two methods provide structural information about the overall shape and dimensionality of biomolecules, and has been utilized to refine protein and RNA structures in combination with NMR or X-ray structural restraints (Zuo 2010) (Zuo 2008). The SAXS data are particularly useful for nucleic acids because the phosphate-sugar in backbones of nucleic acids contain large electron densities, which dominates the intensity of SAXS

data and therefore, SAXS data reflects the shapes of overall structures of nucleic acids. The information contained in SAXS and RDC is complementary to the routine NMR experiments, and RMSDs of many protein structures have been improved and the global structures are more close to the actual experimental data after combining SAXS and NMR methods (Wang 2009).

The different types of structural information obtained from NMR are listed in Table 2.3.

Table 2.3. Structural information extracted from NMR spectra.

Observable	Information
Chemical shifts (^{1}H, ^{13}C, ^{15}N, ^{31}P)	**Assignments**
J-coupling (^{1}H-^{13}C, ^{1}H-^{1}H)	**Dihedral angles**
NOE	**Local inter-atomic distances**
TROSY	**Decrease over-lap for large molecules**
RDC	**Global bond vector orientation**
SAXS	**Global shapes**

2.4. Homology Modeling

Besides experimental methods, homology modeling methods can also be used to obtain 3D structures of bio-molecules. Homology modeling method are used to build up models of unknown structures based on a known structure obtained from experimental data based on the fact that evolutionarily related oligonucleotides or proteins share similar structures (Wallner 2005). The efficient routine sequencing of entire genomes creates an overwhelming flood of sequence data. Compared with the number of new nucleotide sequences that appear every day, the determination rate for protein or RNA, or DNA by NMR and X-ray crystallography methods are much slower and more expensive. Homology modeling method is an effective technique that is used to narrow the gap between the known sequences and the 3D structures. Homology modeling is widely used for obtaining 3D coordinates for proteins, DNA, or RNA because the time spent on homology modeling is much less than NMR and X-ray experiments, although the quality of structural models constructed by homology modeling is not clear. The regularly used steps in homology modeling method are: first, choosing the known structure as the template for modeling the target structure; second, aligning the query sequence (whose structure is unkown) to the template sequence; finally, building a 3D model of the target by using homology modeling method to account for substitutions, insertions, and deletions (Wallner 2005). The modeling programs that used for homology modeling can be divided into three types: rigid-body assembly, segment-matching, or modeling by satisfaction of spatial restraints (Wallner 2005). The rigid-body assembly programs, such as

SWISS-MODEL (Schwede 2004), NEST (Petrey 2003), 3D-JIGSAW (Bates 2001) and Builder (Koehl 1994), fit the rigid bodies (the rigid bodies are from the core of aligned regions) onto the frame-work and rebuilding the non-conserved parts. SWISS-MODEL method shows advantage in stereochemistry, but it causes many crashes and convergence. The NEST method is more reliable than SWISS-MODEL in many examples, but models from NEST often contain bad stereochemistry. The models built by 3D-JIGSAW usually have several problems, such as missing residues, bad side chains and bad stereochemistry (Wallner 2005). The segment-matching programs, including SegMod/ENCAD, use a subset of atomic positions as a guide to find matching segments in all known structures (Levitt 1992). MODELLER belongs to the class of modeling programs that use restraints extracted from the alignment, and then, the model is obtained by minimizing the violations of the restraints, which is a program based on "modeling by satisfaction of spatial restraints" (Šali 1993). SegMod/ENCAD, MODELLER, and NEST have better overview results than other methods (Wallner 2005), while SegMod/ENCAD method often produces poor backbone conformation, NEST build models shows weakness at chemistry. For non-optimal alignments, MODELLER has an advantage is that the modeled structures are almost not affected in some cases that the target structures are possibly distorted when poor alignment between a template and a query may cause gaps between two adjacent residues because MODELLER is based on spatial restraints, and the gap just adds a few additional spatial restraints to the final optimization procedure (Wallner 2005).

Homology modeling can provide some useful structural information, but in many cases, homology modeling miss coordinates, fail to converge (RMSD > 3Å between the model and the simple backbone model), and some programs crash for some structures, and some programs can not model some regions, especially for some loops, and the possible reasons for the failures are either too long loops or not able to find the right loop in the loop library. All the listed problems make homology modeling not that reliable (Wallner 2005). Therefore, the main use of homology modeling method is that it can provide possible structures that are suggestive and useful for experiments although some structures obtained from homology modeling method are not reliable.

CHAPTER 3

NMR Study for the Two Functional Mutants of 690 Hairpin of *E. coli* 16S rRNA

3.1. Introduction of the 690 Loop

Recent high resolution X-ray crystal structures of 70S ribosomes and seperate subunits show that ribosomal RNA (rRNA) forms the central core of the ribosome (Cate 1999; Wimberly 2000; Yusupov 2001; Selmer 2006). The eight-nucleotide 690 hairpin (nucleotides 690~697 in *E. coli*) terminates the helix 23b and is located at the surface of the platform region of the 30S subunit, and the nucleotides in the hairpin (from 690 to 697) are highly conserved in all three domains of life (Van de Peer 1999). The 690 loop has many functional roles, such as interaction with P-site bound tRNA, binding to protein S11, and participating in ribosomal subunit association as part of bridge B7a (Powers and Noller 1995; Mueller 1997; Agalarov 2000). In addition, nucleotides 693-697 interact with IF3, as evidenced by protection from cleavage by hydroxyl radicals (Dallas 2001; Nguyenle 2006). The 690 loop has also been found to be one of the primary binding sites for several antibiotics including pactamycin and edeine, which both protect nucleotide G693 from chemical modification (Woodcock 1991; Mankin 1997).

Conservation of sequence in RNA may be due to direct interaction of the conserved nucleotides with functionally important ligands or to the necessity to maintain the three-dimensional positions of functionally important nucleotides. Nucleotides in the 690 hairpin were subjected to saturation mutagenesis via the

instant evolution method and 101 mutants that retained the ability to synthesize protein *in vivo* were selected, sequenced and compared to identify the sequence and structural motifs required for ribosome function (Morosyuk 2000). The study of mutants has been used to identify the functions related to the 690 hairpin. For example, the mutation of A694G confers resistance to pactamycin (Mankin 1997). Morosyuk *et al.* determined the NMR structures of wild-type of 690 hairpin, and the NMR structure revealed H-bonding and base stacking stabilize the mismatches formed at G690•U697 and G691•A696 (Morosyuk 2001). Morosyuk *et al.* also studied the mutants of the 690 hairpin by site-directed mutagenesis methods. Morosyuk *et al.* homology modeled the structures of functional mutations at positions 690 and 697 and proposed that the functional mutants and wild-type form isomorphous structures (Morosyuk, Cunningham et al. 2001). Though homology modeling of mutant structures is certainly useful and suggestive of experiments, this method is still not as reliable as experimental methods for structure determination. Thus, to test the hypothesis that mutants form isomorphous structures with the wild-type sequence, we present NMR structures of two of the functional mutants and compare the functional mutant structures to the available NMR structure of the wild-type.

The two mutants we chose for our structural studies differ significantly from the wild-type sequence but retain high protein synthesis activity *in vivo*. We focused on sequences with mutations at the positions 690 and 697 since a functionally important interaction between the residues at these two positions was identified by covariation analysis of the 690 hairpin mutation library

(Morosyuk 2000). The following functional double-mutants match these requirements: G690A•U697A (function: 74.1 %), 690A•697U (function: 89.7 %), G690C•U697A (function: 66.8 %), 690C•697U (function: 79.6 %), 690G•697A (function: 68.3 %), G690U•U697A (function: 61.8 %), 690U•697U (function: 66.1 %) and G690U•U697C (function: 70.1 %) (Morosyuk 2000). The mutants at positions 690 and 697 with their functions are shown in Figure 3.1.

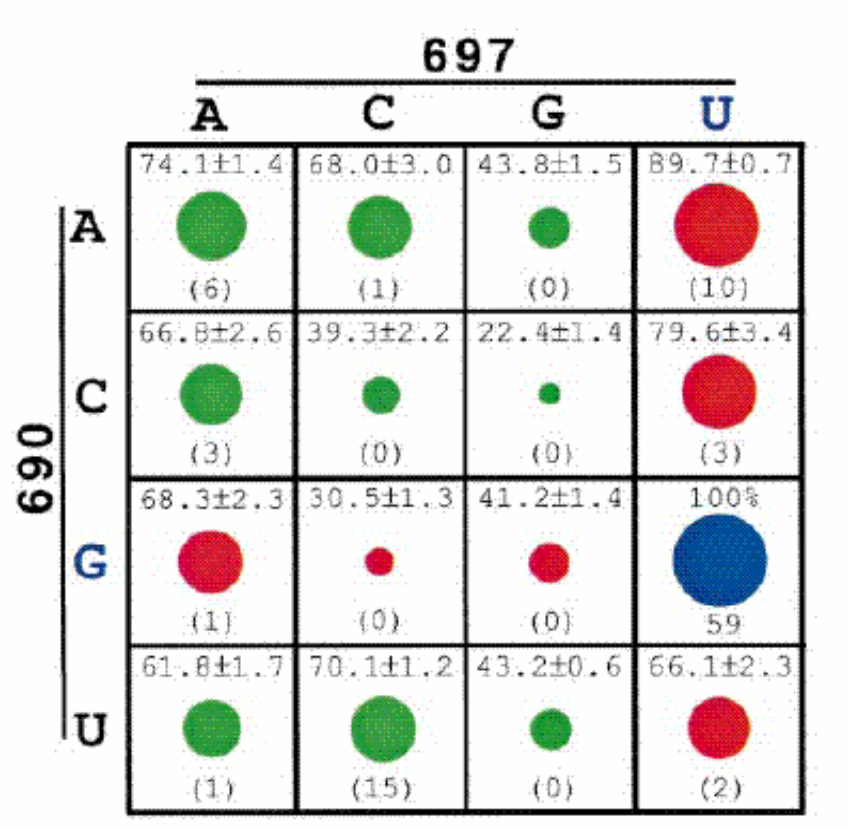


Figure 3.1. The mutants at the position 690 and 697 and the wild-type of the 690 hairpin. Double mutants are in green, single mutants are in red, and the wild-type (G690•U697) is blue. The sizes of the dots are proportional to the level of functions of the mutants and the wild-type. The functions of the mutants and the wild-type are located at the top of each dot and the number of occurences of each pair of mutations shown in the instant evolution study are shown on the bottom of the dots. The functions and the number of occurences for the UC mutant are 70.1 ± 1.2% and 15, respectively (Morosyuk 2001).

In order to study a functional mutant at 690 and 697 that is as different as possible from the wild-type, the four double mutants: 690A•697A, 690C•697A, 690U•697A and 690U•697C match these requirements. We chose the G690U, U697C double mutant sequence (the UC mutant) as one sample in our study since it is the most commonly occurring mutant in the instant evolution study (Morosyuk 2001). The second mutant chosen for the NMR studies is the quadruple mutant (G690A, G693C, A695C, U697A, the QM mutant). Several characteristics of the QM mutant are attractive for structural studies: it has four mutations yet still retains high function (94.6%). Interestingly, the function of the QM mutant is higher than the G690A, U697A double mutant (74.1% function) evaluated by site-directed mutagenesis. Lastly, the nucleotide identities at positions 690 and 697 are different from those in the UC mutant, which allowed us to explore more of the functional sequence space of the 690 loop.

The sequences used for NMR studies of the wild-type and the two chosen mutants are shown in Figure 3.2. The terminal base pair (G687-C700) was added to each NMR construct to stabilize hairpin formation.

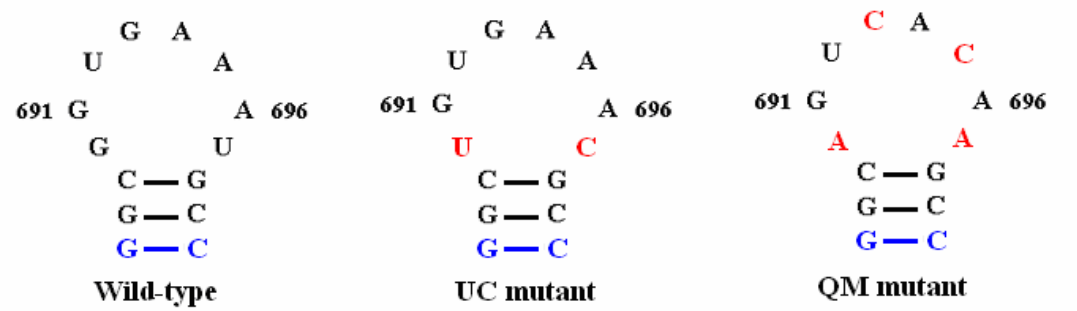


Figure 3.2. The secondary structures of the wild-type, the UC mutant and the QM mutant. The mutations in the UC mutant and the QM mutant are shown in red. The mutations are located at position 690 and 697 in the UC mutant, and at position 690, 693, 695, and 697 in the QM mutant. The terminal base-pair on the wild-type, the UC mutant and the QM mutant were changed to stabilize NMR constructs (in blue).

3.2. Sample Preparation and NMR Data Collection

RNA samples of the UC and QM mutants were ordered from Dharmacon, Inc. The preparation procedures of the two mutants are similar. The 2'-protecting groups of these samples were removed using the buffers provided by the manufacturer (100 mM acetic acid buffer at pH 3.8). The deprotected samples were purified by urea denaturing 20% PAGE, the slowest band cut out, and the RNA was recovered by electroelution (Schleicher & Schuell, Inc.). Next, the RNA was dialyzed against a salt buffer (50 mM NaCl and 0.2 mM EDTA), then dialyzed against double distilled deionized water, and then further purified by G-10 Sephadex column. The sample was dried by centrifugal evaporation. Finally, the dried RNA was dissolved in NMR buffer (50 mM NaCl, 0.2 mM EDTA, 10 mM disodium phosphate at pH 6.8). 90% H_2O and 10% D_2O were used as the solvent for exchangeable proton NMR spectroscopy. The final volume of the NMR sample was 0.3 mL in a Shighemi NMR tube. Samples for non-exchangeable proton spectroscopy were prepared in the same NMR buffer in 99.96% D_2O (Cambridge Isotope Labs). TSP (3-trimethylsilyl propionate) is used as the internal chemical shift reference in NMR experiments, whose resonance is set to 0.0 ppm.

The NMR experiments were carried out on unlabelled 14-mer constructs of the UC and the QM mutants (Figure 3.3). The 1D-NMR, NOESY, DQF-COSY, HMQC and 3D TOCSY-NOESY experiments were acquired using a BRUKER AVANCE 700 MHz spectrometer equipped with ^{1}H, ^{13}C, and ^{15}N triple resonance TXI cryoprobe with z-axis gradient coil and the spectra were processed by XWIN-

NMR 3.1 software. The ^{31}P-^{1}H HETCOR experiments were run on the same spectrometer except equipped with a ^{1}H, ^{13}C, ^{15}N, and ^{31}P quadruple resonance QXI probe. 1D-H_2O NMR and 2D H_2O-NOESY spectra provide information on the exchangeable imino and amino protons. All other NMR experiments were collected using the samples in NMR buffer containing 99.96% D_2O. ^{13}C-^{1}H HMQC experiments were collected with carbon carrier frequencies set at either 95 ppm (for C1' and C5) or 140 ppm (for C8, C6, and C2). All 2D and 3D NMR spectra were transferred to .ucsf files and analyzed by using SPARKY software (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco). The structure calculations were done using CNS software with the use of the extracted structural restraints (Cech 2009). The procedure of using NMR method to obtain the structures of the UC mutant and the QM mutant is shown in Figure 3.3.

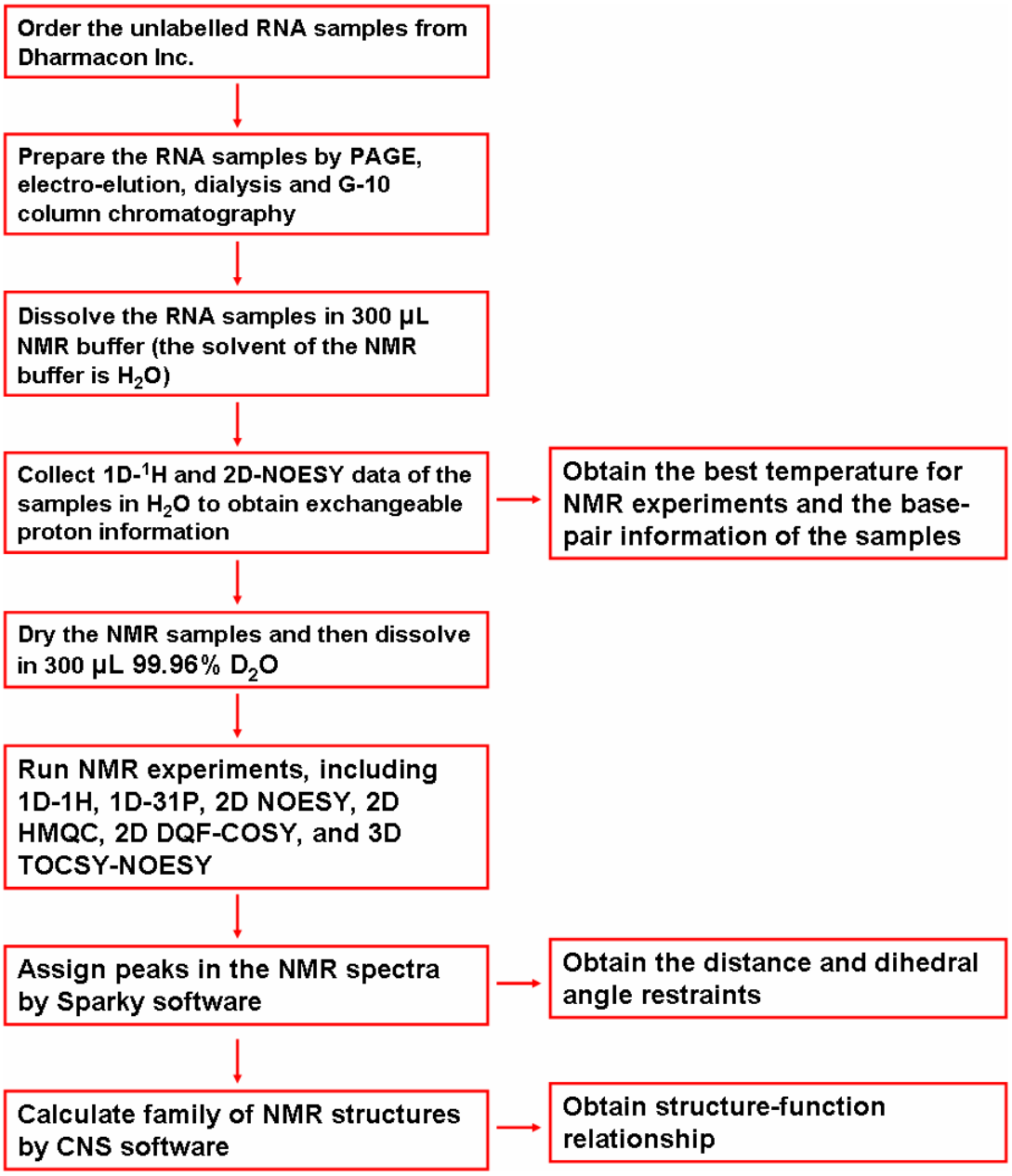


Figure 3.3. The flow chart for the process of obtaining NMR structures of the UC and the QM mutants.

3.3 NMR Experimental Results

The imino and amino regions (10~15 ppm) of 1D H_2O-NMR at various temperatures for the UC mutant (shown in Figure 3.4) and the QM mutant (data is not shown) were collected. 1D-^{1}H NMR spectra of the UC mutant in H_2O were obtained at different temperatures in Figure 3.4. The 1D-^{1}H NMR spectra are used to identify the hydrogen bonded residues, and also, are helpful for determining the most suitable temperature for acquiring the full NMR data set for structure determination. For the UC mutant, three resolved peaks between 12~14 ppm at temperatures of 5 °C, 10 °C, 15 °C, 20 °C and 25 °C support that three base-pairs formed in the stem of the designed secondary structure of the UC mutant as shown in Figure 3.2. The peaks observed between 11~12 ppm at 5 °C, which includes the imino peaks of U690 and U692 have sharper linewidths than at other temperatures, which are useful for obtaining additional NOE restraints. The 1D NMR spectra also indicate that the UC mutant is possibly more dynamic than the wild-type sequence based on the observation that the 1D NMR peaks of the wild-type sequence are sharper and have greater chemical shift dispersion than the UC mutant at 15 °C (Morosyuk 2001). For the QM mutant, the highest quality spectra were obtained at 15 °C. We also compared the fingerprint regions (i.e. crosspeaks between H1' and H6/H8) of the 400ms mixing time D_2O-NOESY of the UC mutant at 5 ^oC, 10 ^oC, 25 ^oC, and 35 ^oC and the QM mutant at 5 ^oC, 15 ^oC and 25 ^oC (Aduri 2007). The best resolution and spectral quality was obtained at 5 °C for the UC mutant and at 15 °C for the QM mutant.

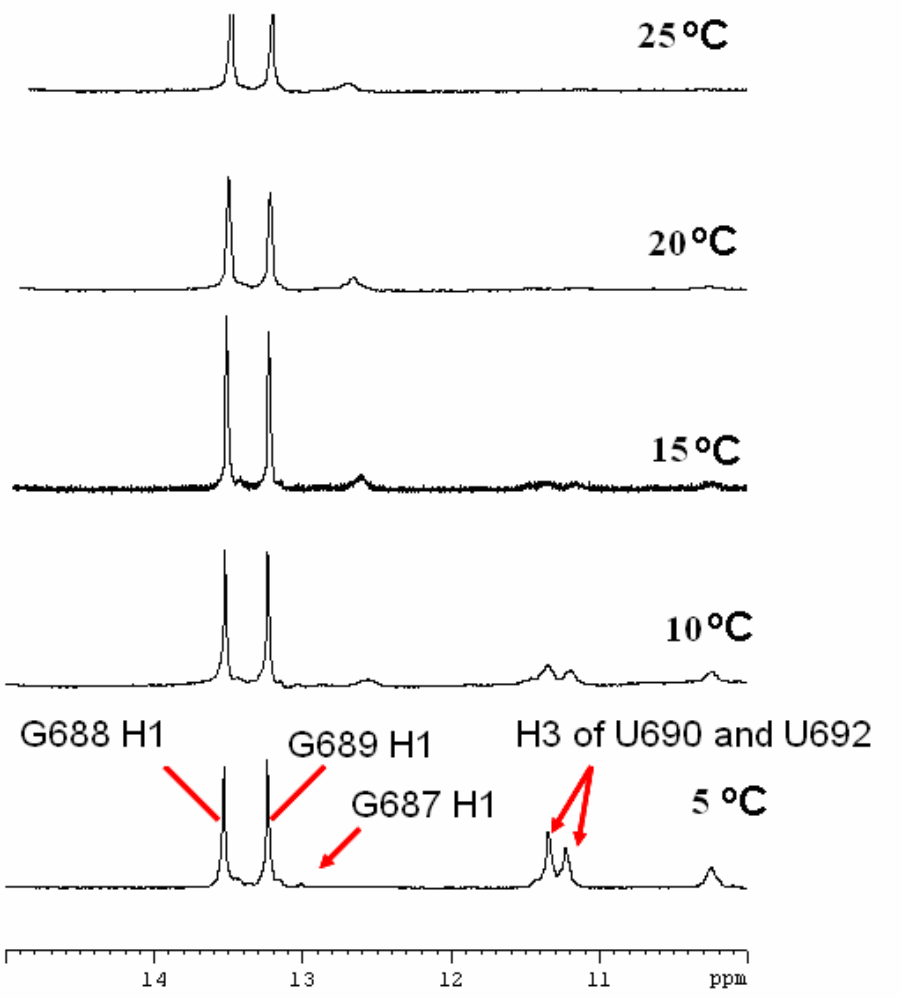


Figure 3.4. 1D-^{1}H NMR spectra of the UC mutant in H_2O at 5 °C, 10 °C, 15 °C, 20 °C, and 25 °C. For the spectra at all the temperatures, three resolved peaks show up in the region between 12 and 14 ppm, which indicate that three hydrogen bonded G or U should exist in the stem of the UC mutant. This matches three G-C base pairs in the designed structure of the UC mutant. And difference of the peak resolution can be observed at different temperatures. For the peaks of U690H3 and U692H3, the resolutions are optimum at 5 °C. Based on the peak resolution, 5 °C is chosen for the best temperature for collecting NMR data of the UC mutant.

The imino protons in G and U residues can be assigned by using the cross-peaks of imino and amino protons in the 2D H_2O-NOESY spectrum. The 2D H_2O-NOESY of the UC mutant is shown in Figure 3.5. The H1 protons of residues G1 (687), G2 (688) and G12 (698) are observed with the highest chemical shifts, which indicates that H1 of these residues are involved in hydrogen bonds. The crosspeaks of G2 (688) H1 to C13 (699) H42, G12 (698) H1 to C3 (689) H42 confirm the hydrogen bonds formed among these residues. The assignments of NH_2 or NH in H_2O-NOESY helped us to identify the hydrogen bond interactions between G687 and C700, G688 and C699, and C689 and G698 in the stems of the two mutants, which confirmed the three G-C base pairs formed in the stem of the UC mutant as shown in the secondary structure in Figure 3.2. The assignments of NH_2 or NH in H_2O-NOESY of the QM mutant were also used to identify the hydrogen bond interactions between G687 and C700, G688 and C699, and C689 and G698 in the stem of the UC mutant (Aduri 2007).

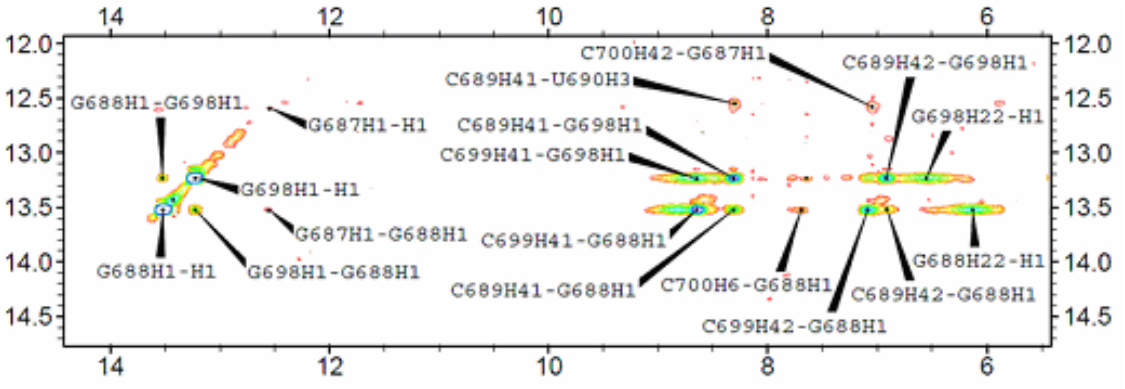


Figure 3.5. The imino and amino proton regions of the 2D H_2O-NOESY spectra of the UC mutant with assignments indicated. The peaks of C700H42-C687H1, C699H41-G688H1 and C689H41-G698H1 indicate the hydrogen bonds formed between these nucleotides.

The first step for a complete resonance assignment is the identification of base spin systems (base protons: H2, H5, H6 and H8) and sugar proton H1'. Secondly, protons of the sugar moiety (H2', H3', H4', H5', and H5'') are assigned. Assigning base protons and sugar H1' is straightforward in helical regions, based on the weak intra-residue or inter-residue NOE interaction between H6/H8 and H1'. The assignment process is divided into the identification of individual spin systems, residue type, and the connection of neighboring nucleotides to obtain sequential assignments. The identification of individual spin systems are obtained based on the chemical shift ranges that are different between sugar protons and base protons. For identification the residue type, purines can be distinguished from pyrimidines based on correlation between H5 and H6 aromatic resonances. The pyrimidine H5 and H6 resonances are relatively easy to identify because of the strong NOE magnetization transfer between between intra-residue H5 and H6 protons due to the short and constant distance of 2.4 Å. Next, the sequential assignments in the "NOE-walk" region are used to confirm the secondary structures of the samples.

Most base protons and H1' are identified by sequential links between intra-residual and inter-residual H6/H8 and H1' peaks in the "NOE-walk" region of 2D NOESY spectra of the UC mutant. Figure 3.6 shows the "NOE-walk" regions of the 2D NOESY spectra at different temperatures, the peak resolution and completion of the "NOE-walk" are used to confirm the suitable temperature chosen based on 1D-^{1}H spectra and to obtain chemical shifts of H1', H2, H5, H6 and H8 of the samples. 5 ^{o}C was confirmed to be best temperature for best peak

resolution by NOESY spectra at the temperatures of 35 ^{o}C, 25 ^{o}C, 10 ^{o}C and 5 ^{o}C (shown in Figure 3.6). At 35 ^{o}C, sequential assignments are observed from G1 to U6, the “NOE-walk” is broken due to resonance overlap and poor sensitivity and thus can not be assigned. At 25 ^{o}C and 10 ^{o}C, the H1’-H6/8 peaks of more residues can be assigned. The “NOE-walk” is complete at 5 ^{o}C, and the complete “NOE-walk” from G1 to C14 confirms the secondary structure of the UC mutant is as designed in Figure 3.2. Therefore, 5^{o}C is the optimal temperature to assign the H1’, H2, H5 and H6/8 peaks and was chosen as the temperature to run other NMR experiments based on the fact that more peaks show up in the “NOE-walk” regions of the NOESY spectra at lower temperature, and the complete “NOE-walk” is obtained only at 5 ^{o}C (Figure 3.6, D).

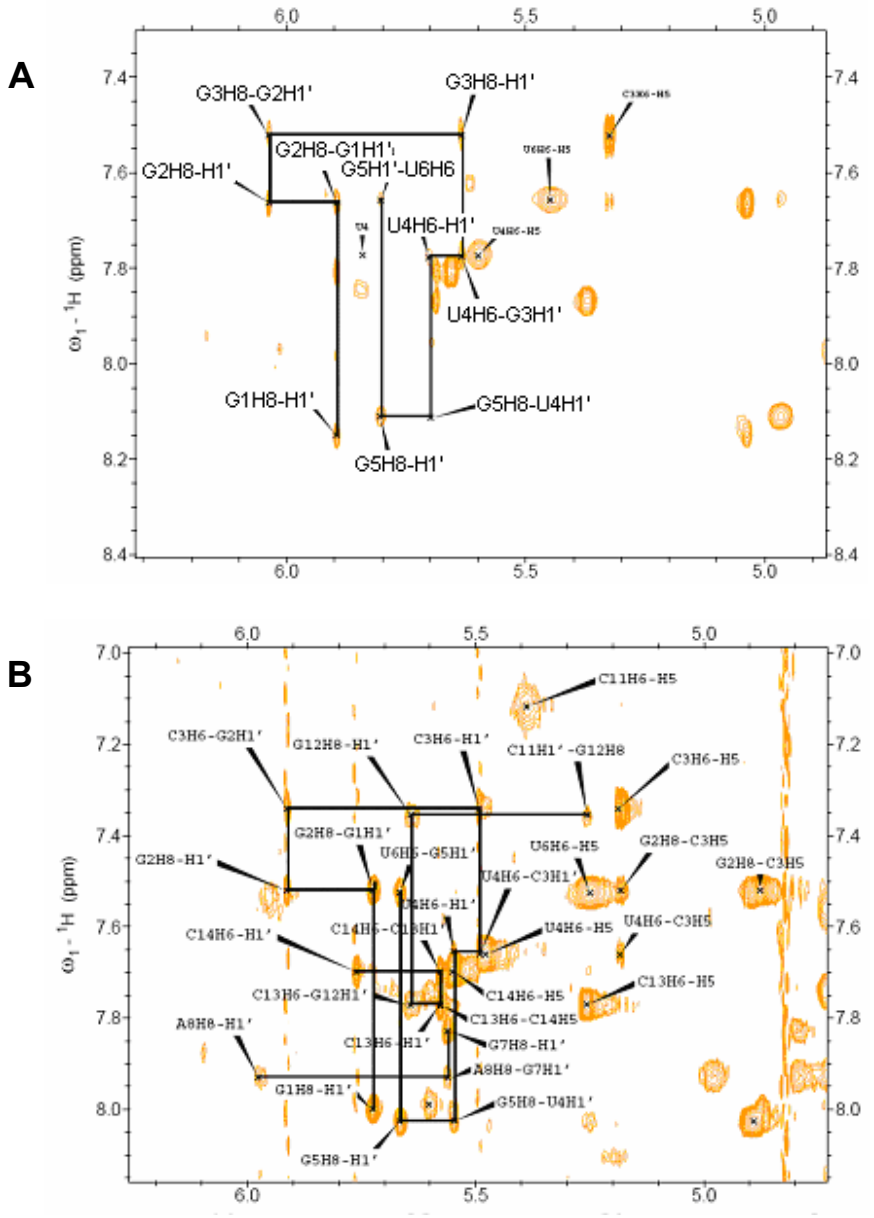


Figure 3.6. "NOE-walk" regions of NOESY spectra of the UC mutant in D_2O at 35 °C (A), 25 °C (B), 10 °C (C) and 5 °C (D). See next page for figure caption.

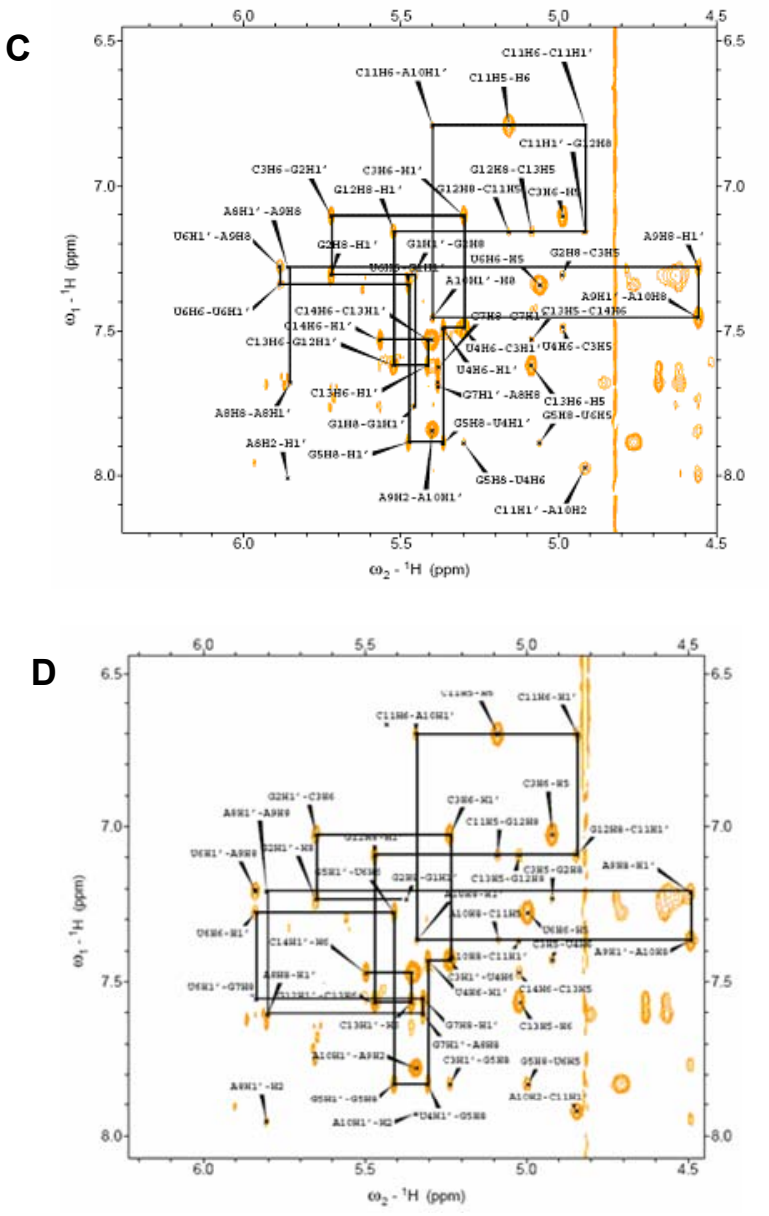


Figure 3.6. "NOE-walk" regions of NOESY spectra of the UC mutant in D_2O at 35 °C (A), 25 °C (B), 10 °C (C) and 5 °C (D). At 35 °C (A), the sequential assignments from G1 (687) to U6 (692) can be assigned. At 25 °C (B), more peaks show up in the "NOE-walk" region, and the sequential assignments are continuous from G1 (687) to A8 (694), and from C11 (697) to C14 (700). The sequential assignments at 10 °C (C) and 5 °C (D) are continuous from G1 (687) to C14 (700). The cross-peaks are resolved better at 5 °C than 10 °C.

The NOESY fingerprint regions of the UC and QM mutants were used to obtain sequential assignments of H1' and H6/H8 resonances (Figure 3.7). In addition, the fingerprint regions of the UC and QM mutants have several structurally diagnostic crosspeaks. Strong inter-residue NOE crosspeaks involving A695H2-A696H1' and A696H2-C697H1' of the UC mutant and similar crosspeaks involving A694H2-C695H1' and A696H2-A697H1' of the QM mutant (labeled in blue in Figure 3.7) indicate the base stacking of these nucleotides. Weak crosspeaks are observed for A694H8-U692H1' and A695H8-U692H1' of the UC mutant and A694H8-U692H1' and C695H6-U692H1' of the QM mutant (labeled in red in Figure 3.7), which are diagnostic of a "U-turn" motif at position 692 (Morosyuk 2001). Corresponding NOE crosspeaks with similar intensities were also observed in the fingerprint region of the wild-type 690 sequence obtained by Morosyuk *et al.* (Morosyuk 2001). The similarities of these unusual peaks in the spectra of the two mutants and the wild-type are a preliminary indication that the two functional mutants are likely to form isomorphous structures with the wild-type.

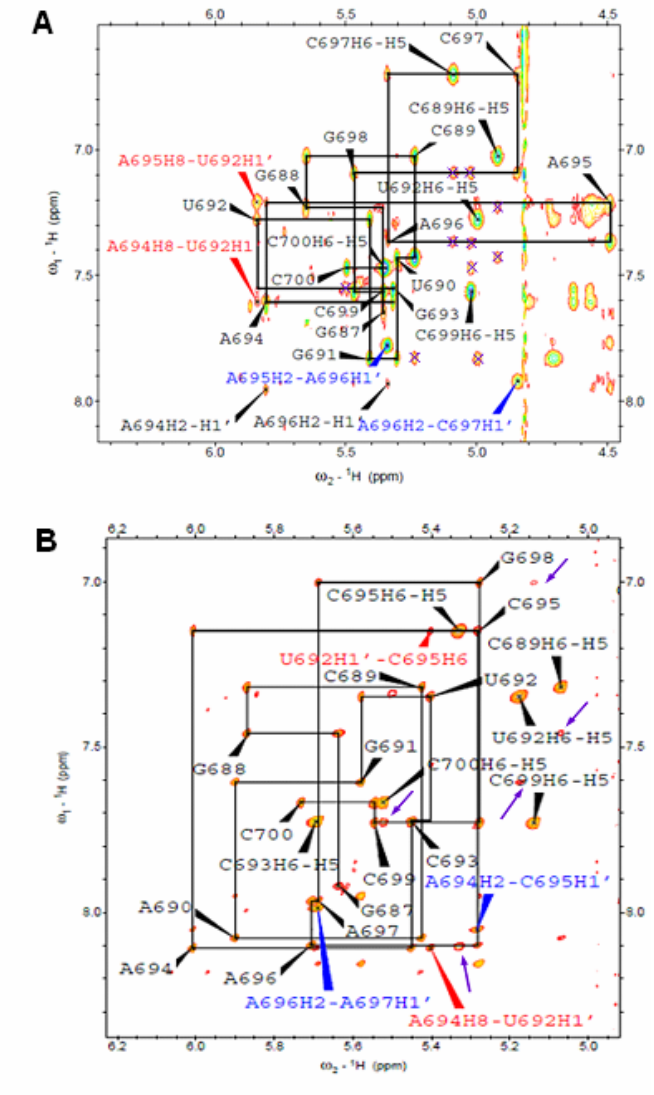


Figure 3.7. The fingerprint regions of the D_2O-NOESY spectra of the UC mutant (A) and the QM mutant (B). Weak peaks of A695H8-U692H1' and A694H8-U692H1' (in A) and C695H6-U692H1' and A694H8-U692H1' (in B) are marked in red, and these peaks indicate the unusual orientation of 692, resulting in the distances between U692H1' and 694H8 or 695H8 that are close enough to be detected by NOESY. The peaks of A695H2-A696H1' and A696H1'-C697H1' (in A), and A694H2-C695H1' and A696H2-A697H1' (in B) are labeled in blue, which indicate base stacking among residues 694, 695, 696 and 697. The cross-peaks of inter-residue base protons are labeled by purple crosses in A and purple arrows in B.

The peaks shown in the “NOE-walk” region also provide information about the *chi* dihedral angles, In helical A-from RNA, the intra-residue distances between the H6 of pyrimidines or H8 of purines and H1’ of their own sugars are around 3.7 Å, and the inter-residual distances between H6 or H8 and H1’ of the preceding nucleotide (i.e. in the 5’-direction) are 4.3 Å (although the distances between H1’ (n-1)-H6/H8 (n) is larger than 4 Å, a cross-peak can be observed partially due to spin diffusion via H2’. Therefore, H6 or H8 have two NOE cross-peaks, to the H1’ protons of their own sugar and to the preceding nucleotide in the 5’ direction, except the 5’-ending nucleotide. This is the anchoring point to start the NOE-based assignment walk. Besides 5’-ending nucleotide, 3’-terminal nucleotide can also be used as a starting point for NOE-walk assignments based on the fact that only 3’-terminal nucleotide has one NOE cross-peak between H6/H8 and H1’, H6/H8 (n) and H1’ (n), and other residues all have one more NOE-peak involving H6/H8-H1’ (n) and H2/H8-H1’ (n-1). The chemical shifts of the proton in the 5’- or 3’- terminal residues are usually shifted, for example, the H2’ and the H3’ at 3’-terminal are moved to lower field because of the missing phosphate group. And all the intra-residual H1’-H6/H8 peaks are weak, which indicates that all residues in the UC and the QM mutant are in *anti* conformations.

In A-form RNA helices, two relatively strong NOE peaks (H2’(n-1) -H6/H8 (n), H3’ (n) – H6/H8 (n), their distances are shorter than 3 Å) and two relatively weak peak (H2’ (n) – H6/H8 (n), H1’ (n) –H6/H8 (n)) are supposedly observed between H6/H8 and sugar protons for every residue if the NOESY spectra are collected at shorter mixing time (such as 150 ms in my experiments), and by

using longer mixing time (such as 400 ms), weak NOE peaks between H6/H8 (n) – H4', H5' or H5'' (n) can also be observed due to spin diffusion. Based on the peak strength comparison, most sugar protons, H2' ~H5'' of the UC mutant are assigned. The sugar protons, H2' /H3'/H4'/H5'/H5''-H6/H8 region of the UC mutant is shown in Figure 3.8, B. In the region of H1'/H5-H2' ~H5'', the strongest peaks are intra-residual H1'-H2', and H1'-H4' are usually stronger than H1'-H3'. The region of H1'/H5-H2' ~H5'' of the UC mutant is shown in Figure 3.8, A.

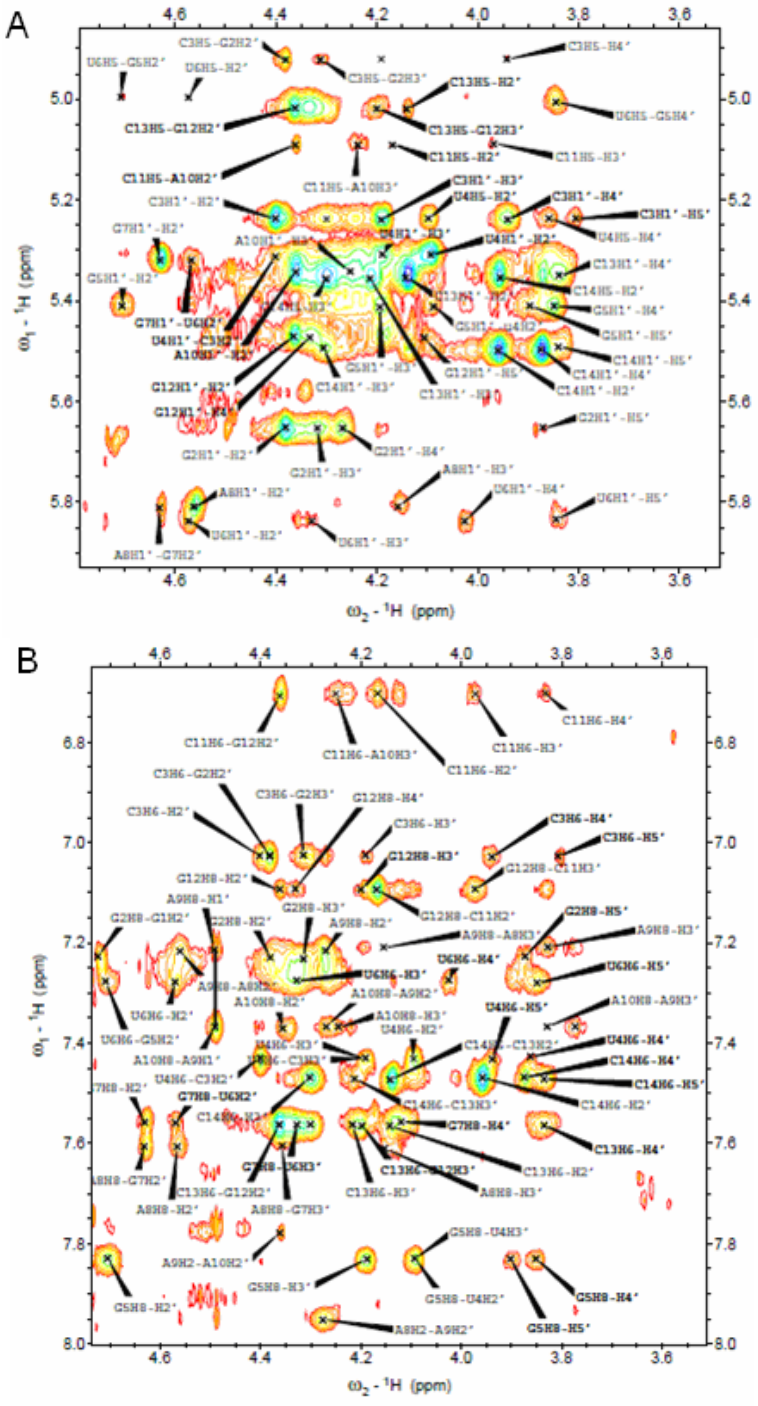


Figure 3.8. The regions of H1'/H5-H2'~H5'' (A) and H6/H8-H2'~H5'' (B) in the NOESY spectrum of the UC mutant. Most sugar protons are assigned in the two regions.

To resolve assignment ambiguities in NOESY spectra caused by spectral overlap (especially the sugar proton regions shown in Figure 3.8), we combined information from other spectra, including HMQC, DQF-COSY, and 3D TOCSY-NOESY. Although ^{13}C-^{1}H HMQC is not as sensitive as NOESY for unlabeled samples, because the ^{13}C is only 1.11% in the natural abundance sample and the gyromagnetic ratio of ^{13}C is only one fourth of ^{1}H, natural abundance ^{13}C-^{1}H HMQC experiments are valuable for simplifying some peak assignments and confirm initial assignments of NOESY spectra for RNA samples (Kruger 1982). ^{1}H-^{13}C HMQC spectra at different carbon frequencies were used to identify chemical shifts of H1', H2, H5, H6 or H8 by their attached carbon chemical shifts (Varani 1991). HMQC spectra simplify peak assignments. 2D ^{1}H-^{13}C HMQC (Figure 3.9) at different carrier frequency settings of 95 ppm and 145 ppm were used to confirm the assignments of H1', H6 of pyrimidines or H8 of purines, H2 of Adenines, and H5 resonances of pyrimidines. H2 chemical shifts can be resolved in HMQC spectra, because C2 region (around 150 ppm) is well separated from other carbon atoms in the RNA. Distinguishing adenine H2's from purine H8's and distinguishing cytosine from uracil base protons can simplify the proton assignment. The H1'-C1' resonances are usually well dispersed, the frequencies of the H1' resonances can be obtained and separated from H5 from HMQC. And also, the H5 protons of cytosine can be differentiated from uracil based on the separation of C5 of cytosine from uracil in HMQC. The HMQC spectra of the QM mutant are also very useful for confirming the chemical shifts of base protons (H2, H5, H6, and H8) and H1' (Varani 1991).

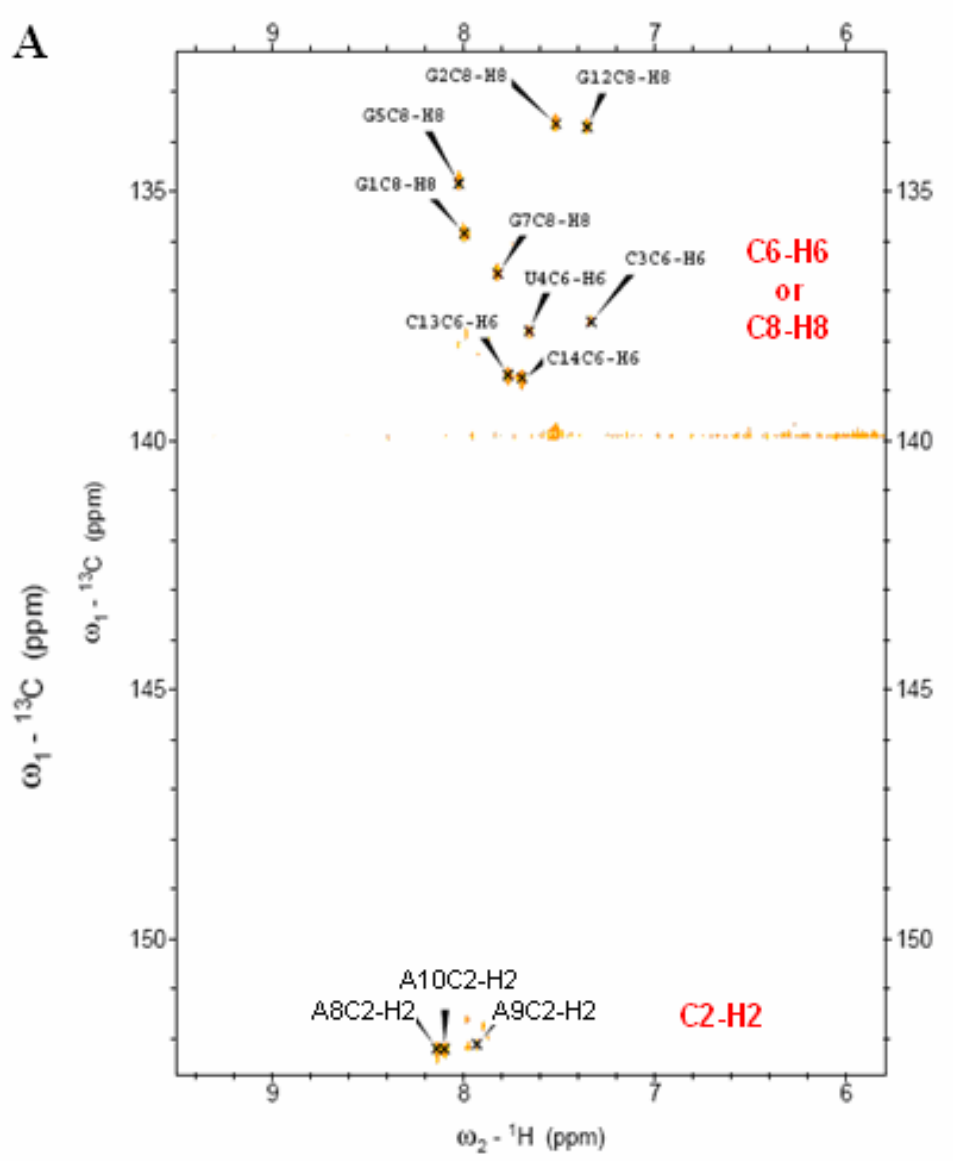


Figure 3.9. HMQC spectra of the UC mutant. C2-H2 and C6/8-H6/8 region is shown in (A), and C1'-H1' and C5-H5 region is shown in (B). The initial NOESY assignments of most base protons in the UC mutant, including H2 of adenine, H6 of pyrimidine, and H8 of purine are confirmed in A. And most assignments of H1' and all H5 of pyrimidine based on NOESY spectrum are confirmed in B.

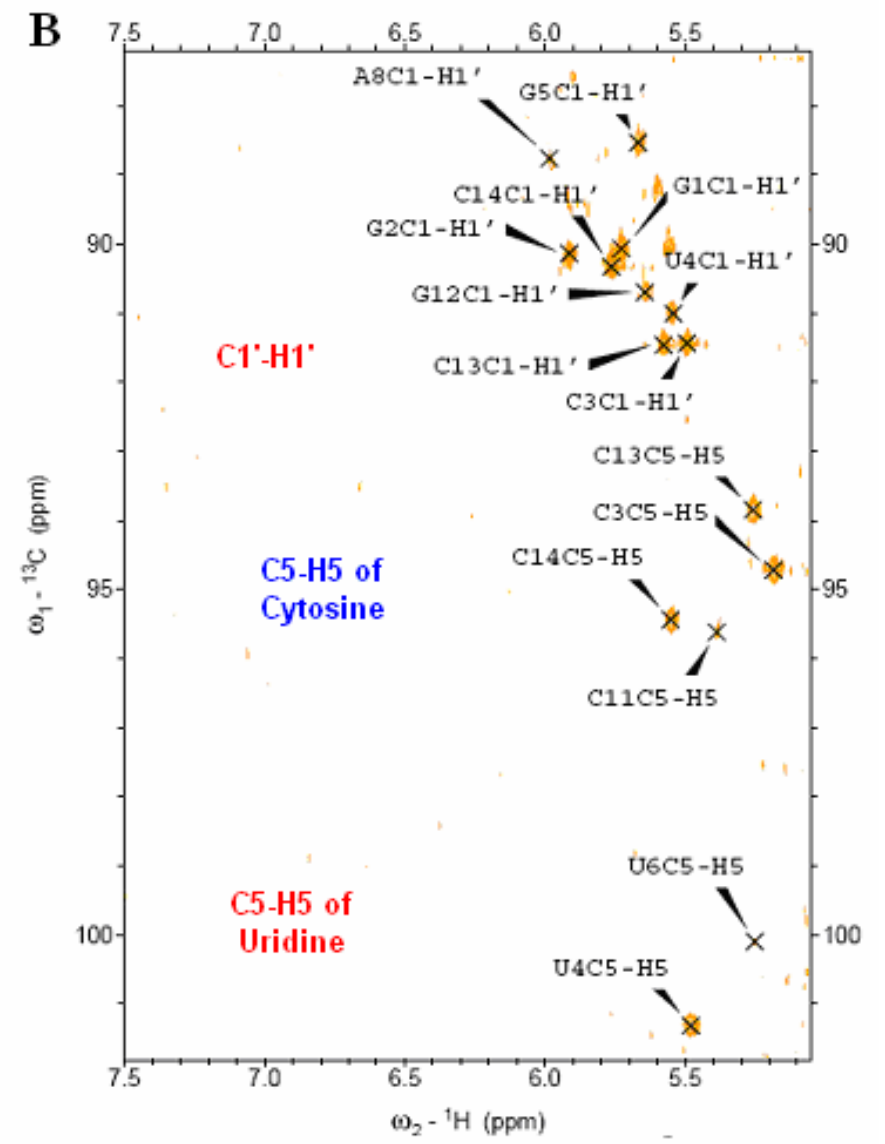


Figure 3.9. HMQC spectra of the UC mutant. C2-H2 and C6/8-H6/8 region is shown in A, and C1'-H1' and C5-H5 region is shown in B. The initial NOESY assignments of most base protons in the UC mutant, including H2 of adenine, H6 of pyrimidine, and H8 of purine are confirmed in A. And most assignments of H1' and all H5 of pyrimidine based on NOESY spectrum are confirmed in B.

Most assignments in NOESY, especially the sugar proton assignments that are partially overlapped in 2D spectra, are resolved in the 3D TOCSY-NOESY spectrum. Nearly complete assignments of the protons in sugar moiety (H1', H2', H3', H4', H5' and H5") of the D_2O-NOESY spectrum were confirmed by 3D TOCSY-NOESY. The chemical shifts of the protons are listed in Table 3.1.

Table 3.1. The chemical shifts of the protons in the UC mutant (A) and the QM mutant (B).

A

Residues	H1'	H2'	H3'	H4'	H5'/H5''p	H2/H5	H6/H8	NH/NH_2
	ppm	ppm	ppm	ppm	pm	ppm	ppm	ppm
G687	5.372	4.722	4.220	ND	3.644	NA	7.692	12.57
G688	5.653	4.384	4.313	4.267	3.873/4.178	NA	7.230	13.52/6.129
C689	5.237	4.404	4.191	3.941	3.810/3.714	4.921	7.027	8.308/6.915
U690	5.307	4.095	4.184	3.859	3.938	5.234	7.430	ND
G691	5.410	4.707	4.194	3.848	3.898	NA	7.829	ND
U692	5.838	4.570	4.327	4.027	3.846	4.997	7.278	ND
G693	5.320	4.631	4.359	4.125	3.831	NA	7.557	ND
A694	5.805	4.561	4.155	3.995	3.637/3.574	7.952	7.606	ND
A695	4.492	4.270	3.831	3.919	3.768	7.776	7.209	ND
A696	5.342	4.360	4.244	4.253	ND	7.922	7.336	ND
C697	4.845	4.167	3.970	3.834	ND	5.087	6.702	ND
G698	5.470	4.361	4.199	4.337	4.103	NA	7.093	13.23/6.552
C699	5.354	4.142	4.213	3.841	ND	5.023	7.656	8.646/7.086
C700	5.497	3.958	4.302	3.873	3.843	5.351	7.470	7.048

B

Residues	H1'	H2'	H3'	H4'	H5'/H5''p	H2/H5	H6/H8	NH/NH_2
	ppm	ppm	ppm	ppm	pm	ppm	ppm	ppm
G687	5.634	4.813	4.503	4.075	3.95/3.87	NA	7.921	12.12
G688	5.871	4.601	4.549	4.277	4.13/4.09	NA	7.459	13.38
C689	5.425	4.706	4.415	4.502	4.09	5.07	7.319	7.75/6.69
A690	5.9	4.521	4.664	4.51	4.2/4.1	7.949	8.077	ND
G691	5.579	4.26	4.675	4.454	4.12	NA	7.606	12.54
U692	5.403	4.611	4.365	4.224	4.02	5.178	7.347	11.95
C693	5.45	4.384	4.512	4.222	ND	5.696	7.726	ND
A694	6.008	4.615	4.832	4.393	4.08	8.05	8.107	ND
C695	5.283	4.469	4.248	4.402	4.09	5.33	7.148	ND
A696	5.706	4.514	4.624	4.4	ND	7.986	8.10	ND
A697	5.689	4.592	4.569	4.556	ND	8.154	7.963	ND
G698	5.278	4.557	4.357	4.415	4.04/4.02	NA	7.003	12.47
C699	5.546	4.335	4.436	4.426	ND	5.141	7.728	8.40/6.87
C700	5.733	4.073	4.184	4.505	ND	5.522	7.669	8.25/6.79

ND-Not determined; NA-Not applicable

Structure determination of nucleic acids by NMR needs to quantify NMR parameters that are sensitive to the conformation of the molecule, such as scalar coupling, chemical shifts, relative distances. The distance restraints are mainly obtained from NOESY and 3D-TOCSY-NOESY spectra. Dihedral angle restraints are important for defining geometry of RNA residues and accurate structure calculations. For RNA, there are seven dihedral angles (*α, β, γ, δ, ε, ζ, χ*) can be used to define the geometry of each RNA residue (Saenger 1984). In the NOESY spectra of the UC mutant and the QM mutants, the intra-residue H8/H6-H1' cross-peaks of all fourteen nucleotides are observed to be weaker than the intra-residue H8/H6-H3' cross-peaks, therefore, all bases in the UC mutant and the QM mutant have χ dihedral angles that are predominantly in the *anti* conformation. DQF-COSY spectra provided useful information regarding the sugar pucker conformation based on the intensity of intra-residue H1'-H2' cross-peaks (Varani 1996). In DQF-COSY and TOCSY-NOESY, strong H1'-H2' intra-residue cross-peaks indicate that the sugar puckers are predominantly S-type (C2'-endo), conversely a weak or non-existent H1'-H2' cross-peak indicates an N-type (C3'-endo) sugar. For the DQF-COSY of the UC mutant (shown in Figure 3.10), only one weak peak is observed for A694H1'-H2'. This suggests the sugar pucker of A694 is dynamic, while the sugar puckers for all other residues are essentially pure C3'-endo conformation. This conclusion is confirmed by the 3D TOCSY-NOESY of the UC mutant, which shows a weak cross-peak for A694H1'-H2', but no H1'-H2' correlations for any other residues (data not shown). The DQF-COSY spectrum of the QM mutant has weak intra-residue H1'-H2' cross-

peaks for residues U692, C693, A694 and C700, which suggests dynamic but predominantly C3'-endo sugar conformations for these nucleotides (data not shown). Therefore, the sugars in both the UC and the QM mutants mainly show in C3'-endo conformation.

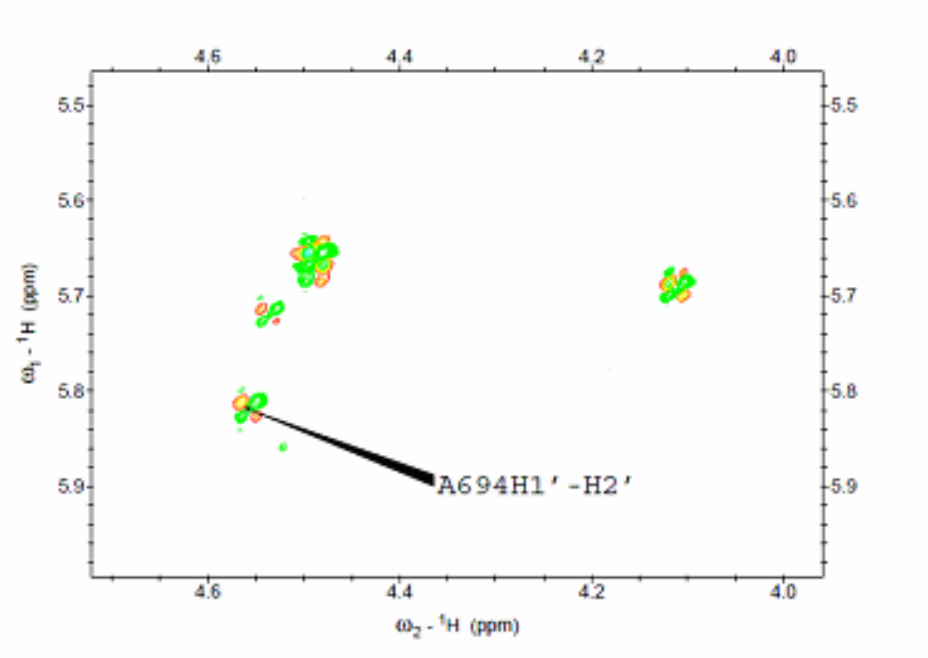


Figure 3.10. The H1'-H2' region of the DQF-COSY spectra of the UC mutant. The H1'-H2' peaks of A694 indicate the partial C2'-endo formation in the residue. Three more peaks can be observed in the spectrum, however, the chemical shifts of the three peaks do not match any residue in the UC mutant, which may be caused by dynamic of some residues in the UC mutant.

No unusual shift of ^{31}P resonances are observed in the ^{31}P-^{1}H HETCOR spectrum of the QM mutant (data not shown), which suggests that no residues in the QM mutant have *trans* conformation for either α and ζ dihedral angles. The 1D-^{31}P spectrum of the UC mutant is shown in Figure 3.11. The sharp peak with the largest peak area is set at 0 ppm, which is peak from solvent. The overlapped peaks with chemical shifts from -1.0 to -2.5 ppm are grouped together, which suggests that the phosphorus of these residues have similar chemical shifts. Two peaks located at -0.4 ppm and 0.8 ppm have unique chemical shifts compared with the grouped peaks in the range of -1.0 to -2.5 ppm. The peak area for the two single peaks and the grouped peaks are 1.0, 1.6 and 28, respectively. And the peaks located at -0.4 and 0.8 ppm weigh about 45% and 71% of one residue, which indicates that the dihedral angles α and ζ of a small percentage of residues (the peaks with area 1.0 and 1.6) are different from most residues in the UC mutant.

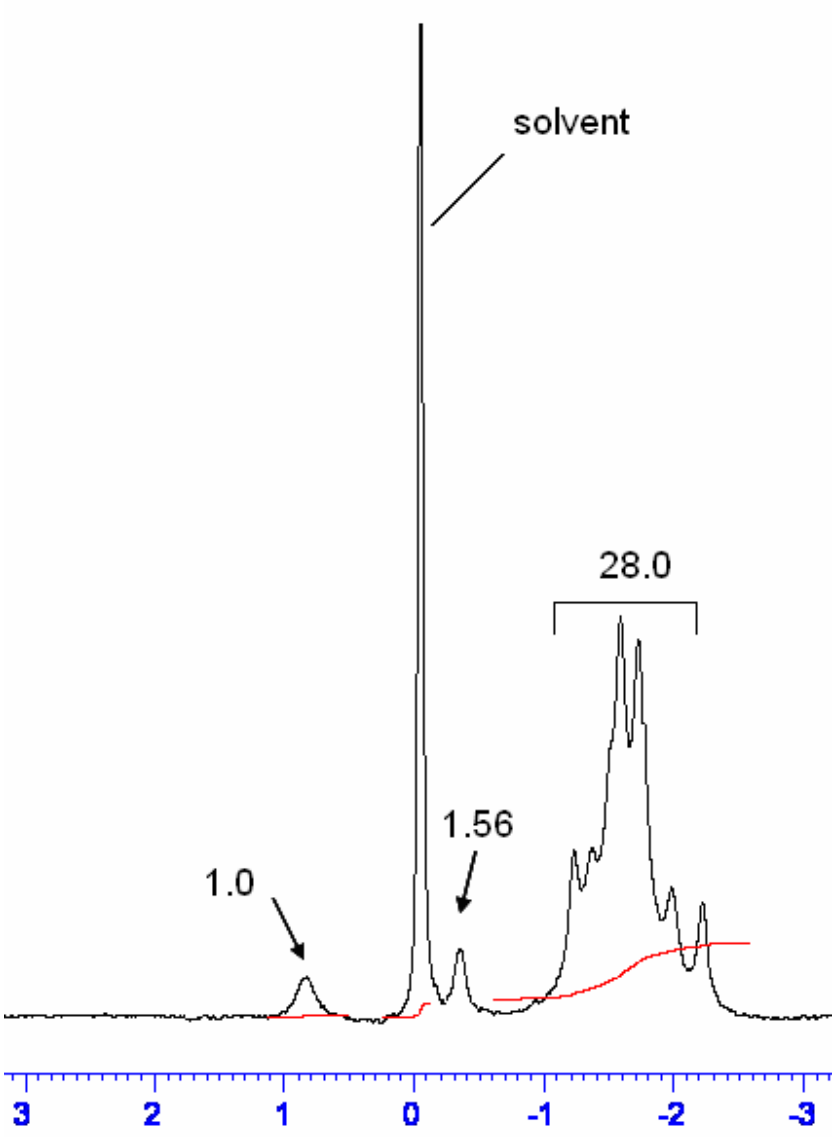


Figure 3.11. The 1H-^{31}P spectrum of the UC mutant.

3.4. Structure Calculations

Distances between protons in the UC and QM mutants have been measured by integrating the cross-peak volumes of NOESY spectra at different mixing times (60ms, 90ms, 120ms, and 150ms) to assess spin diffusion and determine quantitative distance restraints for modeling studies. The volume of the intra-residue H5-H6 cross-peaks of pyrimidines, whose distance is fixed as 2.43 Å and is used as a spectroscopic ruler to scale measured NOE volumes according to the $1/r^6$ relationship (assuming the 2-spin approximation) to deduce estimated distances (Varani 1991). Error bars of 0.5 to 1.0 Å were used in structure calculations to account for quantification errors and spin-diffusion. Cross-peaks with spectral overlap were not used to determine distances. Eighteen hydrogen bond distance restraints were used to ensure the proper formation of the three G-C base pairs in the stem of the UC or the QM mutants. Seven dihedral angles (*α, β, γ, δ, ε, ζ* and *χ*) for each residue were used to constrain the helical residues (G1, G2, C3, G12, C13, C14) to A-form geometry

The structural modeling methods used for the UC and the QM mutants are essentially same, therefore, the UC mutant is used to explain the calculation details. For the UC mutant, 53 intra-residue and 91 inter-residue distance restraints were extracted by integrating D_2O-NOESY spectra at low mixing times (60, 90, 120 and 150 ms) at 5 °C, and 80 dihedral angle constraints were input into CNS 1.1 software. 50 global-folded structures were obtained by simulated annealing calculations using CNS (Brünger 1998), using a protocol similar to that described previously (Morosyuk 2001). Some of these global-folded structures

displayed poor geometry, such as distorted base pairing, positive overall energy, and steric clashes. Distance restraints were added to each base pair to ensure proper geometry in the stem. Some of the preliminary global-folded structures contained inter-proton distances that did not match the corresponding peak volumes in the NMR spectra, such structures with observed vs. restraint distance conflicts greater than 0.5Å were rejected from further refinement. However, further analysis of both accepted and rejected structures revealed cases where inter-proton distances where too short (i.e. <4 Å) to be consistent with the actual NOESY spectral data that displayed no corresponding cross-peak at any mixing time. For such cases, we imposed a “Un-NOE” (Unobserved NOE) restraint of > 4.5 Å was added to force the two atoms apart (Diener 1998). A total 162 Un-observed NOE restraints were added to improve the structures of the UC mutant. The global-folded calculations were then rerun with inclusion of the Un-NOE restraints. The result was a higher percentage of global-folded structures with low energy and proper base pair formation. In addition, the RMSD overlap improved significantly upon addition of the Un-NOE restraints. The twelve lowest energy structures with no restraint violations were chosen from the global-folded structures and submitted into the refinement calculation steps. The restraints used in structure calculations for the UC mutant and the QM mutant are summarized in Table 3.2.

Table 3.2. Summary of distance restraints, dihedral angle restrains and the RMSDs for the UC mutant (A) and the QM mutant (B).

A	
Base pair restraints	34
Total NOE distance restraints	144
Intra-residue	53
Inter-residue	91
Mean NOEs per restraint	10.3
Dihedral angle restraints	84
Mean dihedrals per restraint	6.0
Unobserved NOEs	162
Total restraints	424
Total restraints per residue	30.3
All-atom RMSD (Å)	0.80

B	
Base pair restraints	34
Total NOE distance restraints	144
Intra-residue	49
Inter-residue	95
Mean NOEs per restraint	10.3
Dihedral angle restraints	90
Mean dihedrals per restraint	6.4
Unobserved NOEs	2
Total restraints	270
Total restraints per residue	19.3
All-atom RMSD (Å)	0.42

The average pair-wise RMSD among the twelve final structures of the UC and the QM mutants are 0.80 and 0.42 Å, respectively. The refined structures with the lowest force field energy of the UC and the QM mutants are shown in Figure 3.12, A and B. All structures were visualized using PYMOL software (DeLano, W.L. The PyMOL Molecular Graphics System (2002), http://www.pymol.org).

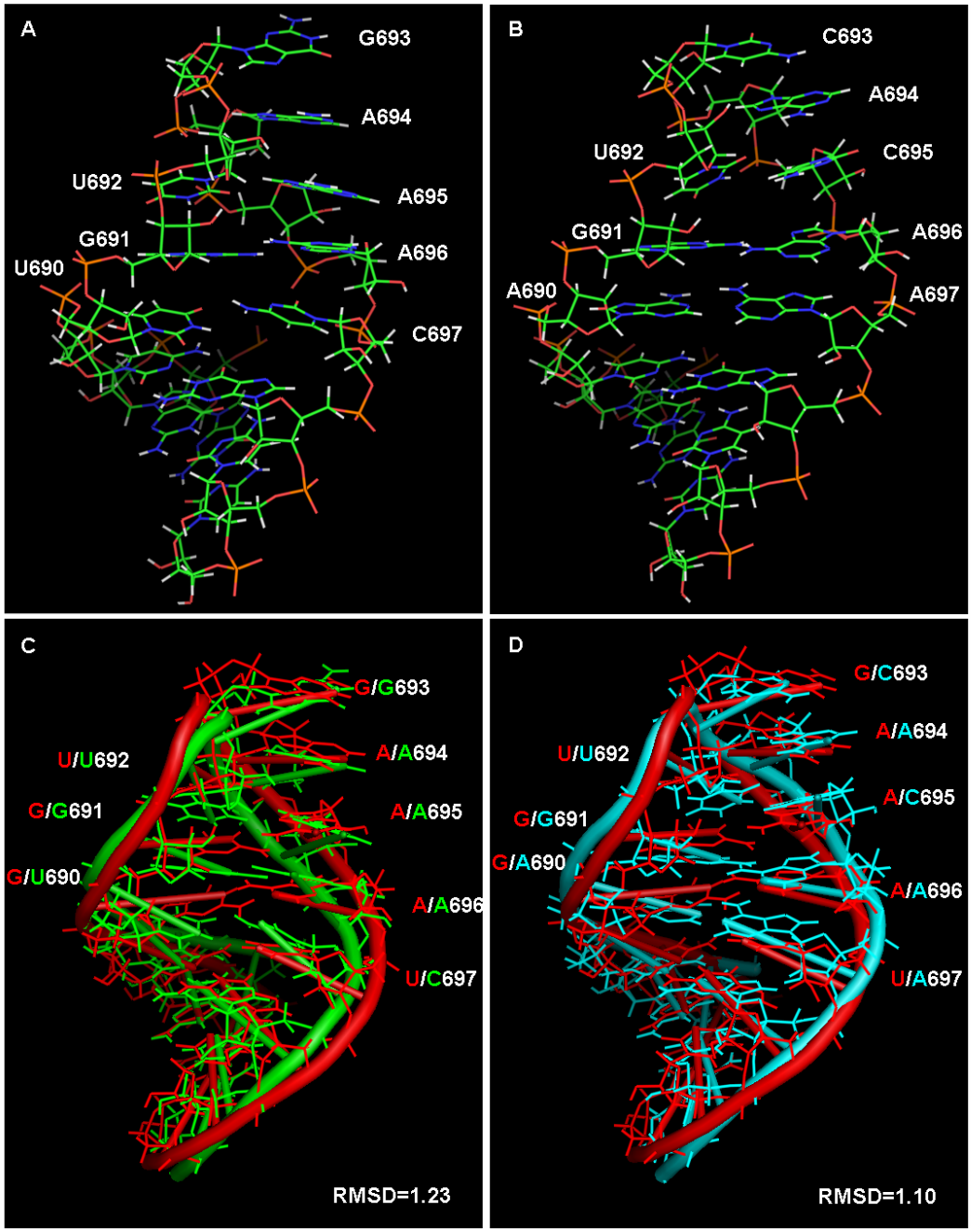


Figure 3.12. Lowest energy refined structures of the UC mutant (A) and of the QM mutant (B). The overlap of the wild-type (in red) and the UC mutant (in green) structures (C), the wild-type (in red) and the QM mutant (in cyan) structures (D). The structures of the wild-type, the UC mutant and the QM mutant have similar overall structural folds.

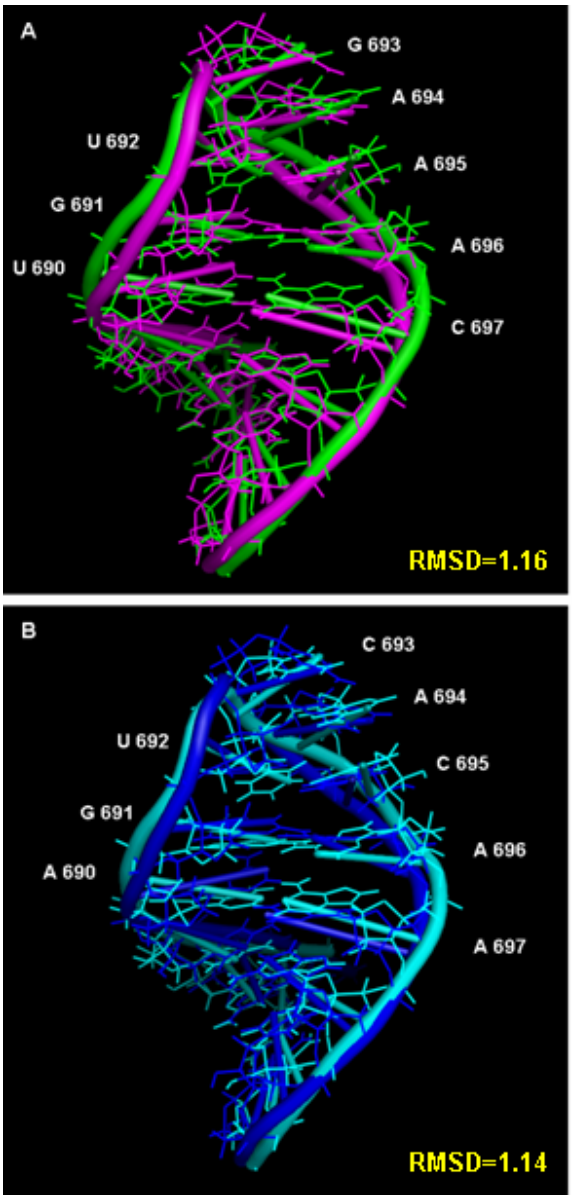


Figure 3.13. The overlap of the NMR structure and the homology model of the UC mutant (A), and the overlap of the NMR structure and homology model of the QM mutant (B). The NMR structure of the UC mutant, the homology model structure of the UC mutant, the NMR structure of the QM mutant, and the homology model of the QM mutant are shown in green, magenta, cyan, and blue. The NMR structures of the UC mutant and the QM mutant are similar as their homology model structures. The homology model structures of the UC and the QM mutants were obtained by RNA-123 software.

3.5. Discussion of the NMR Structures of the UC and the QM Mutants

Figure 3.12, C and D show the superposition of the wild-type 690 loop (PDB ID: 1FHK), with the UC and the QM mutants with respective RMSDs of 1.23 and 1.10 (for the mutated positions, only the sugar and phosphate atoms are included in the RMSD calculation). Small differences are observed between the structures of the wild-type, the UC mutant, and the QM mutant. The structural similarity of the UC mutant, the QM mutant and the wild-type supports the hypothesis that the functional mutants form isomorphous structures as the wild-type.

Key features of the wild-type structure are also observed in the structures of the UC and the QM mutants: mismatches at the 690 and 697 positions (Figure 3.14); mismatches at 691 and 696 (Figure 3.15, D and E); a “U-turn” formed at position 692 (Figure 3.16, B and C); a continuous stacking at the 3’ end of the loop (Figure 3.16, D, E and F). Since these structural characteristics are present in the wild-type sequences, we infer that they are possibly crucial for the high function of the 690 hairpin.

Positions 690 and 697 have a high degree of covariation according to the instant evolution studies, and structures of the functional mutants at these two positions have been proposed to form a sheared mismatch pair as G690•U697 mismatch in the wild-type by modeling studies (Figure 3.14, C and D) (Morosyuk 2001; Morosyuk 2001). For the UC mutant, the geometry of U690•C967 mismatch predicted by homology modeling (Figure 3.14, A) is slightly different from the experimental structure. The NMR structure shows H-bonding from

U690-O4 to C967-H42, rather than the bifurcated H-bonds from U690-O2 to C967-H41/H42 in the homology model (Leontis 2001). For the QM mutant, a medium strength cross-peak of A690H2-A697H2 was observed in the NOESY spectra. The final NMR structure of the QM mutant shows a Watson-Crick type A690•A697 mismatch that is stabilized by an H-bond from A690-N1 to A697-H61 (Figure 3.14, B). However, the homology model predicts A690•A697 interacts by a sheared pair stabilized by a hydrogen bond between A690-N3 and A697-H62 (Figure 3.14, D).

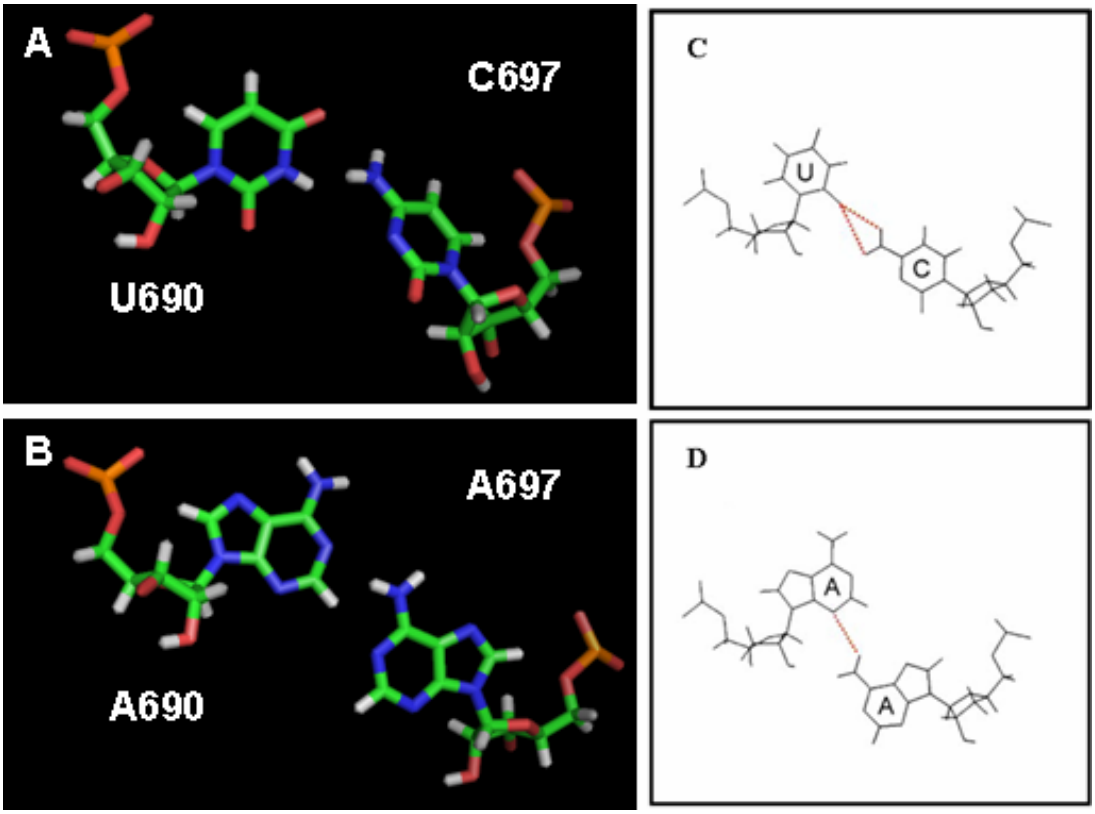


Figure 3.14. The 690-697 mismatches of the UC mutant (A) and the QM mutant (B). The homology models of the UC mutant (C) and the QM mutant (D) at 690 and 697 built by Morosyuk *et. al.* (Morosyuk 2001; Morosyuk 2001). The distance between O2 (U690) and H41 (C697) is 2.58 Å in A, and the distance between N1 (A690) and H61 (A697) is 2.79 Å, and the possible hydrogen bonds are shown in yellow dotted lines in A and B.

Although the mismatch structures of the 690 and 697 of the UC and the QM mutants are not the same as the expected models, the hypothesized critical hydrogen-bond acceptor atoms, the O2 of C697 in the UC mutant, and N3 of A697 in the QM mutant, are located at essentially the same locations as the hydrogen-bond acceptor atom O2 (U697) in the wild-type (the atoms are labeled by magenta arrows in Figure 3.15). Maintainance of hydrogen-bond acceptor atoms at similar positions in the two functional mutants and the wild-type suggests that the presence of the hydrogen-bond acceptor atoms is essential for maintaining the functions the 690 loop (Morosyuk 2001). This hypothesis is supported by the observed hydrogen bond formed between residues U697O2 and G785H22 in the crystal structure of the small subunit of the ribosome of *E. coli* (Wimberly, Brodersen et al. 2000). The perturbed geometries of the UC and AA mismatches may be related to lower functions of the two mutants than the wild-type.

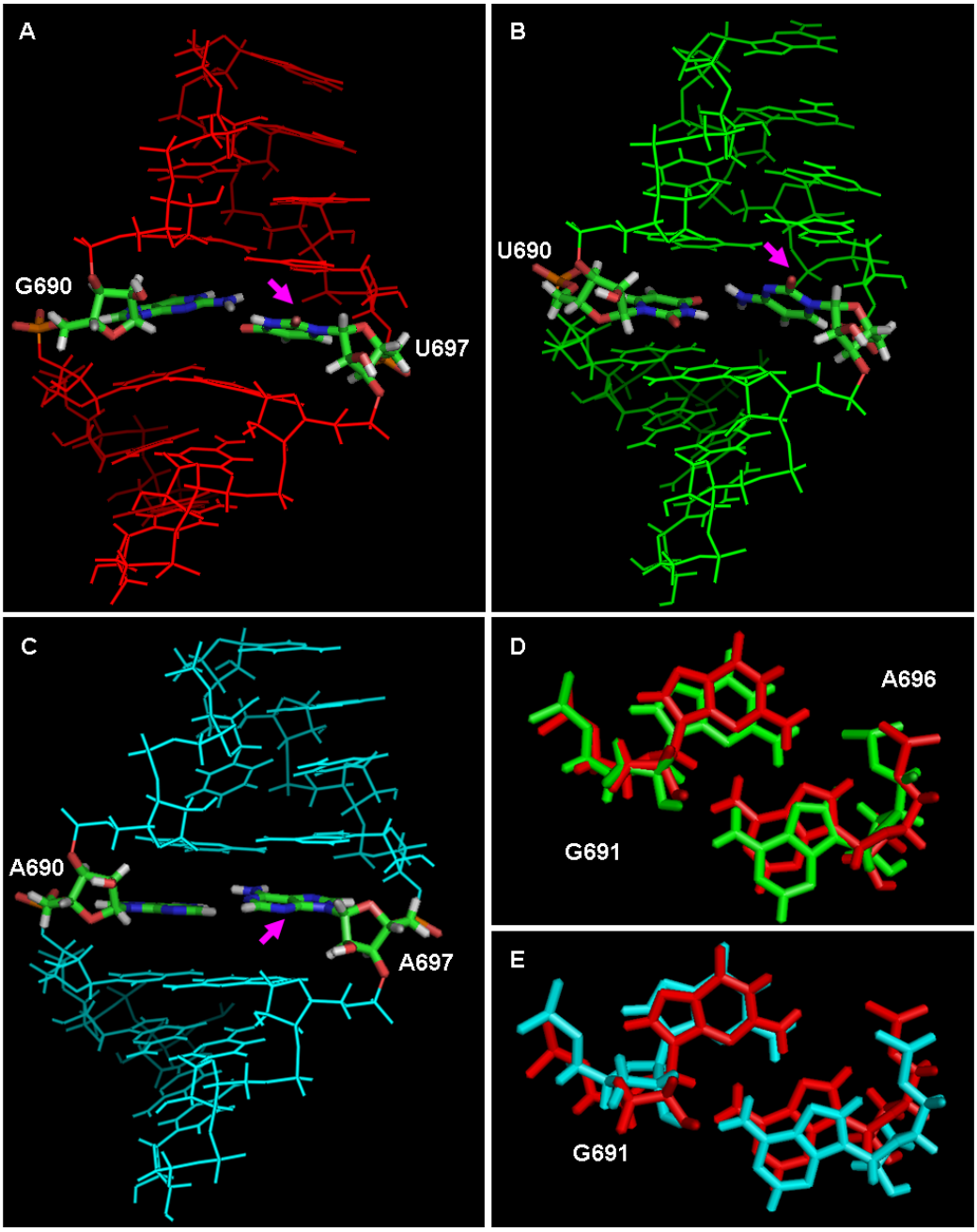


Figure 3.15. The structures of the wild-type (A), the UC mutant (B) and the QM mutant (C). The overlapping of the wild-type (in red) and the UC mutant (in green) is shown in D, and the overlapping of the wild-type (in red) and the QM mutant (in cyan) at 690 and 696 in shown in E.

The geometry of the G691•A696 in the NMR structure of the UC and the QM mutants are similar to the wild-type (Figure 3.15, D and E). The perturbed positions of the residue 696 in the two mutants are possibly due to the mutations at 690 and 697. The structural differences at position 696 in the wild-type and mutants may also be a factor that contributes to the lower functions of the UC and the QM mutants.

The existence of the “U-turn” motifs in NMR structures of the UC mutant, the QM mutant, and the wild-type sequences are to change the direction of the backbones (shown in Figure 3.16, A, B, and C) and to place residue 693 at the top of the 690 loop in proximity of the E-site mRNA codon (Morosyuk 2001) (discussed more below). This structure suggests that nucleotide 693 may act to slow down translation, thereby, improving the fidelity of translation. The dynamic nature of C693 is evident by the very weak peak of U692H6-C693H1’ shown in NOESY spectra, and the C693H1’-H2’ peak in DQF-COSY in the QM mutant (Aduri 2007). In the instant evolution study, there is no nucleotide preference at position 693, and some mutants at 693 have greater than wild-type level of function, such as the G693A, A694U double mutant with function of 124%, and the G693U, A694G, A695C triple mutant with function of 112%. Furthermore, the QM mutant (G690A, G693C, A695C, U697A) is more functional than the G690A, 697A double mutant (the AA mutant, 74.1% function) (Morosyuk 2000). According to these results, we hypothesize that nucleotide at 693 is important for the functions of the 690 loop. The dynamic feature of position 693 may facilitate the interactions with tRNA and mRNA during the protein synthesis processes,

and the mutations at this position have been shown to affect protein synthesis. The X-ray crystal structure of the *Thermus thermophilus* 30S subunit (2HGR) shows stacking interaction between G693 and the mRNA (Yusupova, Jenner et al. 2006). Therefore, we hypothesize that the QM mutant has higher function than the AA double mutant (G690A, 697A) because it is more dynamic and flexible than the wild-type sequence, which weakens the stacking interaction between C693 and mRNA, possibly increase the translocation rate of mRNA at E-site, and thereby causing more protein to be produced, though likely with a higher predicted error rate. The higher error rate may also explain why the site-directed G693U single mutant has high function (122%) in an *in vivo* assay where only two small proteins need to be synthesized (i.e. CAT and GFP), but yet his mutant is not observed in nature where stringent demands for fidelity would be strongly selected for in evolution (Morosyuk 2001).

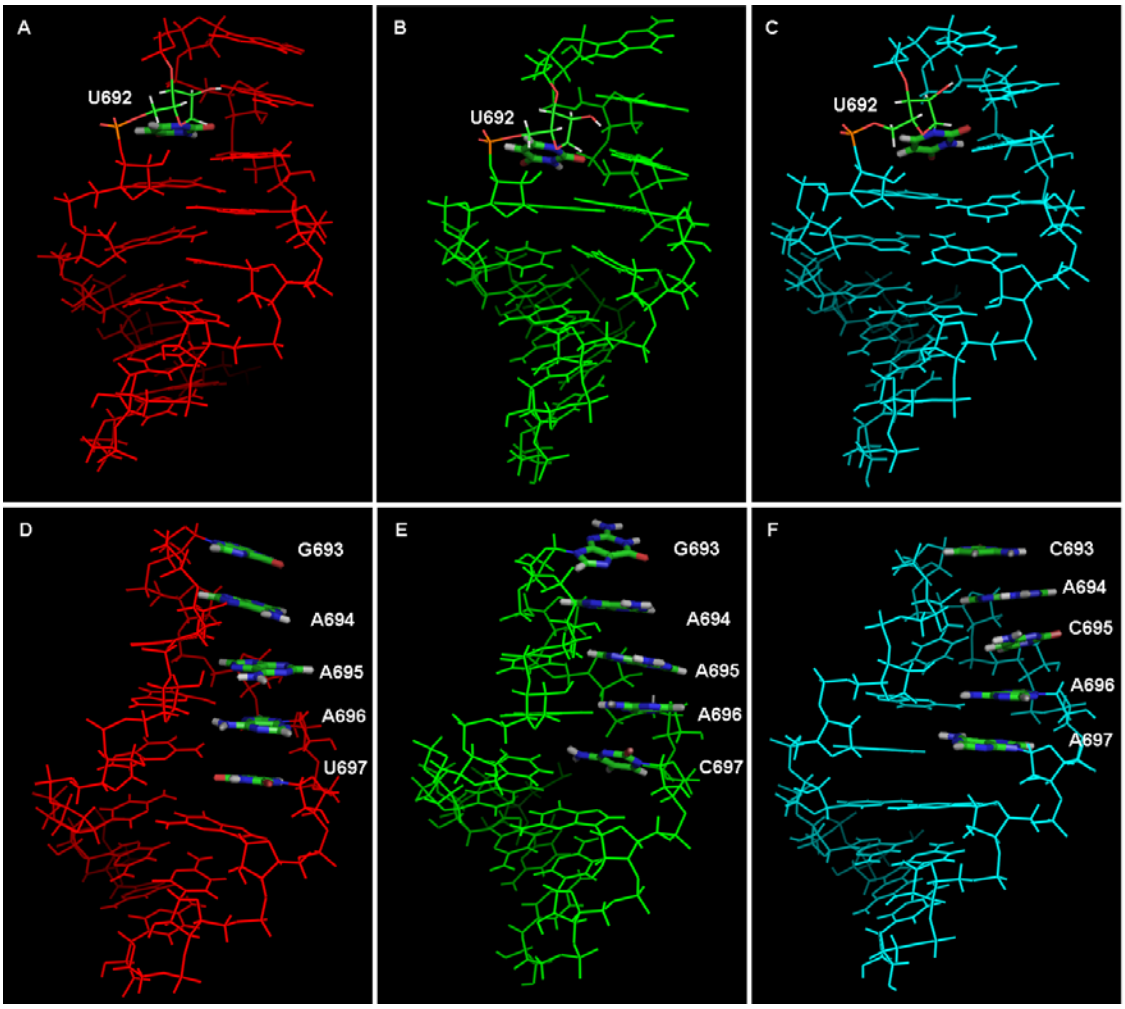


Figure 3.16. The “U-turn” is formed in the structures of the wild-type (A), the UC mutant (B), and the QM mutant (C). The 3’-end continuous stacking of the wild-type (D), the UC mutant (E) and the QM mutant (F) is observed. The C693 does not stack with the A694 as well as the G693 in the wild-type or the UC mutant.

Our NMR structures show that both the UC mutant and the QM mutant structures maintain the key characteristics of the 690 loop and form isomorphous structures as the wild-type. Comparison of the mutant structures with the wild-type allows for the identification of the functional groups in the 690 loop, such as the placements of the hydrogen-bond acceptor atoms at position 697 and stacking of nucleotide 693 are crucial for the functions of the 690 loop. The altered structures of the mutated UC and QM mismatches at 690 and 697 compared to the wild-type GU may provide an explanation for the lower functions than the wild-type due to the shifted mismatches may weaken the interactions between 690 and 697, the “U-turn” is important for maintaining the backbone directions of the loop and also, make 693 available for interacting with other RNA, which positively affects the function of the mutants.

3.6. Conclusions

The NMR structures of two functional mutants, the UC mutant and the QM mutant are isomorphous with the wild-type. The NMR structures of the two functional mutants have been used to identify the key functional groups and structural characteristics for maintaining the function of the 690 loop. The key groups for maintaining the structure and function of the 690 loop include the formation of a “U-turn” at position 692, mismatch pairs involving residues 691-696 and 690-697, the placement of a hydrogen bond acceptor atom in the minor groove at 697, and continuous stacking of five bases at the 3’-terminus of the hairpin.

CHAPTER 4

NMR Study for the 970 Hairpins with and without Modified Nucleotides of *E. coli* 16S rRNA

4.1. Introduction of the 970 Loop

Post-transcriptional modifications are common in natural RNA in all organisms, and a large amount of modifications have been found to exist in natural RNA (Crain 1997) (Czerwoniec 2009). Modifications have been related to some important functions of cells. For example, modifications in tRNA of bacterial thermophiles and in archaeal thermophiles show different effect in stability of tRNA structures (Watanabe 1976; Edmonds 1991). And removal of such modifications have been shown to affect cell functions, such as cell growth rate, ribosome interaction with some amino acids, and antibacterial resistance ability (Piekna-Przybylska 2008; Liang 2009).

11 and 24 modified nucleotides have been observed in 16S rRNA and 23S rRNA of *E. coli* ribosomal RNA. Among the 35 modifications, more than half are methylated modification in base (mN) or ribose (Nm). Another modification popular in ribosomal RNA is pseudouridine (ψ) (Piekna-Przybylska 2008) (Desaulniers 2008). Dihydrouridine is also observed in 23S rRNA. The location of the modifications in *E. coli* ribosomal RNA is shown in Table 4.1 [http://people.biochem.umass.edu/fournierlab/3dmodmap/ecseqframes.php].

Table 4.1. Location of modifications in *E. coli* ribosomal RNA.

16S	
Location	**Modification**
516	Ψ
527	m^7G
966	m^2G
967	m^5C
1207	m^2G
1402	m^4Cm
1407	m^5C
1498	m^3U
1516	m^2G
1518	$m^6{}_2A$
1519	$m^6{}_2A$

23S	
Location	**Modification**
745	m^1G
746	Ψ
747	m^5U (T)
955	Ψ
1618	m^6A
1835	m^2G
1911	Ψ
1915	$m^3Ψ$
1917	Ψ
1939	m^5U (T)
1962	m^5C
2030	m^6A
2069	m^7G
2251	Gm
2445	m^2G
2449	D
2457	Ψ
2498	Cm
2501	C^{*} (a)
2503	m^2A
2504	Ψ
2552	Um
2580	Ψ
2605	Ψ

C^{*} is a substoichiometric modification

The distribution of the modifications in the 3D structure of the ribosome is shown in Figure 4.1 (Decatur 2002). The modified nucleotides are shown in full atomic volume (van der Waals radii), and the pseudouridine (ψ), the 2'-O-methylated nucleotides, and the base methylated nucleotides are shown in red, green, and orange, respectively. The distributions of 11 modifications in *E. coli* modeled in the crystal structure of the small subunit of *T. thermophilus* ribosome (PDB: 1fjf) (Figure 4.1, A). The distributions of the modifications in yeast are shown in the structure of large subunit of *H. marismortui* ribosome (PDB: 1ffk and 1ffz) (Figure 4.1, B). From the 3D structure of the ribosome, it is apparent that many modified nucleotides are located around the functional centers (as shown in Figure 4.1) (Decatur 2002). This suggests that the modifications are functionally important.

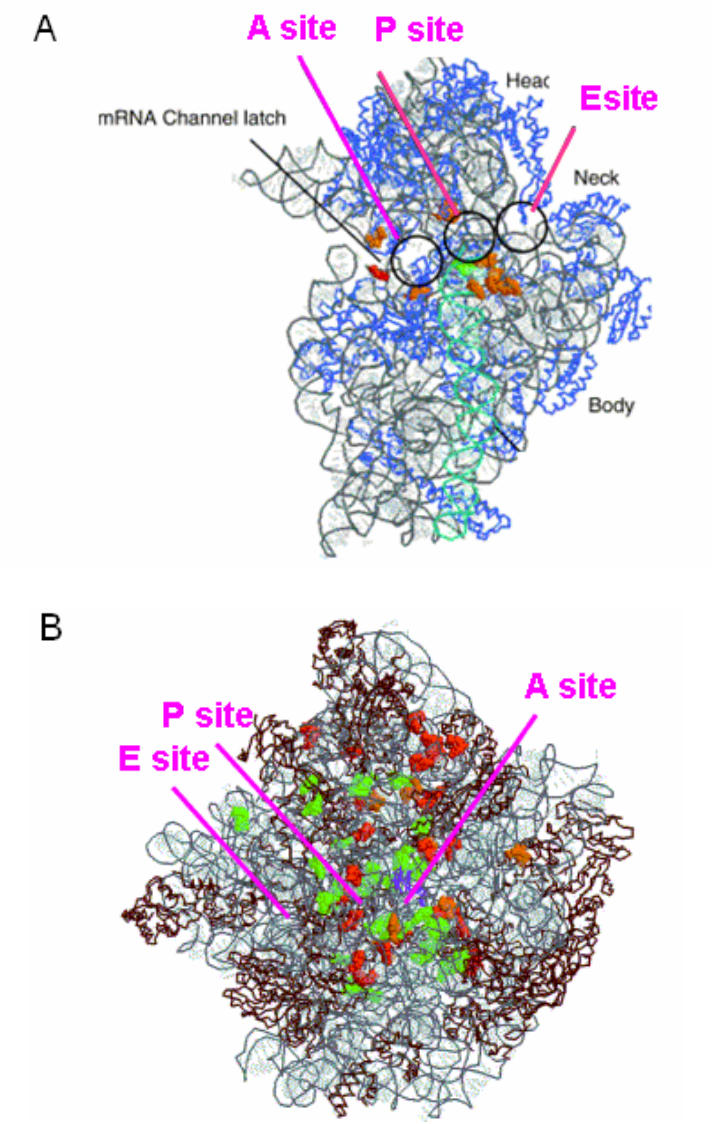


Figure 4.1. The models of the small and large subunits of ribosome with modifications. The distributions of *E. coli* modifications modeled in the crystal structure of SSU of *T. thermophilus* ribosome (PDB: 1fjf) (A). The distributions of the modifications in yeast modeled in the structure of LSU of *H. marismortui* ribosome (PDB: 1ffk and 1ffz) (B). The pseudouridine (ψ), the 2'-O-methylated nucleotides, and the base methylated nucleotides are shown in red triangles, green circles, and orange squares, respectively. (Decatur 2002).

The 970 loop of *E. coli* is located at the central domain of the secondary structure of 16S rRNA (Figure 4.2) (Decatur 2002). And it contains two modified nucleotides, m^2G966 and m^5C967. The sequence of the 970 loop are highly conserved, residue 964, 969 and 970 are conserved among the three life domains, other residues in the 970 loop are conserved within each domain, and modified nucleotides are found at the position 966 in the three life domains (Cannone 2002). In addition, the sequences of the 970 loop in *E. coli* and in *H. Sapiens* are different from each other (shown in Figure 4.3) (Cannone 2002), which make the 970 loop a potential target for anti-bacterials. The 970 loop has been found to be one of the primary interaction sites of tetracycline (Noah 1999).

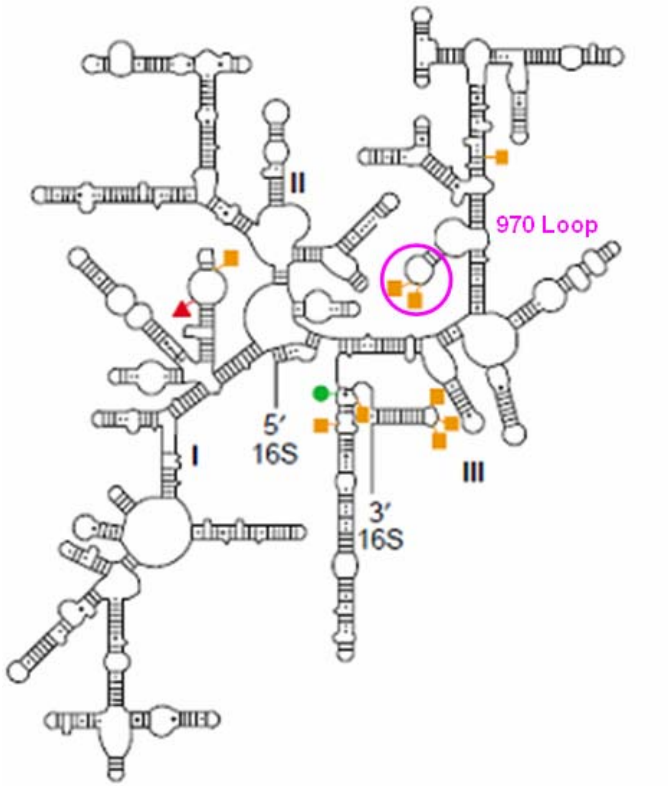


Figure 4.2. Distribution of *E. coli* rRNA modifications in the ribosome. The 970 loop is circled. Pseudouridines and modified pseudouridines, 2'-O-methylations, and base methylations are highlighted in red triangles, green circles, and orange squares (Decatur 2002).

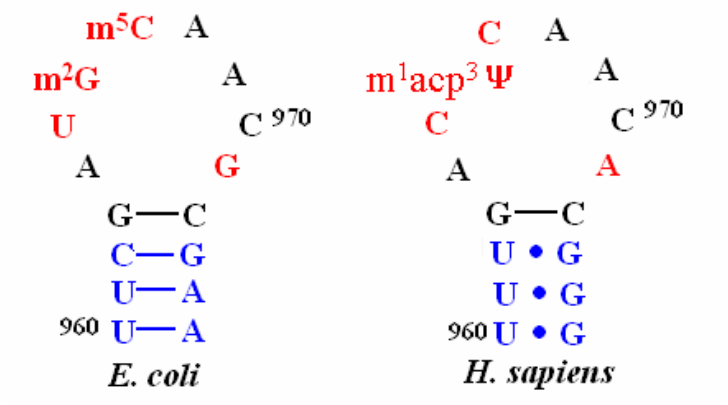


Figure 4.3. The secondary structures of the 970 loop in *E. coli* and in *H. Sapiens*. The residues in red are the residues that are different in *E. coli* and in *H. sapiens*.

The two modified residues in the 970 loop are related closely to some biological functions of the 970 loop. The crystal structure of ribosome shows that the base of m^2G966 of the 970 loop may interact with the ribose of C34 of P-site bounded tRNA (shown in Figure 4.4), suggesting that the 970 loop may be involved in the decoding process (Selmer 2006). The modified cytosine at position 967 is close to reside 1400 of the decoding region in ribosome (Noah 1999). And the methyl groups in the modified nucleotides increase the interaction surface area between the 970 loop and small subunit protein S9, which may stabilize the binding between the 970 loop and S9 through hydrophobic interactions (Aduri 2007). Dr. Chow's lab focused on the thermodynamic stability of the 970 loop with and without modifications, and from the UV-melting experimental data, the modifications slightly destabilize the stability of 970 loop (Abeydeera 2009). Cunningham's lab studied the level of mutant functions of the 970 loop, and they found that single mutation at the positions 966 and 967 have no obvious effect on ribosome function, however, deletion of m^5C causes lethal phenotype (Saraiya 2008). In addition, no statistically significant nucleotide covariations are detected among the residues in the hairpin, suggesting that little intra-hairpin H-bonding is present in the loop (Saraiya 2008). The published crystal structures (2AVY and 1J5E) of the 970 loop in *E. coli* and in *T. thermophilus* have some structural differences, indicating that the 970 loop hairpin can adopt multiple functionally important conformations in different organisms. The structural role of the modifications, however, is not obvious, and therefore, the NMR structures with and without modifications are

obtained to explore the functional effect of the chemical modifications in the structures of the 970 loop. The NMR structures of the 970 loop with and without modifications are useful to reveal the key functional groups that maintain the 970 loop functions. As mentioned earlier in the introduction, the sequence differences of the 970 loop in *E. coli* and in *H. sapiens* make the 970 loop a potential target for anti-bacterials, the primary information of the interaction sites between the 970 loop and the 7mer-peptide that is discovered to be able to bind with the 970 loop by phage display method in Dr. Cunningham's and Dr. Chow's labs, are provided based on the change of NOESY spectra between the 970 loop alone and the complex of the 970 loop and the 7mer-peptide. The molecular interaction information that is useful for development of new classes of anti-bacterials.

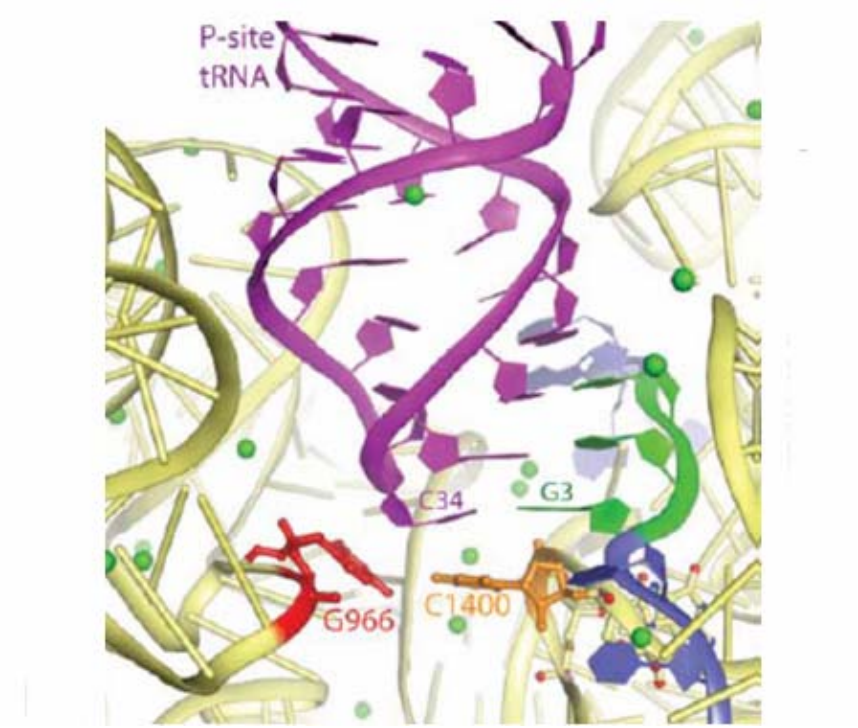


Figure 4.4. The stacking interaction between anticodon loop of the P-site bound tRNA (in purple), m^5C1400 (in orange) and m^2G966 (in red) (Selmer 2006).

4.2. Sample Preparation

The NMR constructs of the 970 loop without modifications (unmodified-970) and the 970 loop with modifications (modified-970) are shown in Figure 4.5. The three G-C base-pairs (shown in blue) at the stem ending are added to stabilize the sequences.

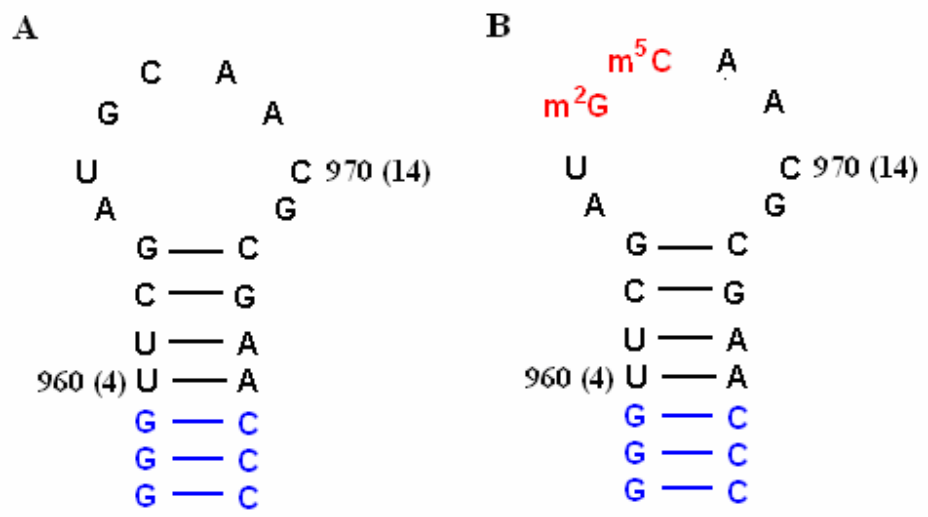


Figure 4.5. The secondary structure of the unmodified RNA of the 970 loop (A) and the modified RNA of the 970 loop (B). The nucleotides shown in red are the two modified nucleotides in the wild-type 970 loop in *E. coli* ribosomal RNA. The three base-pairs at the stem ending are added to stabilize the hairpin structures of the two samples (shown in blue).

The RNA sample of the 970 loop without chemical modifications (unmodified-970) was synthesized by using bacteriophage T7 RNA polymerase catalyzed *in-vitro* transcription (Milligan 1987). The house-made T7 RNA polymerase was used in the transcription reactions. The T7 polymerase enzyme was prepared by over-expression in *E.coli* and then purified by Ni^{2+}-column and dialysis after separation from RNA or DNA of *E.coli*. In the *in vitro* transcription reactions, the starting chemicals that are used for RNA transcription include DNA templates, DNA primers, T7 polymerase enzyme, Mg^{2+}, DTT (reducing agents to stabilize the T7 polymerase enzyme), Triton X-100 (detergents), polyethylene glycol, spermidine, and NTPs. Before the transcription reactions, the primer and template were annealed beforehand (the annealing process is heat the mixture of the templates and primers at 85 ºC for 2 minutes, then immediately cooled by ice). And in some cases, inorganic pyrophosphatase is used to improve synthetic yields by hydrolysis of pyrophosphates (inorganic pyrophosphates inhibit the polymerase reaction), but the pyrophosphatase didn't show obvious improvement in the transcription yields of the unmodified-970. Mg^{2+} concentrations and T7 amounts needed to be optimized by a series of small scale transcription reactions before carrying on large scale of transcription reactions (data not shown). And the amount of produced RNA reached maximum with magnesium concentration at 30 mM among the transcription reactions under the same conditions except the concentrations of magnesium were 24, 26, 28, 30, 32 and 34 mM. Therefore, the magnesium concentration chosen for running large-scale of transcription reactions is 30mM. And the amount of T7 RNA polymerase was also tested by

adding different volumes of the T7 RNA polymerase, 1, 2, 3, 4, 5, 6, 8, 10 μL into 25μL-scale transcription reactions, and product amount didn't increase significantly when the volumes of the T7 RNA polymerase are larger than 2 μL, therefore, 2 μL of T7 RNA polymerase is chosen for the 25 μL-scale transcription reactions. After transcription, the unmodified-970 RNA was separated and purified from other transcription products and impurities by using 20% denaturing PAGE, dialysis, and G-10 Sephadex column (as shown in Figure 4.6, A). To separate the needed unmodified-970 RNA from the other side products from the transcription reactions, it is suggestive to avoid the conditions that the gels are heavily loaded and to allow the polyacryamide gel running as long as possible until the RNA products with the desired sequences almost arrives the edge of the gel. After loading all the transcribed products in the 20% denaturing polyacrylamide gel (PAGE), the gel band with the correct RNA sequences that are designed for NMR study were cut from the gel and electroeluted from the gel strips. Dialysis were used to remove impurities, dialysis is executed for one day by using sodium chloride and EDTA as dialysis buffer and another day by using double deionized water. After dialysis, further desalting needed by using size exclusion chromatography (G-10 column) or reverse-phase chromatography, Sep-pak Column Chromatography. The MALDI-MS (Matrix Assisted Laser Desorption/ionization-Mass Spectroscopy) method was used to measure the molecular mass of the obtained RNA purification, and the molecular mass of the purified RNA samples are 7135.1, which is very close to the molecular mass of the designed unmodified-970 RNA (7112.1, calculated by HyTher software

(ozone3.chem.wayne.edu)) with the consideration of the unmodified-970 RNA may bind with sodium ion.

The unlabelled modified-970 RNA was ordered from Yale University with the use of the phosphoramidite that was provided by Dr. Chow's lab. The protecting group 2'-TOM of the ordered sample was removed according to the steps shown in Dr. Nuwan Abeydeera's thesis (Abeydeera 2009). The modified-970 RNA was then purified by using reverse-phase chromatography, Poly-pak Column Chromatography, High Performance Liquid Chromatography (HPLC), 20% denaturing PAGE, Sep-pak Column Chromatography, and dialysis against ddH_2O. The procedure of preparing the modified-970 RNA suitable for NMR study is shown in Figure 4.6, B.

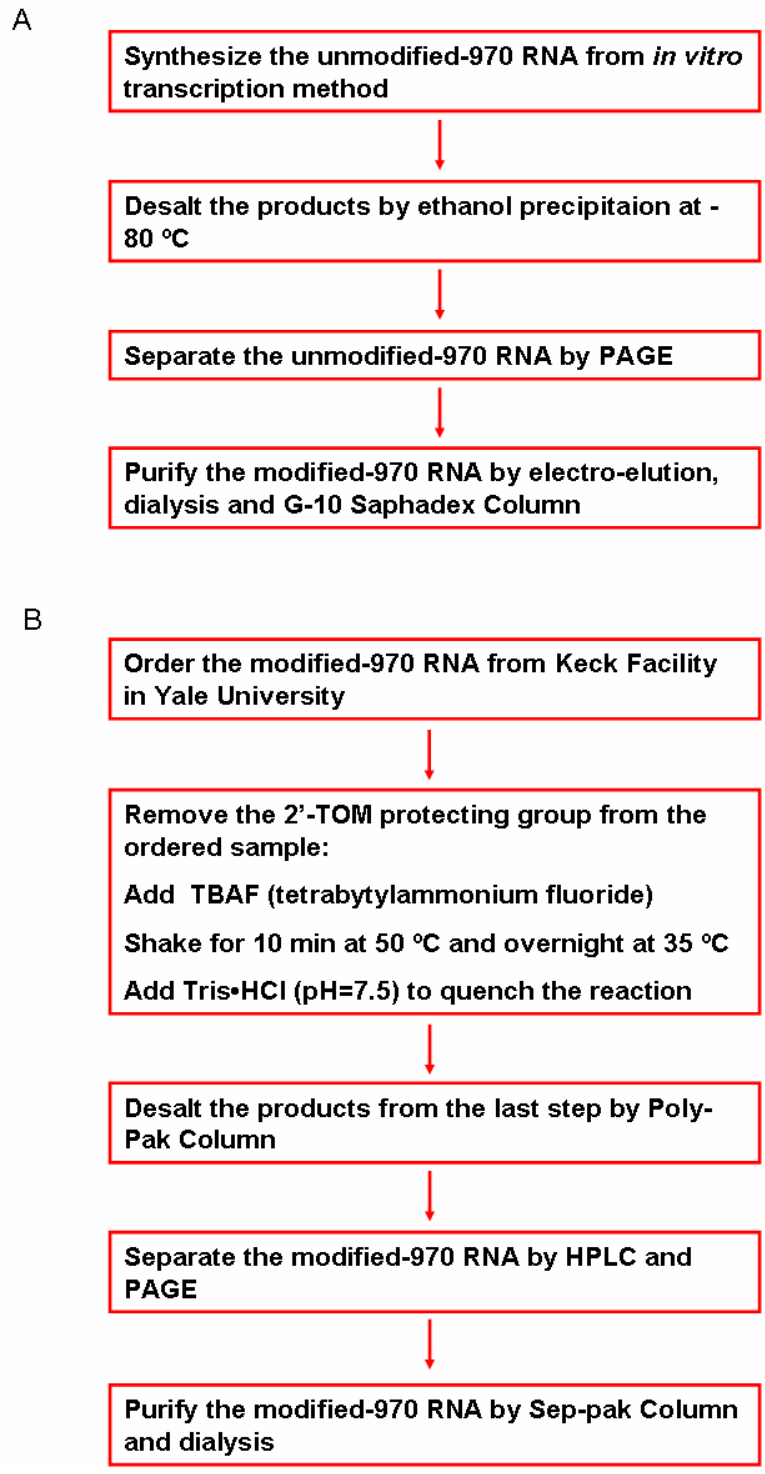


Figure 4.6. The preparation steps for the unmodified-970 RNA (A) and the modified-970 RNA (B).

The MALDI-MS spectrum confirmed the molecular mass of the purified modified-970 RNA is 7082.8, matching the designed sequence of the modified-970 RNA, which indicates that the obtained RNA sample contain two modified residues. The peaks from the 2D-NOESY spectrum at 37 ºC (shown in Figure 4.7), including G10H_{32}-C11H6, G10H_{31}-C11H1', G10H_{33}-H1', C11H_{13}-G10H8, C11H_{12}- A12H2, C11H_{13}-G10H2' and C10H_{13}-G10H3' can confirm the two methyl groups in the modified-970 RNA are located at the right positions.

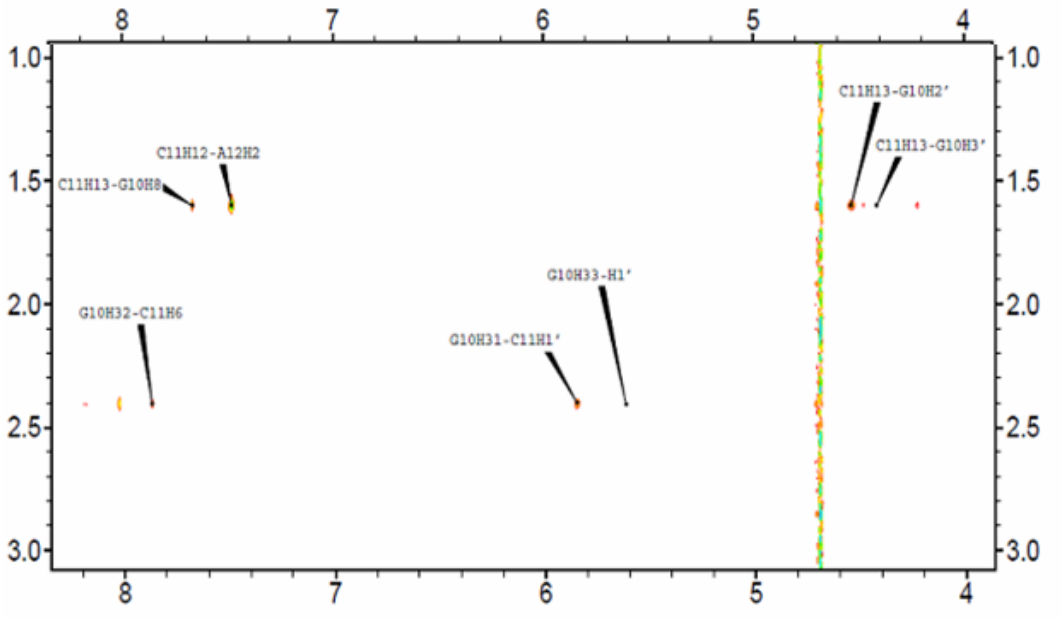


Figure 4.7. The methyl proton region of the NOESY spectrum of the modified-970 RNA. The labeled peaks are G10H_{32}-C11H6, G10H_{31}-C11H1', G10H_{33}-H1', C11H_{13}-G10H8, C11H_{12}-A12H2, C11H_{13}-G10H2' and C10H_{13}-G10H3'. These peaks support that the methyl groups are located properly in the N2-methyl guanine and 5-methyl cytosine of the modified-970 RNA.

4.3. Experimental Results of the Unmodified-970 RNA

The UV-melting experiments provide information about secondary structures of RNA. For hairpin-RNA, as the temperature increases, the structure becomes denatured to random-coil states, which hence increase the UV absorbance. For hairpin-RNA, melting temperatures are independent of the total strand concentration (SantaLucia 2000). While the melting temperatures vary with concentrations if bimolecular or higher order conformations of RNA are present in solution (SantaLucia 2000). Therefore, the melting temperatures of RNA sample of different concentrations can be used to identify whether the predominant structure of a RNA sample is hairpin formation or other molecular structures, such as duplex formation.

From the UV-melting curves of the unmodified-970 RNA (Figure 4.8), the melting temperatures of the unmodified-970 RNA are between 70 ºC and 72 ºC when unmodified-970 RNA are dissolved in the NMR buffer with the concentrations at 0.5 μM, 1.0 μM, 1.5 μM, 2.5 μM and 5.0μM. This indicates that the melting temperatures of the unmodified-970 RNA don’t vary with concentrations in the range of 0.5 ~ 5.0 μM, and the unmodified-970 RNA may exist in a hairpin structure in NMR buffer.

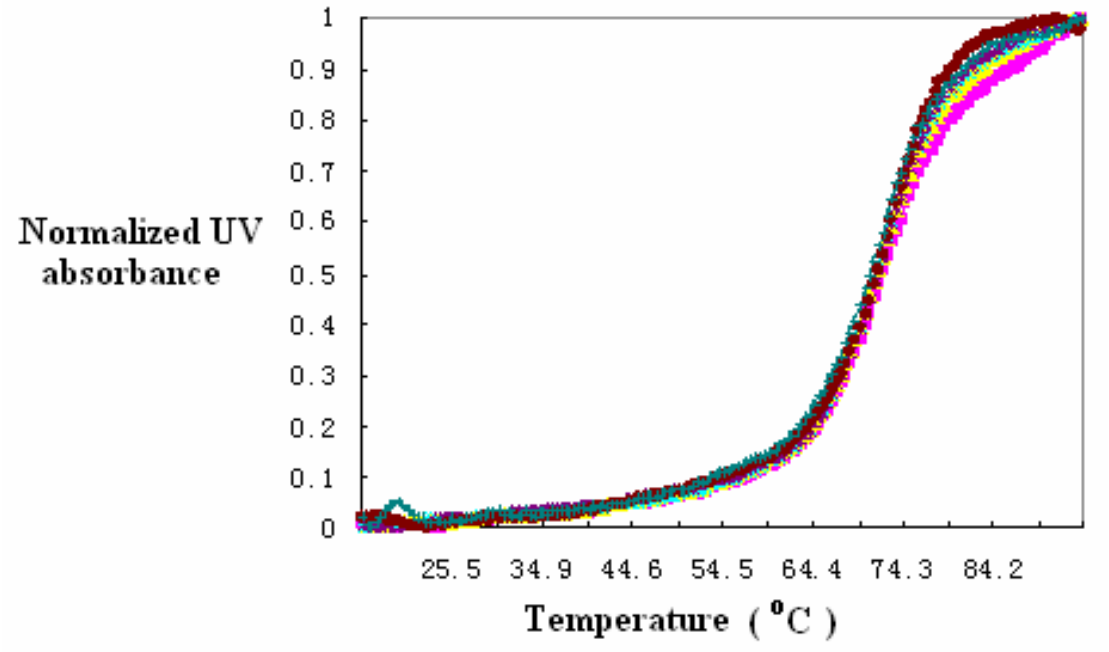


Figure 4.8. The normalized UV-melting curves of the unmodified 970 loop at different concentrations in NMR buffer, the sample concentrations change from 0.5 to 5 μM.

The unmodified-970 RNA was dissolved into 0.3 mL NMR buffer and transferred into a Shighemi microvolume NMR tube. 1D-^{1}H NMR data in D_2O for the unmodified-970 RNA at different temperatures is shown in Figure 4.9. At 10 ºC, 15 ºC, 20 ºC, 25 ºC, 30 ºC, 35 ºC, 37 ºC, 40 ºC, 42 ºC and 45 ºC, the peaks in the spectra at different temperatures are located at similar positions, which indicate that the structures of the unmodified-970 RNA are stable at these various temperatures (Figure 4.9). However, resolution of some peaks, especially peaks in the 5-8 ppm, is better at higher temperatures, but no significant improvements in peak resolution can be observed when the temperatures are higher than 37 °C. Therefore, 37 °C is chosen for later NMR experiments for better peak resolution and since it is the physiological temperature for biological macromolecules. And later NMR experiments for the unmodified-970 RNA were all collected at 37 ºC.

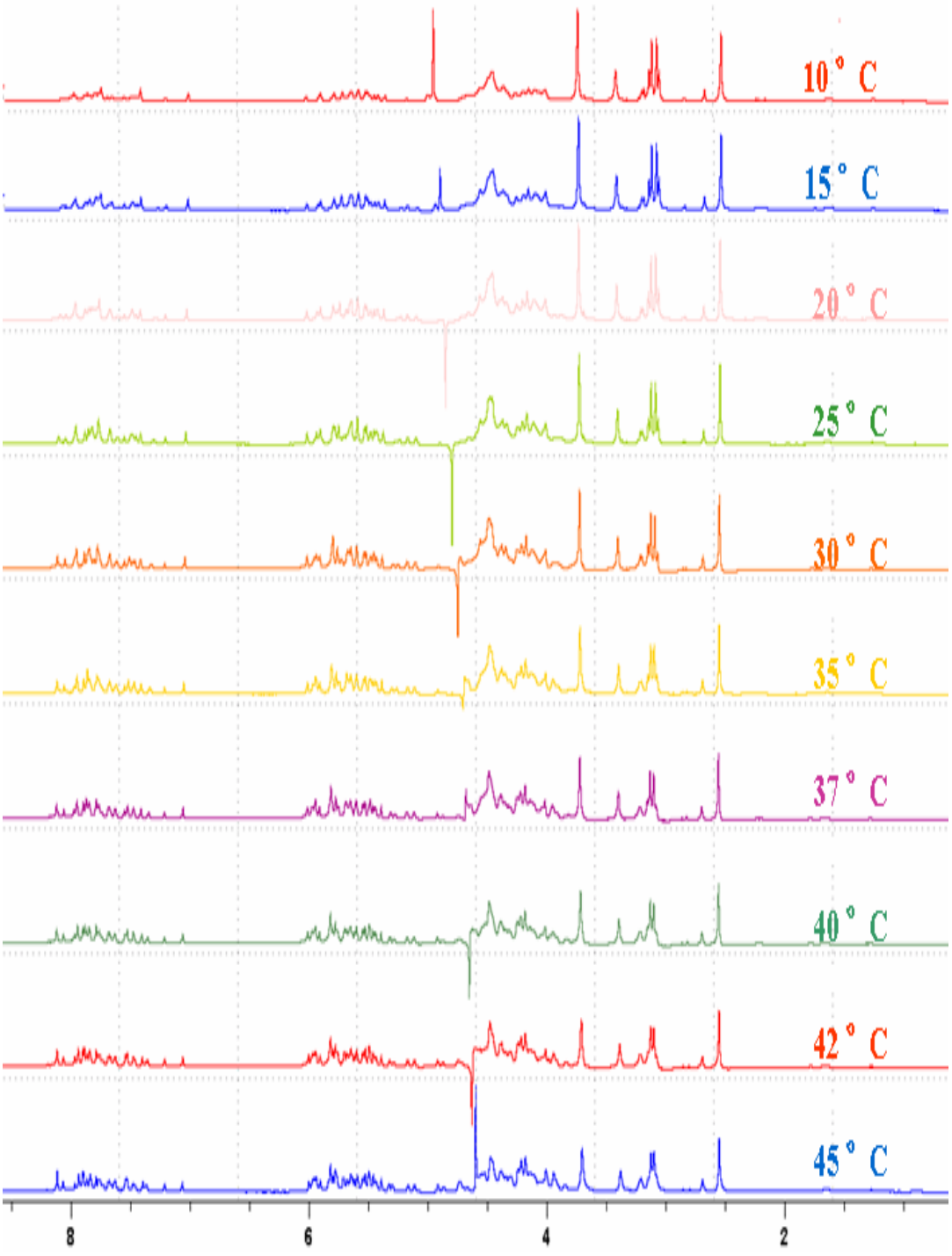


Figure 4.9. The 1D-1H spectra of the unmodified-970 RNA in D_2O at 10 ºC, 15 ºC, 20 ºC, 25 ºC, 30 ºC, 35 ºC, 37 ºC, 40 ºC, 42 ºC and 45 ºC.

.

The 1D-^{1}H NMR spectrum in H_2O of the unmodified-970 RNA at 37 ºC is shown in Figure 4.10. The hydrogen-bonded imino protons have higher chemical shifts than those non hydrogen-bonded imino protons or other base protons or sugar protons in RNA, and the hydrogen-bonded imino protons are usually located in the region between 11 to 15 ppm. In the 11~15ppm region of the spectrum, 6 sharp peaks are observed, which indicates there are 6 base pairs formed in the structure of the unmodified-970 RNA. The imino protons of the G residue at the 5'-end or 3'-end of RNA are usually not apparent or with low resolution in the 1D-^{1}H NMR spectrum, the number of hydrogen-bonded imino peaks shown in Figure 4.10 matches the secondary structure of the unmodified-970 RNA shown in Figure 4.3. Although the hydrogen-bonded imino proton peaks are with high resolution in the 1D-^{1}H NMR spectrum (Figure 4.10), not enough information provided in the 1D-^{1}H NMR spectrum for assigning the imino proton peaks. The assignments of the imino peaks of the unmodified-970 RNA are based on the cross-peaks between the imino protons and amino protons or base protons in the 2D NOESY spectrum in H_2O of the unmodified-970 RNA (Figure 4.11).

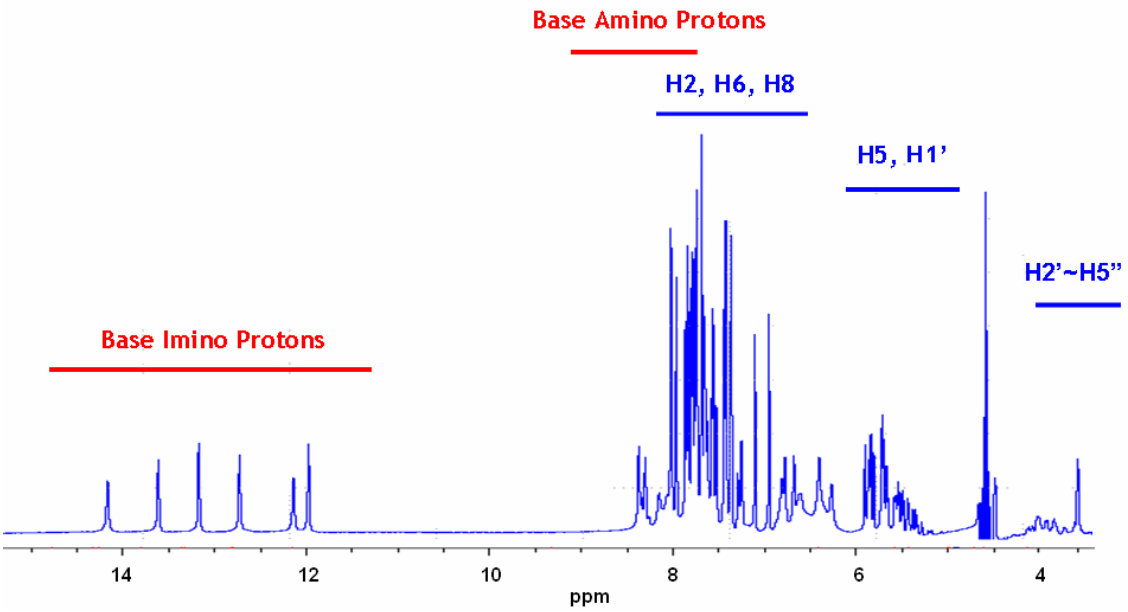


Figure 4.10. The ^{1}D-^{1}H spectra of the unmodified-970 RNA in H_2O at 37 ºC. The imino and amino protons are exchangeable protons (labeled in red), and non-exchangeable protons are labeled in blue. There are 6 sharp imino protons located between 11 and 15 ppm, which indicates 6 hydrogen-bonded imino protons are formed in the structure of the unmodified-970 RNA, and these hydrogen-bonded imino protons are shown in the 1D-^{1}H spectrum with high resolution.

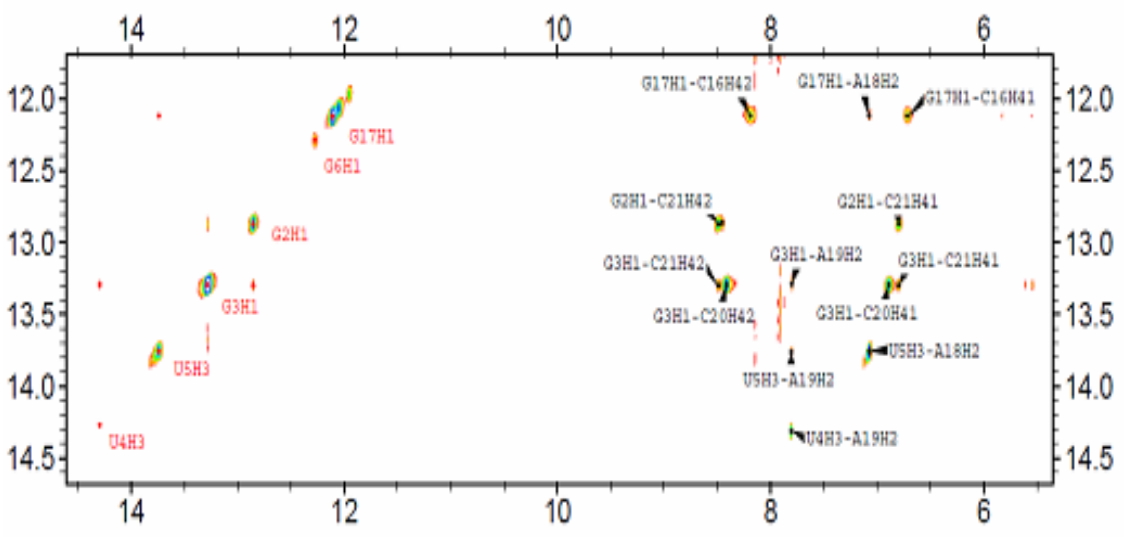


Figure 4.11. The imino region of the NOESY spectra of the unmod-970 loop in H_2O at 37 ºC. All the hydrogen-bonded imino protons except G1H1 are assigned, mainly based on the cross-peaks between imino and amino protons in the base pairs.

The cross-peaks shown in NOESY represent for two atomic resonances that are closer less than 5.0 Å in space (Fürtig 2003). And the H1', H2, H5, H6 and H8 of the unmod-970 RNA are assigned in the "NOE-walk" region of the unmodified-970 RNA. Figure 4.12 shows the fingerprint region of the NOESY of the unmodified-970 RNA. All the cross-peaks between H1' and H6/H8 of the residues in the unmodified-970 RNA that are involved in the sequential walk are connected by red arrows. The NOE walk is complete from the 5'-end to the 3'-end of the unmodified-970 RNA, which indicates that the secondary structure of the unmodified-970 RNA is as expected in Figure 4.3, A. The H5-H6 peaks of the pyrimidines are the strongest peaks and are assigned in green. Other cross-peaks are labeled in purple, which contain important information for restraining the unmodified 970 loop in calculations. And the blue labeled strong peak, A8H2-U9H1', indicates the stacking formed between A8 and U9.

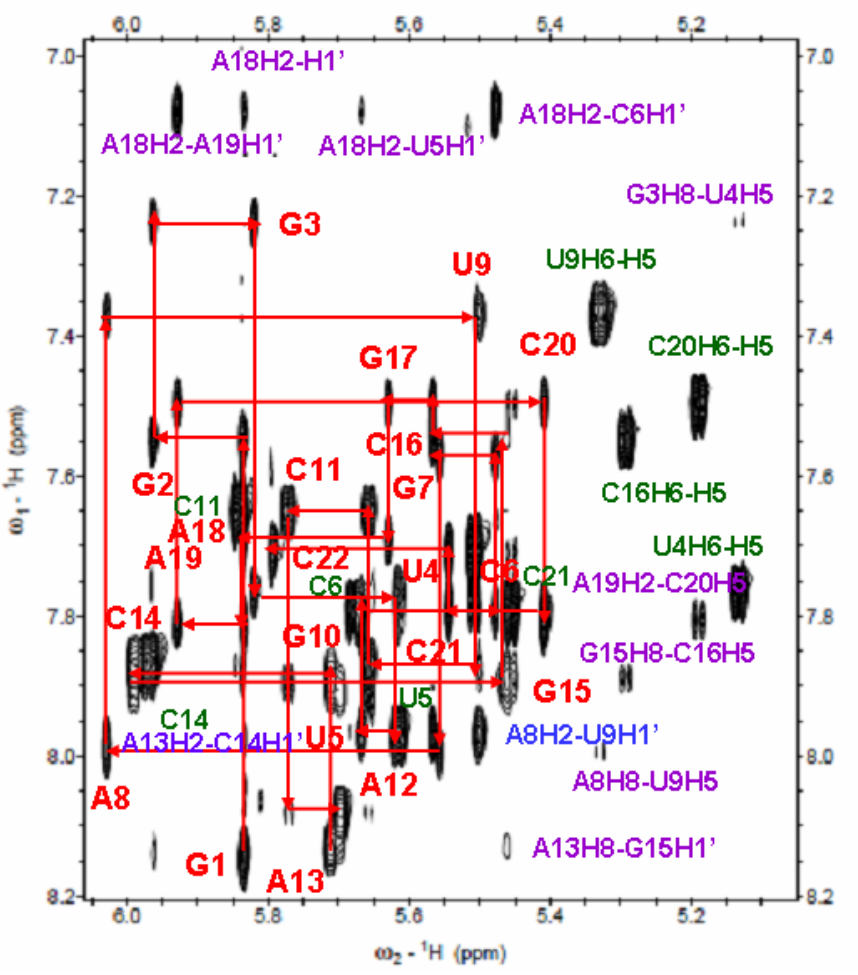


Figure 4.12. The sequential walk region of the NOESY spectrum of the unmodified-970 loop. The sequential walk between residues is collected by red arrows. And the "NOE-walk" is complete from G1 to C22, which indicates the secondary structure of the unmodified-970 RNA is as designed in the NMR construct.The H5-H6 peaks of pyrimidines are labeled in green, the H5-H6 peaks are the strongest peaks in the region. The relatively strong peak A8H2-U9H1' is labeled in blue. Other intra-residual or inter-residual cross-peaks are shown in purple.

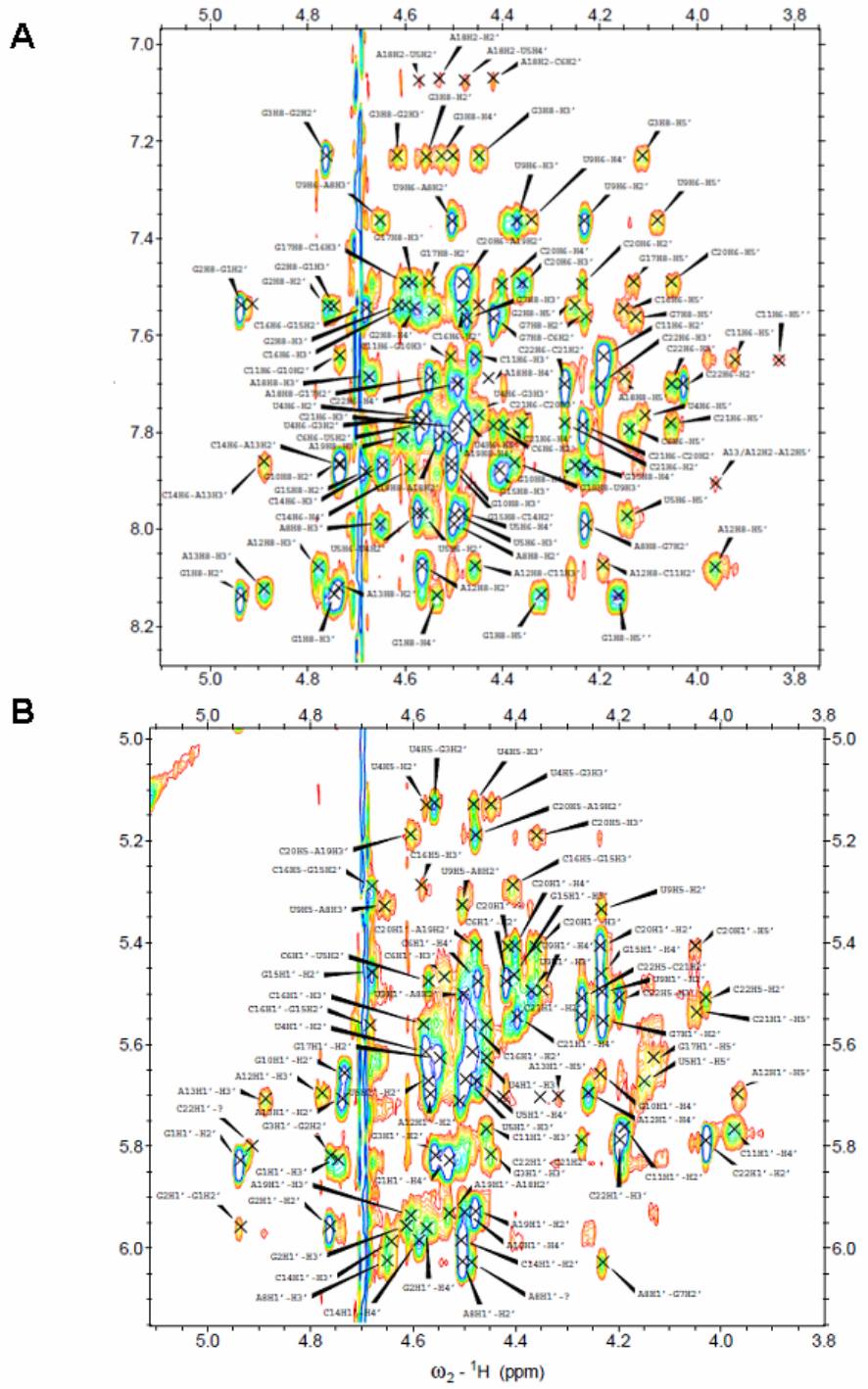


Figure 4.13. The sugar proton regions of the NOESY spectrum of the unmodified-970 loop.

The peak assignments in the “NOE-walk” region of the 2D NOESY spectrum are very useful for obtaining the chemical shifts of H1’, H2, H5, H6 and H8 of the unmodified-970 RNA (Figure 4.12), and the assignments of the peaks in the sugar proton regions of the 2D NOESY spectrum provide chemical shift information for H2’, H3’, H4’ H5’ and H5’’ (Figure 4.13). However, some initial assignments of the unmodified-970 RNA based on the NOESY spectrum need to be confirmed by other spectra because some peaks in the NOESY are overlapped or very close to each other, for example, C20H1’ and C21H5, U4H1’ and U5H5, C5H1’ and U6H5. HMQC spectra are useful for confirming H1’, H2, H5, H6 and H8 assignments (as shown in Figure 4.14). To improve the decoupling efficiency of ^{13}C, three HMQC spectra were collected with different offset frequency for ^{13}C. The chemical shift regions of the three HMQC spectra cover all the C-H peaks in the RNA sample. The separation of H5 and H1’ in HMQC spectrum simplify partial assignments of H1’ and H5 (Figure 4.14, A). The chemical shifts of H1’ from the total 22 residues, and H5 from the 7 cytosines and 3 uracils are confirmed in Figure 4.14, A. Based on the chemical shift range of C6/8 is different from that of C2, all the peaks of C2-H2 are obviously separated from C6/8-H6/8 in Figure 4.14, B. In Figure 4.14, C, C4’-H4’ peak group is different from the C2’/3’-H2’/3’ and C5’/5’’-H5’/5’’ peaks. However, C2’-H2’ can not be separated from C3’-H3’ peaks. Partial sugar assignments can be confirmed in Figure 4.14, C. Most chemical shifts of sugar protons are needed to be confirmed by 3D TOCSY-NOESY spectra. The chemical shifts of ^{13}C and ^{1}H of the unmod-970 RNA are listed in Table 4.2.

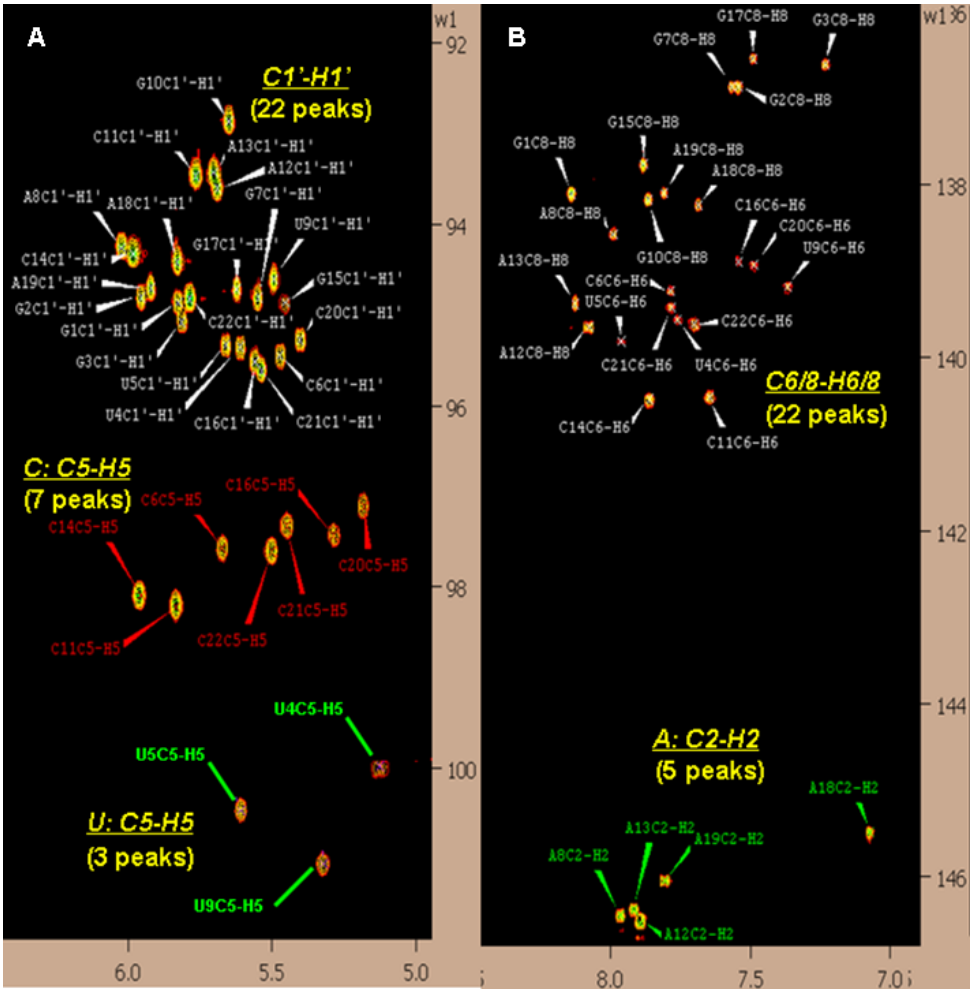


Figure 4.14. The HMQC spectra of the unmodified-970 RNA. The peaks of C1'-H1', C5-H5 of cytosines, and the C5-H5 of uracils are labeled in white, red, and green, respectively (A). The peaks of C6/8-H6/8, C2-H2 of adenosines are labeled in white and green, respectively (B). The peaks of C4'-H4', C2'/3'-H2'/3' and C5'/5''-H5'/5'' are labeled in magenta, cyan and green, respectively (C).

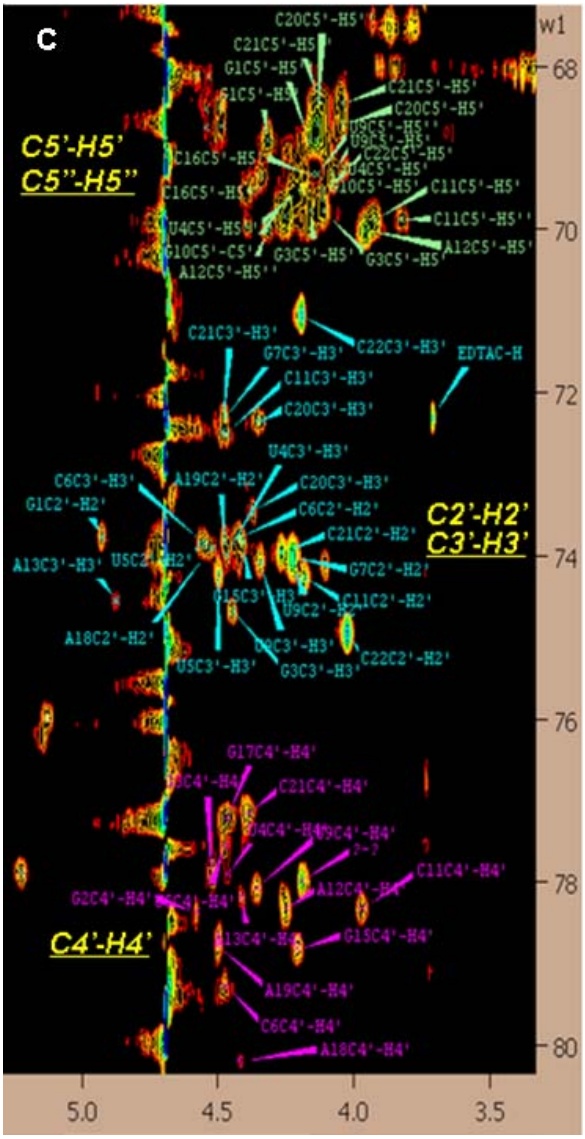


Figure 4.14. The HMQC spectra of the unmodified-970 RNA. The peaks of C1'-H1', C5-H5 of cytosines, and the C5-H5 of uracils are labeled in white, red, and green, respectively (A). The peaks of C6/8-H6/8, C2-H2 of adenosines are labeled in white and green, respectively (B). The peaks of C4'-H4', C2'/3'-H2'/3' and C5'/5''-H5'/5'' are labeled in magenta, cyan and green, respectively (C).

Table 4.2. The chemical shifts of carbon and hydrogen atoms in the unmodified-970 RNA at 37 °C.

	C1'	C2/C5	C6/C8	H1'	H2/H5	H2'	H3'	H4'	H5'/H5''	H6/H8
G1	94.86	N/A	138.1	5.834	N/A	4.938	4.751	4.534	4.327/4.166	8.143
G2	94.81	N/A	136.9	5.963	N/A	4.768	4.615	4.573	4.256	7.544
G3	95.08	N/A	136.6	5.822	N/A	4.558	4.451	4.52	4.116	7.235
U4 (960)	95.35	99.98	139.6	5.619	5.133	4.577	4.485	-	4.112	7.767
U5 (961)	95.33	100.5	139.8	5.672	5.618	4.57	4.497	4.477	4.146	7.97
C6 (962)	95.45	97.57	139.2	5.478	5.681	4.42	4.543	4.475	4.143	7.793
G7 (963)	94.81	N/A	136.9	5.558	N/A	4.233	4.474	4.401	4.128	7.569
A8 (964)	94.25	146.5	138.6	6.031	7.966	4.505	4.653	4.548	4.148	7.991
U9 (965)	94.6	101.1	139.2	5.501	5.332	4.235	4.374	4.347	4.084	7.37
G10 (966)	92.86	N/A	138.2	5.66	N/A	4.726	4.507	4.247	4.125	7.869
C11 (967)	93.46	98.22	140.5	5.771	5.845	4.194	4.463	3.975	3.929/3.836	7.649
A12 (968)	93.6	146.5	139.6	5.699	7.895	4.567	4.779	4.261	3.968/3.943	8.08
A13 (969)	93.42	146.4	139.4	5.711	7.914	4.737	4.89	4.51	4.326/4.166	8.128
C14 (970)	94.32	98.09	140.5	5.989	5.968	4.508	4.647	4.59	4.264/4.258	7.867
G15 (971)	94.86	N/A	137.8	5.462	N/A	4.688	4.407	4.227	3.954	7.883
C16 (972)	95.52	97.43	138.9	5.567	5.292	4.488	4.587	-	4.152	7.548
G17 (973)	94.71	N/A	136.5	5.629	N/A	4.547	4.584	4.465	4.129	7.494
A18 (974)	94.41	145.5	138.2	5.834	7.077	4.531	4.673	4.429	4.151	7.69
A19 (975)	94.7	146.1	138.1	5.93	7.809	4.481	4.603	4.502	4.141	7.813
C20	95.27	97.11	138.9	5.412	5.194	4.238	4.365	4.403	4.053	7.496
C21	95.62	97.33	139.4	5.546	5.455	4.273	4.494	4.396	4.052	7.79
C22	94.83	97.61	139.6	5.793	5.512	4.034	4.202	4.498	4.051	7.705

Some dihedral angel restraints for the unmodified-970 RNA are obtained from DQF-COSY, 1D-^{31}P and HETCOR. In the H5-H6 region of the DQF-COSY (Figure 4.15) spectrum, the chemical shifts of the H5 and H6 protons of all 10 pyrimidines match the assignments of the NOESY spectrum (Figure 4.12) and the HMQC spectra (Figure 4.14). The H1'-H2' region of the DQF-COSY spectrum is very useful to provide information about sugar pucker conformations of the residues in the unmodified-970 RNA (Figure 4.15). The strong peaks in the H1'-H2' region of the DQF-COSY spectrum (Figure 4.15) indicate the sugar pucker conformations predominantly in C2'-endo. Eight peaks show up in the H1'-H2' region of the DQF-COSY spectrum, and five peaks are relatively strong among the eight peaks (labeled in blue), which indicate that the sugar pucker conformations of the five residues G10, C11, A12, A13 and A14 are predominantly C2'-endo. H1'-H2' peaks with medium strength of the residues G1, U9 and C22 are observed, which indicates that the three residues are dynamic with a partial population of C2'-endo conformation.

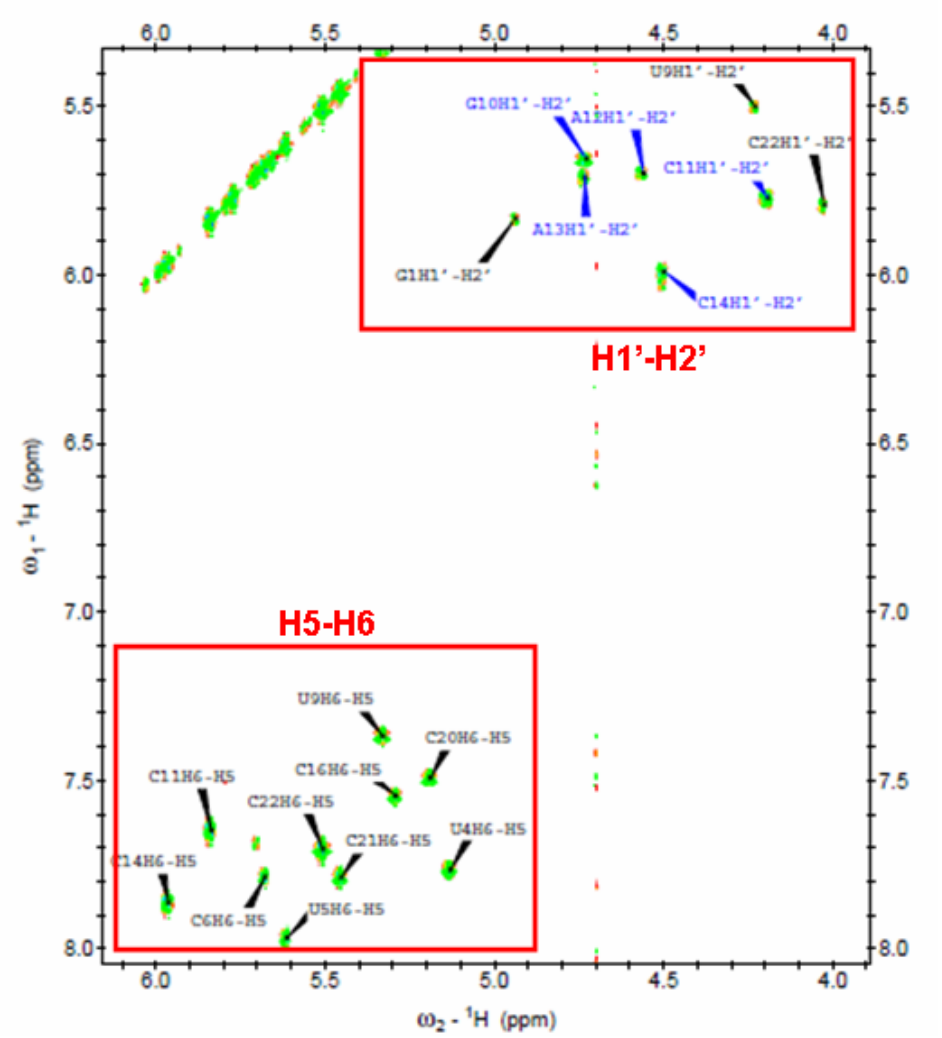


Figure 4.15. H1'-H2' and H5-H6 regions of the DQF-COSY spectrum of the unmodified-970 RNA at 37 ºC. G10H1'-H2', C11H1'-H2', A12H1'-H2', A13H1'-H2' and C14H1'-H2' are relatively strong peaks (labeled in blue), which indicates that the sugar pucker conformations of the residues G10, C11, A12, A13 and C14 are predominantly C2'-endo. The relatively weak peaks, G1H1'-H2', U9H1'-H2' and C22H1'-H2' indicate the sugar pucker conformations of the residues G1, U9 and C22 are dynamic with partial population in C2'-endo.

The 1D-^{31}P spectrum shows that the chemical shifts of ^{31}P of most residues are in the similar range (Figure 4.16, A), only one peak is located apart from most peaks, which indicates that one residue in the unmodified-970 RNA contains ^{31}P with unusual chemical shift. The2D ^{1}H-^{31}P HETCOR spectrum (shown in Figure 4.16, B) identifies the phosphorous with unusual chemical shift is from the residue G10 (the difference between the phosphorous chemical shift of G10 and other residues is around 1ppm). This indicates that the α dihedral angle of G10 and/or the ζ dihedral angle of U9 are possibly different from other residues, and thus, the two dihedral angles are left unrestrained in modeling computations of the unmodified-970 RNA.

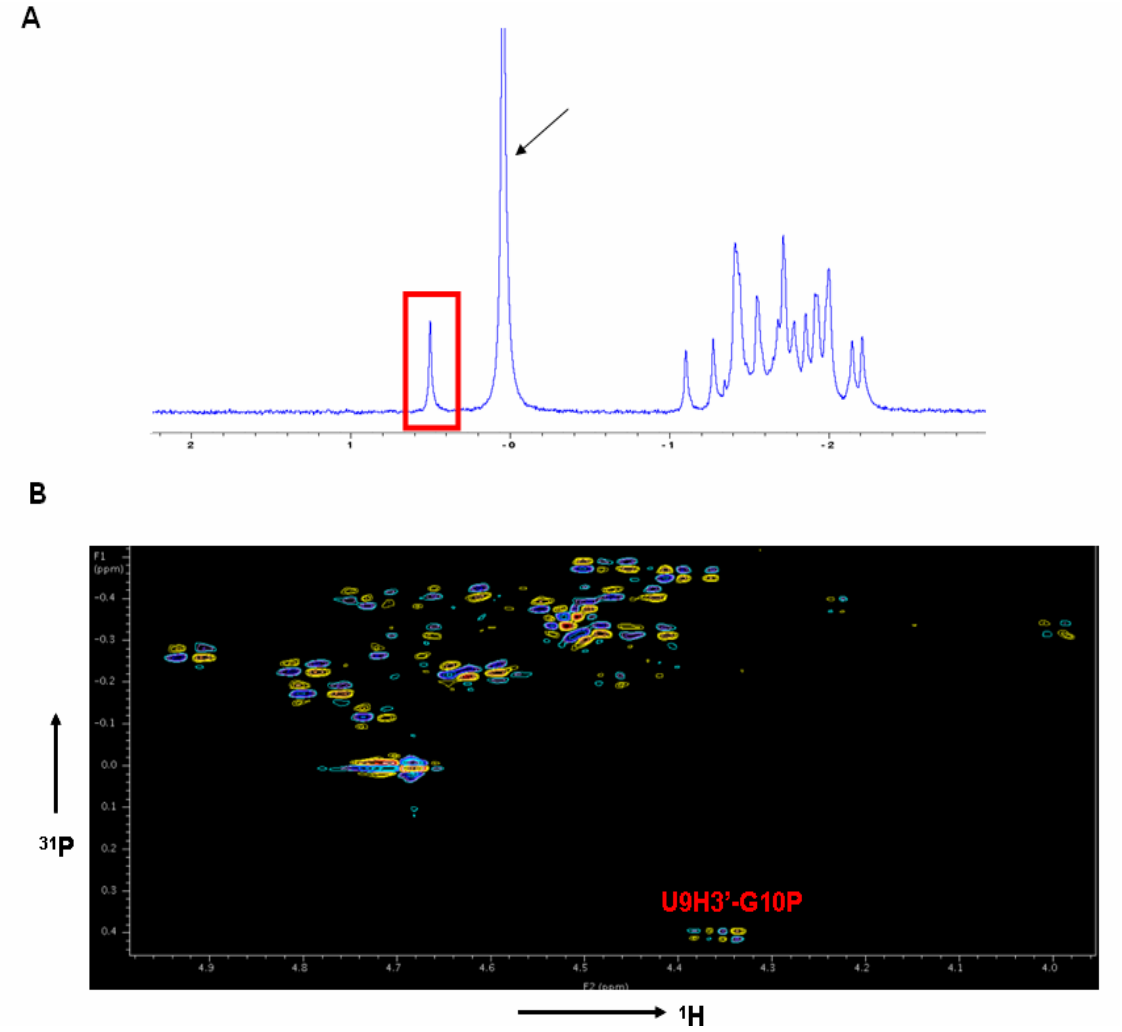


Figure 4.16. 1D-^{31}P (A) and 2D HETCOR (B) spectra of the unmodified-970 RNA at 37 ºC. The G10P has unusual 31P-chemical shift, indicating that the dihedral angel α of G10 and ζ of U9 are different from other residues, and the two dihedral angels are unrestrained in calculations.

4.4. Calculation Results of the Unmodified-970 RNA

Current modeling is based upon a total of 333 distance restraints, 127 dihedral angle restraints, and 94 unobserved-NOEs. One refined NMR structure of the unmodified-970 RNA is shown in Figure 4.17, A. The crystal structure of the 970 loop from the 70S *E. coli* ribosome is shown in Figure 4.17, B (PDB: 2AVY). The two structures in Figure 4.17 share some common features, for example, the main interactions in the 970 loop of the two structures are base stacking, and the residues interact by base stacking in the two structures are shown in the same color. The NMR structure of the unmodified 970 loop shares some characteristics as the crystal structure (Figure 4.17), such as base stacking formed between A964 (8) and U965 (9) (labeled in red), between A969 (13) and C970 (14) and triple-base stacking is observed among G 966 (10), C 967 (11), and A 968 (12) (in cyan). Base stacking is the main interaction in both structures of the 970 loop match the mutation study result obtained in Dr. Cunningham's lab that no significant covariations exist in the 970 loop (Saraiya 2008). However, the NMR structure and the crystal structure have some differences. A964 and U965 (in red) in the NMR structure interact by parallel base stacking, while the base surfaces of the two residues are perpendicular to each other in the crystal structure. Similarly, the stacking formats of A969 and C970 in the two structures are in parallel and perpendicular, respectively. And the triple-base stacking interactions among G966, C967 and A968 (in cyan) are not the same in the two structures. In the NMR structure, G966 forms perpendicular stacking with the other two residues, and base surfaces of C967 and A968 are parallel to each

other. In the crystal structure, the base surfaces of the three residues, 966, 967 and 968 are parallel. The most significant difference between the two structures are the placements of G971 (15) (labeled in green), G971 flips out the loop in the crystal structure, no interactions of the base of G971 with other residues in the loop can be observed in the crystal structure.

Several possible factors contribute to the structural differences between the NMR structure (Figure 4.17, A) and the crystal structure (Figure 4.17, B). One factor that may cause the structural differences is that the NMR structure is the unmodified 970 loop alone, while the crystal structure is the 970 loop in the whole 70S *E. coli* ribosome, and the interactions between the 970 loop and other ribosomal RNA or proteins in the 70S ribosome may contribute to stabilize the flipping and G971 (shown in Figure 4.18). Another possible reason causing the structural differences between the NMR structure and the crystal structure of the 970 loop is the chemical modifications in the 970 loop. That's because the NMR structure (Figure 4.17, A) is for the unmodified-970 RNA, however, the crystal structure (Figure 4.17, B) is from the 70S *E. coli* ribosome and therefore, the crystal structure is for the wild-type 970 loop with two modifications. The chemical modifications in the 970 loop may cause the structural change of the 970 loop. To clarify if the chemical modifications contribute to the structural differences between the NMR and the crystal structures in Figure 4.17, the NMR study for the modified-970 is needed.

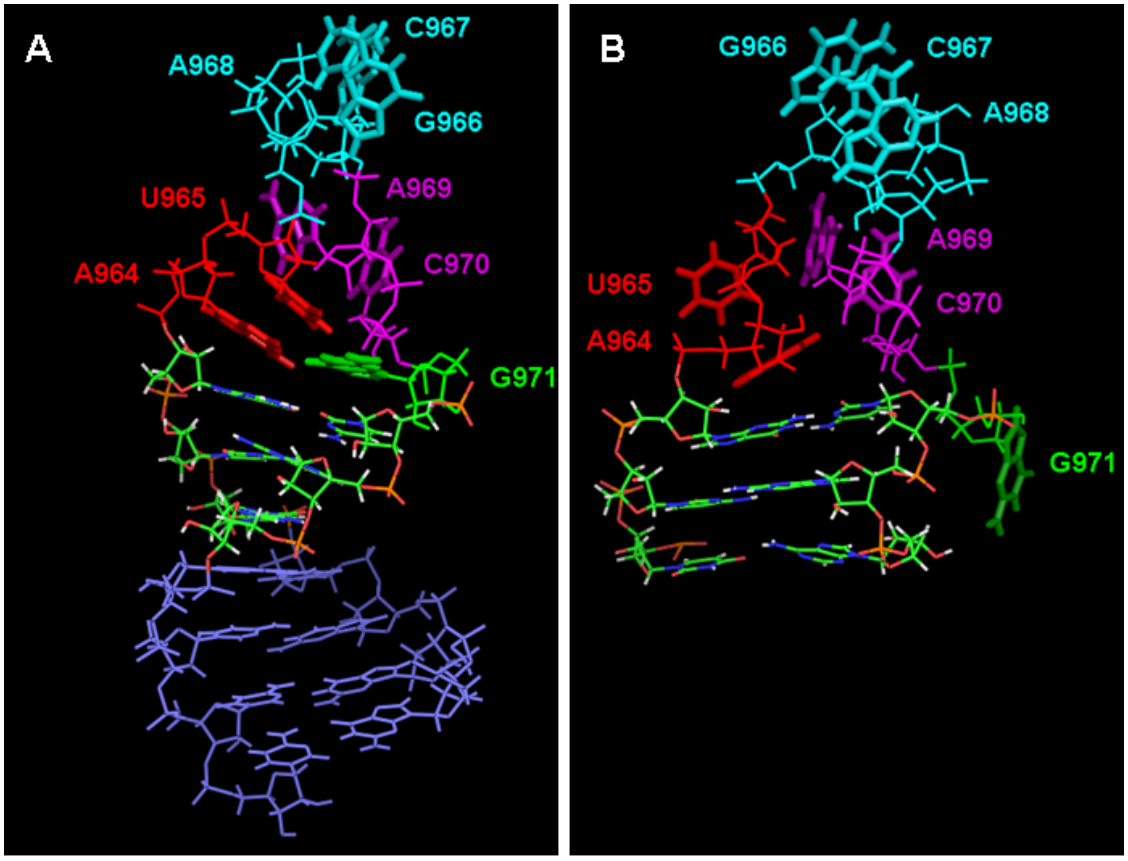


Figure 4.17. The NMR structure of the unmodified 970 loop (A) and the crystal structure of the 970 loop from 2AVY (B). The four extra base pairs in the NMR structure are shown in slate. And the carbon, nitrogen, oxygen, phosphorus, hydrogen atoms of the shared based pairs of the NMR and the crystal structures are shown in green, blue, red, orange, and white, respectively. A964 and U965 (in red) form parallel base-stacking in the NMR structure and perpendicular base-stacking in the crystal structures. A969 and C970 (in magenta) also interact by different base-stacking formats in the two structures. 966-967-968 form triple base-stacking in both structures (in cyan).

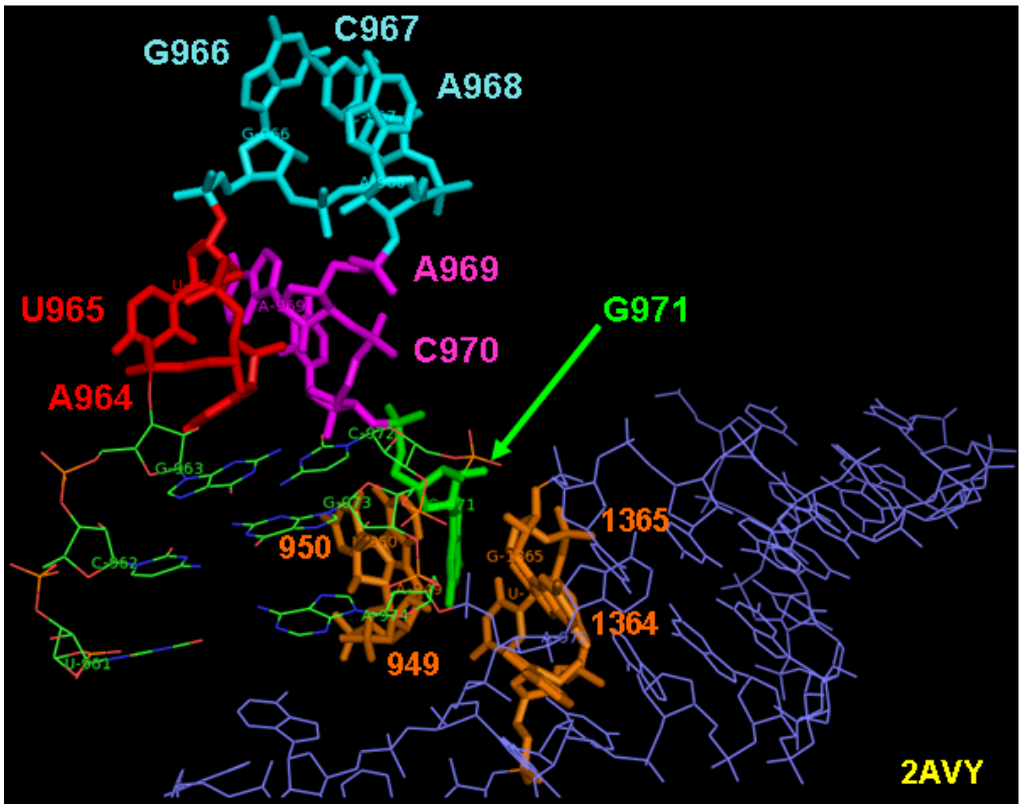


Figure 4.18. The crystal structure of the 970 loop. G971 (in green) flips out of the 970 loop and binds into a pocket formed by the residues 949, 950, 1364 and 1365 (in orange) in the 70S *E. coli* ribosome.

4.4. NMR Experiment Results for the Modified-970 RNA

The 1D-^{1}H NMR spectra of the modified 970 loop at different temperatures (5 ºC, 10 ºC, 15 ºC, 20 ºC, 25 ºC, 30 ºC, 37 ºC and 45 ºC) are shown in Figure 4.19. In the 11~15ppm of the spectrum, there are 6 peaks appear in the region, similar as the unmodified-970 RNA, which indicates the number of resolved hydrogen-bonded imino protons from the base-pairs are 6. And also similarly as the unmodified-970 loop, the peak resolution in 5 ~ 8ppm is improved at higher temperatures, and the peak resolution improvement is not significant when the temperature is higher than 37 ºC, therefore, the temperature chosen for running other NMR experiments is 37 ºC, which is the same temperature as the unmodified-970 RNA, and easily for comparing the spectral similarities and differences of the modified-970 RNA and unmodified-970 RNA.

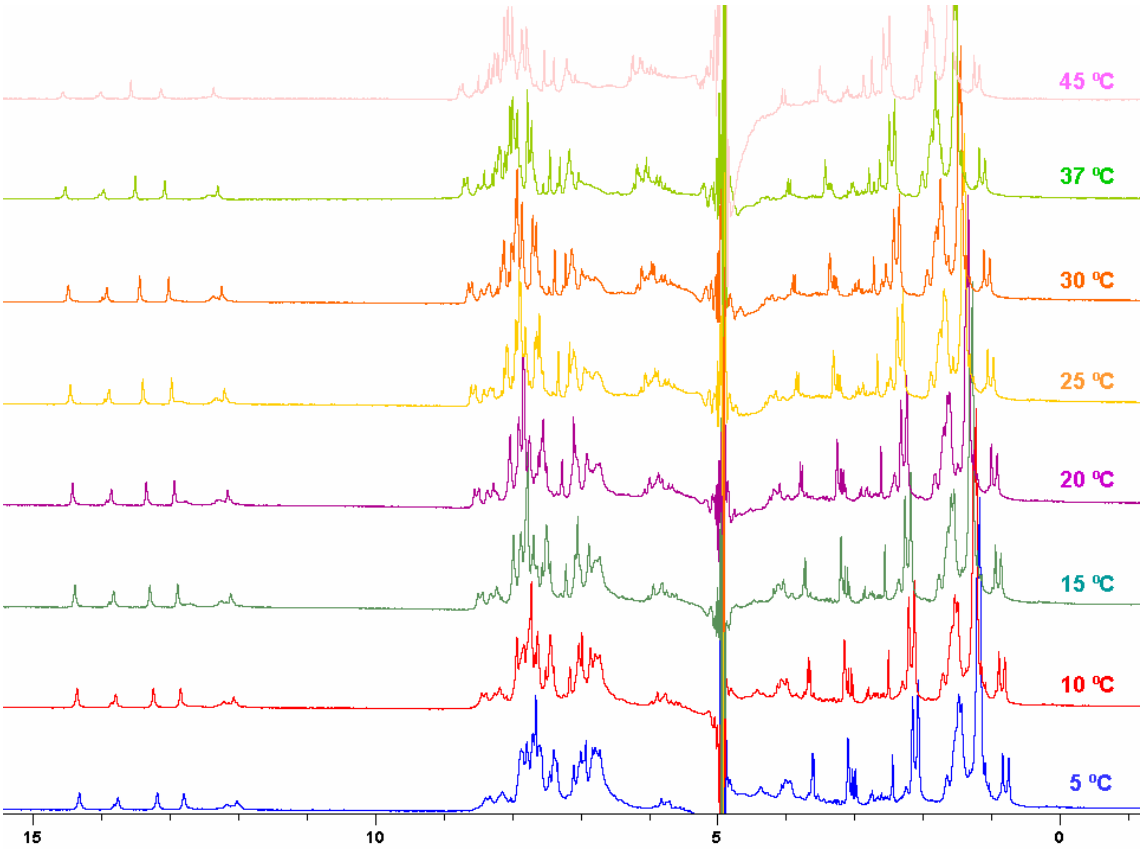


Figure 4.19. 1D-^{1}H spectra of the modified-970 RNA at 5 ºC, 10 ºC, 15 ºC, 20 ºC, 25 ºC, 30 ºC, 37 ºC and 45 ºC are shown in blue, red, blue, purple, yellow, orange, green and pink, respectively. The peak resolution is improved at higher temperature when the temperatures are from 5 ºC to 37 ºC, and no significant improvement in peak resolution can be observed between 37 ºC and 45 ºC. And 37 ºC is chosen as the optimum temperature for running the later NMR experiments of the modified-970 RNA.

The “NOE-walk” regions of NOESY spectra contain structural information. Therefore, the comparison of the NOESY spectra of the unmodified-970 RNA and the modified-970 RNA can be used to identify the structural similarities and differences between the 970 loop with and without modifications (shown in Figure 4.20). Chemical shifts of most H1’, H5, H6 and H8 are very similar between the unmodified-970 RNA and the modified-970 RNA , and most cross-peaks involved in the sequential NOE walk of the two spectra are at similar positions, which indicates that the structures of the unmodified-970 RNA and the modified-970 RNA have similarities. However, the chemical shifts of the modified residues and the residue around the modified residues, including A8 (964), U9 (965), m^2G10 (966) and m^5C11 (967) are perturbed (labeled in blue in Figure 4.20, A). The most significant change in the two spectra is that the NOE walk of C14 (C970)-G15 (G971)-C16 (C972) is broken in the modified-970 (the broken NOE walk is shown by dotted red lines in Figure 4.20, A), which indicates that the G15 (971) possibly flips out of the 970 hairpin and the distance between the base of G15 (971) and other residues are too far to be detected by NOESY, similarly as in the crystal structure (Figure 4. 17, B). And the chemical modifications in the 970 loop may contribute to the structural differences between the NMR and the crystal structures of the 970 loop in Figure 4.17.

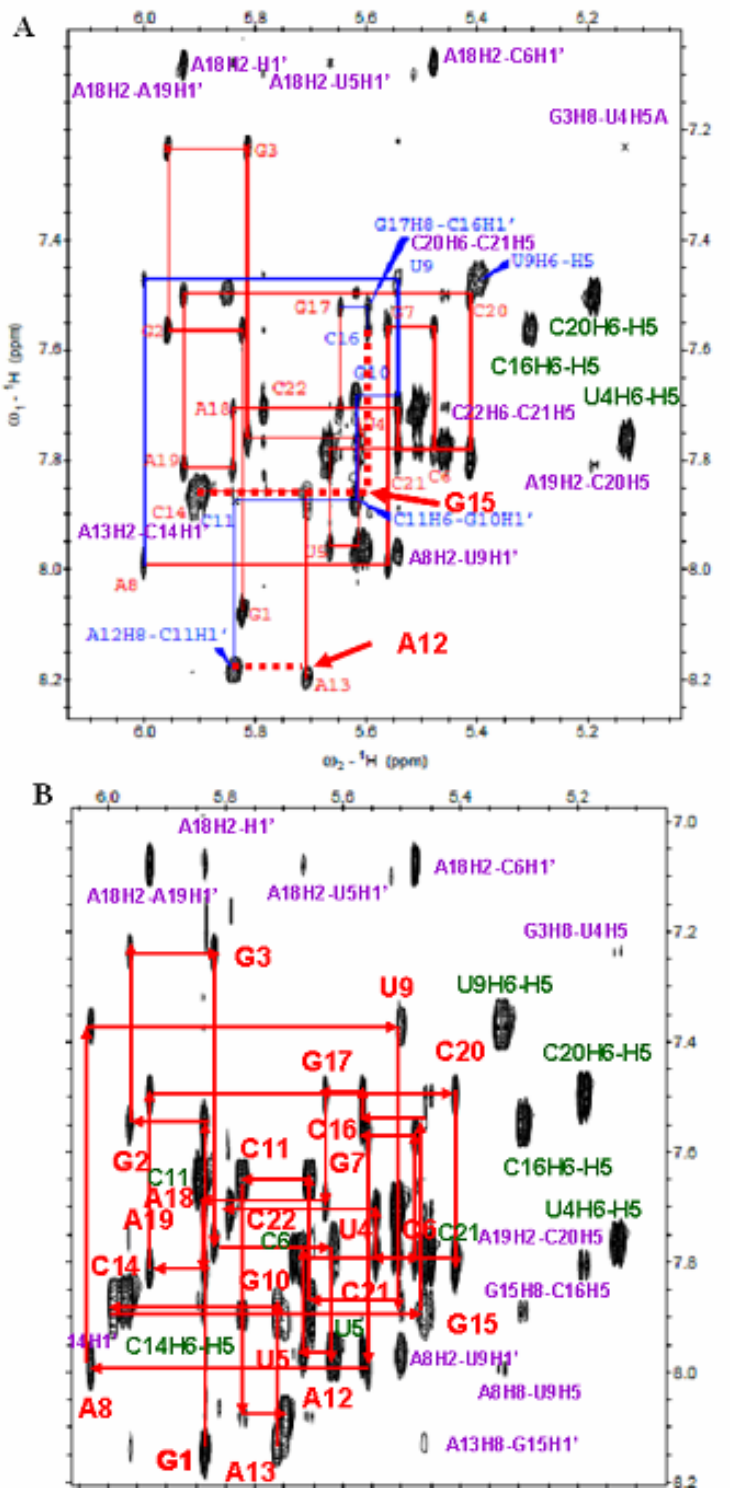


Figure 4.20. The "NOE-walk" regions of the NOESY spectra of the modified-970 RNA (A) and the unmodified-970 RNA (B).

Table 4.3. The chemical shifts of the protons in the modified-970 RNA at 37 ºC.

	H1'	H2	H2'	H3'	H4'	H5	H5'/H5''	H6/H8	H11/12/13 H31/32/33
G1	5.83	NA	4.93	4.62	4.53	NA	4.07/3.96	8.07	NA
G2	5.96	NA	4.77	4.65	4.56	NA	4.23	7.57	NA
G3	5.82	NA	4.56	4.45	4.50	NA	4.11	7.23	NA
U4 (960)	5.61	NA	4.56	4.48	ND	5.14	ND	7.76	NA
U5 (961)	5.67	NA	4.49	4.48	4.47	5.61	4.14	7.99	NA
C6 (962)	5.48	NA	4.42	4.55	4.47	5.67	4.13	7.78	NA
G7 (963)	5.56	NA	4.28	4.47	ND	NA	ND	7.56	NA
A8 (964)	6.00	7.97	4.55	4.63	4.48	5.40	ND	8.00	NA
U9 (965)	5.55	NA	4.43	ND	4.47	NA	4.10	7.48	NA
m^2G (966)	5.62	NA	4.55	4.43	4.30	NA	4.11	7.69	2.40
m^5C (967)	5.85	NA	4.24	4.52	4.17	NA	ND	7.87	1.60
A12 (968)	ND	7.50	4.74	4.63	4.34	NA	4.10/4.00	8.18	NA
A13 (969)	5.71	7.89	4.80	4.87	4.65	NA	ND	8.19	NA
C14 (970)	5.91	NA	4.47	4.66	4.40	5.90	4.25	7.86	NA
G15 (971)	ND	NA	4.31	4.39	4.55	NA	ND	7.90	NA
C16 (972)	5.60	NA	4.51	4.64	4.49	5.31	4.11	7.56	NA
G17 (973)	5.65	NA	4.58	ND	ND	NA	ND	7.52	NA
A18 (974)	5.84	7.08	4.53	4.49	4.34	NA	ND	7.71	NA
A19 (975)	5.93	7.84	4.47	4.61	4.50	NA	4.13	7.82	NA
C20	5.41	NA	4.24	4.34	4.40	5.19	4.05/3.95	7.50	NA
C21	5.44	NA	4.27	4.40	4.39	5.46	4.05	7.78	NA
C22	5.79	NA	4.03	4.20	ND	5.51	ND	7.71	NA

Some differences can also be observed in the DQF-COSY spectra of the unmodified-970 RNA and the modified-970 RNA. Ten peaks in the H5-H6 regions of the unmodified-970 RNA (Figure 4.21, A), which matches the ten pyrimidines in the sequence of the unmodified-970 RNA. The peaks in the H5-H6 region of the DQF-COSY spectrum of the modified-970 RNA (Figure 4.21, B) are very similar as those in the unmodified-970 RNA, and the only significant difference is the C11 (967) H5-H6 peak is absent in the modified-970 RNA. The absence of C11 (967) H5-H6 in the modified-970 RNA is reasonable with the consideration that the methyl group replaces H5 proton in the 5-methyl cytosine at the position 11 (967) in the modified-970 RNA. The missing H5-H6 peak in the residue 11 (967) is another evidence to confirm that the methyl group is located properly at the residue 11 (967). The peaks in the H1'-H2' region of the unmodified-970 and the modified-970 are also different. One relatively strong peak, C11H1'-H2', is observed, and it indicates that the sugar pucker conformation of the residue C11 in the modified-970 RNA is predominantly C2'-endo (Figure 4.21, B). Two relatively weak peaks, C14H1'-H2' and C22H1'-H2' in the H1'-H2' region, which indicates the sugar pucker conformations of the two residues C14 and C22 are dynamic, with partial population of C2'-endo (Figure 4.21, B). The sugar pucker conformations of A8, U9, C11, G15 are similar in the unmodified-970 RNA and in the modified-970 RNA, and the sugar pucker conformations of the four residues, G10, A12, A13 and C14 are different in the unmodified-970 RNA and the modified-970 RNA.

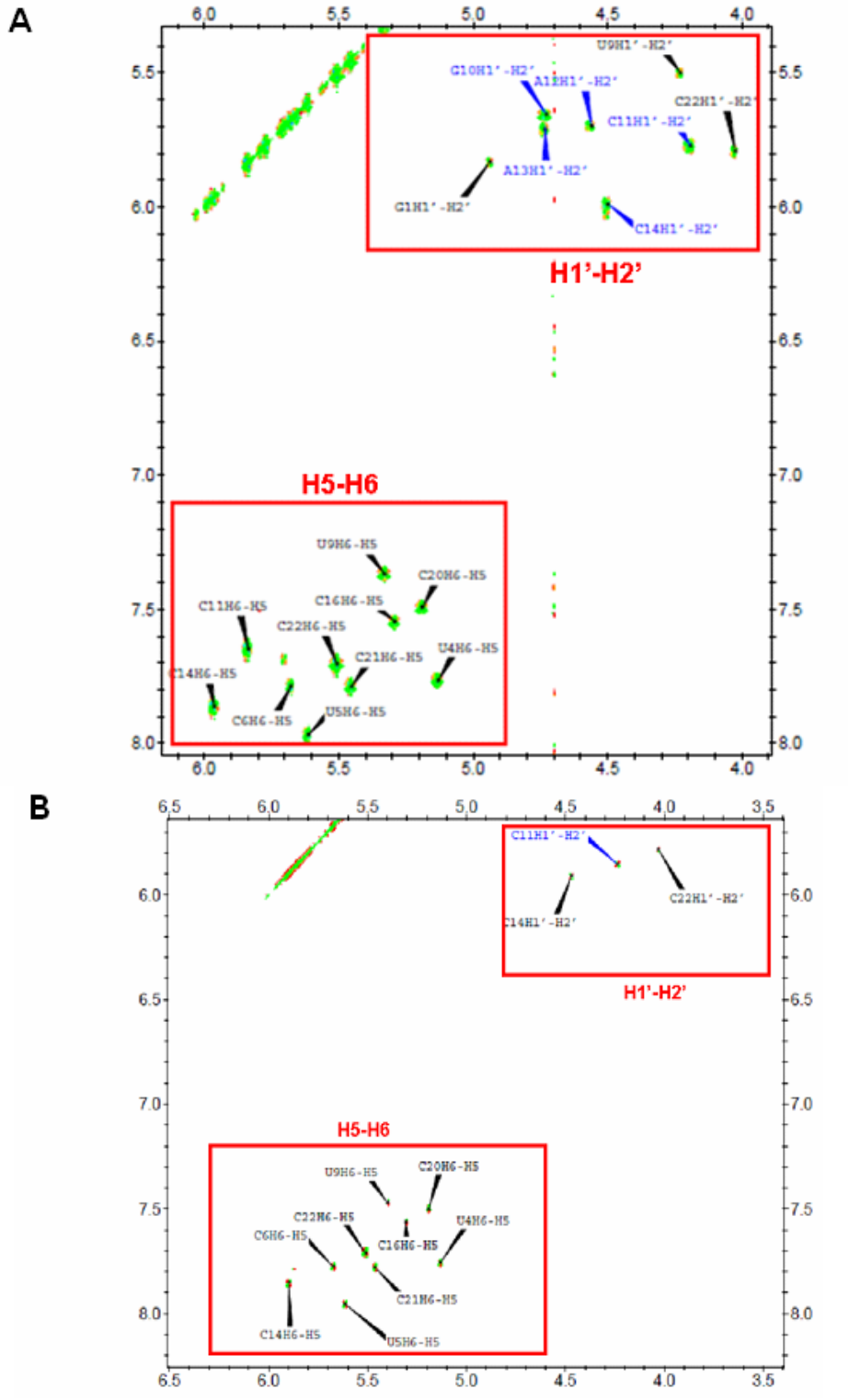


Figure 4.21. The DQF-COSY spectra of the unmodified-970 RNA (A) and the modified-970 RNA (B).

1D-^{31}P spectrum of the unmodified-970 RNA (Figure 4.22, A) is also different from the 1D-^{31}P spectrum of the modified-970 RNA (Figure 4.22, B). In the 1D-^{31}P spectrum of the unmodified-970 RNA, the chemical shift of G10P is identified to be different from phosphorus of other residues, and the dihedral angle α of G10 and ζ of U9 are different from other residues (Figure 4.22, A). The chemical shifts of phosphorus in the modified-970 RNA are all similar, which indicates that no unique α or ζ in the modified-970 RNA (Figure 4.22, B).

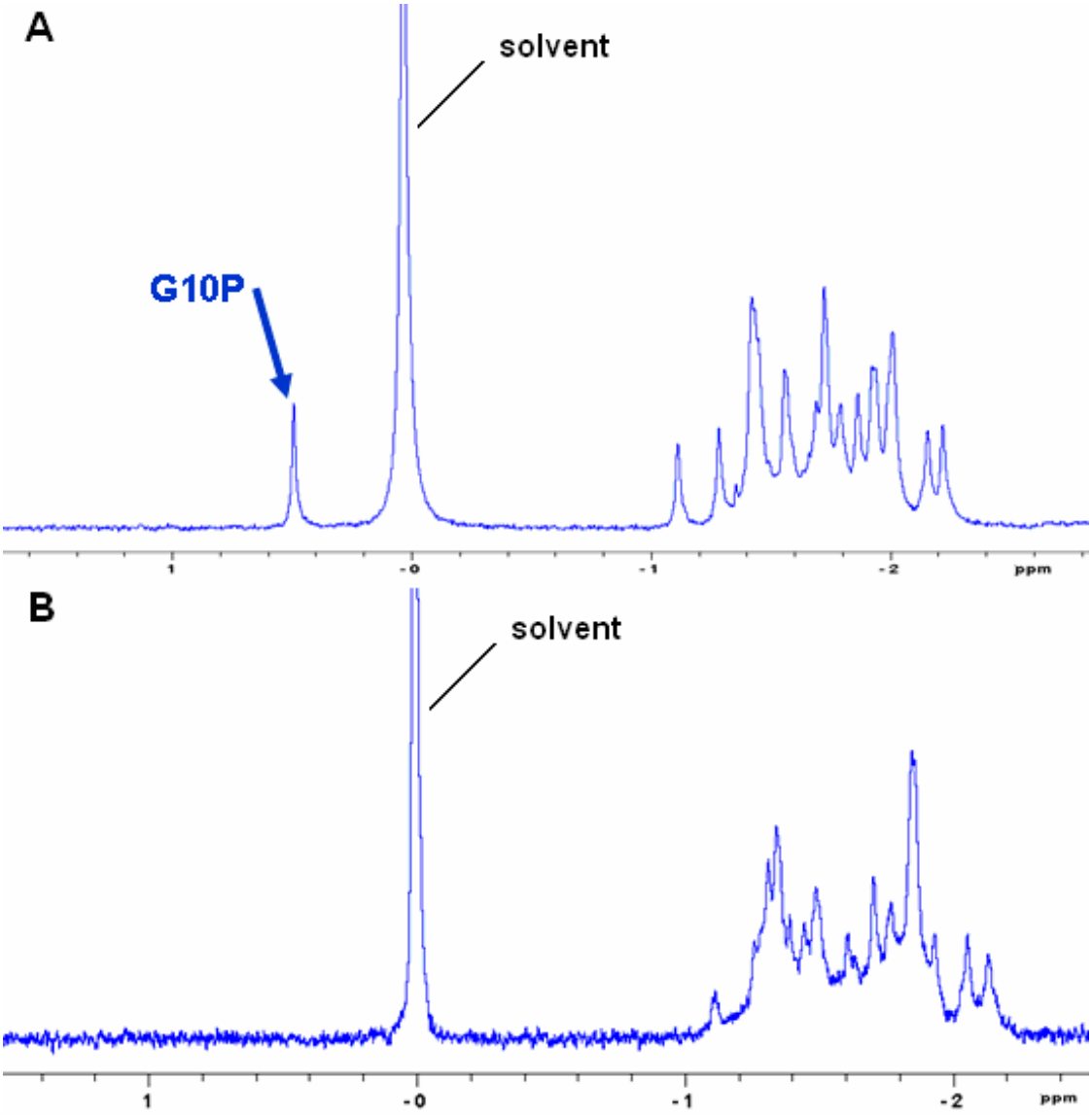


Figure 4.22. The 1D-^{31}P spectra of the unmodified-970 RNA (A) and the modified-970 RNA (B). The significant difference between the unmodified-970 and the modified-970 is that the phosphorus chemical shift of the residue G10. In the unmodified-970 RNA, the dihedral angle α of G10 and ζ of U9 are different from other residues, while in the modified-970 RNA, no unique dihedral angles are observed.

4.5. Comparison of the NMR and Crystal Structures of the 970 Loop

One refined NMR structure of the modified 970 loop is shown in Figure 4.23, A. The NMR structures of the unmodified-970 RNA and the modified-970 RNA share some common features. The residues in the two structures mainly interact by base stacking, for example, the triple base stacking exist among the position 966, 967 and 968 (in cyan), the base stacking formed between the position 964 and 965 (in red), and between the position 969 and 970 (in magenta). Some differences in the base stacking formats in some residues of the two structures are observed. For example, the three residues involved in the triple-base stacking (966-967-968) form parallel base-stacking in the modified-970 RNA, however, the base of residue 966 form perpendicular with the residue 967 and 968 in the unmodified-970 RNA. And the base-stacking formats between the position 969 and 970 are different in the unmodified-970 RNA and the modified-970 RNA, parallel base-stacking and perpendicular base-stacking between the two positions exist in the unmodified-970 RNA and in the modified-970 RNA, respectively. And the most significant difference between the 970 loop with modifications and without modifications is the placement of residue G971. As expected from the NOESY differences between the unmodified-970 and modified-970 loop, the base of residue G971 flips out of the loop in the NMR structure of the modified 970 loop, similarly as in the crystal structure. And the chemical modifications contribute to the structural change, including the flipping of G971 based on the comparison.

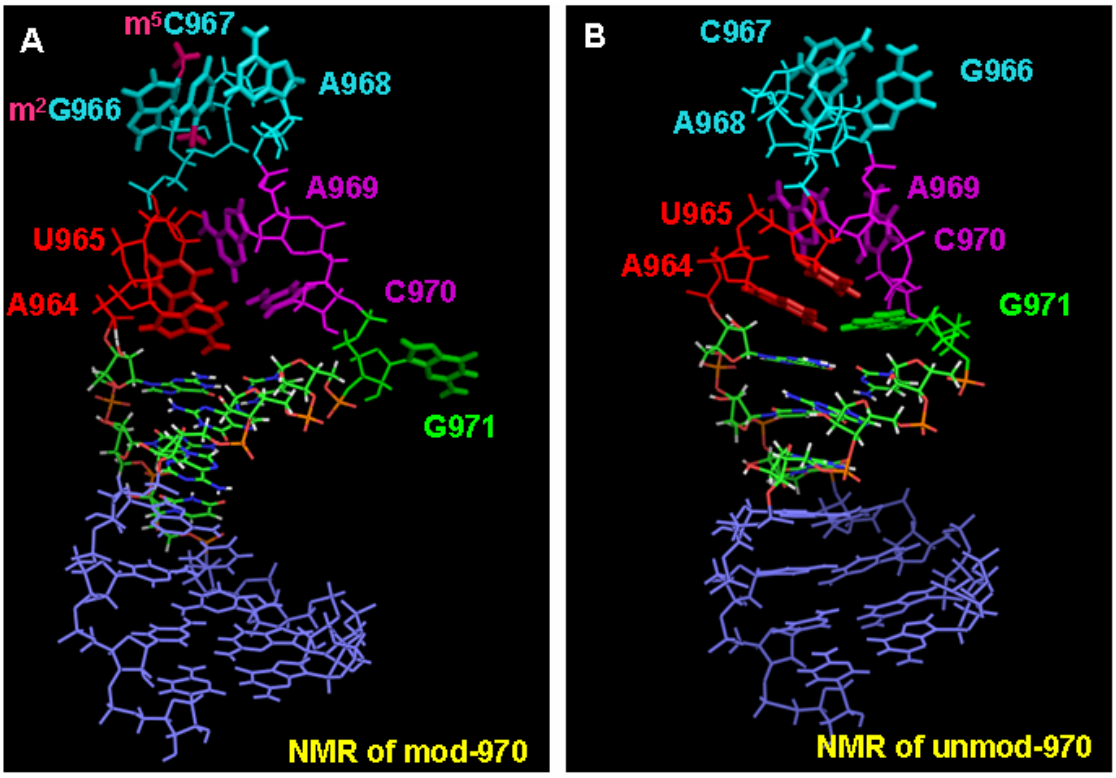


Figure 4.23. The NMR structures of the modified-970 RNA (A) and of the unmodified-970 RNA (B). A964 and U965 form parallel base-stacking in the NMR structures of the two structures (in red). A969 and C970 interact by base-stacking in the two structures (in magenta). And 966-967-968 form triple base-stacking in both the modified-970 RNA and the modified-970 RNA (in cyan). The significant difference between the two structures is the placement of G971.

The NMR structure and the crystal structure of the modified-970 RNA are shown in Figure 4.24. The NMR structure and the crystal structure (PDB: 2AVY) are very similar. For example, each residue in the loop of the two structures are involved in base stacking with other residues, and the residues interact by base stacking are shown in the same color. And the residues involved in triple-base stacking (in cyan) form parallel base-stacking in the two structures, and the two residues shown in magenta interact by perpendicular base-stacking in both structures. In addition, the base of G971 flips out the loop in the two structures. Some structural differences can be observed in the two structures, for example, 964 and 965 form parallel base stacking and perpendicular stacking in the NMR structure and the crystal structure of the modified-970 RNA, respectively. Although 969 and 970 form perpendicular base-stacking in the two structures shown in Figure 4.24, the relative positions of 970 are different in the two structures.

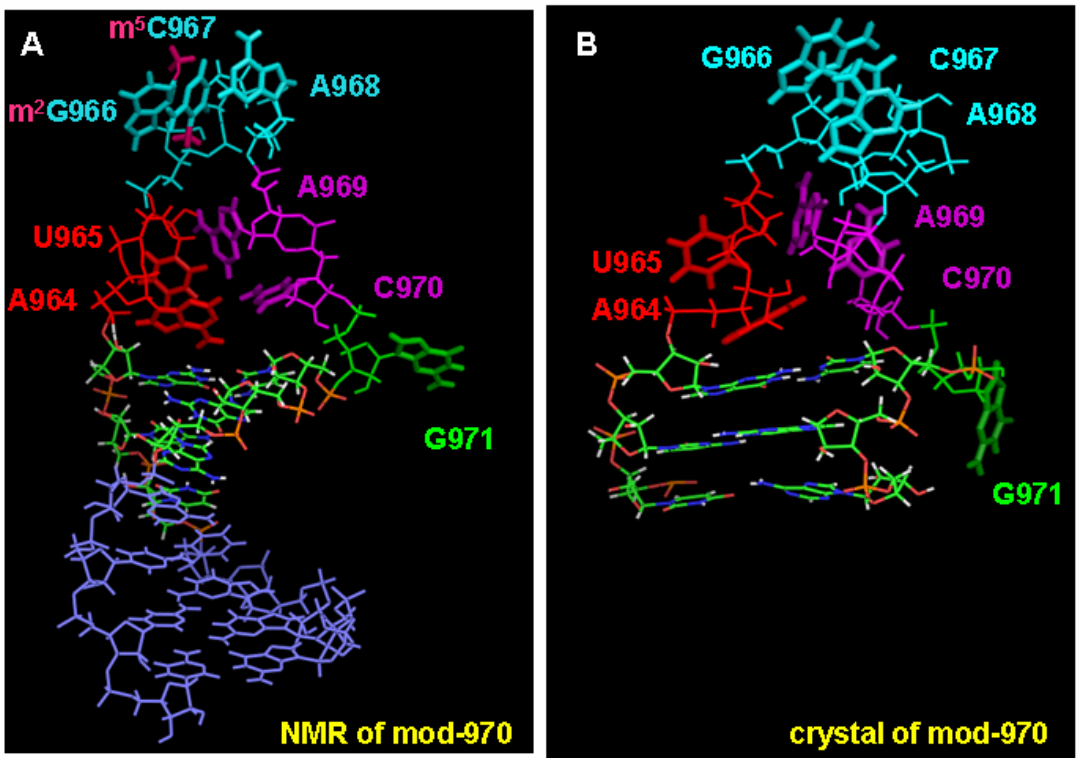


Figure 4.24. The NMR structure (A) and the crystal structure (B) of the modified-970 RNA. A964 and U965 form base-stacking in the NMR structures of the two structures in different formats (in red). A969 and C970 interact by perpendicular base-stacking in the two structures (in magenta). The triple base-stacking of 966-967-968 (in cyan) is observed. Flipping G971 is another common feature of the two structures (in green).

4.6. Structure-function Relationship of the NMR Structure of the Modified-970 Loop

One refined NMR structure of the 970 loop is shown in Figure 4.25. The main interactions in the 970 loop with two modifications of the NMR structure are base-stacking. The triple base stacking among position 966, 967 and 968 locate the three residues on the top of the 970 loop, which are accessible for molecular interaction, including interacting with P-site bounded tRNA, Initiation Factor 3 (IF3), and tetracycline (Brodersen 2000 and Pioletti 2001). Position 966, 967 and 968 form parallel base stacking in the triple-base stacking, and the methyl groups in 966 and 967 increase the stacking surface area, which can stabilize the stability of the 970 loop. What's more, the mutation study done in Dr. Cunningham's lab shows that purines are preferred at position 966 (Saraiya 2008), possibly because purines have larger stacking surface area than pyrimidines, which stabilize the triple-base stacking. And cytosine and adenine are preferred at position 967, uracil is preferred at position 968, the preferences of these nucleotides at 967 and 968 may be due to the additional dipole interactions are formed in the triple-base stacking (Saraiya 2008). Base stacking interactions among position 964, 965, 969 and 970 are shown in red (Figure 4. 25). A969 forms parallel base-stacking with U965 and perpendicular base-stacking with C970. The base stacking interactions between 969 and other residues support the mutation study result obtained in Dr. Cunningham's lab that purines are preferred at position 969 (Saraiya 2008) because that larger stacking surface area of purines may stabilize the stacking interactions of the red residues.

The possible role of flipping G971 in the 970 loop is to put position 971 in a better position for interactions with a pocket formed by the backbones of ribosomal RNA at position 949, 950, 1363, 1364 and 1365 in the crystal structure of the ribosome (shown in Figure 4.18), and the interactions between 971 and other ribosomal RNA may be related to the movement and flexibility of ribosome during protein synthesis, and possibly affect translocation rate of mRNA in translation, and some mutations at 971 can affect the formation of hydrogen bonds and the mobility of the 970 loop, which cause the lower function of the mutant G971U (Saraiya 2008).

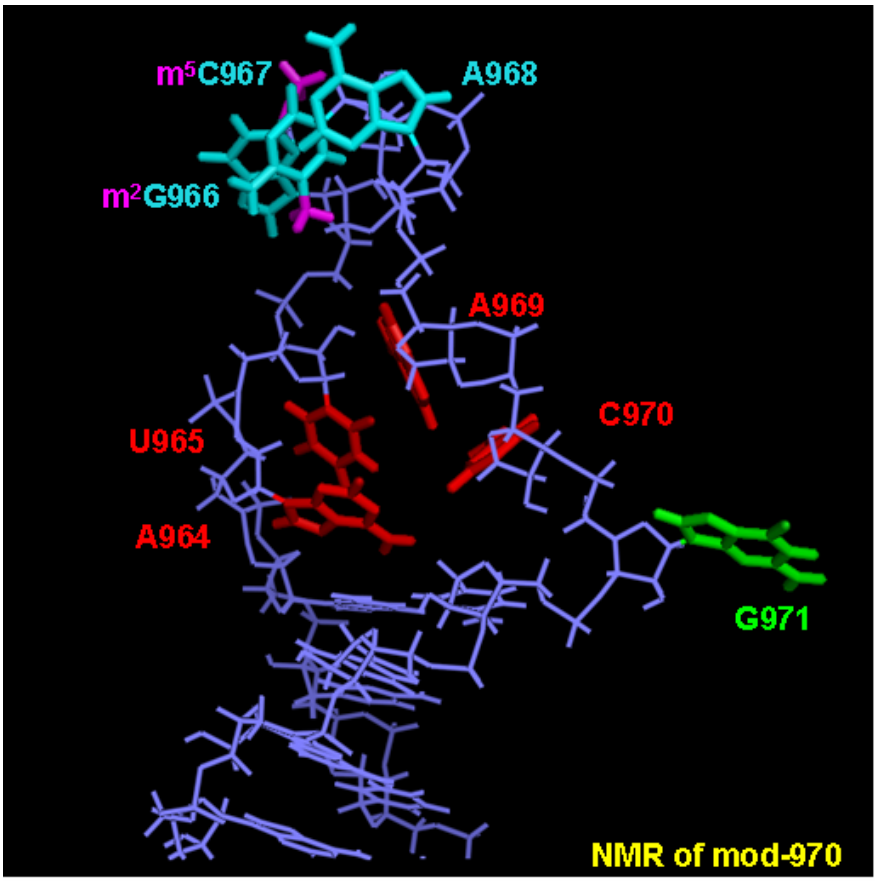


Figure 4.25. The NMR structure of the modified-970 RNA. The triple base-stacking of 966-967-968 is shown in cyan. A964, U965, A969 and C970 form base-stacking (in red). Flipping G971 is shown in green.

4.7. Interactions between the Unmodified-970 RNA and the 7mer-peptide

The peptide ligand with sequence TLWDLIP was found to bind with the 970 loop by phage display method and synthesized in Dr. Chow's and Dr. Cunningham's labs (Abeydeera 2009) (Lamichhane 2009). The primary sequence of the 7mer-peptide is shown in Figure 4.26. The 1D-^{1}H NMR spectrum of the peptide is shown in Figure 4.27. Most peaks in the NMR spectrum are located in the range 0.5~5ppm, these peaks are too crowded to be identified. And resolved peaks in 7~8ppm are protons in the ring of tryptophan (W), and these peaks are assigned based on peak splitting due to J-coupling interactions with their surrounded hydrogen atoms and peak area that are proportional to the hydrogen atom numbers.

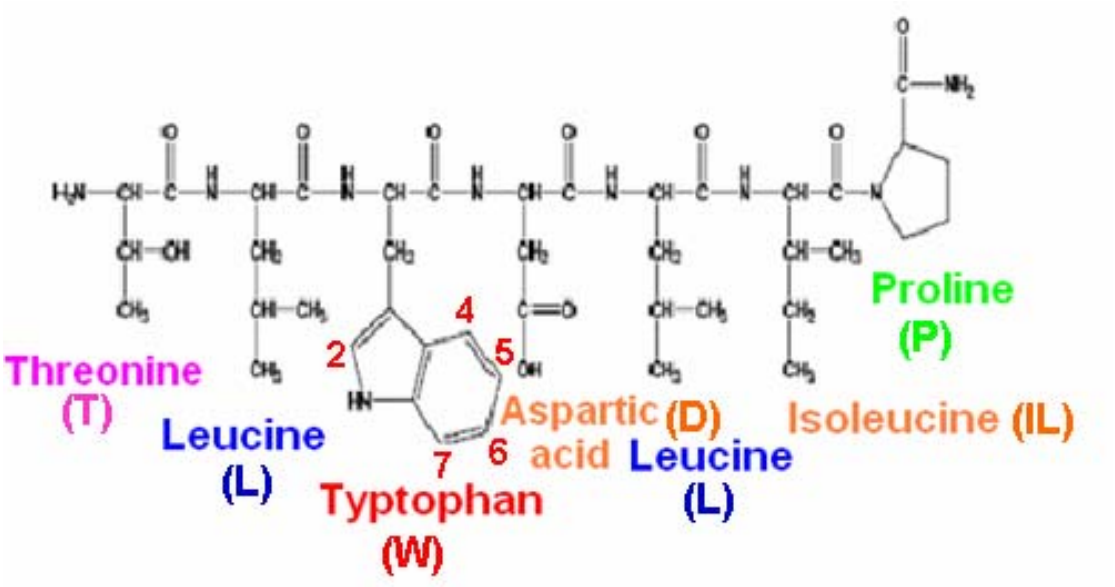


Figure 4.26. The primary sequence of the peptide.

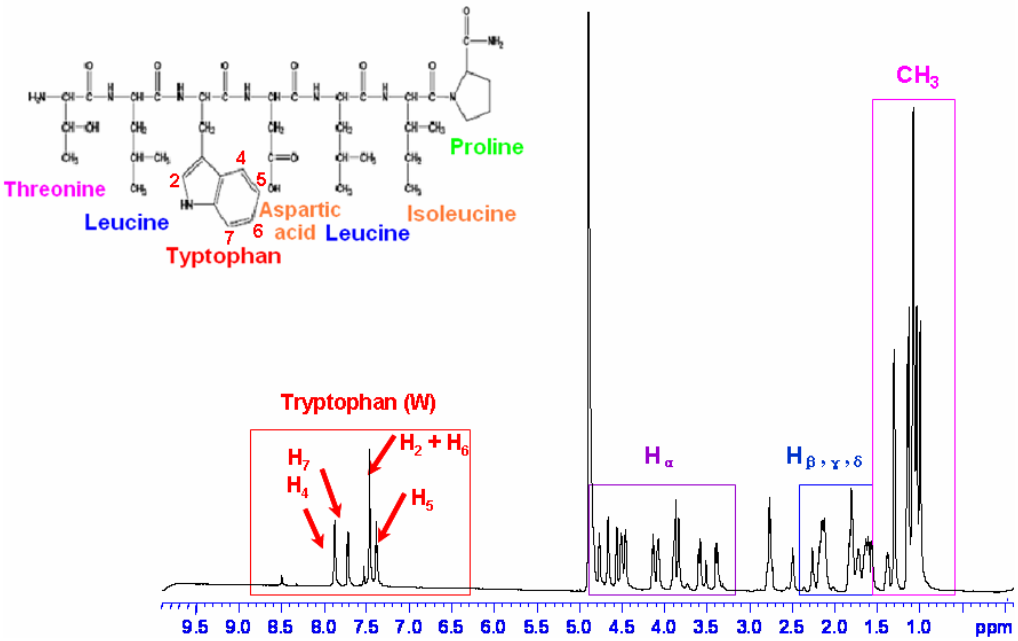


Figure 4.27. 1D-^{1}H NMR spectra of the peptide in D_2O. The peaks labeled in red are the base protons from the ring of tryptophan (W).

The amino acid components in the peptide are confirmed by TOCSY spectrum mainly based on their J-coupling interactions with their neighboring hydrogen atoms, and all the non-exchangeable protons of the peptide are assigned by using the TOCSY spectrum (Figure 4.28), and the assignments for the protons of different amino acids are labeled by different colors matching the color labeled amino acids in Figure 4.26.

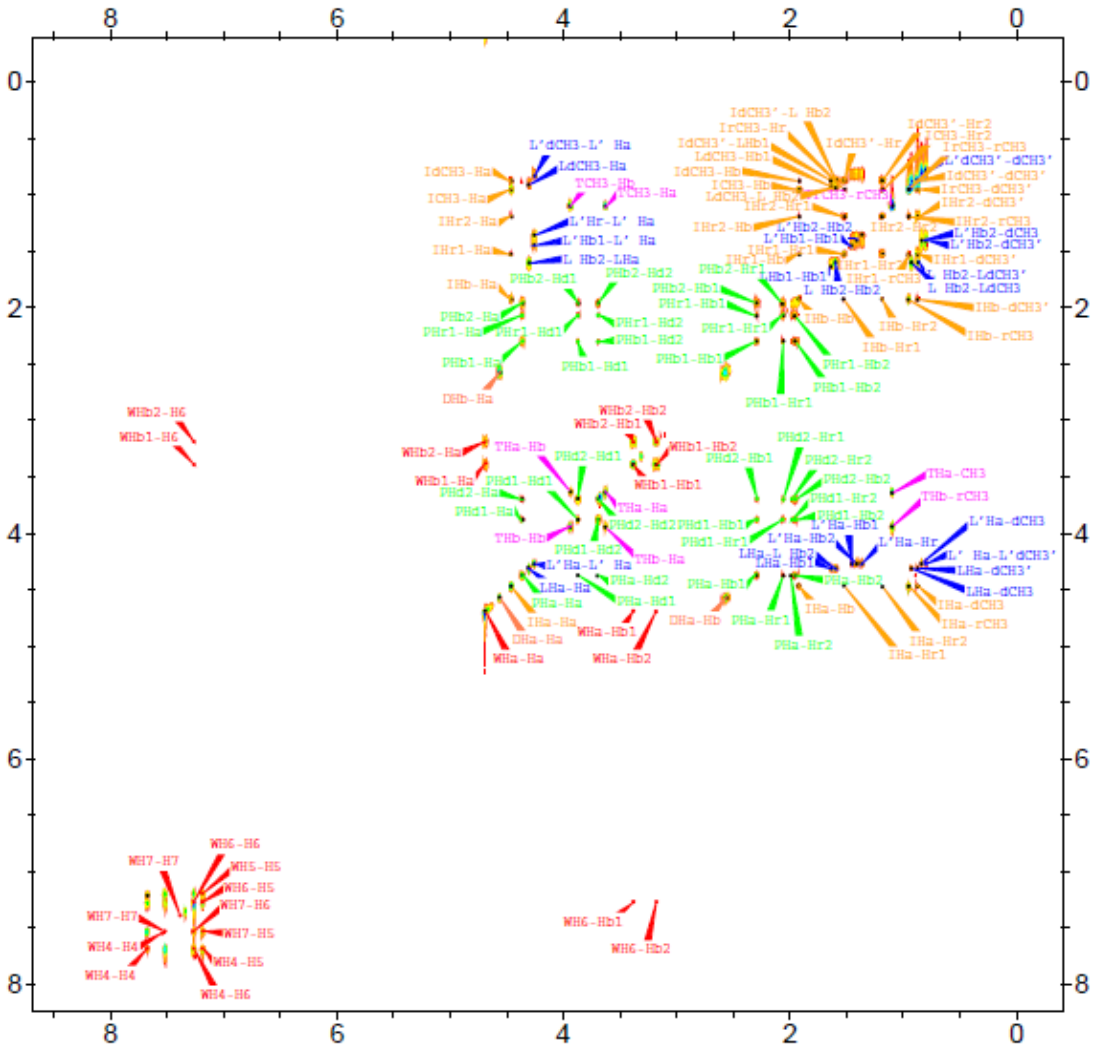


Figure 4.28. The TOCSY spectrum of the peptide in D_2O. The assignments in the TOCSY of the amino acid components in the peptide are labeled by different colors as in the primary sequence (Figure 4-25).

After confirming the component of the peptide, the peptide is added to the unmodified-970 to test if the peptide and the unmodified-970 interact. From the Circular Dichroism spectra, the complex of the unmodified-970 RNA and the peptide is not the sum of the unmodified-970 RNA alone and the peptide alone, which indicates that the unmodified-970 RNA and the peptide interact, and the interactions may cause conformational change when the unmod-970 RNA mixed with the peptide at 1:1 ratio (Figure 4.29).

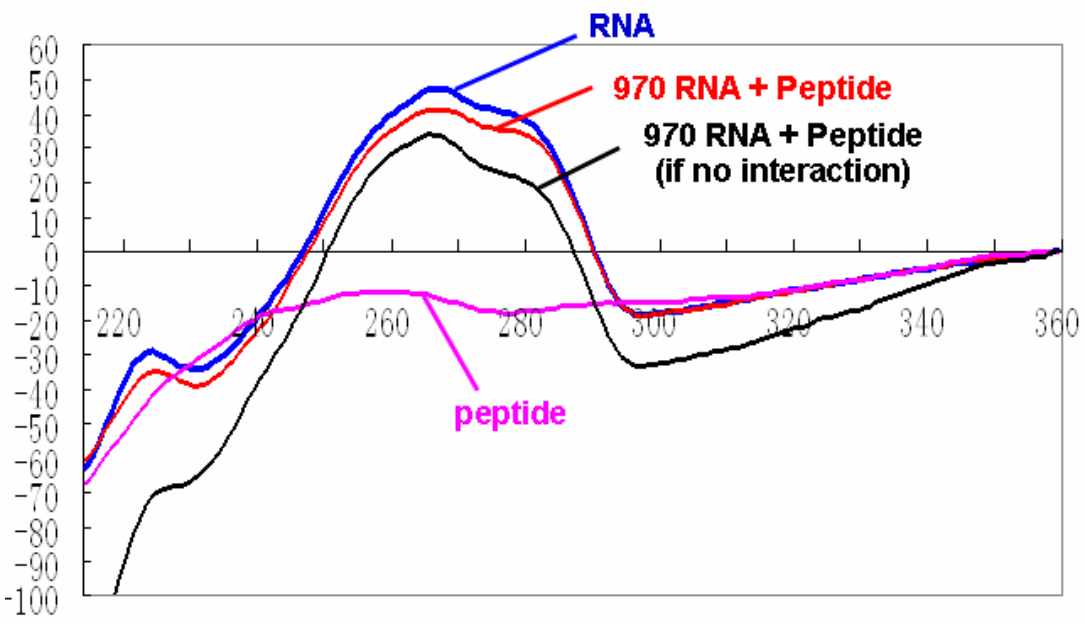


Figure 4.29. The CD spectra of the unmod-970 RNA alone, the peptide alone, sum of the unmod-970 and the peptide, and complex of the unmod-970 and the peptide are shown in blue, magenta, black and red lines, respectively. The differences between the red line and black line indicate that the unmod-970 RNA may interact with the peptide.

The NMR spectra are also used to identify the interactions between the peptide and the unmodified-970 RNA. The exchange between two compounds, mostly in biology, free and ligands bound of proteins or nucleic acids, usually gives rise to two distinct NMR signals for a given spin affected by binding because the chemical environment of the free proteins or nucleic acids are different from the bound proteins or nucleic acids (Clos 2007). Some differences can be observed between the NOESY spectra of the unmod-970 alone and that of the complex of the unmodified-970 RNA and the peptide (in Figure 4.30). Several new peaks shown in the regions of the NOESY spectrum support the interactions between the unmodified-970 RNA and the peptide. And the assignments of the new peaks provide useful information about interaction sites between the peptide and the unmodified-970 RNA are: H4 of tryptophan and A12H1' of the RNA, Hβ of aspartic acid and C11H6, and Hβ of aspartic acid and A8H2 of the RNA (data not shown). But the interactions between the peptide and the unmodified-970 cause little structural change based on the fact that most NOESY peaks of the unmodified-970 RNA alone and those of the complex of the unmodified-970 RNA and the peptide are overlapped (Figure 4.30). The NMR study provide primary information about binding between the unmodified-970 RNA and the 7mer-peptide, however, the mechanism of how the peptide molecules bind with the unmodified-970 RNA molecules is still not clear.

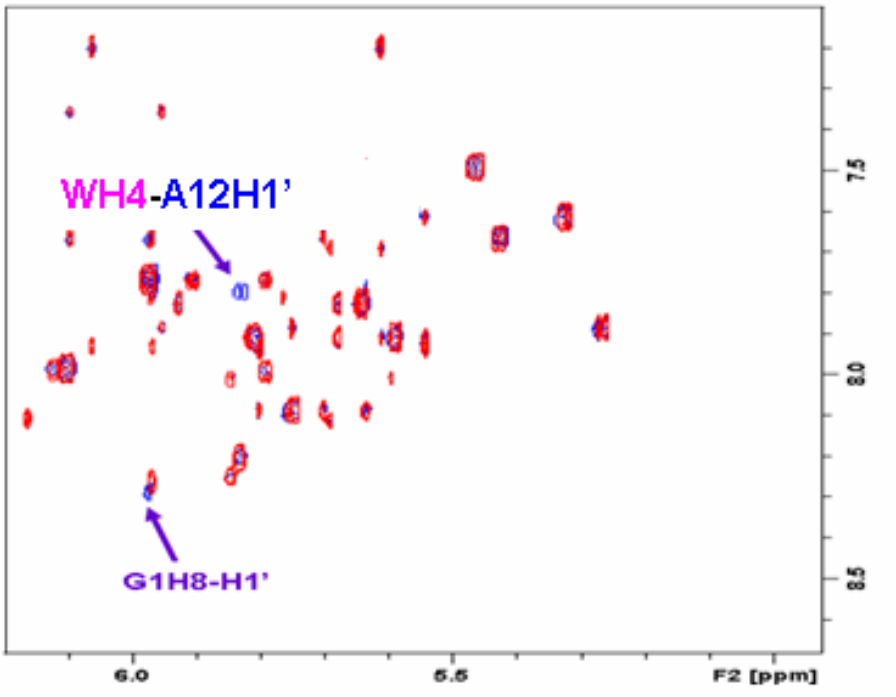


Figure 4.30. The "NOE-walk" regions of the unmodified-970 alone (red peaks) and the the complex of the unmodified-970 RNA and the peptide (blue peaks). The peak WH4-A12H1' indicates the interaction between the tryptophan and A12 in the unmodified-970 RNA. The overlapping of most red peaks and blue peaks indicate that no significant structure change of the unmodified-970 RNA caused by the possible interactions between the modified-970 RNA and the 7mer-peptide.

4.8. Conclusions

The NMR structures of the modified-970 RNA and the unmodified-970 have similarities, for example, the position involved in base-stacking interactions are similar in the two structures. And some differences are also observed in the base stacking formats of some residues in the NMR structures of the 970 loop with and without modifications, and the most significant difference is the placement of G971. The structural differences between the two NMR structures indicate that the chemical modifications in the 970 loop cause the structural change. And the modified nucleotides in the 970 loop may contribute to the structural stability of the 970 loop because methyl groups in the two modified residues 966 and 967 increase the stacking surface area of the triple-base stacking formed among 966, 967 and 967. Some features of the NMR structure of the modified-970 RNA are also supported by the experimental data of mutation study in Dr. Cunningham's lab. In addition, the difference of NOESY spectra between the unmodified-970 alone and the complex of the unmodified-970 RNA and the 7mer-peptide identify the interaction sites between the unmodified-970 RNA and the peptide.

APPENDIX 1

Structural Restraints for the 690UC Mutant

Distance restraints for the 690UC mutant

```
!stem stacking distances
assign (residue 1 and name h1) (residue 2 and name h1) 3.5 0.5 0.5
assign (residue 2 and name h1) (residue 12 and name h1) 4.0 1.0 0.0

! base-pair
assign (residue 1 and name C1') (residue 14 and name C1') 10.6 0.2 0.2
assign (residue 1 and name C8) (residue 14 and name C6) 9.8 0.3 0.3
assign (residue 2 and name C1') (residue 13 and name C1') 10.6 0.2 0.2
assign (residue 2 and name C8) (residue 13 and name C6) 9.8 0.3 0.3
assign (residue 3 and name C1') (residue 12 and name C1') 10.6 0.2 0.2
assign (residue 3 and name C6) (residue 12 and name C8) 9.8 0.3 0.3
assign (residue 1 and name O6) (residue 2 and name O6) 3.0 0.3 0.3
assign (residue 1 and name N2) (residue 2 and name N2) 4.5 0.5 0.5
assign (residue 2 and name O6) (residue 3 and name N4) 3.0 0.6 0.6
assign (residue 2 and name N2) (residue 3 and name O2) 4.5 0.5 0.5
assign (residue 12 and name O6) (residue 13 and name N4) 3.0 0.5 0.5
assign (residue 12 and name N2) (residue 13 and name O2) 4.5 0.5 0.5
assign (residue 13 and name N4) (residue 14 and name N4) 3.0 0.3 0.3
assign (residue 13 and name O2) (residue 14 and name O2) 4.5 0.5 0.5

! Base Pair restraints
```

```
! for G1/ C14  base pair
assign (resid 1 and name N1) (resid 14 and name N3) 2.91 0.3 0.3
assign (resid 1 and name O6) (resid 14 and name N4) 2.71 0.3 0.3
assign (resid 1 and name N2) (resid 14 and name O2) 3.00 0.3 0.3
assign (resid 1 and name H1) (resid 14 and name N3) 1.89 0.3 0.3
assign (resid 1 and name O6) (resid 14 and name H42) 1.71 0.3 0.3
assign (resid 1 and name H22) (resid 14 and name O2) 1.90 0.3 0.3

! for G2/ C13  base pair
assign (resid 2 and name N1) (resid 13 and name N3) 2.91 0.3 0.3
assign (resid 2 and name O6) (resid 13 and name N4) 2.71 0.3 0.3
assign (resid 2 and name N2) (resid 13 and name O2) 3.00 0.3 0.3
assign (resid 2 and name H1) (resid 13 and name N3) 1.89 0.3 0.3
assign (resid 2 and name O6) (resid 13 and name H42) 1.71 0.3 0.3
assign (resid 2 and name H22) (resid 13 and name O2) 1.90 0.3 0.3

! for C3/ G12  base pair
assign (resid 3 and name N3) (resid 12 and name N1) 2.91 0.3 0.3
assign (resid 3 and name N4) (resid 12 and name O6) 2.71 0.3 0.3
assign (resid 3 and name O2) (resid 12 and name N2) 3.00 0.3 0.3
assign (resid 3 and name N3) (resid 12 and name H1) 1.89 0.3 0.3
assign (resid 3 and name H42) (resid 12 and name O6) 1.71 0.3 0.3
assign (resid 3 and name O2) (resid 12 and name H22) 1.90 0.3 0.3

! NOE from H2O at 25C
assign (resid 2 and name H1) (resid 13 and name H5) 4.5 0.5 0.5
assign (resid 3 and name H41) (resid 4 and name H3) 4.5 0.0 0.5
assign (resid 3 and name H42) (resid 11 and name H42) 4.0 0.5 0.5
 assign (resid 4 and name H3) (resid 4 and name H2') 4.5 0.5 0.5
```

```
assign (resid 4 and name H6) (resid 12 and name H1) 4.5 0.0 1.5
assign (resid 4 and name H1') (resid 12 and name H1) 3.5 0.5 0.5
assign (resid 5 and name H2') (resid 6 and name H3) 4.5 0.5 0.5
assign (resid 11 and name H42) (resid 12 and name H1) 4.5 0.5 0.5
assign (resid 11 and name H5) (resid 11 and name H41) 2.5 0.5 0.5

! for G5/ A10 mismatch
assign (resid 5 and name H21) (resid 10 and name N7) 1.80 0.3 0.3

! Base to Base NOEs (change to small error bars on Nov 20.2007)
assign (residue 2 and name H8) (residue 3 and name H5) 4.5 0.5 0.5
assign (residue 2 and name H8) (residue 3 and name H6) 4.5 0.5 0.5
assign (residue 3 and name H5) (residue 4 and name H5) 3.5 0.5 1.0
assign (residue 3 and name H5) (residue 4 and name H6) 4.5 0.5 1.0
assign (residue 3 and name H6) (residue 4 and name H5) 4.5 0.5 0.5
assign (residue 3 and name H6) (residue 4 and name H6) 4.0 0.5 1.0
assign (residue 4 and name H5) (residue 5 and name H8) 4.5 0.5 0.5
assign (residue 4 and name H6) (residue 5 and name H8) 4.0 0.5 1.0
assign (residue 5 and name H8) (residue 6 and name H5) 3.5 0.5 1.5
assign (residue 5 and name H8) (residue 6 and name H6) 4.0 0.5 1.0
assign (residue 7 and name H8) (residue 8 and name H8) 3.5 1.0 1.5
assign (residue 9 and name H8) (residue 10 and name H8) 4.0 0.5 1.0
assign (residue 10 and name H8) (residue 11 and name H5) 4.5 0.5 0.5
assign (residue 10 and name H8) (residue 11 and name H6) 4.5 0.5 0.5
assign (residue 11 and name H5) (residue 12 and name H8) 4.0 0.5 0.5
assign (residue 11 and name H6) (residue 12 and name H8) 3.5 0.5 1.0
assign (residue 12 and name H8) (residue 13 and name H5) 3.5 0.5 1.0
assign (residue 12 and name H8) (residue 13 and name H6) 4.5 0.5 0.5
assign (residue 13 and name H6) (residue 14 and name H6) 4.0 0.5 1.0
```

```
! Intra Residue
assign (residue 1 and name H8)(residue 1 and name H1')  4.0 0.5 1.0
assign (residue 2 and name H8) (residue 2 and name H1') 3.0 0.5 1.0
assign (residue 2 and name H8) (residue 2 and name H2') 3.0 0.5 1.5
assign (residue 2 and name H8) (residue 2 and name H3') 2.5 0.5 1.0
assign (residue 3 and name H5) (residue 3 and name H3') 4.0 0.5 1.0
assign (residue 3 and name H6) (residue 3 and name H1') 3.5 0.5 1.5
assign (residue 3 and name H6) (residue 3 and name H2') 3.0 0.5 1.0
assign (residue 3 and name H6) (residue 3 and name H3') 3.0 0.5 0.8
assign (residue 4 and name H6) (residue 4 and name H1') 3.5 0.5 0.5
assign (residue 4 and name H6) (residue 4 and name H2') 2.5 0.5 1.0
assign (residue 4 and name H6) (residue 4 and name H3') 2.5 0.5 1.0
assign (residue 5 and name H8) (residue 5 and name H1') 3.5 0.5 0.5
assign (residue 5 and name H8) (residue 5 and name H2') 2.5 0.5 1.0
assign (residue 5 and name H8) (residue 5 and name H3') 3.0 0.5 0.8
assign (residue 6 and name H5) (residue 6 and name H2') 4.5 0.5 1.0
assign (residue 6 and name H5) (residue 6 and name H3') 3.5 0.5 1.5
assign (residue 6 and name H6) (residue 6 and name H1') 4.0 0.5 0.5
assign (residue 6 and name H6) (residue 6 and name H2') 3.5 0.5 0.5
assign (residue 6 and name H6) (residue 6 and name H3') 2.5 0.5 0.9
assign (residue 6 and name H6) (residue 6 and name H5') 3.5 0.5 1.0
assign (residue 6 and name H6) (residue 6 and name H5'') 3.5 0.5 1.0
assign (residue 7 and name H8) (residue 7 and name H1') 3.0 0.5 0.8
assign (residue 7 and name H8) (residue 7 and name H2') 3.5 0.5 0.5
assign (residue 7 and name H8) (residue 7 and name H3') 3.0 0.5 0.8
assign (residue 8 and name H2) (residue 8 and name H1') 4.5 0.5 0.5
assign (residue 8 and name H2) (residue 8 and name H2') 4.5 0.5 0.5
assign (residue 8 and name H8) (residue 8 and name H1') 4.0 0.5 1.0
```

```
assign (residue 8 and name H8) (residue 8 and name H2') 3.5 0.5 1.5
assign (residue 8 and name H8) (residue 8 and name H3') 3.5 0.5 1.5
assign (residue 9 and name H2) (residue 9 and name H1') 3.0 0.5 1.0
assign (residue 9 and name H8) (residue 9 and name H1') 3.5 0.5 1.5
assign (residue 9 and name H8) (residue 9 and name H2') 3.0 0.5 1.2
assign (residue 9 and name H8) (residue 9 and name H3') 3.0 0.5 1.0
assign (residue 10 and name H8) (residue 10 and name H1') 3.0 0.5 1.0
assign (residue 10 and name H8) (residue 10 and name H2') 3.5 1.0 1.0
assign (residue 10 and name H8) (residue 10 and name H3') 3.0 0.5 0.8
assign (residue 11 and name H5) (residue 11 and name H3') 3.5 0.5 1.5
assign (residue 11 and name H6) (residue 11 and name H1') 4.5 0.5 0.5
assign (residue 11 and name H6) (residue 11 and name H2') 3.0 0.5 0.8
assign (residue 11 and name H6) (residue 11 and name H3') 3.0 0.5 0.8
assign (residue 12 and name H8) (residue 12 and name H1') 3.5 0.5 1.5
assign (residue 12 and name H8) (residue 12 and name H2') 3.5 0.5 1.5
assign (residue 12 and name H8) (residue 12 and name H3') 3.0 0.5 0.8
assign (residue 13 and name H6) (residue 13 and name H1') 3.5 0.5 1.5
assign (residue 13 and name H6) (residue 13 and name H2') 3.0 0.5 1.0
assign (residue 13 and name H6) (residue 13 and name H3') 3.0 0.5 1.0
assign (residue 13 and name H5) (residue 13 and name H3') 4.5 0.5 0.5
assign (residue 14 and name H6) (residue 14 and name H1') 3.5 0.5 1.5
assign (residue 14 and name H6) (residue 14 and name H2') 3.0 0.5 0.9
assign (residue 14 and name H6) (residue 14 and name H3') 2.5 0.5 0.9
assign (residue 14 and name H5) (residue 14 and name H3') 4.5 0.5 0.5

! Inter Residue
assign (residue 1 and name H1') (residue 2 and name H8) 4.5 0.5 0.5
assign (residue 1 and name H2') (residue 2 and name H8) 2.0 0.2 0.9
assign (residue 2 and name H2') (residue 3 and name H5) 3.0 0.5 1.5
```

```
assign (residue 2 and name H3') (residue 3 and name H5) 3.0 0.5 1.5
assign (residue 2 and name H1') (residue 3 and name H6) 3.5 0.5 1.5
assign (residue 2 and name H2') (residue 3 and name H6) 2.5 0.5 0.8
assign (residue 2 and name H3') (residue 3 and name H6) 3.0 0.5 0.9
assign (residue 3 and name H2') (residue 4 and name H5) 3.5 0.5 0.5
assign (residue 3 and name H3') (residue 4 and name H5) 3.5 0.5 0.5
assign (residue 3 and name H1') (residue 4 and name H6) 4.0 0.5 0.9
assign (residue 3 and name H2') (residue 4 and name H1') 4.0 1.0 1.0
assign (residue 3 and name H2') (residue 4 and name H6) 2.5 0.5 0.8
assign (residue 3 and name H3') (residue 4 and name H6) 3.0 0.5 0.9
assign (residue 4 and name H1') (residue 5 and name H8) 4.0 0.5 1.5
assign (residue 4 and name H2') (residue 5 and name H1') 4.5 0.5 0.5
assign (residue 4 and name H2') (residue 5 and name H8) 3.0 0.5 0.8
assign (residue 5 and name H2') (residue 6 and name H5) 4.0 1.0 1.0
assign (residue 5 and name H1') (residue 6 and name H1') 4.5 0.0 0.5
assign (residue 5 and name H1') (residue 6 and name H6) 3.5 0.5 1.5
assign (residue 5 and name H2') (residue 6 and name H6) 3.5 1.5 1.5
assign (residue 5 and name H3') (residue 6 and name H6) 4.0 0.5 1.0
assign (residue 6 and name H2') (residue 7 and name H8) 4.0 0.5 0.5
assign (residue 6 and name H3') (residue 7 and name H8) 4.0 0.5 1.0
assign (residue 6 and name H1') (residue 8 and name H8) 4.5 0.5 0.5
assign (residue 6 and name H2') (residue 8 and name H8) 4.0 1.0 1.0
assign (residue 6 and name H3') (residue 8 and name H8) 4.0 0.5 1.5
assign (residue 6 and name H1') (residue 9 and name H8) 2.5 0.5 0.5
assign (residue 6 and name H2') (residue 9 and name H8) 3.5 1.0 1.0
assign (residue 7 and name H2') (residue 8 and name H1') 3.5 0.5 1.5
assign (residue 7 and name H1') (residue 8 and name H8) 4.0 0.5 1.0
assign (residue 7 and name H2') (residue 8 and name H8) 3.0 1.5 1.0
assign (residue 7 and name H3') (residue 8 and name H8) 3.0 1.0 1.0
```

```
assign (residue 8 and name H2) (residue 9 and name H1') 3.0 0.5 1.5
assign (residue 8 and name H2') (residue 9 and name H1') 4.0 1.0 1.0
assign (residue 8 and name H2') (residue 9 and name H8) 4.0 0.5 1.0
assign (residue 8 and name H3') (residue 9 and name H8) 3.5 0.5 0.8
assign (residue 8 and name H8) (residue 9 and name H8) 4.5 0.5 0.5
assign (residue 9 and name H2) (residue 10 and name H1') 2.5 0.5 0.5
assign (residue 9 and name H2) (residue 10 and name H2')  4.0 0.5 1.0
assign (residue 9 and name H2) (residue 10 and name H3') 4.5 0.5 0.5
assign (residue 9 and name H2') (residue 10 and name H1') 4.0 0.5 1.0
assign (residue 9 and name H1') (residue 10 and name H8) 3.5 0.5 1.5
assign (residue 9 and name H2') (residue 10 and name H8) 2.5 0.5 0.9
assign (residue 9 and name H3') (residue 10 and name H8) 3.5 1.0 1.0
assign (residue 10 and name H2) (residue 11 and name H1') 3.0 0.5 0.8
assign (residue 10 and name H2') (residue 11 and name H5) 4.5 0.5 0.5
assign (residue 10 and name H3') (residue 11 and name H5) 3.5 0.5 1.0
assign (residue 10 and name H1') (residue 11 and name H6) 4.0 0.5 1.5
assign (residue 10 and name H2') (residue 11 and name H6) 2.5 0.5 0.8
assign (residue 10 and name H3') (residue 11 and name H6) 3.0 0.5 1.5
assign (residue 11 and name H1') (residue 12 and name H8) 4.0 1.0 1.0
assign (residue 11 and name H2') (residue 12 and name H8) 2.5 0.5 0.5
assign (residue 11 and name H3') (residue 12 and name H8) 2.5 0.5 0.9
assign (residue 12 and name H2') (residue 13 and name H1') 3.0 0.5 0.5
assign (residue 12 and name H1') (residue 13 and name H5) 4.5 0.5 1.0
assign (residue 12 and name H2') (residue 13 and name H5) 3.0 0.5 0.9
assign (residue 12 and name H3') (residue 13 and name H5) 3.0 0.5 1.5
assign (residue 12 and name H1') (residue 13 and name H6) 3.5 0.5 1.5
assign (residue 12 and name H2') (residue 13 and name H6) 2.0 0.2 0.9
assign (residue 12 and name H3') (residue 13 and name H6) 3.0 0.5 0.5
assign (residue 13 and name H2') (residue 14 and name H1') 3.5 0.5 1.5
```

```
assign (residue 13 and name H1') (residue 14 and name H6) 3.5 0.5 1.5
assign (residue 13 and name H2') (residue 14 and name H6) 2.5 0.5 0.9
assign (residue 13 and name H3') (residue 14 and name H6) 3.0 0.5 0.9

! added Unobserved-NOEs
assign (residue 1 and name H5') (residue 6 and name H1') 4.5 0.0 50.0
assign (residue 1 and name H5') (residue 6 and name H5) 4.5 0.0 50.0
assign (residue 1 and name H5') (residue 8 and name H1') 4.5 0.0 50.0
assign (residue 1 and name H5') (residue 8 and name H8) 4.5 0.0 50.0
assign (residue 1 and name H5') (residue 6 and name H2') 4.5 0.0 50.0
assign (residue 1 and name H5'') (residue 8 and name H1') 4.5 0.0 50.0
assign (residue 1 and name H8) (residue 8 and name H4') 4.5 0.0 50.0
assign (residue 1 and name H8) (residue 8 and name H1') 4.5 0.0 50.0
assign (residue 1 and name H8) (residue 9 and name H5') 4.5 0.0 50.0
assign (residue 2 and name H8) (residue 4 and name H5) 4.5 0.0 50.0
assign (residue 3 and name H4') (residue 5 and name H8) 4.5 0.0 50.0
assign (residue 3 and name H6) (residue 4 and name H1') 4.5 0.0 50.0
assign (residue 3 and name H6) (residue 5 and name H5') 4.5 0.0 50.0
assign (residue 3 and name H5) (residue 5 and name H8) 4.5 0.0 50.0
assign (residue 3 and name H5) (residue 5 and name H4') 4.5 0.0 50.0
assign (residue 3 and name H5) (residue 5 and name H5') 4.5 0.0 50.0
assign (residue 3 and name H5) (residue 6 and name H5) 4.5 0.0 50.0
assign (residue 3 and name H5) (residue 11 and name H4') 4.5 0.0 50.0
assign (residue 4 and name H3') (residue 6 and name H5) 4.5 0.0 50.0
assign (residue 4 and name H4') (residue 9 and name H2) 4.5 0.0 50.0
assign (residue 4 and name H4') (residue 10 and name H2) 4.5 0.0 50.0
assign (residue 4 and name H5) (residue 6 and name H5) 4.5 0.0 50.0
assign (residue 4 and name H6) (residue 5 and name H2') 4.5 0.0 50.0
assign (residue 4 and name H6) (residue 6 and name H5) 4.5 0.0 50.0
```

```
assign (residue 4 and name H2') (residue 6 and name H6) 4.5 0.0 50.0
assign (residue 4 and name H1') (residue 6 and name H5)  4.5 0.0 50.0
assign (residue 4 and name H1') (residue 10 and name H2) 4.5 0.0 50.0
assign (residue 4 and name H2') (residue 9 and name H2) 4.5 0.0 50.0
assign (residue 4 and name H2') (residue 10 and name H2) 4.5 0.0 50.0
assign (residue 4 and name H5) (residue 10 and name H2) 4.5 0.0 50.0
assign (residue 4 and name H5) (residue 10 and name H4') 4.5 0.0 50.0
assign (residue 4 and name H5) (residue 11 and name H1') 4.5 0.0 50.0
assign (residue 5 and name H1') (residue 6 and name H1') 4.5 0.0 50.0
assign (residue 5 and name H1') (residue 9 and name H2) 4.5 0.0 50.0
assign (residue 5 and name H1') (residue 10 and name H2) 4.5 0.0 50.0
assign (residue 5 and name H1') (residue 10 and name H8) 4.5 0.0 50.0
assign (residue 5 and name H1') (residue 11 and name H1') 4.5 0.0 50.0
assign (residue 5 and name H2') (residue 6 and name H2') 5.0 0.0 50.0
assign (residue 5 and name H2') (residue 7 and name H3') 4.5 0.0 50.0
assign (residue 5 and name H2') (residue 9 and name H2) 4.5 0.0 50.0
assign (residue 5 and name H2') (residue 10 and name H2) 4.5 0.0 50.0
assign (residue 5 and name H4') (residue 9 and name H2) 4.5 0.0 50.0
assign (residue 5 and name H4') (residue 10 and name H2) 4.5 0.0 50.0
assign (residue 5 and name H4') (residue 11 and name H1') 4.5 0.0 50.0
assign (residue 5 and name H5') (residue 6 and name H1') 4.5 0.0 50.0
assign (residue 5 and name H5') (residue 6 and name H5) 4.5 0.0 50.0
assign (residue 5 and name H8) (residue 6 and name H1') 4.5 0.0 50.0
assign (residue 5 and name H8) (residue 8 and name H2) 4.5 0.0 50.0
assign (residue 5 and name H8) (residue 9 and name H2) 4.5 0.0 50.0
assign (residue 5 and name H8) (residue 10 and name H2) 4.5 0.0 50.0
assign (residue 5 and name H5') (residue 9 and name H8) 4.5 0.0 50.0
assign (residue 6 and name H1') (residue 7 and name H4') 4.5 0.0 50.0
assign (residue 6 and name H1') (residue 7 and name H8) 4.5 0.0 50.0
```

```
assign (residue 6 and name H2') (residue 8 and name H2) 4.5 0.0 50.0
assign (residue 6 and name H1') (residue 8 and name H5'') 4.5 0.0 50.0
assign (residue 6 and name H1') (residue 8 and name H3') 4.5 0.0 50.0
assign (residue 6 and name H1') (residue 8 and name H2) 4.5 0.0 50.0
assign (residue 6 and name H1') (residue 9 and name H2') 4.5 0.0 50.0
assign (residue 6 and name H1') (residue 10 and name H1') 4.5 0.0 50.0
assign (residue 6 and name H1') (residue 10 and name H8) 4.5 0.0 50.0
assign (residue 6 and name H1') (residue 10 and name H2) 4.5 0.0 50.0
assign (residue 6 and name H1') (residue 9 and name H5') 4.5 0.0 50.0
assign (residue 6 and name H1') (residue 9 and name H4') 4.5 0.0 50.0
assign (residue 6 and name H1') (residue 9 and name H1') 4.5 0.0 50.0
assign (residue 6 and name H1') (residue 12 and name H5') 4.5 0.0 50.0
assign (residue 6 and name H3') (residue 10 and name H2) 4.5 0.0 50.0
assign (residue 6 and name H4') (residue 8 and name H8) 4.5 0.0 50.0
assign (residue 6 and name H4') (residue 9 and name H8) 4.5 0.0 50.0
assign (residue 6 and name H4') (residue 9 and name H2) 4.5 0.0 50.0
assign (residue 6 and name H4') (residue 10 and name H2) 4.5 0.0 50.0
assign (residue 6 and name H5) (residue 10 and name H2) 4.5 0.0 50.0
assign (residue 6 and name H5') (residue 9 and name H2) 4.5 0.0 50.0
assign (residue 6 and name H5') (residue 10 and name H2) 4.5 0.0 50.0
assign (residue 6 and name H5') (residue 11 and name H1') 4.5 0.0 50.0
assign (residue 6 and name H5) (residue 8 and name H8) 4.5 0.0 50.0
assign (residue 6 and name H5) (residue 9 and name H2) 4.5 0.0 50.0
assign (residue 6 and name H6) (residue 8 and name H1') 4.5 0.0 50.0
assign (residue 6 and name H6) (residue 8 and name H2') 4.5 0.0 50.0
assign (residue 6 and name H6) (residue 8 and name H8) 4.5 0.0 50.0
assign (residue 6 and name H6) (residue 9 and name H2) 4.5 0.0 50.0
assign (residue 6 and name H6) (residue 9 and name H8) 4.5 0.0 50.0
assign (residue 6 and name H6) (residue 10 and name H2) 4.5 0.0 50.0
```

```
assign (residue 6 and name H6) (residue 11 and name H5) 4.5 0.0 50.0
assign (residue 7 and name H8) (residue 6 and name H4') 4.5 0.0 50.0
assign (residue 7 and name H8) (residue 8 and name H2) 4.5 0.0 50.0
assign (residue 7 and name H8) (residue 8 and name H4') 4.5 0.0 50.0
assign (residue 7 and name H1') (residue 9 and name H8) 4.5 0.0 50.0
assign (residue 7 and name H2') (residue 9 and name H8) 4.5 0.0 50.0
assign (residue 7 and name H2') (residue 10 and name H2) 4.5 0.0 50.0
assign (residue 7 and name H4') (residue 8 and name H1') 4.5 0.0 50.0
assign (residue 7 and name H4') (residue 9 and name H2') 4.5 0.0 50.0
assign (residue 7 and name H8) (residue 9 and name H8) 4.5 0.0 50.0
assign (residue 7 and name H8) (residue 9 and name H4') 4.5 0.0 50.0
assign (residue 7 and name H8) (residue 10 and name H8) 4.5 0.0 50.0
assign (residue 7 and name H8) (residue 10 and name H2) 4.5 0.0 50.0
assign (residue 8 and name H1') (residue 9 and name H1') 4.5 0.0 50.0
assign (residue 8 and name H1') (residue 9 and name H2') 4.5 0.0 50.0
assign (residue 8 and name H1') (residue 9 and name H4') 4.5 0.0 50.0
assign (residue 8 and name H1') (residue 9 and name H5') 4.5 0.0 50.0
assign (residue 8 and name H1') (residue 10 and name H8) 4.5 0.0 50.0
assign (residue 8 and name H2') (residue 10 and name H2) 4.5 0.0 50.0
assign (residue 8 and name H4') (residue 9 and name H8) 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 9 and name H2) 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 9 and name H8) 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 10 and name H1') 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 10 and name H2') 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 10 and name H8) 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 11 and name H1') 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 11 and name H2') 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 11 and name H5) 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 11 and name H6) 4.5 0.0 50.0
```

```
assign (residue 8 and name H2) (residue 12 and name H8) 4.5 0.0 50.0
assign (residue 8 and name H8) (residue 9 and name H1') 4.5 0.0 50.0
assign (residue 8 and name H8) (residue 9 and name H2') 4.5 0.0 50.0
assign (residue 8 and name H8) (residue 9 and name H3') 4.5 0.0 50.0
assign (residue 8 and name H8) (residue 9 and name H4') 4.5 0.0 50.0
assign (residue 8 and name H8) (residue 9 and name H5') 4.5 0.0 50.0
assign (residue 9 and name H2') (residue 11 and name H5) 4.5 0.0 50.0
assign (residue 9 and name H3') (residue 11 and name H5) 4.5 0.0 50.0
assign (residue 9 and name H2') (residue 11 and name H6) 4.5 0.0 50.0
assign (residue 9 and name H2) (residue 11 and name H5) 4.5 0.0 50.0
assign (residue 9 and name H2) (residue 12 and name H8) 4.5 0.0 50.0
assign (residue 9 and name H2) (residue 12 and name H5') 4.5 0.0 50.0
assign (residue 9 and name H2) (residue 12 and name H4') 4.5 0.0 50.0
assign (residue 9 and name H8) (residue 12 and name H5') 4.5 0.0 50.0
assign (residue 9 and name H4') (residue 10 and name H8) 4.5 0.0 50.0
assign (residue 9 and name H4') (residue 11 and name H5) 4.5 0.0 50.0
assign (residue 9 and name H5') (residue 10 and name H1') 4.5 0.0 50.0
assign (residue 9 and name H8) (residue 10 and name H1') 4.5 0.0 50.0
assign (residue 9 and name H2) (residue 11 and name H5') 4.5 0.0 50.0
assign (residue 10 and name H2) (residue 11 and name H5') 4.5 0.0 50.0
assign (residue 10 and name H2) (residue 11 and name H2') 4.5 0.0 50.0
assign (residue 10 and name H2) (residue 11 and name H6) 4.5 0.0 50.0
assign (residue 10 and name H2) (residue 12 and name H3') 4.5 0.0 50.0
assign (residue 10 and name H2) (residue 13 and name H5) 4.5 0.0 50.0
assign (residue 10 and name H2) (residue 13 and name H6) 4.5 0.0 50.0
assign (residue 10 and name H2) (residue 14 and name H5) 4.5 0.0 50.0
assign (residue 10 and name H1') (residue 11 and name H1') 4.5 0.0 50.0
assign (residue 10 and name H3') (residue 11 and name H1') 4.5 0.0 50.0
assign (residue 10 and name H3') (residue 12 and name H8) 4.5 0.0 50.0
```

```
assign (residue 10 and name H1') (residue 12 and name H8) 4.5 0.0 50.0
assign (residue 10 and name H8) (residue 12 and name H8) 4.5 0.0 50.0
assign (residue 10 and name H2) (residue 11 and name H3') 4.5 0.0 50.0
assign (residue 10 and name H2) (residue 11 and name H5) 4.5 0.0 50.0
assign (residue 10 and name H2) (residue 12 and name H5') 4.5 0.0 50.0
assign (residue 10 and name H2) (residue 12 and name H4') 4.5 0.0 50.0
assign (residue 10 and name H2) (residue 12 and name H8) 4.5 0.0 50.0
assign (residue 11 and name H4') (residue 12 and name H8) 4.5 0.0 50.0
assign (residue 11 and name H3') (residue 13 and name H5) 4.5 0.0 50.0
assign (residue 11 and name H5) (residue 12 and name H4') 4.5 0.0 50.0
assign (residue 11 and name H5) (residue 12 and name H5') 4.5 0.0 50.0
assign (residue 11 and name H5) (residue 12 and name H1') 4.5 0.0 50.0
assign (residue 11 and name H6) (residue 12 and name H4') 4.5 0.0 50.0
assign (residue 11 and name H6) (residue 12 and name H5') 4.5 0.0 50.0
assign (residue 11 and name H6) (residue 13 and name H5) 4.5 0.0 50.0
assign (residue 11 and name H6) (residue 12 and name H1') 4.5 0.0 50.0
assign (residue 11 and name H1') (residue 14 and name H5) 4.5 0.0 50.0
assign (residue 11 and name H5) (residue 12 and name H3') 4.5 0.0 50.0
assign (residue 11 and name H5) (residue 13 and name H6) 4.5 0.0 50.0
assign (residue 11 and name H6) (residue 13 and name H5') 4.5 0.0 50.0
assign (residue 12 and name H5') (residue 13 and name H5) 4.5 0.0 50.0
assign (residue 12 and name H1') (residue 13 and name H1') 4.5 0.0 50.0
```

Dihedral angle restraints for the 690UC mutant

```
!sugar pucker delta
assign (resid 1 and name c5') (resid 1 and name c4')
       (resid 1 and name c3') (resid 1 and name o3') 1 80 5 2
assign (resid 2 and name c5') (resid 2 and name c4')
       (resid 2 and name c3') (resid 2 and name o3') 1 80 5 2
assign (resid 3 and name c5') (resid 3 and name c4')
```

```
       (resid 3 and name c3') (resid 3 and name o3') 1 80 5 2
assign (resid 4 and name c5') (resid 4 and name c4')
       (resid 4 and name c3') (resid 4 and name o3') 1 80 20 2
assign (resid 5 and name c5') (resid 5 and name c4')
       (resid 5 and name c3') (resid 5 and name o3') 1 80 20 2
assign (resid 6 and name c5') (resid 6 and name c4')
       (resid 6 and name c3') (resid 6 and name o3') 1 80 20 2
assign (resid 7 and name c5') (resid 7 and name c4')
       (resid 7 and name c3') (resid 7 and name o3') 1 80 20 2
assign (resid 9 and name c5') (resid 9 and name c4')
       (resid 9 and name c3') (resid 9 and name o3') 1 80 20 2
assign (resid 10 and name c5') (resid 10 and name c4')
       (resid 10 and name c3') (resid 10 and name o3') 1 80 20 2
assign (resid 11 and name c5') (resid 11 and name c4')
       (resid 11 and name c3') (resid 11 and name o3') 1 80 20 2
assign (resid 12 and name c5') (resid 12 and name c4')
       (resid 12 and name c3') (resid 12 and name o3') 1 80 5 2
assign (resid 13 and name c5') (resid 13 and name c4')
       (resid 13 and name c3') (resid 13 and name o3') 1 80 5 2
assign (resid 14 and name c5') (resid 14 and name c4')
       (resid 14 and name c3') (resid 14 and name o3') 1 80 5 2

!zeta and alpha
assign (resid 1 and name c3') (resid 1 and name o3')
       (resid 2 and name p) (resid 2 and name o5') 1 -68 5 2
assign (resid 1 and name o3') (resid 2 and name p)
       (resid 2 and name o5') (resid 2 and name c5') 1 -71 5 2

assign (resid 2 and name c3') (resid 2 and name o3')
```

```
       (resid 3 and name p) (resid 3 and name o5') 1 -68 5 2
assign (resid 2 and name o3') (resid 3 and name p)
       (resid 3 and name o5') (resid 3 and name c5') 1 -71 5 2

assign (resid 3 and name c3') (resid 3 and name o3')
       (resid 4 and name p) (resid 4 and name o5') 1 0 120 2
assign (resid 3 and name o3') (resid 4 and name p)
       (resid 4 and name o5') (resid 4 and name c5') 1 0 120 2

assign (resid 4 and name c3') (resid 4 and name o3')
       (resid 5 and name p) (resid 5 and name o5') 1 0 120 2
assign (resid 4 and name o3') (resid 5 and name p)
       (resid 5 and name o5') (resid 5 and name c5') 1 0 120 2

assign (resid 5 and name c3') (resid 5 and name o3')
       (resid 6 and name p) (resid 6 and name o5') 1 0 120 2
assign (resid 5 and name o3') (resid 6 and name p)
       (resid 6 and name o5') (resid 6 and name c5') 1 0 120 2

assign (resid 6 and name c3') (resid 6 and name o3')
       (resid 7 and name p) (resid 7 and name o5') 1 0 120 2
assign (resid 6 and name o3') (resid 7 and name p)
       (resid 7 and name o5') (resid 7 and name c5') 1 0 120 2

assign (resid 7 and name c3') (resid 7 and name o3')
       (resid 8 and name p) (resid 8 and name o5') 1 0 120 2
assign (resid 7 and name o3') (resid 8 and name p)
       (resid 8 and name o5') (resid 8 and name c5') 1 0 120 2
```

```
assign (resid 8 and name c3') (resid 8 and name o3')
       (resid 9 and name p) (resid 9 and name o5') 1 0 120 2
assign (resid 8 and name o3') (resid 9 and name p)
       (resid 9 and name o5') (resid 9 and name c5') 1 0 120 2

assign (resid 9 and name c3') (resid 9 and name o3')
       (resid 10 and name p) (resid 10 and name o5') 1 0 120 2
assign (resid 9 and name o3') (resid 10 and name p)
       (resid 10 and name o5') (resid 10 and name c5') 1 0 120 2

assign (resid 10 and name c3') (resid 10 and name o3')
       (resid 11 and name p) (resid 11 and name o5') 1 0 120 2
assign (resid 10 and name o3') (resid 11 and name p)
       (resid 11 and name o5') (resid 11 and name c5') 1 0 120 2
assign (resid 11 and name c3') (resid 11 and name o3')
       (resid 12 and name p) (resid 12 and name o5') 1 0 120 2
assign (resid 11 and name o3') (resid 12 and name p)
       (resid 12 and name o5') (resid 12 and name c5') 1 0 120 2

assign (resid 12 and name c3') (resid 12 and name o3')
       (resid 13 and name p) (resid 13 and name o5') 1 -68 5 2
assign (resid 12 and name o3') (resid 13 and name p)
       (resid 13 and name o5') (resid 13 and name c5') 1 -71 5 2

assign (resid 13 and name c3') (resid 13 and name o3')
       (resid 14 and name p) (resid 14 and name o5') 1 -68 5 2
assign (resid 13 and name o3') (resid 14 and name p)
       (resid 14 and name o5') (resid 14 and name c5') 1 -71 5 2
```

```
!beta
assign (resid 2 and name p) (resid 2 and name o5')
       (resid 2 and name c5') (resid 2 and name c4') 1 180 5 2
assign (resid 3 and name p) (resid 3 and name o5')
       (resid 3 and name c5') (resid 3 and name c4') 1 180 5 2
assign (resid 4 and name p) (resid 4 and name o5')
       (resid 4 and name c5') (resid 4 and name c4') 1 180 60 2
assign (resid 5 and name p) (resid 5 and name o5')
       (resid 5 and name c5') (resid 5 and name c4') 1 180 60 2
assign (resid 6 and name p) (resid 6 and name o5')
       (resid 6 and name c5') (resid 6 and name c4') 1 180 60 2
assign (resid 7 and name p) (resid 7 and name o5')
       (resid 7 and name c5') (resid 7 and name c4') 1 180 60 2
assign (resid 8 and name p) (resid 8 and name o5')
       (resid 8 and name c5') (resid 8 and name c4') 1 180 60 2
assign (resid 9 and name p) (resid 9 and name o5')
       (resid 9 and name c5') (resid 9 and name c4') 1 180 60 2
assign (resid 10 and name p) (resid 10 and name o5')
       (resid 10 and name c5') (resid 10 and name c4') 1 180 60 2
assign (resid 11 and name p) (resid 11 and name o5')
       (resid 11 and name c5') (resid 11 and name c4') 1 180 60 2
assign (resid 12 and name p) (resid 12 and name o5')
       (resid 12 and name c5') (resid 12 and name c4') 1 180 60 2
assign (resid 13 and name p) (resid 13 and name o5')
       (resid 13 and name c5') (resid 13 and name c4') 1 180 5 2
assign (resid 14 and name p) (resid 14 and name o5')
       (resid 14 and name c5') (resid 14 and name c4') 1 180 5 2

!gamma
```

```
assign (resid 1 and name o5') (resid 1 and name c5')
       (resid 1 and name c4') (resid 1 and name c3') 1 54 10 2
assign (resid 2 and name o5') (resid 2 and name c5')
       (resid 2 and name c4') (resid 2 and name c3') 1 54 10 2
assign (resid 3 and name o5') (resid 3 and name c5')
       (resid 3 and name c4') (resid 3 and name c3') 1 54 10 2
assign (resid 4 and name o5') (resid 4 and name c5')
       (resid 4 and name c4') (resid 4 and name c3') 1 60 30 2
assign (resid 5 and name o5') (resid 5 and name c5')
       (resid 5 and name c4') (resid 5 and name c3') 1 60 30 2
assign (resid 6 and name o5') (resid 6 and name c5')
       (resid 6 and name c4') (resid 6 and name c3') 1 60 30 2
assign (resid 7 and name o5') (resid 7 and name c5')
       (resid 7 and name c4') (resid 7 and name c3') 1 60 30 2
assign (resid 8 and name o5') (resid 8 and name c5')
       (resid 8 and name c4') (resid 8 and name c3') 1 60 30 2
assign (resid 10 and name o5') (resid 10 and name c5')
       (resid 10 and name c4') (resid 10 and name c3') 1 60 30 2
assign (resid 11 and name o5') (resid 11 and name c5')
       (resid 11 and name c4') (resid 11 and name c3') 1 60 30 2
assign (resid 12 and name o5') (resid 12 and name c5')
       (resid 12 and name c4') (resid 12 and name c3') 1 60 30 2
assign (resid 13 and name o5') (resid 13 and name c5')
       (resid 13 and name c4') (resid 13 and name c3') 1 54 10 2
assign (resid 14 and name o5') (resid 14 and name c5')
       (resid 14 and name c4') (resid 14 and name c3') 1 54 10 2

!epsilon
assign (resid 1 and name c4') (resid 1 and name c3')
```

```
       (resid 1 and name o3') (resid 2 and name p) 1 -155 5 2
assign (resid 2 and name c4') (resid 2 and name c3')
       (resid 2 and name o3') (resid 3 and name p) 1 -155 5 2
assign (resid 12 and name c4') (resid 12 and name c3')
       (resid 12 and name o3') (resid 13 and name p) 1 -155 5 2
assign (resid 13 and name c4') (resid 13 and name c3')
       (resid 13 and name o3') (resid 14 and name p) 1 -155 5 2
assign (resid 14 and name c4') (resid 14 and name c3')
       (resid 14 and name o3') (resid 15 and name p) 1 -155 5 2

! Chi
assign (resid 1 and name O4') (resid 1 and name C1')
       (resid 1 and name N9) (resid 1 and name C4) 1 -155 5 2

assign (resid 2 and name O4') (resid 2 and name C1')
       (resid 2 and name N9) (resid 2 and name C4) 1 -155 5 2

assign (resid 3 and name O4') (resid 3 and name C1')
       (resid 3 and name N1) (resid 3 and name C2) 1 -155 30 2

assign (resid 4 and name O4') (resid 4 and name C1')
       (resid 4 and name N1) (resid 4 and name C2) 1 -155 45 2

assign (resid 5 and name O4') (resid 5 and name C1')
       (resid 5 and name N9) (resid 5 and name C4) 1 -155 45 2

assign (resid 6 and name O4') (resid 6 and name C1')
       (resid 6 and name N1) (resid 6 and name C2) 1 -155 45 2
```

```
assign (resid 7 and name O4') (resid 7 and name C1')
       (resid 7 and name N9) (resid 7 and name C4) 1 -155 45 2

assign (resid 8 and name O4') (resid 8 and name C1')
       (resid 8 and name N9) (resid 8 and name C4) 1 -155 45 2

assign (resid 9 and name O4') (resid 9 and name C1')
       (resid 9 and name N9) (resid 9 and name C4) 1 -155 45 2

assign (resid 10 and name O4') (resid 10 and name C1')
       (resid 10 and name N9) (resid 10 and name C4) 1 -155 45 2

assign (resid 11 and name O4') (resid 11 and name C1')
       (resid 11 and name N1) (resid 11 and name C2) 1 -155 45 2

assign (resid 12 and name O4') (resid 12 and name C1')
       (resid 12 and name N9) (resid 12 and name C4) 1 -155 30 2

assign (resid 13 and name O4') (resid 13 and name C1')
       (resid 13 and name N1) (resid 13 and name C2) 1 -155 5 2

assign (resid 14 and name O4') (resid 14 and name C1')
       (resid 14 and name N1) (resid 14 and name C2) 1 -155 5 2
```

APPENDIX 2

Structural Restraints for the 690QM Mutant

Distance restraints for the 690QM mutant (modified based on Ravi's distance restraints for the 690QM mutant)

```
!stem stacking distances
assign (residue 1 and name h1) (residue 2 and name h1) 3.5 0.5 0.5
assign (residue 2 and name h1) (residue 12 and name h1) 4.0 1.0 0.0

! Base pair Restraints
assign (residue 1 and name C1') (residue 14 and name C1') 10.6 0.2 0.2
assign (residue 1 and name C8) (residue 14 and name C6) 9.8 0.3 0.3
assign (residue 2 and name C1') (residue 13 and name C1') 10.6 0.2 0.2
assign (residue 2 and name C8) (residue 13 and name C6) 9.8 0.3 0.3
assign (residue 3 and name C1') (residue 12 and name C1') 10.6 0.2 0.2
assign (residue 3 and name C6) (residue 12 and name C8) 9.8 0.3 0.3
assign (residue 1 and name O6) (residue 2 and name O6) 3.0 0.3 0.3
assign (residue 1 and name N2) (residue 2 and name N2) 4.5 0.5 0.5
assign (residue 2 and name O6) (residue 3 and name N4) 3.0 0.6 0.6
assign (residue 2 and name N2) (residue 3 and name O2) 4.5 1.0 1.0
assign (residue 12 and name O6) (residue 13 and name N4) 3.0 0.5 0.5
assign (residue 12 and name N2) (residue 13 and name O2) 4.5 1.0 1.0
assign (residue 13 and name N4) (residue 14 and name N4) 3.0 0.3 0.3
assign (residue 13 and name O2) (residue 14 and name O2) 4.5 0.5 0.5
```

```
!Base pair H-bond distances
! for G1/ C14  base pair
assign (resid 1 and name N1) (resid 14 and name N3) 2.91 0.3 0.3
assign (resid 1 and name O6) (resid 14 and name N4) 2.71 0.3 0.3
assign (resid 1 and name N2) (resid 14 and name O2) 3.00 0.3 0.3
assign (resid 1 and name H1) (resid 14 and name N3) 1.89 0.3 0.3
assign (resid 1 and name O6) (resid 14 and name H42) 1.71 0.3 0.3
assign (resid 1 and name H22) (resid 14 and name O2) 1.90 0.3 0.3

! for G2/ C13  base pair
assign (resid 2 and name N1) (resid 13 and name N3) 2.91 0.3 0.3
assign (resid 2 and name O6) (resid 13 and name N4) 2.71 0.3 0.3
assign (resid 2 and name N2) (resid 13 and name O2) 3.00 0.3 0.3
assign (resid 2 and name H1) (resid 13 and name N3) 1.89 0.3 0.3
assign (resid 2 and name O6) (resid 13 and name H42) 1.71 0.3 0.3
assign (resid 2 and name H22) (resid 13 and name O2) 1.90 0.3 0.3

! for C3/ G12  base pair
assign (resid 3 and name N3) (resid 12 and name N1) 2.91 0.3 0.3
assign (resid 3 and name N4) (resid 12 and name O6) 2.71 0.3 0.3
assign (resid 3 and name O2) (resid 12 and name N2) 3.00 0.3 0.3
assign (resid 3 and name N3) (resid 12 and name H1) 1.89 0.3 0.3
assign (resid 3 and name H42) (resid 12 and name O6) 1.71 0.3 0.3
assign (resid 3 and name O2) (resid 12 and name H22) 1.90 0.3 0.3

! NOEs from H2O NOESY at 5C
assign (residue 1 and name H1) (residue 14 and name H42) 3.0 1.0 0.5
assign (residue 2 and name H1) (residue 12 and name H1) 4.0 0.5 0.5
assign (residue 2 and name H1) (residue 13 and name H42) 3.0 1.0 0.5
```

```
assign (residue 2 and name H1) (residue 13 and name H41) 4.0 0.5 0.5
assign (residue 3 and name H41) (residue 12 and name H1)4.0 0.5 0.5
assign (residue 3 and name H42) (residue 12 and name H1)3.0 1.0 0.5
assign (residue 5 and name H1) (residue 6 and name H3) 4.0 0.5 0.8
assign (residue 12 and name H1) (residue 13 and name H41) 4.5 0.5 0.5
assign (residue 12 and name H1) (residue 13 and name H42) 3.0 0.5 0.5

! Residue 1 GUA
! Intra residue
assign (residue 1 and name H8) (residue 1 and name H1') 4.5 0.5 0.5
assign (residue 1 and name H8) (residue 1 and name H2') 4.5 0.5 0.5
assign (residue 1 and name H8) (residue 1 and name H3') 3.0 0.5 0.5
!Inter residue
assign (residue 1 and name H1') (residue 2 and name H8) 4.0 0.5 0.5
assign (residue 1 and name H2') (residue 2 and name H8) 2.5 0.5 0.5
assign (residue 1 and name H3') (residue 2 and name H8) 2.5 0.5 0.5
assign (residue 1 and name H8) (residue 2 and name H8) 4.5 0.5 0.5

! Residue 2 GUA
! Intra Residue
assign (residue 2 and name H8) (residue 2 and name H1') 3.0 0.5 0.8
assign (residue 2 and name H8) (residue 2 and name H2') 3.5 0.5 1.0
assign (residue 2 and name H8) (residue 2 and name H3') 3.0 0.5 0.5
! Inter Residue
assign (residue 2 and name H1') (residue 3 and name H1') 4.5 0.5 0.5
assign (residue 2 and name H1') (residue 3 and name H5) 4.5 0.5 0.5
assign (residue 2 and name H1') (residue 3 and name H6) 4.0 0.0 0.8
assign (residue 2 and name H2') (residue 3 and name H1') 4.0 0.5 0.5
assign (residue 2 and name H2') (residue 3 and name H5) 4.0 0.5 0.5
```

```
assign (residue 2 and name H2') (residue 3 and name H6) 2.0 0.0 1.0
assign (residue 2 and name H3') (residue 3 and name H5) 4.5 0.5 0.5
assign (residue 2 and name H3') (residue 3 and name H6) 3.5 0.5 0.5
assign (residue 2 and name H8) (residue 3 and name H5) 4.5 0.5 0.5
assign (residue 2 and name H8) (residue 3 and name H6) 4.5 0.5 0.5

! Residue 3 CYT
! Intra Residue
assign (residue 3 and name H6) (residue 3 and name H3') 2.5 0.5 0.5
assign (residue 3 and name H6) (residue 3 and name H1') 3.5 0.5 0.5
assign (residue 3 and name H6) (residue 3 and name H2') 3.0 0.5 0.8
! Inter Residue
assign (residue 3 and name H1') (residue 4 and name H1') 4.5 0.5 0.5
assign (residue 3 and name H1') (residue 4 and name H8) 4.0 0.5 1.0
assign (residue 3 and name H2') (residue 4 and name H1') 4.0 0.5 0.5
assign (residue 3 and name H2') (residue 4 and name H8) 3.0 0.5 0.8
assign (residue 3 and name H3') (residue 4 and name H8) 3.0 0.5 0.5
assign (residue 3 and name H5) (residue 4 and name H8) 4.5 0.5 1.0
assign (residue 3 and name H6) (residue 4 and name H8) 4.5 0.5 0.5

! Hairpin Loop begins
! Residue 4 ADE
! Intra Residue
assign (residue 4 and name H2) (residue 4 and name H1') 4.5 0.5 0.5
assign (residue 4 and name H8) (residue 4 and name H1') 4.0 0.5 0.5
assign (residue 4 and name H8) (residue 4 and name H2') 3.0 0.5 0.8
assign (residue 4 and name H8) (residue 4 and name H3') 2.5 0.5 0.5
! Inter Residue
assign (residue 4 and name H1') (residue 5 and name H1') 4.5 0.5 0.5
```

```
assign (residue 4 and name H1') (residue 5 and name H8) 4.0 0.5 0.8
assign (residue 4 and name H2') (residue 5 and name H8) 2.5 0.5 0.8
assign (residue 4 and name H2) (residue 5 and name H1') 3.5 0.5 0.5
assign (residue 4 and name H2) (residue 5 and name H8) 4.5 0.0 1.0
assign (residue 4 and name H2) (residue 11 and name H2) 4.5 0.5 0.5
assign (residue 4 and name H2) (residue 12 and name H1') 4.5 0.5 1.0
assign (residue 4 and name H8) (residue 5 and name H8) 4.5 0.5 0.5

! Residue 5 GUA
! Intra Residue
assign (residue 5 and name H8) (residue 5 and name H1') 3.5 0.5 0.5
assign (residue 5 and name H8) (residue 5 and name H2') 3.0 0.5 0.8
assign (residue 5 and name H8) (residue 5 and name H3') 2.0 0.0 1.0
! Inter Residue
assign (residue 6 and name H1') (residue 5 and name H2') 4.0 0.5 0.5
assign (residue 5 and name H1') (residue 6 and name H5) 4.5 0.5 0.5
assign (residue 5 and name H1') (residue 6 and name H6) 4.5 0.5 0.5
assign (residue 5 and name H1') (residue 11 and name H2) 4.5 0.5 0.8
assign (residue 5 and name H2') (residue 6 and name H5) 4.5 0.5 0.5
assign (residue 5 and name H2') (residue 6 and name H6) 3.0 0.5 0.5
assign (residue 5 and name H3') (residue 6 and name H5) 4.5 0.5 0.5
assign (residue 5 and name H3') (residue 6 and name H6) 4.5 0.7 0.7
assign (residue 5 and name H8) (residue 6 and name H5) 4.5 0.5 1.0
assign (residue 5 and name H8) (residue 6 and name H6) 4.5 0.5 0.5

! Residue 6 URI
! Intra Residue
assign (residue 6 and name H5) (residue 6 and name H2') 4.5 0.5 0.5
assign (residue 6 and name H5) (residue 6 and name H3') 4.5 0.5 0.5
```

```
assign (residue 6 and name H6) (residue 6 and name H1') 4.0 0.5 0.5
assign (residue 6 and name H6) (residue 6 and name H2') 3.0 0.0 1.0
assign (residue 6 and name H6) (residue 6 and name H3') 2.5 0.5 0.5
! Inter Residue
assign (residue 6 and name H1') (residue 8 and name H8) 4.0 0.5 0.5
assign (residue 6 and name H1') (residue 9 and name H5) 2.0 0.0 1.0
assign (residue 6 and name H1') (residue 9 and name H6) 4.0 1.0 1.0
assign (residue 6 and name H2') (residue 7 and name H6) 4.0 0.5 0.5
assign (residue 6 and name H6) (residue 8 and name H8) 4.5 0.5 1.0

! Residue 7 CYT
! Intra Residue
assign (residue 7 and name H6) (residue 7 and name H1') 3.5 0.5 0.5
assign (residue 7 and name H6) (residue 7 and name H2') 3.5 0.5 0.5
assign (residue 7 and name H6) (residue 7 and name H3') 3.5 0.5 0.5
! Inter Residue
assign (residue 7 and name H1') (residue 8 and name H8) 4.0 0.5 0.8
assign (residue 7 and name H2') (residue 8 and name H8) 2.5 0.5 0.5
assign (residue 7 and name H6) (residue 8 and name H8) 4.0 0.5 0.8

! Residue 8 ADE
! Intra Residue
assign (residue 8 and name H2) (residue 8 and name H1') 4.5 0.5 0.5
assign (residue 8 and name H8) (residue 8 and name H1') 3.5 0.5 0.5
assign (residue 8 and name H8) (residue 8 and name H2') 3.0 0.5 1.0
assign (residue 8 and name H8) (residue 8 and name H3') 3.0 0.5 0.5
! Inter Residue
assign (residue 8 and name H1') (residue 9 and name H1') 4.5 0.5 1.0
assign (residue 8 and name H1') (residue 9 and name H6) 4.5 0.5 1.0
```

```
assign (residue 8 and name H2') (residue 9 and name H1') 3.5 0.5 0.5
assign (residue 8 and name H2) (residue 9 and name H6) 4.5 0.5 1.0
assign (residue 8 and name H2') (residue 9 and name H5) 4.5 0.5 1.0
assign (residue 8 and name H2') (residue 9 and name H6) 3.0 0.5 0.5
assign (residue 8 and name H3') (residue 9 and name H6) 4.0 0.5 0.5
assign (residue 8 and name H2) (residue 9 and name H1') 3.0 0.5 0.5
assign (residue 8 and name H8) (residue 9 and name H6) 4.5 0.5 0.5

! Residue 9 CYT
! Inter residue
assign (residue 9 and name H5) (residue 9 and name H2') 4.5 0.5 0.5
assign (residue 9 and name H5) (residue 9 and name H3') 4.5 0.5 0.5
assign (residue 9 and name H6) (residue 9 and name H1') 3.0 0.5 0.8
assign (residue 9 and name H6) (residue 9 and name H2') 3.0 0.5 0.8
assign (residue 9 and name H6) (residue 9 and name H3') 2.0 0.0 1.0
! Inter Residue
assign (residue 9 and name H1') (residue 10 and name H1')4.5 0.5 1.0
assign (residue 9 and name H1') (residue 10 and name H8) 4.0 0.5 1.0
assign (residue 9 and name H2') (residue 10 and name H1') 4.0 0.5 0.5
assign (residue 9 and name H2') (residue 10 and name H8) 3.0 0.5 0.5
assign (residue 9 and name H6) (residue 10 and name H8) 4.0 0.5 0.5

! Residue 10 ADE
! Intra Residue
assign (residue 10 and name H8) (residue 10 and name H1') 3.0 0.5 0.8
assign (residue 10 and name H8) (residue 10 and name H2') 3.5 0.5 1.0
assign (residue 10 and name H8) (residue 10 and name H3') 3.0 0.5 1.0
! Inter Residue
assign (residue 10 and name H1') (residue 11 and name H8) 4.0 0.5 1.0
```

```
assign (residue 10 and name H2') (residue 11 and name H8) 3.0 0.5 0.5
assign (residue 10 and name H3') (residue 11 and name H8) 2.5 0.5 0.5
assign (residue 10 and name H2) (residue 11 and name H1') 3.5 0.5 0.5
assign (residue 10 and name H2) (residue 11 and name H2) 4.5 0.5 1.0
assign (residue 10 and name H8) (residue 11 and name H8) 4.0 0.5 0.5

! Residue 11 ADE
! Intra Residue
assign (residue 11 and name H2) (residue 11 and name H1') 4.5 0.5 0.5
assign (residue 11 and name H8) (residue 11 and name H1') 3.5 0.5 0.5
assign (residue 11 and name H8) (residue 11 and name H2') 3.5 0.5 0.5
assign (residue 11 and name H8) (residue 11 and name H3') 3.5 0.5 0.5
! Inter Residue
assign (residue 11 and name H1') (residue 12 and name H1') 4.5 0.5 0.5
assign (residue 11 and name H1') (residue 12 and name H8) 4.0 0.5 0.8
assign (residue 11 and name H2') (residue 12 and name H1') 4.0 0.5 0.5
assign (residue 11 and name H2') (residue 12 and name H8) 3.0 0.5 0.8
assign (residue 11 and name H2) (residue 12 and name H1') 3.0 0.5 0.5
assign (residue 11 and name H2) (residue 12 and name H8) 4.5 0.5 1.0
assign (residue 11 and name H8) (residue 12 and name H8) 4.5 0.5 0.5

! Hairpin Loop Ends
! Residue 12 GUA
! Intra Residue
assign (residue 12 and name H8) (residue 12 and name H1') 3.5 0.5 0.5
assign (residue 12 and name H8) (residue 12 and name H2') 4.0 0.5 0.5
assign (residue 12 and name H8) (residue 12 and name H3') 3.5 0.5 0.5
! Inter Residue
assign (residue 12 and name H1') (residue 13 and name H6) 4.0 0.5 0.5
```

```
assign (residue 13 and name H1') (residue 12 and name H2') 3.5 0.5 0.5
assign (residue 12 and name H2') (residue 13 and name H5) 3.5 0.5 0.5
assign (residue 12 and name H2') (residue 13 and name H6) 2.5 0.5 0.5
assign (residue 12 and name H8) (residue 13 and name H5) 4.5 0.5 0.5
assign (residue 12 and name H8) (residue 13 and name H6) 4.5 0.5 0.5

! Residue 13 CYT
! Intra Residue
assign (residue 13 and name H6) (residue 13 and name H1') 3.5 0.5 0.5
assign (residue 13 and name H6) (residue 13 and name H2') 3.0 0.5 1.0
assign (residue 13 and name H6) (residue 13 and name H3') 2.5 0.5 0.5
! Inter Residue
assign (residue 13 and name H6) (residue 14 and name H5) 4.0 0.5 0.5 !
400mix

! Residue 14 CYT
! Intra Residue
assign (residue 14 and name H6) (residue 14 and name H1') 3.5 0.5 0.5
assign (residue 14 and name H6) (residue 14 and name H2') 3.0 0.5 0.5
assign (residue 14 and name H6) (residue 14 and name H3') 2.5 0.5 0.5
! Inter Residue
assign (residue 14 and name H1') (residue 13 and name H2') 3.5 0.5 1.0
assign (residue 14 and name H5) (residue 13 and name H2') 3.5 0.5 0.5
assign (residue 14 and name H6) (residue 13 and name H1') 3.5 0.5 1.0
assign (residue 14 and name H6) (residue 13 and name H2') 2.5 0.5 0.5
assign (residue 14 and name H6) (residue 13 and name H6) 4.5 0.5 0.5

! Unobserved-NOEs:
assign  (residue 12 and name H1) (residue 11 and name H2) 4.0 0.0 50.0
```

```
assign (resid 5 and name H1) (resid 10 and name H2) 4.5 0.0 50.0
```

Dihedral angle restraints for the 690QM mutant (modified based on Ravi's dihedral angel restraints for the 690QM mutant)

```
!sugar pucker delta
assign (resid 1 and name c5') (resid 1 and name c4')
       (resid 1 and name c3') (resid 1 and name o3') 1 80 5 2
assign (resid 2 and name c5') (resid 2 and name c4')
       (resid 2 and name c3') (resid 2 and name o3') 1 80 5 2
assign (resid 3 and name c5') (resid 3 and name c4')
       (resid 3 and name c3') (resid 3 and name o3') 1 80 5 2
assign (resid 4 and name c5') (resid 4 and name c4')
       (resid 4 and name c3') (resid 4 and name o3') 1 80 20 2
assign (resid 5 and name c5') (resid 5 and name c4')
       (resid 5 and name c3') (resid 5 and name o3') 1 80 20 2
assign (resid 6 and name c5') (resid 6 and name c4')
       (resid 6 and name c3') (resid 6 and name o3') 1 80 20 2
assign (resid 7 and name c5') (resid 7 and name c4')
       (resid 7 and name c3') (resid 7 and name o3') 1 80 20 2
assign (resid 8 and name c5') (resid 8 and name c4')
       (resid 8 and name c3') (resid 8 and name o3') 1 80 20 2
assign (resid 9 and name c5') (resid 9 and name c4')
       (resid 9 and name c3') (resid 9 and name o3') 1 80 20 2
assign (resid 10 and name c5') (resid 10 and name c4')
       (resid 10 and name c3') (resid 10 and name o3') 1 80 20 2
assign (resid 11 and name c5') (resid 11 and name c4')
       (resid 11 and name c3') (resid 11 and name o3') 1 80 20 2
assign (resid 12 and name c5') (resid 12 and name c4')
       (resid 12 and name c3') (resid 12 and name o3') 1 80 5 2
```

```
assign (resid 13 and name c5') (resid 13 and name c4')
       (resid 13 and name c3') (resid 13 and name o3') 1 80 5 2
assign (resid 14 and name c5') (resid 14 and name c4')
       (resid 14 and name c3') (resid 14 and name o3') 1 80 5 2

!zeta and alpha
assign (resid 1 and name c3') (resid 1 and name o3')
       (resid 2 and name p) (resid 2 and name o5') 1 -68 5 2
assign (resid 1 and name o3') (resid 2 and name p)
       (resid 2 and name o5') (resid 2 and name c5') 1 -71 5 2

assign (resid 2 and name c3') (resid 2 and name o3')
       (resid 3 and name p) (resid 3 and name o5') 1 -68 5 2
assign (resid 2 and name o3') (resid 3 and name p)
       (resid 3 and name o5') (resid 3 and name c5') 1 -71 5 2

assign (resid 3 and name c3') (resid 3 and name o3')
       (resid 4 and name p) (resid 4 and name o5') 1 0 120 2
assign (resid 3 and name o3') (resid 4 and name p)
       (resid 4 and name o5') (resid 4 and name c5') 1 0 120 2

assign (resid 4 and name c3') (resid 4 and name o3')
       (resid 5 and name p) (resid 5 and name o5') 1 0 120 2
assign (resid 4 and name o3') (resid 5 and name p)
       (resid 5 and name o5') (resid 5 and name c5') 1 0 120 2

assign (resid 5 and name c3') (resid 5 and name o3')
       (resid 6 and name p) (resid 6 and name o5') 1 0 120 2
assign (resid 5 and name o3') (resid 6 and name p)
```

```
       (resid 6 and name o5') (resid 6 and name c5') 1 0 120 2

assign (resid 6 and name c3') (resid 6 and name o3')
       (resid 7 and name p) (resid 7 and name o5') 1 0 120 2

assign (resid 7 and name c3') (resid 7 and name o3')
       (resid 8 and name p) (resid 8 and name o5') 1 0 120 2
assign (resid 7 and name o3') (resid 8 and name p)
       (resid 8 and name o5') (resid 8 and name c5') 1 0 120 2

assign (resid 8 and name c3') (resid 8 and name o3')
       (resid 9 and name p) (resid 9 and name o5') 1 0 120 2

assign (resid 9 and name c3') (resid 9 and name o3')
       (resid 10 and name p) (resid 10 and name o5') 1 0 120 2
assign (resid 9 and name o3') (resid 10 and name p)
       (resid 10 and name o5') (resid 10 and name c5') 1 0 120 2

assign (resid 10 and name c3') (resid 10 and name o3')
       (resid 11 and name p) (resid 11 and name o5') 1 0 120 2
assign (resid 10 and name o3') (resid 11 and name p)
       (resid 11 and name o5') (resid 11 and name c5') 1 0 120 2

assign (resid 11 and name c3') (resid 11 and name o3')
       (resid 12 and name p) (resid 12 and name o5') 1 0 120 2
assign (resid 11 and name o3') (resid 12 and name p)
       (resid 12 and name o5') (resid 12 and name c5') 1 0 120 2

assign (resid 12 and name c3') (resid 12 and name o3')
```

```
       (resid 13 and name p) (resid 13 and name o5') 1 -68 5 2
assign (resid 12 and name o3') (resid 13 and name p)
       (resid 13 and name o5') (resid 13 and name c5') 1 -71 5 2

assign (resid 13 and name c3') (resid 13 and name o3')
       (resid 14 and name p) (resid 14 and name o5') 1 -68 5 2
assign (resid 13 and name o3') (resid 14 and name p)
       (resid 14 and name o5') (resid 14 and name c5') 1 -71 5 2

!beta
assign (resid 1 and name p) (resid 1 and name o5')
       (resid 1 and name c5') (resid 1 and name c4') 1 180 5 2
assign (resid 2 and name p) (resid 2 and name o5')
       (resid 2 and name c5') (resid 2 and name c4') 1 180 5 2
assign (resid 3 and name p) (resid 3 and name o5')
       (resid 3 and name c5') (resid 3 and name c4') 1 180 20 2
assign (resid 5 and name p) (resid 5 and name o5')
       (resid 5 and name c5') (resid 5 and name c4') 1 180 60 2  !JSL
assign (resid 6 and name p) (resid 6 and name o5')
       (resid 6 and name c5') (resid 6 and name c4') 1 180 60 2  !JSL
assign (resid 7 and name p) (resid 7 and name o5')
       (resid 7 and name c5') (resid 7 and name c4') 1 180 60 2  !JSL
assign (resid 8 and name p) (resid 8 and name o5')
       (resid 8 and name c5') (resid 8 and name c4') 1 180 60 2  !JSL
assign (resid 10 and name p) (resid 10 and name o5')
       (resid 10 and name c5') (resid 10 and name c4') 1 180 60 2  !JSL
assign (resid 11 and name p) (resid 11 and name o5')
       (resid 11 and name c5') (resid 11 and name c4') 1 180 30 2
assign (resid 12 and name p) (resid 12 and name o5')
```

```
       (resid 12 and name c5') (resid 12 and name c4') 1 180 20 2
assign (resid 13 and name p) (resid 13 and name o5')
       (resid 13 and name c5') (resid 13 and name c4') 1 180 5 2
assign (resid 14 and name p) (resid 14 and name o5')
       (resid 14 and name c5') (resid 14 and name c4') 1 180 5 2

!gamma
assign (resid 1 and name o5') (resid 1 and name c5')
       (resid 1 and name c4') (resid 1 and name c3') 1 54 10 2
assign (resid 2 and name o5') (resid 2 and name c5')
       (resid 2 and name c4') (resid 2 and name c3') 1 54 10 2
assign (resid 3 and name o5') (resid 3 and name c5')
       (resid 3 and name c4') (resid 3 and name c3') 1 54 10 2
assign (resid 4 and name o5') (resid 4 and name c5')
       (resid 4 and name c4') (resid 4 and name c3') 1 60 60 2
assign (resid 5 and name o5') (resid 5 and name c5')
       (resid 5 and name c4') (resid 5 and name c3') 1 60 60 2
assign (resid 6 and name o5') (resid 6 and name c5')
       (resid 6 and name c4') (resid 6 and name c3') 1 60 60 2
assign (resid 7 and name o5') (resid 7 and name c5')
       (resid 7 and name c4') (resid 7 and name c3') 1 60 60 2
assign (resid 8 and name o5') (resid 8 and name c5')
       (resid 8 and name c4') (resid 8 and name c3') 1 60 60 2
assign (resid 10 and name o5') (resid 10 and name c5')
       (resid 10 and name c4') (resid 10 and name c3') 1 60 60 2
assign (resid 11 and name o5') (resid 11 and name c5')
       (resid 11 and name c4') (resid 11 and name c3') 1 60 60 2
assign (resid 12 and name o5') (resid 12 and name c5')
       (resid 12 and name c4') (resid 12 and name c3') 1 60 60 2
```

```
assign (resid 13 and name o5') (resid 13 and name c5')
       (resid 13 and name c4') (resid 13 and name c3') 1 54 10 2
assign (resid 14 and name o5') (resid 14 and name c5')
       (resid 14 and name c4') (resid 14 and name c3') 1 54 10 2

!epsilon
assign (resid 1 and name c4') (resid 1 and name c3')
       (resid 1 and name o3') (resid 2 and name p) 1 -155 5 2
assign (resid 2 and name c4') (resid 2 and name c3')
       (resid 2 and name o3') (resid 3 and name p) 1 -155 5 2
assign (resid 3 and name c4') (resid 3 and name c3')
       (resid 3 and name o3') (resid 4 and name p) 1 -155 45 2
assign (resid 4 and name c4') (resid 4 and name c3')
       (resid 4 and name o3') (resid 5 and name p) 1 -155 45 2
assign (resid 5 and name c4') (resid 5 and name c3')
       (resid 5 and name o3') (resid 6 and name p) 1 -155 45 2
assign (resid 6 and name c4') (resid 6 and name c3')
       (resid 6 and name o3') (resid 7 and name p) 1 -155 45 2
assign (resid 7 and name c4') (resid 7 and name c3')
       (resid 7 and name o3') (resid 8 and name p) 1 -155 45 2
assign (resid 8 and name c4') (resid 8 and name c3')
       (resid 8 and name o3') (resid 9 and name p) 1 -155 45 2
assign (resid 9 and name c4') (resid 9 and name c3')
       (resid 9 and name o3') (resid 10 and name p) 1 -155 45 2
assign (resid 10 and name c4') (resid 10 and name c3')
       (resid 10 and name o3') (resid 11 and name p) 1 -155 45 2
assign (resid 11 and name c4') (resid 11 and name c3')
       (resid 11 and name o3') (resid 12 and name p) 1 -155 5 2
assign (resid 12 and name c4') (resid 12 and name c3')
```

```
       (resid 12 and name o3') (resid 13 and name p) 1 -155 5 2
assign (resid 13 and name c4') (resid 13 and name c3')
       (resid 13 and name o3') (resid 14 and name p) 1 -155 5 2

! chi
assign (resid 1 and name O4') (resid 1 and name C1')
       (resid 1 and name N9 ) (resid 1 and name C4 ) 1 -155 5 2

assign (resid 2 and name O4') (resid 2 and name C1')
       (resid 2 and name N9 ) (resid 2 and name C4 ) 1 -155 5 2

assign (resid 3 and name O4') (resid 3 and name C1')
       (resid 3 and name N1 ) (resid 3 and name C2 ) 1 -155 30 2

assign (resid 4 and name O4') (resid 4 and name C1')
       (resid 4 and name N9 ) (resid 4 and name C4 ) 1 -155 45 2

assign (resid 5 and name O4') (resid 5 and name C1')
       (resid 5 and name N9 ) (resid 5 and name C4 ) 1 -155 45 2

assign (resid 6 and name O4') (resid 6 and name C1')
       (resid 6 and name N1 ) (resid 6 and name C2 ) 1 -155 45 2

assign (resid 7 and name O4') (resid 7 and name C1')
       (resid 7 and name N1 ) (resid 7 and name C2 ) 1 -155 45 2

assign (resid 8 and name O4') (resid 8 and name C1')
       (resid 8 and name N9 ) (resid 8 and name C4 ) 1 -155 45 2
```

```
assign (resid 9 and name O4') (resid 9 and name C1')
       (resid 9 and name N1 ) (resid 9 and name C2 ) 1 -155 45 2

assign (resid 10 and name O4') (resid 10 and name C1')
       (resid 10 and name N9 ) (resid 10 and name C4 ) 1 -155 45 2

assign (resid 11 and name O4') (resid 11 and name C1')
       (resid 11 and name N9 ) (resid 11 and name C4 ) 1 -155 45 2

assign (resid 12 and name O4') (resid 12 and name C1')
       (resid 12 and name N9 ) (resid 12 and name C4 ) 1 -155 30 2

assign (resid 13 and name O4') (resid 13 and name C1')
       (resid 13 and name N1 ) (resid 13 and name C2 ) 1 -155 5 2

assign (resid 14 and name O4') (resid 14 and name C1')
       (resid 14 and name N1 ) (resid 14 and name C2 ) 1 -155 5 2
```

APPENDIX 3

Structural Restraints of the Unmodified-970 RNA

Distance restraints for the modified-970

```
!Base-pair restraints
assign (residue 2 and name O6) (residue 1 and name O6) 4.0 1.0 0.5
assign (residue 2 and name N2) (residue 1 and name N2) 4.0 1.0 0.5
assign (residue 2 and name O6) (residue 3 and name O6) 4.0 1.0 0.5
assign (residue 2 and name N2) (residue 3 and name N2) 4.0 1.0 0.5
assign (residue 3 and name O6) (residue 4 and name O4) 4.0 1.0 0.5
assign (residue 3 and name N2) (residue 4 and name O2) 4.0 1.0 0.5
assign (residue 4 and name O4) (residue 5 and name O4) 4.0 1.0 0.5
assign (residue 4 and name O2) (residue 5 and name O2) 4.0 1.0 0.5
assign (residue 5 and name O4) (residue 6 and name N4) 4.0 1.0 0.5
assign (residue 5 and name O2) (residue 6 and name O2) 4.0 1.0 0.5
assign (residue 6 and name N4) (residue 7 and name O6) 4.0 1.0 0.5
assign (residue 6 and name O2) (residue 7 and name N2) 4.0 1.0 0.5
assign (residue 16 and name N4) (residue 17 and name O6) 4.0 1.0 0.5
assign (residue 16 and name O2) (residue 17 and name N2) 4.0 1.0 0.5
assign (residue 17 and name O6) (residue 18 and name N6) 4.0 1.0 0.5
assign (residue 17 and name N2) (residue 18 and name H2) 4.0 1.0 0.5
assign (residue 18 and name N6) (residue 19 and name N6) 4.0 1.0 0.5
assign (residue 18 and name H2) (residue 19 and name H2) 4.0 1.0 0.5
assign (residue 19 and name N6) (residue 20 and name N4) 4.0 1.0 0.5
assign (residue 19 and name H2) (residue 20 and name O2) 4.0 1.0 0.5
assign (residue 20 and name N4) (residue 21 and name N4) 3.5 1.0 0.0
```

```
assign (residue 20 and name O2) (residue 21 and name O2) 4.0 1.0 0.5
assign (residue 21 and name N4) (residue 22 and name N4) 3.5 1.0 0.0
assign (residue 21 and name O2) (residue 22 and name O2) 4.0 1.0 0.5

!stem stacking distances
assign (residue 1 and name H1) (residue 2 and name H1) 3.5 1.5 0.5
assign (residue 2 and name H1) (residue 3 and name H1) 3.5 1.5 0.5
assign (residue 3 and name H1) (residue 4 and name H3) 3.5 1.5 0.5
assign (residue 4 and name H3) (residue 5 and name H3) 3.5 1.5 0.5
assign (residue 5 and name H3) (residue 17 and name H1) 3.5 1.5 0.5
assign (residue 7 and name H1) (residue 17 and name H1) 3.5 1.5 0.5

! Base pair distances
assign (residue 1 and name C1') (residue 22 and name C1') 10.8 0.2 0.2
assign (residue 1 and name C8) (residue 22 and name C6) 9.9 0.2 0.2
assign (residue 2 and name C1') (residue 21 and name C1') 10.8 0.2 0.2
assign (residue 2 and name C8) (residue 21 and name C6) 9.9 0.2 0.2
assign (residue 3 and name C1') (residue 20 and name C1') 10.8 0.2 0.2
assign (residue 3 and name C8) (residue 20 and name C6) 9.9 0.2 0.2
assign (residue 4 and name C1') (residue 19 and name C1') 10.5 0.2 0.2
assign (residue 4 and name C6) (resid 19 and name C8) 9.7 0.2 0.2
assign (residue 5 and name C1') (residue 18 and name C1') 10.5 0.2 0.2
assign (residue 5 and name C6) (resid 18 and name C8) 9.7 0.2 0.2
assign (residue 6 and name C1') (residue 17 and name C1') 10.8 0.2 0.2
assign (residue 6 and name C6) (residue 17 and name C8) 9.9 0.2 0.2
assign (residue 7 and name C1') (residue 16 and name C1') 10.8 0.2 0.2
assign (residue 7 and name C8) (residue 16 and name C6) 9.9 0.2 0.2

! Base Pair restraints
```

```
! for G1/ C22  base pair
assign (resid 1 and name N1) (resid 22 and name N3) 2.91 0.2 0.2
assign (resid 1 and name O6) (resid 22 and name N4) 2.71 0.2 0.2
assign (resid 1 and name N2) (resid 22 and name O2) 3.08 0.2 0.2
assign (resid 1 and name H1) (resid 22 and name N3) 1.89 0.2 0.2
assign (resid 1 and name O6) (resid 22 and name H42) 1.71 0.2 0.2
assign (resid 1 and name H22) (resid 22 and name O2) 2.08 0.2 0.2

! for G2/ C21  base pair
assign (resid 2 and name N1) (resid 21 and name N3) 2.91 0.2 0.2
assign (resid 2 and name O6) (resid 21 and name N4) 2.71 0.2 0.2
assign (resid 2 and name N2) (resid 21 and name O2) 3.08 0.2 0.2
assign (resid 2 and name H1) (resid 21 and name N3) 1.89 0.2 0.2
assign (resid 2 and name O6) (resid 21 and name H42) 1.71 0.2 0.2
assign (resid 2 and name H22) (resid 21 and name O2) 2.08 0.2 0.2

! for G3/ C20  base pair
assign (resid 3 and name N1) (resid 20 and name N3) 2.91 0.2 0.2
assign (resid 3 and name O6) (resid 20 and name N4) 2.71 0.2 0.2
assign (resid 3 and name N2) (resid 20 and name O2) 3.08 0.2 0.2
assign (resid 3 and name H1) (resid 20 and name N3) 1.89 0.2 0.2
assign (resid 3 and name O6) (resid 20 and name H42) 1.71 0.2 0.2
assign (resid 3 and name H22) (resid 20 and name O2) 2.08 0.2 0.2

! for U4/A19 base pair
assign (resid 19 and name N1) (resid 4 and name N3) 2.95 0.2 0.2
assign (resid 19 and name N6) (resid 4 and name O4) 2.83 0.2 0.2
assign (resid 19 and name N1) (resid 4 and name H3) 1.93 0.2 0.2
assign (resid 19 and name H62) (resid 4 and name O4) 1.82 0.2 0.2
```

```
! for U5/A18 base pair
assign (resid 18 and name N1) (resid 5 and name N3) 2.95 0.2 0.2
assign (resid 18 and name N6) (resid 5 and name O4) 2.83 0.2 0.2
assign (resid 18 and name N1) (resid 5 and name H3) 1.93 0.2 0.2
assign (resid 18 and name H62) (resid 5 and name O4) 1.82 0.2 0.2

! for C6/G17 base pair
assign (resid 17 and name N1) (resid 6 and name N3) 2.91 0.2 0.2
assign (resid 17 and name O6) (resid 6 and name N4) 2.71 0.2 0.2
assign (resid 17 and name N2) (resid 6 and name O2) 3.08 0.2 0.2
assign (resid 17 and name H1) (resid 6 and name N3) 1.89 0.2 0.2
assign (resid 17 and name O6) (resid 6 and name H42) 1.71 0.2 0.2
assign (resid 17 and name H22) (resid 6 and name O2) 2.08 0.2 0.2

! for G7/ C16 base pair  ! change on 10/10/10
assign (resid 7 and name N1) (resid 16 and name N3) 2.91 0.2 0.5
assign (resid 7 and name O6) (resid 16 and name N4) 2.71 0.2 0.5
assign (resid 7 and name N2) (resid 16 and name O2) 3.08 0.2 0.5
assign (resid 7 and name H1) (resid 16 and name N3) 1.89 0.2 0.5
assign (resid 7 and name O6) (resid 16 and name H42) 1.71 0.2 0.5
assign (resid 7 and name H22) (resid 16 and name O2) 2.08 0.2 0.5

! Add from NOE-H2O
assign (residue 3 and name H1) (residue 4 and name H1') 4.5 0.5 0.5
assign (residue 3 and name H1) (residue 21 and name H1') 4.5 0.5 0.5
assign (residue 3 and name H1) (residue 21 and name H41) 4.0 0.2 0.5
assign (residue 3 and name H1) (residue 21 and name H42) 4.0 0.2 0.5
assign (residue 4 and name H3) (residue 19 and name H2) 3.0 0.5 0.5
```

```
assign (residue 5 and name H3) (residue 18 and name H2) 3.0 0.5 0.5
assign (residue 5 and name H3) (residue 19 and name H2) 4.0 0.5 0.5
assign (residue 7 and name H1') (residue 17 and name H1) 4.5 0.5 0.5
assign (residue 7 and name H1) (residue 16 and name H41) 4.0 0.5 0.5
assign (residue 7 and name H1) (residue 16 and name H42) 4.0 0.5 0.5
assign (residue 16 and name H41) (residue 17 and name H1) 3.5 0.5 0.5
assign (residue 16 and name H42) (residue 17 and name H1) 4.5 0.5 0.5

! Residue 1GUA
assign (residue 1 and name H1') (residue 1 and name H8) 4.0 0.5 0.5
assign (residue 1 and name H2') (residue 1 and name H8) 3.0 0.5 0.8
assign (residue 1 and name H3') (residue 1 and name H8) 3.0 0.5 0.5
assign (residue 1 and name H5') (residue 1 and name H8) 3.5 0.5 0.5
assign (residue 1 and name H5'') (residue 1 and name H8) 3.5 0.5 1.0
! Inter residue
assign (residue 1 and name H1') (residue 2 and name H8) 4.0 0.5 0.8
assign (residue 1 and name H2') (residue 2 and name H1') 4.5 0.5 0.5
assign (residue 1 and name H2') (residue 2 and name H8) 2.0 0.2 0.9
assign (residue 1 and name H3') (residue 2 and name H8) 4.0 0.5 1.0
assign (residue 1 and name H8) (residue 2 and name H8) 4.0 0.5 0.8

! Residue 2GUA
! Intra residues
assign (residue 2 and name H1') (residue 2 and name H8) 4.0 0.5 0.5
assign (residue 2 and name H2') (residue 2 and name H8) 4.0 0.5 0.5
assign (residue 2 and name H3') (residue 2 and name H8) 2.0 0.2 0.9
assign (residue 2 and name H4') (residue 2 and name H8) 3.5 0.5 1.0
assign (residue 2 and name H5') (residue 2 and name H8) 3.0 0.5 0.8
! Inter residues
```

```
assign (residue 2 and name H1') (residue 3 and name H8) 4.0 0.5 0.8
assign (residue 2 and name H2') (residue 3 and name H8) 2.0 0.2 0.9
assign (residue 2 and name H3') (residue 3 and name H8) 3.5 0.5 1.0
assign (residue 2 and name H8) (residue 3 and name H8) 4.0 0.5 0.8

! Residue 3GUA
! Intra residues
assign (residue 3 and name H1') (residue 3 and name H8) 3.5 0.5 0.5
assign (residue 3 and name H2') (residue 3 and name H8) 3.5 0.5 1.0
assign (residue 3 and name H3') (residue 3 and name H8) 2.0 0.2 1.0
assign (residue 3 and name H4') (residue 3 and name H8) 3.5 0.5 1.0
assign (residue 3 and name H5') (residue 3 and name H8) 3.0 0.5 0.8
! Inter residues
assign (residue 3 and name H1') (residue 4 and name H6) 4.0 0.5 1.5
assign (residue 3 and name H2') (residue 4 and name H5) 3.5 0.5 1.0
assign (residue 3 and name H2') (residue 4 and name H6) 3.0 1.0 1.0
assign (residue 3 and name H3') (residue 4 and name H5) 3.5 0.5 1.5
assign (residue 3 and name H3') (residue 4 and name H6) 3.5 0.5 1.5
assign (residue 3 and name H8) (residue 4 and name H6) 4.5 0.5 0.5

! Residue 4URI
! Intra residues
assign (residue 4 and name H2') (residue 4 and name H5) 3.5 0.5 1.5
assign (residue 4 and name H3') (residue 4 and name H5) 4.0 0.5 1.5
assign (residue 4 and name H1') (residue 4 and name H6) 2.5 0.3 1.0
assign (residue 4 and name H2') (residue 4 and name H6) 2.5 0.3 1.5
assign (residue 4 and name H3') (residue 4 and name H6) 2.0 0.2 1.5
assign (residue 4 and name H5') (residue 4 and name H6) 3.5 0.5 1.5
! Inter Residues
```

```
assign (residue 4 and name H2') (residue 5 and name H6) 2.0 0.2 0.9
assign (residue 4 and name H5) (residue 5 and name H5) 3.5 0.5 0.5
assign (residue 4 and name H6) (residue 5 and name H6) 4.5 0.5 0.5

! Residue 5URI
! Intra residues
assign (residue 5 and name H1') (residue 5 and name H6) 3.5 0.5 1.5
assign (residue 5 and name H2') (residue 5 and name H6) 3.5 0.5 1.5
assign (residue 5 and name H3') (residue 5 and name H6) 3.0 0.5 0.8
assign (residue 5 and name H5') (residue 5 and name H6) 3.0 0.5 1.5
! Inter residues
assign (residue 5 and name H2') (residue 6 and name H1') 3.5 0.5 1.5
assign (residue 5 and name H2') (residue 6 and name H5) 4.0 0.5 0.5
assign (residue 5 and name H2') (residue 6 and name H6) 2.5 0.5 0.8
assign (residue 5 and name H3') (residue 6 and name H6) 4.0 0.5 0.5
assign (residue 5 and name H1') (residue 18 and name H2) 3.5 0.5 1.5
assign (residue 5 and name H3) (residue 17 and name H1) 4.0 0.5 0.5
assign (residue 5 and name H6) (residue 6 and name H6) 4.5 0.5 0.5

! Residue 6CYT
! Intra residues
assign (residue 6 and name H1') (residue 6 and name H5') 4.0 0.5 1.5
assign (residue 6 and name H1') (residue 6 and name H5) 4.0 0.5 1.5
assign (residue 6 and name H2') (residue 6 and name H5) 4.0 0.5 1.5
assign (residue 6 and name H1') (residue 6 and name H6) 3.5 0.5 1.5
assign (residue 6 and name H2') (residue 6 and name H6) 3.5 0.5 1.5
assign (residue 6 and name H3') (residue 6 and name H6) 3.0 0.5 1.5
assign (residue 6 and name H5') (residue 6 and name H6) 3.0 0.5 0.8
! Inter residues
```

```
assign (residue 6 and name H1') (residue 7 and name H8) 4.0 0.5 1.0
assign (residue 6 and name H1') (residue 18 and name H2) 3.5 0.5 0.8
assign (residue 6 and name H2') (residue 7 and name H8) 2.5 0.5 0.8
assign (residue 6 and name H3') (residue 7 and name H8) 3.5 0.5 1.0
assign (residue 6 and name H6) (residue 7 and name H8) 4.5 0.5 0.5

! Residue 7GUA
! Intra residues
assign (residue 7 and name H1') (residue 7 and name H2') 2.5 0.5 0.5
assign (residue 7 and name H1') (residue 7 and name H5') 3.5 0.5 1.0
assign (residue 7 and name H1') (residue 7 and name H8) 4.0 0.5 0.5
assign (residue 7 and name H2') (residue 7 and name H8) 4.0 0.5 0.8
assign (residue 7 and name H3') (residue 7 and name H8) 3.5 0.5 1.5
assign (residue 7 and name H5') (residue 7 and name H8) 4.5 0.5 0.5
! Inter residues
assign (residue 7 and name H1') (residue 8 and name H8) 4.0 0.5 1.0
assign (residue 7 and name H1') (residue 8 and name H2) 3.5 0.5 0.8
assign (residue 7 and name H1') (residue 8 and name H1') 4.5 0.5 0.5
assign (residue 7 and name H2') (residue 8 and name H1') 4.0 0.5 0.8
assign (residue 7 and name H2') (residue 8 and name H8) 2.5 0.5 0.5
assign (residue 7 and name H8) (residue 8 and name H8) 4.0 0.5 0.8

! Residue 8ADE
! Intra residues
assign (residue 8 and name H1') (residue 8 and name H2') 2.5 0.5 1.5
assign (residue 8 and name H1') (residue 8 and name H3') 4.0 0.5 0.5
assign (residue 8 and name H1') (residue 8 and name H8) 4.0 0.5 0.5
assign (residue 8 and name H2') (residue 8 and name H8) 3.5 0.5 0.8
assign (residue 8 and name H3') (residue 8 and name H8) 3.5 0.5 0.5
```

```
! Inter residues
assign (residue 8 and name H1') (residue 9 and name H1') 4.5 0.5 0.5
assign (residue 8 and name H1') (residue 9 and name H5) 4.5 0.5 0.5
assign (residue 8 and name H1') (residue 9 and name H6) 4.5 0.2 0.5
assign (residue 8 and name H2') (residue 9 and name H5) 4.5 0.5 0.5
assign (residue 8 and name H2') (residue 9 and name H6) 2.5 0.5 0.8
assign (residue 8 and name H3') (residue 9 and name H6) 4.0 0.5 0.5
assign (residue 8 and name H3') (residue 9 and name H5) 4.0 0.5 0.8
assign (residue 8 and name H2) (residue 9 and name H1') 3.5 0.5 0.5
assign (residue 8 and name H2) (residue 16 and name H1') 4.5 0.5 1.0
assign (residue 8 and name H8) (residue 9 and name H5) 4.5 0.5 0.8
assign (residue 8 and name H8) (residue 9 and name H6) 4.5 0.5 0.8

! Residue 9URI
! Intra residues
assign (residue 9 and name H1') (residue 9 and name H2') 2.5 0.5 0.8
assign (residue 9 and name H1') (residue 9 and name H3') 4.0 0.5 1.0
assign (residue 9 and name H1') (residue 9 and name H5) 4.0 0.5 0.8
assign (residue 9 and name H2') (residue 9 and name H5) 4.0 0.5 1.0
assign (residue 9 and name H3') (residue 9 and name H5) 4.0 0.5 1.0
assign (residue 9 and name H1') (residue 9 and name H6) 3.5 0.5 1.5
assign (residue 9 and name H2') (residue 9 and name H6) 3.0 0.5 0.8
assign (residue 9 and name H3') (residue 9 and name H6) 3.0 0.5 0.5
assign (residue 9 and name H4') (residue 9 and name H6) 4.0 0.5 1.0
assign (residue 9 and name H5') (residue 9 and name H6) 3.5 0.5 0.5
! Inter residues
assign (residue 9 and name H4') (residue 13 and name H1') 3.5 0.5 0.8
assign (residue 9 and name H4') (residue 13 and name H8) 4.5 0.5 0.8
```

```
! Residue 10GUA
! Intra residues
assign (residue 10 and name H1') (residue 10 and name H2') 3.0 0.5 0.8
assign (residue 10 and name H1') (residue 10 and name H8) 4.0 0.5 0.5
assign (residue 10 and name H2') (residue 10 and name H8) 3.0 0.5 0.5
assign (residue 10 and name H3') (residue 10 and name H8) 3.0 0.3 0.5
assign (residue 10 and name H4') (residue 10 and name H8) 4.0 0.5 0.8
! Inter residues
assign (residue 10 and name H1') (residue 11 and name H5) 3.5 0.5 0.5
assign (residue 10 and name H1') (residue 11 and name H6) 4.0 0.5 0.5
assign (residue 10 and name H2') (residue 11 and name H6) 3.0 0.5 0.5
assign (residue 10 and name H3') (residue 11 and name H6) 4.0 0.5 1.0
assign (residue 10 and name H5') (residue 10 and name H8) 4.0 0.5 0.5
assign (residue 10 and name H8) (residue 11 and name H5) 4.0 0.5 0.5
assign (residue 10 and name H8) (residue 11 and name H6) 4.5 0.5 0.8

! Residue 11CYT
! Intra residues
assign (residue 11 and name H1') (residue 11 and name H3') 3.5 0.5 1.0
assign (residue 11 and name H1') (residue 11 and name H4') 3.0 0.5 0.8
assign (residue 11 and name H5'') (residue 11 and name H3') 3.0 0.5 0.5
assign (residue 11 and name H1') (residue 11 and name H6) 3.0 0.5 0.8
assign (residue 11 and name H2') (residue 11 and name H6) 3.0 0.2 0.8
assign (residue 11 and name H3') (residue 11 and name H6) 3.0 0.5 1.0
assign (residue 11 and name H4') (residue 11 and name H6) 4.5 0.5 0.5
assign (residue 11 and name H5') (residue 11 and name H6) 4.0 0.5 0.5
assign (residue 11 and name H5'') (residue 11 and name H6) 4.0 0.5 0.5
! Inter residues
assign (residue 11 and name H1') (residue 12 and name H2) 4.0 0.5 1.0
```

```
assign (residue 11 and name H6) (residue 12 and name H8) 4.5 0.5 1.0

! Residue 12ADE
! Intra residues
assign (residue 12 and name H1') (residue 12 and name H3') 4.0 0.5 1.0
assign (residue 12 and name H1') (residue 12 and name H4') 3.0 0.5 0.8
assign (residue 12 and name H1') (residue 12 and name H2) 4.0 0.5 1.0
assign (residue 12 and name H1') (residue 12 and name H8) 3.0 0.5 0.8
assign (residue 12 and name H2') (residue 12 and name H8) 3.0 0.5 0.8
assign (residue 12 and name H2') (residue 12 and name H3') 3.5 0.5 0.5
assign (residue 12 and name H3') (residue 12 and name H8) 4.5 0.5 0.8
assign (residue 12 and name H3') (residue 12 and name H5') 3.0 0.5 0.8
! Inter residues
assign (residue 12 and name H1') (residue 13 and name H8) 4.5 0.5 1.5

! Residue 13ADE
! Intra residues
assign (residue 13 and name H1') (residue 13 and name H3') 3.5 0.5 0.5
assign (residue 13 and name H1') (residue 13 and name H4') 3.5 0.5 0.5
assign (residue 13 and name H1') (residue 13 and name H8) 3.5 0.5 0.5
assign (residue 13 and name H2') (residue 13 and name H8) 3.5 0.5 0.5
assign (residue 13 and name H3') (residue 13 and name H8) 3.0 0.5 0.5
assign (residue 13 and name H1') (residue 13 and name H2) 4.0 0.5 1.0
assign (residue 13 and name H3') (residue 13 and name H5') 3.5 0.5 0.5
assign (residue 13 and name H8) (residue 13 and name H5') 4.0 0.5 0.5
assign (residue 13 and name H3') (residue 13 and name H5'') 3.5 0.5 0.5
! Inter residues
assign (residue 13 and name H1') (residue 14 and name H6) 4.5 0.5 0.5
assign (residue 13 and name H2') (residue 14 and name H6) 4.0 0.5 1.0
```

```
assign (residue 13 and name H2) (residue 14 and name H1') 4.0 0.5 1.0

! Residue 14CYT
! Intra residues
assign (residue 14 and name H1') (residue 14 and name H3') 4.0 0.5 0.5
assign (residue 14 and name H1') (residue 14 and name H4') 4.0 0.5 0.5
assign (residue 14 and name H1') (residue 14 and name H6) 3.0 0.5 0.8
assign (residue 14 and name H2') (residue 14 and name H6) 2.5 0.5 0.5
assign (residue 14 and name H3') (residue 14 and name H6) 2.5 0.5 1.0
! Inter residues
assign (residue 14 and name H1') (residue 15 and name H8) 4.0 0.5 0.5

! Residue 15GUA
! Intra residues
assign (residue 15 and name H1') (residue 15 and name H8) 3.5 0.5 0.5
assign (residue 15 and name H2') (residue 15 and name H8) 3.5 0.8 0.8
assign (residue 15 and name H3') (residue 15 and name H8) 3.5 0.5 0.5
assign (residue 15 and name H4') (residue 15 and name H8) 4.0 0.5 1.0
! Inter residues
assign (residue 15 and name H1') (residue 16 and name H6) 4.5 0.5 0.5
assign (residue 15 and name H2') (residue 16 and name H6) 2.5 0.5 0.5
assign (residue 15 and name H3') (residue 16 and name H6) 3.5 0.5 0.5
assign (residue 15 and name H3') (residue 16 and name H5) 4.0 0.5 0.5
assign (residue 15 and name H8) (residue 16 and name H5) 4.5 0.5 0.5
assign (residue 15 and name H8) (residue 16 and name H6) 4.5 0.5 0.5

! Residue 16CYT
! Intra residues
assign (residue 16 and name H2') (residue 16 and name H5) 4.0 0.5 1.0
```

```
assign (residue 16 and name H3') (residue 16 and name H5) 4.5 0.5 0.5
assign (residue 16 and name H1') (residue 16 and name H6) 3.5 0.5 0.5
assign (residue 16 and name H2') (residue 16 and name H6) 3.5 0.5 0.5
assign (residue 16 and name H3') (residue 16 and name H6) 3.0 0.5 1.0
assign (residue 16 and name H5') (residue 16 and name H6) 4.5 0.5 0.5
! Inter residues
assign (residue 16 and name H1') (residue 17 and name H8) 4.0 0.5 0.8
assign (residue 16 and name H2') (residue 17 and name H8) 2.0 0.2 0.9

! Residue 17GUA
! Intra residues
assign (residue 17 and name H1') (residue 17 and name H8) 3.5 0.5 1.5
assign (residue 17 and name H2') (residue 17 and name H8) 4.0 0.5 1.5
assign (residue 17 and name H3') (residue 17 and name H8) 2.5 0.2 1.0
assign (residue 17 and name H4') (residue 17 and name H8) 4.0 0.5 1.5
assign (residue 17 and name H5') (residue 17 and name H8) 3.0 0.5 0.8
! Inter residues
assign (residue 17 and name H1') (residue 18 and name H8) 4.0 0.5 1.0
assign (residue 17 and name H2') (residue 18 and name H8) 2.0 0.2 0.9
assign (residue 17 and name H8) (residue 18 and name H8) 4.5 0.5 0.5

! residue 18ADE
! Intra residues
assign (residue 18 and name H1') (residue 18 and name H2) 3.5 0.5 1.5
assign (residue 18 and name H2') (residue 18 and name H2) 3.5 0.5 1.5
assign (residue 18 and name H1') (residue 18 and name H8) 4.0 0.5 1.5
assign (residue 18 and name H3') (residue 18 and name H8) 3.0 0.5 0.8
assign (residue 18 and name H5') (residue 18 and name H8) 3.0 0.5 0.8
! inter residues
```

```
assign (residue 18 and name H1') (residue 19 and name H8) 3.5 0.5 1.5
assign (residue 18 and name H2') (residue 19 and name H8) 2.0 0.2 0.9
assign (residue 18 and name H2) (residue 19 and name H1') 3.5 0.5 1.5
assign (residue 18 and name H2') (residue 19 and name H1') 3.0 0.5 0.8
assign (residue 18 and name H2) (residue 19 and name H2) 4.0 0.5 0.5
assign (residue 19 and name H2) (residue 20 and name H5) 4.0 0.5 1.0

! residue 19ADE
! Intra residue
assign (residue 19 and name H1') (residue 19 and name H8) 4.0 0.5 1.5
assign (residue 19 and name H3') (residue 19 and name H8) 3.0 0.5 0.8
! Inter residue
assign (residue 19 and name H1') (residue 20 and name H5) 4.0 0.5 1.5
assign (residue 19 and name H2) (residue 20 and name H1') 3.5 0.5 1.5
assign (residue 19 and name H2') (residue 20 and name H5) 3.5 0.5 1.5
assign (residue 19 and name H3') (residue 20 and name H5) 3.5 0.5 1.5
assign (residue 19 and name H1') (residue 20 and name H6) 3.5 0.5 1.5
assign (residue 19 and name H2') (residue 20 and name H6) 3.0 0.5 0.8
assign (residue 19 and name H3') (residue 20 and name H6) 3.5 0.5 1.5
assign (residue 19 and name H2') (residue 20 and name H1') 3.5 0.5 1.5
assign (residue 19 and name H2) (residue 20 and name H6) 4.5 0.5 0.5

! residue 20CYT
! Intra residues
assign (residue 20 and name H1') (residue 20 and name H3') 3.5 0.5 1.5
assign (residue 20 and name H1') (residue 20 and name H4') 3.5 0.5 1.5
assign (residue 20 and name H1') (residue 20 and name H5) 3.5 0.5 1.5
assign (residue 20 and name H3') (residue 20 and name H5) 4.0 0.5 1.5
assign (residue 20 and name H1') (residue 20 and name H6) 3.5 0.5 1.5
```

```
assign (residue 20 and name H2') (residue 20 and name H6) 3.5 0.5 1.5
assign (residue 20 and name H3') (residue 20 and name H6) 3.0 0.5 0.8
assign (residue 20 and name H4') (residue 20 and name H6) 3.5 0.5 1.5
assign (residue 20 and name H5') (residue 20 and name H6) 3.5 0.5 1.5
! inter residues
assign (residue 20 and name H1') (residue 21 and name H6) 4.0 0.5 1.5
assign (residue 20 and name H2') (residue 21 and name H6) 2.0 0.2 0.9
assign (residue 20 and name H3') (residue 21 and name H6) 3.0 0.5 0.8
assign (residue 20 and name H5) (residue 21 and name H5) 3.5 0.5 1.0
assign (residue 20 and name H6) (residue 21 and name H5) 3.5 0.5 1.0
assign (residue 20 and name H6) (residue 21 and name H6) 4.5 0.5 0.5

! residue 21CYT
! intra residues
assign (residue 21 and name H1') (residue 21 and name H4') 3.0 0.5 0.8
assign (residue 21 and name H1') (residue 21 and name H5') 4.0 0.5 1.5
assign (residue 21 and name H2') (residue 21 and name H5) 4.0 0.5 1.5
assign (residue 21 and name H1') (residue 21 and name H6) 3.5 0.5 1.5
assign (residue 21 and name H2') (residue 21 and name H6) 4.0 0.5 1.5
assign (residue 21 and name H3') (residue 21 and name H6) 3.0 0.5 0.8
assign (residue 21 and name H4') (residue 21 and name H6) 3.5 0.5 1.5
assign (residue 21 and name H5') (residue 21 and name H6) 3.5 0.5 1.5
! inter residues
assign (residue 21 and name H2') (residue 22 and name H1') 4.0 0.5 1.5
assign (residue 21 and name H2') (residue 22 and name H5) 4.0 0.5 1.5
assign (residue 21 and name H1') (residue 22 and name H6) 3.5 0.5 1.5
assign (residue 21 and name H2') (residue 22 and name H6) 2.0 0.2 0.9
assign (residue 21 and name H3') (residue 22 and name H6) 2.5 0.5 1.5
assign (residue 21 and name H4') (residue 22 and name H6) 4.0 0.5 1.5
```

```
assign (residue 21 and name H5) (residue 22 and name H6) 4.0 0.5 1.0

! residue 22CYT
! intra residues
assign (residue 22 and name H1') (residue 22 and name H3') 3.0 0.5 0.8
assign (residue 22 and name H1') (residue 22 and name H4') 3.0 0.5 0.8
assign (residue 22 and name H1') (residue 22 and name H5) 4.0 0.5 1.5
assign (residue 22 and name H3') (residue 22 and name H5) 3.5 0.5 1.5
assign (residue 22 and name H1') (residue 22 and name H6) 3.5 0.5 1.5
assign (residue 22 and name H2') (residue 22 and name H6) 3.5 0.5 1.5
assign (residue 22 and name H3') (residue 22 and name H6) 2.0 0.2 0.9
assign (residue 22 and name H4') (residue 22 and name H6) 3.0 0.5 0.8
assign (residue 22 and name H5') (residue 22 and name H6) 3.5 0.5 1.5

! Un-NOE's
assign (residue 1 and name H1') (residue 6 and name H1') 4.5 0.0 50.0
assign (residue 1 and name H8) (residue 6 and name H1') 4.5 0.0 50.0
assign (residue 1 and name H2') (residue 18 and name H2) 4.5 0.0 50.0
assign (residue 1 and name H8) (residue 18 and name H4') 4.5 0.0 50.0
assign (residue 2 and name H1') (residue 4 and name H1') 4.5 0.0 50.0
assign (residue 2 and name H1') (residue 6 and name H1') 4.5 0.0 50.0
assign (residue 2 and name H2') (residue 5 and name H6) 4.5 0.0 50.0
assign (residue 2 and name H2') (residue 18 and name H2) 4.5 0.0 50.0
assign (residue 2 and name H3') (residue 4 and name H3') 4.5 0.0 50.0
assign (residue 2 and name H3') (residue 19 and name H1') 4.5 0.0 50.0
assign (residue 2 and name H3') (residue 19 and name H2) 4.5 0.0 50.0
assign (residue 2 and name H8) (residue 18 and name H2) 4.5 0.0 50.0
assign (residue 3 and name H1') (residue 5 and name H6) 4.5 0.0 50.0
assign (residue 3 and name H1') (residue 6 and name H5) 4.5 0.0 50.0
```

```
assign (residue 3 and name H1') (residue 6 and name H6) 4.5 0.0 50.0
assign (residue 3 and name H2') (residue 15 and name H1') 4.5 0.0 50.0
assign (residue 3 and name H3') (residue 5 and name H6) 4.5 0.0 50.0
assign (residue 3 and name H8) (residue 5 and name H1') 4.5 0.0 50.0
assign (residue 3 and name H8) (residue 6 and name H6) 4.5 0.0 50.0
assign (residue 4 and name H5) (residue 6 and name H1') 4.5 0.0 50.0
assign (residue 4 and name H5) (residue 6 and name H5) 4.5 0.0 50.0
assign (residue 4 and name H5) (residue 18 and name H2) 4.5 0.0 50.0
assign (residue 4 and name H5) (residue 21 and name H1') 4.5 0.0 50.0
assign (residue 4 and name H5) (residue 22 and name H2') 4.5 0.0 50.0
assign (residue 5 and name H5) (residue 21 and name H2') 4.5 0.0 50.0
assign (residue 5 and name H5) (residue 21 and name H4') 4.5 0.0 50.0
assign (residue 5 and name H5) (residue 22 and name H1') 4.5 0.0 50.0
assign (residue 5 and name H6) (residue 22 and name H1') 4.5 0.0 50.0
assign (residue 6 and name H5) (residue 22 and name H1') 4.5 0.0 50.0
assign (residue 6 and name H6) (residue 22 and name H1') 4.5 0.0 50.0
assign (residue 7 and name H8) (residue 8 and name H2) 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 13 and name H1') 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 12 and name H1') 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 12 and name H4') 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 13 and name H8) 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 16 and name H6) 4.5 0.0 50.0
assign (residue 8 and name H3') (residue 11 and name H5) 4.5 0.0 50.0
assign (residue 8 and name H5') (residue 11 and name H5) 4.5 0.0 50.0
assign (residue 8 and name H8) (residue 11 and name H5) 4.5 0.0 50.0
assign (residue 9 and name H1') (residue 12 and name H3') 4.5 0.0 50.0
assign (residue 9 and name H2') (residue 11 and name H5) 4.5 0.0 50.0
assign (residue 9 and name H2') (residue 14 and name H5) 4.5 0.0 50.0
assign (residue 9 and name H3') (residue 10 and name H2') 4.5 0.0 50.0
```

```
assign (residue 9 and name H3') (residue 11 and name H5) 4.5 0.0 50.0
assign (residue 9 and name H4') (residue 10 and name H8) 4.5 0.0 50.0
assign (residue 9 and name H5) (residue 10 and name H2') 4.5 0.0 50.0
assign (residue 9 and name H5) (residue 11 and name H5) 4.5 0.0 50.0
assign (residue 9 and name H5) (residue 11 and name H6) 4.5 0.0 50.0
assign (residue 9 and name H5') (residue 10 and name H2') 4.5 0.0 50.0
assign (residue 9 and name H5') (residue 10 and name H8) 4.5 0.0 50.0
assign (residue 9 and name H6) (residue 10 and name H2') 4.5 0.0 50.0
assign (residue 9 and name H6) (residue 11 and name H5) 4.5 0.0 50.0
assign (residue 10 and name H1') (residue 11 and name H2') 4.5 0.0 50.0
assign (residue 10 and name H1') (residue 12 and name H2) 4.5 0.0 50.0
assign (residue 10 and name H1') (residue 12 and name H8) 4.5 0.0 50.0
assign (residue 10 and name H4') (residue 11 and name H5) 4.5 0.0 50.0
assign (residue 10 and name H8) (residue 13 and name H8) 4.5 0.0 50.0
assign (residue 10 and name H8) (residue 17 and name H1') 4.5 0.0 50.0
assign (residue 11 and name H1') (residue 12 and name H4') 4.5 0.0 50.0
assign (residue 11 and name H1') (residue 13 and name H8) 4.5 0.0 50.0
assign (residue 11 and name H5) (residue 12 and name H2) 4.5 0.0 50.0
assign (residue 11 and name H5'') (residue 12 and name H8) 4.5 0.0 50.0
assign (residue 12 and name H1') (residue 13 and name H3') 4.5 0.0 50.0
assign (residue 12 and name H1') (residue 13 and name H5') 4.5 0.0 50.0
assign (residue 12 and name H1') (residue 15 and name H2') 4.5 0.0 50.0
assign (residue 12 and name H1') (residue 15 and name H4') 4.5 0.0 50.0
assign (residue 12 and name H2) (residue 14 and name H5) 4.5 0.0 50.0
assign (residue 12 and name H2) (residue 16 and name H5) 4.5 0.0 50.0
assign (residue 13 and name H1') (residue 14 and name H2') 4.5 0.0 50.0
assign (residue 13 and name H1') (residue 14 and name H4') 4.5 0.0 50.0
assign (residue 13 and name H2) (residue 14 and name H5) 4.5 0.0 50.0
assign (residue 13 and name H2) (residue 16 and name H5) 4.5 0.0 50.0
```

```
assign (residue 13 and name H2') (residue 15 and name H1') 4.5 0.0 50.0
assign (residue 13 and name H4') (residue 15 and name H1') 4.5 0.0 50.0
assign (residue 13 and name H5') (residue 15 and name H1') 4.5 0.0 50.0
assign (residue 13 and name H2) (residue 14 and name H2') 4.5 0.0 50.0
assign (residue 13 and name H8) (residue 14 and name H2') 4.5 0.0 50.0
assign (residue 13 and name H8) (residue 14 and name H3') 4.5 0.0 50.0
assign (residue 13 and name H8) (residue 14 and name H5) 4.5 0.0 50.0
assign (residue 13 and name H8) (residue 14 and name H5') 4.5 0.0 50.0
assign (residue 13 and name H8) (residue 15 and name H1') 4.5 0.0 50.0
assign (residue 14 and name H4') (residue 15 and name H1') 4.5 0.0 50.0
assign (residue 14 and name H4') (residue 15 and name H4') 4.5 0.0 50.0
assign (residue 14 and name H4') (residue 15 and name H5') 4.5 0.0 50.0
assign (residue 14 and name H1') (residue 16 and name H5) 4.5 0.0 50.0
assign (residue 15 and name H1') (residue 16 and name H5) 4.5 0.0 50.0
assign (residue 16 and name H5) (residue 22 and name H2') 4.5 0.0 50.0
assign (residue 16 and name H5) (residue 22 and name H3') 4.5 0.0 50.0
assign (residue 17 and name H8) (residue 22 and name H5) 4.5 0.0 50.0
assign (residue 18 and name H2) (residue 22 and name H1') 4.5 0.0 50.0
assign (residue 18 and name H8) (residue 19 and name H3') 4.5 0.0 50.0
assign (residue 18 and name H8) (residue 21 and name H5) 4.5 0.0 50.0
assign (residue 19 and name H1') (residue 21 and name H5) 4.5 0.0 50.0
assign (residue 19 and name H3') (residue 21 and name H5) 4.5 0.0 50.0
```

Dihedral angle restraints for the unmodified-970 RNA

```
assign (resid 1 and name c5') (resid 1 and name c4')
       (resid 1 and name c3') (resid 1 and name o3') 1 80 20 2
assign (resid 2 and name c5') (resid 2 and name c4')
       (resid 2 and name c3') (resid 2 and name o3') 1 80 20 2
assign (resid 3 and name c5') (resid 3 and name c4')
       (resid 3 and name c3') (resid 3 and name o3') 1 80 20 2
assign (resid 4 and name c5') (resid 4 and name c4')
       (resid 4 and name c3') (resid 4 and name o3') 1 80 20 2
assign (resid 5 and name c5') (resid 5 and name c4')
       (resid 5 and name c3') (resid 5 and name o3') 1 80 20 2
assign (resid 6 and name c5') (resid 6 and name c4')
       (resid 6 and name c3') (resid 6 and name o3') 1 80 20 2
assign (resid 7 and name c5') (resid 7 and name c4')
       (resid 7 and name c3') (resid 7 and name o3') 1 80 20 2
assign (resid 8 and name c5') (resid 8 and name c4')
       (resid 8 and name c3') (resid 8 and name o3') 1 80 20 2
assign (resid 9 and name c5') (resid 9 and name c4')
       (resid 9 and name c3') (resid 9 and name o3') 1 80 20 2
assign (resid 10 and name c5') (resid 10 and name c4')
       (resid 10 and name c3') (resid 10 and name o3') 1 80 20 2
assign (resid 11 and name c5') (resid 11 and name c4')
       (resid 11 and name c3') (resid 11 and name o3') 1 80 20 2
assign (resid 12 and name c5') (resid 12 and name c4')
       (resid 12 and name c3') (resid 12 and name o3') 1 157 20 2
assign (resid 13 and name c5') (resid 13 and name c4')
       (resid 13 and name c3') (resid 13 and name o3') 1 80 20 2
```

```
assign (resid 14 and name c5') (resid 14 and name c4')
       (resid 14 and name c3') (resid 14 and name o3') 1 80 20 2
assign (resid 15 and name c5') (resid 15 and name c4')
       (resid 15 and name c3') (resid 15 and name o3') 1 80 20 2
assign (resid 16 and name c5') (resid 16 and name c4')
       (resid 16 and name c3') (resid 16 and name o3') 1 80 20 2
assign (resid 17 and name c5') (resid 17 and name c4')
       (resid 17 and name c3') (resid 17 and name o3') 1 80 20 2
assign (resid 18 and name c5') (resid 18 and name c4')
       (resid 18 and name c3') (resid 18 and name o3') 1 80 20 2
assign (resid 19 and name c5') (resid 19 and name c4')
       (resid 19 and name c3') (resid 19 and name o3') 1 80 20 2
assign (resid 20 and name c5') (resid 20 and name c4')
       (resid 20 and name c3') (resid 20 and name o3') 1 80 20 2
assign (resid 21 and name c5') (resid 21 and name c4')
       (resid 21 and name c3') (resid 21 and name o3') 1 80 20 2
assign (resid 22 and name c5') (resid 22 and name c4')
       (resid 22 and name c3') (resid 22 and name o3') 1 80 20 2


!zeta and alpha
assign (resid 1 and name c3') (resid 1 and name o3')
       (resid 2 and name p) (resid 2 and name o5') 1 -71 10 2
assign (resid 1 and name o3') (resid 2 and name p)
       (resid 2 and name o5') (resid 2 and name c5') 1 -68 10 2

assign (resid 2 and name c3') (resid 2 and name o3')
       (resid 3 and name p) (resid 3 and name o5') 1 -71 10 2
assign (resid 2 and name o3') (resid 3 and name p)
```

```
       (resid 3 and name o5') (resid 3 and name c5') 1 -68 10 2

assign (resid 3 and name c3') (resid 3 and name o3')
       (resid 4 and name p) (resid 4 and name o5') 1 -71 10 2
assign (resid 3 and name o3') (resid 4 and name p)
       (resid 4 and name o5') (resid 4 and name c5') 1 -68 10 2

assign (resid 4 and name c3') (resid 4 and name o3')
       (resid 5 and name p) (resid 5 and name o5') 1 -71 10 2
assign (resid 4 and name o3') (resid 5 and name p)
       (resid 5 and name o5') (resid 5 and name c5') 1 -68 10 2

assign (resid 5 and name c3') (resid 5 and name o3')
       (resid 6 and name p) (resid 6 and name o5') 1 -71 10 2
assign (resid 5 and name o3') (resid 6 and name p)
       (resid 6 and name o5') (resid 6 and name c5') 1 -68 10 2

assign (resid 6 and name c3') (resid 6 and name o3')
       (resid 7 and name p) (resid 7 and name o5') 1 -71 10 2
assign (resid 6 and name o3') (resid 7 and name p)
       (resid 7 and name o5') (resid 7 and name c5') 1 -68 10 2

assign (resid 7 and name c3') (resid 7 and name o3')
       (resid 8 and name p) (resid 8 and name o5') 1 0 120 2
assign (resid 7 and name o3') (resid 8 and name p)
       (resid 8 and name o5') (resid 8 and name c5') 1 0 120 2

assign (resid 8 and name c3') (resid 8 and name o3')
       (resid 9 and name p) (resid 9 and name o5') 1 0 120 2
```

```
assign (resid 8 and name o3') (resid 9 and name p)
       (resid 9 and name o5') (resid 9 and name c5') 1 0 120 2

assign (resid 10 and name c3') (resid 10 and name o3')
       (resid 11 and name p) (resid 11 and name o5') 1 0 120 2
assign (resid 10 and name o3') (resid 11 and name p)
       (resid 11 and name o5') (resid 11 and name c5') 1 0 120 2

assign (resid 11 and name c3') (resid 11 and name o3')
       (resid 12 and name p) (resid 12 and name o5') 1 0 120 2
assign (resid 11 and name o3') (resid 12 and name p)
       (resid 12 and name o5') (resid 12 and name c5') 1 0 120 2

assign (resid 12 and name c3') (resid 12 and name o3')
       (resid 13 and name p) (resid 13 and name o5') 1 0 120 2
assign (resid 12 and name o3') (resid 13 and name p)
       (resid 13 and name o5') (resid 13 and name c5') 1 0 120 2

assign (resid 13 and name c3') (resid 13 and name o3')
       (resid 14 and name p) (resid 14 and name o5') 1 0 120 2
assign (resid 13 and name o3') (resid 14 and name p)
       (resid 14 and name o5') (resid 14 and name c5') 1 0 120 2

assign (resid 14 and name c3') (resid 14 and name o3')
       (resid 15 and name p) (resid 15 and name o5') 1 0 120 2
assign (resid 14 and name o3') (resid 15 and name p)
       (resid 15 and name o5') (resid 15 and name c5') 1 0 120 2

assign (resid 15 and name c3') (resid 15 and name o3')
```

```
       (resid 16 and name p) (resid 16 and name o5') 1 0 120 2
assign (resid 15 and name o3') (resid 16 and name p)
       (resid 16 and name o5') (resid 16 and name c5') 1 0 120 2

assign (resid 16 and name c3') (resid 16 and name o3')
       (resid 17 and name p) (resid 17 and name o5') 1 -71 10 2
assign (resid 16 and name o3') (resid 17 and name p)
       (resid 17 and name o5') (resid 17 and name c5') 1 -68 10 2

assign (resid 17 and name c3') (resid 17 and name o3')
       (resid 18 and name p) (resid 18 and name o5') 1 -71 10 2
assign (resid 17 and name o3') (resid 18 and name p)
       (resid 18 and name o5') (resid 18 and name c5') 1 -68 10 2

assign (resid 18 and name c3') (resid 18 and name o3')
       (resid 19 and name p) (resid 19 and name o5') 1 -71 10 2
assign (resid 18 and name o3') (resid 19 and name p)
       (resid 19 and name o5') (resid 19 and name c5') 1 -68 10 2

assign (resid 19 and name c3') (resid 19 and name o3')
       (resid 20 and name p) (resid 20 and name o5') 1 -71 10 2
assign (resid 19 and name o3') (resid 20 and name p)
       (resid 20 and name o5') (resid 20 and name c5') 1 -68 10 2

assign (resid 20 and name c3') (resid 20 and name o3')
       (resid 21 and name p) (resid 21 and name o5') 1 -71 10 2
assign (resid 20 and name o3') (resid 21 and name p)
       (resid 21 and name o5') (resid 21 and name c5') 1 -68 10 2
```

```
assign (resid 21 and name c3') (resid 21 and name o3')
       (resid 22 and name p) (resid 22 and name o5') 1 -71 10 2
assign (resid 21 and name o3') (resid 22 and name p)
       (resid 22 and name o5') (resid 22 and name c5') 1 -68 10 2

!beta
assign (resid 2 and name p) (resid 2 and name o5')
       (resid 2 and name c5') (resid 2 and name c4') 1 180 20 2
assign (resid 3 and name p) (resid 3 and name o5')
       (resid 3 and name c5') (resid 3 and name c4') 1 180 20 2
assign (resid 4 and name p) (resid 4 and name o5')
       (resid 4 and name c5') (resid 4 and name c4') 1 180 20 2
assign (resid 5 and name p) (resid 5 and name o5')
       (resid 5 and name c5') (resid 5 and name c4') 1 180 20 2
assign (resid 6 and name p) (resid 6 and name o5')
       (resid 6 and name c5') (resid 6 and name c4') 1 180 20 2
assign (resid 7 and name p) (resid 7 and name o5')
       (resid 7 and name c5') (resid 7 and name c4') 1 180 30 2
assign (resid 16 and name p) (resid 16 and name o5')
       (resid 16 and name c5') (resid 16 and name c4') 1 180 30 2
assign (resid 17 and name p) (resid 17 and name o5')
       (resid 17 and name c5') (resid 17 and name c4') 1 180 20 2
assign (resid 18 and name p) (resid 18 and name o5')
       (resid 18 and name c5') (resid 18 and name c4') 1 180 20 2
assign (resid 19 and name p) (resid 19 and name o5')
       (resid 19 and name c5') (resid 19 and name c4') 1 180 20 2
assign (resid 20 and name p) (resid 20 and name o5')
       (resid 20 and name c5') (resid 20 and name c4') 1 180 20 2
assign (resid 21 and name p) (resid 21 and name o5')
```

```
       (resid 21 and name c5') (resid 21 and name c4') 1 180 20 2
assign (resid 22 and name p) (resid 22 and name o5')
       (resid 22 and name c5') (resid 22 and name c4') 1 180 20 2

!gamma
assign (resid 1 and name o5') (resid 1 and name c5')
       (resid 1 and name c4') (resid 1 and name c3') 1 54 10 2
assign (resid 2 and name o5') (resid 2 and name c5')
       (resid 2 and name c4') (resid 2 and name c3') 1 54 10 2
assign (resid 3 and name o5') (resid 3 and name c5')
       (resid 3 and name c4') (resid 3 and name c3') 1 54 10 2
assign (resid 4 and name o5') (resid 4 and name c5')
       (resid 4 and name c4') (resid 4 and name c3') 1 54 10 2
assign (resid 5 and name o5') (resid 5 and name c5')
       (resid 5 and name c4') (resid 5 and name c3') 1 54 10 2
assign (resid 6 and name o5') (resid 6 and name c5')
       (resid 6 and name c4') (resid 6 and name c3') 1 54 10 2
assign (resid 7 and name o5') (resid 7 and name c5')
       (resid 7 and name c4') (resid 7 and name c3') 1 60 30 2
assign (resid 8 and name o5') (resid 8 and name c5')
       (resid 8 and name c4') (resid 8 and name c3') 1 60 30 2
assign (resid 9 and name o5') (resid 9 and name c5')
       (resid 9 and name c4') (resid 9 and name c3') 1 60 30 2
assign (resid 10 and name o5') (resid 10 and name c5')
       (resid 10 and name c4') (resid 10 and name c3') 1 60 30 2
assign (resid 14 and name o5') (resid 14 and name c5')
       (resid 14 and name c4') (resid 14 and name c3') 1 60 30 2
assign (resid 15 and name o5') (resid 15 and name c5')
       (resid 15 and name c4') (resid 15 and name c3') 1 60 30 2
```

```
assign (resid 16 and name o5') (resid 16 and name c5')
       (resid 16 and name c4') (resid 16 and name c3') 1 60 30 2
assign (resid 17 and name o5') (resid 17 and name c5')
       (resid 17 and name c4') (resid 17 and name c3') 1 60 30 2
assign (resid 18 and name o5') (resid 18 and name c5')
       (resid 18 and name c4') (resid 18 and name c3') 1 60 30 2
assign (resid 19 and name o5') (resid 19 and name c5')
       (resid 19 and name c4') (resid 19 and name c3') 1 60 30 2
assign (resid 20 and name o5') (resid 20 and name c5')
       (resid 20 and name c4') (resid 20 and name c3') 1 60 30 2
assign (resid 21 and name o5') (resid 21 and name c5')
       (resid 21 and name c4') (resid 21 and name c3') 1 54 10 2
assign (resid 22 and name o5') (resid 22 and name c5')
       (resid 22 and name c4') (resid 22 and name c3') 1 54 10 2


!epsilon
assign (resid 1 and name c4') (resid 1 and name c3')
       (resid 1 and name o3') (resid 2 and name p) 1 -155 10 2
assign (resid 2 and name c4') (resid 2 and name c3')
       (resid 2 and name o3') (resid 3 and name p) 1 -155 10 2
assign (resid 3 and name c4') (resid 3 and name c3')
       (resid 3 and name o3') (resid 4 and name p) 1 -155 10 2
assign (resid 4 and name c4') (resid 4 and name c3')
       (resid 4 and name o3') (resid 5 and name p) 1 -155 10 2
assign (resid 5 and name c4') (resid 5 and name c3')
       (resid 5 and name o3') (resid 6 and name p) 1 -155 10 2
assign (resid 6 and name c4') (resid 6 and name c3')
       (resid 6 and name o3') (resid 7 and name p) 1 -155 10 2
```

```
assign (resid 17 and name c4') (resid 17 and name c3')
       (resid 17 and name o3') (resid 18 and name p) 1 -155 10 2
assign (resid 18 and name c4') (resid 18 and name c3')
       (resid 18 and name o3') (resid 19 and name p) 1 -155 10 2
assign (resid 19 and name c4') (resid 19 and name c3')
       (resid 19 and name o3') (resid 20 and name p) 1 -155 10 2
assign (resid 20 and name c4') (resid 20 and name c3')
       (resid 20 and name o3') (resid 21 and name p) 1 -155 10 2
assign (resid 21 and name c4') (resid 21 and name c3')
       (resid 21 and name o3') (resid 22 and name p) 1 -155 10 2

! chi
assign (resid 1 and name O4') (resid 1 and name C1')
       (resid 1 and name N9 ) (resid 1 and name C4 ) 1 -160 10 2

assign (resid 2 and name O4') (resid 2 and name C1')
       (resid 2 and name N9 ) (resid 2 and name C4 ) 1 -160 10 2

assign (resid 3 and name O4') (resid 3 and name C1')
       (resid 3 and name N9 ) (resid 3 and name C4 ) 1 -160 10 2

assign (resid 4 and name O4') (resid 4 and name C1')
       (resid 4 and name N1 ) (resid 4 and name C2 ) 1 -160 10 2

assign (resid 5 and name O4') (resid 5 and name C1')
       (resid 5 and name N1 ) (resid 5 and name C2 ) 1 -160 10 2

assign (resid 6 and name O4') (resid 6 and name C1')
       (resid 6 and name N1 ) (resid 6 and name C2 ) 1 -160 10 2
```

```
assign (resid 7 and name O4') (resid 7 and name C1')
       (resid 7 and name N9 ) (resid 7 and name C4 ) 1 -160 10 2

assign (resid 8 and name O4') (resid 8 and name C1')
       (resid 8 and name N9 ) (resid 8 and name C4 ) 1 -160 50 2

assign (resid 9 and name O4') (resid 9 and name C1')
       (resid 9 and name N1 ) (resid 9 and name C2 ) 1 -160 50 2

assign (resid 10 and name O4') (resid 10 and name C1')
       (resid 10 and name N9 ) (resid 10 and name C4 ) 1 -160 50 2

assign (resid 11 and name O4') (resid 11 and name C1')
       (resid 11 and name N1 ) (resid 11 and name C2 ) 1 -160 50 2

assign (resid 12 and name O4') (resid 12 and name C1')
       (resid 12 and name N9 ) (resid 12 and name C4 ) 1 -160 50 2

assign (resid 13 and name O4') (resid 13 and name C1')
       (resid 13 and name N9 ) (resid 13 and name C4 ) 1 -160 50 2

assign (resid 14 and name O4') (resid 14 and name C1')
       (resid 14 and name N1 ) (resid 14 and name C2 ) 1 -160 50 2

assign (resid 15 and name O4') (resid 15 and name C1')
       (resid 15 and name N9 ) (resid 15 and name C4 ) 1 -160 50 2

assign (resid 16 and name O4') (resid 16 and name C1')
```

```
       (resid 16 and name N1 ) (resid 16 and name C2 ) 1 -160 10 2

assign (resid 17 and name O4') (resid 17 and name C1')
       (resid 17 and name N9 ) (resid 17 and name C4 ) 1 -160 10 2

assign (resid 18 and name O4') (resid 18 and name C1')
       (resid 18 and name N9 ) (resid 18 and name C4 ) 1 -160 10 2

assign (resid 19 and name O4') (resid 19 and name C1')
       (resid 19 and name N9 ) (resid 19 and name C4 ) 1 -160 10 2

assign (resid 20 and name O4') (resid 20 and name C1')
       (resid 20 and name N1 ) (resid 20 and name C2 ) 1 -160 10 2

assign (resid 21 and name O4') (resid 21 and name C1')
       (resid 21 and name N1 ) (resid 21 and name C2 ) 1 -160 10 2

assign (resid 22 and name O4') (resid 22 and name C1')
       (resid 22 and name N1 ) (resid 22 and name C2 ) 1 -160 10 2
```

APPENDIX 4

Structural Restraints of the Modified-970 RNA

Distance restraints for the modified-970

```
! Base-pairs
assign (residue 2 and name O6) (residue 1 and name O6) 4.0 1.0 0.5
assign (residue 2 and name N2) (residue 1 and name N2) 4.0 1.0 0.5
assign (residue 2 and name O6) (residue 3 and name O6) 4.0 1.0 0.5
assign (residue 2 and name N2) (residue 3 and name N2) 4.0 1.0 0.5
assign (residue 3 and name O6) (residue 4 and name O4) 4.0 1.0 0.5
assign (residue 3 and name N2) (residue 4 and name O2) 4.0 1.0 0.5
assign (residue 4 and name O4) (residue 5 and name O4) 4.0 1.0 0.5
assign (residue 4 and name O2) (residue 5 and name O2) 4.0 1.0 0.5
assign (residue 5 and name O4) (residue 6 and name N4) 4.0 1.0 0.5
assign (residue 5 and name O2) (residue 6 and name O2) 4.0 1.0 0.5
assign (residue 6 and name N4) (residue 7 and name O6) 4.0 1.0 0.5
assign (residue 6 and name O2) (residue 7 and name N2) 4.0 1.0 1.0
assign (residue 16 and name N4) (residue 17 and name O6) 4.0 1.0 0.5
assign (residue 16 and name O2) (residue 17 and name N2) 4.0 1.0 1.0
assign (residue 17 and name O6) (residue 18 and name N6) 4.0 1.0 0.5
assign (residue 17 and name N2) (residue 18 and name H2) 4.0 1.0 0.5
assign (residue 18 and name N6) (residue 19 and name N6) 4.0 1.0 0.5
assign (residue 18 and name H2) (residue 19 and name H2) 4.0 1.0 0.5
assign (residue 19 and name N6) (residue 20 and name N4) 4.0 1.0 0.5
assign (residue 19 and name H2) (residue 20 and name O2) 4.0 1.0 0.5
assign (residue 20 and name N4) (residue 21 and name N4) 3.5 1.0 0.0
```

```
assign (residue 20 and name O2) (residue 21 and name O2) 4.0 1.0 0.5
assign (residue 21 and name N4) (residue 22 and name N4) 3.5 1.0 0.0
assign (residue 21 and name O2) (residue 22 and name O2) 4.0 1.0 0.5

!stem stacking distances
assign (residue 1 and name H1) (residue 2 and name H1) 3.5 1.5 0.5
assign (residue 2 and name H1) (residue 3 and name H1) 3.5 1.5 0.5
assign (residue 3 and name H1) (residue 4 and name H3) 3.5 1.5 0.5
assign (residue 4 and name H3) (residue 5 and name H3) 3.5 1.5 0.5
assign (residue 5 and name H3) (residue 17 and name H1) 3.5 0.5 1.5
assign (residue 7 and name H1) (residue 17 and name H1) 3.5 0.5 1.5

! Base pair distances
assign (residue 1 and name C1') (residue 22 and name C1') 10.8 0.2 0.2
assign (residue 1 and name C8) (residue 22 and name C6) 9.9 0.2 0.2
assign (residue 2 and name C1') (residue 21 and name C1') 10.8 0.2 0.2
assign (residue 2 and name C8) (residue 21 and name C6) 9.9 0.2 0.2
assign (residue 3 and name C1') (residue 20 and name C1') 10.8 0.2 0.2
assign (residue 3 and name C8) (residue 20 and name C6) 9.9 0.2 0.2
assign (residue 4 and name C1') (residue 19 and name C1') 10.5 0.2 0.2
assign (residue 4 and name C6) (resid 19 and name C8) 9.7 0.2 0.2
assign (residue 5 and name C1') (residue 18 and name C1') 10.5 0.2 0.2
assign (residue 5 and name C6) (resid 18 and name C8) 9.7 0.2 0.2
assign (residue 6 and name C1') (residue 17 and name C1') 10.8 0.2 0.2
assign (residue 6 and name C6) (residue 17 and name C8) 9.9 0.2 0.2
assign (residue 7 and name C1') (residue 16 and name C1') 10.8 0.2 0.2
assign (residue 7 and name C8) (residue 16 and name C6) 9.9 0.2 0.2

! Base Pair restraints
```

```
! for G1/ C22  base pair
assign (resid 1 and name N1) (resid 22 and name N3) 2.91 0.2 0.2
assign (resid 1 and name O6) (resid 22 and name N4) 2.71 0.2 0.2
assign (resid 1 and name N2) (resid 22 and name O2) 3.08 0.2 0.2
assign (resid 1 and name H1) (resid 22 and name N3) 1.89 0.2 0.2
assign (resid 1 and name O6) (resid 22 and name H42) 1.71 0.2 0.2
assign (resid 1 and name H22) (resid 22 and name O2) 2.08 0.2 0.2

! for G2/ C21  base pair
assign (resid 2 and name N1) (resid 21 and name N3) 2.91 0.2 0.2
assign (resid 2 and name O6) (resid 21 and name N4) 2.71 0.2 0.2
assign (resid 2 and name N2) (resid 21 and name O2) 3.08 0.2 0.2
assign (resid 2 and name H1) (resid 21 and name N3) 1.89 0.2 0.2
assign (resid 2 and name O6) (resid 21 and name H42) 1.71 0.2 0.2
assign (resid 2 and name H22) (resid 21 and name O2) 2.08 0.2 0.2

! for G3/ C20  base pair
assign (resid 3 and name N1) (resid 20 and name N3) 2.91 0.2 0.2
assign (resid 3 and name O6) (resid 20 and name N4) 2.71 0.2 0.2
assign (resid 3 and name N2) (resid 20 and name O2) 3.08 0.2 0.2
assign (resid 3 and name H1) (resid 20 and name N3) 1.89 0.2 0.2
assign (resid 3 and name O6) (resid 20 and name H42) 1.71 0.2 0.2
assign (resid 3 and name H22) (resid 20 and name O2) 2.08 0.2 0.2

! for U4/A19 base pair
assign (resid 19 and name N1) (resid 4 and name N3) 2.95 0.2 0.2
assign (resid 19 and name N6) (resid 4 and name O4) 2.83 0.2 0.2
assign (resid 19 and name N1) (resid 4 and name H3) 1.93 0.2 0.2
assign (resid 19 and name H62) (resid 4 and name O4) 1.82 0.2 0.2
```

```
! for U5/A18 base pair
assign (resid 18 and name N1) (resid 5 and name N3) 2.95 0.2 0.2
assign (resid 18 and name N6) (resid 5 and name O4) 2.83 0.2 0.2
assign (resid 18 and name N1) (resid 5 and name H3) 1.93 0.2 0.2
assign (resid 18 and name H62) (resid 5 and name O4) 1.82 0.2 0.2

! for C6/G17 base pair
assign (resid 17 and name N1) (resid 6 and name N3) 2.91 0.2 0.2
assign (resid 17 and name O6) (resid 6 and name N4) 2.71 0.2 0.2
assign (resid 17 and name N2) (resid 6 and name O2) 3.08 0.2 0.2
assign (resid 17 and name H1) (resid 6 and name N3) 1.89 0.2 0.2
assign (resid 17 and name O6) (resid 6 and name H42) 1.71 0.2 0.2
assign (resid 17 and name H22) (resid 6 and name O2) 2.08 0.2 0.2

! for G7/ C16 base pair  ! change on 10/10/10
assign (resid 7 and name N1) (resid 16 and name N3) 2.91 0.2 0.5
assign (resid 7 and name O6) (resid 16 and name N4) 2.71 0.2 0.5
assign (resid 7 and name N2) (resid 16 and name O2) 3.08 0.2 0.5
assign (resid 7 and name H1) (resid 16 and name N3) 1.89 0.2 0.5
assign (resid 7 and name O6) (resid 16 and name H42) 1.71 0.2 0.5
assign (resid 7 and name H22) (resid 16 and name O2) 2.08 0.2 0.5

! Add from NOE-H2O
assign (residue 3 and name H1) (residue 21 and name H41) 4.0 0.2 0.5
assign (residue 3 and name H1) (residue 21 and name H42) 4.0 0.2 0.5
assign (residue 4 and name H3) (residue 19 and name H2) 3.0 0.5 0.5
assign (residue 5 and name H3) (residue 18 and name H2) 3.0 0.5 0.5
assign (residue 5 and name H3) (residue 19 and name H2) 4.0 0.5 0.5
```

```
assign (residue 7 and name H1') (residue 17 and name H1) 4.0 0.5 1.0
assign (residue 7 and name H1) (residue 16 and name H41) 4.0 0.5 0.5


! Residue 1GUA
! 1GUA
!Intra
assign (residue 1 and name H8) (residue 1 and name H1') 3.5 0.5 0.5
assign (residue 1 and name H8) (residue 1 and name H2') 3.5 0.5 0.5
assign (residue 1 and name H8) (residue 1 and name H3') 3.0 0.5 0.5
assign (residue 1 and name H8) (residue 1 and name H5') 3.5 0.5 0.5
!Inter
assign (residue 1 and name H8) (residue 2 and name H8) 4.5 0.5 0.5
assign (residue 1 and name H1') (residue 2 and name H8) 4.5 0.5 0.5
assign (residue 1 and name H2') (residue 2 and name H8) 2.0 0.2 0.5

! 2GUA
! Intra
assign (residue 2 and name H8) (residue 2 and name H1') 3.5 0.5 0.5
assign (residue 2 and name H8) (residue 2 and name H2') 4.0 0.5 0.5
assign (residue 2 and name H8) (residue 2 and name H3') 2.5 0.5 0.5
assign (residue 2 and name H8) (residue 2 and name H5') 3.5 0.5 0.5
! Inter
assign (residue 2 and name H8) (residue 3 and name H8) 4.5 0.5 0.5
assign (residue 2 and name H2') (residue 3 and name H1') 4.0 0.5 0.5
assign (residue 2 and name H1') (residue 3 and name H8) 4.5 0.5 0.5
assign (residue 2 and name H2') (residue 3 and name H8) 2.0 0.2 0.5

! 3GUA
```

```
! Intra
assign (residue 3 and name H8) (residue 3 and name H1') 3.5 0.5 0.5
assign (residue 3 and name H8) (residue 3 and name H2') 4.0 0.5 0.5
assign (residue 3 and name H8) (residue 3 and name H3') 3.0 0.5 0.5
assign (residue 3 and name H8) (residue 3 and name H5') 3.5 0.5 0.5
! Inter
assign (residue 3 and name H8) (residue 4 and name H5) 4.5 0.5 0.5
assign (residue 3 and name H2') (residue 4 and name H5) 3.5 0.5 0.5
assign (residue 3 and name H3') (residue 4 and name H5) 3.5 0.5 0.5
assign (residue 3 and name H1') (residue 4 and name H6) 4.5 0.5 0.5
assign (residue 3 and name H2') (residue 4 and name H6) 2.0 0.2 0.5

! 4URI
! Intra
assign (residue 4 and name H5) (residue 4 and name H3') 3.5 0.5 0.5
assign (residue 4 and name H6) (residue 4 and name H1') 3.5 0.5 0.5
assign (residue 4 and name H6) (residue 4 and name H3') 2.5 0.5 0.5
! Inter
assign (residue 4 and name H2') (residue 5 and name H1') 3.5 0.5 0.5
assign (residue 4 and name H5) (residue 5 and name H5) 3.5 0.5 0.5
assign (residue 4 and name H2') (residue 5 and name H6) 2.0 0.2 0.5

! 5URI
! Intra
assign (residue 5 and name H1') (residue 5 and name H4') 3.5 0.5 0.5
assign (residue 5 and name H6) (residue 5 and name H1') 3.5 0.5 0.5
assign (residue 5 and name H6) (residue 5 and name H3') 2.5 0.5 0.5
assign (residue 5 and name H6) (residue 5 and name H5') 3.0 0.5 0.5
! Inter
```

```
assign (residue 5 and name H2') (residue 6 and name H6) 2.0 0.2 0.5

! 6CYT
! Intra
assign (residue 6 and name H1') (residue 6 and name H4') 3.5 0.5 0.5
assign (residue 6 and name H6) (residue 6 and name H1') 4.5 0.5 0.5
assign (residue 6 and name H6) (residue 6 and name H2') 3.5 0.5 0.5
! Inter
!assign (residue 6 and name H2') (residue 7 and name H1') 3.5 0.5 0.5
assign (residue 6 and name H1') (residue 7 and name H8) 4.5 0.5 0.5
assign (residue 6 and name H2') (residue 7 and name H8) 2.0 0.5 0.8

! 7GUA
! Intra
assign (residue 7 and name H8) (residue 7 and name H1') 3.5 0.5 0.5
assign (residue 7 and name H8) (residue 7 and name H2') 4.0 0.5 0.5
assign (residue 7 and name H8) (residue 7 and name H3') 3.5 0.5 0.5
! Inter
assign (residue 7 and name H1') (residue 8 and name H8) 4.5 0.5 0.5
assign (residue 7 and name H2') (residue 8 and name H8) 2.0 0.5 0.5

! 8ADE
! Intra
assign (residue 8 and name H1') (residue 8 and name H2') 3.0 0.5 0.5
assign (residue 8 and name H1') (residue 8 and name H4') 3.0 0.5 0.5
assign (residue 8 and name H8) (residue 8 and name H1') 3.5 0.5 0.5
assign (residue 8 and name H2) (residue 8 and name H1') 4.5 0.5 0.5
assign (residue 8 and name H8) (residue 8 and name H2') 3.5 0.5 1.0
```

```
! Inter
assign (residue 8 and name H2) (residue 9 and name H1') 4.5 1.0 1.0
assign (residue 8 and name H2') (residue 9 and name H1') 4.5 0.5 0.8
assign (residue 8 and name H2) (residue 14 and name H1') 4.5 0.5 0.5
assign (residue 8 and name H2) (residue 16 and name H1') 4.5 0.5 1.5
assign (residue 8 and name H1') (residue 9 and name H6) 4.5 0.5 1.0
assign (residue 8 and name H2') (residue 9 and name H6) 3.5 0.5 1.0
assign (residue 8 and name H2') (residue 9 and name H5) 4.5 0.5 0.8

! 9URI
! Intra
assign (residue 9 and name H1') (residue 9 and name H2') 2.5 0.5 0.5
assign (residue 9 and name H1') (residue 9 and name H4') 3.5 0.5 0.5
assign (residue 9 and name H1') (residue 9 and name H6) 3.5 0.5 0.5
assign (residue 9 and name H5) (residue 9 and name H2') 3.5 0.5 0.8
assign (residue 9 and name H6) (residue 9 and name H2') 2.0 0.2 0.5
! Inter
assign (residue 9 and name H1') (residue 13 and name H2) 4.0 0.5 0.5
assign (residue 9 and name H4') (residue 13 and name H8) 3.5 0.5 1.0

! 10 2MG
! Intra
assign (residue 10 and name H33) (residue 10 and name H1') 4.5 0.5 0.5
assign (residue 10 and name H8) (residue 10 and name H1') 3.5 0.5 0.5
assign (residue 10 and name H8) (residue 10 and name H3') 3.0 0.5 0.5
assign (residue 10 and name H8) (residue 10 and name H4') 4.0 0.5 0.5
assign (residue 10 and name H8) (residue 10 and name H5') 4.0 0.5 0.5
assign (residue 10 and name H1') (residue 10 and name H2') 2.5 0.5 0.5
! Inter
```

```
assign (residue 10 and name H31) (residue 11 and name H1') 3.5 0.5 0.5
assign (residue 10 and name H32) (residue 11 and name H6) 4.0 0.5 0.8
assign (residue 10 and name H8) (residue 11 and name H13) 3.5 0.5 0.5
assign (residue 10 and name H1') (residue 11 and name H6) 4.5 0.5 0.5
assign (residue 1O and name H3') (residue 11 and name H11) 3.0 0.5 0.5
assign (residue 10 and name H3') (residue 11 and name H6) 4.0 0.5 0.5

! 11 5MC
! Intra
assign (residue 11 and name H1') (residue 11 and name H2') 3.0 0.5 0.5
assign (residue 11 and name H1') (residue 11 and name H3') 3.5 0.5 0.5
assign (residue 11 and name H1') (residue 11 and name H4') 3.5 0.5 0.5
assign (residue 11 and name H6) (residue 11 and name H1') 3.5 0.5 0.5
! Inter
assign (residue 11 and name H12) (residue 12 and name H2) 3.5 0.5 1.0
assign (residue 11 and name H1') (residue 12 and name H8) 4.5 0.5 1.0
assign (residue 11 and name H2') (residue 12 and name H2) 2.5 0.5 0.5
assign (residue 11 and name H3') (residue 12 and name H8) 4.5 0.5 1.5

! 12ADE
! Intra
assign (residue 12 and name H8) (residue 12 and name H2') 3.5 0.5 0.5
assign (residue 12 and name H8) (residue 12 and name H3') 4.5 0.5 1.5
assign (residue 12 and name H8) (residue 12 and name H5') 4.5 0.5 1.0

! 13ADE
! Intra
assign (residue 13 and name H1') (residue 13 and name H2') 2.5 0.5 0.5
assign (residue 13 and name H1') (residue 13 and name H3') 3.5 0.5 0.5
```

```
assign (residue 13 and name H2) (residue 13 and name H1') 4.0 0.5 0.8
assign (residue 13 and name H8) (residue 13 and name H1') 3.5 0.5 0.5
assign (residue 13 and name H8) (residue 13 and name H2') 3.5 0.5 0.5
assign (residue 13 and name H8) (residue 13 and name H3') 3.5 0.5 0.8
! Inter
assign (residue 13 and name H2) (residue 14 and name H1') 4.5 0.5 1.5
assign (residue 13 and name H1') (residue 14 and name H1') 4.5 1.0 1.0
assign (residue 13 and name H1') (residue 14 and name H6) 4.5 0.5 1.0
assign (residue 13 and name H2') (residue 14 and name H6) 4.5 0.5 1.0

! 14CYT
! Intra
assign (residue 14 and name H5) (residue 14 and name H3') 4.5 0.5 0.5
assign (residue 14 and name H6) (residue 14 and name H2') 2.0 0.2 0.5
assign (residue 14 and name H6) (residue 14 and name H3') 3.0 0.5 0.5
assign (residue 14 and name H6) (residue 14 and name H4') 4.5 0.5 0.5

! 15GUA
! Intra
assign (residue 15 and name H8) (residue 15 and name H2') 2.5 0.5 1.0
assign (residue 15 and name H8) (residue 15 and name H3') 4.0 0.5 0.5
! Inter
assign (residue 15 and name H4') (residue 16 and name H5) 4.5 0.5 0.5
assign (residue 15 and name H4') (residue 16 and name H6) 4.0 0.8 0.5

! 16 CYT
! Intra
assign (residue 16 and name H6) (residue 16 and name H1') 3.5 0.5 0.5
assign (residue 16 and name H6) (residue 16 and name H2') 3.5 0.5 0.5
```

```
assign (residue 16 and name H6) (residue 16 and name H3') 3.0 0.5 0.5
assign (residue 16 and name H6) (residue 16 and name H5') 3.0 0.5 0.5
! Inter
assign (residue 16 and name H1') (residue 17 and name H8) 4.5 0.5 0.5
assign (residue 16 and name H2') (residue 17 and name H8) 2.0 0.5 0.8

! 17 GUA
! Intra
assign (residue 17 and name H8) (residue 17 and name H1') 4.0 0.5 0.5
assign (residue 17 and name H8) (residue 17 and name H2') 4.0 0.5 0.5
! Inter
assign (residue 17 and name H1') (residue 18 and name H8) 4.5 0.5 0.5
assign (residue 17 and name H2') (residue 18 and name H8) 2.0 0.2 0.5

! 18 ADE
! Intra
assign (residue 18 and name H1') (residue 18 and name H4') 3.5 0.5 0.5
assign (residue 18 and name H2) (residue 18 and name H1') 4.5 0.5 0.5
assign (residue 18 and name H8) (residue 18 and name H1') 3.5 0.5 0.5
assign (residue 18 and name H8) (residue 18 and name H2') 4.0 0.5 0.5
assign (residue 18 and name H8) (residue 18 and name H3') 2.5 0.5 0.5
! Inter
assign (residue 18 and name H2) (residue 19 and name H1') 3.5 0.5 0.5
assign (residue 18 and name H1') (residue 19 and name H8) 4.5 0.5 0.5
assign (residue 18 and name H2') (residue 19 and name H8) 2.0 0.2 0.5

! 19 ADE
! Intra
assign (residue 19 and name H1') (residue 19 and name H4') 3.5 0.5 0.5
```

```
assign (residue 19 and name H8) (residue 19 and name H1') 3.5 0.5 0.5
assign (residue 19 and name H8) (residue 19 and name H3') 3.0 0.5 0.5
! Inter
assign (residue 19 and name H2') (residue 20 and name H1') 4.0 0.5 0.5
assign (residue 19 and name H2') (residue 20 and name H5) 3.5 0.5 0.5
assign (residue 19 and name H1') (residue 20 and name H6) 4.5 0.5 0.5
assign (residue 19 and name H2') (residue 20 and name H6) 2.0 0.2 0.5

! 20 CYT
! Intra
assign (residue 20 and name H1') (residue 20 and name H4') 3.5 0.5 0.5
assign (residue 20 and name H6) (residue 20 and name H1') 3.5 0.5 0.5
assign (residue 20 and name H6) (residue 20 and name H2') 3.5 0.5 0.5
assign (residue 20 and name H6) (residue 20 and name H3') 2.5 0.5 0.5
assign (residue 20 and name H6) (residue 20 and name H4') 4.0 0.5 0.5
assign (residue 20 and name H5) (residue 20 and name H3') 4.5 0.5 0.5
! Inter
assign (residue 20 and name H6) (residue 21 and name H5) 4.5 0.5 0.5
assign (residue 20 and name H3') (residue 21 and name H5) 4.5 0.5 0.5
assign (residue 20 and name H1') (residue 21 and name H6) 4.5 0.5 0.5
assign (residue 20 and name H2') (residue 21 and name H6) 2.0 0.2 0.5

! 21 CYT
! Intra
assign (residue 21 and name H5) (residue 21 and name H3') 4.5 0.5 0.5
assign (residue 21 and name H6) (residue 21 and name H1') 3.5 0.5 0.5
assign (residue 21 and name H6) (residue 21 and name H2') 3.5 0.5 0.5
assign (residue 21 and name H6) (residue 21 and name H3') 2.5 0.5 0.5
! Inter
```

```
assign (residue 21 and name H6) (residue 22 and name H5) 4.5 0.5 0.5
assign (residue 21 and name H2') (residue 22 and name H1') 3.5 0.5 0.5
assign (residue 21 and name H2') (residue 22 and name H5) 3.5 0.5 0.5
assign (residue 21 and name H1') (residue 22 and name H6) 4.5 0.5 0.5
assign (residue 21 and name H2') (residue 22 and name H6) 2.0 0.2 0.5
assign (residue 21 and name H5) (residue 22 and name H6) 4.5 0.5 0.5

! 22 CYT
! Intra
assign (residue 22 and name H5) (residue 22 and name H3') 4.5 0.5 0.5
assign (residue 22 and name H6) (residue 22 and name H1') 3.5 0.5 0.5
assign (residue 22 and name H6) (residue 22 and name H2') 3.5 0.5 0.5
assign (residue 22 and name H6) (residue 22 and name H3') 2.5 0.5 0.5

! Un-NOE's
assign (residue 1 and name H1') (residue 6 and name H1') 4.5 0.0 50.0
assign (residue 1 and name H8) (residue 6 and name H1') 4.5 0.0 50.0
assign (residue 1 and name H2') (residue 18 and name H2) 4.5 0.0 50.0
assign (residue 1 and name H8) (residue 18 and name H4') 4.5 0.0 50.0
assign (residue 2 and name H1') (residue 4 and name H1') 4.5 0.0 50.0
assign (residue 2 and name H1') (residue 6 and name H1') 4.5 0.0 50.0
assign (residue 2 and name H2') (residue 5 and name H6) 4.5 0.0 50.0
assign (residue 2 and name H2') (residue 18 and name H2) 4.5 0.0 50.0
assign (residue 2 and name H3') (residue 4 and name H3') 4.5 0.0 50.0
assign (residue 2 and name H3') (residue 19 and name H1') 4.5 0.0 50.0
assign (residue 2 and name H3') (residue 19 and name H2) 4.5 0.0 50.0
assign (residue 2 and name H8) (residue 18 and name H2) 4.5 0.0 50.0
assign (residue 3 and name H1') (residue 5 and name H6) 4.5 0.0 50.0
assign (residue 3 and name H1') (residue 6 and name H5) 4.5 0.0 50.0
```

```
assign (residue 3 and name H1') (residue 6 and name H6) 4.5 0.0 50.0
assign (residue 3 and name H2') (residue 15 and name H1') 4.5 0.0 50.0
assign (residue 3 and name H3') (residue 5 and name H6) 4.5 0.0 50.0
assign (residue 3 and name H8) (residue 5 and name H1') 4.5 0.0 50.0
assign (residue 3 and name H8) (residue 6 and name H6) 4.5 0.0 50.0
assign (residue 4 and name H5) (residue 6 and name H1') 4.5 0.0 50.0
assign (residue 4 and name H5) (residue 6 and name H5) 4.5 0.0 50.0
assign (residue 4 and name H1') (residue 7 and name H8) 4.5 0.0 50.0
assign (residue 4 and name H5) (residue 18 and name H2) 4.5 0.0 50.0
assign (residue 4 and name H5) (residue 21 and name H1') 4.5 0.0 50.0
assign (residue 4 and name H5) (residue 22 and name H2') 4.5 0.0 50.0
assign (residue 5 and name H5) (residue 7 and name H8) 4.5 0.0 50.0
assign (residue 5 and name H6) (residue 7 and name H8) 4.5 0.0 50.0
assign (residue 5 and name H5) (residue 22 and name H1') 4.5 0.0 50.0
assign (residue 5 and name H5) (residue 21 and name H2') 4.5 0.0 50.0
assign (residue 5 and name H5) (residue 21 and name H4') 4.5 0.0 50.0
assign (residue 5 and name H6) (residue 7 and name H8) 4.5 0.0 50.0
assign (residue 5 and name H6) (residue 22 and name H1') 4.5 0.0 50.0
assign (residue 6 and name H5) (residue 22 and name H1') 4.5 0.0 50.0
assign (residue 6 and name H6) (residue 22 and name H1') 4.5 0.0 50.0
assign (residue 7 and name H1') (residue 18 and name H2) 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 14 and name H5) 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 14 and name H6) 4.5 0.0 50.0
assign (residue 8 and name H2) (residue 16 and name H6) 4.5 0.0 50.00
assign (residue 8 and name H2') (residue 13 and name H2) 4.5 0.0 50.0
assign (residue 8 and name H3') (residue 13 and name H2) 4.5 0.0 50.0
assign (residue 9 and name H2') (residue 13 and name H2) 4.5 0.0 50.0
assign (residue 9 and name H2') (residue 14 and name H5) 4.5 0.0 50.0
assign (residue 9 and name H3') (residue 10 and name H2') 4.5 0.0 50.0
```

```
assign (residue 9 and name H4') (residue 10 and name H8) 4.5 0.0 50.0
assign (residue 9 and name H4') (residue 12 and name H2) 4.5 0.0 50.0
assign (residue 9 and name H5) (residue 10 and name H2') 4.5 0.0 50.0
assign (residue 9 and name H5) (residue 11 and name H6) 4.5 0.0 50.0
assign (residue 9 and name H5) (residue 12 and name H2) 4.5 0.0 50.0
assign (residue 9 and name H5) (residue 13 and name H2) 4.5 0.0 50.0
assign (residue 9 and name H5) (residue 14 and name H5) 4.5 0.0 50.0
assign (residue 9 and name H5') (residue 10 and name H8) 4.5 0.0 50.0
assign (residue 9 and name H6) (residue 10 and name H2') 4.5 0.0 50.0
assign (residue 9 and name H6) (residue 12 and name H2) 4.5 0.0 50.0
assign (residue 9 and name H6) (residue 13 and name H2) 4.5 0.0 50.0
assign (residue 9 and name H5') (residue 10 and name H2') 4.5 0.0 50.0
assign (residue 10 and name H1') (residue 11 and name H2') 4.5 0.0 50.0
assign (residue 10 and name H1') (residue 12 and name H2) 4.5 0.0 50.0
assign (residue 10 and name H1') (residue 12 and name H8) 4.5 0.0 50.0
assign (residue 10 and name H3') (residue 12 and name H2) 4.5 0.0 50.0
assign (residue 10 and name H8) (residue 13 and name H2) 4.5 0.0 50.0
assign (residue 10 and name H31) (residue 12 and name H2) 4.5 0.0 50.0
assign (residue 10 and name H32) (residue 12 and name H2) 4.5 0.0 50.0
assign (residue 10 and name H33) (residue 12 and name H2) 4.5 0.0 50.0
assign (residue 11 and name H1') (residue 12 and name H4') 4.5 0.0 50.0
assign (residue 11 and name H1') (residue 13 and name H8) 4.5 0.0 50.0
assign (residue 12 and name H1') (residue 13 and name H3') 4.5 0.0 50.0
assign (residue 12 and name H1') (residue 13 and name H5') 4.5 0.0 50.0
assign (residue 12 and name H1') (residue 15 and name H2') 4.5 0.0 50.0
assign (residue 12 and name H1') (residue 15 and name H4') 4.5 0.0 50.0
assign (residue 12 and name H2) (residue 13 and name H2) 4.5 0.0 50.0
assign (residue 12 and name H2) (residue 16 and name H5) 4.5 0.0 50.0
assign (residue 13 and name H1') (residue 14 and name H4') 4.5 0.0 50.0
```

```
assign (residue 13 and name H2) (residue 14 and name H5) 4.5 0.0 50.0
assign (residue 13 and name H2) (residue 16 and name H5) 4.5 0.0 50.0
assign (residue 13 and name H2') (residue 15 and name H1') 4.5 0.0 50.0
assign (residue 13 and name H4') (residue 15 and name H1') 4.5 0.0 50.0
assign (residue 13 and name H5') (residue 15 and name H1') 4.5 0.0 50.0
assign (residue 13 and name H2) (residue 14 and name H2') 4.5 0.0 50.0
assign (residue 13 and name H8) (residue 14 and name H2') 4.5 0.0 50.0
assign (residue 13 and name H8) (residue 14 and name H3') 4.5 0.0 50.0
assign (residue 13 and name H8) (residue 14 and name H5) 4.5 0.0 50.0
assign (residue 13 and name H8) (residue 14 and name H5') 4.5 0.0 50.0
assign (residue 13 and name H8) (residue 15 and name H1') 4.5 0.0 50.0
assign (residue 14 and name H4') (residue 15 and name H1') 4.5 0.0 50.0
assign (residue 14 and name H4') (residue 15 and name H4') 4.5 0.0 50.0
assign (residue 14 and name H4') (residue 15 and name H5') 4.5 0.0 50.0
assign (residue 14 and name H1') (residue 16 and name H5) 4.5 0.0 50.0
assign (residue 14 and name H6) (residue 15 and name H8) 4.5 0.0 50.0
assign (residue 15 and name H1') (residue 16 and name H5) 4.5 0.0 50.0
assign (residue 15 and name H2') (residue 16 and name H5') 4.5 0.0 50.0
assign (residue 16 and name H5) (residue 22 and name H2') 4.5 0.0 50.0
assign (residue 16 and name H5) (residue 22 and name H3') 4.5 0.0 50.0
assign (residue 17 and name H8) (residue 22 and name H5) 4.5 0.0 50.0
assign (residue 18 and name H2) (residue 22 and name H1') 4.5 0.0 50.0
assign (residue 18 and name H8) (residue 19 and name H3') 4.5 0.0 50.0
assign (residue 18 and name H8) (residue 21 and name H5) 4.5 0.0 50.0
assign (residue 19 and name H1') (residue 21 and name H5) 4.5 0.0 50.0
assign (residue 19 and name H3') (residue 21 and name H5) 4.5 0.0 50.0
```

Dihedral angle restraints for the modified-970

```
!sugar pucker delta
assign (resid 1 and name c5') (resid 1 and name c4')
       (resid 1 and name c3') (resid 1 and name o3') 1 80 20 2
assign (resid 2 and name c5') (resid 2 and name c4')
       (resid 2 and name c3') (resid 2 and name o3') 1 80 20 2
assign (resid 3 and name c5') (resid 3 and name c4')
       (resid 3 and name c3') (resid 3 and name o3') 1 80 20 2
assign (resid 4 and name c5') (resid 4 and name c4')
       (resid 4 and name c3') (resid 4 and name o3') 1 80 20 2
assign (resid 5 and name c5') (resid 5 and name c4')
       (resid 5 and name c3') (resid 5 and name o3') 1 80 20 2
assign (resid 6 and name c5') (resid 6 and name c4')
       (resid 6 and name c3') (resid 6 and name o3') 1 80 20 2
assign (resid 7 and name c5') (resid 7 and name c4')
       (resid 7 and name c3') (resid 7 and name o3') 1 80 20 2
assign (resid 8 and name c5') (resid 8 and name c4')
       (resid 8 and name c3') (resid 8 and name o3') 1 80 20 2
assign (resid 9 and name c5') (resid 9 and name c4')
       (resid 9 and name c3') (resid 9 and name o3') 1 80 20 2
assign (resid 10 and name c5') (resid 10 and name c4')
       (resid 10 and name c3') (resid 10 and name o3') 1 80 20 2
assign (resid 11 and name c5') (resid 11 and name c4')
       (resid 11 and name c3') (resid 11 and name o3') 1 157 20 2
assign (resid 12 and name c5') (resid 12 and name c4')
       (resid 12 and name c3') (resid 12 and name o3') 1 80 20 2
assign (resid 13 and name c5') (resid 13 and name c4')
```

```
       (resid 13 and name c3') (resid 13 and name o3') 1 80 20 2
assign (resid 14 and name c5') (resid 14 and name c4')
       (resid 14 and name c3') (resid 14 and name o3') 1 80 20 2
assign (resid 15 and name c5') (resid 15 and name c4')
       (resid 15 and name c3') (resid 15 and name o3') 1 80 20 2
assign (resid 16 and name c5') (resid 16 and name c4')
       (resid 16 and name c3') (resid 16 and name o3') 1 80 20 2
assign (resid 17 and name c5') (resid 17 and name c4')
       (resid 17 and name c3') (resid 17 and name o3') 1 80 20 2
assign (resid 18 and name c5') (resid 18 and name c4')
       (resid 18 and name c3') (resid 18 and name o3') 1 80 20 2
assign (resid 19 and name c5') (resid 19 and name c4')
       (resid 19 and name c3') (resid 19 and name o3') 1 80 20 2
assign (resid 20 and name c5') (resid 20 and name c4')
       (resid 20 and name c3') (resid 20 and name o3') 1 80 20 2
assign (resid 21 and name c5') (resid 21 and name c4')
       (resid 21 and name c3') (resid 21 and name o3') 1 80 20 2
assign (resid 22 and name c5') (resid 22 and name c4')
       (resid 22 and name c3') (resid 22 and name o3') 1 80 20 2


!zeta and alpha
assign (resid 1 and name c3') (resid 1 and name o3')
       (resid 2 and name p) (resid 2 and name o5') 1 -71 10 2
assign (resid 1 and name o3') (resid 2 and name p)
       (resid 2 and name o5') (resid 2 and name c5') 1 -68 10 2

assign (resid 2 and name c3') (resid 2 and name o3')
       (resid 3 and name p) (resid 3 and name o5') 1 -71 10 2
```

```
assign (resid 2 and name o3') (resid 3 and name p)
       (resid 3 and name o5') (resid 3 and name c5') 1 -68 10 2

assign (resid 3 and name c3') (resid 3 and name o3')
       (resid 4 and name p) (resid 4 and name o5') 1 -71 10 2
assign (resid 3 and name o3') (resid 4 and name p)
       (resid 4 and name o5') (resid 4 and name c5') 1 -68 10 2

assign (resid 4 and name c3') (resid 4 and name o3')
       (resid 5 and name p) (resid 5 and name o5') 1 -71 10 2
assign (resid 4 and name o3') (resid 5 and name p)
       (resid 5 and name o5') (resid 5 and name c5') 1 -68 10 2

assign (resid 5 and name c3') (resid 5 and name o3')
       (resid 6 and name p) (resid 6 and name o5') 1 -71 10 2
assign (resid 5 and name o3') (resid 6 and name p)
       (resid 6 and name o5') (resid 6 and name c5') 1 -68 10 2

assign (resid 6 and name c3') (resid 6 and name o3')
       (resid 7 and name p) (resid 7 and name o5') 1 -71 10 2
assign (resid 6 and name o3') (resid 7 and name p)
       (resid 7 and name o5') (resid 7 and name c5') 1 -68 10 2

assign (resid 7 and name c3') (resid 7 and name o3')
       (resid 8 and name p) (resid 8 and name o5') 1 0 120 2
assign (resid 7 and name o3') (resid 8 and name p)
       (resid 8 and name o5') (resid 8 and name c5') 1 0 120 2

assign (resid 8 and name c3') (resid 8 and name o3')
```

```
       (resid 9 and name p) (resid 9 and name o5') 1 0 120 2
assign (resid 8 and name o3') (resid 9 and name p)
       (resid 9 and name o5') (resid 9 and name c5') 1 0 120 2

assign (resid 9 and name c3') (resid 9 and name o3')
       (resid 10 and name p) (resid 10 and name o5') 1 0 120 2
assign (resid 9 and name o3') (resid 10 and name p)
       (resid 10 and name o5') (resid 10 and name c5') 1 0 120 2

assign (resid 10 and name c3') (resid 10 and name o3')
       (resid 11 and name p) (resid 11 and name o5') 1 0 120 2
assign (resid 10 and name o3') (resid 11 and name p)
       (resid 11 and name o5') (resid 11 and name c5') 1 0 120 2

assign (resid 11 and name c3') (resid 11 and name o3')
       (resid 12 and name p) (resid 12 and name o5') 1 0 120 2
assign (resid 11 and name o3') (resid 12 and name p)
       (resid 12 and name o5') (resid 12 and name c5') 1 0 120 2

assign (resid 12 and name c3') (resid 12 and name o3')
       (resid 13 and name p) (resid 13 and name o5') 1 0 120 2
assign (resid 12 and name o3') (resid 13 and name p)
       (resid 13 and name o5') (resid 13 and name c5') 1 0 120 2

assign (resid 13 and name c3') (resid 13 and name o3')
       (resid 14 and name p) (resid 14 and name o5') 1 0 120 2
assign (resid 13 and name o3') (resid 14 and name p)
       (resid 14 and name o5') (resid 14 and name c5') 1 0 120 2
```

```
assign (resid 14 and name c3') (resid 14 and name o3')
       (resid 15 and name p) (resid 15 and name o5') 1 0 120 2
assign (resid 14 and name o3') (resid 15 and name p)
       (resid 15 and name o5') (resid 15 and name c5') 1 0 120 2

assign (resid 15 and name c3') (resid 15 and name o3')
       (resid 16 and name p) (resid 16 and name o5') 1 0 120 2
assign (resid 15 and name o3') (resid 16 and name p)
       (resid 16 and name o5') (resid 16 and name c5') 1 0 120 2

assign (resid 16 and name c3') (resid 16 and name o3')
       (resid 17 and name p) (resid 17 and name o5') 1 -71 10 2
assign (resid 16 and name o3') (resid 17 and name p)
       (resid 17 and name o5') (resid 17 and name c5') 1 -68 10 2

assign (resid 17 and name c3') (resid 17 and name o3')
       (resid 18 and name p) (resid 18 and name o5') 1 -71 10 2
assign (resid 17 and name o3') (resid 18 and name p)
       (resid 18 and name o5') (resid 18 and name c5') 1 -68 10 2

assign (resid 18 and name c3') (resid 18 and name o3')
       (resid 19 and name p) (resid 19 and name o5') 1 -71 10 2
assign (resid 18 and name o3') (resid 19 and name p)
       (resid 19 and name o5') (resid 19 and name c5') 1 -68 10 2

assign (resid 19 and name c3') (resid 19 and name o3')
       (resid 20 and name p) (resid 20 and name o5') 1 -71 10 2
assign (resid 19 and name o3') (resid 20 and name p)
       (resid 20 and name o5') (resid 20 and name c5') 1 -68 10 2
```

```
assign (resid 20 and name c3') (resid 20 and name o3')
       (resid 21 and name p) (resid 21 and name o5') 1 -71 10 2
assign (resid 20 and name o3') (resid 21 and name p)
       (resid 21 and name o5') (resid 21 and name c5') 1 -68 10 2

assign (resid 21 and name c3') (resid 21 and name o3')
       (resid 22 and name p) (resid 22 and name o5') 1 -71 10 2
assign (resid 21 and name o3') (resid 22 and name p)
       (resid 22 and name o5') (resid 22 and name c5') 1 -68 10 2

!beta
assign (resid 2 and name p) (resid 2 and name o5')
       (resid 2 and name c5') (resid 2 and name c4') 1 180 20 2
assign (resid 3 and name p) (resid 3 and name o5')
       (resid 3 and name c5') (resid 3 and name c4') 1 180 20 2
assign (resid 4 and name p) (resid 4 and name o5')
       (resid 4 and name c5') (resid 4 and name c4') 1 180 20 2
assign (resid 5 and name p) (resid 5 and name o5')
       (resid 5 and name c5') (resid 5 and name c4') 1 180 20 2
assign (resid 6 and name p) (resid 6 and name o5')
       (resid 6 and name c5') (resid 6 and name c4') 1 180 20 2
assign (resid 7 and name p) (resid 7 and name o5')
       (resid 7 and name c5') (resid 7 and name c4') 1 180 30 2
assign (resid 16 and name p) (resid 16 and name o5')
       (resid 16 and name c5') (resid 16 and name c4') 1 180 30 2
assign (resid 17 and name p) (resid 17 and name o5')
       (resid 17 and name c5') (resid 17 and name c4') 1 180 20 2
assign (resid 18 and name p) (resid 18 and name o5')
```

```
       (resid 18 and name c5') (resid 18 and name c4') 1 180 20 2
assign (resid 19 and name p) (resid 19 and name o5')
       (resid 19 and name c5') (resid 19 and name c4') 1 180 20 2
assign (resid 20 and name p) (resid 20 and name o5')
       (resid 20 and name c5') (resid 20 and name c4') 1 180 20 2
assign (resid 21 and name p) (resid 21 and name o5')
       (resid 21 and name c5') (resid 21 and name c4') 1 180 20 2
assign (resid 22 and name p) (resid 22 and name o5')
       (resid 22 and name c5') (resid 22 and name c4') 1 180 20 2

!gamma
assign (resid 1 and name o5') (resid 1 and name c5')
       (resid 1 and name c4') (resid 1 and name c3') 1 54 10 2
assign (resid 2 and name o5') (resid 2 and name c5')
       (resid 2 and name c4') (resid 2 and name c3') 1 54 10 2
assign (resid 3 and name o5') (resid 3 and name c5')
       (resid 3 and name c4') (resid 3 and name c3') 1 54 10 2
assign (resid 4 and name o5') (resid 4 and name c5')
       (resid 4 and name c4') (resid 4 and name c3') 1 54 10 2
assign (resid 5 and name o5') (resid 5 and name c5')
       (resid 5 and name c4') (resid 5 and name c3') 1 54 10 2
assign (resid 6 and name o5') (resid 6 and name c5')
       (resid 6 and name c4') (resid 6 and name c3') 1 54 10 2
assign (resid 7 and name o5') (resid 7 and name c5')
       (resid 7 and name c4') (resid 7 and name c3') 1 60 30 2
assign (resid 8 and name o5') (resid 8 and name c5')
       (resid 8 and name c4') (resid 8 and name c3') 1 60 30 2
assign (resid 9 and name o5') (resid 9 and name c5')
       (resid 9 and name c4') (resid 9 and name c3') 1 60 30 2
```

```
assign (resid 10 and name o5') (resid 10 and name c5')
       (resid 10 and name c4') (resid 10 and name c3') 1 60 30 2
assign (resid 14 and name o5') (resid 14 and name c5')
       (resid 14 and name c4') (resid 14 and name c3') 1 60 30 2
assign (resid 15 and name o5') (resid 15 and name c5')
       (resid 15 and name c4') (resid 15 and name c3') 1 60 30 2
assign (resid 16 and name o5') (resid 16 and name c5')
       (resid 16 and name c4') (resid 16 and name c3') 1 60 30 2
assign (resid 17 and name o5') (resid 17 and name c5')
       (resid 17 and name c4') (resid 17 and name c3') 1 60 30 2
assign (resid 18 and name o5') (resid 18 and name c5')
       (resid 18 and name c4') (resid 18 and name c3') 1 60 30 2
assign (resid 19 and name o5') (resid 19 and name c5')
       (resid 19 and name c4') (resid 19 and name c3') 1 60 30 2
assign (resid 20 and name o5') (resid 20 and name c5')
       (resid 20 and name c4') (resid 20 and name c3') 1 60 30 2
assign (resid 21 and name o5') (resid 21 and name c5')
       (resid 21 and name c4') (resid 21 and name c3') 1 54 10 2
assign (resid 22 and name o5') (resid 22 and name c5')
       (resid 22 and name c4') (resid 22 and name c3') 1 54 10 2

!epsilon
assign (resid 1 and name c4') (resid 1 and name c3')
       (resid 1 and name o3') (resid 2 and name p) 1 -155 10 2
assign (resid 2 and name c4') (resid 2 and name c3')
       (resid 2 and name o3') (resid 3 and name p) 1 -155 10 2
assign (resid 3 and name c4') (resid 3 and name c3')
       (resid 3 and name o3') (resid 4 and name p) 1 -155 10 2
assign (resid 4 and name c4') (resid 4 and name c3')
```

```
       (resid 4 and name o3') (resid 5 and name p) 1 -155 10 2
assign (resid 5 and name c4') (resid 5 and name c3')
       (resid 5 and name o3') (resid 6 and name p) 1 -155 10 2
assign (resid 6 and name c4') (resid 6 and name c3')
       (resid 6 and name o3') (resid 7 and name p) 1 -155 10 2
assign (resid 17 and name c4') (resid 17 and name c3')
       (resid 17 and name o3') (resid 18 and name p) 1 -155 10 2
assign (resid 18 and name c4') (resid 18 and name c3')
       (resid 18 and name o3') (resid 19 and name p) 1 -155 10 2
assign (resid 19 and name c4') (resid 19 and name c3')
       (resid 19 and name o3') (resid 20 and name p) 1 -155 10 2
assign (resid 20 and name c4') (resid 20 and name c3')
       (resid 20 and name o3') (resid 21 and name p) 1 -155 10 2
assign (resid 21 and name c4') (resid 21 and name c3')
       (resid 21 and name o3') (resid 22 and name p) 1 -155 10 2

! chi
assign (resid 1 and name O4') (resid 1 and name C1')
       (resid 1 and name N9 ) (resid 1 and name C4 ) 1 -160 10 2

assign (resid 2 and name O4') (resid 2 and name C1')
       (resid 2 and name N9 ) (resid 2 and name C4 ) 1 -160 10 2

assign (resid 3 and name O4') (resid 3 and name C1')
       (resid 3 and name N9 ) (resid 3 and name C4 ) 1 -160 10 2

assign (resid 4 and name O4') (resid 4 and name C1')
       (resid 4 and name N1 ) (resid 4 and name C2 ) 1 -160 10 2
```

```
assign (resid 5 and name O4') (resid 5 and name C1')
       (resid 5 and name N1 ) (resid 5 and name C2 ) 1 -160 10 2

assign (resid 6 and name O4') (resid 6 and name C1')
       (resid 6 and name N1 ) (resid 6 and name C2 ) 1 -160 10 2

assign (resid 7 and name O4') (resid 7 and name C1')
       (resid 7 and name N9 ) (resid 7 and name C4 ) 1 -160 10 2

assign (resid 8 and name O4') (resid 8 and name C1')
       (resid 8 and name N9 ) (resid 8 and name C4 ) 1 -160 50 2

assign (resid 9 and name O4') (resid 9 and name C1')
       (resid 9 and name N1 ) (resid 9 and name C2 ) 1 -160 50 2

assign (resid 10 and name O4') (resid 10 and name C1')
       (resid 10 and name N9 ) (resid 10 and name C4 ) 1 -160 50 2

assign (resid 11 and name O4') (resid 11 and name C1')
       (resid 11 and name N1 ) (resid 11 and name C2 ) 1 -160 50 2

assign (resid 12 and name O4') (resid 12 and name C1')
       (resid 12 and name N9 ) (resid 12 and name C4 ) 1 -160 50 2

assign (resid 13 and name O4') (resid 13 and name C1')
       (resid 13 and name N9 ) (resid 13 and name C4 ) 1 -160 50 2

assign (resid 14 and name O4') (resid 14 and name C1')
       (resid 14 and name N1 ) (resid 14 and name C2 ) 1 -160 50 2
```

```
assign (resid 15 and name O4') (resid 15 and name C1')
       (resid 15 and name N9 ) (resid 15 and name C4 ) 1 -160 50 2

assign (resid 16 and name O4') (resid 16 and name C1')
       (resid 16 and name N1 ) (resid 16 and name C2 ) 1 -160 10 2

assign (resid 17 and name O4') (resid 17 and name C1')
       (resid 17 and name N9 ) (resid 17 and name C4 ) 1 -160 10 2

assign (resid 18 and name O4') (resid 18 and name C1')
       (resid 18 and name N9 ) (resid 18 and name C4 ) 1 -160 10 2

assign (resid 19 and name O4') (resid 19 and name C1')
       (resid 19 and name N9 ) (resid 19 and name C4 ) 1 -160 10 2

assign (resid 20 and name O4') (resid 20 and name C1')
       (resid 20 and name N1 ) (resid 20 and name C2 ) 1 -160 10 2

assign (resid 21 and name O4') (resid 21 and name C1')
       (resid 21 and name N1 ) (resid 21 and name C2 ) 1 -160 10 2

assign (resid 22 and name O4') (resid 22 and name C1')
       (resid 22 and name N1 ) (resid 22 and name C2 ) 1 -160 10 2
```

REFERENCES

Abeydeera, D., and Chow, C. (2009). "Synthesis and characterization of modified nucleotides in the 970 hairpin loop of *Escherichia coli* 16S ribosomal RNA." Bioorg Med Chem. **17**: 5887-5893.

Abeydeera, N. D. R. (2009). Synthesis, biophysical studeis, and ligand interactions of helix 31 from Escherichia coli ribosomes. Chemistry Department, Wayne State University. **PhD**.

Aduri, R. (2007). Force field parameters and NMR spectroscopy of modified nucleic acids. Chemistry Department, Wayne State University. **PhD**.

Aduri, R., Psciuk, B. T., Saro, P., Taniga, H., Schlegel, H. B., and SantaLucia, J., Jr. (2007). "AMBER Force Field Parameters for the Naturally Occurring Modified Nucleosides in RNA." J. Chem. Theory and Comput. **3**: 1464 - 1475.

Agalarov, S. C., and Williamson, J. R. (2000). "A hierarchy of RNA subdomains in assembly of the central domain of the 30 S ribosomal subunit." RNA **6**: 402-408.

Agris, P. F. (2004). "Decoding the genome: a modified view." Nucleic Acids Res. **32**: 223-238.

Agris, P. F., Vendeix, F. A., and Graham, W. D. (2007). " tRNA's wobble decoding of the genome: 40 years of modification." J. Biol. Chem. **366**: 1-13.

Allain, F., and Varani, G. (1995). "Structure of the P1 helix from group I self-splicing introns." J. Mol. Biol. **250**: 333-353.

Andersen, T. E., Porse, B. T., and Kirperkar, F. (2004). "A novel partial modification at C2501 in *Escherichia coli* 23S ribosomal RNA." RNA **10**: 907-913.

Andersson, D. I., Andersson, G. E., and Kurland, C. G. (1986). "Functional interactions between mutated forms of ribosomal proteins S4, S5 and S12." Biochimie **68**: 705-173.

Aue, W. P., Bartholdi, E., and Ernst, R. R. (1976). "Two-dimensional spectroscopy. Application to nuclear magnetic resonance." J. Chem. Phys. **64**: 2229-2246.

Bakin, A., and Ofengand, J. (1993). "Four newly located pseudouridylate residues in *Escherichia coli* 23S ribosomal RNA are all at the peptidyltransferase center: Analysis by the application of a new sequencing technique. ." Biochemistry **32**: 9754–9762.

Ban, N., Nissen, P., Hansen, J., Moore, P. B. and Steitz, T. A. (2000). "The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution." Science **289**: 905-920.

Bates, P. A., Kelley, L. A., MacCallum, R. M., and Sternberg, M. J. (2001). "Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM." Proteins **45 (suppl 5)**: 39-46.

Batey, R., Battiste, J. L. and Williamson, J. R. (1995). "Preparationi of isotopically enriched RNAs fro heteronuclear NMR." Methods Enzymol. **261**: 300-322.

Batey, R., Inada, M., Kujawinski, E., Puglisi, J. D. and Williamson, J. R. (1992). "Preparation of isotopically labeled ribonucleotides for multidimensional NMR spectroscopy of RNA." Nucleic Acids Res. **20**: 4515-4523.

Batey, R. T., Inada, M., Kujawinski, E., Puglisi, J. D., and Williamson, J. R. (1992). "Preparation of isotopically labeled ribonucleotides for multidimensional NMR spectroscopy of RNA." Nucleic Acids Res. **20**: 4515-4523.

Batey, R. T., Rambo, R. P. and Doudna, J. A. (1999). "Tertiary motifs in RNA structure and folding." Angew. Chem. Int. Ed. **38**: 2326-2343.

Bax, A., and Davis, D. G. (1985). "MLEV-17 based two-dimensional homonuclear magnetization transfer spectroscopy." J. Magn. Reson. **65**: 355-360.

Beyer, P., Varani, L., and Varani, G. (1999). "Refinement of the structure of protein–RNA complexes by residual dipolar coupling analysis. ." J. Biomol. NMR **14**: 149-155.

Böttger, E. C., Springer, B., Prammananan, T., Kidan, Y. and Sander, P. (2001). "Structural basis for selectivity and toxicity of ribosomal antibiotics." EMBO Rep. **2**: 318-323.

Breaker, R. R. (2008). "Complex riboswitches." Science **319** 1795-1797.

Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G. J. (2007). "Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila." Cell **128**: 1089-1103.

Bruce, G. A., and Uhlenbeck, O. C. (1982). "Enzymatic replacement of the anticodon of yeast phenylalanine transfer ribonucleic acid." Biochemistry **1982**: 855-861.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M,, Simonson, T., and Warren, G. L. (1998). "Crystallography & NMR system: a new software suite for macromolecular structure determination." Acta Cryst. **D54**: 905-921.

Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L. V., Müller, K. M., Pande, N., Shang, Z., and Gutell, R. R. (2002). "The Comparative RNA Web (CRW) Site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs." BMC Bioinformatics **3**: 2.

Carter, A. P., Clemons, W. M. Jr., Brodersen, D. E., Morgan-Warren, R. J., Wimberly, B. T., and Ramakrishnan, V. (2000). "Functional insights from the structure of the 30S ribosomal subunit with and its interactions with antibiotics." Nature **407**: 340-348.

Caskey, C. T., Beaudet, A., and Nirenberg, M. (1968). "RNA codons and protein synthesis. 15. Dissimilar responses of mammalian and bacterial transfer RNA fractions to messenger RNA codons " J. Mol. Biol. **37**: 99–118.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996). "Crystal structure of a group I ribozyme domain: principles of RNA packing." Science **273**: 1678-1685.

Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N., and Noller, H. F. (1999). " X-ray crystal structures of 70S ribosome functional complexes." Science **285**: 2095-2101.

Cater, A. P., Clemons, W. M., Brodersen, D. E., Morgan-Warren, R. J., Hartsch, T., Wimberly, B. T., and Ramakrishnan, V. (2001). "Crystal structure of an initiation factor bound to the 30S ribosomal subunit." Science **291**: 498-501.

Cech, T. (2009). "Crawling out of the RNA world." Cell **136**: 599-602.

Clos, L. J. (2007). NMR solution studied of 16S rRNA A - site functional mutants and binding of novel ligands. Chemistry Department, Wayne State University. **PhD**.

Crain, P. F., and McCloskey, J. A. (1997). "The RNA modification database." Nucleic Acids Res. **25**: 126-127.

Crick, F. (1970). "Central Dogma of Molecular Biology." Nature **227**: 561-563.

Culver, G. M. (2003). "Assembly of the 30S Ribosomal Subunit." Biopolymers **68**: 234-249.

Czerwoniec, A., Dunin-Horkawicz, S., Purta, E., Kaminska, K. H., Kasprzak, J. M., Bujnicki, J. M., Grosjean, H., and Rother, K. (2009). "MODOMICS: a database of RNA modification pathways: 2008 update." Nucleic Acids Res. **37**: D118-D121.

Dallas, A., and Noller, H. F. (2001). "Interaction of translation intitiation factor 3 with the 30S ribosomal subunit." Molecular Cell **8**: 855-864.

Das, R., and Baker, D. (2007). "Automated de novo prediction of native-like RNA tertiary structures." Proc. Natl. Acad. Sci. USA **104**: 14664-14669.

Davis, D. G., and Bax, A. (1985). "Assignment of complex 'H NMR Spectra via two-dimensional homonuclear hartmann-hahn spectroscopy." J. Am. Chem. Soc. **107**: 2820-2821.

de Alba, E., and Tjandra, N. (2000). "Protein backbone ^{15}N relaxation rates as a tool for the diagnosis of structure quality." J. Magn. Reson. **144**: 367-371.

Decatur, W. A., and Fournier, M. J. (2002). "rRNA modifications and ribosome function." Trends Biochem. Sci. **27**: 344–351.

Deltour, R., and Jeener, J. (1965). "One photon-two spins processes." *Proc. Coll. AMPERE* **13**: 271-272.

Desaulniers, J., Chang, Y., Aduri, R., Abeysirigunawardena, S. C., SantaLucia, J., Jr., and Chow, C. S. (2008). "Pseudouridines in rRNA helix 69 play a role in loop stacking interactions." Org. Biomol. Chem. **6**: 3892-3895.

Diener, J. L., and Moore, P. B. (1998). "Solution structure of a substrate for the archaeal pre-tRNA splicing endonucleases: the bulge-helix-bulge motif." Molecular Cell **1**: 883-894.

Ding, Y., Chan, C. Y., and Lawrence, C. E. (2004). "Sfold web server for statistical folding and rational design of nucleic acids." Nucleic Acids Res. **32**: W135-W141.

Dingley, A. J., and Grzesiek, S. (1998). "Direct Observation of Hydrogen Bonds in Nucleic Acid Base Pairs by Internucleotide ^{2}JNN couplings." J. Am. Chem. Soc. **120**: 8293-8297.

Doran, G. (2007). "RNAi-is one suffix sufficient? ." Journal of RNAi and gene silencing **3**: 217-219.

Edmonds, C. G., Crain, P. F., Gupta, R., Hashizume, T., Hocart, C. H., Kowalak, J. A., Pomerantz, S. C., Stetter, K. O., and McCloskey, J. A. (1991). "Posttranscriptional modification of tRNA in thermophilic archaea (Archaebacteria)." J. Bacteriol. **173**: 3138–3148.

Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J. P., and Ehresmann, B. (1987). "Probing the structure of RNAs in solution." Nucleic Acids Res. **15**: 53-72.

Ernst, R. R., and Anderson, W. A. (1966). "Application of Fourier transform spectroscopy to magnetic resonance." Rev. Sci. Inst. **37**: 93-102.

Fernández, C., and Wider, G. (2003). "TROSY in NMR studies of the structure and function of large biological macromolecules." Curr. Opin. Struct. Biol. **13**: 570-580.

Fiaux, J., Bertelsen, E. B., Horwich, A. L., and Wüthrich, K. (2002). "NMR analysis of a 900K groEL-groES complex." Nature **418**: 207-211.

Fire, A., Xu, S. Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). "Potent and specific genetic interference by double-stranded RNA in caenorhabditis elegans." Nature **391** 806-811.

Fox, G. W., and Woese, C. R. (1975). "5S RNA secondary structure." Nature **256**: 505-507.

Frank, J. (2003). "Electron microscopy of functional ribosome complexes." Biopolymers **68**: 223-233.

Frank, J., and Agrawal, R. K. (2000). "A ratchet-likc inter-subunit reorganization of the ribosome during translocation." Nature **406**: 318-322.

Fürtig, B., Richter, C., Wöhnert, J. and Schwalbe, H. (2003). " NMR spectroscopy of RNA." ChemBioChem **4**: 936-962.

Gabashvili, I. S., Agrawal R. K., Spahn, C. M., Grassucci, R. A., Svergun, D. I., Frank, J., and Penczek, P. (2000). "Solution structure of the *E. coli* 70S ribosome at 11.5 A° resolution." Cell **100**: 537-549.

Geis, M., Flamm, C., Wolfinger, M. T., Tanzer, A., Hofacker, I. L., Middendorf, M., Mandl, C., Stadler, P. F., and Thurner, C. (2008). "Folding kinetics of large RNAs " J. Mol. Biol. **379**: 160-173.

Gesteland, R. E., and Atkins, J. F. (1993). The RNA World. Plainview, NY, Cold Spring Harbor Laboratory Press, .

Gilbert, W. (1986). "Origin of life: the RNA world." Nature **319**: 618.

Gorenstein, D. (1984). Principles and applications in phosphorus-31 NMR Orlando, FL, Academic Press.

Grüne, M., Fürste, J. P., Klussmann, S., Erdmann, V. A., and Brown, L. R. (1996). "Detection of multiple conformations of the E-domain of 5S rRNA from Escherichia coli in solution and in crystals by NMR spectroscopy." Nucleic Acids Res. **24**: 2592-2596.

Grzesiek S, A. J., Ren H, and Bax A. (1993). "13C line narrowing by 2H decoupling in 2H/13C/15N-enriched proteins-application to triple-resonance 4D J-connectivity of sequential amides. ." J. Am. Chem. Soc. **115**: 4369-4370.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983). "The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme." Cell **35**: 849-857.

Gutell, R. R., Noller, H. F. and Woese, C. R. (1986). " Higher order structure in ribosomal RNA. ." EMBO J. **5**: 1111-1113.

Gutell, R. R., Weiser, B., Woese, C. R. and Noller, H. F. (1985). "Comparative anatomy of 16-S-like ribosomal RNA." Prog. Nucleic Acids Res. Mol. Biol. **32**: 155-216.

Hansen, J. L., Schmeing, T. M., Moore, P. B. and Steitz, T. A (2002). "Structural insights into peptide bond formation." Proc. Natl. Acad. Sci. USA **99**: 11670-11675.

Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., and Yonath, A. (2001). "A high resolution structure of the large ribosomal subunit from a mesophilic eubacterium." Cell **107**: 679-688.

Hashem, Y., and Auffinger, P. (2009). "A short guide for molecular dynamics simulations of RNA systems." Methods **47**: 187-197.

Hauptman, H. A. (1991). "The phase problem of x-ray crystallography." Rep. Prog. Phys. **54**: 1427-1454.

Held, W. A., Ballou, B., Mizushima, S. and Nomura, M. (1974). "Assembly mapping of 30 S ribosomal proteins from *Escherichia coli*: future studies." J. Biol. Chem. **249**: 3103-3111.

Herr, A. J., Jensen, M. B., Dalmay, T., and Baulcombe, D. C. (2005). "RNA polymerase IV directs silencing of endogenous DNA." Science **308**: 118-120.

Heus, H. A., and Pardi, A. (1991). "Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops." Science **253**: 191-194.

Hines, J. V., Landry, S. M., Varani, G., and Tinoco, I. Jr. (1994). "Carbon-proton scalar couplings in RNA: 3D heteronuclear and 2D isotope-edited NMR of a 13C-labeled extra-stable hairpin." J. Am. Chem. Soc. **116**: 5823-5831.

Hines, J. V., Landry, S. M., Varani, G., and Tinoco, I., Jr. (1994). "Carbon-Proton scalar couplings in RNA: 3D heteronuclear and 2D isotope-ddited NMR of a 13C-labeled extra-stable hairpin." J. Am. Chem. Soc. **116**: 5823-5831.

Hines, J. V., Varani, G., Landry, S. M., and Tinoco, I. Jr. (1993). "The stereospecific assignment of H5' and H5" in RNA using the sign of two-bond carbon-carbon scalar couplings." J. Am. Chem. Soc. **115**: 11002-11003.

Holbrook, S. R., Cheong, C., Tinoco, I. Jr, and Kim. S. H. (1991). "Crystal structure of an RNA double helix in incorporating a track of non-Watson-Crick base pairs." Nature **353**: 579-581.

Jaeger, J. A., SantaLucia, J., Jr., and Tinoco, I., Jr. (1993). "Determination of RNA Structure and Thermodynamics " Ann. Rev. Biochem. **62**: 255-287.

Karimi, R., and Ehrenberg, M. (1994). "Dissociation rate of cognate peptidyl-tRNA from the A-site of hyper-accurate and error prone ribosomes." Eur. J. Biochem. **226**: 355–360.

Karplus, M. (1959). "Contact electron-spin coupling of nuclear magnetic moments." J. Chem. Phys. **30**: 11-15.

Kazantsev, A. V., Krivenko, A. A., Harrington, D. J., Holbrook, S. R., Adams, P. D., and Pace, N. R. (2005). "Crystal structure of a bacterial ribonuclease P RNA." Proc. Natl. Acad. Sci. USA **102**: 13392-13397.

Khaitovich, P., Mankin, A. S., Green, R., Lancaster, L. and Noller, H. F. (1999). "Characterization of functionally active subribosomal particles from Thermus aquaticus." Proc. Natl. Acad. Sci. USA **96**: 85-90.

Kim, S. H. (1976). "Three-dimenstional structure of transfer RNA." Prog. Nucleic Acid Res. Mol. Biol. **17**: 181-216.

Klein, D. J., Moore, P. B., and Steitz, T. A. (2004). "The roles of ribosomal proteins in the structure assembly, and evolution of the large ribosomal subunit." J. Mol. Biol. **340**: 141-177.

Koehl, P., and Delarue, M. (1994). "J. Mol. Biol. Application of a self-consistent mean field theory to predict protein side-chains conformation and estimate their conformational entropy." J. Mol. Biol. **239**: 249-275.

Kowalak, J. A., Pomerantz, S. C., Crain, P. F., and McCloskey, J. A. (1993). "A novel method for the determination of post-transcriptional modification in RNA by mass spectrometry." Nucleic Acids Res. **21**: 4577–4585.

Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982). "Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena." Cell **31**: 147-157.

Lamichhane, T. N. (2009). Nucleotide modificatios in helix 31 (970 loop) of 16S ribosomal RNA of *ESCHERICHIA COLI*: biological roles and identification of inhibitory ligands. Chemistry Department, Wayne State Univeristy. **PhD**.

Lapham, J., Rife, J. P., Moore, P. B., and Crothers, D. M. (1997). "Measurement of diffusion constants for nucleic acids by NMR." J. Biomol. NMR **10**: 255-262.

Lee, A. J., and Crothers, D. M. (1998). "The solution structure of an RNA loop-loop complex: the ColE1 inverted loop sequence." Structure **6**: 993-1005.

Lee, R. C., Feinbaum, R. L. and Ambros, V. (1993). "The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14." Cell **75** 843-854.

LeMaster, D. M., and Richards, F. M. (1988). "NMR sequential assignment of Escherichia coli thioredoxin utilizing random fractional deuteriation." Biochemistry **27**: 142-150.

Leontis, N. B., and Westhof, E. (2001). "Geometric nomenclature and classification of RNA base pairs." RNA **7**: 499-512.

Lescoute, A., and Westhof, E. (2006). "The interaction networks of structured RNAs." Nucleic Acids Res. **34**: 6587-6604.

Leulliot, N., and Varani, G. (2001). "Current topics in RNA-protein recognition: Control of specificity and biological function through induced fit and conformational capture." Biochemistry **40**: 7947–7956.

Levitt, M. (1969). "Detailed molecular model for transfer ribonucleic acid." Nature **224** 759-763.

Levitt, M. (1992). "Accurate modeling of protein conformation by automatic segment matching." J. Mol. Biol. **226**: 507-533.

Liang, X., Liu, Q. and Fournier, M. (2009). "Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing." RNA **15**: 1716-1728

Lida, T., Nakayama, J., and Moazed, D. (2008). "siRNA-mediated heterochromatin establishment requires HP1 and is associatied with antisense transcription." Mol. Cell **31**: 178-189.

Lincoln, T. A., and Joyce, G. F. (2009). "Self-sustained replication of an RNA enzyme." Science **323** 1229-1232.

Lipsett, M. N., Heppell, L. A. and Bradley, D. F. (1960). "Complex formation between adenine oligonucleotides and polyuridylic acid." Biochim. Biophys. Acta **41**: 175-177.

Lynch, S. R., Gonzalez, R. L., Jr., and Puglisi, J. D. (2003). "Comparison of X-ray crystal structure of the 30S subunit-antibiotic complex with NMR structures of decoding site oligonucleotide-paromomycin complex." Structure **11**: 43-53.

Mankin, A. S. (1997). "Pactamycin resistance mutations in functional sites of 16S rRNA." J. Mol. Biol. **274**: 8-15.

Marino, J. P., Gregorian, R. S., Csankovszki, G. Jr., Crothers, D. M. (1995). "Bent helix formation between RNA hairpins with complementary loops." Science **268**: 1448-1454.

McLaughlin, C. S., Dondon, J., Grunberg-Manago, M., Michelson, A. M., and Saunders, G. (1966). "Stability of the messenger RNA-sRNA-ribosome complex." Quant. Biol. **31**: 601-610.

Milligan, J., Groebe, D., Witherell, G., and Uhlenbeck, O. (1987). "Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates." Nucleic Acids Res. **15**: 8783-8798.

Mizushima, S., and Nomura, M. (1970). "Assembly mapping of 30S ribosomal proteins from *E. coli*. ." Nature **226**: 1214-1218.

Moazed, D., and Noller, H. F. (1989). "Intermediate states in the movement of transfer RNA in the ribosome." Nature **342**: 142-148.

Moazed, D., and Noller, H.F. (1987). "Interaction of antibiotics with functional sites in 16S ribosomalRNA." Nature **327**: 389-394.

Moore, D. M., and Reynolds, R. C., Jr. (1997). X-ray diffraction and the identification and analysis of clay minerals New York, Oxford University Press

Moore, P. B., and Steitz, T. A. (2002). "The involvement of RNA in ribosome function." Nature **418**: 229-235.

Morosyuk, S. V., P. R. Cunningham, et al. (2001). "Structure and function of the conserved 690 hairpin in Escherichia coli 16 S ribosomal RNA. II. NMR solution structure." J Mol Biol **307**(1): 197-211.

Morosyuk, S. V., Cunningham, P. R., and SantaLucia, J., Jr. (2001). "Structure and function of the conserved 690 hairpin in Escherichia coli 16 S ribosomal RNA. II. NMR solution structure." J. Mol. Biol. **307**: 197-211.

Morosyuk, S. V., Lee, K., SantaLucia, J., Jr., and Cunningham, P. R. (2000). "Structure and function of the conserved 690 hairpin in Escherichia coli 16 S ribosomal RNA: analysis of the stem nucleotides." J. Mol. Biol. **300**: 113-126.

Morosyuk, S. V., SantaLucia, J., Jr., and Cunningham, P. R. (2001). "Structure and function of the conserved 690 hairpin in Escherichia coli 16 S ribosomal RNA. III. Functional analysis of the 690 loop." J. Mol. Biol. **307**(1): 213-228.

Mueller, F., and Brimacombe, R. (1997). "A new model for the three-dimensional folding of *Escherichia coli* 16S ribosomal RNA." J. Mol. Biol. **271**: 545-565.

Napoli, C., Lemieux, C. and Jorgensen, R. (1990). "Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in tran." Plant Cell **2**: 279-289.

Nelson, D. L., and Cox, M. M. (2004). Lehninger principles of biochemistry. New York, W. H. Freeman.

Neuhaus, D., and Williamson, M. P. (2000). The Nuclear Overhauser Effect in Structural and Conformational Analysis. New York, Wiley-VCH.

Nguyenle, T., Laurberg, M., Brenowitz, M., and Noller, H. F. (2006). "Following the dynamics of changes in solvent accessibiliyt of 16S and 23S rRNA during ribosomal subunit association using synchrotron-generated hydroxyl radicals." J. Mol. Biol. **359**: 1235-1248.

Nikonowicz, E. P., Sirr, A., Legaut, F. M., Jucker, F. M., Baer, L. M., and Pardi, A. (1992). "Preparation of 13C and 15N labeled RNAs for heteronuclear multi-dimensional NMR studies." Nucleic Acids Res. **20**: 4507-4513.

Nissen, P., Hansen, J., Ban, N., Moore, P. B. and Steitz, T. A. (2000). "The structural basis of ribosome activity in peptide bond synthesis." Science **289**: 920-930.

Noah, J. W., Dolan, M. A., Babin, P., and Wollenzien, P. (1999). "Effects of tetracycline and spectinomycin on the tertiary structure of ribosomal RNA in the *Escherichia coli* 30 S ribosomal subunit." J. Biol. Chem. **274**: 16576-16581.

Noller, H. F. (2005). "RNA structure: reading the ribosome." Science **309**: 1508-1514.

Noller, H. F., and Baucom, A. (2002). "Structure of the 70 S ribosome: implications for movement." Biochem. Soc. Trans. **30**: 1159-1161.

Noller, H. F., Kop, J., Wheaton, V., Brosius, J., Gutell, R. R., Kropylov, A. M., Dohme, F., Herr, W., Stahl, D. A., Gupta, R., and Waese, C. R. (1981). "Secondary structure model for 23S ribosomal RNA." Nucleic Acids Res. **9**: 6167-6189.

Noller, H. F., Yusupov, M. M., Yusupova, G. Z., Baucom, A., and Cate, J. H. (2002). "Translocation of tRNA during protein synthesis." FEBS Lett. **514**: 11-16.

Nomura, M. J. (1974). "Assembly of bacterial ribosomes." Supramol Struct. **2**: 163-165.

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

Olson, W. (1982). "Computational studies of polynucleotide flexibility." Nucleic Acids Res. **10**: 777-787.

Olson, W. K. (1982). "How flexible is the furanose ring? II. An updated potential energy estimate." J. Am. Chem. Soc. **104**: 278-284.

Page, R., Grzechnik, S. K., Canaves, J. M., Spraggon, G., Kreusch, A., Kuhn, P., Stevens, R. C., and Lesley, S. A. (2003). "Shotgun crystallization strategy for structural genomics: an optimized two-tiered crystallization screen against the thermotoga maritima proteome." Acta Cryst. **D59**: 1028–1037.

Page, R., Peti W., Wilson, I. A., Stevens R. C., and Wüthrich, K. (2005). "NMR screening and crystal quality of bacterially expressed prokaryotic and eukaryotic proteins in a structural genomics pipeline." Proc. Natl. Acad. Sci. USA **102**: 1901-1905.

Pape, T., Wintermeyer, W., and Rodnina, M.V. (2000). "Conformational switch in the decoding region of 16S rRNA during aminoacyl-tRNA selection on the ribosome. ." Nat. Struct. Biol. **7**: 104–107.

Pervushin, K., Riek, R., Wider, G. and Wüthrich, K. (1997). "Attenuated T2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. ." Proc. Natl. Acad. Sci. USA **94**: 12366-12371.

Petersheim, M., and Turner, D. H. (1983). "Base-stacking and base-pairing contributions to helix stability: thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp and ACCGGUp " Biochemistry **22**: 256-263.

Petrey, D., Xiang, Z., Tang, C. L., Xie L., Gimpelev, M., Mitros, T., Soto, C. S., Goldsmith-Fischman, S., Kernytsky, A., Schlessinger, A., Koh, I. Y., Alexov, E., and Honig, B. (2003). "Using multiple structure alignments, fast model building, and energetic analysis in fold recognition and homology modeling." Proteins **53 (Suppl. 6)**: 430-435.

Piekna-Przybylska, D., Decatur, W. and Fournier, M. (2008). "The 3D rRNA modification maps database: with interactive tools for ribosome analysis." Nucleic Acids Res. **36**: D178-D183.

Piepersberg, W., BÖck, A. and Wittmann, H. G. (1975). "Effect of different mutations in ribosomal protein S5 of Escherichia coli on translational fidelity." Mol. Gen. Genet. **140**: 91-100.

Pioletti, M., Schlünzen, F., Harms, J., Zarivach, R., Glühmann, M., Avila, H., Bashan, A., Bartels, H., Auerbach, T., Jacobi, C., Hartsch, T., Yonath, A., and Franceschi, F (2001). "Crystal structures of complexes of the small

ribosomal subunit with tetracycline, edeine and IF3 " EMBO J. **20**: 1829-1839.

Pley, H. W., Flaherty, K. M. and Mckay, D. B. (1974). "Three-dimensional structure of a hammerhead ribozyme." Nature **372**: 68-74.

Powers, T. and H. F. Noller (1995). "Hydroxyl radical footprinting of ribosomal proteins on 16S rRNA." RNA **1**: 194-207.

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

Pushparaj, P. N., Aarthi, J. J., Kumar, S. D. and Manikandan, J. (2008). "RNAi and RNAa-the Yin and Yang of RNAome." Bioinformation **2**: 235-237.

Riek, R., Wider, G., Pervushin, K., and Wüthrich, K. (1999). "Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules." Proc. Natl. Acad. Sci. USA **96**: 4918-4923.

Robertson, M. P., and Scott, W. G. (2007). "The structural basis of ribozymecatalyzed RNA assembly." Science **315**: 1549-1553.

Rozenski, J., Crain, P. F., and McCloskey, J. A. (2000). "The RNA Modification Database: 2000 update." Nucleic Acids Res. **27**: 196-197.

Rüdisser, S., and Tinoco, I. Jr. (2000). "Solution structure of cobalt (III) hexamine complexed to the GAAA tetraloop, and metal-ion binding to G•A mismatches." J. Mol. Biol. **295**: 1211-1223.

Saenger, W. (1984). Principles of Nucleic Acid Structure New York, Springer-Verlag.

Šali, A., and Blundell, T. L. (1993). "Comparative protein modeling by satisfaction of spatial restraints." J. Mol. Biol. **234**: 779–815.

SantaLucia, J., Jr. (2000). The use of spectroscopic techniques in the study of DNA stability. Chapter 13 in Spectrophotometry and spectrofluorimetry Oxford, Oxford University Press.

SantaLucia, J., Jr., and Turner, D. H. (1993). "Structure of (rGGCGAGCC)2 in solution from NMR and restrained molecular dynamics " Biochemistry **32**: 12612-12623.

SantaLucia, J., Jr., Shen, L., Cai, Z., Lewis, H., and Tinoco, I., Jr. (1995). "Synthesis and NMR of RNA with isotopic enrichment in the bases." Nucleic Acids Res. **23**: 4913-4921.

Santoro, J., and King, G. C. (1992). "A constant-time 2D overbodenhausen experiment for inverse correlation of isotopically enriched species." J. Magn. Reson. **97**: 202-207.

Saraiya, A., Lamichhane, T., Chow, C., SantaLucia, J., and Cunningham, P. R. (2008). "Identification and role of functionally important motifs in the 970 loop of Escherichia coli 16S ribosomal RNA." J. Mol. Biol. **376**: 645-657.

Saunders, M., Wishnia, A., and Kirkwodd, J. G. (1957). "The nuclear magnetic resonance spectrum of ribonuclease." J. Am. Chem. Soc. **79**: 3289-3290.

Schlünzen, F., Zarivach, R., Harms, J., Bashan, A., Tocilj, A., Albrecht, R., Yonath, A., and Franceschi, F. (2001). "Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria " Nature **413**: 814-821.

Schmitz, M., and Tinoco, I. Jr. (2000). "Solution structure and metal-ion binding of the P4 element from bacterial RNase P RNA." RNA **6**: 1212-1225.

Schwede, T., Kopp, J., Guex, N., and Peitsch, M.C. (2004). "SWISS-MODEL: An automated protein homology-modeling server." Nucleic Acids Res **31**: 3381-3385.

Scott, W. G., Finch, J. T. and Klug, A. (1995). "The crystal structure of an all-RNAhammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage." Cell **81**: 991-1002.

Selmer, M., Dunham, C. M., Murphy, F. V., Weixlbaumer, A., Petry, S., Kelley, A. C., Weir, J. R., and Ramakrishnan, V. (2006). "Structure of the 70S ribosome complexed with mRNA and tRNA." Science **313**: 1935-1942.

Selmer, M., Dunham, C., Murphy, F., Weixlbaumer, A., Petry, S., Kelley, A. C., Weir, J. R., and Ramakrishnan, V. (2006). "Structure of the 70S ribosome complexed with mRNA and tRNA." Science **313**: 1935-1942.

Snyder, D. A., Chen, Y., Denissova, N. G., Acton, T., Aramini, J. M., Ciano, M., Karlin, R., Liu, J., Manor, P., Rajan, P. A., Rossi, P., Swapna, G. V. T., Xiao, R., Rost, B., Hunt, J., and Montelione, G. T. (2005). "Comparisons of NMR spectral quality and success in crystallization demonstrate that NMR and X-ray crystallography are complementary methods for small protein structure determination." J. Am. Chem. Soc. **127**: 16505-16511.

Sontheimer, E. J., and Carthew, R. W. (2005). "Silence from within: endogenous siRNAs and miRNAs." Cell **122**: 9-12.

Spahn, C. M. T., Beckmann, R., Eswar, N., Penczek, P., Sali, A., Blobel, G., and Frank J. (2001). "Structure of the 80S ribosome from saccharomyces cerevisiae-tRNA-ribosome and subunit-subunit interactions." Cell **107**: 373-386.

Stallings, S. C., and Moore, P. B. (1997). "The structure of an essential splicing element: stem loop IIa from yeast U2 snRNA." Structure **5**: 1173-1185.

Steitz, T. (2008). "A structural understanding of the dynamic ribosome machine." Molecular Cell Biology **9**: 242-253.

Stern, S., Powers, T., Changchien, L. M., and Noller, H. F. (1989). " RNA-protein interactions in 30S ribosomal subunits: folding and functions of 16S rRNA." Science **244**: 783-790.

Stöffler, G., Deusser, E., Wittmann, H. G. and Apirion, D. (1971). "Ribosomal proteins. XIX. Altered S5 ribosomal protein in an Escherichia coli revertant from strptomycin dependence to independence." Mol. Gen. Genet. **111**: 334-341.

Strobel, S. A., and Cochrane, J. C. (2007). "RNA catalysis: ribozymes, ribosomes, and riboswitches." Curr. Opin. Struct. Biol. **11**: 636-643.

Szathmáry, E. (1999). "The origin of the genetic code: amino acids as cofactors in an RNA world." Trends Genet **15**: 223-229.

Szewczak, A. A., Moore, P. B. , Chang, Y. L. , and Wool, I. G. (1993). "The conformation of the sarcin/ricin loop from 28S ribosomal RNA." Proc. Natl. Acad. Sci. USA **90**: 9581-9585.

Tinoco, I., and Bustamante, C. (1999). "How RNA folds." J. Mol. Biol. **293**: 271-281.

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

Torchia, D. A., Sparks, S. W, and Bax, A. (1988). "NMR signal assignments of amide protons in the alpha-helical domains of staphylococcal nuclease." Biochemistry **27**: 5135-5141.

Torres-Larios, A., Swinger, K. K., Krasilnikov, A. S., Pan, T., and Mondragón, A. (2005). "Crystal structure of the RNA component of bacterial ribonuclease P." Nature **437**: 584-587.

Traub, P., and Nomura, M. (1969). "Studies on the assembly of ribosomes in vitro." Cold Spring Harbor Symp Quant Biol **34**: 63-67.

Traub, P., and Nomura, M. J. (1969). "Structure and function of Escherichia coli ribosomes: VI. Mechanism of assembly of 30 S ribosomes studied in vitro." J. Mol. Biol. **40**: 391-413.

Turner, D. H., Sugimoto, N., and Freir, S. M. (1988). "RNA structure prediction." Annul. Rev. Biophys. Biophys. Chem **17**: 167-192.

Valle, M., Zavialov, A., Sengupta, J., Rawat, U., Ehrenberg, M., and Frank, J. (2003). "Locking and unlocking of ribosomal motions." Cell **114**: 123-134.

Van de Peer, Y., Robbrecht, E., de Hoog, S., Caers, A., De Rijk, P., and De Wachter, R. (1999). "Database on the structure of small subunit ribosomal RNA." Nucleic Acids Res **27**(1): 179-183.

Varani, G., Aboul-ela, F., and Allain, H. T. (1996). "NMR investigation of RNA structure." Prog. Nucl. Magn. Reson. Spectrosc. **29**: 51-127.

Varani, G., and Tinoco, I., Jr. (1991). "Carbon assignments and heteronuclear coupling constants for an RNA oligonucleotide from natural abundance ^{13}C-^{1}H Correlated Experiments." J. Am. Chem. Soc. **113**: 9349-9354.

Varani, G., Cheong, C., and Tinoco, I. Jr. (1991). "Structure of an unusually stable RNA hairpin." Biochemistry **30**: 3280-3289.

Wahl, M. C., Rao, S. T. and Sundaralingam, M. (1996). "The structure of r(UUCGCG) has a 5'-UU-overhang exhibiting Hoogsteen-like trans U•U base pairs." Nature Struct. Biol. **3**: 24-31.

Wallner, B., and Elofsson, A. (2005). "All are not equal: a benchmark of different homology modeling programs." Protein Science **14**: 1315-1327.

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P., and Burge, C. B. (2008). "Alternative isoform regulation in human tissue transcriptomes." Nature **456**: 470-476.

Wang, J., Zuo, X., Yu, P., Byeon, I. L., Jung, J., Wang, X., Dyba, M., Seifert, S., Schwieters, C. D., Qin, J., Gronenbrn, A. M., and Wang, Y. (2009). "Determination of multicomponent protein structures in solution using global orientation and shape restraints." J. Am. Chem. Soc. **131**: 10507-10515.

Watanabe, K., Shinma, M., and Oshima, T. (1976). "Heat-induced stability of tRNA from an extreme thermophile, Thermus thermophilus." Biochem. Res. Commun. **72**: 1137-1144.

Watts, J. M., Dang, K. K., Gorelick, R. J., Leonard, C. W., Bess, J. W., Swanstrom, R., Burch, C. L., and Weeks, K. M. (2009). "Architecture and secondary structure of an entire HIV-1 RNA genome." Nature **460**: 711-716.

Westhof, E., and Auffinger, P. (2000). Encyclopedia of Analytical Chemistry, John Wiley & Sons Ltd, .

Wijmenga, S. S., Heus, H. A., Werten, B., van der Marel, G. A., van Boom, J. H., Hilbers, C. W. (1994). "Assignment strategies and analysis of cross-peak patterns and intensities in the three-dimensional homonuclear TOCSY-NOESY of RNA." J. Magn. Reson. B. **103**: 131-141.

Williamson, J. R. (2000). "Induced fit in RNA-protein recognition." Nat. Struct. Biol. **7**: 834–837.

Willionaon, M. P., Hovel, T. F. and Wüthrich, K. (1985). "Solution conformation of proteinase inhibitor IIA from bull seminal plasma by 1H nuclear magnetic resonance and distance geometry." J. Mol. Biol. **182**: 295-315.

Wimberly, B. T., D. E. Brodersen, et al. (2000). "Structure of the 30S ribosomal subunit." **407**(21): 327-339.

Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000). "Structure of the 30S ribosomal subunit." Nature **407**: 327-339.

Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000). "Structure of the 30S ribosomal subunit." Nature **407**(21): 327-339.

Woese, C. R., and Pace, N. R. (1993). The RNA World. New York, Cold Spring Harbor Laboratory Press.

Woese, C. R., Magrum, L. J., Gupta, R., Siegel, R. B., Stahl, D. A., Kop, J., Crawford, N., Brosius, J., Gutell, R., Hogan, J. J., and Noller, H. F. (1980). "Secondary structure model for bacterial 16S ribosomal RNA: phylogenetic, enzymatic and chemical evidence." Nucleic Acids Res. **8** 2275-2283.

Wöhnert, J., Dingley, A. J., Stoldt, M., Görlach, M., Grzesiek, S., and Brown, L. R. (1999). "Direct identification of NH···N hydrogen bonds in noncanonical base pairs of RNA by NMR spectroscopy." Nucleic Acids Res. **27**: 3104-3110.

Woodcock, J., Moazed, D., Cannon, M., Davies, J., and Noller, H. F (1991). "Interaction of antibiotics with A- and P-site-specific bases in 16S ribosomal RNA." EMBO J. **10**: 3099-3103.

Woodcock, J., Moazed, D., Cannon, M., Davies, J., and Noller, H.F (1991). "Interaction of antibiotics with A- and P-site-specific bases in 16S ribosomal RNA." EMBO J. **10**: 3099–3103.

Wüthrich, K. (1976). NMR in Biological Research: Peptides and Proteins. Oxford, Elsevier.

Wüthrich, K. (1986). NMR of proteins and nucleic acids. New York, Wiley.

Wüthrich, K. (1990). "Protein structure determination in solution by NMR spectroscopy." J. Biol. Chem. **265**: 22059-22062.

Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D. R., and Kay, L. E. (1994). "A suite of triple-resonance NMR experiments for the backbone assignment of 15N, 13C, 2H- labeled proteins with high-sensitivity." J. Am. Chem. Soc. **116**: 11655-11666.

Ye, K. (2007). "H/ACA guide RNAs, proteins and complexes." Curr. Opin. Struct. Biol. **17**: 287–292.

Yee, A. A., Savchenko, A., Ignachenko, A., Lukin, J., Xu, X., Skarina, T., Evdokimova, E., Liu, C., Semesi, A., Guido, V., Edwards, A. M., and Arrowsmith, C. H. (2005). "NMR and X-ray crystallography, complementary tools in structural proteomics of small proteins." J. Am. Chem. Soc. **127**: 16512-16517.

Yu, L. P., Petros, A. M., Schnuchel, A., Zhong, P., Severin, J. M., Walter, K., Holzman, T. F., and Fesik, S. W. (1997). "Solution structure of an rRNA methyltransferase (ErmAM) that confers macrolide-lincosamidestreptogramin antibiotic resistance." Nat. Struct. Biol. **4**: 483-489.

Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H. D., and Noller, H. F. (2001). "Crystal structure of the ribosome at 5.5 Å resolution." Science **292**: 883-896.

Yusupova, G., L. Jenner, et al. (2006). "Structural basis for messenger RNA movement on the ribosome." Nature **444**: 391-394.

Zaher, H. S., and Green, R. (2009). "Fidelity at the molecular level: lessons from protein synthesis." Cell **316** 746-762.

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

Zuo, X., Wang, J., Foster, T. R., Schwiesters, C. D., Tiede, D. M., Butcher, S. E., and Wang, Y. (2008). "Global molecular structure and interfaces: refining an RNA: RNA complex structure using solution X-ray scattering data." J. Am. Chem. Soc. **130**: 3292-3293.

Zuo, X., Wang, J., Yu, P., Eyler, D., Xu, H., Starich, M. R., Tiede, D. M., Simon, A. E., Kasprzak, W., Schwieters, C. D., Shapiro, B. A., and Wang, Y. (2010). "Solution structure of the cap-independent translational enhancer and ribosome-binding element in the 3'-UTR of turnip crinkle virus." Proc. Natl. Acad. Sci. USA **107**: 1385-1390.

ABSTRACT

NMR SOLUTION STRUCTURES OF TWO HAIRPINS OF *E. coli* 16S rRNA: THE EFFECTS OF MUTATIONS AND CHEMICAL MODIFICATIONS ON STRUCTURE AND FUNCTION OF rRNA

by

YU LIU

May 2011

Advisor: Prof. John SantaLucia Jr.

Major: Chemistry (Analytical Chemistry)

Degree: Doctor of Philosophy

The structures of two functional mutants, the UC (G690U, U697C) and the QM mutants (G690A, G693C, A695C, U697A) of the 690 hairpin of *E. coli* 16S ribosomal RNA were determined by NMR. The UC mutant and the QM mutant with high biological function are able to fold into structures that are isomorphous with the wild-type 690 hairpin sequence. The structural comparisons among the functional mutants and the wild-type provides structural validation for previously identified specific functional groups that are crucial for maintaining function of the 690 hairpin. The key groups for maintaining the structure and function of the 690 loop include the formation of a “U-turn” at position 692, mismatch pairs involving residues 691-696 and 690-697, the placement of a hydrogen bond acceptor atom in the minor groove at 697, and continuous stacking of five bases at the 3’-terminus of the hairpin.

The 970 hairpin of *E. coli* 16S rRNA contains two modified nucleotides, m^2G966 and m^5C967. The two modified nucleotides in the 970 loop are involved

in some important biological functions, for example, the m^2G966 of the 970 hairpin interact with C34 of P-site bounded tRNA in the crystal structure. The structural role of the modifications, however, is not obvious in the published crystal structures (2AVY and 1J5E), and the 970 loop structures in the two crystal structures have some differences, possibly because that the 970 loop hairpin can form multiple functionally important conformations. We determined NMR structures of the 970 hairpin with and without modifications to identify the functional roles of the modifications on the structures of the 970 loop. And the experimental results indicate that the chemical modifications in the 970 loop cause the structural differences between the NMR structures of the 970 loop with and without modifications, for example, the flipping G971. And the modified nucleotides in the 970 loop may contribute to the structural stability of the 970 loop because methyl groups in the two modified residues 966 and 967 increase the stacking surface area of the triple-base stacking formed among 966, 967 and 967. In addition, the NMR structure of the modified-970 RNA support the mutation study results obtained in Dr. Cunningham's lab.

AUTOBIOGRAPHICAL STATEMENT

YU LIU

I was born in Jishui, a beautiful small town in south of China. After living in my hometown for sixteen years, I went to Beijing, the capital of China, for pursuing my study in Beijing Normal University, where I got my Bachelor and Master degree of Science. I worked on the "Detected H_2O_2 concentrations by using chemiluminescence spectroscopy" and "Determination trace amount of tosufloxacin (TFLX) by using electro-chemiluminescence (ECL) method" as my thesis for Bachelor and Master Degree. After finishing the undergraduate and graduate study in Beijing Normal University, I came to the United States in 2005, and joined Dr. John SantaLucia's lab to learn RNA structural biology.

Education

Doctor of Philosophy (Aug 2005-Nov 2010): Department of Chemistry, Wayne State University, Michigan. Dissertation Title: "NMR solution structures of two hairpins of E. coli 16S rRNA: the effects of mutations and chemical modifications on structure and function of rRNA*".* Advisor: Prof. John SantaLucia Jr.

Master of Science (Sep 2002-Jul 2005): Beijing Normal University, Beijing, China. Specialization: Analytical Chemistry

Bachelor of Science (Sep 1998-Jul 2002): Beijing Normal University, Beijing, China. Specialization: Chemistry

Publications

1. **Liu, Y.**, Aduri, R., Cunningham, P. and SantaLucia, J., Jr. "NMR solution structures of functional mutants of the 690 hairpin of *E. coli* 16S ribosomal RNA". (To be submitted).
2. **Liu, Y.**, Chow, C., Cunningham, P., and SantaLucia, J., Jr. "Structural and functional effects of modifications to the 970 hairpin of *E.coli* 16S rRNA by using NMR method". (In preparation).
3. Chen, S., **Liu, Y.**, Zhao, H., Jin, L., Zhang, Z., and Zheng, Y. (2006), "Determination of norfloxacin using a terbium-sensitized electrogenerated chemiluminescence method". *Luminescence* **21**: 20-25.
4. Chen, S., Ding, F., **Liu, Y.**, and Zho, H. (2006) "Electrochemiluminescence of terbium (III)-two fluoroquinolones-sodium sulfite system in aqueous solution." *Spectrochimica Acta Part A* **64**: 130-135.
5. **Liu, Y.**, Chen, S., Zhao, H., and Jin, L. (2005) "Terbium sensitized determination of tosufloxacin using electrochemiluminescence method." *Analytical Sciences* **21**: 373-375.
6. Feng, Q., Fen, D., and **Liu, Y**. (2005) "The determination of fleroxacin by photochemical fluorescence of zinc-fleroxacin complex." *Spectroscopy and Spectral Analysis (Chinese)* **25**: 1468-1470.
7. **Liu,Y.,** Zhao, H., Yi, L., and Sun, C., and Chen, S. (2003) "Determination of H_2O_2 by $KMnO_4$-Luminol System". *Journal of Instrumental Analysis (Chinese)* **22**: 91-93.