1-3-2012

"fine-tuning" of ribosomal structure and functions by pseudouridylation and rna-protein interactions

Jun Jiang
Wayne State University,

"FINE-TUNING" OF RIBOSOME STRUCTURE AND FUNCTIONS BY PSEUDOURIDYLATION AND RNA-PROTEIN INTERACTIONS

by

JUN JIANG

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

2012

MAJOR: CHEMISTRY (BIOCHEMISTRY)

Approved by:

_____
Advisor   Date

_____

_____

_____

DEDICATION

To my Beloved Family.

ACKNOWLEDGEMENTS

been as actively involved in my training process as they would have for their own students.

I am also grateful for the experience of working with the members of the SantaLucia, Chow, and Cunningham labs. Dr. Fred Sijenyi of the SantaLucia lab has been my "shadow" trainer of NMR experiments and provides me with a lot of great advice even after his graduation. Jenna and Vidisha of the Cunningham lab coached me in molecular biotechnology. Yogo and Daya of the Chow lab fulfilled their roles as project collaborators excellently with instructive discussions and convenience in bench work. I could not list the details of how much I benefit from each of these brilliant minds, while I want to have their names in my acknowledgement: Larry, Ravi, Marcus, Yu, Li, Laura, Andrew, Julia, Pirro, Chandani of the SantaLucia lab; Anne-Cécile, Dinuka, Peiwen, Yucheng, Minako, Sanjaya, Moninderpal, Keshab, Santosh, Tek, Papa Nii, Yuqin, Xun, Gayani, Danielle, Hyosuk, Beth of the Chow lab; Wesley, Ami, Sathya, Pankti, Mark of the Cunningham lab.

I would like thank my committee members, who have been helping me at every step I take towards graduation. During the one and a half years in the life science building, Dr. Pflum and the Pflum lab members were my happy nice neighbors, providing both companionship and access to lab instruments. Dr. Dutta helped me with planning and registration of the two pharmaceutical courses in my first year of study.

Many other people in the Department of Chemistry and Department of Biological Sciences offered their generous help to me too. I would like to thank

iv

best in both academic pursuits and music performance. To my parents-in-law Chunmei Chen and Dashui Liu, I would like to express my gratitude for their support and help which make it possible for both my wife and I to focus on our careers.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

xiv

CHAPTER 1

INTRODUCTION TO THE RIBONUCLEIC ACIDS AND RIBOSOMES

1.1 Introduction

In 1956, Francis Crick, whose name is imprinted into history due to his role in the discovery of Watson Crick base pairing, first described the concept of central dogma (Figure 1.1). It says "Once information has got into a protein it can't get out again"(Crick 1956). This doctrine of biology has been widely accepted and proven through the years. In his original draft, RNA is positioned between DNA, the most commonly used genetic information storage medium, and protein, the most widely used cellular structural and functional component. This picture provides us with a direct illustration of the crucial role that RNA is playing in the flow of genetic information from the genome to protein expression.

RNA is an abbreviation for RiboNulceic Acid. RNA is a group of molecules with extremely high diversity in structures and functions. There are three major groups of RNA directly involved in translation, messenger RNA (mRNA), ribosomal RNA (rRNA), and transfer RNA (tRNA). mRNA is the product synthesized according to the protein-encoding DNA sequence by a process called transcription. The genetic information carried in mRNA is subsequently translated into protein as amino acid sequence. rRNA is the major component of ribosomes, the universally conserved protein synthesis machinery in all living cells. It has been shown that, protease treated 50S subunit retains the capability to catalyze peptide bond formation reaction (Noller, Hoffarth *et al*. 1992). This observation demonstrates an essential catalytic role of rRNAs in the translation mechanism. Each tRNA carries a specific amino acid, the building blocks of
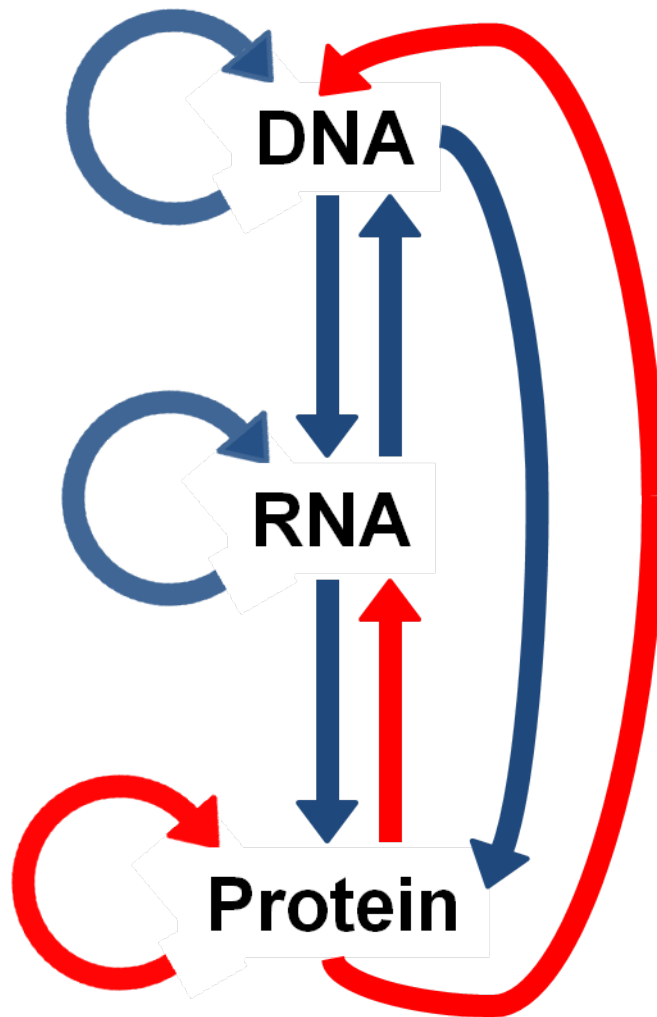
Figure 1.1 The Central Dogma by Francis Crick. Dark blue arrows show the possible directions of genetic information flow, and the red arrows show the impossible directions of genetic information flow.

proteins, to the ribosome in the translation process. The amino acid is covalently attached to the 3'-terminus of a tRNA by an ester bond formed between the carboxylic acid group of the amino acid and the 2'- or 3'- hydroxyl group of the tRNA 3'-terminal nucleotide residue. This reaction is catalyzed by a family of aminoacyl tRNA synthetases, with high recognition specificity for the cognate amino acid and the corresponding tRNA. On the other end of the folded tRNA structure, there is a group of three continuous nucleotide residues forming the anticodon, which specifically recognizes the codon on the mRNA. Both the codon-anticodon recognition and aminoacyl tRNA synthesis determine the accuracy of translation (Ibba and Soll 1999).

Many other RNA species are indirectly involved in the translation process, and are essential for the correct tailoring of the mRNA sequences. As part of the pre-mRNA in eukaryotes, some introns have self splicing activities (Altman 1989; Cech 1990). It was also reported that group II introns, capable of carrying out self-splicing with the help of an intron encoded protein, was identified in bacteria (Ferat and Michel 1993). This discovery changed the opinion that pre-mRNA splicing only exists in eukaryotic cells. In eukaryotic cells, small nuclear RNA (snRNA) containing spliceosomes are more widely employed in pre-mRNA splicing. For example, U2/U6 snRNAs of the spliceosomes were found to show partial splicing activity *in vitro*, while all other snRNAs and spliceosomal proteins ensure the normal complete splicing of the pre-mRNA (Valadkhan and Manley 2001).

Other functions of RNA include regulation on gene expression, maintenance of telomeres, and genetic information storage. MicroRNA (miRNA) complexed with Augonaute protein can repress translation initiation and/or destabilize mRNA (Bartel

2009). An essential RNA component of the telomerase functions as the template to dictate the sequence of DNA repeats added in the telomeres (Greider and Blackburn 1989). In some viruses, single-stranded or double-stranded RNA is used as the carrier of genetic information (Dodds, Morris *et al*. 1984; Melzer, Sether *et al*. 2008).

In 2011, the Holliger group reported that an engineered RNA polymerase ribozyme successfully replicated RNA templates of a variety of sequences (Wochner, Attwater *et al*. 2011). This observation provides additional evidence of the versatility of RNA molecules. Taken together, it is not difficult to envision a long lost world of "living" RNAs, as depicted by Walter Gilbter in his article *The RNA World,* where RNA molecules were replicating, splicing, folding, and passing on genetic information to the next generation of RNA molecules (Gilbert 1986).

1.2 Ribonucleotides

1.2.1 Structures and modifications

The diverse functionalities of RNA molecules originate from the different chemical structures of their building blocks, the ribonucleotides. In the codon table of mRNA, different arrangements of A, C, G, and U into a three letter sequence give rise to coding of a specific amino acid. Each of the single letter listed above stands for a chemical moiety called ribonucleotide residue, composed of three moieties: a purine or pyrimidine base, a ribose sugar, and a phosphate. The ribose and phosphate groups are connected by phosphodiester bonds to form the backbone of RNA molecules, in which the phosphate group is connected to the O5' of the ribose in the same residue through a phosphester bond, and to the O3' of the ribose of the 5'- neighbouring residue through another phosphester bond. The connectivity of phosphodiester bond and the

chirality of the sugars give rise to the directionality of a RNA molecule. If the 3'-O of residue X ribose is connected to the phosphate of residue Y, then residue X is said to reside at the 5' end of the chain and residue Y .The purine or pyrimidine base specifies the identity of the ribonulceotide. A pyrimidine base, i.e. Cytosine and Uracil (thymidine in DNA), is an aromatic ring structure similar to benzene, with two nitrogen atoms replacing two carbon atoms at positions 1 and 3. And a purine base, i.e. Adenine and Guanine, is a heterocyclic structure formed by fusion of a pyrimidine ring and an imidazole ring on the carbon-carbon double-bond edges.   A purine base is covalently attached to the ribose moiety through a C-N glycosidic bond formed between the N9 of the base and C1' of the ribose. In most cases (pseudouridine is an exception), the pyrimidine base is attached to the ribose moiety through a C-N glycosidic bond formed between the N1 of the base and C1' of the ribose (Figure 1.2).

Besides the four most commonly occurring nucleotides discussed above, RNA molecules undergo intensive post-transcriptional modifications, and show much more variety in the structures of nucleotides than DNA (Figure 1.3). Pseudouridine is introduced into RNA post-transcriptionally by a family of enzymes called psudouridine synthases (Ofengand 2002). Unlike the N1-C1' glycosidic bond found in uridine, pseudouridine is attached to the ribose through a C5-C1' glycosidic bond (Figure 1.3). This isomerization of uridine to pseudouridine occurs widely spread in tRNA and rRNA (Singer and Smith 1972; Chow, Larnichhane *et al*. 2007). In other forms of modifications, the connection between the base and ribose sugar is kept, while the base and/or ribose structure are modified by removing or adding functional groups
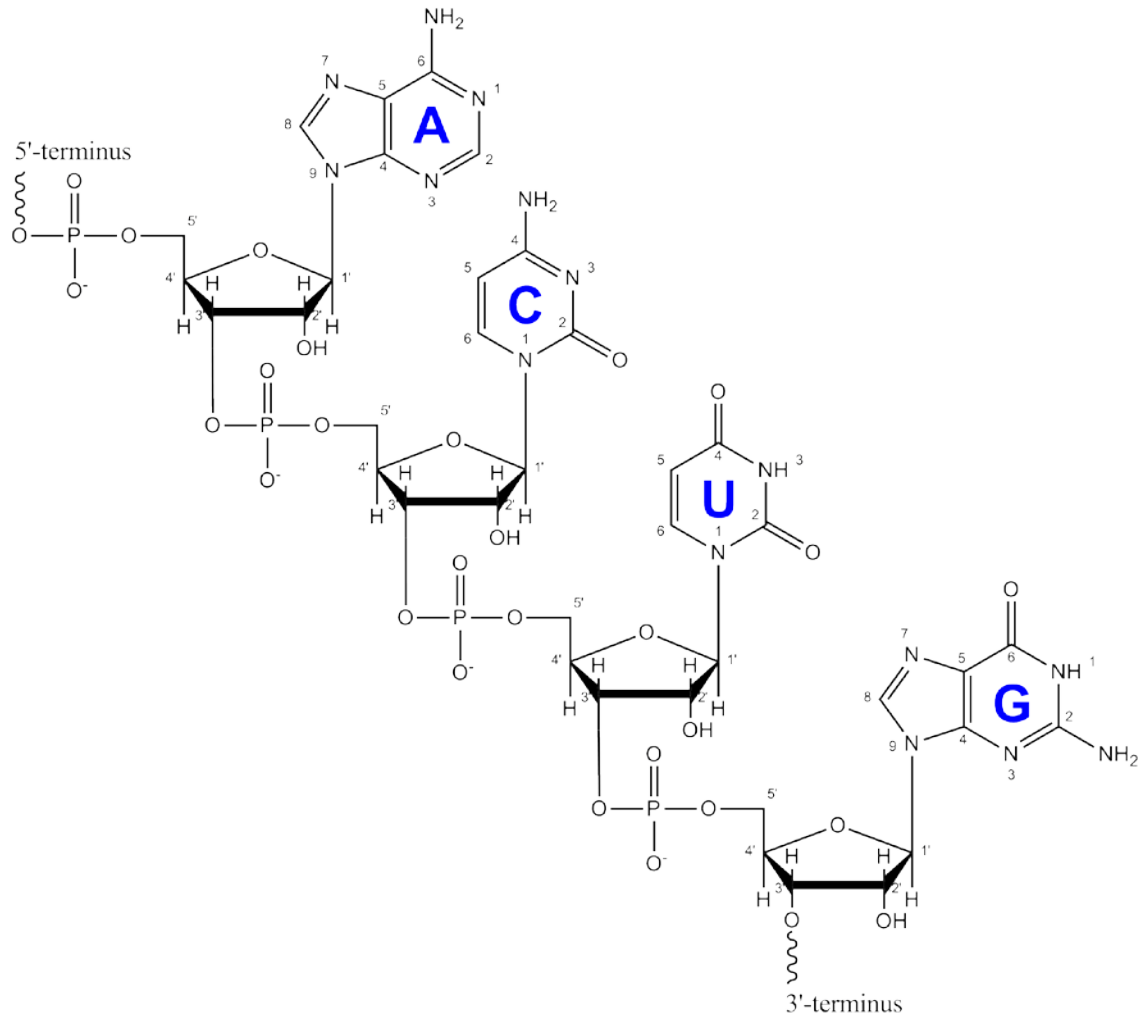
Figure 1.2 Chemical structures of the four standard ribonucleic acid residues connected with phosphodiester linkages.

catalyzed by dedicated enzymes. Inosine is the product from deamination of adenosine catalyzed by adenosine deaminase, and it has been found to exist in both mRNA and tRNA (Holley, Everett *et al*. 1965; Paul and Bass 1998). Another modification changing the base ring scaffold is dihydrouridine, in which the double-bond of uracil between C5 and C6 is reduced by the addition of two hydrogen atoms. This modification occurs with the second highest abundance, right after psudouridinylation, found in tRNA and rRNA (Kowalak, Bruenger *et al*. 1995; Sprinzl, Horn *et al*. 1998). Nucleotide structures are also modified by a variety of methylations. In the mRNA cap structure, the N7 position of the Guanosine is methylated in the mRNA, while in some snRNAs, two additional methyl groups are introduced onto the N2 of the cap guanosine to replace the two hydrogen atoms (Busch, Reddy *et al*. 1982). Other positions in the nucleotide are also eligible for methylation, including the nitrogen atoms of all the base amino and imino groups or with a neighboring carbon covalently attached to an amino group, oxygen atoms of all the sugar 2'-hydroxyl groups, and specific carbon atoms, e.g. C2 and C8 of adenine and C5 of the pyrimidines (Motorin and Helm 2011). As one of the hypermodification examples identified in rRNA, 1-methyl-3-y-(α-amino-α-carboxypropyl)pseudouridine has three different modifications in its structure, i.e. pseudouridylation, methylation at N1 position, and addition of an α-amino acid group onto the N3 position through a C-N bond formed between the Cγ of the amino acid side chain and N3 nitrogen of the base (Figure 1.3) (Saponara and Enger 1974).

1.2.2 Biomolecular interactions in RNA

Hydrogen bonding is a category of biomolecular interaction essential to the folding of RNA and interactions between RNA and other biomolecules, and it is the

Pseudouridine 5'-monophohsphate

Inosine 5'-monophohsphate
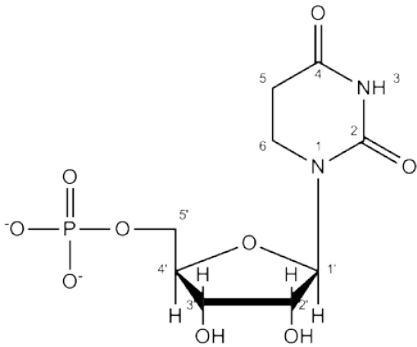
Dihydrouridine 5'-monophohsphate

N2, N2, N7-trimethylguanosine

5'-monophohsphate
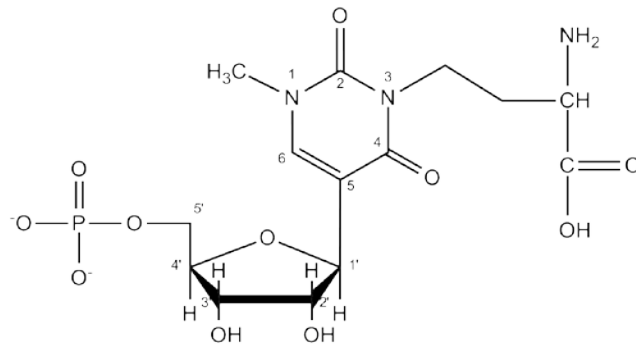
N1-methyl-*N3*-(3-amino-3-carboxylpropyl)pseudouridine 5'-monophosphate

Figure 1.3 Chemical structures of several modified ribonucleic acids.

basic driving force to the formation of base pairing interactions (Santalucia, Kierzek *et al*. 1991). A hydrogen bond can be shown as "X-H···X/Y", where X and Y are heavy atoms with high electronegativity, and H stands for hydrogen. The "X-H" group is the hydrogen bond donor, and the "X/Y" atom is the accepter. The two heavy atoms involved in one hydrogen bond can be identical or different, e.g. "N-H···N" or "N-H···O", which are two of the most commonly observed hydrogen bonds in RNA molecules. It is easy to understand the electrostatic property of hydrogen bond, since difference in the electronegativity creates dipole-dipole interaction, essentially electrostatic, where the partial positive charge on the proton is attracted by the partial negative charge on the accepter atom to align all the three atoms into a linear geometry (Pauling 1982). However, this is only part of the truth, with hyperconjugation of pi and sigma orbitals of all the three atoms involved in a hydrogen bond get into close vicinity, since covalent characteristic has been observed in RNA HNN-COSY experiment (Dingley and Grzesiek 1998). In this experiment, direct J-coupling ($^2J_{NN}$) was used to transfer the magnetization between the two nitrogen atoms involved in a hydrogen bond, and J-coupling is characteristic to atoms covalently connected. Taken together, both electrostatic interaction and covalent bonding are contributing to the formation of a hydrogen bond.

Hydrogen bonding is a primary driving force for base pairing in RNA molecules. Each base structure has three edges available for hydrogen bonding interactions (Figure 1.4 A and 1.4 B)(Leontis and Westhof 1998). Generally, any two nucleotide residues in the RNA can base-pair with each other, while Watson-Crick base pairs are the most widely recognized for their indispensible role in the genetic information

replication and transfer. In RNA, an adenosine residue base pairs with a uridine residue by two hydrogen bonds and a guanosine residue base pairs with a cytidine residue by three hydrogen bonds (Figure 1.4 C and 1.4 D). In both cases, only groups on the "Watson-Crick Edge" of the bases are involved in hydrogen bonding interactions. In DNA, the only difference is that, a thymidine residue, instead of uradine residue, base pairs with an adenosine residue (Figure 1.4 E).

Other common base pairing interactions between nucleotide residues using "Watson-Crick Edge" are wobble base pairs GU (Figure 1.4 F), IU, IC (Figure 1.4 G), and IA (Figure 1.4 H). Francis Crick proposed the "wobble hypothesis" in 1966 (Crick 1966). These wobble base pairs result in base-base recognition pattern different from Watson-Crick base pairs, and potentially affect the identity of amino acid introduced into protein sequences. For instance, if I is in the first position of an anticodon, the anticodon can base pair to U, C, and A, and XYU, XYC, and XYA will code for the same amino acid. This is an example of codon degeneracy.

Edges other than the Watson-Crick edge can be used to form base pair interactions too, e.g. AU Hoogsteen pair (Figure 1.4 I), AU reverse-Hoogsteen pair (Figure 1.4 J), and A/rG pair (Figure 1.4 K). With more than one edge available for hydrogen bonding interactions, one nucleotide residue is readily able to base pair to more than one partner at the same time, and this creates the possibility of base pair triplets and quartets, which contribute to stabilize non-helical RNA structures. For example, in tRNA^{phe}, a base pair triplet is formed between the Watson-Crick GC loop-closing base pair in the D-arm and a G nucleotide residue in the variable loop (Figure 1.4 L), which helps bring the two "paddles" of the "clover leave" together to fold into the

"L" shape tertiary structure (Green and Jones 1986; Gautheret and Gutell 1997). An A-minor motif is formed when an adenine base of a nucleotide residue binds into the minor groove of a base pair, by single or multiple hydrogen bonding interactions (Figure 1.4 M)(Klein, Schmeing *et al*. 2001; Razga, Koca *et al*. 2005). There are four types of A-minor motifs, and all of them are important to stabilize long range rRNA tertiary structures (Nissen, Ippolito *et al*. 2001).

Another category of interactions contributing to the stabilization of RNA structures is stacking (Cate, Gooding *et al*. 1996). The stabilizing effect of stacking is originated primarily from London dispersion energy, an instantaneous induced dipole-induced dipole interaction energy term (Devoe and Tinoco 1962). Theoretically two factors of the interacting groups determine the magnitude of this energy contribution, e.g. the polarizabilities of the groups and the distance between the groups (Cerny and Hobza 2007). All RNA bases have an aromatic ring structure with highly delocalized electron cloud and are by RNA folding, which positions the bases in optimal geometry for stacking. Forcefield calculations on aromatic ring stacking shows that energy minimization is achieved when the two rings are in a "displaced parallel" or a "T-shaped" stacking conformation, dependent on the distance (Sun and Bernstein 1996). In RNA molecules, the "displaced parallel" stacking interactions are the most frequently observed conformation in duplex regions, and the "T-shaped" stacking are usually assumed by residues in the single-stranded region, e.g. an internal bulge region connecting two helices (Diener and Moore 1998). An empirical Coulomb law combined with Lennard-Jones van der Waals equation, involving the summation of energy contributions from each pair of atoms with partial charges on the two interacting

A

B
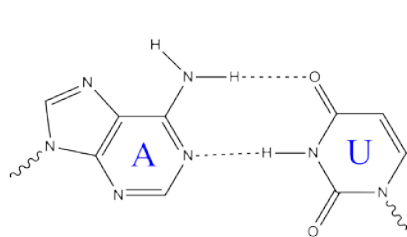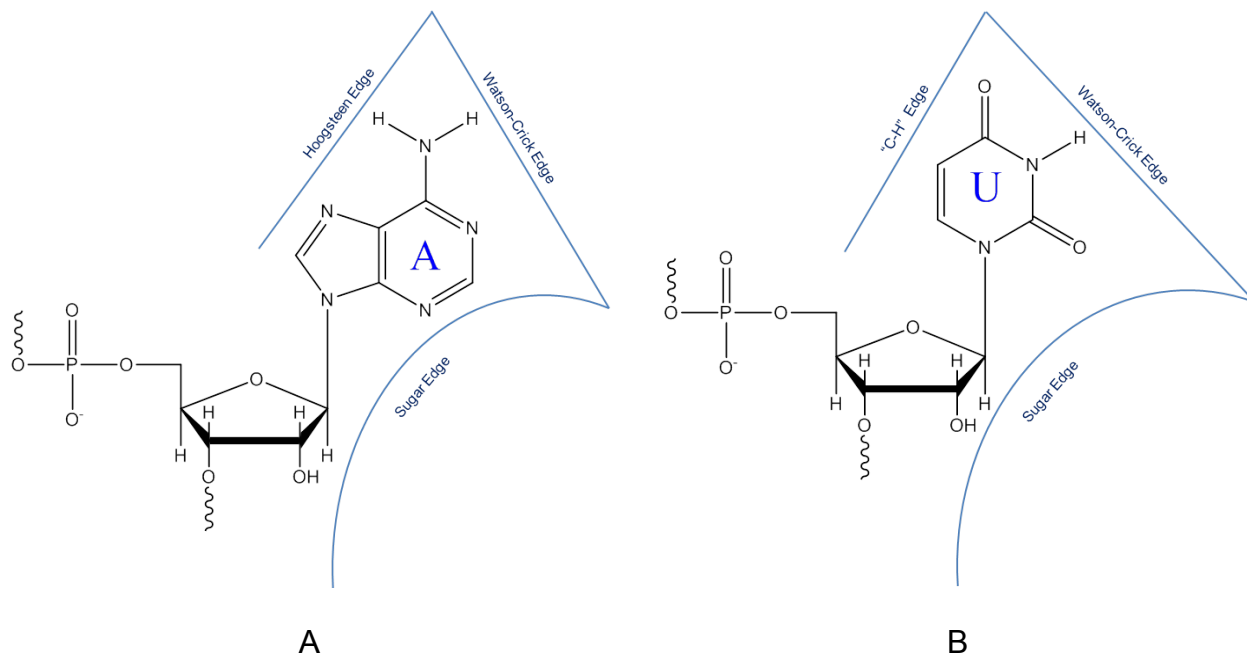
C
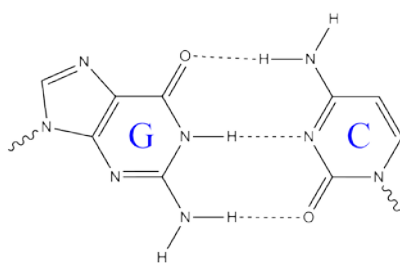
D
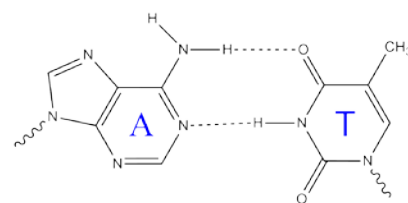
E
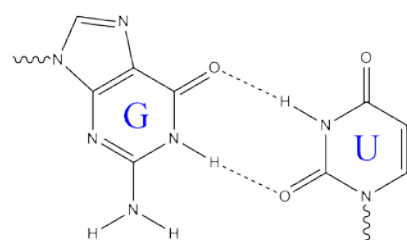
F
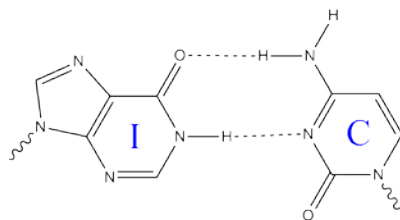
G
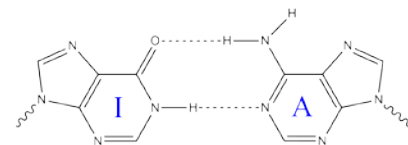
H
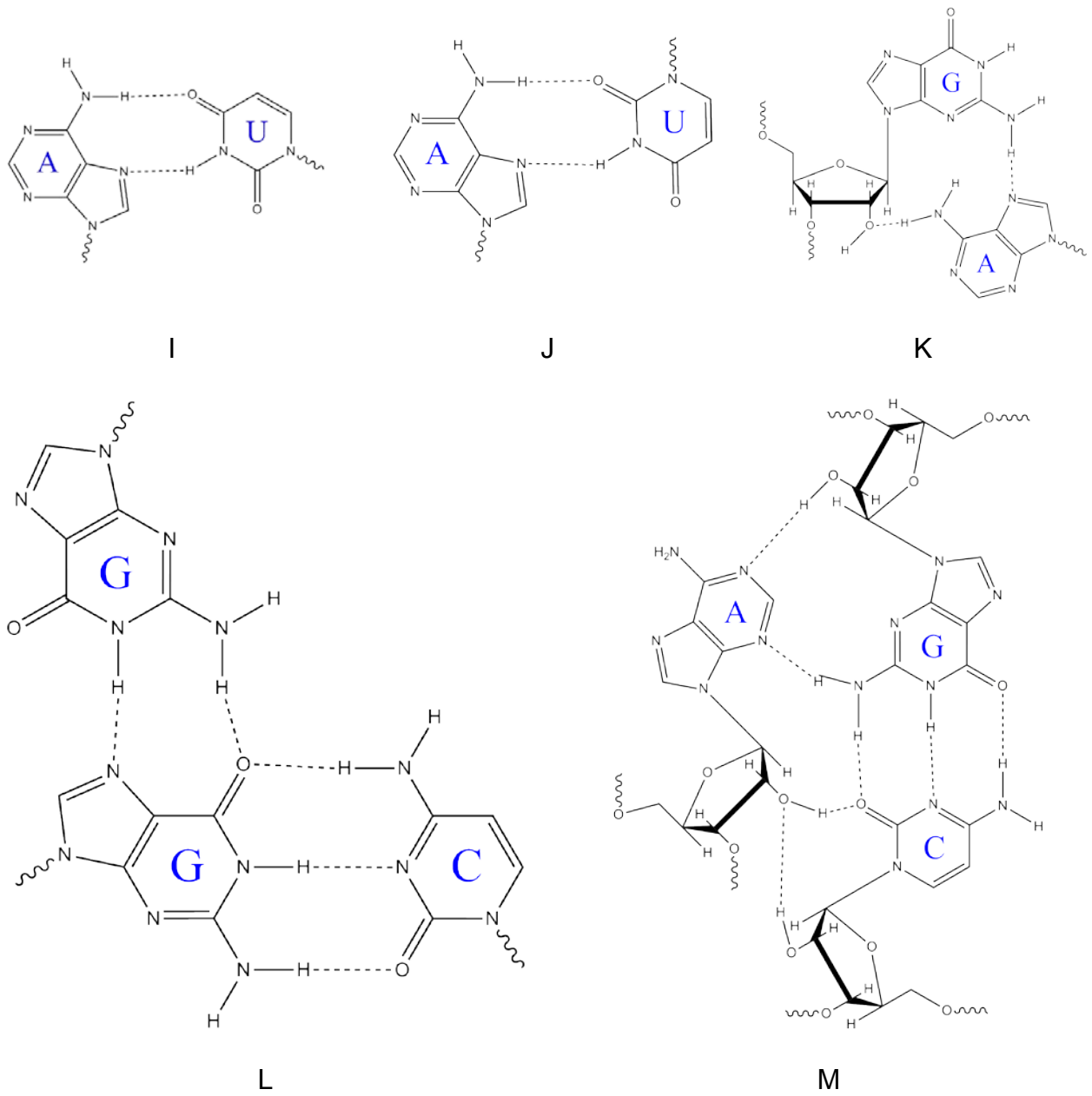
Figure 1.4 Base pair interaction patterns observed in several RNAs.

aromatic rings, is used to calculate the energy contribution of base stacking (Jorgensen and Severance 1990).

Hydrogen bond and stacking contain electrostatic contributions, while electrostatic interactions could be singled out as a third interaction mode when both partners of the interactions carry "full" charges. A fundamental intramolecular electrostatic interaction in RNA is the repulsion between phosphate groups in the backbone. At neutral pH, each phosphodiester linkage carries a negative charge due to the low pK_a of the phosphate moiety, and the electrostatic repulsion favors unfolding of the RNA molecules. In some biophysical experiments with RNA oilgonucleotides, monovalent cations, e.g. Na^+, NH_4^+, and K^+, are enough to relieve the repulsion effect and keep the RNA molecules folded, while a divalent metal ion, i.e. Mg^{2+}, is always employed to mimic the physiological environment, especially in studies on folding of large RNA molecules and huge ribonucleoprotein complexes, e.g. ribosomal subunit assembly and ribosome association (Feig and Ulenbeck 1999; Klein, Moore *et al*. 2004). Attractive electrostatic interactions exist in intermolecular interactions involving RNA and other binding partners, e.g. proteins and small molecule ligands. For instance, most of the ribosomal proteins from yeast are alkaline, and in crystal structures of prokaryotic ribosomes, basic amino acid side chains are shown to stay in close proximity to rRNA backbone phosphate groups (Berk, Zhang *et al*. 2006; Kamita, Kimura *et al*. 2011). These observations provide evidence that electrostatic interactions stabilize the binding between ribosomal proteins and ribosomal RNAs, and may partially determine the binding specificity together with hydrogen bonding and stacking interactions. Interactions between aminoglycoside antibiotic molecules and ribosomal

RNAs also benefit from electrostatic interactions (Chow and Bogdan 1997). The positively charged amino groups on aminoglycoside molecules are shown to have direct contact with backbone phosphate moieties in crystal structures, and one major mechanism of resistance acquisition is to lower the positive charges of the aminoglycoside molecules by nucleotidylation, acetylation, and phosphorylation through enzymatic modifications (Vicens and Westhof 2003; Francois, Russell *et al*. 2005).

1.3 Ribosomes

1.3.1 Ribosome composition

As to RNA modifications, ribosomes are classic treasure trove to show the important roles that modified nucleotides play in the maintenance of ribosome structural integrity and flexibility to assure the normal translation of proteins, and acquisition of antibiotic resistance in bacteria (Decatur and Fournier 2002).

Ribosomes are the universally conserved machinery where translation takes place in all living cells (Nierhaus and Wilson 2004). Ribosomes of bacteria and eukaryotes are similar in that in both cases, the ribosomes are formed by association of a large subunit and a small subunit, but the mass of a eukaryotic ribosome is about 4.3 MDa, much larger than that of a bacterial ribosome, whose mass is "only" about 2.3 MDa (Videler, Ilag *et al*. 2005; Steitz 2008). In a bacterial small subunit (30S) ribosome, there is one rRNA of about 1540 nucleotide long named 16S rRNA and 21 ribosomal proteins. In a bacterial large subunit (50S), there are two rRNAs, a 5S rRNA of 120 nucleotides long and a 23S rRNA of about 2900 nucleotide long, and 33 ribosomal proteins (Figure 1.5). While in a eukaryotic small subunit (40S), there are one 18S rRNA and about 30 ribosomal proteins. In a eukaryotic large subunit (60S), there are
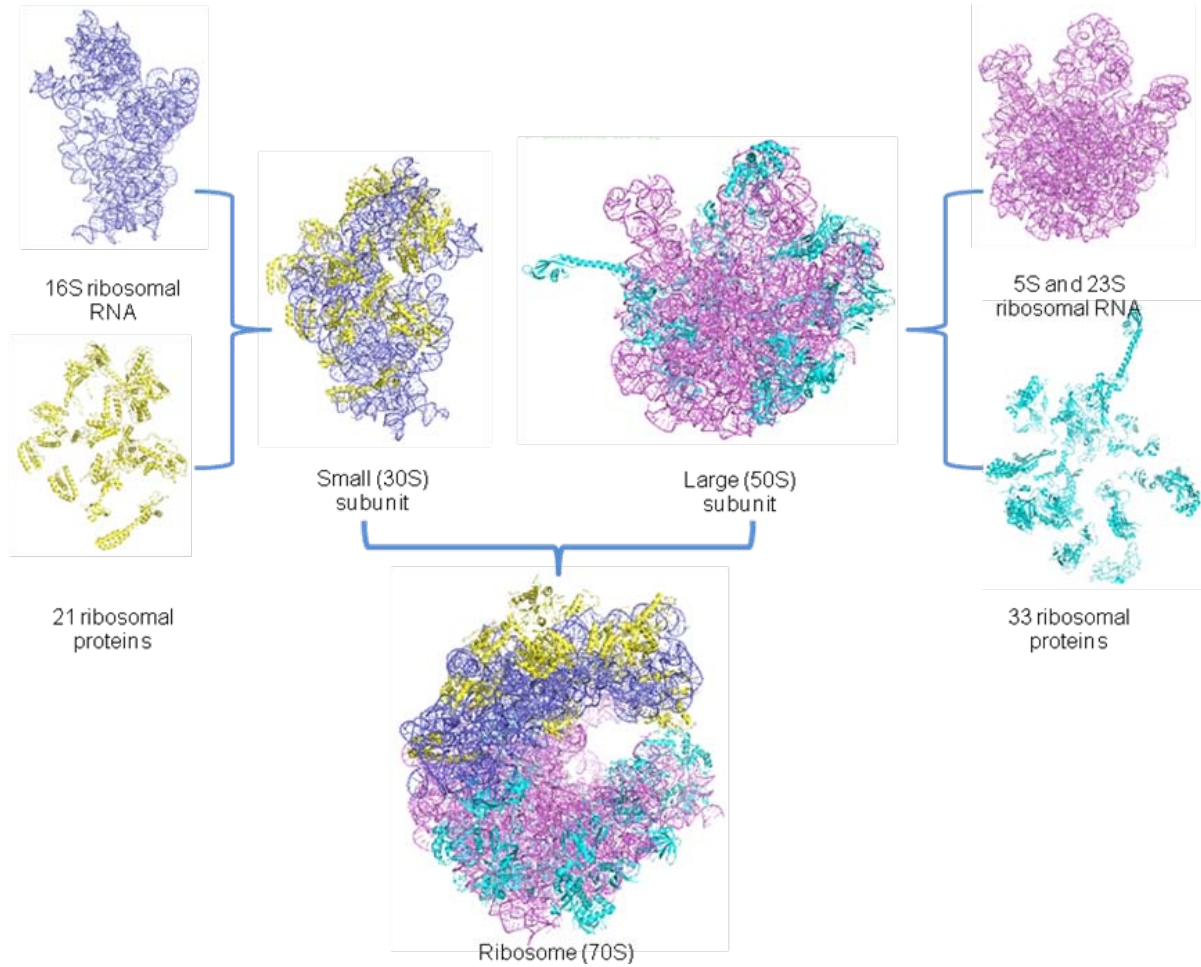
Figure 1.5 Composition of a bacterial 70S ribosome. The pictures were edited from PDB 3R8S and 3R8N (Dunkle, Wang *et al*. 2011).
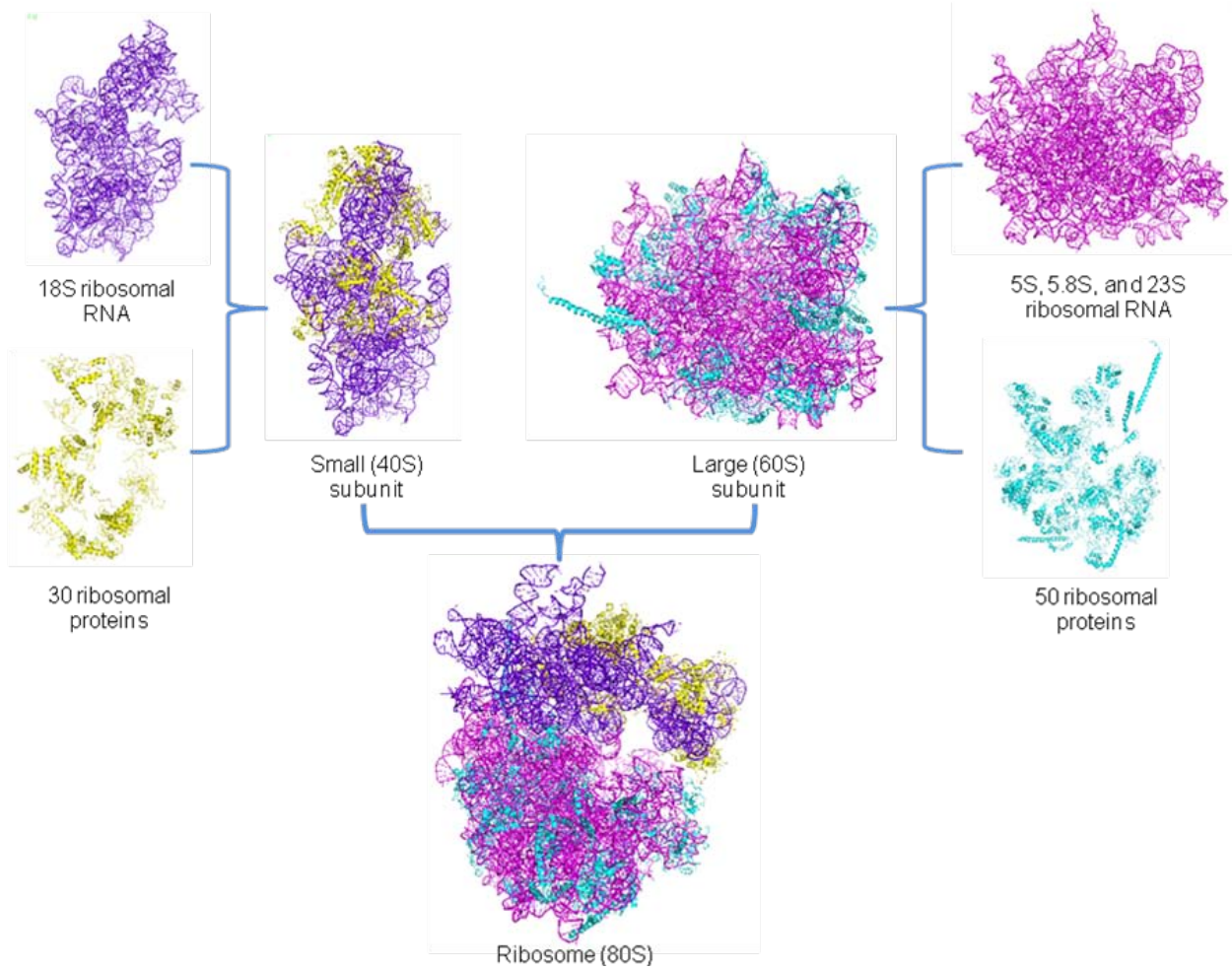
Figure 1.6 Composition of a eukaryotic 80S ribosome. The pictures were edited from PDB 3O2Z and 3O58 (Ben-Shem, Jenner *et al*. 2010).

three rRNAs, i.e. 5S, 5.8S, and 28S rRNAs, and about 50 ribosomal proteins (Figure 1.6) (Armache, Jarasch *et al*. 2010; Ben-Shem, Jenner *et al*. 2010).

As the third domain, archaea have unique properties in composition of ribosomes. Ribosomes of archaea are composed of a 30S small subunit and a 50S large subunit, with a 16S rRNA in the small subunit and two copies of ribosomal RNAs, the 23S rRNA and 5S rRNA, in the large subunit (Teichner, Londei *et al*. 1986). Though the compositions of bacterial ribosomes and archaeal ribosomes are similar, they are different in both the ribosomal RNA sequences and ribosomal protein compositions (Olsen and Woese 1993; Lecompte, Ripp *et al*. 2002). The sequences of 16S rRNA (18S rRNA in eukaryotes) are used in phylogenetic analysis, for the quantity of information they carry and constrained evolution due to their essential functions, and this phylogenetic analysis splits the "prokaryotes" into two domains, bacteria and archaea (Lane, Pace *et al*. 1985; Olsen and Woese 1993). The analysis on ribosomal protein families revealed that the number of archaeal ribosomal proteins is between those of bacteria and eukaryotes, and archaea and eukaryotes share ribosomal proteins absent in bacteria while bacteria do not share ribosomal proteins with either eukaryotes or archaea, respectively, beyond the scope of ribosomal proteins shared by all three domains (Lecompte, Ripp *et al*. 2002; Marquez, Frohlich *et al*. 2011).

From the first partially solved X-ray crystal structure of *Thermus thermophilus* 70S ribosome complexes with tRNAs and mRNA was released in protein data bank on Oct 4, 1999, to the latest release of a *Thermus thermophilus* 70S ribosome complexes with the GTPase release factor 3 on Sept 28, 2011, there are now over 230 individual entries (as of December 2011) for ribosome crystal structures (Cate, Yusupov *et al*.

1999; Jin, Kelley *et al*. 2011). The development of X-ray crystallography and cryo-EM provides insight into structural details of ribosomes at atomic resolution (Spahn, Jan *et al*. 2004).

1.3.2 Structures of the small subunit

Even though the sequence alone of the 16S rRNA is not able to direct the folding of the RNA into its fully native conformation, 16S rRNA does provide a scaffold for binding of the ribosomal proteins to assemble the 30S subunit, with the ribosomal proteins inserted into or patched on the surface of the 16S rRNA structure (Moazed, Stern *et al*. 1986). This is clearly observed when the 16S rRNA tertiary structure and the 30S subunit structure are compared side by side (Figure 1.5 and 1.6) (Schuwirth, Borovinskaya *et al*. 2005; Rabl, Leibundgut *et al*. 2011). The secondary structure of the 16S rRNA is conventionally divided into four domains: 5', central, 3' major, and 3' minor (Gutell, Lee *et al*. 2002). Schuwirth *et al.* dissected the bacterial small subunit structure into several structural features respective to the 16S rRNA domains (Figure 1.7 A and 1.7 B). The small subunit structure corresponding to the 5' domain of the 16S rRNA is associated with ribosomal proteins that are assigned to "shoulder", "body", and "spur". The central domain of the 16S rRNA frames the formation of "platform" region. "Beak", "head", and "neck" are included in the small subunit structure with the 3' major domain of the 16S rRNA. The 3' minor domain of the 16S rRNA, starting from the bottom of the "neck", folds onto the "platform" and "body".

Similar structural features are identified in the small subunit of eukaryotic ribosomes (Figure 1.7 C and 1.7 D) (Rabl, Leibundgut *et al*. 2011). Compared to the 16S rRNA in prokaryotic ribosomes, the eukaryotic 18S rRNA is about 260 nucleotides
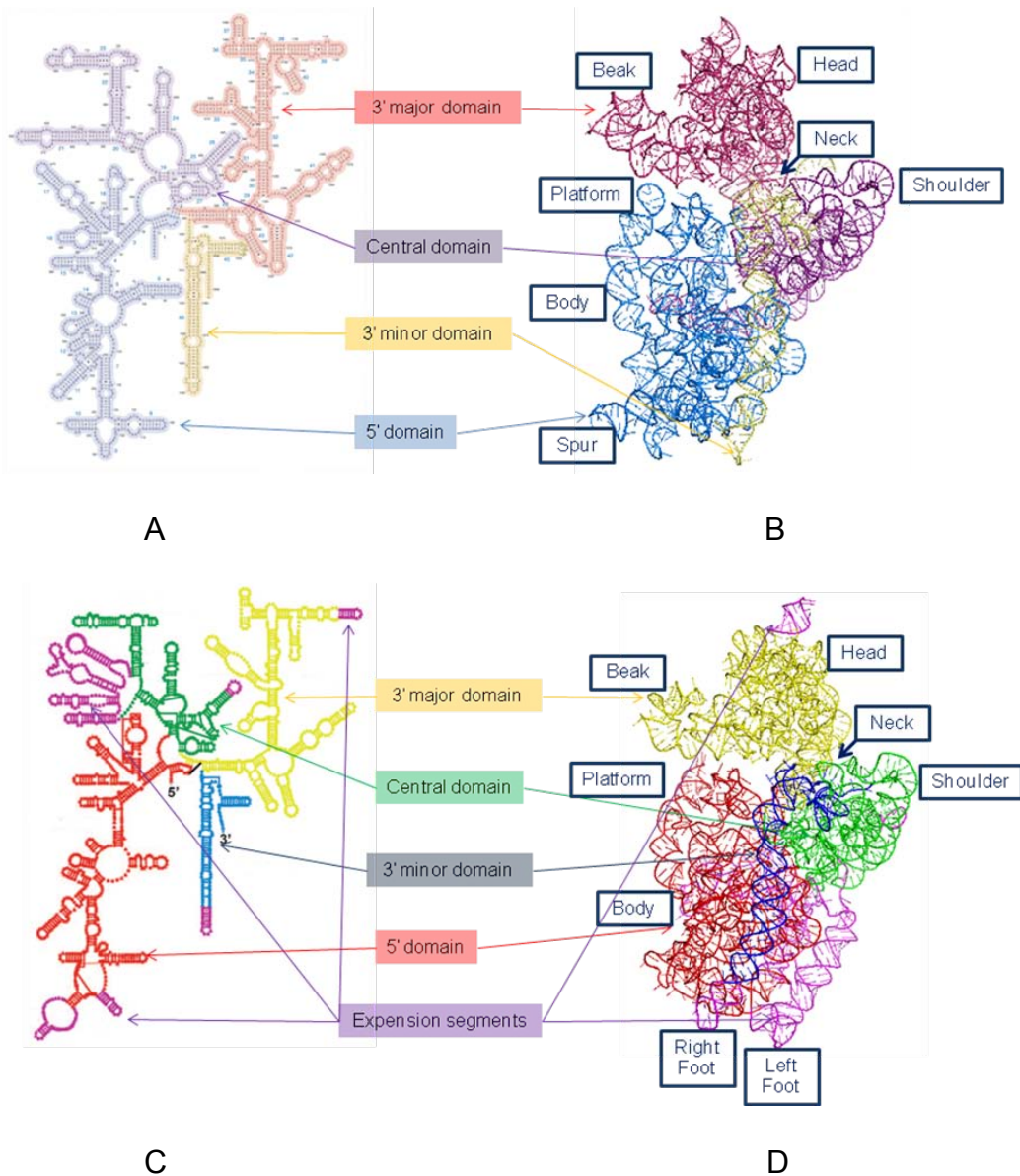
Figure 1.7 Comparison of secondary structures and crystal structures of the small subunit ribosomal RNAs from *Escherichia coli* (A and B) and *Tetrahymena thermophilaI* (C and D)*. The pictures were edited from PDB 3R8N and 2XZM (Dunkle, Wang *et al*. 2011; Rabl, Leibundgut *et al*. 2011). RNA domains from the secondary structures and the morphological features are shown in the crystal structures. The pictures of the secondary structures are from the Noller's lab website. http://rna.ucsc.edu/rnacenter/ribosome_images.html

longer, while the global foldings of the two RNAs are very similar. The "extra" nucleotide residues are distributed into all the four domains of the 18S rRNA, termed eukaryotic expansion segments (ES), when aligned to the 16S rRNA secondary structure, with one huge cluster of about 200 nucleotide (ES6) inserted into the central domain of the 18S rRNA. This cluster of nucleotide residues are folded into three major helices, stretching from the "platform" onto the back of the "body". One of the helices runs parallel with the 3' minor domain, and potentially interacts with L19 of the large ribosomal subunit (Ben-Shem, Jenner *et al*. 2010). The ES6 together with the ribosomal proteins associated with it push ES3, located in the 5' domain of the 18S rRNA, giving rise to two additional features to the eukaryotic small ribosomal subunit, the "left foot" and "right foot" (Rabl, Leibundgut *et al*. 2011).

1.3.3 Structures of the large subunit

In 2010, the 3D crystal structure of the 80S ribosome from *Saccharomyces cerevisiae* was determined and this is the first eukaryotic ribosome structure calculated from X-ray diffraction experiment, with a resolution of 4Å (Ben-Shem, Jenner *et al*. 2010). The global structure of the large subunit looks very similar to its counterpart in the prokaryotes, even though there are 12 eukaryotic expansion segments observed on the 25S rRNA, compared to bacterial 23S rRNA.

The large subunit structure does not show as much as segmentation as the small subunit structure, even though in both prokaryotic and eukaryotic ribosomes, the ribosomal RNAs (23S for *E. coli* and 25S for *S. cerevisiae*) of the large subunit are about 1000 nucleotide longer than the small subunit ribosomal RNAs (16S for *E. coli* and 18S for *S. cerevisiae*), with two extra domains (Figure 1.8 A). Viewed from the

small subunit side, four features lined up on the top of the structures are visible in the large subunits, i.e. L1 stalk on the left side, L7/L12 stalk on the right side, the central protuberance in the middle, and an A site finger wrapping around the neck of the central protuberance stretching from the back to the front of the subunit, composed of the longest helix in Domain II of the large ribosomal RNA (Figure 1.8 B and C) (Ben-Shem, Jenner *et al*. 2010; Dunkle, Wang *et al*. 2011).

The catalytic activity of the peptide bond formation, called peptidyl transferase activity, is an essential function of ribosomes, and it has been considered an activity of the large subunit ribosomal RNA since the early 1990s (Noller, Hoffarth *et al*. 1992). In the secondary structure of *E. coli* 23S rRNA, nucleotide residues essential to the structure and activity of peptidyltransferase are distributed in Domain V around the multi-branch loop region (Figure 1.8 A). Such a secondary structure is also observed in the yeast 25S rRNA, with no eukaryotic expansion segment in this domain (Polacek and Mankin 2005). When highlighted in the crystal structure of the large ribosomal subunit of *E. coli*, these functional nucleotide residues are shown to be clustered at the center region of the subunit, forming a sphere devoid of ribosomal proteins (Figure 1.8 D). Taken together, catalysis of peptide bond formation appears to be an activity of the 23S rRNA in bacteria and archaea, and it is highly possible that the structure and activity are conserved in eukaryotic ribosomes too.

Different from the bacterial and archeal ribosomal large subunit, there is one extra copy of ribosomal RNA, the 5.8S rRNA, in the eukaryotic large ribosomal subunit, and this rRNA binds into Domain I of yeast 25S rRNA shown in the secondary structure (Gutell). In the crystal structure of yeast ribosome solved by Ben-Shem *et al*, the yeast

5.8S rRNA is positioned on the opposite side of the large subunit relative to the small subunit, with three structural segments. Among the 158 nucleotide residues in the 5.8S rRNA, 32 nucleotide residues on the 5'-terminus participate in formation of two helical structures with residues 332-340 and 404-419 of the 25S rRNA, 20 residues on the 3'-terminus form a duplex structure with the 5'-terminus of the 25S rRNA, and all the residues in the central segment are folded into two major helical regions perpendicular to each other (Figure 1.8 E and F).

1.3.4 Intersubunit bridges

The association of ribosomal subunits happens in the initiation phase of the translation process, with multiple interaction regions between the two subunits to stabilize the association state of the ribosome (Frank, Verschoor *et al*. 1995; Yusupov, Yusupova *et al*. 2001; Schuwirth, Borovinskaya *et al*. 2005; Ben-Shem, Jenner *et al*. 2010). These interaction regions are called intersubunit bridges. In the elongation phase, to accommodate the movement of mRNA through the ribosome, ribosomal subunits undergo a ratcheting style motion relative to each other, with breakage and re-establishment of the intersubunit bridges going in cycles (Frank, Zhu *et al*. 1995; Gabashvili, Agrawal *et al*. 2000). A total of 12 intersubunit bridges were reported by Yusupov *et al* in 2001, in a 70S ribosome crystal structure of *Thurmus thermophilus*, with mRNA, P site tRNA, and E site tRNA bound. Interactions between the two ribosomal RNAs (16S and 23S) are proposed in five intersubunit bridges, i.e. B2a, B2b, B2c, B3, and B7a, as the sole interaction pattern. Only one intersubunit bridge, B1b, is stabilized by protein-protein interaction. In intersubunit bridges B1a, B7b, and B8, interactions are only established between RNA-protein. Interactions involving both

A

B

C

D

E   F
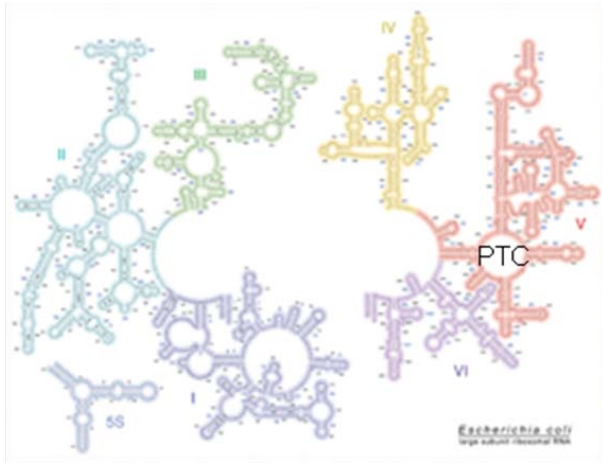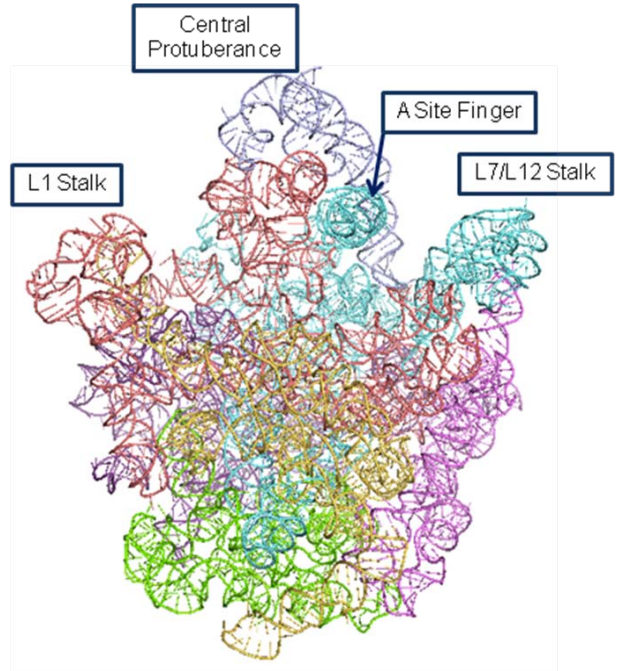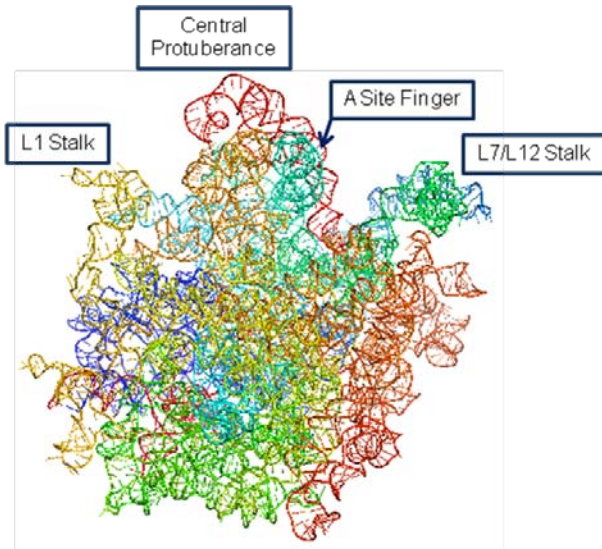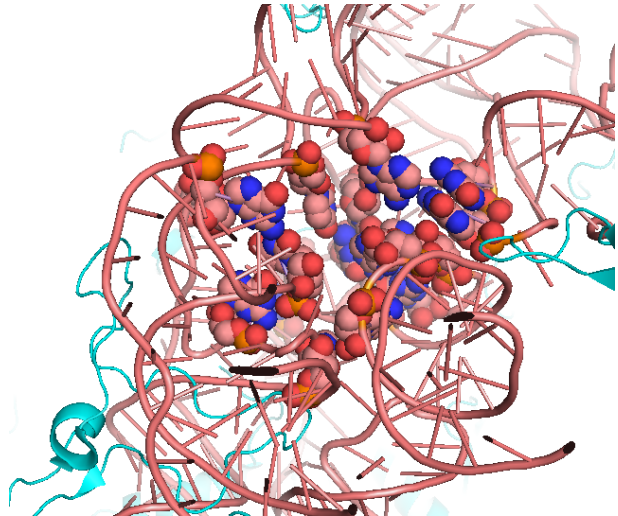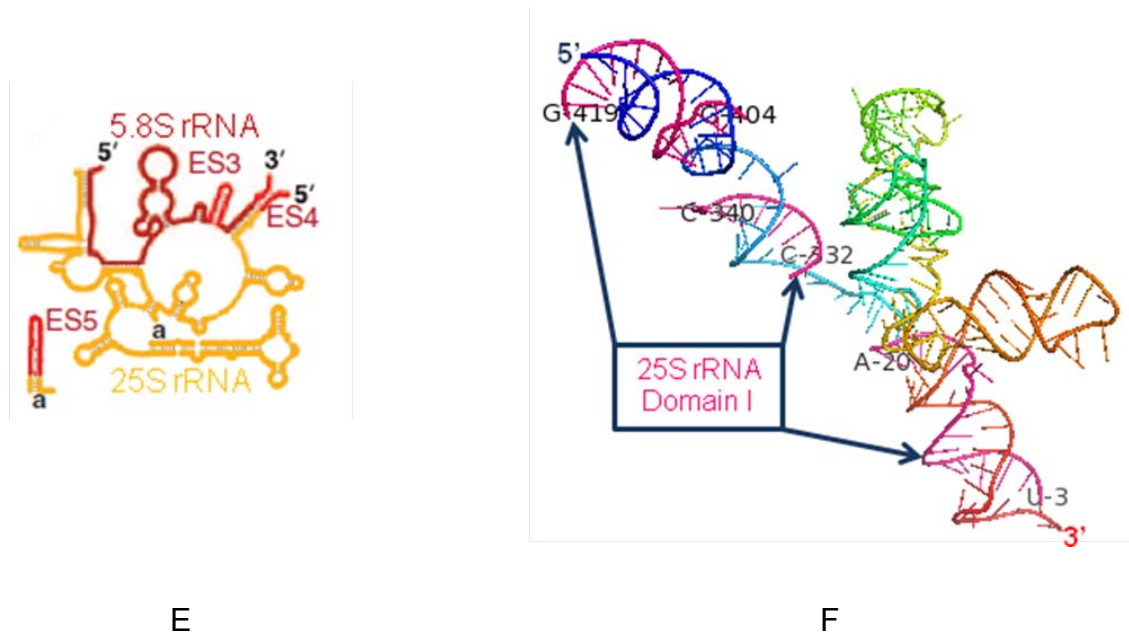
Figure 1.8 Structures of the large subunit ribosomal RNAs from *E. coli* and *Saccharomyces cerevisiae*, and the peptidyl transferase center of the large ribosomal subunit of *E. coli*. The pictures were edited from PDB 3R8S and 3O58 (Ben-Shem, Jenner *et al*. 2010; Dunkle, Wang *et al*. 2011). (A) is the secondary structure of the *E. coli* 23S rRNA with 6 domains shown in different colors, and the 5S rRNA. Similar color pattern is used to identify the domains in the crystal structure (B). Secondary structure of the yeast cytoplasmic 25S rRNA is not available, so "rainbow" is used to color the RNA sequence to identify the direction from 5'-terminus to 3'-terminus in the crystal structure (C). In D, Nucleotide residues in the *E. coli* 23S rRNA with identified functions in the peptidyltransferase center are shown in space filling CPK model (Polacek and Mankin 2005). Secondary structure of 5.8S rRNA from yeast is shown in (E), with the crystal structure shown in (F). 5.8S rRNA sequence is colored in "rainbow" with the termini labeled with corresponding colors, and nucleotide residues on 25S rRNA involved in formation of three helical structures with 5.8S rRNA are colored in hotpink.

RNA-RNA and RNA-protein are seen in intersubunit bridges B4, B5, and B6 (Figure 1.9 A). During the translation process, most intersubunit bridges distributed in the "body" of the small subunit do not show much motion, while B1b in the "head" moves for about 20 Å, and B7a in the "platform" and B8 close to the "spur" move for about 10 Å, where all the three intersubunit bridges are potentially broken by ratcheting motion (Frank, Zhu *et al*. 1995; Gabashvili, Agrawal *et al*. 2000; Zhang, Dunkle *et al*. 2009).

With improvement in crystal structure resolution, more structural details of the intersubunit bridges are identified. In intersubunit bridge B3, interactions between h44 of the 16S rRNA and H71 of the 23S rRNA are stabilized by intricate hydrogen bond networks, including both intrastrand and interstrand hydrogen bonds. One of the networks involves G1959-C1947 Watson-Crick base pair in H71, and A1483-G1417 sheared base pair in h44 (Figure 1.10 A). Among the four intersubunit hydrogen bonds, three are formed between the sugar edge of A1483 and the minor groove of G1959-C1947 base pair, and one is formed between the sugar 2' hydroxyl group of G1959 and N1 of A1483 on the Watson-Crick edge. This is a typical A minor motif (Nissen, Ippolito *et al*. 2001; Razga, Koca *et al*. 2005). Intersubunit bridge B4 is established between S15 of the small subunit and H34 loop region of the 23S rRNA in the large subunit (Figure 1.10 B). It is clearly shown that the interface on S15 is lined up with at least three amino acid residues with hydrophobic side chains, i.e. Leu55, Leu56, and Val59. They form a hydrophobic binding pocket to accommodate the base of A715 in 23S rRNA. Weak as this intersubunit bridge appears, it is essential for subunit association (Maivali and Remme 2004). If A715 is methylated on the N1 position, one positive charge introduced to this position can potentially form electrostatic repulsion against

Arg52 on S15, and destabilize B4 intersubunit brige, resulting in severely compromised subunit association (Maivali and Remme 2004). A third type of intersubunit bridges are formed between ribosomal proteins from the two subunits. Benefiting from the various properties of the amino acid side chains, proteins are able to participate in electrostatic, hydrogen bonding, and hydrophobic interactions. In intersubunit bridge B1b, consists of salt bridges between three pairs of oppositely charged amino acid side chains dominate the stabilizing contribution (Figure 1.10 C). The residues on S13 are Arg 69, Arg 90, and Asp81, paired with Asp143, Glu133, and Arg111 on L5 respectively, within a distance of 6 Å. In the crystal structure of yeast ribosome, four extra eukaryote-specific intersubunit bridges are identified (Figure 1.9 B). All of them are distributed close to the "foot" regions of the small ribosomal subunit. Due to the limitation of resolution, some small subunit ribosomal proteins located in this region were not well resolved. In this case, the crystal structure of the small ribosomal subunit from *Tetrahymena thermophilal* is taken to postulate the possible intersubunit bridges potentially involving eukaryotic small subunit ribosomal proteins. When the "foots" and adjacent regions in the eukaryotic small subunit are compared to their counterparts in the bacterial small subunit, two major differences involving expansion segments are readily observed: 1) ES3 and ES6 form additional RNA structure, which provides a framework for three eukaryotic small subunit ribosomal proteins (S6E, S8E, and S17E) to bind, and 2) the three ribosomal proteins cover up most of the RNA surface area on the subunit interface. In contrast, there is only one ribosomal protein, S20, in the corresponding region of bacterial small subunit, and it is buried inside the RNA structure (Figure 1.10 D and E). It is highly possible that these eukaryotic ribosomal proteins (S6E, S8E, and S17E) mediate some
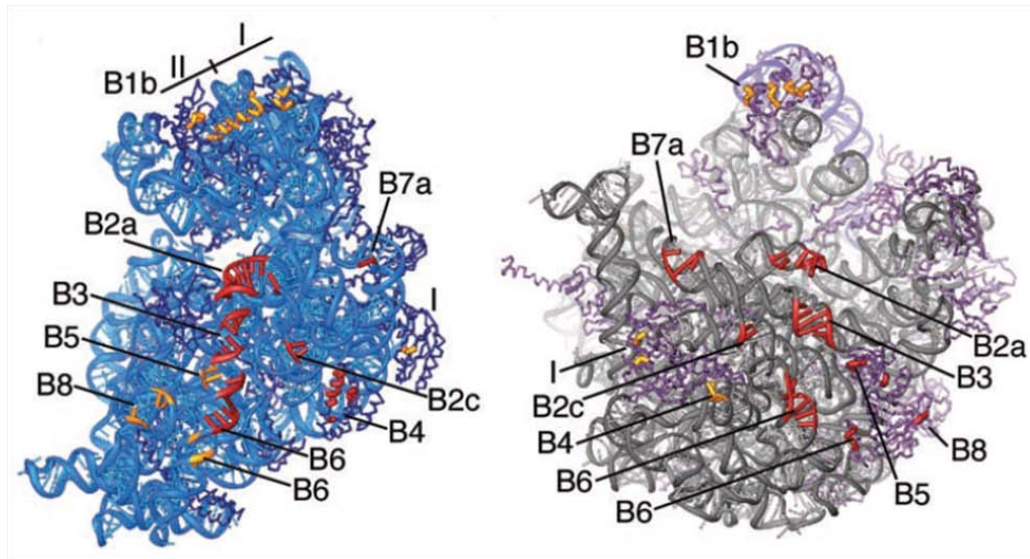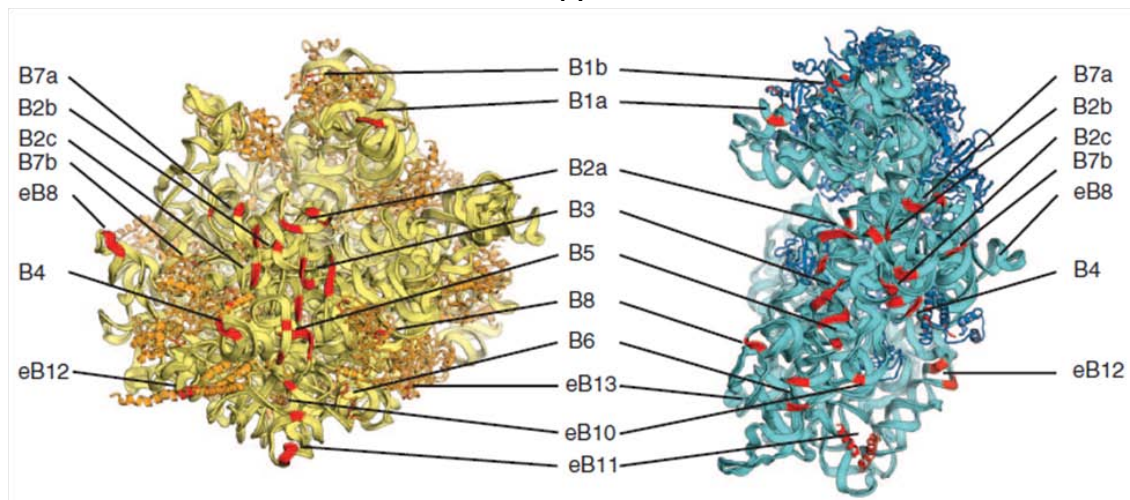
Figure 1.9 Comparison of ribosomal intersubunit bridges from *E. coli* and *S cerevisiae* (Schuwirth, Borovinskaya *et al*. 2005; Ben-Shem, Jenner *et al*. 2010). In A, the *E. coli* small ribosomal subunit is shown on the left side and large subunit on the right side. In B, the *S cerevisiae* small subunit is shown on the right side and the large subunit on the left side.
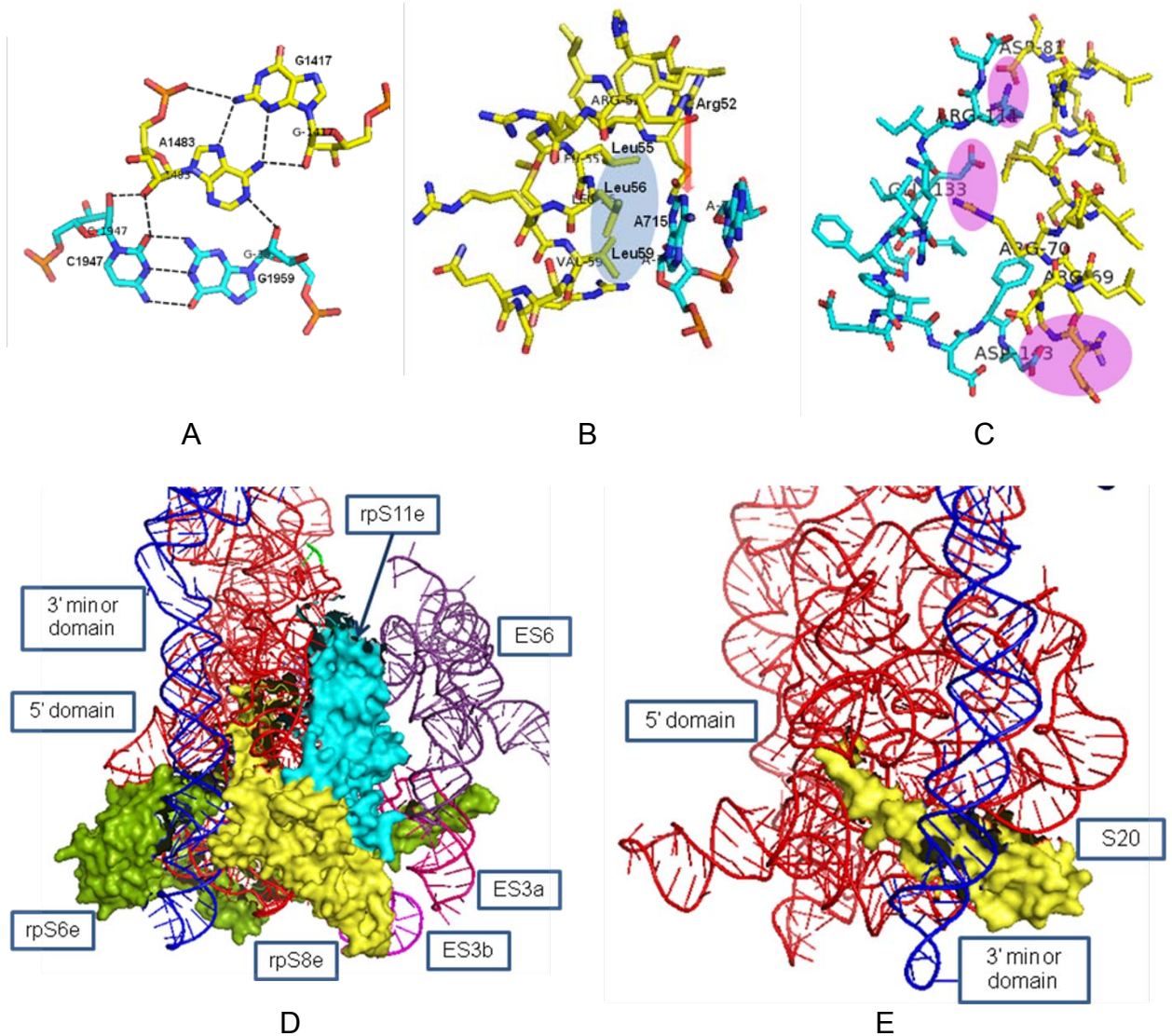
Figure 1.10 Interaction types of intersubunit bridges from *E. coli* and comparison of small ribosomal subunit lower regions from *T. thermophilal* and *E. coli* (Dunkle, Wang *et al*. 2011; Rabl, Leibundgut *et al*. 2011). Intersubunit bridge B3 is shown in (A), where the A1483 of the 16S rRNA forms A-minor interaction with G1959-C1947 base pair of the 23S rRNA, and hydrogen bonds between the A1483 and G1959 further stabilize this intersubunit bridge. All hydrogen bonds are shown in black dashed lines. Intersubunit bridge B4 is shown in (B), where the A715 on H34 loop region of the 23S rRNA binds

into a hydrophobic pocket formed by Leu55, Leu56, and Val59 on S15. The red arrow shows (B), if A715 is methylated, the positive charge on N1 of A715 forms electrostatic repulsion against Arg52 on S15, which could severely compromise the stability of B4 intersubunit bridge. Intersubunit bridge B1b is shown in (C), where Arg69, Arg70, and Asp81 on S13 (yellow) form three salt bridges with Asp143, Glu133, and Arg111 on L5 (Cyan). Each pair of the side chains involved in the salt bridge interactions are shadowed in one pink ellipse. Compared with the small subunit crystal structure from *E. coli* (E), with only one ribosomal protein S20 buried into the lower "body" region, three ribosomal proteins are found in the corresponding region in the crystal structure from the *T. thermophilaI* small subunit (D). Ribosomal RNAs are shown in "cartoon", and ribosomal proteins are shown in space filling CPK.

of the eukaryotic-specific intersubunit bridges.

1.3.5 tRNA binding and the peptidyltransferase activity

In the bacterial 70S ribosome crystal structure from *T. thermophilus*, three tRNAs along with the mRNA are clearly observed (Figure 1.11 A) (Jenner, Demeshkina *et al*. 2010). Following the directionality of the mRNA from 5' terminus to 3' terminus, the three tRNA binding sites on ribosomes are termed A-site (Aminoacyl), P-site (Peptidyl), and E-site (Exit). A wide spread of interaction surface areas between the tRNAs and ribosome subunits are shown in Figure 1.11 B and C, for frontal and rear views, respectively. During translation, mRNA threads through a tunnel around the "Neck" of the small subunit, and possibly forms no direct contact with the large subunit. As the result, most contact surface between the three tRNAs and the ribosome close to the tRNA anticodon loops are on the small subunit, with one exception that of H69 of the 23S rRNA shows direct contacts to both the A-site tRNA and P-site tRNA with its loop region and helix base region, respectively.

A variety of nucleotide residues of the ribosomal RNAs and amino acid residues of small subunit ribosomal proteins interact with anticodon stem loops of the tRNAs and the codon-anticodon minihelix to keep the tRNA registered. One of the best elucidated structural features for tRNA binding is the hydrogen bonding network, involving seven hydrogen bonds outside of the codon-anticodon minihelix, between nucleotide residues G530, A1492, and A1493 of 16S rRNA and the A-site codon-anticodon base pairs (Figure 1.11, E, F, and G). Over 40 years ago, Davies and Davis reported that aminoglycoside antibiotics could induce misreading of mRNA codons, and 10 years ago Ramakrishnan suggested the pharmacological relevance of this structural feature

(Davies and Davis 1968; Ramakrishnan 2002). With crystal structures of improved resolution put side-by-side, it is readily shown that, when neomycin, an aminoglycoside antibiotic, binds into the A-site helix of 16S rRNA, the nucleotide residues A1492 and A1493 are displaced from the interior of the helix, and flipped out in a configuration very similar to that when the correct codon-anticodon helix is formed (Figure 1.11 H)(Borovinskaya, Pai *et al*. 2007; Jenner, Demeshkina *et al*. 2010; Dunkle, Wang *et al*. 2011). Tolerance of codon-anticodon recognition to the wobble base pair in the third position is also illustrated in the figures (Figure 1.11 E, F, and G). Instead of a Watson-Crick base pair in the third position, a GU wobble base pair is formed between U3 of the mRNA and G34 of the tRNA, and the base pair morphology is not scanned by G530. Due to the similar stability and morphology of GU wobble base pair and the Watson-Crick base pair, the codon-anticodon minihelix is not interrupted (Varani and McClain 2000).

Besides H69 base region, P-site tRNA shows direct contact to the h24 loop residue A790 and two residues, G1338 and A1339, on a multibranch loop connecting h29, h30, h41, and h42. These three nucleotide residues form a molecular check point for the translocation of tRNA from P-site to E-site (Schuwirth, Borovinskaya *et al*. 2005). Ribosomal proteins S12 and S13 are potentially involved in interactions with P-site tRNA too, even though no clear interaction pattern is observed in 3I3G (Carter, Clemons *et al*. 2000). While in the same crystal structure, amino acid side chains on S7 and S11 are positioned in close proximity to the E-site tRNA anticodon stem loop, with multiple positively charged residues. And three nucleotides, G693, A694, and A695, on h23 loop participate in E-site tRNA binding in the same region.

There are several interaction surface areas on the large subunit form tRNA binding pockets too, and the most distinguished is the PeptidylTransferase Center (PTC), where the 3' termini of A-site tRNA and P-site tRNAs are positioned within 10 Å of one another. Several nucleotide residues have direct interactions with the two tRNAs, while others either form the peptide exit tunnel or embrace the primary binding nucleotide residues (Figure 1.11 I) (Polacek and Mankin 2005). Besides the nucleotide residues shown in Figure 1.11 I, other nucleotide residues and amino acid residues are shown to be close to the PTC, within 6 Å from the tRNAs (Figure 1.11 J). His3 on L27 is sandwiched between the 3'-termini of A-site tRNA and P-site tRNA, and the distances between His3 NE2 and all the four 3'-terminal sugar hydroxyl-O range from 6.3 Å to 12.0 Å. In another crystal structure (2WRO) with His3 coordinates solved, where the A-site is vacant, and only one tRNA binds into the P-site, the distances between the NE2 of His3 and two 3'-terminal sugar hydroxyl-O of the tRNA are longer than 12 Å, and both hydroxyl-O are in an orientation to push the potential aminoacyl group away from the His3. Previous studies show L27 is essential to the assembly of large ribosomal subunits, and may be involved in the peptidyltransferase activity (Wower, Wower *et al*. 1998). It has been suggested that 23S rRNA is able to catalyze peptide bond formation, and the function of L27 in peptidyltransferase center is largely to position the A-site tRNA. It would be interesting to identify the function and importance of L27 amino acid residues close to the N-termnus (Schmeing, Seila *et al*. 2002; Trobro and Aqvist 2008).

In 1992, Noller *et al* reported that the 50S ribosomal subunit from *Thermus aquaticus* was able to catalyze peptide bond formation at similar efficiency as the complete *E. coli* 70S ribosome, even after rigorous treatment with 0.5% SDS, 1 mg/mL

A

B

C

D

E

F
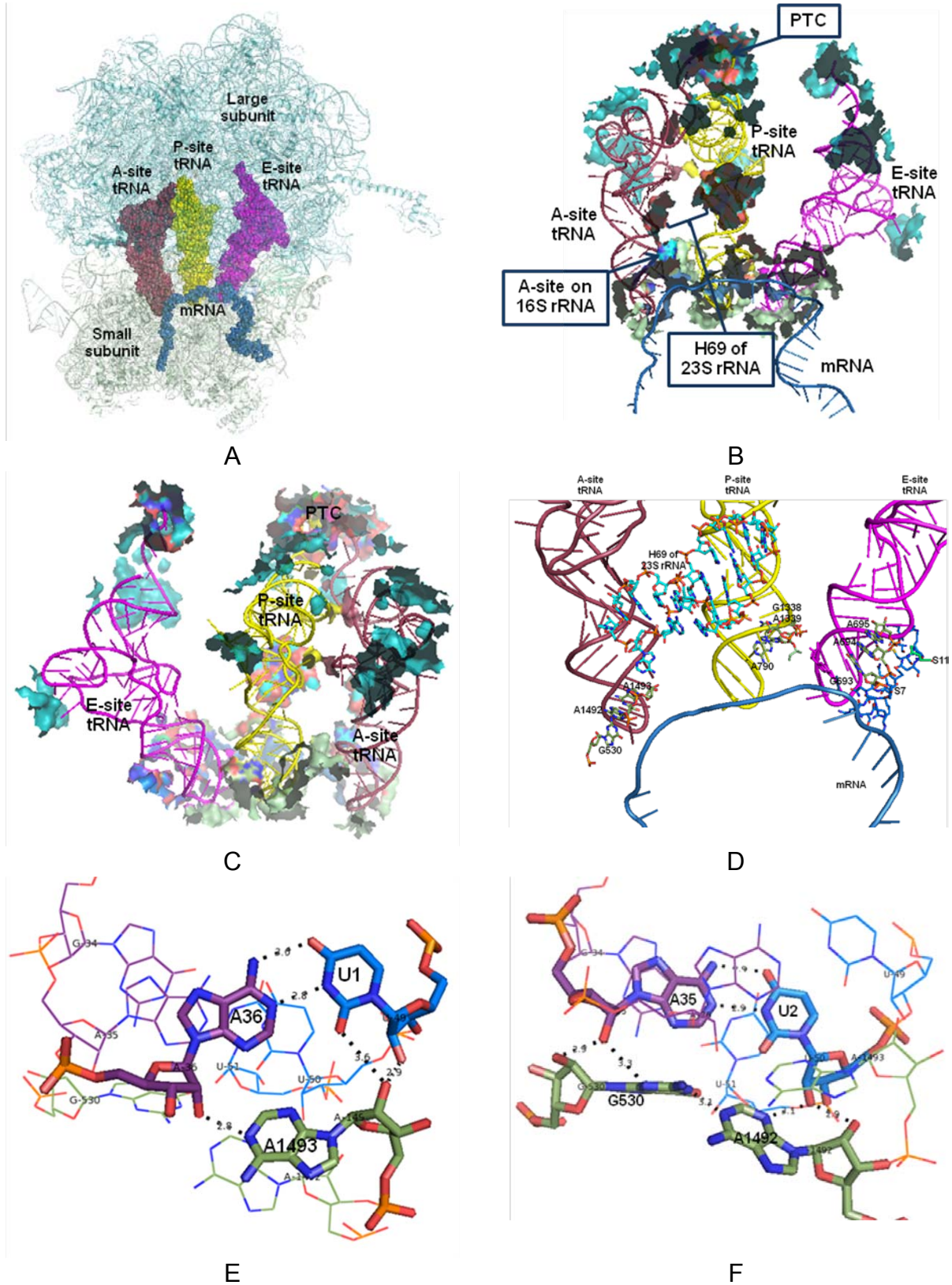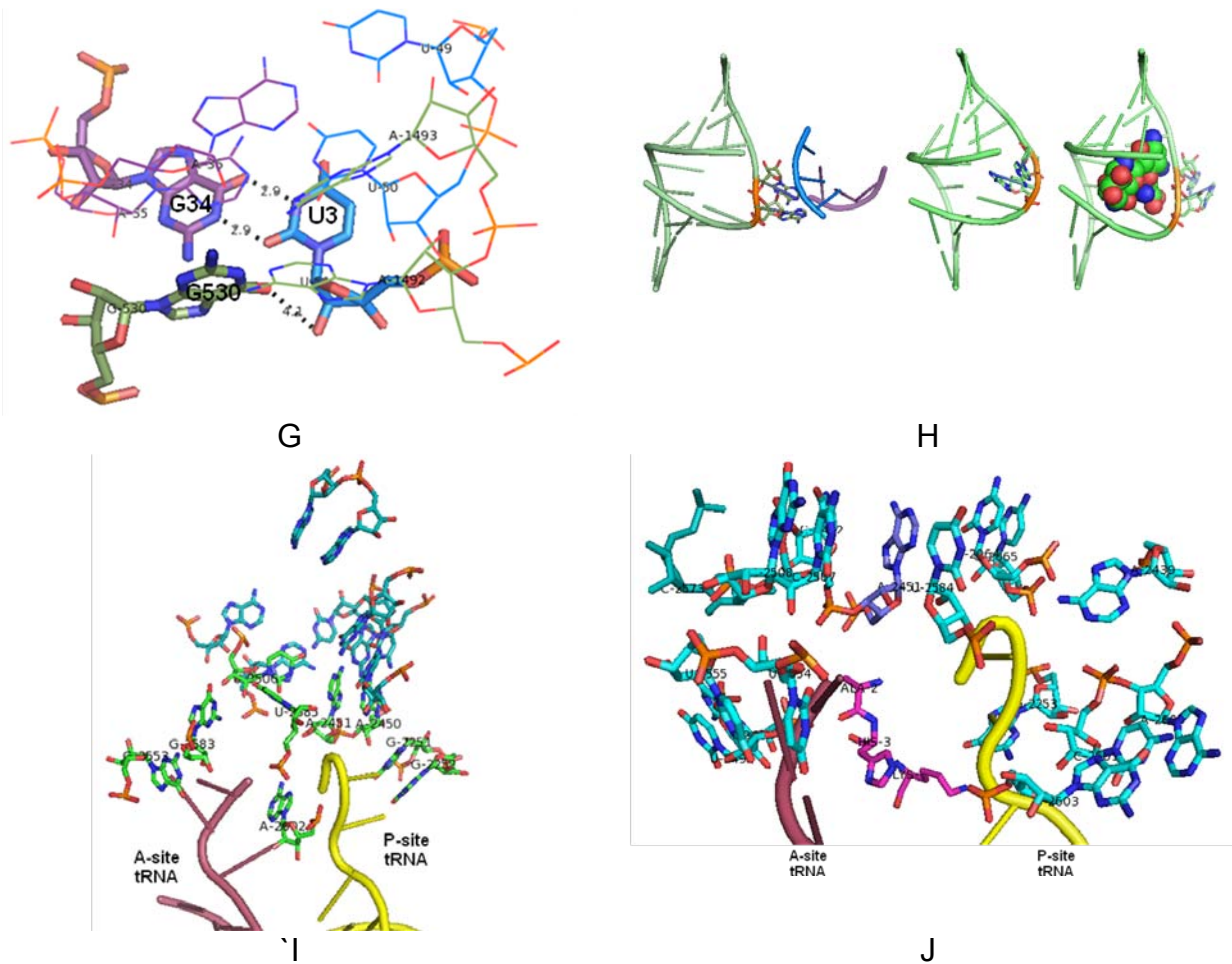
Figure 1.11 Interactions between the three tRNAs, mRNA, and the ribosome. The pictures are adapted from crystal structures 3I8G, 3I8F, 3R8N, and 2QAL (Borovinskaya, Pai *et al*. 2007) (Jenner, Demeshkina *et al*. 2010; Dunkle, Wang *et al*. 2011).Three tRNAs are shown to be bound in A-site, P-site, and E-site (A). Anticodon loops of the tRNAs form base pairs with the mRNA codons, oriented toward the small subunit, while the majority of the tRNA structures are buried into the large subunit. This is also illustrated in B and C, with the distribution of the tRNA-ribosome contact surface areas on each subunit. To define the possible contact surface areas, residues in the two subunits within a distance of 6 Å from atoms on the three tRNAs were chosen, and shown in "surface" rendering. Contact areas on the small subunit are colored in light

green and those on the large subunit are colored in cyan. In B, A-site residues on the 16S rRNA, H69 on the 23S rRNA, and the PeptidylTransferase Center (PTC) in the large subunit are highlighted in the frontal view. Residues on the ribosome participating in critical tRNA-ribosome interactions close to the tRNA anticodon loops are shown in "stick" rendering (D). Residues interacting with A-site tRNA include G530, A1492, and A1493, on the 16S rRNA and the loop residues of H69 on the 23S rRNA. On the P-site tRNA-ribosome interface, A790, G1338, and A1339 on the 16S rRNA and the stem residues of H69 on the 23S rRNA are shown. Ribosomal proteins S7 and S11, together with the G693, A694, and A695 on 16S rRNA are involved in E-site tRNA anticodon loop binding. A hydrogen bond network forms between the A-site 16S rRNA residues G530, A1492, and A1493, and the codon-anticodon minihelix is shown in E, F, and G, with the hydrogen bonds shown as in black dashes. In H, three structures of the A-site Helix on 16S rRNA are shown. A1492 and A1493 (shown in "stick" rendering) are flipped out of the A-site helix when cognate codon-anticodon minihelix is formed (left), and when no tRNA binds in the A-site, the two residues are stacked into the A-site helix (middle). The aminoglycoside antibiotic neomycin can displace the A1492 and A1493 out of the A-site helix (right), which assumes a conformation very similar to that when the cognate codon-anticodon minihelix is formed. In G, nucleotide residues on 23S rRNA directly involved in tRNA binding are shown in green, and residues forming exterior structure of the PTC are shown in cyan. Many other nucleotide and amino acid residues are positioned within 6 Å from the A-site and P-site tRNAs (H), including the N-terminal segment of L27 (in pink).

of proteinase K, and phenol extraction, with a minimal aminoacyl-tRNA mimic and puromucin as substrates (Noller, Hoffarth *et al*. 1992). While a brief treatment with RNase T1 completely abolished the catalysis activity. This finding laid down solid experimental evidence to assign the peptidyltransferase activity to the 23S ribosomal RNA. And, the chemistry mechanism behind this catalysis activity has been debated for a long time (Thompson, Kim *et al*. 2001; Hiller, Singh *et al*. 2011). It has been widely accepted that, peptidyltransferase center of the large ribosomal subunit helps bring the A-site tRNA and P-site tRNA into close proximity, in an orientation favoring the peptide bond formation. The 2-hydroxyl of the P-site tRNA A76 participates in the hydrogen bond network to conduct the proton transfer (Zaher, Shaw *et al*. 2011). Hiller *et al.* used isotope kinetic effect to identify reactions involved in the rate-limiting step. They proposed that formation of the peptide bond and the deprotonation of the amino/amide group are *concerted*, which constitutes the rate-limiting reaction, followed by a fast breakdown of the negatively charged intermediate, and ribosomes help destabilize the substrate and pass the proton to other coordinated base (Hiller, Singh *et al*. 2011).

1.3.6 Ribosomal subunit assembly

Ribosomal RNA transcription, folding, and modification ribosomal protein expression, modification, and binding are coordinated in the ribosomal subunit assembly and subunit association processes (Mizushima and Nomura 1970; Williamson 2005; Shajani, Sykes *et al*. 2011). For the first time, the general dimension and shape of the 16S rRNA in isolated form and in the 30S ribosomal subunit were compared by Serdyutk *et al.* in space 1982 by neutron scattering experiments, and they found that in both cases, 16S rRNA were folded into similar dimension and shape, with the

assembled 16S rRNA more compact (Serdyuk, Agalarov *et al*. 1983). In the case of the 23S rRNA, Nitta *et al.* discovered that even 23S rRNA prepared from *in vitro* T7 polymerase transcription could catalyze the peptide bond formation reaction, without any ribosomal protein component or post-transcriptional modification (Nitta, Ueda *et al*. 1998). This body of evidence shows that ribosomal RNAs's nucleotide sequences dictate the folding of ribosomal RNAs themselves, and the function of ribosomes. While other factors, i.e. ribosomal proteins, modification enzymes, assembly factors, and translation factors are indispensible in the biogenesis of active ribosomes and the translation process(Kaczanowska and Ryden-Aulin 2007).

Ribosomal proteins demonstrate their importance in biogenesis of ribosomes by binding into ribosomal RNAs before the transcription is completed. In bacterial cells, ribosomal proteins are detected in the nucleoid (de Narvaez and Schaup 1979). In eukaryotic cells, ribosomal proteins are imported, by a family of importins, into the nucleolus, a sub-compartment of the nucleus, where ribosomal RNAs are transcribed (Jakel and Gorlich 1998). The investigation of the co-transcriptional ribosome assembly was started with pulse labeling in *in vivo* conditions, where a series of ribosomal subunit precursors were identified. They are 21S, 26S, and 30S precursors for bacterial 30S subunit, and 30S, 43S, and 50S precursors for bacterial 50S subunit (McCarthy, Britten *et al*. 1962; Mangiarotti, Apirion *et al*. 1968; Osawa, Otaka *et al*. 1969; Homann and Nierhaus 1971; Lindahl 1973). Two dimensional gel electrophoresis was used to identify the ribosomal protein composition in each precursor qualitatively (Table 1.1) (Nierhaus, Bordasch *et al*. 1973). Until recently, ribosomal protein compositions were quantified in the precursors accumulated in cells with the help of growth perturbation

Table 1.1 Protein composition of the precursors of the 30S and 50S subunits

30S precursor		50S precursors					
Protein	21S	Protein	32S	43S	Protein	32S	43S
S1	+	L1	+	+	L21	+	+
S2	-	L2	-	-	L22	+	+
S3	-	L3	(±)	+	L23	(±)	+
S4	+	L4	+	+	L24	+	+
S5	+	L5	+	(±)	L25	+	+
S6	-	L6	-	-	L27	+	(+)
S7	-	L7	(±)	+	L28	-	-
S8	+	L8 }	+	} (+)	L29	+	+
S9	-	L9			L30	+	+
S10	-	L10	(±)	+	L31	-	-
S11	-	L11	(±)	+	L32	-	-
S12	-	L12	-	-	L33	-	+
S13	+	L13	+	+			
S14	-	L14	-	+			
S15	+	L15	-	+			
S16	+	L16	-	-			
S17	+	L17	+	+			
S18	-	L18	+	+			
S19	-	L19	-	(+)			
S20	+	L20	+	+			
S21	-						

+. Present in normal amount; (+), present in reduced amount; (±), present in traces; -. not detectable.

This table is remade to improve the appearance, and original format is strictly followed (Nierhaus, Bordasch *et al*. 1973).

(Table 1.1) (Charollais, Pflieger *et al*. 2003; Sharpe Elles, Sykes *et al*. 2009; Sykes, Shajani *et al*. 2010).

It is still impossible to quantify assembly of ribosomal proteins in ribosome precursors at normal growth conditions *in vivo*. As an alternative method, *in vitro* ribosomal subunit reconstitution provides a possibility to investigate the order and interplay of ribosomal proteins binding to the ribosomal RNAs. The Nomura lab started this effort in the early 70's with 30S ribosomal proteins purified from bacterial ribosome preparations. They tested the activity of reconstituted 30S subunit with complete set of ribosomal proteins or a set of ribosomal proteins with one ribosomal protein absent (Mizushima and Nomura 1970; Held, Mizushima *et al*. 1973). The Noller lab introduced recombinant ribosomal proteins into reconstitution experiments (Culver and Noller 2000). The merit of this endeavor is not limited to high yield of pure ribosomal proteins. It makes possible to incorporate mutations and modifications into ribosomal proteins to explore the assembly process in more depth, e.g. Fe(II) tethered ribosomal proteins for directed hydroxyl radical probing on ribosomal RNA structures (Culver and Noller 2000).

Early *in vitro* reconsititution experiments correctly revealed most aspects of the assembly maps of the small and large subunit, named after the major contributors, the Nomura map and the Nierhaus map, respectively (Figure 1.12 A and B) (Mizushima and Nomura 1970; Held, Mizushima *et al*. 1973; Held, Ballou *et al*. 1974; Herold and Nierhaus 1987; Grondek and Culver 2004; Shajani, Sykes *et al*. 2011). The two maps suggest that the assembly of the subunits are modulated in both cases based on the ribosomal RNA sequences. Possibly due to the morphological divisions of the small subunit, there are only 4 arrows showing interplay of ribosomal protein binding events
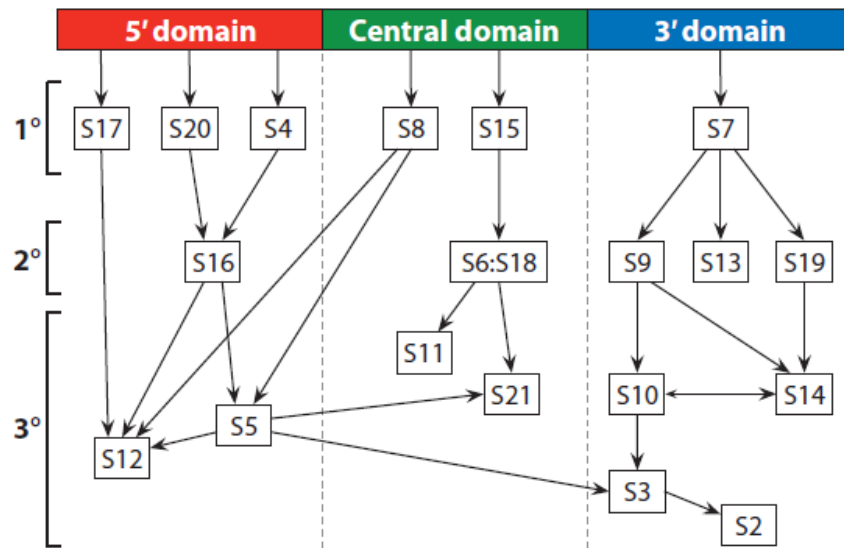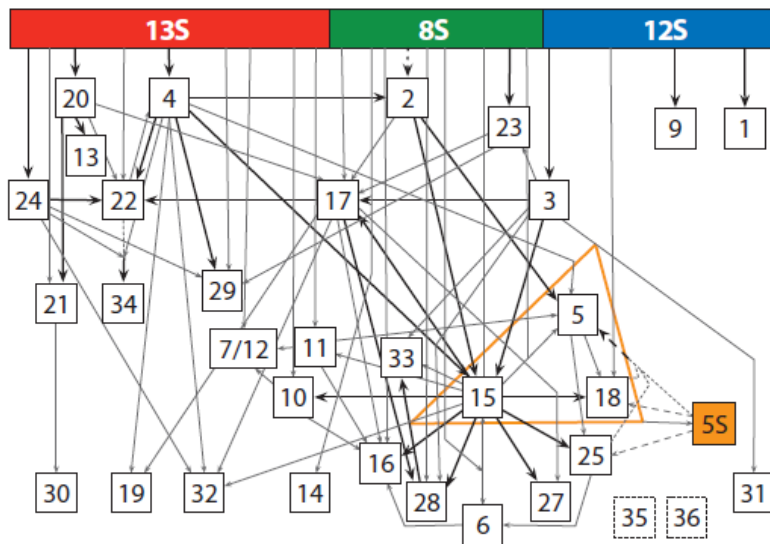
Figure 1.12 Assembly maps of the bacterial ribosomal subunits. A. Assembly map of the 30S subunit. B. Assembly map of the 50S subunit. The Ribosomal RNA sequences are divided into three segments: 5' domain, central domain, and 3' domain for the 16S rRNA and 13S, 8S, and 12S for the 23S rRNA (Zaher, Shaw *et al*. 2011).

across 16S rRNA domains. It is completely different in the case of 50S subunit assembly map, where almost all the binding events are interconnected with each other. As a consequence, usually much lower activity is observed in reconstituted large subunits compared to that of the reconstituted small subunits (Khaitovich, Tenson *et al*. 1999). The possible explanation for this is that complexity of the ribosomal protein binding network decreases the reconstitution efficiency dramatically, e.g. more population of the reconstituted particles are trapped in some sub-optimal inactive conditions. Those results may also indicate the participation of other factors in vivo that aid proper assembly of the large subunit.

RNases, RNA modification enzymes, and assembly factors also play important roles in the biogenesis of bacterial ribosomes. The primary transcript including all three ribosomal RNAs and tRNAs is processed by RNase III, a double-stranded RNA specific RNase, to yield ribosomal RNAs with extra nucleotides on both termini (Ginsburg and Steitz 1975). The precursor of the 16S rRNA is further processed by RNase E and RNase G on the 5' terminus sequentially and an unknown RNase on the 3' terminus (Li, Pandit *et al*. 1999). RNase T removes the extra nucleotides on the 3' termini of both precursors of the 23S rRNA and 5S rRNA, while no RNase is identified to process the 5' termini of the two precursors (Li, Pandit *et al*. 1999). Assembly factors, e.g. DEAD box proteins, are essential for ribosome biogenesis in eukaryotic cells. For bacterial cells growing under normal conditions, deletion of the assembly factor genes do not show significant growth phenotype, while the bacterial cell growth becomes temperature sensitive. It is believed that binding of assembly factors can substantially change the subunit assembly energy landscape, and help subunit maturation at unfavorable

temperatures (Jones, Mitta *et al*. 1996; El Hage, Sbai *et al*. 2001; Charollais, Pflieger *et al*. 2003; Inoue, Alsina *et al*. 2003; Lovgren, Bylund *et al*. 2004; Nord, Bylund *et al*. 2009). RNA modification enzymes are also essential for the biogenesis and activity of ribosomes, and this issue will be reviewed later.

1.4 Translation process

A complete translation process includes initiation, elongation, termination, and ribosome recycling stages.

In bacteria, an initiation complex is formed with mRNA, fMet-tRNA^{fMet}, and initiation factor (IF) 1, 2, and 3 binding onto the 30S subunit. The fMet-tRNA^{fMet}-IF2-GTP ternary complex binds into the P-site of 30S subunit, positioned by the base pairing between the tRNA anticodon and mRNA start codon. IF2 has GTPase activity, while this activity is essential for neither the ternary complex binding, nor IF2 release (Tomsic, Vitali *et al*. 2000). IF1 binds into the A-site on the 30S subunit to block binding of tRNA into this site, and it may help stabilize the binding of fMet-tRNA^{fMet}-IF2-GTP complex by interacting with IF2 (Zucker and Hershey 1986; Carter, Clemons *et al*. 2001). IF3 plays dual functions too, where it prevents 50S subunit from being associated to the 30S subunit, and scans the anticodon stem loop of the tRNA to reject non-cognate tRNA binding (Hartz, Binkley *et al*. 1990). Finally when the 50S subunit binds to the 30S subunit, all the initiation factors are removed from the complex, and only the mRNA and initiator tRNA (P-site) are sequestered between the two subunits.

The elongation stage starts with tRNA-EF-Tu-GTP ternary complex binding into the A-site of the ribosome (Figure 1.13 A) (Schmeing, Voorhees *et al*. 2009). With cognate codon-anticodon formed in the decoding center, multiple hydrogen bonds are

A

B

C
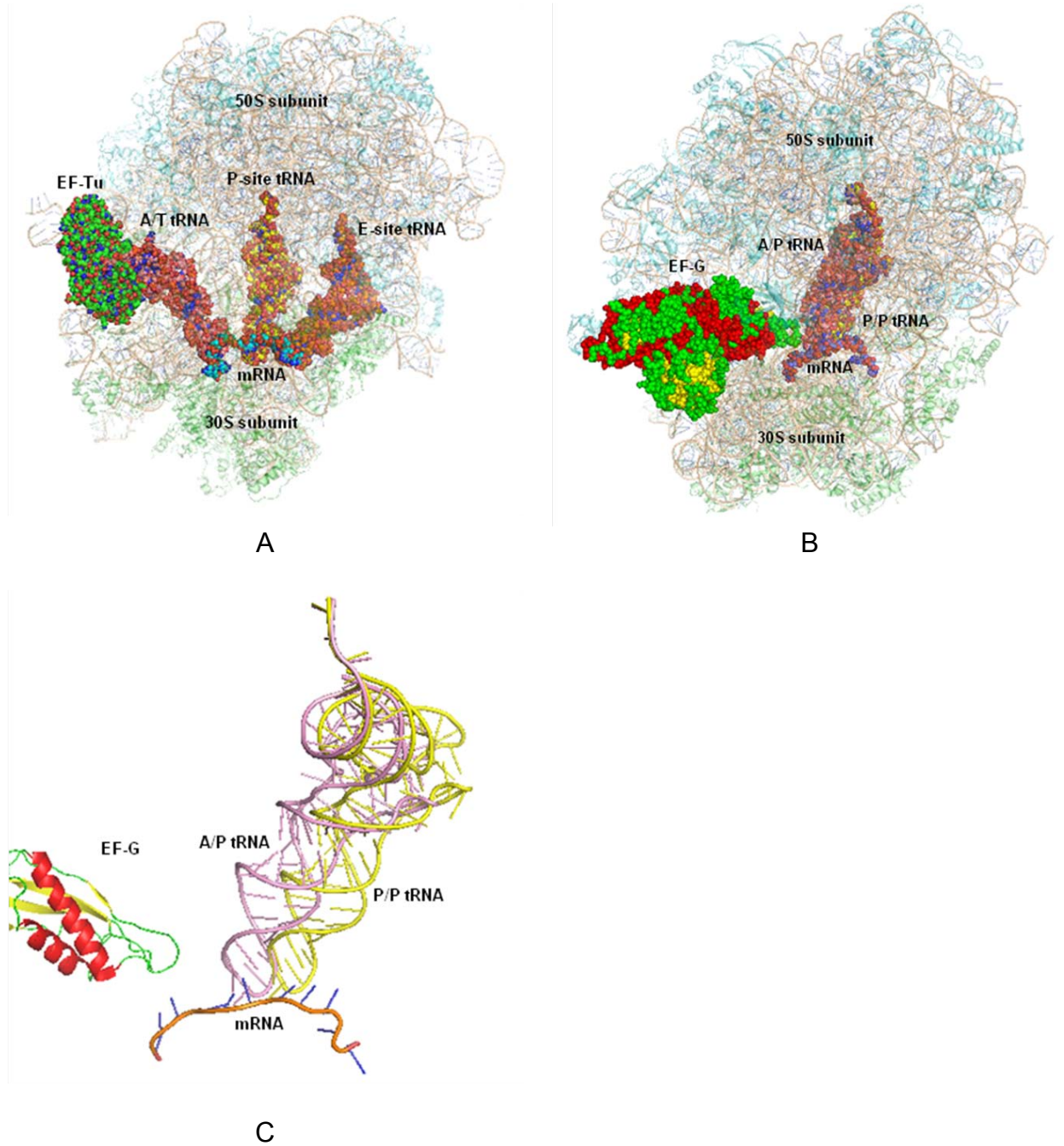
Figure 1.13 Elongation stage of the translation process. A. tRNA-EF-Tu-GTP tertiary complex binds into ribosomal A-site. The picture is made from PDB 2WRN and 2WRO (Schmeing, Voorhees *et al*. 2009). EF-TU binds to the acceptor stem of the A-site tRNA,

and the anticodon stem loop of the tRNA base-pairs with the codon on the mRNA. B and C are made from PDB 2XSY, 2XTG, 2XUX, and 2XUY (Ratje, Loerke *et al*. 2010). B. Comparison of the positions of pre-translocation (A/P) and post-translocation (P/P) of the tRNA in the context of 70S ribosome. Only one set of the ribosomal subunits, Ef-G, and mRNA are shown for clarity. The A/P tRNA has the carbons colored pink, and the P/P tRNA has the carbons colored yellow. C. Close view of the positions of pre-translocation (A/P) and post-translocation (P/P) of the tRNA. The receptor stems of the two tRNA structure overlap, which means the receptor stem does not move in this step, while the anticodon stem loop migrates by approximately one codon distance along the mRNA direction. Note that mRNA co-migrates with the tRNA anticodon stem loop.

established between the codon-anticodon minihelix and three conserved bases on 16S rRNA, G530, A1492, and A1493 (Figure 1.11 E, F, and G). The flipped-out conformation of the A1492 and A1493 is essential for the processing of the downstream events, since aminoglycosides that facilitate A1492 and A1493 moving into the flipped-out conformation constitute misreading (Figure 1.11 H) (Borovinskaya, Pai *et al*. 2007). Once the tRNA binding is tested by this decoding mechanism, the GTPase activity of the EF-Tu is activated, and the receptor stem of the A-site tRNA is released from the EF-Tu-GDP binary complex, and enters the peptidyltransferase center (Figure 1.11 A) (Jenner, Demeshkina *et al*. 2010). The peptide bond formation reaction happens instantly, and the peptide chain is transferred from the P-site tRNA onto the A/P tRNA, and the acceptor stem of the P-site tRNA moves to the E-site in the 50S subunit (Moazed and Noller 1989). EF-G is employed at this stage to facilitate the concomitant movement of mRNA, A/P tRNA, and P/E tRNA with hydrolysis of the GTP bond in EF-G (Rodnina, Savelsbergh *et al*. 1997). After translocation, the two tRNAs are in P/P and E/E positions, and the A site is available for the binding of new tRNA (Figure 1.13 B and C) (Ratje, Loerke *et al*. 2010).

When a stop codon, i.e. UAA, UAG, and UGA, from the mRNA moves into the A-site of the ribosome, release factors are recruited into the ribosome by specific recognition of the stop codons and the synthesized polypeptide is released from the tRNA by hydrolysis. In bacteria, there are three release factors, named RF1, RF2, and RF3 participating translational termination. RF1 recognizes UAA and UGA stop codons, RF2 recognizes UAA and UAG stop codons, and RF3 catalyzes the release of RF1 and RF2 by its inherent GTPase activity (Scolnick, Tompkins *et al*. 1968; Freistroffer, Pavlov

et al. 1997). Specific recognition of the RF1 to UGA stop codon and RF2 to UAG stop codon is determined by the different hydrogen bond patterns (Figure 1.14 A and B) (Weixlbaumer, Jin *et al*. 2008; Korostelev, Zhu *et al*. 2010). RF1 and RF2 are also involved in catalysis of the peptide release reaction. A GGQ tripeptide located in a coil region above helix α7 of the RFs is highly conserved throughout all three kingdoms, and mutation of this tripeptide severely compromises the peptide release rate (Frolova, Tsivkovskii *et al*. 1999; Mora, Heurgue-Hamard *et al*. 2003). In crystal structures, Gln240 (numbering according to *T. Thermophilus* RF2) is positioned close to A76 sugar hydroxyl groups of the P-site tRNA, and the distance between the amide oxygen of Gln240 and 2' hydroxyl oxygen of A76 is 4.1 Å, indicating a potential hydrogen bond mediated by a water molecule (Figure 1.14 C) (Weixlbaumer, Jin *et al*. 2008). This water molecule could be used for hydrolysis and release the peptide from tRNA (Figure 1.14 D).

After the peptide is released, P-site tRNA and mRNA are still trapped between the two ribosomal subunits. To reuse the ribosomal subunits, ribosome recycling factor (RRF) and EF-G-ATP are employed to dissociate the 70S ribosome, and release the tRNA and mRNA. Both the RRF and EF-G-GTP are essential for ribosome recycling (Hirashima and Kaji 1972). It is shown in the crystal structure that, when the RRF binds into the ribosome, H69 of the 23S rRNA moves toward the RRF by about 8Å, resulting in the disruption of B2a intersubunit bridge, while with aminoglycoside antibiotics, e.g. neomycin, gentamicin, and paromomycin bind to H69, this movement is hindered, leading to interrupted ribosome recycling (Borovinskaya, Pai *et al*. 2007). The small

A

B

C

D
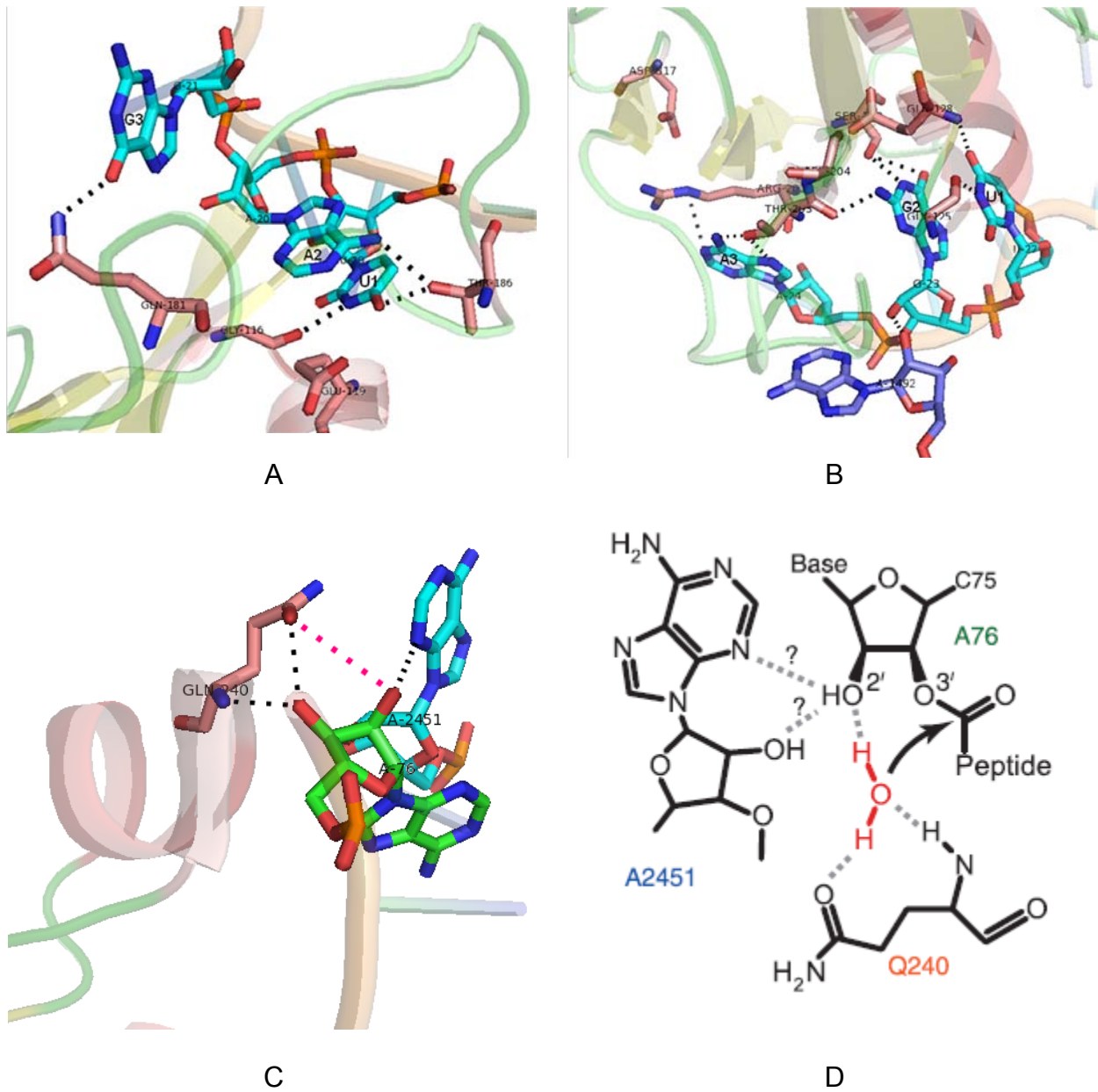
Figure 1.14 Translation termination and ribosome recycling. Hydrogen bonds are shown as black dashes. A is made from PDB 3MS0 (Korostelev, Zhu *et al*. 2010). B and C are made from PDB 2WH1 and 2WH2 (Weixlbaumer, Jin *et al*. 2008). A. *Thermus thermophilus* Release factor (RF1) specifically recognizes UAG stop codon. The hydrogen bonds are shown in black dashes. Amino acid residues Gly116, Glu119, Thr186, and Gln181 of RF1 are involved in hydrogen bond formation. B. *T. thermophilus*

RF2 specifically recognizes UGA stop codon. Amino acid residues Gly125, Glu128, Ser 193, Thr203, and Thr204 of RF2 are involved in hydrogen bond formation. A1492 2'-hydroxl group stabilizes the conformation of G2 in the stop codon through a hydrogen bond with G2 2'-hydroxl group. C. In the peptidyltransferase center of *T. thermophilus* risosome with RF2 and P-site tRNA bound, the pink line connects the side chain amide oxygen of RF2 Gln240 and the 2' hydroxyl oxygen of tRNA A76, with a distance of approximately 4.1 Å. D. Potential catalysis mechanism of peptide release, where the substrate water molecule is hydrogen bonded to RF2 Gln240 and 2' hydroxyl oxygen of tRNA A76. A2451 of the 23S rRNA may play a role in the catalysis by donating it N3 and 2' hydroxyl oxygen as hydrogen bond donors.

ribosomal protein S12 may participate in the interaction between RRF and the ribosome, but no research has been done on the effect of S12 on ribosome recycling.

1.5 Ribosomal protein S20 in bacterial small ribosomal subunit

As shown in Figure 1.25 A, ribosomal protein S20 is a primary binding protein of the bacterial small ribosomal subunit, and the amino acid sequence of S20 is highly conserved from eubacteria to chloroplasts of plants (Marchler-Bauer, Lu *et al*. 2011). *In vitro* reconstitution experiments indicate that S20 works synergistically with S4 in promoting the binding of S16 onto the 16S rRNA (Figure 1.15 A) (Held, Ballou *et al*. 1974).

Bacterial cells lacking S20 show lowered translation activity and slow growth rate phenotype due to the impaired translation initiation (Tobin, Mandava *et al*. 2010). In prokaryotic cell translation process, the initiation complex starts to form with binding of mRNA, fMet-tRNA[fMet], Initiation Factors 1, 2, and 3, and the 30S subunit. With depletion of the S20 protein in the 30S subunit, the rate of mRNA binding onto the 30S subunit is significantly decreased, and possibly as a secondary effect, binding of the fMet-tRNA[fMet] onto the 30S subunit is also dramatically decreased (Gotz, Dabbs *et al*. 1990; Tobin, Mandava *et al*. 2010). It is also observed that the downstream event, association of the 70S ribosome is compromised, even with prolonged preincubation of S20[-] 30S subunit with initiation factors, mRNA, and fMet-tRNA[fMet] before addition of the 50S subunit. All these negative effects could be potentially abolished with addition of S20 protein into the association mixture. The methylation of 16S rRNA is altered with absence of the S20 in the small subunit, which may also contribute to the lowered activity of the mutant small subunit (Ryden-Aulin, Shaoping *et al*. 1993).

It is impossible to directly correlate the function of S20 to mRNA binding, since the distance between the mRNA and S20 protein is at least 100 Å. It is speculated that, absence of the S20 in the small subunit could change the folding of h44 (the dominant component of the 3' minor domain), which runs across the whole "body" of the small subunit and conducts the effect to the mRNA and tRNA binding site (Tobin, Mandava *et al*. 2010). By similar mechanism, multiple intersubunit bridges involving h44 are affected by the absence of S20, and lower the binding affinity between the two subunits. An alternative possible explanation is that, the effect of S20 depletion is diffused through binding of the secondary and tertiary ribosomal proteins and related conformational and modification pattern change of the 16S rRNA, leading to lowered rate of mRNA and tRNA binding on to the small subunit, and impaired association of the 70S ribosome (Figure 1.15 B). However, there is not enough experimental result to support this hypothesis.

1.6 H69 of the 23S rRNA

H69 is a stem loop segment on 23S rRNA ranging from G1906 to C1924 in bacteria (Figure 1.16). The sequence and secondary structure of this stem-loop structure are highly conserved in bacteria. Compared with bacterial sequence of H69, covariation of C1908-G1922 and C1909-G1921 (*E. coli* numbering) base pairs in archaea and eukaryotes ensures the conservation of the Watson-Crick base pairs in these two places (Figure 1.16) (Cannone, Subramanian *et al*. 2002). Another significant difference in the H69 sequences between bacteria and archaea/eukaryotes is the nucleotide identity of position 1918, where an A is highly conserved in bacteria, and a G is highly conserved in archaea and eukaryotes, though no clear phenotype has been

A   B
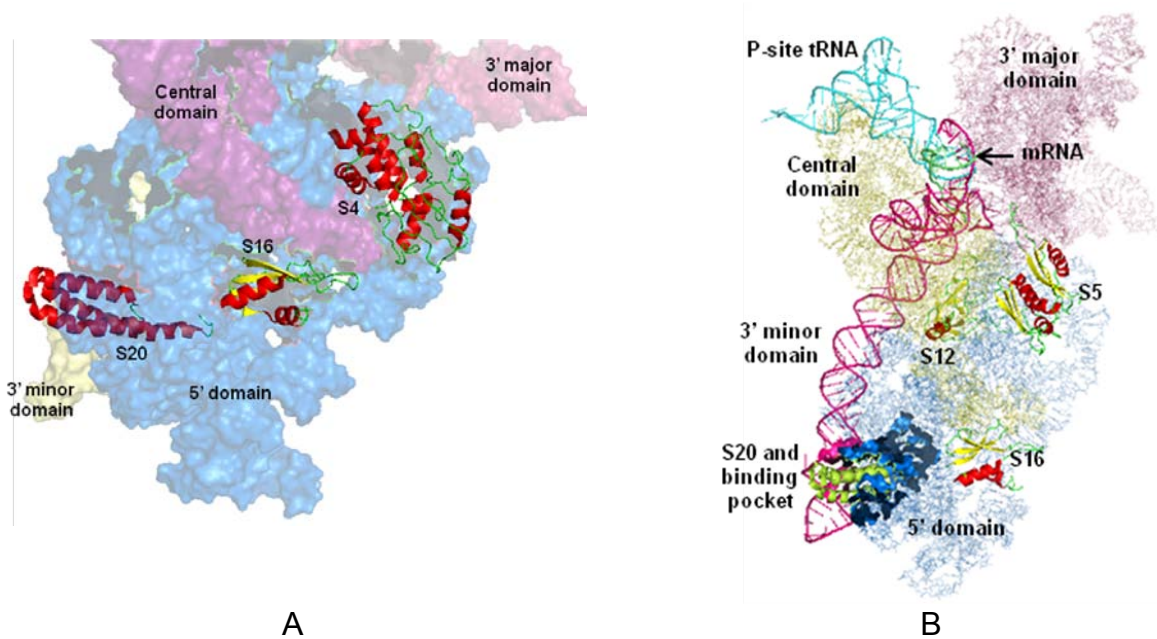
Figure 1.15 Binding of S20 in the *E. coli* small ribosomal subunit. A. Binding sites of ribosomal proteins S4, S16, and S20 in bacterial small ribosomal subunit. The 16S rRNA sequence is colored to show the 5' domain (blue), central domain (Purple), 3' major domain (pink), and the 3' minor domain (yellow). S4 and S16 bind on the surface of 16S rRNA 5' domain, close to the central domain, while S20 is inserted between the 5' domain and the 3' minor domain. B. Locations of S20, S16, S5, S12, mRNA, and P-site tRNA in *E. coli* small ribosomal subunit. S20 is shown in green cartoon, and all other ribosomal proteins are shown in carton with color scheme to show the secondary structures. 3' minor domain of the 16S rRNA, P-site tRNA, and the mRNA are shown in cartoon with pink, cyan, and green color, respectively. All other domains of the 16S rRNA are shown as transparent sticks.
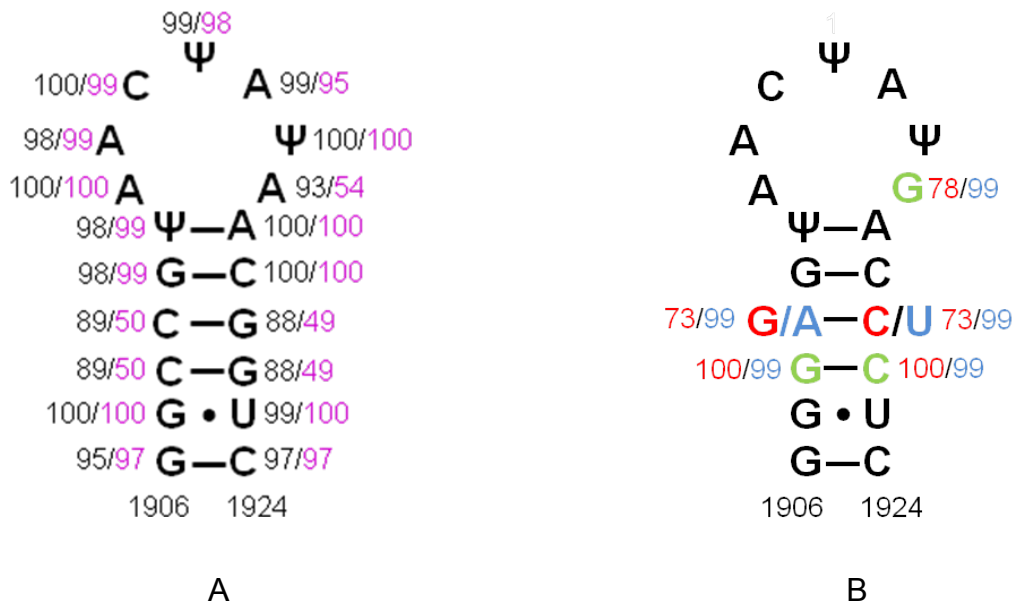
Figure 1.16 Sequence and secondary structure conservation of the large ribosomal subunit RNA H69. A. The sequence of bacterial H69, starting from G1906 to C1924. Numbers in black show the conservation of the nucleotide residues in bacteria, and the numbers in purple show the conservation of the nucleotide residues throughout three domains. *E. coli* has pseudouridine modifications at positions 1911, 1915, and 1917 positions, while other bacteria have different modification profile. The numbers for position 1911, 1915, and 1917 reflect the level of conservation of U and Ψ combined. B. The sequence of H69 from archaea and eukaryotes (*E. coli* numbering system is used for simplicity). Letters in red, blue and green show nucleotide residues in archaea, eukaryotes, and shared variations, respectively. Numbers in yellow and blue show the conservation of the residues in archaea and eukaryotes, respectively. It is clearly shown in the secondary structures that the base pairing patterns are highly conserved, even though there are differences in the nucleotide residue identities.

identified for A1918G mutant in bacteria (Ali, Lancaster *et al*. 2006).

High conservation indicates an essential functional role. Residing on the subunit interface in the 70S ribosome, H69 participates in multiple interactions during the translation process (Figure 1.17 A-I). H69 forms the essential B2a intersubunit bridge with A-site of 16S rRNA, and deletion of the H69 sequence causes complete loss of *in vitro* 70S ribosome formation with 30S and 50S subunits and   dissociation of the 70S ribosome without binding of the Ribosomal Recycling Factor (RRF) (Ali, Lancaster *et al*. 2006). It has been shown that mutations A1912G and A1919G could slow down the IFs-dependent ribosome association, and decrease the processivity of translation, leading to lowered peptide synthesis capability *in vitro* and slow-growth phenotype *in vivo* (Liiv, Karitkina *et al*. 2005; Kipper, Hetenyi *et al*. 2009). Stem residue G1922 is critical for peptide release catalyzed by RF2 recognizing the UGA stop codon, and the single mutation of G1922A, which is able to decrease the rate constant by a factor of 12 (Ortiz-Meoz and Green 2011).

A distinguishing feature of this stem-loop structure is the clustered nucleotide modifications. In *E. coli*, pseudouridylation is identified at positions 1911, 1915, and 1917, with an additional methylation on N3 of Ψ1915. A pseudouridine synthase, RluD, and a methyl transferase, RlmH, are responsible for introducing these modifications into H69. The 50S subunit and 70S ribosome are the best substrates for RluD and RlmH, respectively (Raychaudhuri, Conrad *et al*. 1998; Leppik, Peil *et al*. 2007; Ero, Peil *et al*. 2008). Pseudouridylation is highly conserved at all three positions, with some exceptions where the 1911 or 1915 pseudouridylation is missing (Ofengand and Bakin 1997). Mutations of Ψ1915C and Ψ1917C confer severely hindered 70S ribosome

A

B

C

D

E

F

G

H

I
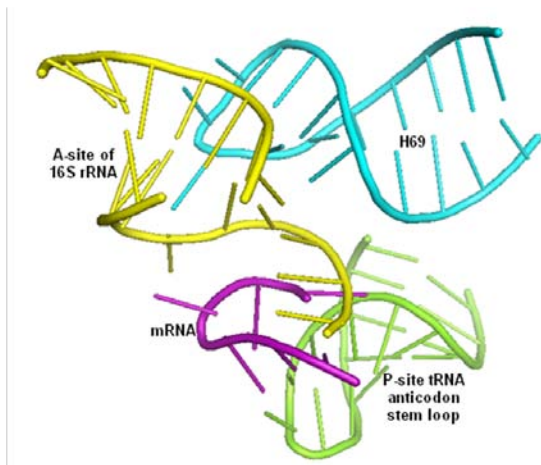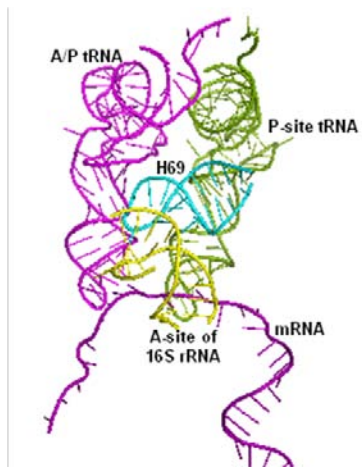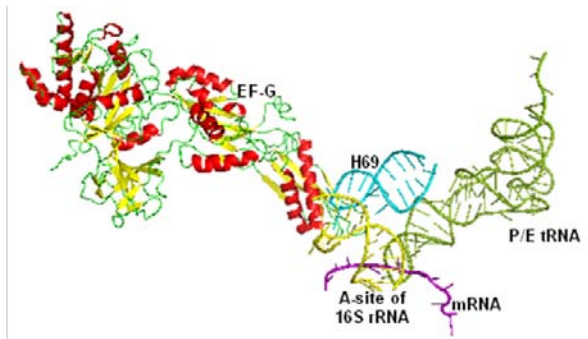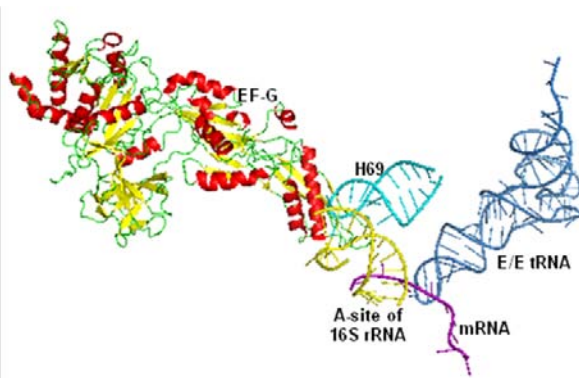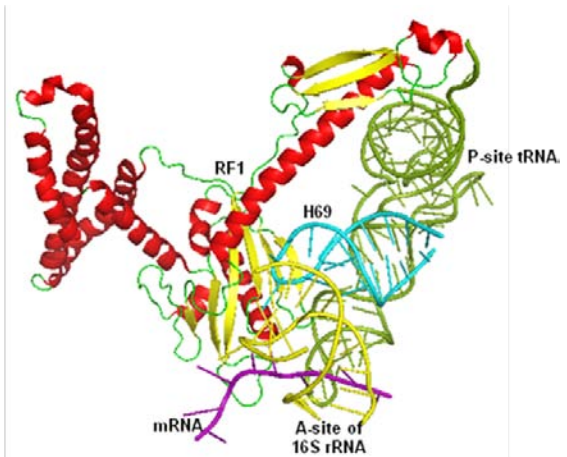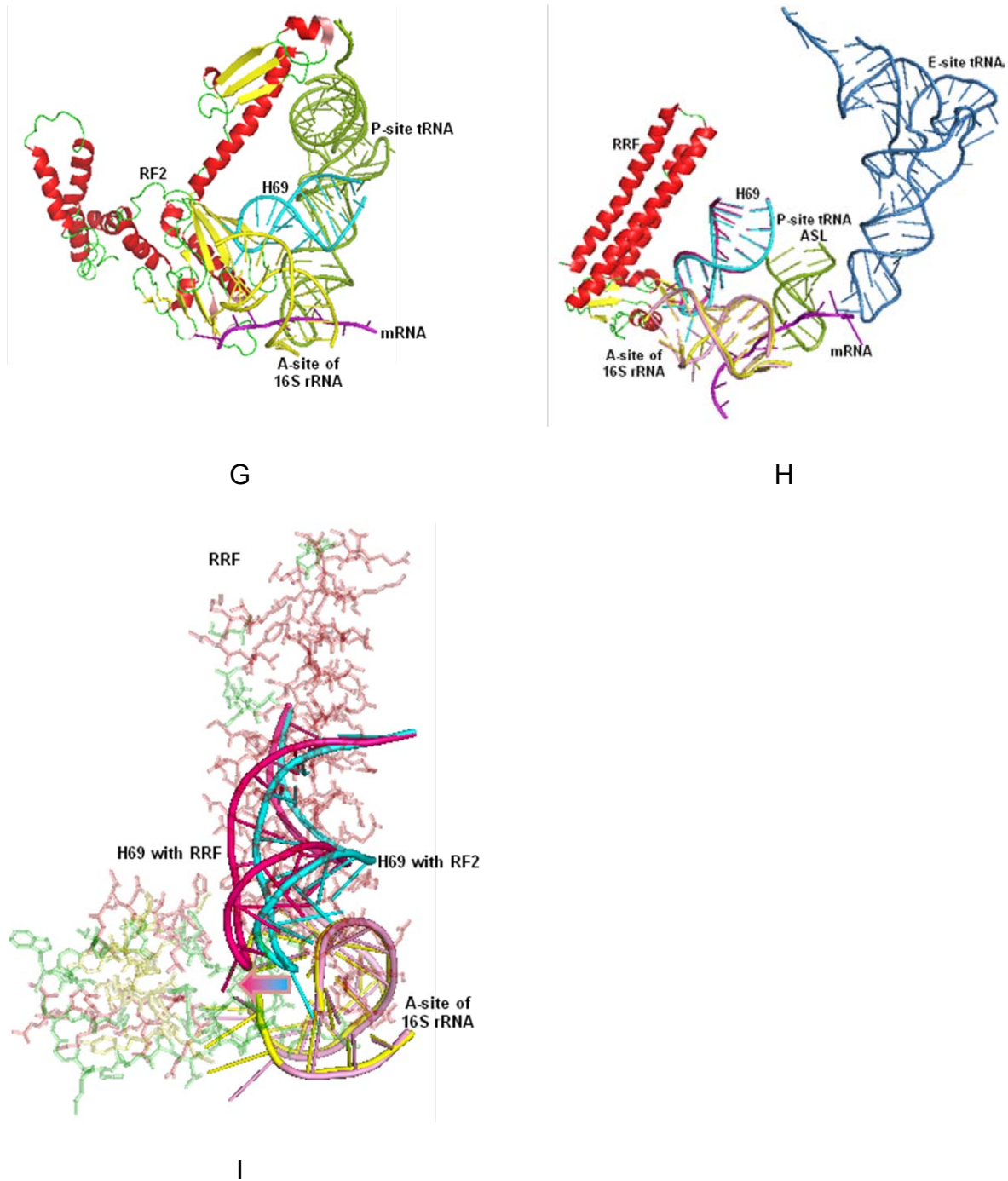
Figure 1.17 H69 participates in all the events during the translation process. A. In a post-initiation complex, H69 forms interaction with the A-site of 16S rRNA, which constitutes the B2a intersubunit bridge (2I2P and 2I2T) (Berk, Zhang *et al*. 2006). B and C show H69 interactions with the A-site of 16S rRNA and the newly recruited tRNA in

pre- and post- accommodation steps (2WRN, 2WR0, 3I8G, and 3I8F) (Schmeing, Voorhees *et al*. 2009; Jenner, Demeshkina *et al*. 2010). The loop of H69 is directly interacting with the A-site of 16S rRNA and the newly-recruited tRNA, and the stem of H69 is interacting with the P-site tRNA. D and E show H69 interactions with the A-site of 16S rRNA and EF-G in pre- and post-translocation steps (2XSY, 2XTG, 2XUY, and 2XUX) (Ratje, Loerke *et al*. 2010). During the translocation stage, the B2a intersubunit bridge was held intact by interactions between the H69 and the A-site of 16S rRNA. H69 is potentially able to form direct interactions with EF-G. F and G show H69 direct interactions with the A-site of 16S rRNA, P-site tRNA, and corresponding RFs (3MS0, 3MRZ, 2WH1, and 2WH2) (Weixlbaumer, Jin *et al*. 2008; Korostelev, Zhu *et al*. 2010). H and I show that H69 moves away (an arrow shows the direction of movement in I) from the A-site of 16S rRNA towards the RRF for about 5 Å in 70S ribosome with RRF bound (pink cartoon), during the breakage of the intersubunit bridge, compared to that when RF2 bound (cyan cartoon), when the B2a bridge is still intact (2WH1, 2WH2, 2V46, and 2V47) (Weixlbaumer, Petry *et al*. 2007; Weixlbaumer, Jin *et al*. 2008). H is the front view and I is the side view seeing through the RRF (transparent "sticks"). The back bond contours of the A-site of 16S rRNA from 2WH1 (RF2) and 2V46 (RRF) are aligned well. The two H69 structures are aligned with the G1906 phosphate, G1907 phosphate, C1924 phosphate, and C1924 C3' hydroxyl oxygen atom coordinates.

association, compromised elongation efficiency *in vivo*, and slow-growth phenotype (Liiv, Karitkina *et al*. 2005). Depletion of the pseudouridylations in H69 resulting from deletion of RluD affects translation termination in *E. coli* K12 derivative strains, and read-through of the UGA stop codon is the mostly influenced (Ejby, Sorensen *et al*. 2007). A temperature-sensitive slow-growth phenotype was also observed in yeast when three loop pseudouridylation modifications were removed (Liang, Liu *et al*. 2007). Biophysical studies on H69 model systems show pseudouridylation affects the pH sensitivity of the stem loop folding (Abeysirigunawardena and Chow 2008). A compact conformation of the loop region with full modification at lower pH was observed in NMR experiments on H69 oligonucleotides, and an apparent pKa was determined to be 6.3 by Circular Dichroism experiments (Abeysirigunawardena and Chow 2008). *In vitro* chemical probing experiments on purified ribosomes show that A1913 is more protected in fully-modified H69 loop region at lower pH (Sakakibara and Chow 2011).

1.7 Project rationale

Experimental work included in this project is designed to elucidate the chemical and structural mechanisms of "fine-tuning" ribosome functions via ribosomal RNA modifications and interactions between ribosomal RNAs and ribosomal proteins. Knowledge obtained in this research is helpful for designing antibiotics targeting bacterial ribosomes with improved efficacy and specificity in treatment of bacterial infection diseases.

First, the effect of pseudouridylation modification on thermodynamic properties and 3D structure of 23S rRNA H69 was studied. As previously discussed, the sequence of H69 is highly conserved and the function of this 23S rRNA segment is essential for

normal ribosomal activity and cell viability. In a long-term collaboration with Dr. Christine Chow's lab, global conformation, thermodynamics, and ligand binding of H69 have been studied with H69 oligonucleotides with and without post-transcriptional modifications (Chow, Ng *et al*. 1999; Meroueh, Grohar *et al*. 2000; Chui, Desaulniers *et al*. 2002; Desaulniers, Chui *et al*. 2005; Sumita, Desaulniers *et al*. 2005; Abeysirigunawardena and Chow 2008; Desaulniers, Chang *et al*. 2008; Duc 2009). Furthermore, the effects of psudouridylation on H69 3D folding was also studied at whole ribosome level with chemical probing (Sakakibara and Chow 2011). All observations show that even though no net global stabilizing effect is gained with pseudouridylation, the modification does affect the stem-loop structure and biophysical properties. This raises questions to be addressed: 1) What is structural basis for thermodynamic properties of the H69 oligonucleotide; 2) What is the difference in 3D structure of the H69 oligonucleotide with and without pseudouridylation modifications. The first question is addressed by UV-melting experiments and the second question by Nuclear Magnetic Resonance experiments.

Secondly, the species-specificity of binding of h9 RNAs and S20 proteins from *E. coli* and *Pseudomonas aeruginosa* was investigated with biophysical methods at the molecular level. H9 is a segment of RNA located in the 16S rRNA 5' domain, right next to ribosomal protein S20. Research in Dr. Philip Cunningham's lab shows that, incorporation of non-cognate S20 protein into ribosomes could potentially compromise the biogenesis and activity of ribosomes, which would be reversed by introduction of the corresponding h9 sequence into the 16S rRNA (Lamichhane 2009). With a plethora of biological research data at hand, the molecular mechanism determining the species-

specificity of h9-S20 interaction is still a mystery. As part of this project, binding affinities of cognate and non-cognate pairs of h9 RNA and S20 protein were determined by Surface Plasmon Resonance experiments, and Circular Dichroism experiments were used to elucidate the conformational change of S20 protein after binding with cognate and non-cognate h9 RNA oligonucleotides.

CHAPTER 2

INSTRUMENTATION

2.1 Preparation of RNA samples

2.1.1 T7 RNA polymerase in vitro transcription

In vitro run-off transcription from a synthesized DNA template (Integrated DNA technologies, IDT®) with T7 RNA polymerase was employed to prepare unmodified H69 oligonucleotides for NMR study of H69 3D structure and h9 oligonucleotides from *E. coli* and *P. aeruginosa* and for Circular Dichroism (CD) experiments on h9-S20 binding (Figure 2.1) (Milligan, Groebe *et al.* 1987; Wyatt, Chastain *et al.* 1991). Natural isotope abundance and ^{13}C, ^{15}N doubly labeled NTPs were purchased from Sigma-Aldrich®. T7 RNA polymerase with a his^6 tag was prepared with induced over-expression in *E. coli* host cells and Ni^{2+} affinity chromatography. The concentration of T7 RNA polymerase in the final transcription mixture was optimized for every batch. In the transcription mixture, there are 40 mM Tris-HCl (pH 8.1), 32 mM MgCl$_2$, 4 mM of each NTP, 0.4 µM of template DNA, 0.4 µM of promoter DNA, 2 mM spermidine, 10 mM DTT, 80 mg/mL PEG8000, 0.02% (v/v) Triton X-100, and T7 RNA polymerase. This reaction mixture was incubated in water bath at 37 ºC for 4 hours, then the magnesium pyrophosphate precipitate was removed by centrifugation. After ethanol precipitation and centrifugation, the pellet, which contains the transcribed RNA, was dried briefly and re-dissolved in loading buffer containing 9M urea and 100mM EDTA. Full length (N) RNA transcript was resolved from other byproducts, including N+1 transcript, by overnight denaturing polyacrylamide gel
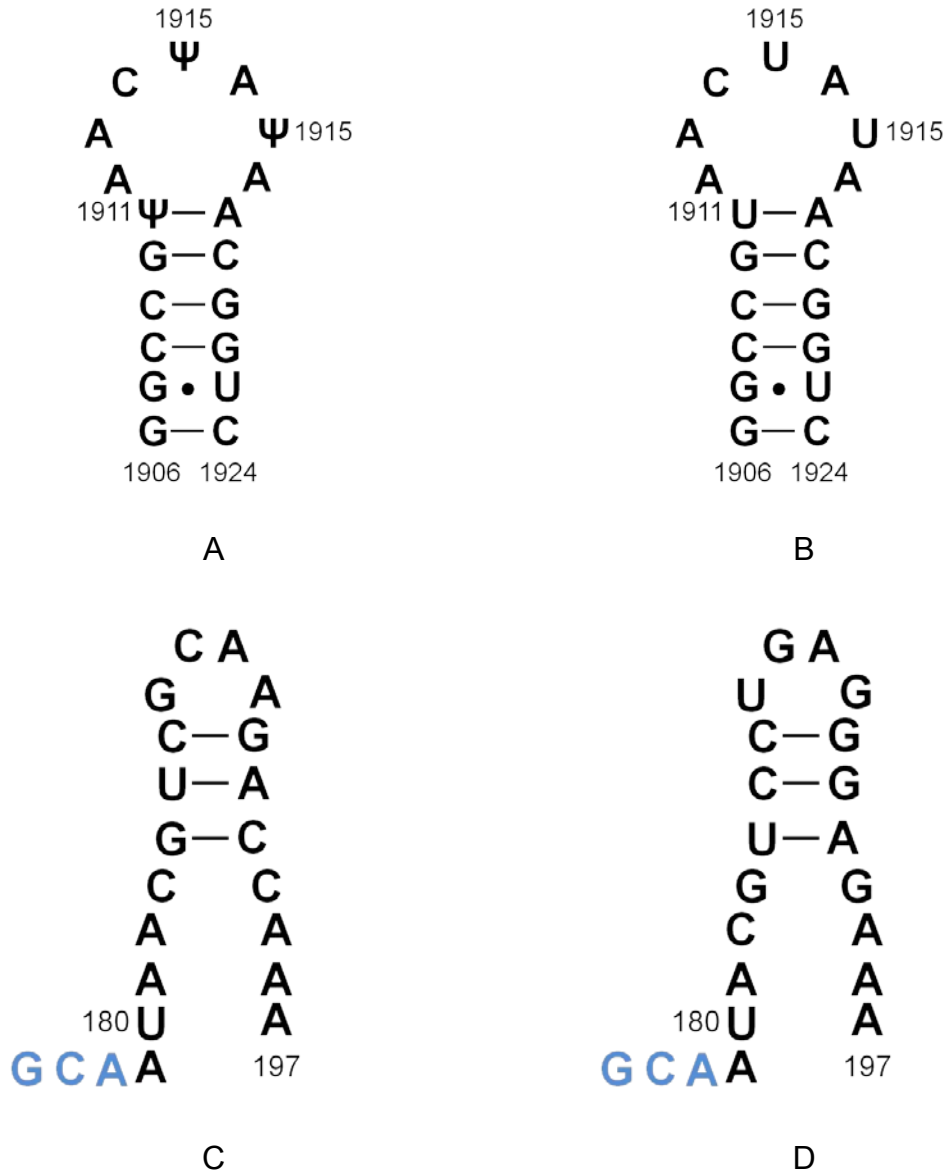
Figure 2.1 Secondary structures of RNA oligonucleotides used in this project. A and B show H69 with and without pseudouridylation at 1911, 1915, and 1917. C and D show h9 from *E. coli* and *P. aeruginosa*, respectively. Blue high-lighted residues are designed to minimize homodimer formation without interrupting the protein contact interface. Construct A was ordered from Dharmacon®, and the rest were prepared by T7 polymerase *in vitro* transcription.

(20%, w/v) electrophoresis at 750V with 1× TBE buffer. RNA bands were visualized with UV shadowing against a fluorescent TLC plate. A Schleicher and Schuell® Elutrap was used to recover the RNA in gel stripes cut out from the polyacrylamide gel. To minimize salt and urea concentrations in the collected RNA-containing fractions, 0.2× TBE buffer was used outside of the electrophoresis chamber, and deionized H_2O was put in the chamber to merge the gel strips. This step takes about 2 hours at 230V. Eluted fractions were combined and loaded onto a Sep-pak® reverse phase chromatography cartridge, washed with 21 mL deionized H_2O, and eluted with 1.5 mL 30% ACN/H2O (v/v). RNA-containing fractions were combined, and dried into white powders in vacuum. The molecular mass of the RNA oligonucleotide was determined by MALDI-tof mass spectrometry, and compared to the calculated molecular mass obtained from www.ozone3.wayne.edu/Hyther for RNA identity confirmation.

2.1.2 HPLC purification of RNA oligonucleotides

Though powerful, T7 RNA polymerase *in vitro* run-off transcription has its limitations. First, T7 RNA polymerase is not able to carry out site-specific incorporation of pseudouridine into the RNA transcript when uridine is also in the sequence and reaction mixture (Figure 2.1 A). Second, T7 RNA polymerase is inefficient in synthesizing very short oligonucleotides, and prefers certain sequences at the 5' terminus (Figure 2.2) (Milligan, Groebe *et al.* 1987). In such cases, solid phase synthesis of RNA oligonucleotides has to be employed, followed by HPLC purification. The oligonucleotides mentioned above were all purchased from Dharmacon®.

A   B
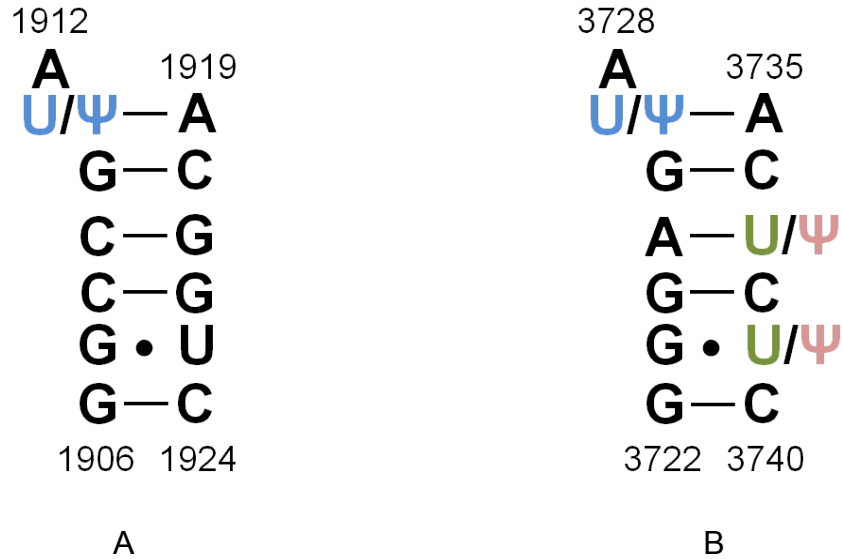
Figure 2.2 RNA oligonucleotide sequences used in UV-melting experiments. A. Secondary structure of the *E. coli* H69 stem duplex sequences. Two sequences, different in the modification at 1911 position (blue letters), correspond to the left strand, so there are two possible combinations for the duplex formation. B. Secondary structure of the *Homo sapiens* H69 stem duplex sequences. Two sequences, different in the modification at 3727 position, correspond to the left strand (blue letters), and two sequences, different in the simultaneous modifications at 3737 and 3739 positions (green and pink), correspond to the right strand (blue letters). So there are four possible duplex combinations.

The modified H69 stem-loop oligonucleotide was purified by HPLC on an Xterra MS C18 column (Waters) with a 6-9% linear gradient of acetonitrile (ACN) over 30 min in 25mM Triethylammonium acetate (pH 6.5) as the mobile phase. After the RNA-containing fractions were dried down, the pellet was re-dissolved in deionized H_2O and desalted with Sep-pak® cartridge (See 2.1.1). Single strands of the *E. coli* and *H. sapiens* H69 stems were purified with the same protocol, with two changes: 1) The ACN gradient used in HPLC purification was 4-7%, and 2) The ACN concentration used to elute RNA from the Sep-pak® cartridge was 10%. Molecular masses of the RNA oligonucleotides were determined by MALDI-tof mass spectrometry, and compared to the molecular masses in the delivery sheet provided by the manufacturer for RNA identity confirmation.

2.2 UV-melting

Hypochromicity is the basic mechanism enabling RNA UV-melting experiment (Puglisi and Tinoco 1989; Breslauer 1995; SantaLucia and Turner 1997). Bases always form stacking in the folded structure, e.g. in the stem region, or sometimes in the loop, of the hairpin structure. The per base absorbance of a stacked base is smaller than that of a base in random coil state. This is termed as hypochromicity (Doty, Boedtker *et al.* 1959; Tinoco 1959; Devoe and Tinoco 1962). If an RNA sample is heated, the UV absorbance of the sample will increase along with the temperature, and when there is half of the population in random coil state, the temperature is called melting temperature (T_M). The population of the folded and random-coil state molecules, hence the

equilibrium constant at that temperature, can be determined from the difference

between the UV absorbance at that temperature, and the calculated absorbance

of the folded and random-coil state by linear extrapolation (Figure 2.3)

(Petersheim and Turner 1983). A van't Hoff analysis (i.e. assuming $\Delta C_p° = 0$) is

then carried out to fit the curve shape, assuming a two state model, to provide

$\Delta G_T°$, $\Delta H°$, and $\Delta S°$ (Eq.1) (Borer, Dengler *et al.* 1974; Albergo, Marky *et al.*

1981; Petersheim and Turner 1983; Freier, Kierzek *et al.* 1986; Puglisi and

Tinoco 1989; Rentzeperis, Ho *et al.* 1993):

$$\ln K = \frac{-\Delta H°}{RT} + \frac{\Delta S°}{R} \qquad \text{Eq.2.1}$$

$$\frac{1}{T_M} = \frac{R}{\Delta H°} \ln C_T + \frac{\Delta S°}{\Delta H°} \qquad \text{Eq.2.2}$$

For a unimolecular transition, an RNA molecule is unfolded from a hairpin

structure, so the melting temperature does not show concentration-dependence

(Jaeger, SantaLucia *et al.* 1993). This experiment is useful in monitoring the

folding of designed RNA model sequences, to determine whether the RNAs

assume a hairpin conformation, with much lower demand for the quantity of the

RNA sample (typically 10^{-4} to 10^{-6} M), compared with the high concentrations

required for NMR RNA sample (typically 10^{-3} M). However, for a bimolecular

transition, RNA molecules are unfolded from a duplex into two random-coil

strands, and the concentration of each strand has an effect on the melting
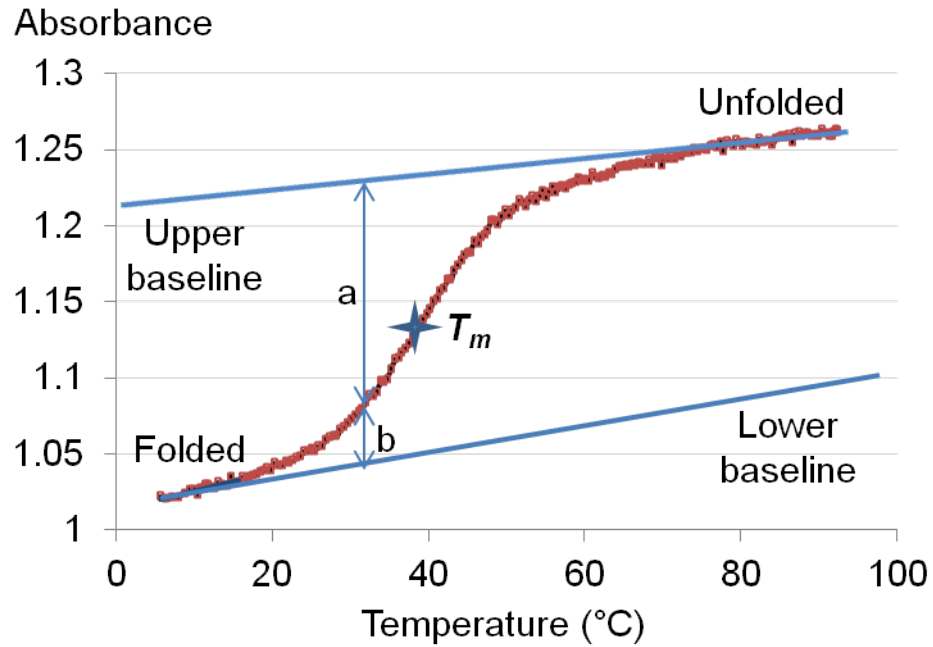
Figure 2.3 A typical UV-melting curve obtained from a two-state conformational change of an RNA duplex (John SantaLucia 2000). At a temperature, the equilibrium constant can be derived from the proportion of the native folded state α, which is the ratio of a/(a+b), where a is the vertical distance between the upper baseline and the data point and b is that between the lower baseline and the data point.

A

B

C
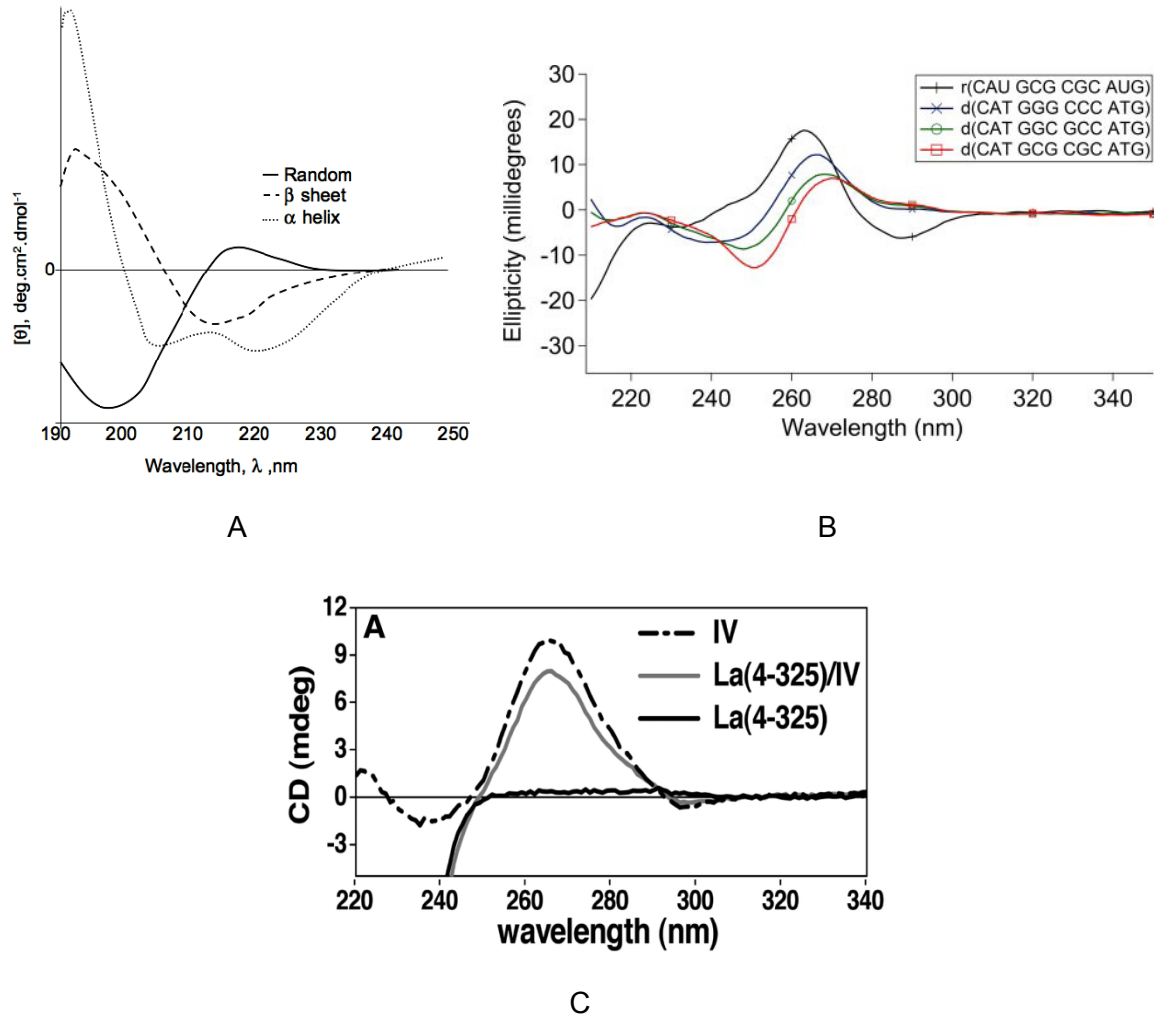
Figure 2.4 Circular Dichroism spectra of proteins and nucleic acids. A. Protein CD spectra overlay shows different absorption properties in the far-UV region of different protein secondary structure (Gratzer and Mendelsohn 1978). B. Nucleic acid CD spectra overlay shows different absorption properties in the near-UV region of B-form DNA duplex (d(CATGCGCGCATG)), A-form RNA duplex (r(CAUGCGCGCAUG)), and two transition forms of DNA duplexes (Tsai, Engelhart *et al.* 2009). C. CD spectra overlay of hepatitis C virus RNA domain IV, La motif (protein), and the complex shows the conformational change of HPC domain IV after binding of La motif (adapted from (Martino, Pennell *et al.* 2011)).

temperature, which enables an alternative method to obtain thermodynamic parameters (i.e. ΔG°_{37}, ΔH°, and ΔS°) from linear-regression of $1/T_M$ vs. $\ln(C_T/4)$, where C_T is the total concentration of RNA molecules, besides the curve shape fitting (Eq.2) (Borer, Dengler *et al.* 1974). Usually, a series of samples covering 20 to 50 fold in concentration range are used. More importantly, agreement between the thermodynamic parameter sets obtained from curve shape fitting and $1/T_M$ vs. $\ln(C_T/4)$ linear regression is essential to justify the assumption of a two state model, and "<15%" deviation is usually desired (Freier, Kierzek *et al.* 1986; SantaLucia and Turner 1997).

An Aviv 14DS UV-Visible spectrometer was used for all UV-melting experiments. The features of this model of spectrometer include a double monochromator, dual-beams, a thermoelectric cuvette holder, and controlled by an interfaced PC computer. The sample temperature was increased from 0 ºC to over 90 ºC at a steady rate of 0.8 ºC per minute. Furthermore, combinations of different cuvette and aluminum spacer can provide four different light path lengths (0.1, 0.2, 0.5, and 0.8 cm). This guarantees accurate measurements of a large range of concentration (up to about 50 fold).

2.3 Circular Dichroism

For nearly half a century, Circular Dichroism (CD) has been a powerful biophysical method to explore the conformation of proteins and nucleic acids individually (Figure 2.4 A and B), and conformational change of complex formation of proteins and nucleic acids (Figure 2.4 C) (Brahms and Mommaerts

1964; Holzwarth and Doty 1965; Gratzer and Mendelsohn 1978; Warrant and Kim 1978; Tsai, Engelhart *et al.* 2009; Martino, Pennell *et al.* 2011).

CD is a spectroscopic method measuring the difference in absorption of left vs. right circularly polarized light. Circularly polarized light is obtained from the superposition of two electromagnetic waves satisfying the following conditions: 1) the two electromagnetic waves are oscillating in two plains perpendicular to each other; 2) the two electromagnetic waves are of the same amplitude and wavelength; 3) there is a difference of π/2 (or - π/2) in phases of the two electromagnetic waves, giving rise to the right and left circularly polarized light. Since most biomolecules built up with chiral residues are asymmetric, they show different behavior in absorption of left and right circularly polarized light, and the difference in extinction coefficient is termed as Circular Dichroism (Cantor and Schimmel 1980).

Absorption of circularly polarized light follows Lambert-Beer law, so the circular dichroism is defined as Δε by Eq. 2.3:

$$\Delta A = A_l - A_r = \varepsilon_l \bullet C \bullet l - \varepsilon_r \bullet C \bullet l = \Delta \varepsilon \bullet C \bullet l \qquad \text{Eq.2.3}$$

where ΔA is the difference in absorption between the left (l) and right (r) circularly polarized light, ε is the molar extinction coefficient, C is the molar concentration of the molecule, l is the light path length. Δε can be further converted to ellipticity (θ, Eq. 2.4) and molar ellipticity ([θ], Eq. 2.5) for an expression of the circular dichroism, [θ].

$$\theta = 32.98 \bullet \Delta \varepsilon \bullet C \bullet l \qquad \text{Eq.2.4}$$

$$[\theta] = \frac{100 \bullet \theta}{C \bullet l} \qquad\qquad \text{Eq 2.5}$$

In protein CD, absorption by the electronic transitions in the hydrogen bounded peptide bond in far-UV region is typically used to investigate the secondary structure of the peptide chain. α-helical structure gives rise to one negative peak at 222 nm originated from the n-π* transition, and two peaks, including one negative peak at 208 nm from the parallel π-π* and one positive peak around 195 nm from perpendicular π-π* , originated from the split π-π* transition due to the exciton coupling (Greenfield and Fasman 1969). β-sheet structure is characterized by one negative peak around 215 nm originated from the n-π* transition and one positive peak around 195 nm from π-π* (Figure 2.4 A) (Whitmore and Wallace 2008; Wallace and Janes 2009). Aromatic side chains and disulfide bonds also contribute to the CD spectra in the near-UV region (Strickland 1974). Even though aromatic ring structures have high absorption in the near-UV region from the π-π* transition, while they are not asymmetric intrinsically, and have to be folded into a defined asymmetric tertiary structure to be CD active. Therefore, CD signals from aromatic residues are not usually used to explore the tertiary folding of proteins.

Similarly, aromatic ring structures of the nucleic acid bases do not have intrinsic CD activity either, but when folded into helical (A-, B-, and Z-form) structures, base rings gain hyperchromicity in CD spectra through base stacking and Coulombic interactions (Wallace and Janes 2009). B-DNA, the most commonly observed DNA double-helix form in solution, shows a positive peak around 280 nm and a negative peak around 240 nm of similar amplitude, with an

intense positive peak around 190 nm. In CD spectra of A-DNA (A-RNA), all the characteristic peaks are blue-shifted, i.e. toward shorter wavelength, compared to B-DNA, and the amplitude of the positive peak around 270 nm is more intense compared to that of the negative peak around 210 nm (Figure 2.4 B). Z-DNA shows very different CD characteristics due to its left-handedness, opposite to that of the A-form and B-form DNAs (Wang, Quigley et al. 1979). The most prominent signatures of Z-DNA CD spectra are a very shallow negative peak around 290 nm, a deep trough around 190 nm, and an intense positive peak around 175 nm (Riazance, Baase *et al.* 1985).

CD spectra of proteins and nucleic acids provide only collective effects of global structure on electron transitions, and no detailed structural information can be directly determined from CD experiments. However, CD is very powerful for monitoring real-time conformational changes of biomolecules in melting, titration and ligand binding experiments (Wang, Huber *et al.* 1998; Abeysirigunawardena and Chow 2008).

2.4 Surface Plasmon Resonance

When there is an interface between two media composed of two materials, whose real part of the dielectric function has opposite sign, coherent electron oscillation across the interface creates surface plasmons (Raether 1988). The excitation of surface plasmons forms the foundation for surface Plasmon resonance. If a light beam is incoming through the medium of higher refractive index, reflection and refraction happen simultaneously at the interface, until the incidence angle is higher than a certain critical angle, termed total

internal reflection angle (θ_{TIR}), when only total internal reflection is detected. Besides reflection, the electromagnetic field component of the incidence light penetrates into the low-refractive-index medium for a distance of 300 nanometers, and produces an evanescent wave. The resonance energy of this evanescent wave can be transferred to surface plasmons when following conditions are satisfied: 1) the interface is coated with a thin layer of metal (gold or silver); 2) incident light is monochromatic; 3) polarization of the incident light occurs parallel to the plane of incident light; 4) incident angle is right on a certain critical angle (Ordal, Long *et al.* 1983). Properties of the material attached to this thin layer of metal influence the resonance conditions, i.e. refractive index, and then lead to the change of the identity of the absorbed angle, termed surface plasmon resonance angle (θ_{SPR}) (Flanagan and Pantell 1984; Kooyman, Kolkman *et al.* 1988; Liedberg, Nylander *et al.* 1995). Shift of θ_{SPR} for 0.1 millidegree corresponds to 1 Response Unit (RU), which is equivalent to binding of 1 pg/mm^2 of protein or 0.8 pg/mm^2 of RNA. (Van Ryk and Venkatesan 1999; Wang and Anslyn 2011). Thus the SPR signal denoted by the RU is a direct measurement of the mass concentration of the biomolecules attached to the sensor chip (Figure 2.5).

In a Biacore SPR chip, four flow cells, each of 2.4×0.5×0.05 mm, are serially connected. To monitor the non-specific binding of analyte onto the chip surface and avoid artifacts introduced by buffer changing, at least one flow cell is used as a reference cell, where no ligand is immobilized on the chip surface. The net binding effect is calculated by subtracting the reference flow cell signal from
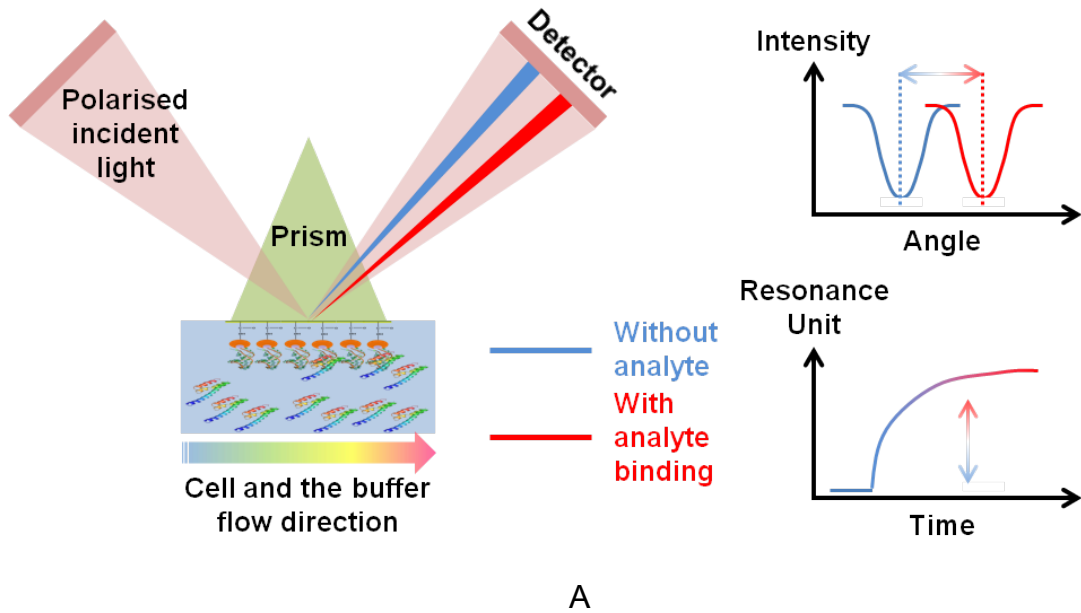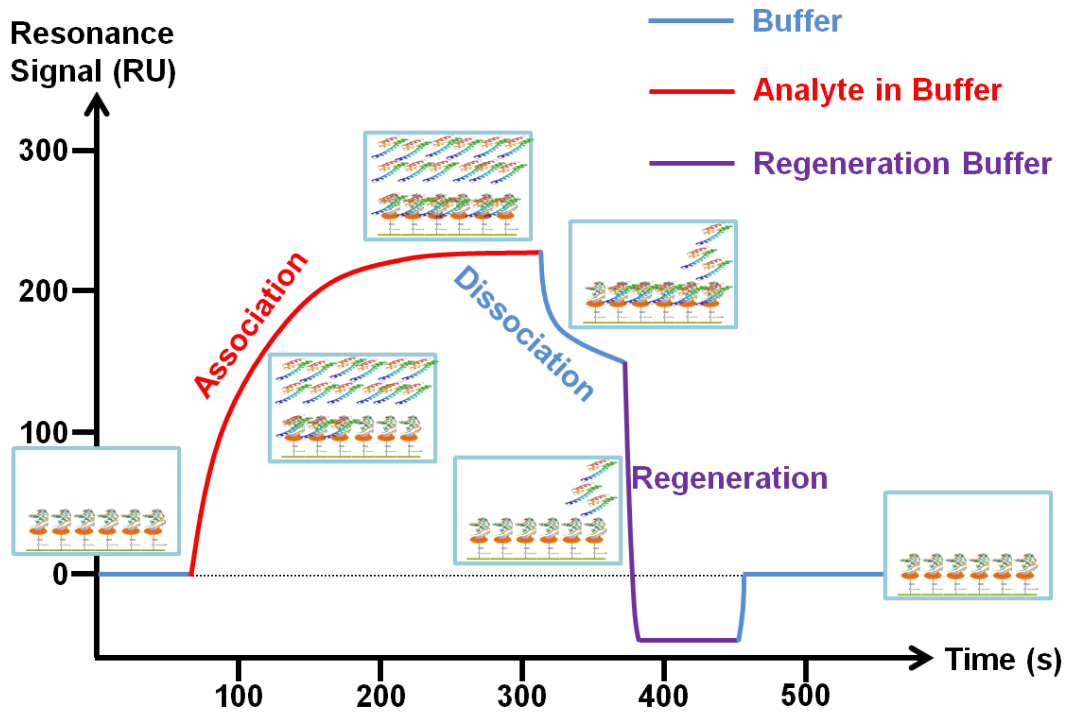
Figure 2.5 Surface Plasmon Resonance diagram and sensorgram (Courtesy of Biacore international AB). A. At θ$_{SPR,}$ conversion of photon energy to plasmon results in a dark band in the reflected light. With binding of analyte (in the mobile

phase) onto the ligand (immobilized on the SPR chip surface), the refractive index of the chip surface changed proportionally, leading to shift of the $\theta_{SPR,}$ and this shift is converted to increase of the resonance signal in response unit along the time when the analyte is injected. B. Once the injection of analyte solution is started, the sensorgram enters "association" phase, when analyte molecules keep binding onto the ligand until equilibrium is established. In the dissociation phase, analyte solution is replaced by buffer, and the bound analyte molecules start dissociating from the ligand, showing decrease in the response unit. A regeneration buffer is always used before re-injection of the analyte solution when the binding between the analyte and ligand is so strong that dissociation is extremely slow and inefficient. In the final phase, buffer is injected again, and the sensorgram profile should be stabilized at the same response unit level as before the injection of analyte solution.

the experimental flow cell signal. The net sensorgram (SPR vs time) is further used to deduce the kinetic and thermodynamic constants of the binding event.

Assumed to be a one to one stoichiometry, binding between the Analyte and Ligand can be treated as a reaction expressed as Eq. 2.6:

$$A + L \underset{k_d}{\overset{k_a}{\rightleftharpoons}} AL \qquad \text{Eq.2.6}$$

where k_a is the association rate constant ($M^{-1} \cdot s^{-1}$) and the k_d is the dissociation rate constant (s^{-1}), as defined by Eq. 2.7 and 2.8:

$$\frac{d[AL]}{dt} = k_a [A][L] \qquad \text{Eq.2.7}$$

$$\frac{d[AL]}{dt} = -k_d [AL] \qquad \text{Eq.2.8}$$

where [AL] is the direct observable in the sensorgram response (R), [A] is the concentration of analyte in the mobile phase, which is assumed to be constant (C), and [L] is the concentration of unbound ligand on the chip surface. In the association phase, both association and dissociation are taking place in the flow cells, the net effect gives rise to Eq. 2.9:

$$\frac{d[AL]}{dt} = k_a [A][L] - k_d [AL] \qquad \text{Eq.2.9}$$

At time point t, the concentration of unbound ligand is expressed as Eq. 2.10:

$$[L]_t = [L]_{total} - [AL]_t \qquad \text{Eq.2.10}$$

Since $[L]_{total}$ dictates the binding capacity of the flow cell, in terms of response of analyte binding, it is equivalent to the maximum response (R_{max}). So the

association phase reaction rate and dissociation phase reaction rate equations

can be re-written as Eq. 2.11 and 2.12 respectively:

$$\frac{dR}{dt} = k_a\,C(R_{max} - R_t) - k_d\,R_t = k_a\,CR_{max} - (k_a\,C + k_d)R_t \quad \text{Eq.2.11}$$

$$\frac{dR}{R_t} = -k_d\,dt \quad \text{Eq.2.12}$$

From integration of the two equations listed above, functions correlating

the real-time response (R_t) and time point t can be obtained for the association

and dissociation phase, respectively:

$$R_t = \frac{k_a\,CR_{max}\,(1 - e^{-(k_a C + k_d)t})}{k_a\,C + k_d} + R_0' \quad \text{Eq.2.13}$$

$$R_t = R_0\,e^{-k_d t} + R_\infty \quad \text{Eq.2.14}$$

Association and dissociation rate constants can be determined from non-

linear regression of the sensorgram in the association phase and dissociation

phase with Eq. 2.13 and 2.14, respectively. The dissociation constant, K_d, is

derived from Eq 2.15 based on the known k_a and k_d from the kinetic calculation:

$$K_d = \frac{k_d}{k_a} \quad \text{Eq.2.15}$$

2.5 Nuclear Magnetic Resonance

2.5.1 Introduction

From the Nobel prize in physics being awarded to Otto Stern in 1943 to

the Nobel prize in physiology or medicine being awarded to Paul C. Lauterbur

and Peter Mansfield in 2003, six Nobel prizes have been awarded to researchers

working in Nuclear Magnetic Resonance (NMR) field. Kurt Wűthrich, the

chemistry Nobel laureate of 2002, established the methodology for solution three-dimensional structure determination of biological macromolecules by nuclear magnetic resonance spectroscopy. The first solution NMR RNA structure deposited in Protein Data Bank (PDB 1ELH) in 1993 was solved by Peter Moore's lab on the helix I from the 5S RNA of *E. coli* (White, Nilges *et al.* 1992). Since then, 402 solution NMR structures of RNA have been deposited into the PDB. The largest RNA structure ever solved by solution NMR was published by the Summers' group in 2004, and the number of residues is 101 (D'Souza, Dey *et al.* 2004). Though limited by the resolving power in terms of residue number, NMR method can be used to elucidate the RNA 3D structure in solution, without being biased by the artifacts introduced by crystallization necessary for X-ray crystallography method, and relevant to physiological conditions. Another unprecedented application of solution NMR is to study the dynamics of RNA structures in solution (Varani and Tinoco 1991; Latham, Brown et al. 2005; Miller, Shajani et al. 2006; Getz, Sun et al. 2007).

Research presented in this thesis is focused on the structure calculation of solution NMR RNA structures. A systematic methodology has been established for this application (Fürtig and Schwalbe 2003). Briefly, a complete set of one, two, and three-dimensional NMR experiments are run on an RNA NMR sample, and structural restraints, including distances between atom pairs, dihedral angles involving 4 consecutive atoms (Figure 2.6), and the global positioning of the atoms relative to each other (Residual Dipolar Coupling, RDC), are generated from the available NMR spectra. Finally, a force-field-based structural calculation
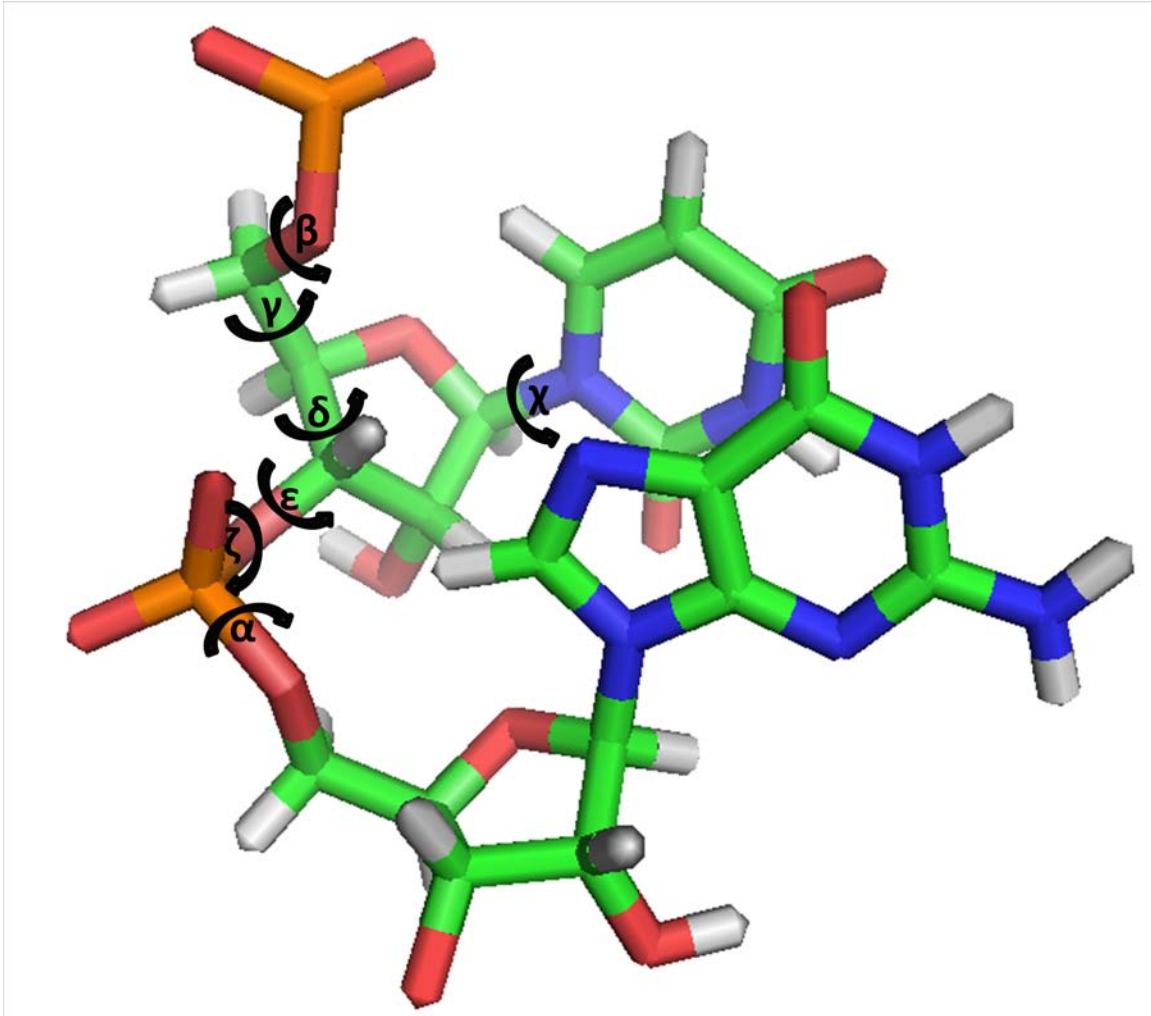
Figure 2.6 Six backbone and one side chain dihedral angles of one residue in RNA molecules. Green, blue, red, yellow, and white colors code for carbon, nitrogen, oxygen, phosphorus, and hydrogen atoms. (Saenger 1984)

and energy minimization is used to calculate the 3D structure of the RNA, with consideration of all the restraints and RNA covalent linkages (Figure 2.7).

2.5.2 One dimensional (1D) experiments

1D NMR experiments are the most basic and important experiments, even though barely any structural information could be pulled out from them. When a new sample is ready or any condition is changed, 1D NMR experiments are used to calibrate the pulse widths of all the nuclei employed in the experiments for optimal performance. For proton (^{1}H), a single pulse experiment is carried out with various pulsing time to determine the 360° pulse width (the second null point of the solvent proton NMR spectrum). For other nuclei (^{13}C and ^{15}N), multi-pulse decoupling experiments are often used to determine the pulse width by varying the decoupling time, and looking for the null time point, corresponding to the creation of undetectable double quantum coherence, or reversion of the peak doublet from the proton covalently attached to the nucleus (Claridge 1999). Another essential application of 1D NMR experiment is to determine the optimal temperature condition for all other multi-dimentional NMR experiments.

Due to the limitation on size of the RNA molecules that can be solved in solution NMR experiments, model systems of oligonucleotides are always constructed from segments of the biomacromolecules, e.g. ribosomal RNAs, transfer RNAs, and telomerase RNA (Morosyuk, Cunningham *et al.* 2001; Kim, Zhang *et al.* 2008; Denmon, Wang *et al.* 2011). The possibility of alternative folding has to be excluded first, by comparing the pattern of imino proton peaks downfield in the 1D proton NMR spectrum of RNA samples dissolved in
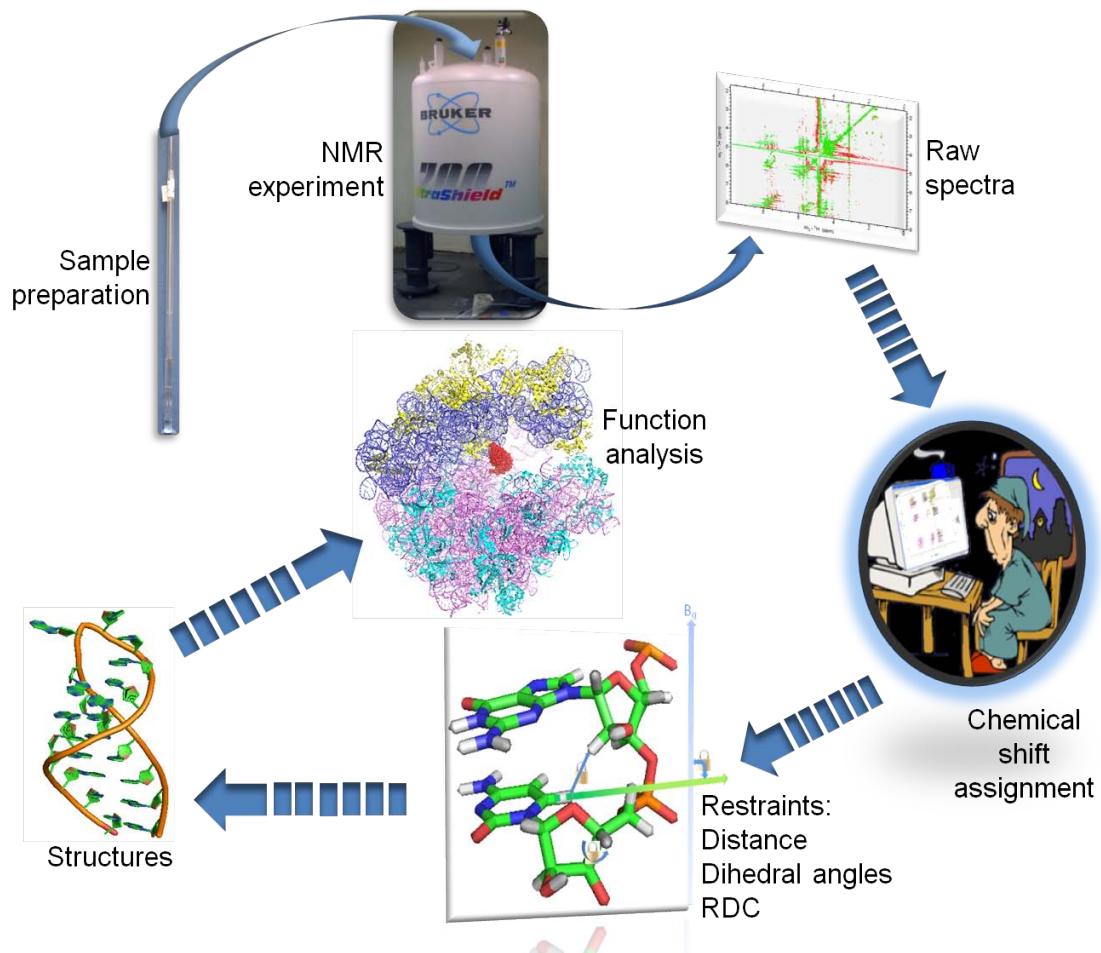
Figure 2.7 A flow chart showing all steps for NMR experiments on RNA oligonucleotides.

90%H_2O•10%D_2O with known base pairing pattern from other knowledge sources, e.g. comparative sequence analysis, crystal structure, chemical probing, and RNA secondary structure prediction.

2.5.3 Two dimensional (2D) experiments

With all the optimization and validation work done with the 1D NMR experiments, the majority of the assignments and structural information have to be obtained from multi-dimensional experiments, starting with 2D NMR experiments.

The mere number and chemical shift distribution of protons in an RNA oligonucleotide render 1D NMR experiments incapable of resolving all the nuclei (Figure 2.8) (Wuthrich 1986). By correlating the two or more nuclei in close spatial proximity or J-coupling through covalent bonding (2-4 bonds typically), a 2nd dimension can help resolve the overlapping chemical shifts. There are two basic categories for the correlation. The first way is to correlate two nuclei in close proximity through space by the Nuclear Overhauser Effect (NOE), and a representative experiment of this type is the 2D NOESY experiment (Schirmer and Noggle 1972; Neuhaus and Williamson 1989). The second transfer mechanism is to correlate two nuclei connected directly or indirectly through covalent bond(s), e.g. HMQC (1 bond), HNN COSY (2 bonds), 2D DQF COSY (3 bonds), TOCSY (up to 6 bonds).

2.5.3.1 2D NOESY experiment

2D NOESY experiment is based on nuclear Overhauser effect, resulted from through space dipolar coupling between one pair of protons within 4 ~ 5 Å in
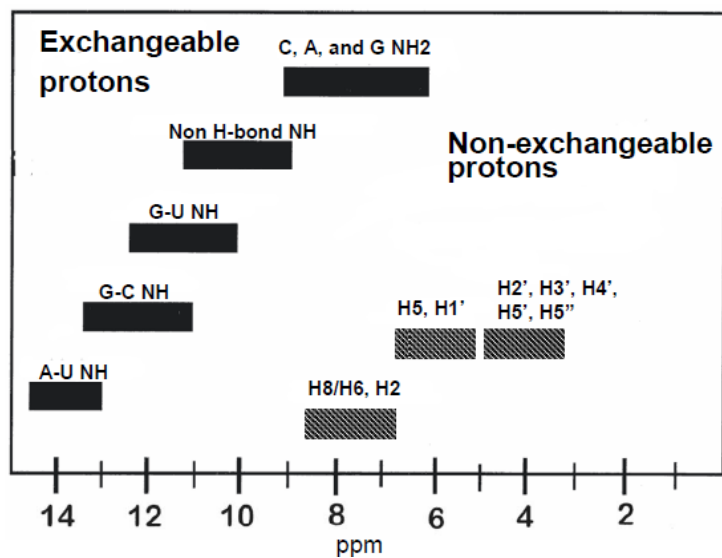
A
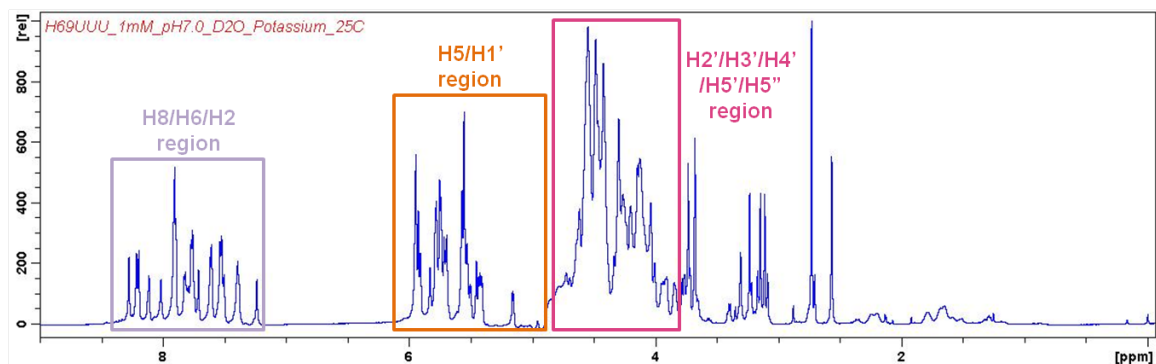
B
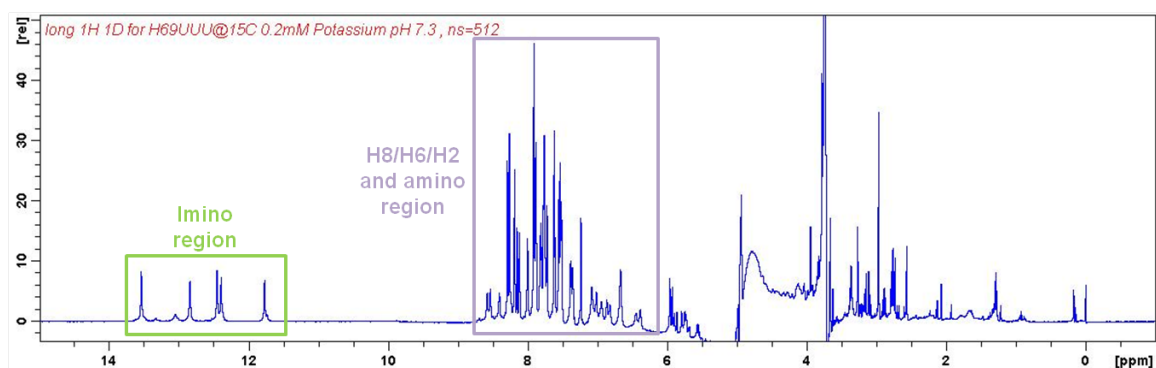
C

Figure 2.8 Chemical shift distributions of protons in RNA molecules. A. Illustration

of chemical shift distribution of protons in RNA molecules. (Wuthrich 1986). B.

1D-^{1}H NMR experiment on H69UUU at 25°C in 99.96% D_2O. C. 1D-^{1}H NMR experiment on H69UUU at 15°C in 90%H_2O•10%D_2O. Different regions of proton distributions are color coded in the two spectra. Overlap of proton resonances are seen in almost all regions other than the imino region.
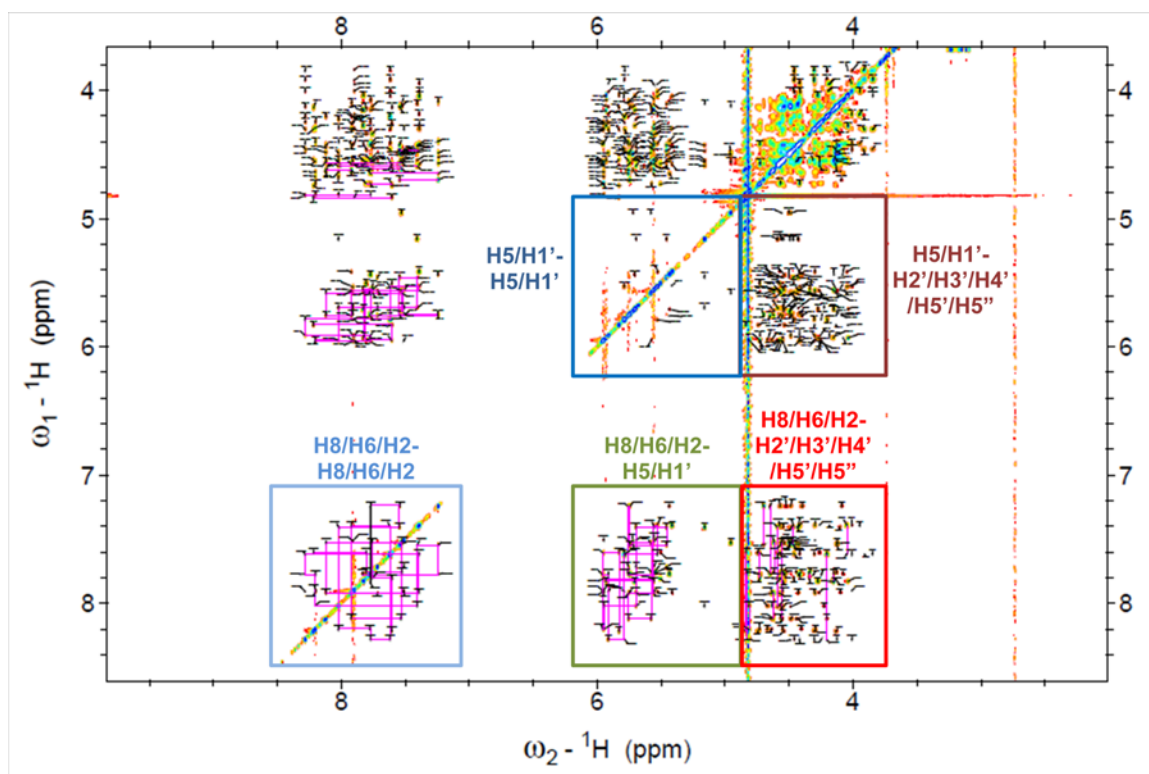
Figure 2.9 2D NOESY spectrum of H69UUU at 25°C in 99.96% D_2O.

3D space (Claridge 1999; Neuhaus 2011). Due to this property, 2D NOESY is most useful in generating inter-proton distance restraints for structure calculations. In this case, a short mixing time, the duration when magnetization transfer is permitted, ranging from 60 to 150 ms is used, to avoid compromising the accuracy from spin diffusion. Another indispensible application of this experiment on RNA oligonucleotides is to assign chemical shifts of the base and sugar protons by a NOESY "walk". This experiment is done on RNA samples dissolved in 99.96% D_2O with a mixing time of 250 to 400 ms to maximize the magnetization transfer (Figure 2.9). Additional information that can be generated from this experiment includes: 1) if the sequence is correct; 2) if there is the modification(s), e.g. psudouridylation, 3) is the χ dihedral angle *syn* or *anti*.

The 2D NOESY experiment on RNA samples dissolved in 90%H_2O•10%D_2O is helpful to assign the imino and amino protons involved in hydrogen bonding. Since base pairs in the stem region of an RNA oligonucleotide are always stacked on top of each other, the imino protons involved in hydrogen bonding are within the NOE detection range, and a crosspeak of two imino protons next to each other shows up in the 2D NOESY spectrum. The base pair interaction also position the imino proton in close proximity to amino protons, and the crosspeaks are useful to assign the amino proton chemical shifts (Fürtig and Schwalbe 2003).

2.5.3.2 2D DQF-COSY experiment

DQF-COSY is the abbreviation of Double Quantum Filtered COrrelation SpectroscopY, which is based on through-bond resonance correlation (Varani

and Tinoco 1991). The 2D DQF-COSY experiment on RNA oligonucleotides, in 99.96% D_2O, is useful to detect the sugar conformation, i.e. C2'-endo or C3'-endo sugar pucker. Sugar protons H1' and H2' are connected by three covalent bonds with C1' and C2' in between. In A-form RNA, e.g. the stem region of a hairpin structure, C3'-endo conformation orients the C1'-H1' and C2'-H2' bonds perpendicular to each other, and the coupling constant is only about 1Hz, which renders the crosspeak between the H1' and H2' unobservable. While some residues in the loop region of a hairpin structure may assume C2'-endo sugar pucker conformation, typical for B-form duplexes, and show strong H1'-H2' crosspeaks due to a large coupling constant of about 8Hz from a near antiparallel geometry of the C1'-H1' and C2'-H2' bonds (Altona 1982).

The 2D DQF-COSY experiment is also helpful in providing information regarding pseudouridylation modifications. First, the signature strong H5-H6 crosspeak typical for pyrimidines is absent due to removal of the H5 in the isomerization from uridine to pseudouridine. This property is used to confirm the modification conveniently, without all the modification and mass-spectrometry experiments. Second, a moderate to weak crosspeak from H6 to H1' could be observed if the H6-C6-C5-C1'-H1' moiety assumes a near planer conformation. Combined with H6-H1' crosspeak in the 2D NOESY spectrum, the χ dihedral angle can be determined.

2.5.3.3 HMQC and HSQC

Resonance correlations between covalently connected ^1H and ^{13}C or ^1H and ^{15}N can be detected by Heteronuclear Multiple (or single) Quantum

Correlation spectroscopy (Nikonowicz and Pardi 1993; Dieckmann and Feigon 1997). These types of 2D experiments are indispensible in assigning the chemical shifts of RNA proton chemical shifts benefiting from the wider distribution of ^{13}C chemical shifts (Figure 2.10). Adenine H2 resonances can be easily distinguished from the H8 and H6 resonances from the 2D HMQC of the base region, facilitating the assignments of H8/H6 chemical shifts in the 2D NOESY spectra. In the sugar moiety, almost all sugar protons (other than the H2' and H3') are further separated by the chemical shifts of carbons they covalently attach to, compared to that in the 2D NOESY spectra. ^{13}C enrichment of the RNA oligonucleotide sample is crucial in improving the sensitivity of this experiment. 2D HMQC or 2D HSQC with the carbon chemical shift ranging from 90 to 110 ppm are used to distinguish the H5 of cytidines, whose C5 chemical shift is around 95, from the H5 of uridines, whose C5 chemical shift is around 105 ppm.

2.5.3.4 Other J-coupling experiments

Other J-coupling based NMR experiments are also applied in RNA structure studies. These methods include TOCSY, HNN-COSY, and ^{1}H-^{31}P HETCOR.

In TOtal Correlation SpectroscopY experiments, a spin lock pulse sequence is applied, when magnetization is transferred throughout the whole scalar-coupled network of spins (Bax, Clore et al. 1990). For instance, in the 2D homonuclear TOCSY experiment, all the proton resonances belonging to the same ribose sugar ring in RNA are correlated to the H1'resonance, and a series of crosspeaks from the H1' to all sugar protons show up in the spectra. This
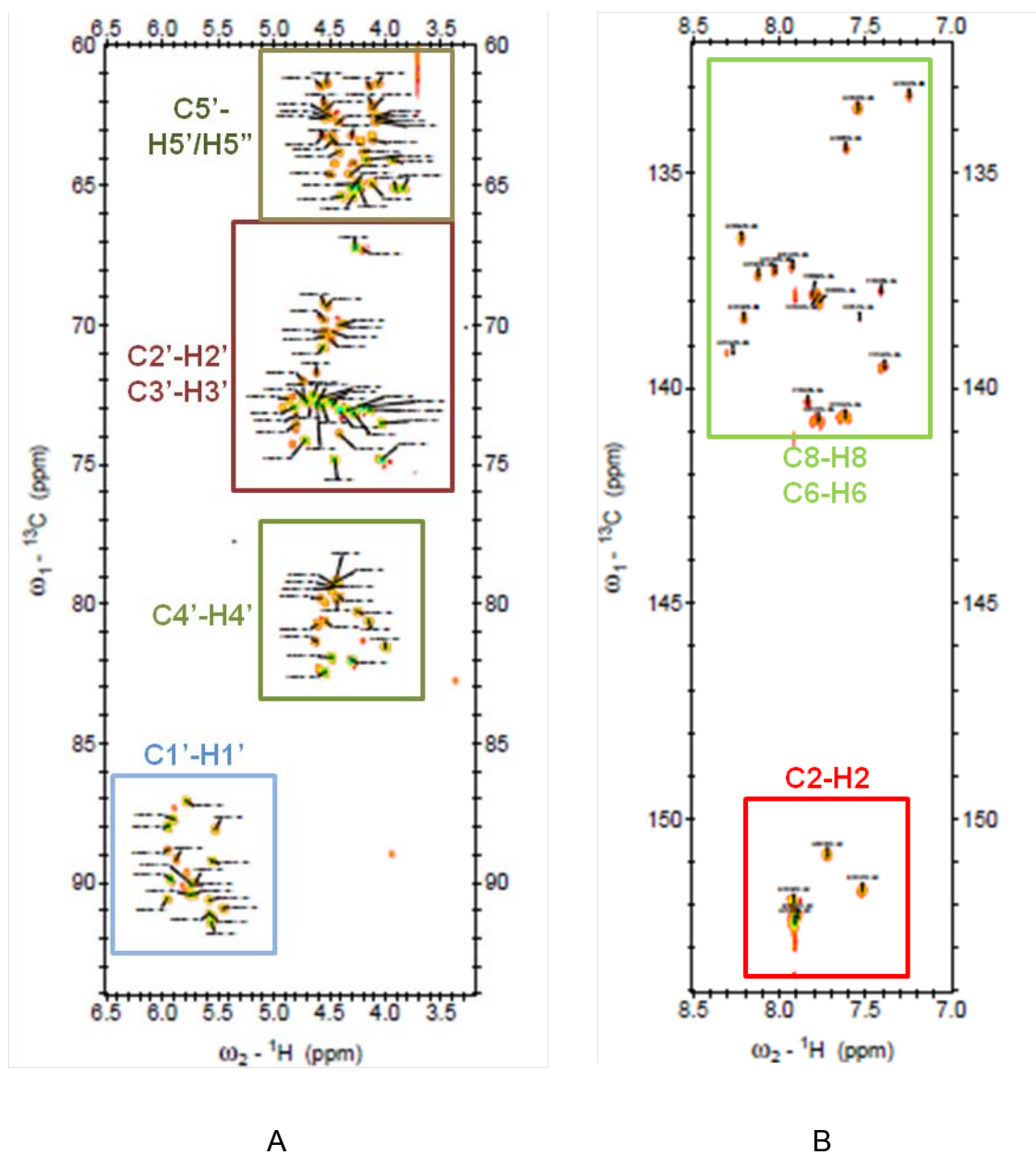
Figure 2.10 HSQC and HMQC. A. 2D CT-HSQC of the sugar region from H69UUUDL (^{13}C/^{15}N doubly labeled) in 99.96% D_2O at 25°C. B. 2D HMQC of the base region from H69UUU in 99.96% D_2O at 25°C.

experiment provides complementary information to group protons in the same residue.

2D HNN-COSY experiments can be used to assign the imino protons involved in base pair interactions (Wohnert, Dingley *et al.* 1999). This type of experiment takes advantage of the $^2J_{NN}$ of the two nitrogen atoms involved in one hydrogen bond, and ^{15}N enrichment is required to improve the sensitivity due to the extremely low abundance of ^{15}N in natural-abundance RNA samples.

In natural-abundance NMR samples, the NMR active ^{31}P is the dominant isotope, and it is the only information source to determine restraints for dihedral angles α, β, ε, and ζ. Since the chemical shift of ^{31}P is sensitive to the RNA conformation, 1D ^{31}P is able to well-define the ranges for α and ζ (Gorenstein 1981). While, 2D ^1H-^{31}P HETCOR can help narrow down the ranges of all the dihedral angles by determination of the coupling between the ^{31}P (n) and H3' (n-1), H4' (n), H5' (n), and H5'' (n) (Varani, Cheong *et al.* 1991).

2.5.4 Three dimensional (3D) experiments

In most cases, 1D and 2D NMR experiments are not enough to assign the chemical shifts of all protons in an RNA oligonucleotide, and it is necessary to introduce an extra dimension to further separate the resonances, and finalize the assignment table. With natural-abundance RNA samples, only a limited number of 3D experiments can be run, e.g. 3D TOCSY-NOESY and 3D NOESY-NOESY (Vuister, Boelens *et al.* 1991; Wijmenga, Heus *et al.* 1994). Utilization of NMR active isotope enrichment could greatly expands the repertoire of 3D NMR experiments. Some of these experiments are solely based on scalar-coupled

magnetization transfer, e.g. 3D HCcH-COSY, 3D HCcH RELAY, 3D HCcH-TOCSY, and 3D hCCH-TOCSY (the upper case letters denote the observable resonances), and others employ through space coupling, e.g. 3D NOESY-HMQC (Nikonowicz and Pardi 1992; Nikonowicz and Pardi 1993; Hall 1995; Pardi 1995; Varani, Aboul-ela *et al.* 1996).

2.5.4.1 3D TOCSY-NOESY

As is directly illustrated in its name, the 3D TOCSY-NOESY experiment can be seen as a hybridization of 2D TOCSY and 2D NOESY experiments (Wijmenga, Heus et al. 1994). The merit making this experiment distinguished from most other 3D NMR experiments on RNA is that no isotope enrichment is needed. It plays a major role when there are modified residues in the RNA oligonucleotide, since isotope-enriched solid phase synthesis is usually prohibitively expensive.

There are two magnetization transfer pathways from one proton, e.g. H2' on a ribose sugar ring (Figure 2.11) (Sijenyi 2008). One is a through bond pathway to transfer the magnetization throughout the ribose ring during the TOCSY spin lock, and all sugar protons are correlated to the H2'. The second is a through space pathway to transfer the magnetization from the H2' to all other protons within NOE distance (4 ~ 5 Å), and the correlations between the H2' and the base H8/H6/H2 of the same residue or the 3' side residue are essential to resolve the overlaps of proton resonances that are observed in a 2D TOCSY.

2.5.4.2 3D experiments employing isotope enrichment
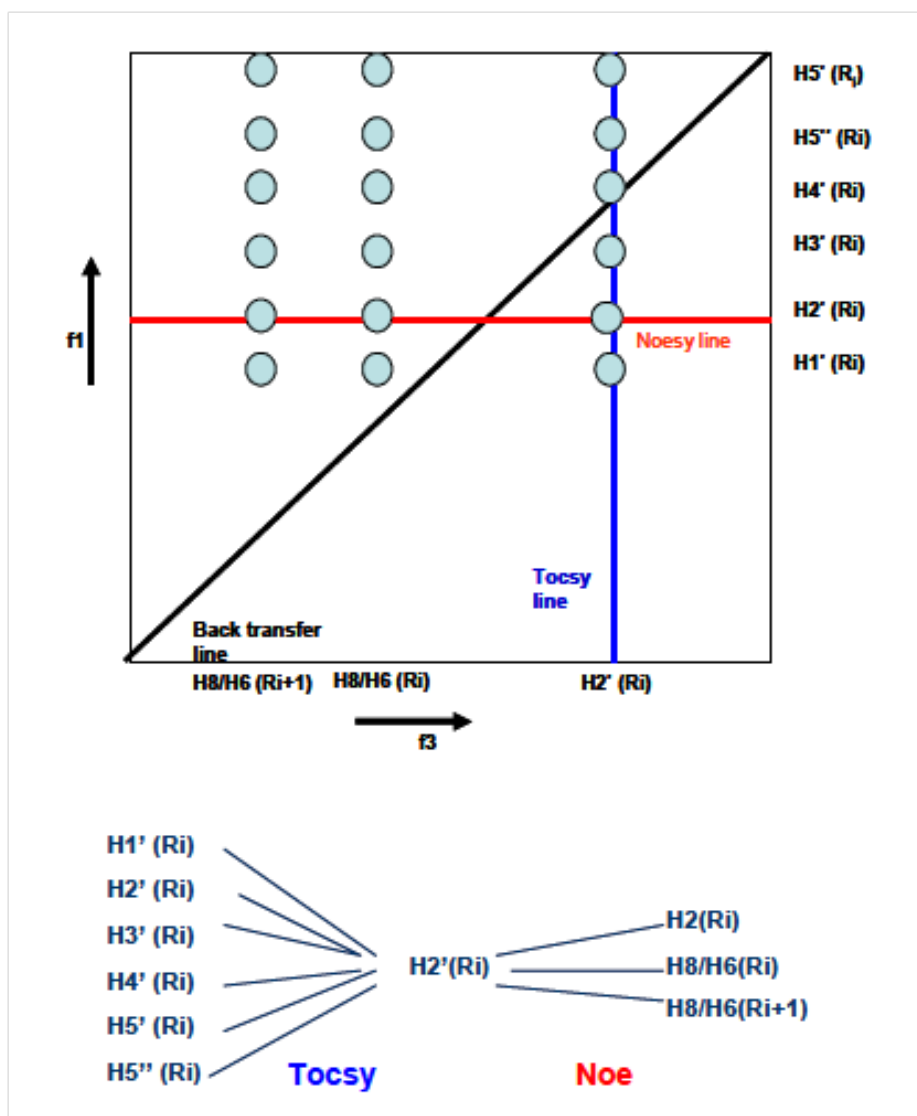
Figure 2.11 3D TOCSY-NOESY spectrum and magnetization transfer pathways. When the H2' of a ribose sugar ring is considered, i.e. the f2 dimension is positioned at the chemical shift of this H2', all the TOCSY crosspeaks from H2' and sugar protons in the same ring are deployed along the f1 dimention on the "Tocsy line", and along the f3 dimension, crosspeaks between the H2' and base H8/H6 protons belonging to the same residue or the 3' side residue show up on the "Noesy line".

Compared to the natural-abundance RNA samples, ^{13}C/^{15}N isotope enriched samples provide much more options of 3D NMR experiments to run (Pardi 1995).

To unambiguously assign all the chemical shifts of proton and hydrogen atoms in one sugar ring, HCcH-COSY, hCCH-TOCSY, and HCcH-TOCSY can be used complementarily (Figure 2.12) (Bax, Clore et al. 1990; Kay, Ikura et al. 2011). Starting with the known chemical shifts of C1' and H1', all the carbon resonances in the same sugar ring appear in a slice in the hCCH-TOCSY. Similarly, in a HCcH-TOCSY, all the proton correlations appear in a slice. With limited choices of chemical shifts of all the carbon and proton resonances in the ring, HCcH-COSY is used to confirm the assignments in a step-wise fashion, since only one-bond magnetization between the carbon atoms is permitted.

Dipolar coupling 2D experiments can be hybridized with scalar coupling to provide the third dimension, e.g. 3D NOESY-HMQC. In 3D NOESY-HMQC experiment, the magnetization is first transferred from a proton to another proton within NOE distance, e.g. a sugar proton to a base H8/H6 or the sugar proton, then to the carbon atom directly attached to the proton by scalar coupling, e.g. H8/H6 to C8/C6 or the sugar proton to the sugar carbon covalently attached to it. In this case, any overlapping in the 2D NOESY due to the close chemical shifts of two H8/H6 is now resolved by the C8/C6 dimension. One application of this technique is to resolve the crosspeaks of H3'-H5'/H5″ and H4'-H5'/H5″. Comparison between the intensities of H3'-H5' with H3'-H5″ and H4'-H5' with H4'-H5″ is diagnostic to determine the conformation of dihedral angle γ,
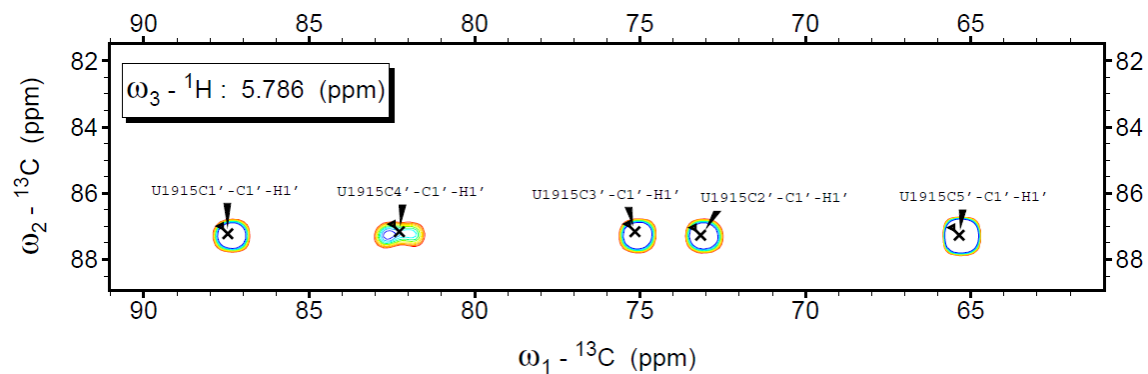
A
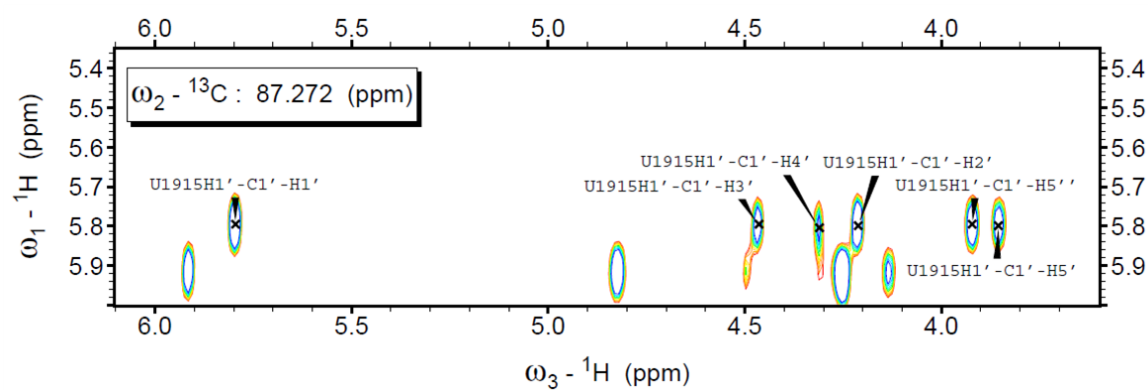
B
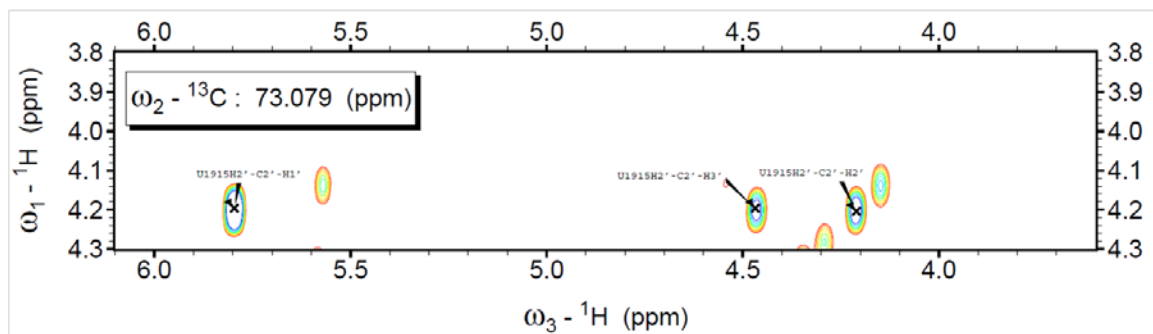
C

Figure 2.12 3D NMR experiments of sugar proton and carbon atoms. A. 3D hCCH-TOCSY correlates all the carbon atoms in a sugar ring. When the chemical shift of f3 dimension is positioned at H1' of U1915, all the carbon resonances in the sugar ring of U1915 show up at C1' chemical shift of f2 dimension. B. 3D HCcH-TOCSY correlates all the protons in a sugar ring. When

the chemical shift of f2 is positioned at C1' of U1915, all the protons in the sugar ring of U1915 show up at H1' chemical shift of f1 dimension. C. 3D HCcH-COSY correlates two protons on adjacent carbon atoms in a sugar ring. When the chemical shift of f2 is positioned at C2' of U1915, only H1', H2' and H3' resonances of U1915 show up at H2' chemical shift of f1 dimension.

while crosspeaks in this region are severely overlapping (Saenger 1984). With addition of chemical shift dimension of C3' or C4', most of those crosspeaks could be resolved to show clear contour intensity levels. This application is also essential to pinpoint the weak NOEs and UNNOEs, i.e. "unobserved NOEs", by employing the third dimension.

2.5.5 Structural restraints

To calculate the structures of an RNA oligonucleotide, restraint files must be generated from quantified geometric parameters determined from NMR experiments. Most commonly used restraint files include distance, dihedral angle, planarity, and Residual Dipolar Coupling (RDC) restraints.

2.5.5.1 Distance restraint file

Interproton distances and the distances between heavy atoms are usually included in the distance restraint file (.noe). 2D NOESY spectra of samples dissolved in D_2O with short mixing times (60, 90, 120, 150 ms) are used to generate interproton distances from integration of the peak volumes. Since the magnetization transfer through the NOE is inversely proportional to the sixth power of the interproton distance, it is possible to calculate the interproton distance in a 2D NOESY spectrum by comparing to the cross peak volumes of pyrimidine H6-H5, whose distance is fixed by covalent linkages in a pyrimidine ring, following Eq. 2.16:

$$r = Dist_{ref} \times \left[\frac{Intensity_{ref}}{Intensity_{obs}} \right]^{1/6} \qquad \text{Eq. 2.16}$$

where *Intensity*$_{ref}$ and *Intensity*$_{obs}$ are the measured peak integration from 2D NOESY, and *Dist*$_{ref}$ is the distance between pyrimidine H5 and H6, which is usually set to 2.41 Å (Varani and Tinoco 1991). An error bar is assigned to each calculated distance according to the intensity of the peak. Strong NOEs (i.e. 1.5 to 2.5 Å) are assigned errors of 0.5 Å, while weak NOEs (i.e. > 4 Å) are assigned errors of 1.0 Å.

Interproton distances involving exchangeable protons are derived from the hydrogen bonding pattern observed in the 2D NOESY carried out on samples dissolved in H_2O/D_2O (9:1). Once a base paring interaction is identified in the 2D NOESY, a distance with an error bar is imposed on the heavy atom pair and proton-acceptor pair involved in the hydrogen bonds of the base pair. In a GC base pair, the distances between the heavy atom pairs are 2.91, 2.71, and 3.08 Å for the N1-N3, O6-N4, and N2-O2 (donor-acceptor, Fig 1.3 D) respectively, with an error bar of 0.5 Å. In an AU base pair, the distances between the heavy atom pairs are 2.71 and 3.00 Å for the N3-N1 and N6-O4 (donor-acceptor, Fig 1.3 C) respectively, with an error bar of 0.5 Å. The distance between the donor proton and the acceptor heavy atom is 1 Å shorter than that of the corresponding heavy atom pair.

A third category of distance restraints, unobserved noes (UNNOEs), are always derived after each round of structure calculation to further refine the structures (Jaeger, SantaLucia *et al.* 1993; Vallurupalli and Moore 2003). If two protons are positioned in the calculated structures to be very close, e.g. 3 Å, while no cross peak between these two protons are observed in the 2D NOESY

with a mixing time of 400 ms, an additional distance restraint should be imposed on these two protons to push them apart to a distance of at least 4 Å. These UNNOEs are helpful in mining out more distance restraints in addition to the observed ones. However, while extra attention should be paid before assigning an UNNOE restraint, since factors, e.g. the chemical shifts of the protons being close to the water proton chemical shift, chemical exchange broadening, and overlap with noise can severely compromise the intensity of cross peaks.

2.5.2.2 Dihedral restraint file

Seven dihedral angles, including six backbone ones (α, β, γ, δ, ε and ζ) and one base orientation (χ), are used in RNA structure calculations (Figure 2.6). In an RNA hairpin structure, the stem and loop regions are usually considered separately, in terms of the dihedral angles, since RNA duplex usually assumes the A-form conformation, and each of the seven dihedral angles falls in a narrow range respectively, independent of the identity of the residues (Saenger 1984; Davis, Tonelli *et al.* 2005). In contrast, additional experimental data are essential to provide dihedral angles of the loop residues.

Due to the energy penalty, a base ring in RNA usually assumes *anti* conformation, with the dihedral angle of χ within -160 ± 40. In this case, a cross peak of medium to weak intensity between the base proton H8/H6 to the sugar proton H1' in the same residue is seen in the 2D NOESY with the mixing time of 60 ms. If an intense cross peak is observed, the base ring must be in a *syn* conformation, which positions the base proton H8/H6 right on top of the sugar

proton H1', and the dihedral angle χ should be restrained in the range of -70 ± 20.

The dihedral angle δ is determined by ^{31}P-decoupled 2D DQF-COSY (Altona 1982; Varani and Tinoco 1991). In A-form RNA, the dominant sugar conformation is C3' *endo*, and the orientations of H1' and H2' are close to perpendicular to each other. In this conformation, the three-bond magnetization transfer between the H1' and H2' is highly unfavorable, with a coupling constant less than 3 Hz, and as a result, the cross peak between the H1' and H2' is weak or missing in the spectrum. If an intense cross peak is observed in the 2D DQF-COSY spectrum between H1' and H2', the coupling constant would be larger than 7 Hz, and the sugar conformation is predominantly C2' *endo* (Hall 1995; Varani 1996). Sugar pucker conformations can be verified by 3D TOCSY-NOESY experiment (Wijmenga, Heus *et al.* 1994). An intense cross peak of H2'-H1'-H1' (f1/f2/f3) is a direct indication of C2' *endo*, and such peak should be absent if the sugar pucker is C3' *endo*. The restraint ranges of dihedral angle δ for C3' *endo* and C2' *endo* are 80 ± 20 and 157 ± 20, respectively.

Dihedral angle γ can be determined either by three-bond *J*-coupling between the H4'-H5'/H5" in ^{31}P-decoupled 2D DQF-COSY or from the NOE dipolar coupling between the H4'-H5'/H5" and H3'-H5'/H5" in 3D NOESY-HMQC or 3D TOCSY-NOESY. As a result of the "gauche effect", the *gauche-* γ conformation is the least favorable, and usually the dihedral angle γ is restrained in the range of 120 ± 120, to exclude the possibility of *gauche-* conformation (Saenger 1984; Chattopadhyay, Thibaudeau *et al.* 1999). In *gauche+* (60 ± 20)

conformation, the H4'-H5'/H5" cross peaks in the 2D DQF-COSY spectrum are absent, and cross peaks of equal intensity show up in the 3D NOESY-HMQC when the carbon chemical shift is set at the corresponding C4'. If the H4'-H5'/H5" cross peaks have unequal intensity, and both the cross peaks of H3'-H5' and H3'-H5" are intense in the 3D NOESY-HMQC, the dihedral angle γ should assume the *trans* conformation.

Dihedral angles ζ and α are determined by the downfield shift of ^{31}P chemical shift (Gorenstein 1981; Varani, Cheong *et al.* 1991). If the downfield shift is observed in the ^{31}P spectrum, both dihedral angles assume *trans* conformation, otherwise, both would be restrained to 0 ± 120. The ^{31}P-^{1}H HETCOR spectrum is used in determination of β and ε dihedral angles. If the dihedral angle β is *trans* conformation, the most commonly seen conformation in RNA, the small coupling constants (less than 5 Hz) would result in the absence of P-H5' and P-H5" cross peaks in the spectrum. Otherwise, dihedral angle β should be restrained to 180 ± 90 to correspond to the sterically allowed range (Altona 1982; Saenger 1984; Smith and Nikonowicz 1998). For dihedral angle ε, *trans* and *gauche-* are the major conformations, with the dihedral angle restrained within 150 ± 30, which can be verified by intense cross peaks between the P-H3' (n-1) (Varani and Tinoco 1991; Cabello-Villegas and Nikonowicz 2005).

2.5.6 Structure calculation

Crystallography & Nuclear Magnetic Resonance System (CNS) 1.2 is employed in solving structure families of RNA oligonucleotides by NMR-derived restraints in this thesis work (Brunger, Adams *et al.* 1998; Brunger 2007). The

molecular dynamics in CNS algorithms sample structural candidates that have the minimum global folding total energy, including covalent geometry, nonbonded contacts, and NMR restraints terms (Clore and Gronenborn 1989; Clore and Gronenborn 1998).

An extended structure is first generated by CNS 1.2 from the sequence of an RNA oligonucleotide, in ".pdb" format. This output file is then used as the input file in the following simulated annealing protocol, including one high temperature annealing, two slow cool-down annealing stages and one final minimization stage, to generate a family of structure candidates, with all NMR-derived restraints implemented. Torsion angle molecular dynamics and Cartesian coordinate space molecular dynamics are employed in the two slow cool-down annealing stages, respectively. In the first three stages, noe, dihedral, and van der Waals (repel) energy terms are all imposed in the molecular dynamic simulation. In the final minimization stage, base pair planarity restraints are taken into consideration, together with the noe and dihedral restraints, to further fine-tune the candidate structures.

Structure candidates generated in the first round of CNS calculation are chosen for further structural refinement by UNNOEs, if no NOE violation (> 0.5 Å) or dihedral angle violation (> 7.5°) is observed. The program RNA 1-2-3 (DNA software, Inc.) is used to estimate the convergence of CNS structure candidates by comparing the pair-wise RMSD in the same batch of results, and further optimize the structures by removing high energy penalty originating from steric clashes, incorrect covalent bonding, and incorrect hydrogen bonding, with the

Discrete Sampling of Torsion Angles (DSTA) algorithm (Saro and SantaLucia 2012).

CHAPTER 3

UV-MELTING STUDIES ON UNMODIFIED AND PSEUDOURIDYLATED HELIX 69 STEM DUPLEX OF 23S RIBOSOMAL RNA FROM E. COLI AND HUAMN

Sumita, M., J. Jiang, *et al*. (2012). "Comparison of solution conformations and stabilities of modified helix 69 rRNA analogs from bacteria and human." Biopolymers **97**(2): 94-106.

3.1 Structural and functional perspectives of Helix 69

As discussed in chapter 1.6, Helix 69 (H69) plays essential roles in ribosome association and protein synthesis. One striking structural feature of this short RNA hairpin structure is that three pseudouridylation modifications are clustered in the loop region on positions 1911, 1915, and 1917 (*E. coli* numbering), and this modification pattern is highly conserved in bacteria, archaea, and eukaryotes (Fig 1.15) (Ofengand and Bakin 1997). On association of a ribosome, H69 is brought close to the helix 44 (h44) of the small ribosomal RNA to establish the B2a intersubunit bridge. Even though the B2a subunit interface is highly dynamic during the translation process, the interaction with h44 is kept intact until the ribosome dissociates (Berk, Zhang *et al*. 2006; Weixlbaumer, Petry *et al*. 2007; Weixlbaumer, Jin *et al*. 2008; Schmeing, Voorhees *et al*. 2009; Jenner, Demeshkina *et al*. 2010; Korostelev, Zhu *et al*. 2010; Ratje, Loerke *et al*. 2010).

Research published by Dr. Chow's group shows that pseudouridylation does affect the structure and stability of H69, in both the oligonucleotide constructs and whole ribosomes (Meroueh, Grohar *et al*. 2000; Sumita,

Desaulniers *et al*. 2005; Abeysirigunawardena and Chow 2008; Sakakibara and Chow 2011). However, so far no research has been done to investigate the effect of pseudouridylation on the stem and loop region of H69 separately. To address the aforementioned fact, research included in this chapter provides a better understanding of thermodynamic effects of pseudouridylation on the stem and loop region of H69 in bacteria and Archaea/Eukaryotes.

3.2 Materials and Methods

3.2.1 Materials

Chemical reagents were purchased from Thermo Fisher Scientific (Waltham, MA). Seven RNA oligonucleotides, synthesized on a scale of 200 nmol with deprotection, were purchased from Dharmacon (Thermo Scientific). Sep-pak columns and HPLC column were purchased from Waters (Milford, MA).

3.2.2 RNA purification

Seven RNA oligonucleotides (Fig 2.2) were subjected to HPLC purification on a Waters Xterra MS C18 column to remove any potential impurities. A gradient of acetonitrile from 4 to 7% in 25 mM triethylammonium acetate (TEAA) buffer (pH 6.5) was employed to purify the RNAs and separate the heptmer oligonucleotides from the hexmer oligonucleotides after UV-melting experiments on duplexes, at a flow rate of 3 mL/min in 30 min. The RNA fractions were further dried and desalted with Sep-Pak columns in the following step, and characterized with MALDI-tof for molecular mass verification.

3.2.3 UV-melting experiment

An Aviv 14DS UV-vis spectrometer equipped with a thermocoupler was used to monitor the absorbance versus temperature cruves. Five microcuvettes of four different volumes, 60, 120, 300, and 480 μL, with a light path length of 1, 2, 5, and 8 mm, respectively, were used to the maximize the concentration range of an RNA duplex monitored under one wavelength. The buffer condition was 15 mM NaCl, 20 mM sodium cacodylate, and 0.5 mM Na_2EDTA, pH 7.0. The concentration of each sample was determined by absorbance (260nm) measured at 90°C. The extinction coefficient of 66500 M^{-1} cm^{-1} was used for *E. coli* duplexes, and 62800 M^{-1} cm^{-1} for human duplexes (Watkins and SantaLucia 2005). Two UV-melting macros (5 to 95 °C at 260 nm and 0 to 90 °C at 280 nm) were run on samples of 20 to 400 μM and 10 μM, respectively. Since the melting point of human unmodified duplex was relatively lower, only the most concentrated sample (394 μM) was run from 5 to 95 °C at 280 nm and all the other samples were run from 0 to 90 °C at 260 nm. The combination of four light path lengths with two absorbance wavelengths expands the range of sample concentrations to a factor of 40, without compromise in the absorbance range. This is essential to improve the accuracy of a van Hoff's equation fitting model. MeltWin 3.5 was employed to derive the thermodynamic parameters from the UV-melting curves (Figure 3.1), and with the assumption of a two-state model, a van't Hoff equation (Eq. 2.2) was used in the linear regression as an alternative method for parameter quantification (McDowell and Turner 1996). Errors from each of the two aforementioned methods were employed in the calculation of error-weighted average (Bevington 1969; Bommarito, Peyret *et al*. 2000).

Eds

Hds
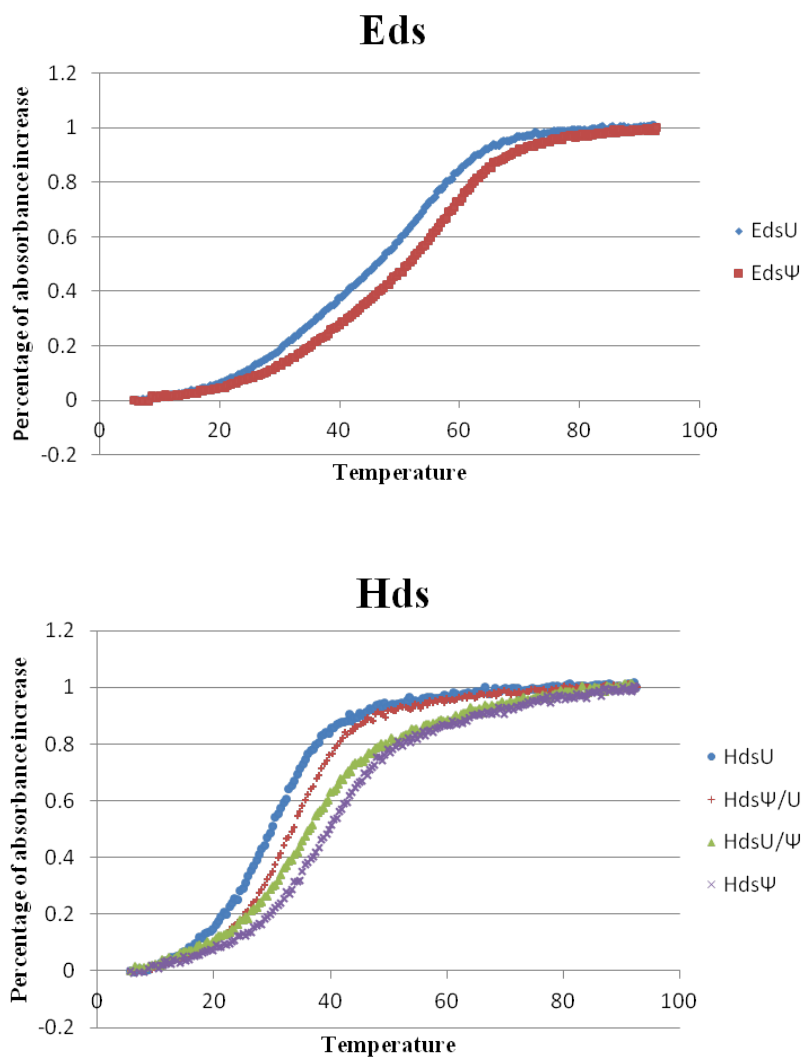
Figure 3.1 UV-melting curves shown above were acquired at 280 nm from the most concentrated samples of each duplex, where the total RNA strand concentration is about 400 µM, with double normalization. E, *E. coli*; H, human, ds, double strand; U, unmodified; Ψ, modified; Ψ/U, modified in the upper strand and unmodified in the lower strand; U/Ψ, unmodified in the upper strand and modified in the lower strand.

3.3 Results and Discussion

When the lengths of the upper strands (RNA oligonucleotide sequences close to the 5' end of the H69 sequence) and lower strands (RNA oligonucleotide sequences close to the 3' end of the H69 sequence) are compared (Figure 2.2), it is readily observed that one extra A is residing on the 3' terminus of the upper strand, following the U/Ψ1911. The reason for that is, solid phase chemical synthesis of RNA oligonucleotide has to start from the 3' terminus, while no Ψ-modified solid support is currently available (Meroueh, Grohar *et al*. 2000; Chui, Desaulniers *et al*. 2002). To keep the unmodified and modified upper strands consistent, A1912 of the H69 (Figure 2.2) was introduced into the upper strand. It has been reported that a dangling A does provide extra stabilizing effect, most likely by stacking interactions, on RNA duplexes (Freier, Burger *et al*. 1983; O'Toole, Miller *et al*. 2005). While the extra stabilizing effect was assumed to be consistent throughout all the duplex combinations, given the fact that only when a Ψ-modification is on the 5' terminus, it has a significant stabilizing effect through stacking than U (Davis 1995). As a favorable byproduct, this design made it possible to recover the RNA oligonucleotides after UV-melting experiments by HPLC. A low salt buffer was used in this study, since direct comparison to previously published data of the hairpin constructs is required to evaluate the thermodynamic effect of pseudouridylation modifications on the loop residues (Meroueh, Grohar *et al*. 2000; Sumita, Desaulniers *et al*. 2005).

Raw data of $\Delta H°$, $\Delta S°$, $\Delta G°_{37}$, and T_m, together with the average and error, were generated by MeltWin 3.5 from UV-melting curves of different
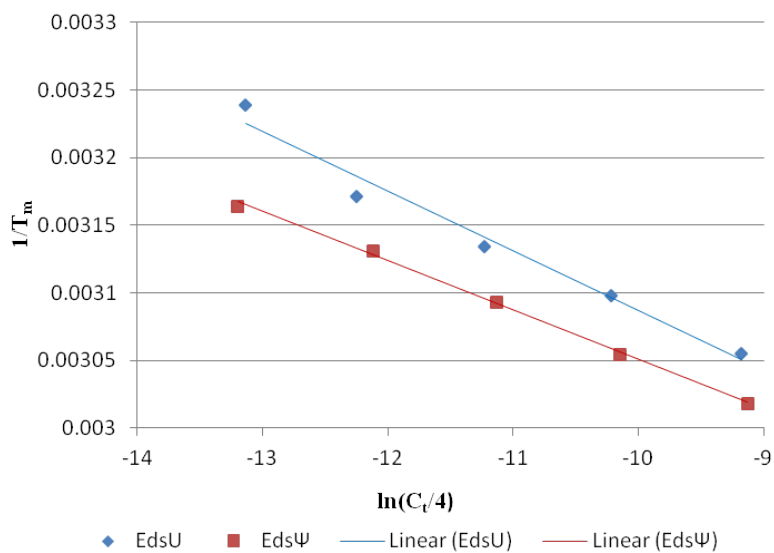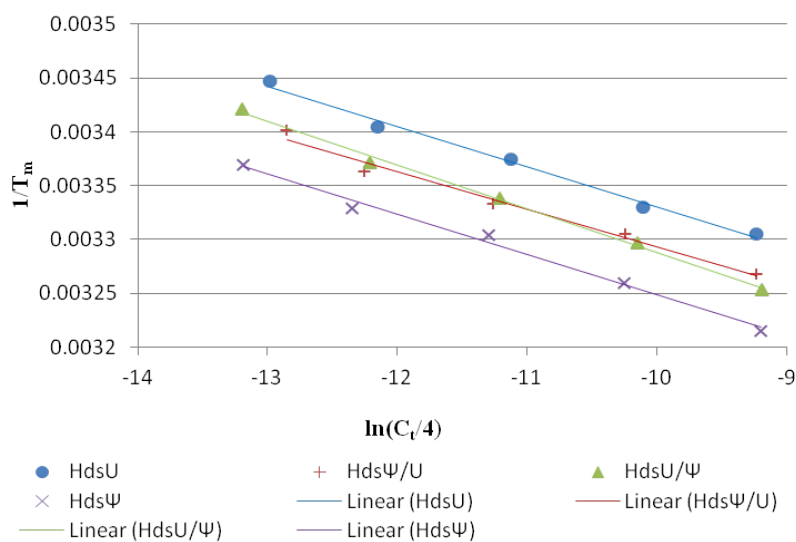
Figure 3.2 A van Hoff's equation fitting of T_m^{-1} vs. natural logarithm of $C_t/4$, where T_m is the melting temperature obtained directly from MeltWin 3.5 curve fitting, and the C_t is the total single-stranded RNA concentration.

concentrations for each set of duplex combination. Then the T_ms and the total single strand concentrations were used as input for a van Hoff's equation fitting to obtain an alternative set of thermodynamic parameters ($\Delta H°$, $\Delta S°$, and $\Delta G°_{37}$) (Figure 3.2). First, $\Delta H°$s from the two methods were compared, and all pairs of $\Delta H°$s showed agreement with each other within a 10% deviation, which indicats that the transition was from a duplex to random coil and the two-state model assumption is valid (Freier, Burger *et al*. 1983). The average and error of $\Delta H°$, $\Delta S°$, and $\Delta G°_{37}$ for each sample set were calculated from the two sets of data obtained with different methods mentioned above. All the thermodynamic parameters are shown in Table 3.1.

Table 3.1 shows that pseudouridylation modification in both *E. coli* and human H69 stem duplexes contributes to the stabilization of the double strand formation. In *E. coli* H69 duplexes, unmodified double strand (EdsU) has a $\Delta G°_{37}$(ds) of -8.1 ± 0.1 kcal/mol, while the modified double strand (EdsΨ) has a more favorable $\Delta G°_{37}$(ds) of -9.1 ± 0.1 kcal/mol. In human H69 duplexes, since there are four possible double strand modification combinations, e.g. unmodified, modification in the upper strand only, modifications in the lower strand only, and the completely modified (modifications in both strands), four $\Delta G°_{37}$(ds)s are shown in Table 3.1. Complete modification duplex (HdsΨ) shows the most favorable $\Delta G°_{37}$(ds) of -5.7 ± 0.1 kcal/mol, followed by the HdsΨ/U ($\Delta G°_{37}$(ds) of -4.9 ± 0.1 kcal/mol) and HdsU/Ψ (($\Delta G°_{37}$(ds) of -5.2 ± 0.1 kcal/mol), where only one of the two strands has the modification, and the unmodified duplex has the least favorable $\Delta G°_{37}$(ds) of -4.3 ± 0.1 kcal/mol.

Table 3.1: Thermodynamic parameters of helix 69 stem-region (ds) RNAs

RNA	$\Delta G°_{37}$(ds) (kcal/mol)	$\Delta H°$ (kcal/mol)	$\Delta S°$ (e.u.)	T_m (°C)
Average of curve fittings				
EdsU	-8.1 ± 0.1	-45.6 ± 2.9	-120.7 ± 9.0	48.3
EdsΨ	-9.2 ± 0.1	-55.3 ± 4.5	-148.7 ± 14.4	52.5
HdsU	-4.1 ± 0.2	-58.5 ± 3.1	-175.4 ± 10.4	24.8
HdsΨ	-5.7 ± 0.1	-54.6 ± 3.2	-157.7 ± 10.3	32.5
HdsΨ/U	-4.9 ± 0.1	-59.6 ± 2.4	-176.4 ± 7.6	28.5
HdsU/Ψ	-5.1 ± 0.1	-51.3 ± 2.4	-148.8 ± 7.6	28.8
A van Hoff's equation fittings				
EdsU	-8.1± 0.2	-45.0 ± 4.1	-119.1 ± 12.9	48.0
EdsΨ	-9.1 ± 0.1	-54.5 ± 1.3	-146.4 ± 4.0	52.3
HdsU	-4.4 ± 0.1	-53.4 ± 2.7	-157.9 ± 9.1	25
HdsΨ	-5.8 ± 0.1	-53.3 ± 3.0	-153.3 ± 9.7	32.5
HdsΨ/U	-4.9 ± 0.2	-56.8 ± 4.1	-167.4 ± 13.6	28.5
HdsU/Ψ	-5.2 ± 0.1	-48.9 ± 1.8	-140.9 ± 5.9	28.7
Average of the two methods				
EdsU	-8.1 ± 0.1	-45.3 ± 2.4	-119.9 ± 7.4	48.2
EdsΨ	-9.2 ± 0.1	-54.9 ± 1.2	-147.6 ± 3.8	52.4
HdsU	-4.3 ± 0.1	-52.9 ± 2.0	-166.6 ± 8.9	24.9
HdsΨ	-5.7 ± 0.1	-54.0 ± 2.2	-155.5 ± 7.1	32.5
HdsΨ/U	-4.9 ± 0.1	-58.2 ± 2.1	-171.9 ± 6.7	28.5
HdsU/Ψ	-5.2 ± 0.1	-50.1 ± 1.4	-144.9 ± 4.7	28.8

To calculate the $\Delta G°_{37}$ of the unmodified or modified loop, $\Delta G°_{37}$ of the unmodified or completely modified duplex was subtracted from that of the corresponding RNA hairpin construct (Meroueh, Grohar *et al*. 2000; Sumita, Desaulniers *et al*. 2005). For the *E. coli* loop segment, the $\Delta G°_{37}$ of the unmodified or modified loops are 3.3 ± 0.1 kcal/mol (ElpU) and 4.1 ± 0.1 kcal/mol (ElpΨ), respectively. For human loop segment, the $\Delta G°_{37}$ of the unmodified or modified loop is 1.6 ± 0.1 kcal/mol (HlpU) and 1.9 ± 0.1 kcal/mol (ElpΨ), respectively. A positive $\Delta G°_{37}$ of loop in both *E. coli* and human H69 means the loop segment has destabilizing effect on the folding of the H69 sequences into a hairpin structure.

To further quantify the stabilizing/destabilizing effect of pseudouridylation modification(s) on the H69 stem and loop segment conformations, the difference between the $\Delta G°_{37}$s of a modified (complete or incomplete) segment and the corresponding unmodified segment is calculated. The equations and results are shown in Table 3.2.

In a previous study on thermodynamic properties of selectively modified *E. coli* H69 sequences, it is reported that H69 (Ψ1911) ($\Delta G°_{37}$, -5.9 kcal/mol) is more stable than unmodified H69 ($\Delta G°_{37}$, -4.9 kcal/mol) by -1.0 kcal/mole (Meroueh, Grohar *et al*. 2000). These results agree well with the stabilizing effect of pseudouridylation modification on *E. coli* H69 stem duplex ($\Delta\Delta G°_{37}$, -1.0 kcal/mol) reported in this research.

In human H69 stem duplex sequences, when the pseudouridylation modification is in the upper strand only, the stabilizing effect is -0.6 kcal/mol (-4.9

Table 3.2: Stabilizing /destabilizing effect of pseudouridylation modification(s) on

the stem and loop segments of H69 from *E. coli* and human.

Modification location	Equation	$\Delta\Delta G°37$ *(kcal/mol)*
E. coli H69		
Stem	$\Delta\Delta G°37$ (Es) = $\Delta G°37$ (EdsΨ) - $\Delta G°37$ (EdsU)	-1.0 ± 0.1
Loop	$\Delta\Delta G°37$ (El) = $\Delta G°37$ (ElpΨ) - $\Delta G°37$ (ElpU)	0.8 ± 0.1
Human H69		
Upper strand	$\Delta\Delta G°37$ (HΨ/U) = $\Delta G°37$ (HdsΨ/U) - $\Delta G°37$ (HdsU)	-0.6 ± 0.1
Lower strand	$\Delta\Delta G°37$ (HU/Ψ) = $\Delta G°37$ (HdsU/Ψ) - $\Delta G°37$ (HdsU)	-0.9 ± 0.1
Both strands	$\Delta\Delta G°37$ (HΨ) = $\Delta G°37$ (HdsΨ) - $\Delta G°37$ (HdsU)	-1.4 ± 0.1
Loop	$\Delta\Delta G°37$ (Hl) = $\Delta G°37$ (HlpΨ) - $\Delta G°37$ (HlpU)	0.3 ± 0.1

- (-4.3) kcal/mol), and when the pseudouridylation modifications are in the lower strand only, the stabilizing effect is -0.9 kcal/mol (-5.2 - (-4.3) kcal/mol). Even though there is one more pseudouridylation site in the lower strand compared to that in the upper strand, the difference between the stabilizing effect of pseudouridylation in the upper strand and the lower strand is only -0.3 kcal/mol (-0.9 - (-0.6) kcal/mol), less than -0.6 kcal/mol. This finding indicates that the pseudouridines stabilize the Human H69 stem duplex differently. When the stabilizing effect of the single pseudouridylation in the upper strand are considered, it is more significant in *E. coli* stem duplex (-1.0 kcal/mol, Es in Table 3.2) than that in Human H69 stem duplex with the modification only in the upper strand (-0.6 kcal/mol, HΨ/U in Table 3.2). This result suggests that pseudouridylation modification does stabilize the stem formation in both cases, while the feedback from *E. coli* H69 stem is more favorable than that from the human H69 stem, due to the sequence difference. When there are pseurouridylation modifications in both the upper and lower strand, human H69 stem duplex obtain an extra -1.4 kcal/mol in $\Delta G°_{37}$. This value is very close to the sum of stabilizing contributions of pseudouridylation modifications in the upper and lower strand individually (-0.9 +(-0.6) kcal/mol = -1.5 kcal/mol), while it is not necessarily true that stabilizing effect is additive. Since a -1.4 kcal/mol in $\Delta G°_{37}$ is almost within error from a -1.0 kcal/mol in $\Delta G°_{37}$ as the case of *E. coli* H69 stems duplex, the conclusion would be complete pseudouridylation has similar stabilizing contributions to the H69 stem duplex in both *E. coli* and human.

Thermodynamic parameters of the H69 loop region from *E. coli* and human were derived from experimental data of complete hairpin structure and the stem duplexes, therefore the stabilizing/destabilizing effect of pseudouridylation modifications on 1915 and 1917 (*E. coli* numbering, 3731 and 3733 for human) in the loop region on the hairpin conformations were determined by comparing the modified and unmodified loop $\Delta G°_{37}$ (Table 3.2). It is shown in previous discussion that loop sequences without pseudouridylation modifications have a destabilizing effect on the hairpin conformations in both *E. coli* and human H69 constructs, and this destabilizing effect is further enhanced by pseudouridylations, to different extent. *E. coli* H69 loop pseudouridylations show more prominent destabilizing effect (0.8 kcal/mol, Table 3.2) than human loop pseudouridylations (0.3 kcal/mol, Table 3.2). This difference can not be explained simply by a difference in loop region sequences of *E. coli* (A1918) and human (G1918), since previous studies on H69 hairpin thermodynamics of *E. coli* wild type (A1918) and the A1918G mutant shows that $\Delta G°_{37}$ s of the two constructs are essentially the same (-4.8 kcal/mol for the wild type *E. coli* H69 hairpin and -4.7 kcal/mol for the A1918G mutant). Thus the origin of this difference could be extended to the context of the whole hairpin, including the stem duplex. We hypothesize that the loop sequences are "framed" onto the stem duplexes, and the conformations of the stem duplexes do affect the thermodynamic effect of pseudouridylation modifications on the loop conformation.

Dr. Sumita in Dr. Chow's group performed Circular Dichroism (CD) experiments on the *E. coli* and human H69 stem duplexes with and without
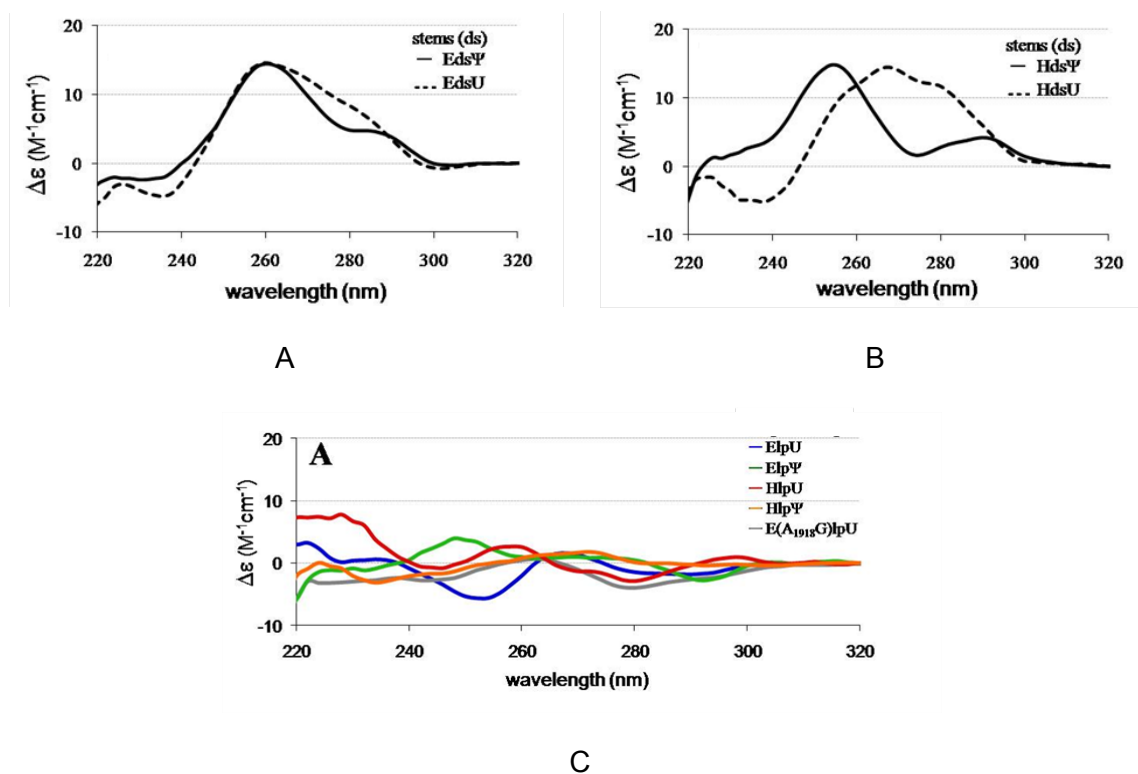
Figure 3.3 CD spectra of H69 stem duplexes and loop regions of *E. coli* and human. A. An overlay of the CD spectra from modified (EdsΨ) and unmodified (EdsU) *E. coli* H69 stem duplexes. B. An overlay of the CD spectra from completely modified (HdsΨ) and unmodified (HdsU) human H69 stem duplexes. C. An overlay of CD spectra from loop region derived from corresponding hairpin CD spectra and stem duplex spectra.

pseudouridylation modifications (Sumita, Jiang *et al*. 2012). In Figure 3.3 A and B, it is clearly shown that the CD spectra of the four stem duplexes are quite different. The CD spectra of EdsU and EdsΨ share a peak of about the same intensity at 260 nm, while EdsU has a more prominent shoulder peak at 284 nm and EdsΨ has a more prominent trough at 238 nm. In contrast, the peak of HdsΨ is blue-shifted to 254 nm, and the peak of HdsU is red-shifted to 268 nm, with barely any overlay between the two curves. The different features of these four CD spectra argue that conformational differences exist between the four stem duplexes when folded. The differences in stem duplex conformations impose differential feedbacks to the loop region pseudouridylation modifications, perhaps either by different covalent connectivity geometry, or different extensive stacking from the stem duplex base pairs, or both. This feedback effect can be seen in the overlay of CD spectra of loop regions (Figure 3.3 C), where all the CD spectra are different from each other indicating different conformations in the loop region.

Conformational differences observed in crystal structures of H69 from bacteria and eukaryotes support the hypothesis that thermodynamic effect of psudouridylation modifications on conformations of H69 is affected by the sequence context of the whole hairpin structure, and might be helpful to understand the difference in destabilizing effect of pseudouridylation modifications in loop residues on the H69 hairpin formation. Four crystal structures were chosen from Protein Data Bank, e.g. 1NKW, 2I2T, 4A18, and 3U5D (Harms, Schluenzen *et al*. 2001; Berk, Zhang *et al*. 2006; Ben-Shem, Garreau de Loubresse *et al*. 2011; Klinge, Voigts-Hoffmann *et al*. 2011). 1NKW

is the crystal structure of the large ribosomal subunit from *Deinococcus radiodurans*, and this structure is representative for H69 in a free bacterial large ribosomal subunit (pre-association of the 70S ribosome). H69 sequence in this structure ranges from residue 1889 to residue 1907, with identical pseudouridylation modifications (Ofengand and Bakin 1997; Harms, Schluenzen *et al*. 2001). 2I2T is the crystal structure of the 70S ribosome, containing both the small and large subunits, from *E. coli*, and this structure is representative for H69 in an assembled bacterial large ribosomal subunit (post-association of the ribosome). H69 sequence in this structure ranges from residue 1906 to residue 1924 (Berk, Zhang *et al*. 2006). 4A18 is the crystal structure of the large ribosomal subunit from *Tetrahymena thermophila*, and this structure is representative for H69 in a free eukaryotic large ribosomal subunit (pre-association of the ribosome). H69 sequence in this structure ranges from residue 2244 to residue 2262. Even though the pseudouridylation modification mapping on the ribosome of this species has not been done, this structure is helpful in that an estimation of the conformational effect of pseudouridylation can be obtained, given the fact that that two pseudouridylation sites are conserved in eukaryotic ribosome H69 loop region (Ofengand and Bakin 1997; Klinge, Voigts-Hoffmann *et al*. 2011). 3U5D is the crystal structure of the 80S ribosome, containing both the small and large subunits, from *yeast*, and this structure is representative for H69 in an assembled eukaryotic large ribosomal subunit (post-association of the ribosome). H69 sequence in this structure ranges from residue 2249 to residue 2267, and the pseudouridylation modifications are determined to exist in both
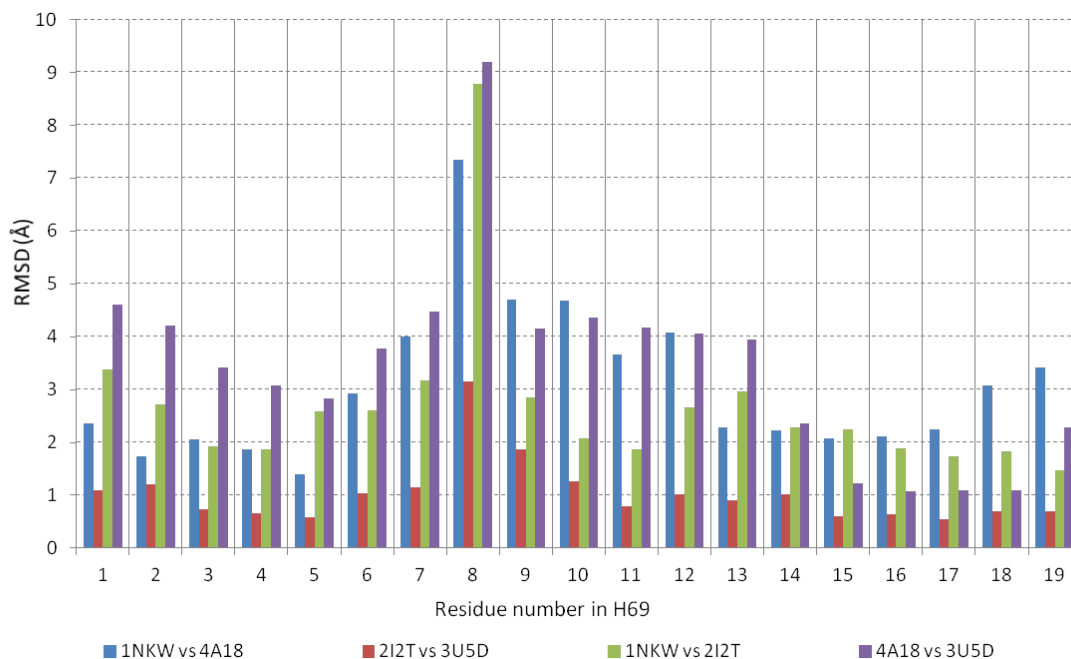
Figure 3.4 Pair-wise RMSD between the four H69 crystal structures (1NKW, 2I2T, 4A18, and 3U5D). The RMSD of each residue was calculated when the two structures were superimposed onto each other with the minimum global RMSD. The RMSD could be resulted from transition of the residue propagated along the structure and/or dihedral angle differences of each residue. This analysis provides an estimation of how different the local conformations are in the pair and where the differences are distributed.
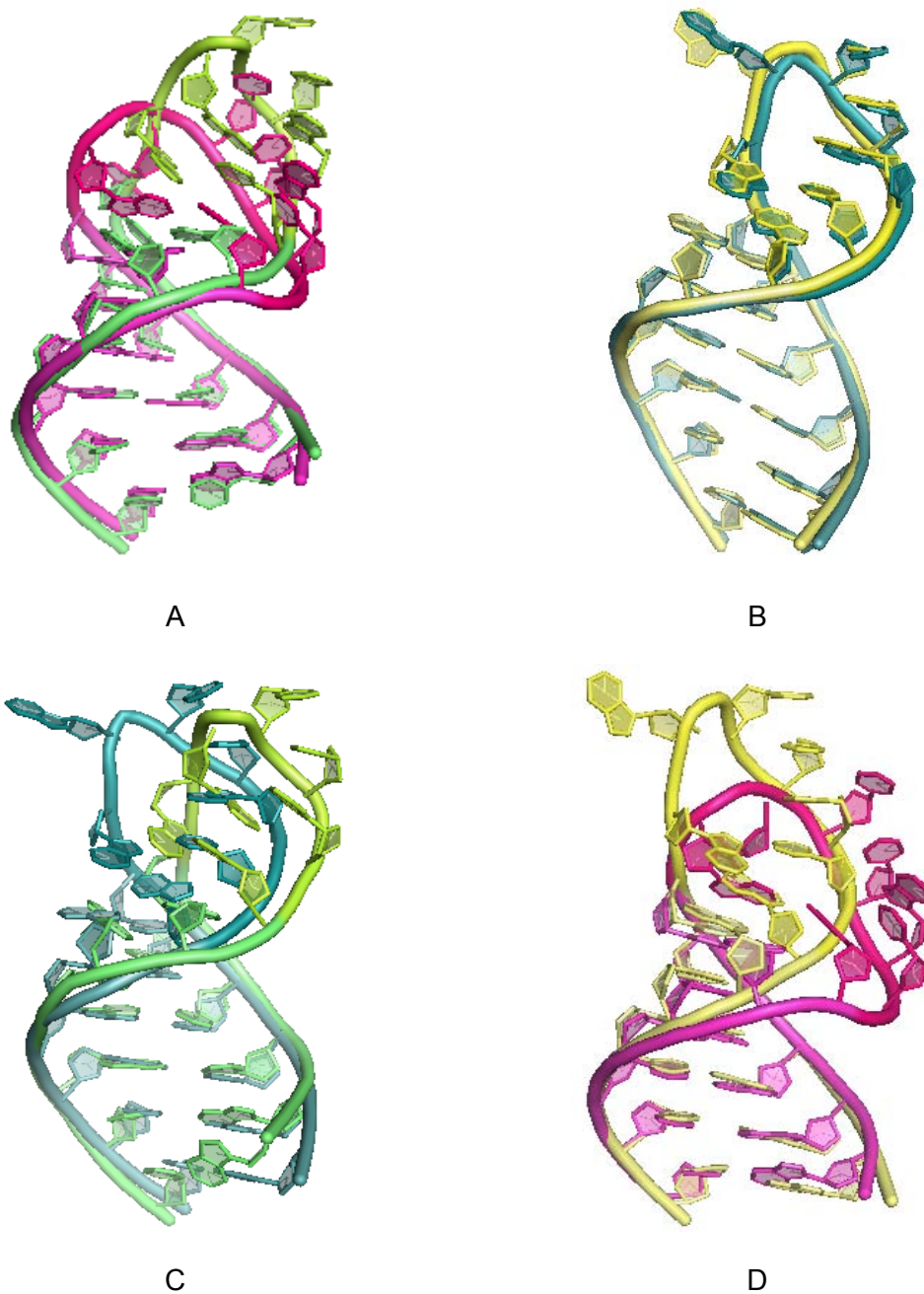
Figure 3.5 Pair-wise comparison of conformations between the four H69 crystal structures (1NKW, 2I2T, 4A18, and 3U5D). Four structures were aligned with phosphorus atoms on residues 2, 10, 16, and 19. A. 1NKW versus 4A18. B. 2I2T versus 3U5D. C. 1NKW versus 2I2T. D. 4A18 versus 3U5D.

loop residues (Ψ2258 and Ψ2260, corresponding to Ψ1915 and Ψ1917 in *E. coli*) (Ofengand and Bakin 1997; Ben-Shem, Garreau de Loubresse *et al*. 2011). Sequences outside of H69 were removed from the pdb file, and hydrogen atoms were added with RNA 1-2-3 (Saro and SantaLucia 2007). Then the global RMSD and residue RMSD of the four H69 structures were compared pairwise.

When structures of H69 in large subunit (pre-association) of bacteria (1NKW) and eukaryotes (4A18) are compared, the global RMSD is 3.07 Å, with a local RMSD of 3.74 Å for the loop region only. In Figure 3.4, it is clearly shown that the blue bars form a high plateau from residue 7 to 12, with a significant peak value on residue 8 (corresponding to residue A1913 in *E. coli* numbering). When the RMSD was calculated, the conformations of both the loop regions and stems were taken into consideration. If the two structures were aligned using only the stem residues, then the RMSD of loop residues would be even larger. This trend is readily distinguished in Figure 3.5 A, where the loop structures are significantly different, while the stem duplexes are almost perfectly superimposed onto each other. In the assembled ribosome, the conformational differences among H69 structures from bacteria (2I2T) and eukaryotes (3U5D) is much smaller, with a global RMSD of 1.03 Å and a local RMSD of 1.38 Å for the loop region only. The only distinguished RMSD value in Figure 3.4 red bar series is on residue 8 (corresponding to residue A1913 in *E. coli* numbering). This is also shown in Figure 3.5 B, where the backbone of two structures agree well with each other, and the only structural difference feature seen is the orientation of the base of residue 8. Structural similarities are usually coupled to the conserved

functionalities. In the associated ribosome, H69 in both bacteria and eukaryotes participates in the formation of intersubunit bridges, and may play similar roles throughout translation process. When conformational changes on association of the whole ribosome are considered in both bacteria (1NKW versus 2I2T, Figure 3.4 green bar series, Figure 3.5 C) and eukaryotes (4A18 versus 3U5D, Figure 3.4 purple bar series, Figure 3.5 D), extensive conformational rearrangements of global structures of H69, especially the loop region, are observed. The global RMSD for H69 structures of bacteria in the pre- and post-associationstates is 2.68 Å, with a local RMSD of 3.82 Å for the loop region, and those for the eukaryotes are 3.44 Å and 4.10 Å, respectively. Residue 8 is still the motion center. The RMSD values indicate that, on association of the ribosome, eukaryotic H69 might undergo more extensive conformational rearrangement than bacterial H69, especially the loop region.

This finding could be potentially interpreted as, on association of a ribosome, either in bacteria or eukaryotes, H69 undergoes a rearrangement process. Since the destabilizing effect of H69 loop from bacteria (3.3 kcal/mol) is stronger than that of eukaryotes (1.6 kcal/mol), which is further enhanced by the loop pseudouridylation modifications (4.1 kcal/mol for bacteria versus 1.9 kcal/mol for eukaryotes), less structural rearrangement would be required for the bacterial H69 loop on association of the ribosome than eukaryotic H69 loop.

In residue RMSD analysis, residue 8 (corresponding to *E. coli* A1913) has the largest RMSD throughout all the four ribosomes in the comparison. It has been reported that conformation of the *E. coli* H69 oligonucleotide with

pseudouridylation modifications is also sensitive to pH changes, while unmodified H69 oligonucleotide does not show this property (Abeysirigunawardena and Chow 2008). This pH-sensitive conformational change has been attributed to the change of A1913 base orientation by monitoring the fluorescence intensity change of 2-aminopurine 1913 along the pH titration, which is determined by the stacking of 2-aminopurine base with neighboring residues (Abeysirigunawardena 2008). Chemical probing experiments carried out on the *E. coli* large ribosomal subunit also shows that A1913 is more sensitive to pH changes when H69 was pseudouridylated by RluD (Sakakibara and Chow 2011).

3.4 Conclusion

The thermodynamic properties of unmodified and modified H69 constructs from bacteria and human are determined in this study. The results show that pseudouridylation modifications in the stem duplex region contribute a stabilizing effect in both bacterial and human H69. Pseudouridylation(s) in the upper and lower strand of human H69 stem duplex may work synergistically to stabilize the stem conformation. Pseudouridylations in the loop region of bacterial H69 have significant destabilizing effect, while this effect is marginal in the case of human H69 loop. This difference might be correlated to the differential conformational changes of H69, especially the loop region, on ribosome association in bacteria and eukaryotes. It is shown in this project that pseudouridylation modification(s) can exert both stabilizing effect and destabilizing effects on RNA structures, and the effects are affected by the sequence/structural context where the pseudouridylation is located. The conclusion and data generated in this project

are helpful to understand the thermodynamic effects of pseudouridylation on local RNA structures, and explore drug candidate binding into a Ψ-containing RNA region. The data could be applied in future research on thermodynamic properties of H69 mutants, especially functionally important loop residue mutants, to establish the correlation between contributions from a residue or a motif to the thermodynamic properties of H69 and the phenotypes with severely compromised ribosome activity or antibiotic resistance.

CHAPTER 4

NMR STUDIES ON SOLUTION STRUCTURES OF UNMODIFIED AND

PSEUDOURIDYLATED HELIX 69 OF 23 S RIBOSOMAL RNA FROM E. COLI

4.1 Introduction

Ribosomes, composed of ribosomal RNAs (rRNAs) and proteins (rproteins) in two discrete subunits, are universally conserved as protein biosynthesis machinery in all living cells (Yusupov, Yusupova *et al.* 2001; Schuwirth, Borovinskaya *et al.* 2005; Berk, Zhang *et al.* 2006; Weixlbaumer, Petry *et al.* 2007; Weixlbaumer, Jin *et al.* 2008; Schmeing, Voorhees *et al.* 2009; Ben-Shem, Jenner *et al.* 2010; Jenner, Demeshkina *et al.* 2010; Korostelev, Zhu *et al.* 2010; Ratje, Loerke *et al.* 2010; Ben-Shem, Garreau de Loubresse *et al.* 2011). Association of the two ribosomal subunits during the dynamic translation process is stabilized by intersubunit bridges (Frank, Verschoor *et al.* 1995; Yusupov, Yusupova *et al.* 2001; Schuwirth, Borovinskaya *et al.* 2005; Ben-Shem, Jenner *et al.* 2010). Intersubunit bridge B2a in bacterial ribosomes is established by interactions between the helix 44 (h44) of the 16S rRNA in the small subunit and helix 69 (H69) of the 23S rRNA in the large subunit (Yusupov, Yusupova *et al.* 2001). This configuration is conserved in eukaryotic ribosomes (Ben-Shem, Jenner *et al.* 2010). Deletion of H69 in bacterial ribosomes causes defects in ribosome association and peptide release, mainly due to disruption of the B2a intersubunit bridge (Ali, Lancaster *et al.* 2006). H69 potentially plays other important roles in the translation process, and it is observed in crystal structures that H69 forms direct contacts to multiple translational factors, including the A-

site and P-site tRNAs, elongation factor G (EF-G), release factors 1 and 2 (RF1 and RF2), and ribosome recycling factor (RRF), in slightly different conformations (Weixlbaumer, Petry *et al.* 2007; Weixlbaumer, Jin *et al.* 2008; Schmeing, Voorhees *et al.* 2009; Jenner, Demeshkina *et al.* 2010; Korostelev, Zhu *et al.* 2010; Ratje, Loerke *et al.* 2010). In addition to its central localization, H69 shows high conservation in the primary and secondary structures, and post-transcriptional modification pattern through at all the three domains (Ofengand and Bakin 1997; Cannone, Subramanian *et al.* 2002).

One distinguishing chemical feature of H69 is that multiple post-transcriptional modifications, including pseudouridylation (Ψ) and base methylation, are concentrated in this small 19mer stem-loop region (Ofengand and Bakin 1997). In *Escherichia coli* ribosomes, residues 1911, 1915, and 1917 in H69 are all pseudouridylated, with one methyl group attached to the N3 position of Ψ1915 ($m^3\Psi$). The modification pattern is slightly different in *Thermus thermophilus*, where U1936 (corresponding to Ψ1915 in *E. coli*) is methylated on the base, without pseudouridylation (mU) (Mengel-Jorgensen, Jensen *et al.* 2006). In *Homo sapiens* 28S ribosomal rRNA H69, all the three aforementioned pseudouridylation sites are conserved (3727, 3731, and 3733 using *H. sapiens* numbering), with two additional pseudouridylation sites in the stem region (3737 and 3739 using *H. sapiens* numbering) and one 2'-*O*-methylation on loop residue A3729 (Am3729 using *H. sapiens* numbering) (Ofengand and Bakin 1997; Baudin-Baillieu, Fabret *et al.* 2009). The only difference between the H69 modification patterns from *H. sapiens* and *Saccharomyces cereviciae* is that

U2254 (corresponding to Ψ1911 in *E. coli* and Ψ3727 in *H. sapiens*) in *S. cereviciae* H69 is not pseudouridylated (Baudin-Baillieu, Fabret *et al.* 2009). RNA modifications are capable of fine-tuning RNA structures and dynamics. Pseudouridylation contributes to stabilization of local structures by enhanced stacking and one additional water-mediated hydrogen bond involving HN1 (Davis 1995; Auffinger and Westhof 1997; Charette and Gray 2000; Newby and Greenbaum 2002; Newby and Greenbaum 2002). It has been reported that 2'-*O*-methylation boosters A-form conformation in single-stranded RNA regions, and methylation on RNA bases favors base stacking in single-stranded RNA regions, potentially by increased London dispersion or hydrophobicity, and blocks formation of base pairing involving the modified imino or amino groups (Rife and Moore 1998; Blanchard and Puglisi 2001; Chow, Lamichhane *et al.* 2007). Research addressing the biological significance of pseudouridylation in *E. coli* H69 indicate that mutations of the pseudouridylation sites (Ψ1915 and Ψ1917) or A1916, which affects the RluD activity of pseudouridylation on Ψ1915 and Ψ1917, and deletion of RluD result in slow growth phenotypes associated with defects in ribosome association and translation efficiency (Liiv, Karitkina *et al.* 2005; Leppik, Peil *et al.* 2007). Growth advantage with pseudouridylation on Ψ2258 and Ψ2260 (corresponding to Ψ1915 and Ψ1917 in *E. coli*) in *S. cerevisiae* was also reported (Badis, Fromont-Racine *et al.* 2003). Recently, different conformational behaviors of H69 with and without pseudouridylation were observed in *in vitro* chemical probing studies on *E. coli* ribosomes and ribosomal subunits at different conditions (Sakakibara and Chow 2011;

Sakakibara and Chow 2012). These observations suggest that pseudouridylation has structural effects on H69 folding, which may be involved in biological phenomena.

Crystal structures of H69 have been visualized in both isolated large ribosomal subunits from *Deinococcus radiodurans* and *Tetrahemena thermophila* and 70S or 80S ribosomes from *E. coli*, *Thermus thermophilus*, and *S. cerevisiae* (Harms, Schluenzen *et al.* 2001; Berk, Zhang *et al.* 2006; Weixlbaumer, Petry *et al.* 2007; Ben-Shem, Garreau de Loubresse *et al.* 2011; Klinge, Voigts-Hoffmann *et al.* 2011). The most striking difference in H69 structures on association of ribosomal subunits is that A1913 from H69 (using *E. coli* numbering) flips out to form direct interactions with h44 of the small subunit, when all other loop residues participate in two base stacking systems separated by A1913. H69 post-transcriptional modification patterns in *D. radiodurans* and *E. coli* are identical (Del Campo, Recinos *et al.* 2005). The modification map of H69 from *T. thermophila* is not currently available. In the H69 from *S. cerevisiae*, pseudouridylation is missing on U2254 (corresponding to Ψ1911 in *E. coli*), and there is a 2'-*O*-methylation on A2256 (Am2256, corresponding to A1913 in *E. coli*) instead (Ofengand and Bakin 1997; Baudin-Baillieu, Fabret *et al.* 2009). All the post-transcriptional modifications listed above are theoretically correlated to stabilization of local base stacking interactions in single-stranded loop region of H69 (Davis 1995; Auffinger and Westhof 1997; Rife and Moore 1998; Charette and Gray 2000; Blanchard and Puglisi 2001; Newby and Greenbaum 2002; Chow, Lamichhane *et al.* 2007). While *in vitro* model studies in solution show that

pseudouridylation in *E. coli* H69 has differential effects on the thermodynamics of stem-loop RNA folding. Ψ1911 has a stabilizing effect, while Ψ1915 and/or Ψ1917 have destabilizing effects (Meroueh, Grohar *et al.* 2000; Sumita, Jiang *et al.* 2012). To address the structural basis for these effects, Nuclear Magnetic Resonance (NMR) spectrometry was employed in this study. NMR has been used to study the structural effects of pseudouridylation on tRNA anticodon stem-loop, splicesomal RNA and telomerase RNA, where either subtle or significant conformational changes were observed on modification at atomic level (Durant and Davis 1999; Newby and Greenbaum 2002; Cabello-Villegas and Nikonowicz 2005; Kim, Theimer *et al.* 2010; Denmon, Wang *et al.* 2011). In this thesis, NMR was used to solve the solution structures of *E. coli* H69 with pseudouridines (H69ΨΨΨ) and without (H69UUU) modification (Figure 4.1), and the direct comparison between the two structures are used to provide insights into the differential effects of pseudouridylation on the thermodynamics and stem-loop structure.

4.2 Materials and Methods

4.2.1 Materials

Unlabeled and ^{13}C, ^{15}N- labeled nucleotide 5'-triphosphates (NTPs) were purchased from Sigma-Aldrich®. T7 RNA polymerase was prepared as described (Davanloo, Rosenberg *et al.* 1984). Template and primer DNA sequences for *in vitro* transcription were purchased from IDT®. Modified H69 RNA oligonucleotides (H69ΨΨΨ), were purchased from Dharmacon (Thermo Scientific). Sep-pak columns and HPLC column were purchased from Waters.
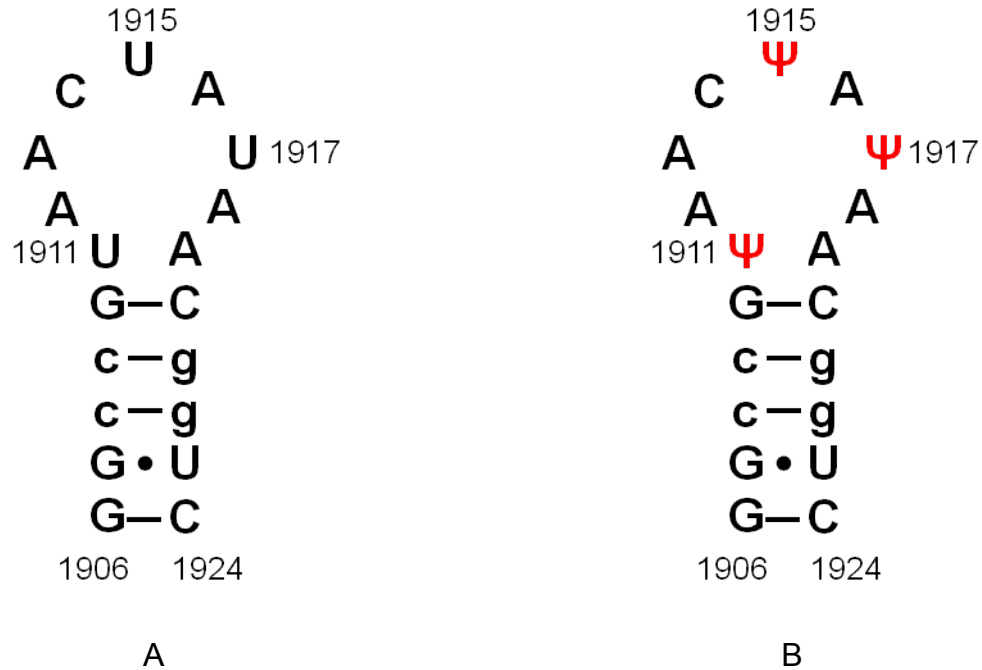
Figure 4.1 Comparison of the secondary structures of unmodified (H69UUU, A) and pseudouridylated (H69ΨΨΨ, B) 23S rRNA H69 in *E. coli*. Letters in upper case and lower case show >90% and ≥88% conserved nucleotide residues in bacteria, respectively (Cannone, Subramanian *et al.* 2002). Pseudouridylated nucleotide residues in *E. coli* H69 are shown in red (right), and the *E. coli* 23S rRNA numbering is used.

4.2.2 Preparation of the unmodified H69 RNA (H69UUU)

The unmodified H69 RNA samples were synthesized by *in vitro* T7 RNA polymerase transcription with unlabeled or ^{13}C, ^{15}N- labeled NTPs, and synthesized template and primer DNA sequences (Wyatt, Chastain *et al.* 1991). Full length H69 RNA transcripts were separated by electrophoresis through a denaturing 20% (w/v) preparative polyacrylamide gel (PAGE), and electroeluted in a 0.2 X TBE buffer with a Schleicher and Schuell® Elutrap. RNA-containing fractions from the elutrap were desalted with a Sep-pak® reverse phase chromatography cartridge, and deionized water was used to wash the column. A 30% acetonitrile/H_2O (v/v) solution was used to elude the RNA, and the eluted fractions were pooled and lyophilized to a powder.

4.2.3 Preparation of the modified H69 RNA (H69ΨΨΨ)

Synthesized H69ΨΨΨ RNA oligonucleotides were subjected to HPLC purification on a Waters Xterra MS C18 column. A gradient of acetonitrile from 6.0 to 7.8% in 25 mM triethylammonium acetate (TEAA) buffer (pH 6.5) was employed to purify the RNA oligonucleotides, at a flow rate of 3 mL/min in 24 min. The RNA-containing fractions were lyophilized and desalted with a Sep-Pak column.

4.2.4 Preparation of RNA NMR samples

Purified H69 oligonucleotides (H69UUU and H69ΨΨΨ) were dissolved in a buffer containing 10 mM potassium phosphate, 50 mM KCl, pH 7.3, and 0.1 mM Na_2-EDTA to a volume of 300 µL. The samples were lyophilized to a powder. D_2O (99.9%) was used to exchange the residual H_2O in the powder twice by

lyophilization, and dissolve the sample in 99.99% D_2O (Cambridge Isotope Labs) to a final volume of 300 µL for NMR experiments on non-exchangeable protons. A H_2O/D_2O (90%/10%) mixture was used to dissolve the sample to a final volume of 300 µL for NMR experiments on exchangeable protons. All the NMR samples contained 0.8-1.0 mM RNA oligonucleotides, and trace amount of 3-(trimethylsilyl) propionate (TSP) was used as an internal proton chemical shift reference.

4.2.5 NMR spectroscopy

All NMR experiments were carried out on a Bruker Avance 700 MHz spectrometer equipped with a HCN cryoprobe and a Varian Mercury 400 MHz equipped with a room temperature QXI probe. Spectra for exchangeable proton resonance assignments and base pairing identification were acquired at 288 K and 298 K. All other spectra of H69UUU and H69PPP samples were acquired at 298 K and 310 K, respectively. The Topspin 2.1 (Bruker) and Sparky 3.114 (University of California, San Francisco, CA, USA) were used for spectral processing and analysis.

2D NOESY spectra of the unlabeled H69 RNA samples (H69UUU and H69ΨΨΨ) dissolved in 99.9% D_2O were initially analyzed to assign the resonances of base protons (H8/H6) and sugar protons (H1') (Varani and Tinoco 1991). For the unmodified H69 RNA samples (H69UUU), a combination of 2D NOESY, 2D DQF-COSY, 2D ^{13}C-^{1}H HMQC, and 3D TOCSY-NOESY run on the unlabeled sample, and 2D ^{13}C-^{1}H CT-HSQC, 3D HCcH-COSY, 3D HCcH-TOCSY, 3D HCCh-TOCSY, and 3D NOESY-HMQC run on the ^{13}C, ^{15}N- labeled

Table 4.1 Resonance assignments of protons in H69UUU

	H6/H8	H5/H2	H1'	H2'	H3'	H4'	H5'	H5''
G1906	8.219	N/A	5.887	4.883	4.750	4.622	4.501	4.333
G1907	7.605	N/A	5.967	4.744	4.555	4.586	4.573	4.310
C1908	7.807	5.506	5.589	4.347	4.559	4.469	4.585	4.148
C1909	7.760	5.542	5.558	4.623	4.563	4.440	4.543	4.127
G1910	7.521	N/A	5.730	4.509	4.493	4.496	4.486	4.129
U1911	7.532	4.955	5.587	4.572	4.483	4.447	4.543	4.120
A1912	8.119	7.509	5.760	4.614	4.631	4.544	4.427	4.203
A1913	7.918	7.909	5.558	4.290	4.572	4.253	4.426	4.099
C1914	7.390	5.414	5.530	4.041	4.421	4.012	4.183	3.948
U1915	7.617	5.788	5.794	4.210	4.463	4.308	3.852	3.918
A1916	8.280	7.899	5.915	4.832	4.842	4.492	4.133	4.252
U1917	7.773	5.743	5.961	4.418	4.728	4.554	4.308	4.248
A1918	8.206	7.911	5.960	4.812	4.673	4.633	4.399	4.297
A1919	8.021	7.717	5.706	4.601	4.563	4.548	4.469	4.225
C1920	7.408	5.160	5.458	4.477	4.421	4.429	4.438	4.108
G1921	7.545	N/A	5.748	4.716	4.559	4.497	4.491	4.130
G1922	7.244	N/A	5.760	4.656	4.387	4.485	4.528	4.117
U1923	7.771	5.435	5.570	4.147	4.540	4.413	4.536	4.077
C1924	7.827	5.706	5.936	4.043	4.281	4.162	4.553	4.054

Table 4.2 Resonance assignment of carbon atoms in H69UUU

	C6/C8	C5/C2	C1'	C2'	C3'	C4'	C5'
G1906	136.4	N/A	89.17	72.61	71.95	80.70	64.42
G1907	134.5	N/A	90.59	72.57	70.16	79.78	63.1
C1908	137.9	94.72	91.13	73.00	69.16	79.21	61.22
C1909	138.0	95.40	91.22	72.78	69.80	79.13	62.67
G1910	133.5	N/A	90.42	72.64	70.54	79.47	63.29
U1911	138.4	100.0	90.63	72.85	70.19	79.60	62.13
A1912	137.3	151.6	90.00	72.59	71.71	80.66	63.78
A1913	137.1	152.2	89.23	73.11	70.76	80.29	62.71
C1914	139.5	95.46	88.10	73.50	73.85	81.54	64.04
U1915	140.5	102.9	87.10	72.90	74.80	82.00	65.00
A1916	139.2	152.3	87.70	72.95	73.20	81.90	64.80
U1917	140.7	102.9	87.96	72.97	74.10	82.40	65.00
A1918	138.4	151.8	88.83	73.16	73.22	81.35	65.33
A1919	137.3	150.8	89.98	72.75	70.82	79.98	63.26
C1920	137.8	94.78	90.93	72.78	69.99	79.23	62.43
G1921	138.4	N/A	90.09	72.68	70.40	79.28	63.18
G1922	133.2	N/A	90.43	72.49	69.87	79.40	62.24
U1923	137.8	101.3	91.45	73.01	69.25	79.72	61.14
C1924	140.4	95.25	89.86	74.77	67.17	80.62	62.49

Table 4.3 Resonance assignments of protons and carbon atoms in H69ΨΨΨ

	C6/C8	C5/C2	C1'	H6/H8	H5/H2	H1'	H2'	H3'	H4'	H5'	H5''
G1906	136.2	N/A	90.13	8.086	N/A	5.795	4.816	4.613	4.406	4.079	3.966
G1907	134.5	N/A	90.58	7.581	N/A	5.936	4.728	4.602	4.572	4.604	4.228
C1908	138.0	94.87	91.18	7.799	5.511	5.585	4.344	4.562	4.470	4.588	4.155
C1909	138.1	95.97	91.31	7.758	5.553	5.572	4.607	4.565	4.459	4.555	4.143
G1910	133.8	N/A	90.19	7.580	N/A	5.758	4.666	4.601	4.582	4.511	4.138
Ψ1911	137.0	n.d.	n.d.	6.978	N/A	4.713	4.492	4.436	4.253	4.445	4.063
A1912	137.3	151.8	90.00	8.064	7.527	5.776	4.569	4.670	4.550	n.d.	n.d.
A1913	137.3	152.2	89.59	7.891	7.875	7.657	4.391	4.516	4.403	n.d.	n.d.
C1914	140.0	95.15	69.68	7.445	5.501	5.588	4.319	4.421	4.276	n.d.	n.d.
Ψ1915	139.9	n.d.	n.d.	7.535	N/A	4.673	4.310	4.400	4.189	n.d.	n.d.
A1916	138.5	152.7	88.76	8.217	8.083	6.019	4.821	4.508	4.390	n.d.	n.d.
Ψ1917	138.8	n.d.	n.d.	7.295	N/A	4.625	4.282	4.540	4.293	n.d.	n.d.
A1918	138.6	152.2	88.80	8.239	8.043	5.928	4.827	4.680	4.423	n.d.	n.d.
A1919	137.3	150.8	90.28	8.059	7.693	5.649	4.626	4.547	4.534	4.457	4.233
C1920	137.9	94.80	91.13	7.377	5.123	5.449	4.457	4.441	4.517	4.446	4.119
G1921	133.6	N/A	80.07	7.569	N/A	5.766	4.708	4.602	4.497	4.516	4.137
G1922	133.3	N/A	90.47	7.234	N/A	5.758	4.657	4.399	4.499	4.521	4.107
U1923	138.0	n.d.	91.44	7.782	5.439	5.569	4.174	4.546	4.434	4.557	4.096
C1924	140.4	95.40	89.99	7.855	5.720	5.938	4.053	4.280	4.182	4.549	4.080

n.d. = not determined

sample was used to assign resonances of proton and carbon atoms (Altona 1982; Griffey, Poulter *et al.* 1983; Varani and Tinoco 1991; Pardi and Nikonowicz 1992; Kay, Xu *et al.* 1993; Nikonowicz and Pardi 1993; Wijmenga, Heus *et al.* 1994; Dieckmann and Feigon 1997; Tjandra and Bax 1997). For the modified H69 RNA sample (H69ΨΨΨ), 2D NOESY, 2D DQF-COSY, 2D ^{13}C-^{1}H HMQC, and 3D TOCSY-NOESY were used for resonance assignments. 1D ^{31}P and 2D ^{31}P-^{1}H HETCOR were run on the unlabeled H69 samples to identify any shift of ^{31}P resonances (H69UUU and H69ΨΨΨ) (Gorenstein 1981; Sklenar, Miyashiro *et al.* 1986). Resonance assignments of H69UUU and H69ΨΨΨ are shown in Tables 4.1, 4.2 and 4.3.

4.2.6 Structure calculation

Interproton distance restraints of non-exchangeable protons were derived from 2D NOESY spectra with a short mixing time ($τ_m$ = 120 ms for the H69UUU sample and $τ_m$ = 150 ms for the H69ΨΨΨ sample), by integrating the volume under each crosspeak, and the average of pyrimidine H6-H5 crosspeaks was used as the standard corresponding to an interproton distance of 2.45 Å (Varani and Tinoco 1991). Interproton distance restraints of exchangeable protons were only applied to residues involved in base pairs in the stem region (from G1906•C1924 to G1910•C1920). The loop closing base pairs (U1911•A1919 for H69UUU and Ψ1911•A1919 for H69ΨΨΨ) were left unrestrained because neither U/Ψ1911 NOE crosspeak to A1919 nor hydrogen bond between the two residues was observed (Sigel, Sashital *et al.* 2004). Unobserved NOEs (unNOEs) were employed in structure refinement (Jaeger, SantaLucia *et al.*

1993; Vallurupalli and Moore 2003). Briefly, if the crosspeak between two protons was not observed in the 2D NOESY (τ_m = 400 ms) or 3D NOESY-HMQC (for the ^{13}C, ^{15}N- labeled sample of H69UUU), the two protons were restrained to be at least 4 Å apart.

The dihedral angles α, β, γ, δ, ε, ζ, and χ of residues in the stem region (from G1906-C1924 to G1910-C1920) of H69UUU and H69ΨΨΨ were restrained according to the A-form RNA helix geometry (-62 ± 20°, 172 ± 20°, 60 ± 20°, 84 ± 20°, -160 ± 20°, -71 ± 20°, and -160 ± 20°, respectively). Loop residue (including U/Ψ1911 and A 1919) dihedral angles α and ζ were restrained (0 ± 120°) to exclude a trans-conformation, due to absence of a downfield shifted ^{31}P resonance in H68UUU and H69ΨΨΨ 1D ^{31}P spectra (Gorenstein 1981; Varani, Cheong *et al.* 1991). Since no crosspeaks between P-H5' or P-H5" were readily observed, the β dihedral angles of loop residues were loosely restrained to -180 ± 90° to cover both A-form and B-form structure with a deviation of 50° on both sides of the range (van Dijk and Bonvin 2009). The γ dihedral angles of loop residues in H69UUU were derived from 3D NOESY-HMQC experiment on the ^{13}C, ^{15}N- labeled sample. If the H4'-H5' and H4'-H5" crosspeaks were of equal intensity, and H3'-H5" crosspeak was more intense than that of H3'-H5', the γ dihedral angle was restrained to be 60 ± 60°, otherwise, the dihedral angles were restrained to be 120 ± 120° to exclude *gauch-* conformation. Crosspeaks from J-coupling of H1' and H2' in the 2D DQF-COSY were used to determine the δ dihedral angles of loop residues, where an intense crosspeak comparable to the pyrimidine H6-H5 crosspeaks indicated a

C2'-endo conformation and the δ dihedral angle was restrained to 157 ± 40°. If the sugar H1'-H2' crosspeak was very weak to invisible, a 84 ± 20° restraint was used to restrain the δ dihedral angle to a *C3'-endo* conformation; otherwise, it was left unrestrained (Hall 1995; Varani, Aboulela *et al.* 1996). The ε dihedral angles of loop residues were restrained to be -120 ± 120° to exclude the *gauche+* conformation. Due to absence of intense crosspeaks between base protons H8/H6 and sugar protons H1' in the 2D NOESY spectra with a short mixing time (τ_m = 60 ms) from H69UUU and H69ΨΨΨ, the χ dihedral angles of loop residues were all restrained to -110 ± 110° to exclude a *syn* conformation. One additional dihedral angle restraint of 180 ± 45° was applied on H1'-C1'-C5-C6 of Ψ1915 in H69ΨΨΨ, if a crosspeak of $^4J_{H1'-H6}$ of moderate intensity was observed in the 2D DQF-COSY spectrum of H69ΨΨΨ.

The Crystallography and NMR System (CNS) 1.2 employing a simulated annealing and restrained molecular dynamics (rMD) protocol was used in structure calculations (Brunger, Adams *et al.* 1998). An extended structure containing only the RNA sequence and covalent linkages of the RNA construct generated using the CNS 1.2 was subjected to a torsion angle molecular dynamics at 20,000 K for 40,000 steps (1 fs/step, 40 ps in total) with molecular dynamic scale factors of 50 kcal/molÅ^2 and 150 kcal/molrad^2 for the distance restraints and dihedral angle restraints, respectively. Molecular dynamics pseudopotential force constants of 100 kcal/molÅ^2 and 250 kcal/molrad^2 were then applied to distance and dihedral restraints, respectively, in the first round of slow cooling down (50,000 steps in 100 ps) to 0 K. In the second round of slow

cooling down (10,000 steps in 35 ps) from 2,000 K to 0 K, a Cartesian molecular dynamics simulation was employed (using pseudopotential force constants 200 kcal/molÅ^2 for the distance restraints and 500 kcal/molrad^2 for the dihedral restraints). A force constant of 50 kcal/molÅ^2 was used for the planarity restraints in the final energy minimization stage consisting of 800 steps for 20 rounds (using pseudopotential force constants 300 kcal/molÅ^2 for the distance restraints and 700 kcal/molrad^2 for the dihedral restraints). Qualified output structures (i.e. with no distance violation > 0.5 Å or dihedral violation > 7.5°) from this global fold were chosen for further torsion angle molecular dynamics refinement. Structure calculation restraints and statistics are shown in Table 4.4.

4.3 Results

4.3.1 NMR spectroscopy of H69UUU and H69ΨΨΨ

In the NMR experiments, two 19mer RNA oligonucleotides corresponding to the H69 (G1906 - C1924) of *E. coli* 23S rRNA were studied. One of the oligonucleotides has all common ribonucleotides, i.e. A, C, U, and G, and pseudouridine (Ψ) was introduced into positions 1911, 1915, and 1917 of the modified oligonucleotide, representing the pseudouridylated H69 stem-loop. More structural restraints were derived for H69UUU structure calculation due to the availability of the ^{13}C, ^{15}N- labeled sample. It is difficult to incorporate ^{13}C, ^{15}N- labeled pseudouridines into the H69ΨΨΨ loop region selectively by *in vitro* transcription catalyzed by T7 RNA polymerase, so only a natural isotopic abundance sample was employed in the NMR experiments for H69ΨΨΨ.

Table 4.4 Summary of structure calculation restraints and statistics of H69UUU

and H69ΨΨΨ

Restraints	H69UUU	H69 ΨΨΨ
NOE distance restraints (total / per residue)	266 / 14.0	214 / 11.3
Intraresidue	112 / 5.9	76 / 4.0
Interresidue	154 / 8.1	138 / 7.3
unNOEs	196	48
Stem restraints[a]	51	51
Base flipping control[b]	16	16
Stem stacking control[c]	6	6
Carbon distance[d]	10	10
Base pair distance[e]	14	14
Base pair planarity	5	5
Dihedral angle restraints (total / per residue)	128 / 6.7	127 / 6.7
Statistics		
Violations		
NOE distance violation > 0.5 Å[f]	0	0
Dihedral angle violation > 7.5 °[g]	0	0
Average pairwise RMSD of all atoms (Å)[h]	1.34	1.23

[a]Five base pairs from G1906•C1924 to G1910•C1920 were considered as the stem region for both H69UUU and H69ΨΨΨ, and distance restraints derived from 1R2P were used together with the A-form RNA and spectrum-derived NOEs to restrain the G1907•U1923 wobble base pair.

[b]A deviation of 0.5 Å was applied on the distance restraints of base excyclic heavy atoms to allow the conformational adjustment of the base pair stacking in the stem region.

[c]A deviation of 0.5 Å was applied on the distance restraints of base imino protons to allow the conformational adjustment of the base pair stacking in the stem region.

[d]A deviation of 0.2 Å was applied on the distance restraints of base H8/H6 proton pair and sugar H1' proton pair within one base pair to allow the conformational adjustment within one base pair in the stem region.

[e]A deviation of 0.2 Å was applied on the distance restraints of heavy atom pairs involved in a hydrogen bond within one base pair to allow the conformational adjustment within one base pair in the stem region.

[f]An NOE violation of 0.5 Å is of an energy penalty of 10 kcal/mol.

[g]A dihedral angle violation of 7.5 ° corresponds to an energy penalty of 8.6 kcal/mol.

[h]A family of ten lowest energy structures from H69UUU and H69ΨΨΨ without NOE violation > 0.5 Å or dihedral angle violation of 7.5 ° were compared, respectively, using the program RNA-123 (Saro and SantaLucia 2007)

Resonances of imino protons involved in Watson-Crick base pairs are unambiguously assigned from the 2D NOSEY spectra of H69UUU and H69ΨΨΨ, except that of the hairpin closing base pair G1906•C1924, which undergoes rapid exchange with the solvent at 298K (Figure 4.2). Formation of G1907•U1923 wobble base is evidenced by appearance of one upfield shifted imino proton resonance from each of the two nucleotides, and the intense crosspeak of the two imino proton resonances. Two medium-weak crosspeaks from G1922 N1H to G1921 N1H, G1907 N1H and U1923 N3H are also shown in the same spectrum, which suggests that G1907•U1923 wobble base pair is formed in an A-form RNA structural context. Only one upfield shifted imino proton resonance from N1H of Ψ1911 is shown to be in close proximity to H6 of Ψ1911 (Figure 4.8 D). While all other imino protons from the pseudouridine residues are not visible in the spectra at 298K, suggesting that none of them are involved in hydrogen bonding interactions. It is highly possible that N1H of Ψ1911 is protected from exchange with solvent by a water-mediated hydrogen bond (Newby and Greenbaum 2002).

Comparison of the base proton H8/6 to sugar proton H1' regions of 2D NOESY spectra from H69UUU and H69ΨΨΨ shows that chemical shifts and crosspeaks of the stem region protons involved in the "walk" are very similar (crosspeaks connected by green lines) and more differences are observed in the loop region, especially the upfield shifted chemical shifts of pseudouridine H1' compared to U H1', due to a more less deshielding C5 covalently bond to the C1' in the pseudouridine, instead of an N1 in U. Other resonances showing
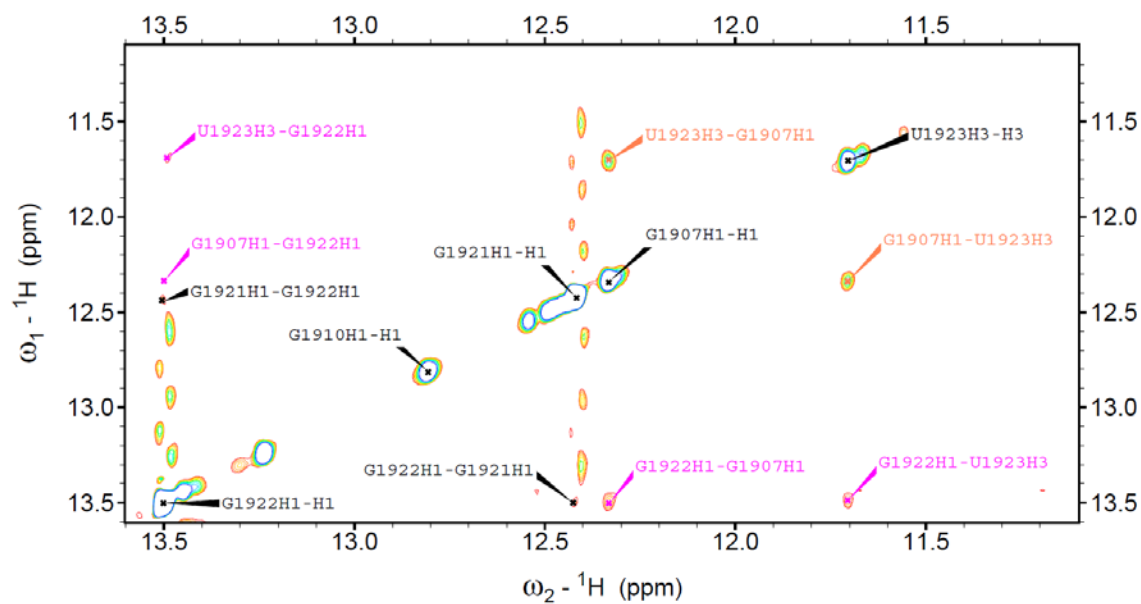
A
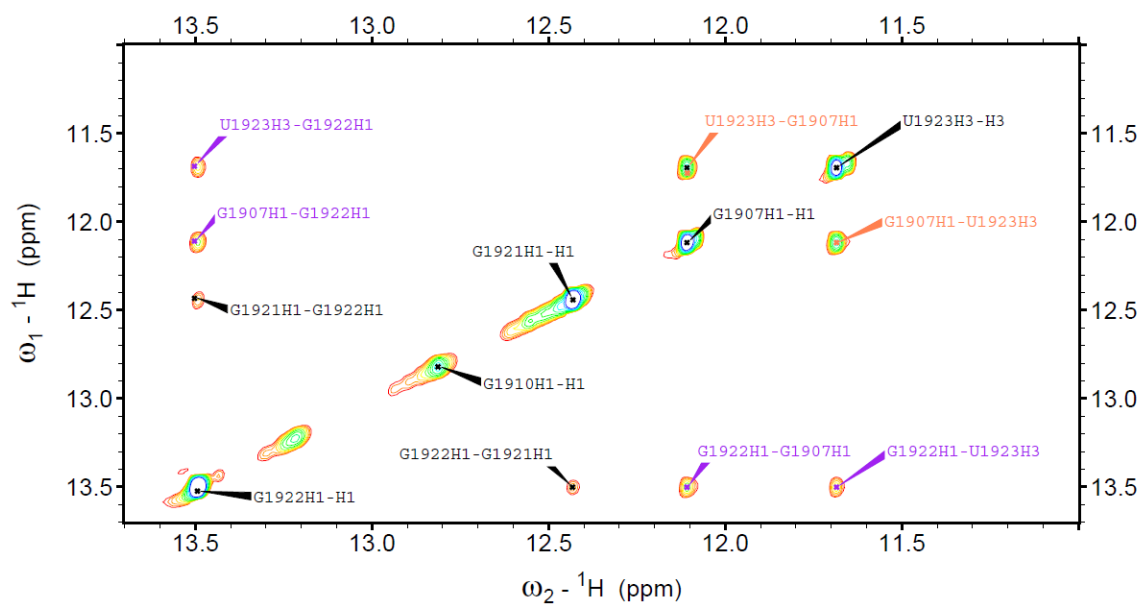
B

Figure 4.2 Comparison of the imino proton region of the 2D NOESY (H_2O/D_2O 90%/10%) spectra of H69UUU (A) and H69ΨΨΨ (B). A strong NOE between G1907 N1H and U1923 (cross diagonal assignments in orange) and two

medium-weak NOEs between G1922 N1H and G1907 N1H/U1923 N3H (cross diagonal peaks in purple) are observed in the spectra. A weak NOE is also observed between G1922 N1H and G1921 N1H, indicating that formation of G1907•U1923 wobble base pair does not disrupt the general A-form RNA structure of the stem regions. 2D NOESY (H_2O/D_2O 90%/10%) spectra of H69ΨΨΨ (B) was done by Dr. Aduri.

significant chemical shift change (δA_{H69UUU}- $\delta A_{H69\Psi\Psi\Psi}$ in ppm) include Ψ1911 H6 (0.554), Ψ1917 H6 (0.478) (Figure 4.3), A1916 H2 (-0.184), and A1918 H2 (-0.132) (Figure 4.4), indicating a subtle conformational change in the A1916-A1918 region.

In the 2D DQFCOSY spectrum of H69UUU (Figure 4.5 A), all the nine H6-H5 crosspeaks are observed in the H6-H5 region. Due to pseudouridylation at positions 1911, 1915, and 1917, three crosspeaks from the modified residues are missing in the H69$\Psi\Psi\Psi$ 2D DQFCOSY (Figure 4.5 C). Another distinguishing feature of this spectrum is that the H6-H5 crosspeak of C1914 H6-H5 is much less intense than all other H6-H5 crosspeaks, and a big blurry crosspeak of C1914 H6-H5, instead of a well defined doublet, is seen in the 2D NOESY spectrum of H69$\Psi\Psi\Psi$ (Figure 4.3 B), indicating a local dynamic conformation with intermediate exchange time regime. In the H1'-H2' region of the 2D DQFCOSY spectrum of H69UUU (Figure 4.5 B), strong crosspeaks from H1'-H2' of C1914, U1915, and U1917 are observed. This observation helped restrain the δ dihedral angles of the three residues to 157° ± 40°, typical for a *C2'-endo* conformation. The δ dihedral angles of Ψ1915, A1916, and Ψ1917 in the H69$\Psi\Psi\Psi$ sample were left unrestrained, since only weak crosspeaks of H1'-H2' of these residues were observed (Figure 4.5 D and E) and the sugar pucker conformation may undergo exchange. One additional $^4J_{HH}$-*coupling* crosspeak of Ψ1915 H6-H1' shows up in the H69$\Psi\Psi\Psi$ 2D DQFCOSY (Figure 4.5 D and E), which can only be explained by a close to co-planar configuration of the H6-C6-C5-C1'-H1' covalent system. If the C6-H6 and C1'-H1' bonds are on

A

B
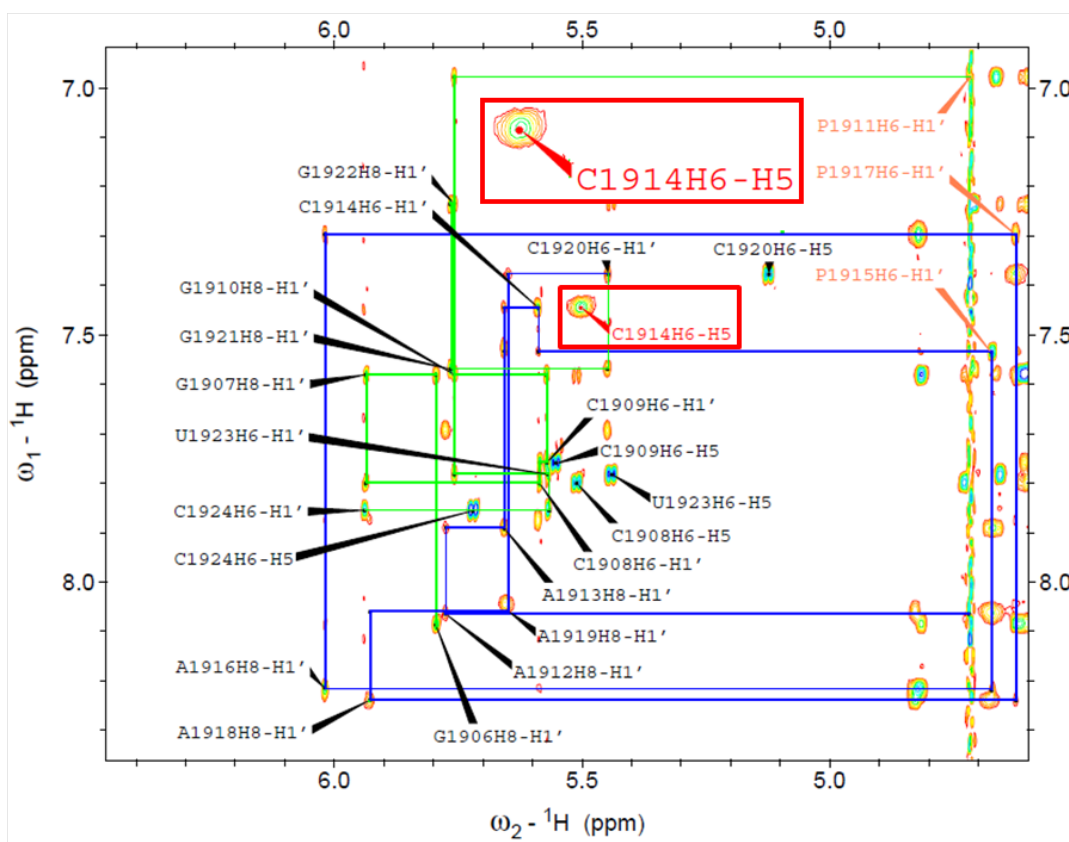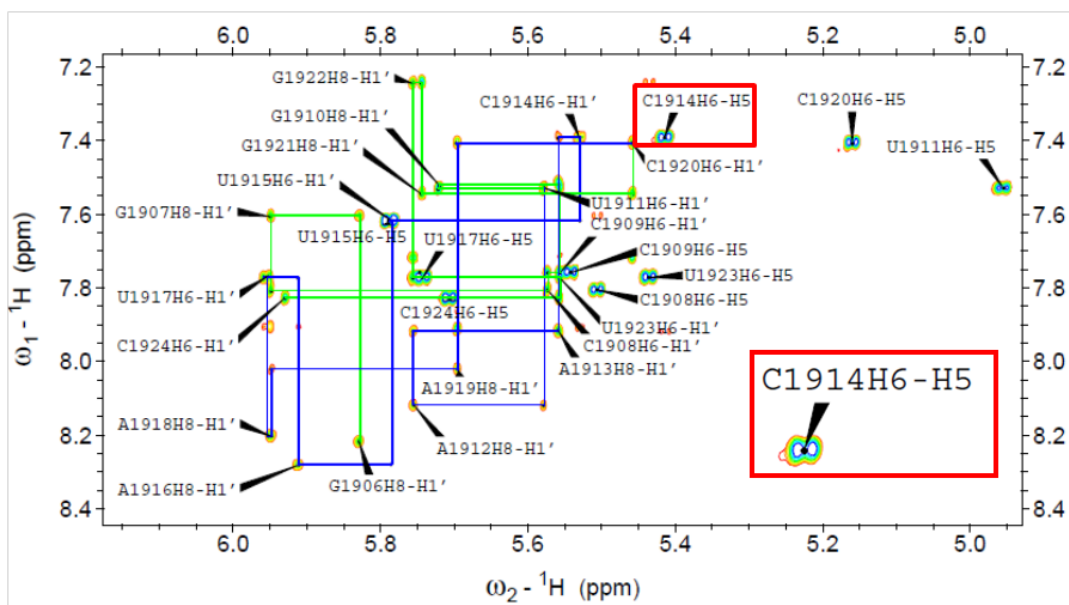
Figure 4.3. The base proton H8/H6 to sugar proton H1' region of the 2D NOESY

spectra of H69UUU (A) and H69ΨΨΨ (B). The pyrimidine H6-H5 crosspeaks and intraresidue H8/H6-H1' crosspeaks are labeled accordingly for visual simplicity. The H8/H6-H1' crosspeaks of the loop region (from U/Ψ1911H6-H1' to C1920H6-H1') are connected by blue lines, and those of the stem region are connected by green lines. Crosspeaks involving H1' of the pseudouridine residues are all upfield shifted significantly compared to those of the uridine residues, due to the isomerization from a N1-C1' covalent bond in a Uridine to a C5-C1' bond in a pseudouridine. In the 2D NOESY spectrum of H69ΨΨΨ, the assignment of C1914 H6-H5 crosspeak is shown in red. Possibly resulted from a local dynamic conformation, this crosspeak shows blurry contour levels, instead of a well defined doublet shape characteristic to a pyrimidine H6-H5 crosspeak in a 2D NOESY spectrum.
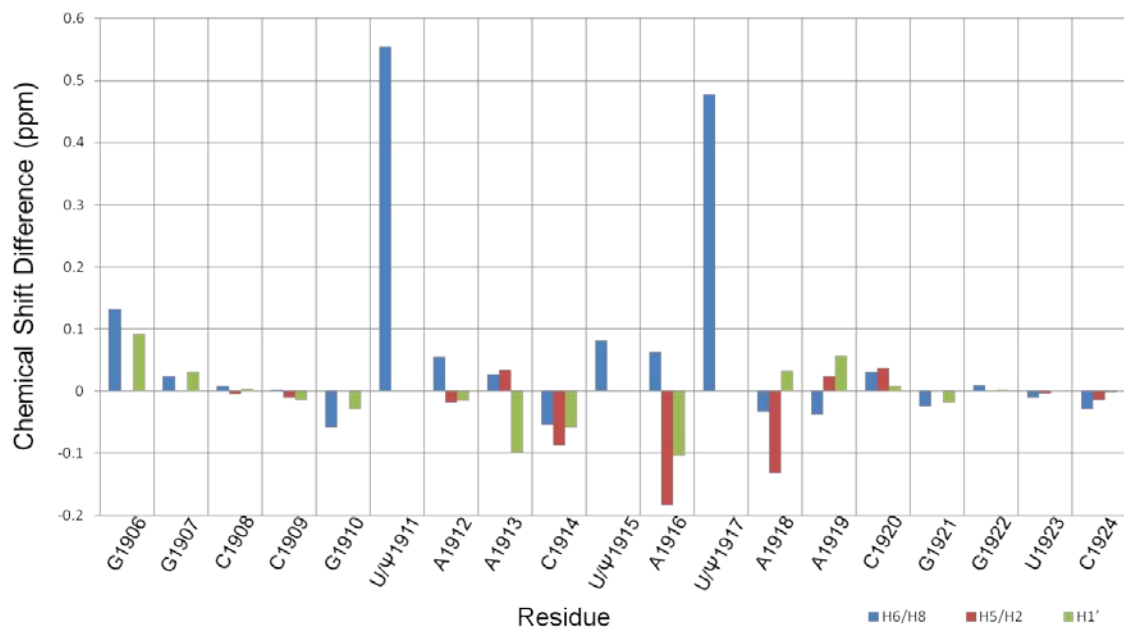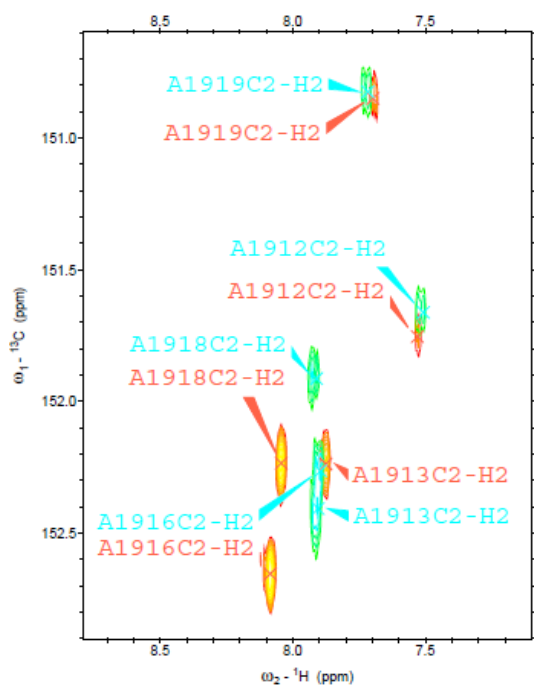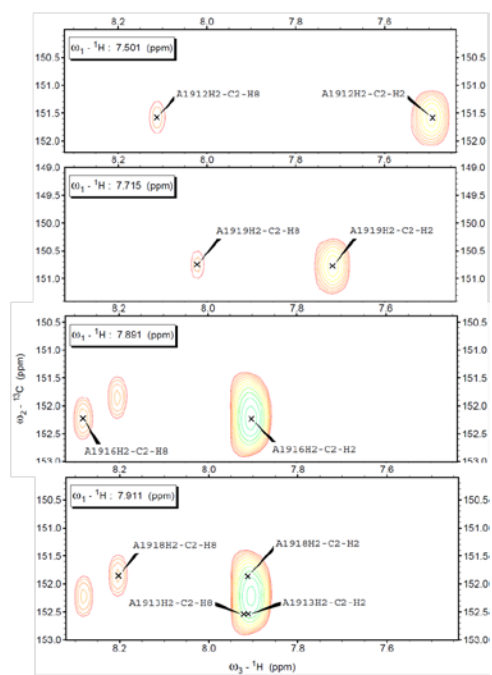
Figure 4.4 Difference in Chemical Shifts of base protons H8/H6/H5/H2 and sugar

proton H1' between H69UUU and H69ΨΨΨ. The chemical shift differences of H1' from U/Ψ1911, U/Ψ1915, and U/Ψ1917 are not shown for visualization of the smaller changes of other protons in the bar chart (A). Differences in chemical shifts of Adenine base proton H2s are directly compared in the C2-H2 region of 2D ^{13}C-^{1}H HMQC spectra of H69UUU (B cyan) and H69ΨΨΨ (B red). Crosspeaks between C2 and H2 of A1916 and A1918 are downfield shifted by 0.1-0.2 ppm in the ^{1}H dimension and 0.3 ppm in the ^{13}C dimension. The other three C2-H2 crosspeaks are barely shifted. 3D HCcH-TOCSY was used to identify and confirm the resonances of C2-H2 by correlating C2-H2 to the H8 in the same adenine ring (C).

A

B

C

D

E
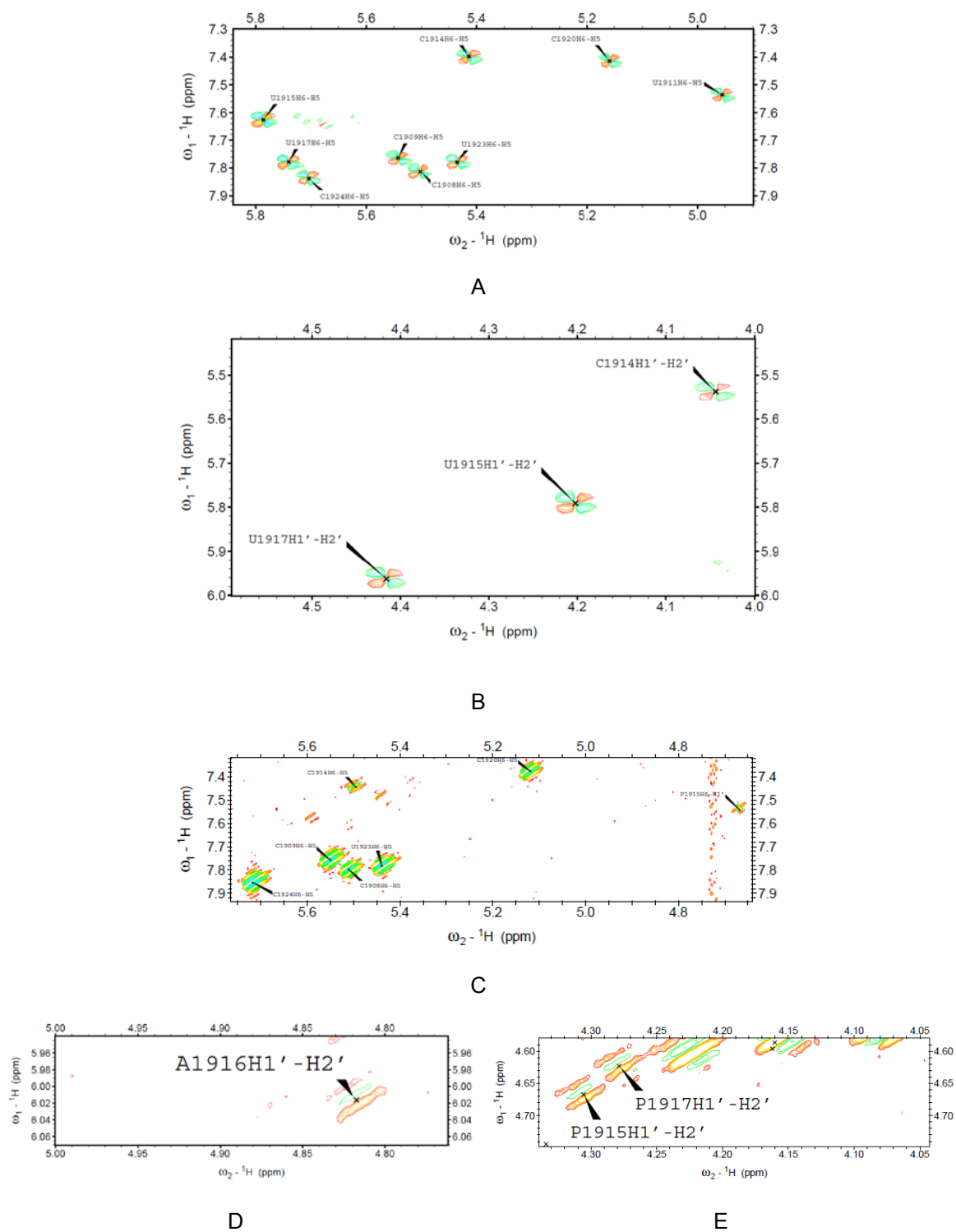
Figure 4.5 2D DQFCOSY spectra of H69UUU (A and B) and H69ΨΨΨ (C, D,

and E). All the H6-H5 crosspeaks were observed from H69UUU (A) and H69ΨΨΨ (C). One additional H6-H1' crosspeak from Ψ1915 in the H69ΨΨΨ spectrum (C upfield crosspeak), together with the medium-weak NOE between the H6-H1' pair, suggests that C6-H6 and C1'-H1' bonds are co-planar and antiparallel, and a "pseudo dihedral angle" was applied to restrain the χ dihedral angle of Ψ1915.

the same side of a plane along the C6-C5-C1', perpendicular to the H6-C6-C5-C1'-H1' plane, a strong NOE from H6-H1' would have been observed in the 2D NOESY spectrum, while an NOE characteristic of *anti* base conformation is observed (Figure 4.3 B), so an additional dihedral angle C6-C5-C1'-H1' was restrained to be 180 ± 45°.

4.3.2 NMR structures of H69UUU and H69ΨΨΨ

More NOE distance restraints and unNOEs restraints were employed in the structure calculation of H69UUU, benefiting from the availability of a ^{13}C, ^{15}N-labeled sample (Table 4.4). After structure calculation and refinement, ten converged structures with no NOE violation > 0.5 Å and dihedral angle violation > 7.5 ° were selected for H69UUU and H69ΨΨΨ, respectively (Figure 4.6 A and B). The RMSDs of these two structure families are comparable (Table 4.4). To compare the effects of pseudouridylation on RNA structures, one of the lowest energy structures was chosen from each family of the structures to represent the NMR solution structure of H69UUU or H69ΨΨΨ (Figure 4.6 C and D).

In solution, H69UUU and H69ΨΨΨ are folded into stem-loop structures, sharing several common structural features due to highly similar sequences (Figure 4.6). Residues G1906-G1910 and C1920-C1924 constitute the stem regions in an A-form RNA conformation (Figure 4.7 A and B). All of the expected intra- and inter-residue H8/6-H1' crosspeaks (of medium intensity) from the stem region were observed in the 2D NOESY spectra (Figure 4.3). Formation of the hydrogen bonds in the stem region is clearly shown in the 2D NOESY spectra acquired with samples dissolved in H_2O/D_2O (90%/10%) (Figure 4.2). Residues
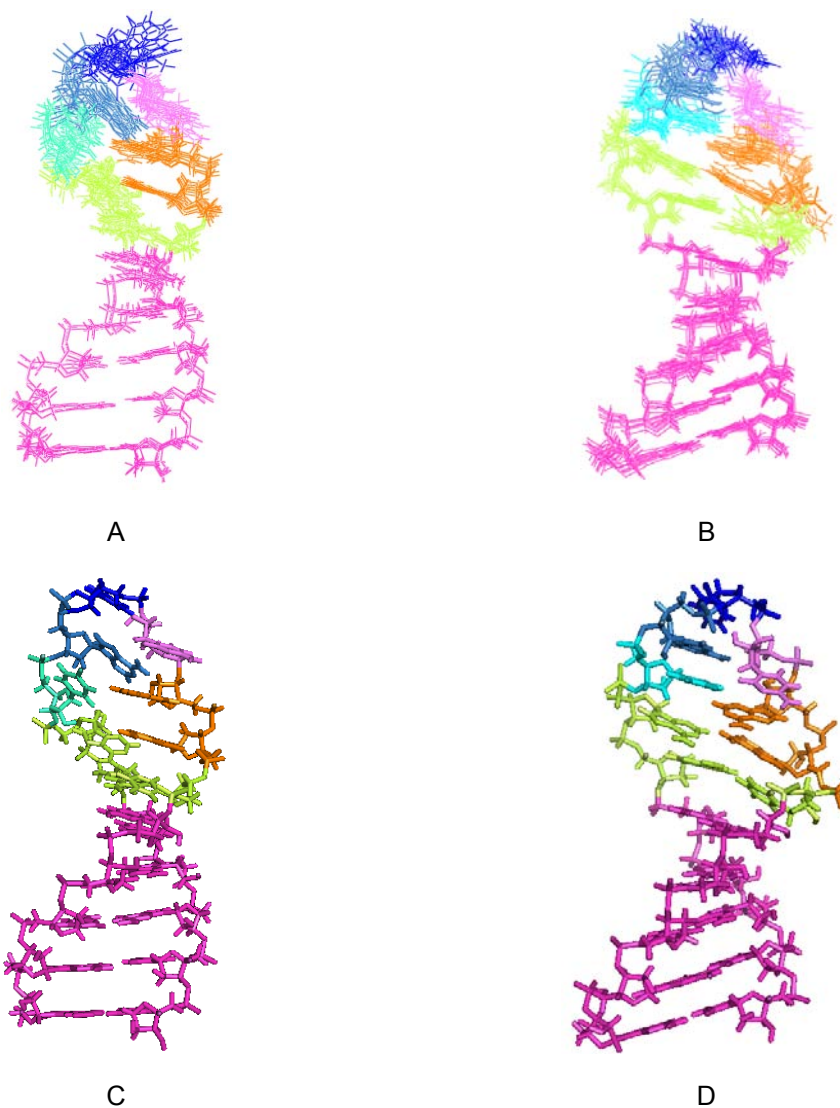
Figure 4.6 NMR solution structures of H69UUU (A and C) and H69ΨΨΨ (B and D). In each of the structure families, ten of the lowest energy structures are superimposed by "alignment" with PyMol (The PyMOL Molecular Graphics System, Schrödinger, LLC.). Stem residues involved in canonical Watson-Crick base pairs are colored in magenta. Loop residues are colored as follows: U/Ψ1911, A1918, and A1919 in lime green, A1912 and A1913 in orange, C1914 in violet, U/Ψ1915 in blue, A1916 in skyblue, and U/Ψ1917 in cyan.

A

B

C

D
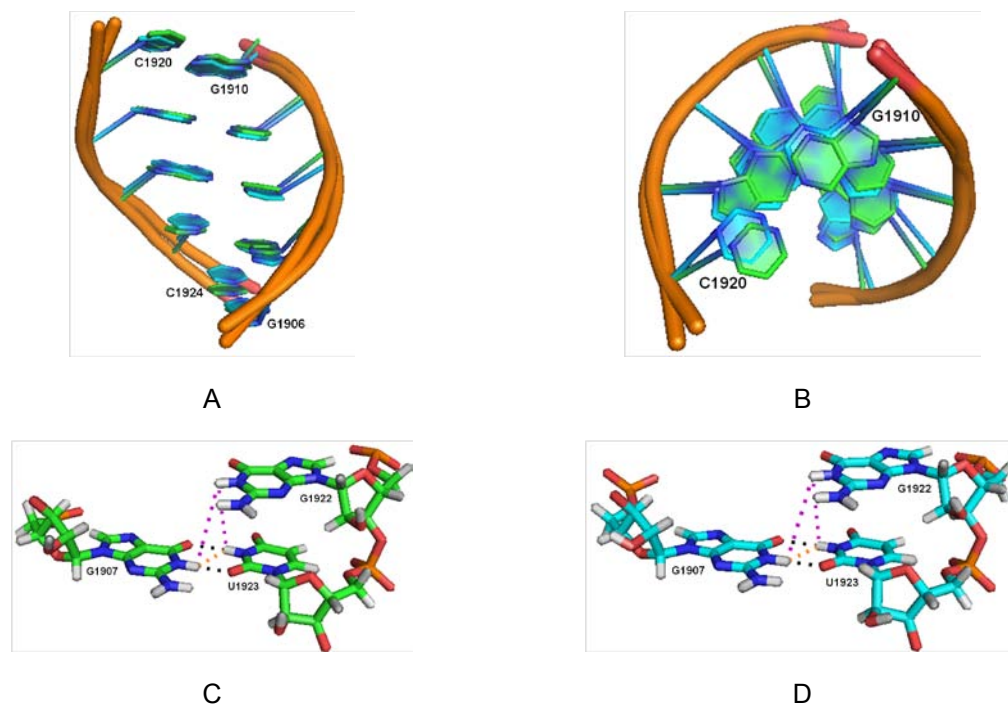
Figure 4.7 Structures of stem regions of H69UUU and H69ΨΨΨ, and the G1907•U1923 wobble base pairs with their positions relative to G1922. The stem regions of H69UUU (green) and H69ΨΨΨ (cyan) were aligned for comparison. The views from stem major groove side (A) and loop region (B) are shown. Two hydrogen bonds, shown as black dashed lines, are established between G1907 O6 with U1923 N3H (2.1 Å) and G1907 N1H with U1923 O2 (1.9 Å) in both structures (C and D). The distance between G1907 N1H and U1923 N3H (orange dashed line) is 2.3 Å, supported by an intense crosspeak (assignments in orange) of G1907 N1H-U1923 N3H observed in 2D NOESY spectra (Figure 4.2). The distances from G1922 N1H to G1907 N1H and U1923 N3H (purple dashed lines) are 4.3 and 4.0 Å, giving rise to two medium-weak crisspeaks (assignments in purple) of G1922 N1H-G1907 N1H and G1922 N1H- U1923 N3H in the 2D NOESY spectra (Figure 4.2).
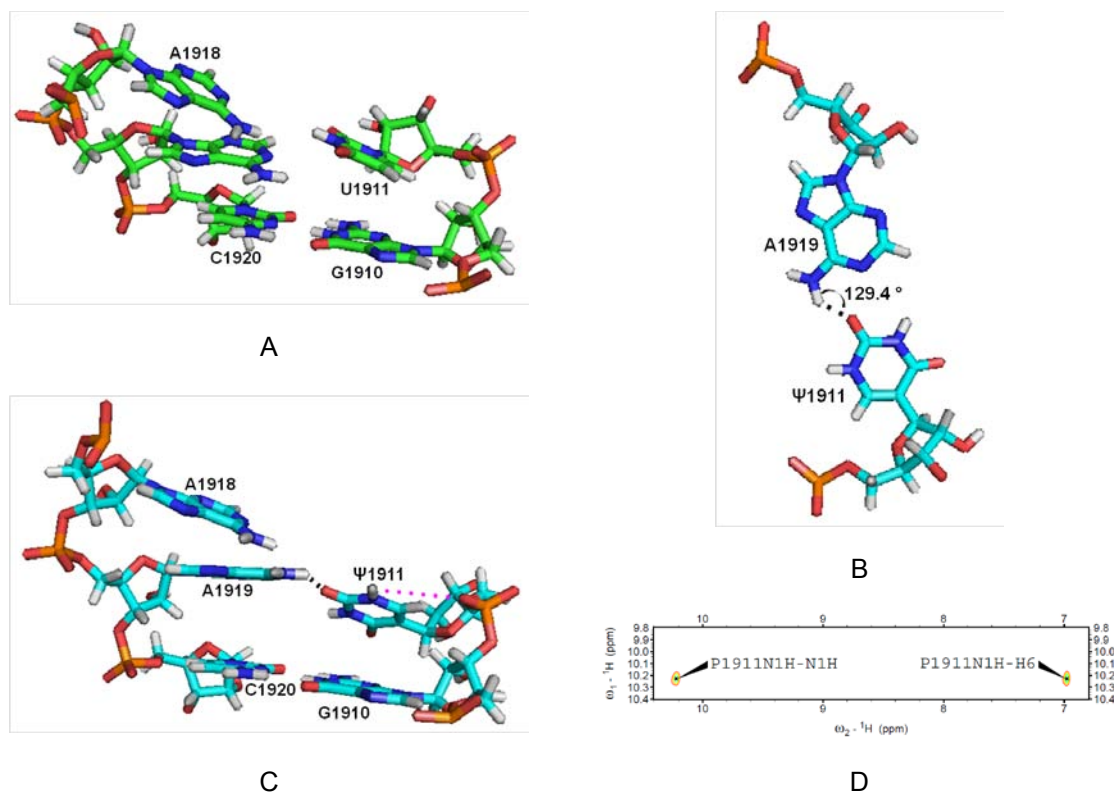
Figure 4.8 Effects of pseudouridylation at position 1911 on the structures of H69. No Watson-Crick base pair is observed with either U1911 (A) or Ψ1911 (C). The distances between atoms, which can participate hydrogen bonding interactions, on U1911 in H69UUU are at least 3.4 Å apart. The distance between Ψ1911 O2 and A1919 1H6 is 1.7 Å (black dashed line), while the angle of N6 and 1H6 of A1919 and O2 of Ψ1911 is 129.4° (B and D), so even if the hydrogen bond is established, its stability would be sacrificed by the unfavorable geometry. The distance between N1H and O2P of Ψ1911 is 4.5 Å (purple dashed line), enabling formation of a water-mediated hydrogen bond, which can protect the N1H from exchange with the solvent (D) (Durant and Davis 1995; Newby and Greenbaum 2002; Kim, Theimer et al. 2010).

G1907 and U1923 in the lower stem region form a wobble base pair (Figure 4.7 C and D). Resonances of the two imino protons of G1907 and U1923, protected from exchange with solvent by hydrogen bonding interactions, are visible in the 2D NOESY spectra (Figure 4.2). The interproton distances between the imino proton of G1922 and the imino protons of G1907 and U1923 support an A-form RNA base stacking involving the G1907•U1923 wobble base pair. The stem region is concluded with a G1910•C1920 base (Figure 4.7 A and B).

Instead of forming a canonical Watson-Crick base pair, U/Ψ1911 and A1919 are stacked onto G1910 and C1920, respectively. The distances between atoms on U1911 and U1918/U1919 exclude the possibility of hydrogen bond formation in H69UUU (Figure 4.8 A). In H69ΨΨΨ, The distance between O2 of Ψ1911 and 1H6 of A1919 is 1.7 Å (Figure 4.8 C), optimal for hydrogen bonds, though the geometry of the three atoms involved could potentially undermine the stability of this hydrogen bond (Figure 4.8 B). In 1D ^{1}H NMR spectrum carried out at 298K, the resonance of N1H Ψ1911 was visible, and the assignment was confirmed by the 2D NOESY spectrum of H69ΨΨΨ dissolved in H_2O/D_2O (90%/10%) by a crosspeak between the N1H and H6 of Ψ1911 (Figure 4.8 D). Protection of this imino proton unique to pseudouridines from exchange with solvent suggests that Ψ1911 N1H participates in hydrogen bond formation. Since no candidate acceptor is within direct hydrogen bond distance, the only explanation is that a water molecule mediates this hydrogen bonding interaction to a backbone oxygen atom (Davis 1995; Newby and Greenbaum 2002; Kim, Theimer *et al.* 2010). The distance between O2P and N1H of Ψ1911 is 4.5Å

(Figure 4.8 C), and both atoms are located in the major groove. This local configuration makes O2P of Ψ1911 suitable as a hydrogen bond acceptor to fix a water molecule.

The A-form RNA characters are extended into the loop regions on both the 5' and the 3' side. Base stacking of A1913, A1912, and U/Ψ1911 onto G1910, and base stacking of A1918 and A1919 onto C1920 are observed (Figure 4.6 C and D). The only difference is that bases from G1910 to C1914 form a regular stacking system in H69UUU (Figure 4.9 A), while the corresponding region in H69ΨΨΨ forms an irregular stacking conformation (Figure 4.9 B). In the crystal structure 1NKW (isolated large ribosomal subunit from *Deinococcus radiodurans*), C1914 is shown to be stacked on top of Ψ1915, while no corresponding crosspeaks in 2D NOESY spectra of H69UUU and H69ΨΨΨ were observed. Instead, crosspeaks of A1913 H8-C1914 H6 were observed in both spectra, indicating that C1914 is in close proximity to A1913 in each structure (Figure 4.6 C and D).

A significant difference in the structures of H69UUU and H69ΨΨΨ is within U/Ψ1915 to A1918 region. In the NMR structure of H69ΨΨΨ, bases of residues Ψ1915, A1916, Ψ1917, and A1918 form a continuous sheared (parallel-displaced) stacking system (Figure 4.10 B), which is also observed in the crystal structure 1NKW (Figure 4.10 C) (Harms, Schluenzen *et al.* 2001). In the NMR structure of H69UUU (Figure 4.10 A), this structural feature is absent, and interruptions of stacking are shown between U1915-A1916 step and U1917-A1918 step. The stepwise distances between the mass centers of bases
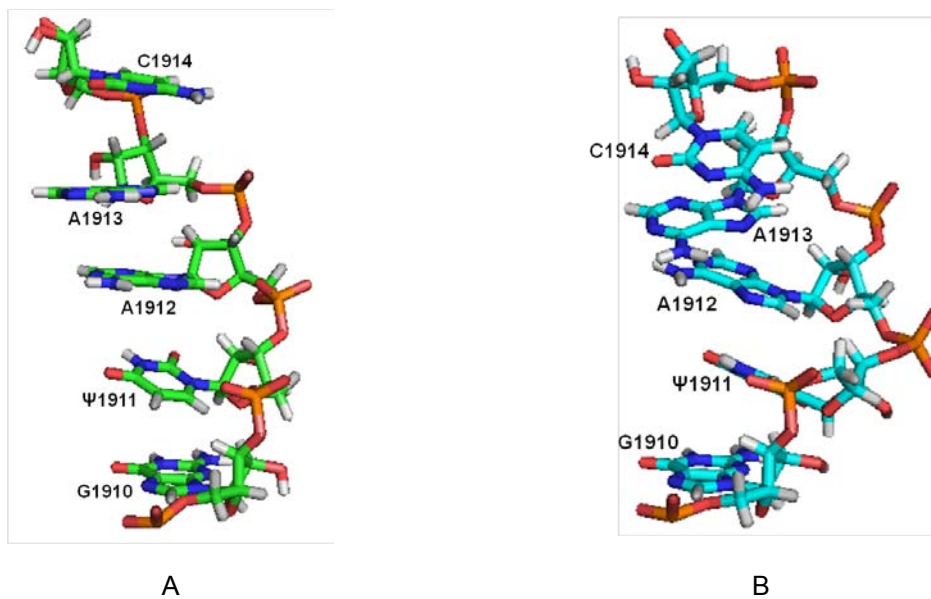
Figure 4.9 Comparison of the structures from G1910 to C1914 of H69UUU (A) and H69ΨΨΨ (B). Regular base stacking of the five bases is observed in the H69UUU sample, and the corresponding region in H69ΨΨΨ Shows more irregular stacking of the bases.

A   B   C
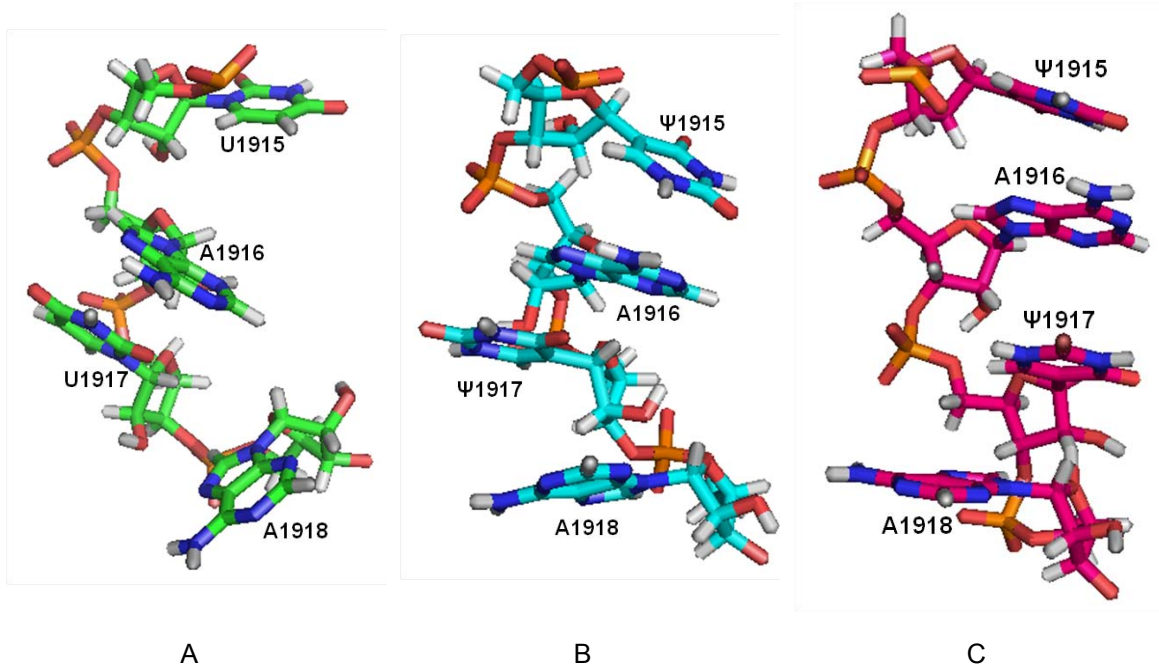
Figure 4.10 Comparison of loop region structures from U/Ψ1915 to A1918. Bases from residue Ψ1915 to A1918 in H69ΨΨΨ form a continuous sheared stacking system (B), which is also observed in the crystal structure of 1NKW (C). In the NMR structure of H69UUU (A), the stacking system is broken between the U1915-A1916 step and the U1917-A1918 step.

from Ψ1915 to 1918 are estimated to be within 5.5 Å in H69ΨΨΨ, compared to the distances of at least 7Å between mass centers of U1915-A1916 and U1917-A1918 in H69UUU. The mere distances can potentially compromise the stability of stacking interactions, even if a parallel-displaced stacking does exist between U1915 and U1916, and a "T-shape" stacking between U1917 and A1918 (Diener and Moore 1998). Residues U1915 and U1917 in the H69UUU structure assume a *C2'-endo* sugar pucker conformation, which is evidenced by appearance of two intense H1'-H2' crosspeaks in the 2D DQFCOSY spectrum (Figure 4.5 B). The *C2'-endo* sugar pucker enable the backbones of U1915 and U1917 to span a longer distance than the *C3'-endo* conformation, creating breakages in the base stacking system (Saenger 1984; Kim, Theimer *et al.* 2010). Even though a *C2'-endo* conformation is observed in A1916 of H69ΨΨΨ, an accompanying change of the χ dihedral angle (-79° vs. -160° in A-form RNA) helps reorient the A1916 base to form parallel-displaced stacking with bases of Ψ1915 and Ψ1917. The stacking interaction between bases of Ψ1917 and A1918 in H69ΨΨΨ brings H8 of A1918 close (≈ 3 Å) to the sugar protons H2' and H3' of Ψ1917, and two crosspeaks of medium intensity in the 2D NOESY ($τ_m$ = 150 ms) spectrum of H69ΨΨΨ were observed. The distances of A1918 H8-U1917 H2' and A1918 H8-U1917 H3' are ≈ 4 Å in H69UUU, where two weak crosspeaks were observed in the 2D NOESY ($τ_m$ = 150 ms) spectrum. The different base stacking patterns in H69UUU and H69ΨΨΨ may be also helpful to explain the resonance shifts (0.1 ~ 0.2 ppm) of base proton H2s of A1916 and A1918, when the two samples are

Table 4.5 Key NOEs from H69ΨΨΨ NMR experiment also observed in 1NKW

Proton 1	Proton 2	NOE distance (Å)[a]	Distance in 1NKW (Å)
A1912H2	A1918H2	5.00 ± 1.00	4.37
A1913H2	A1918H1'	4.50 ± 1.50	6.29
A1913H2	Ψ1917H1'	5.00 ± 1.00	3.82
A1918H2	A1919H1'	3.90 ± 1.00	5.34

a. An error bar of NOE distance of 0.5 Å is added on the range shown below.

Table 4.6 Violations of H69ΨΨΨ crystal structure in 1NKW to the NOEs derived

from the NMR experiments on H69ΨΨΨ

Proton 1	Proton 2	NOE distance (Å)[a]	Distance in 1NKW (Å)
Ψ1911H6	A1912H8	5.00 ± 1.00	6.53
Ψ1911H3'	A1912H8	2.70 ± 0.50	4.46
A1912H2	A1913H8	5.00 ± 1.00	7.06
A1912H1'	A1919H2	3.40 ± 0.60	7.62
A1913H8	C1914H6	5.00 ± 1.00	7.60
A1913H8	C1914H5	5.00 ± 1.00	9.01
A1913H8	C1914H1'	5.00 ± 1.00	9.08
A1913H2	C1914H1'	3.30 ± 0.60	11.30
A1913H2	Ψ1917H6	unNOE	2.84
A1913H1'	C1914H6	5.00 ± 1.00	7.95
A1913H1'	A1918H2	3.80 ± 1.00	8.20
A1913H1'	A1919H2	5.00 ± 1.00	10.74
A1913H2'	C1914H6	2.90 ± 0.60	5.34
A1913H2'	C1914H1'	4.50 ± 1.50	8.10
A1913H3'	C1914H6	3.00 ± 0.60	4.96
A1913H3'	C1914H1'	5.00 ± 1.00	7.08
C1914H1'	A1916H8	5.00 ± 1.00	9.71
C1914H1'	A1916H2'	4.50 ± 1.50	12.90
C1914H2'	C1914H6	2.90 ± 0.50	4.02
C1914H4'	Ψ1915H6	4.00 ± 1.00	5.59
Ψ1915H6	A1916H2	5.00 ± 1.00	7.72
A1916H8	A1916H3'	4.50 ± 1.00	2.21
A1916H2	Ψ1917H1'	3.00 ± 0.50	5.07

Proton 1	Proton 2	NOE distance (Å)[a]	Distance in 1NKW (Å)
A1916H2	Ψ1917H2'	5.00 ± 1.00	7.84
A1916H2	Ψ1917H4'	5.00 ± 1.00	6.62
A1916H1'	Ψ1917H1'	4.60 ± 1.00	6.17
A1916H1'	Ψ1917H6	4.20 ± 1.00	7.30
A1916H1'	Ψ1917H2'	5.00 ± 1.00	8.62
A1916H1'	A1918H8	5.00 ± 1.00	9.25
A1916H2'	Ψ1917H6	2.60 ± 0.50	4.82
Ψ1917H1'	A1918H2	5.00 ± 1.00	9.23
Ψ1917H2'	A1918H1'	5.00 ± 1.00	2.94
Ψ1917H3'	A1918H8	3.00 ± 0.60	1.06
A1918H8	A1919H8	5.00 ± 1.00	6.68
A1919H1'	C1920H5	5.00 ± 1.00	7.13
G1921H1'	G1922H8	3.60 ± 0.80	4.93
G1922H1'	U1923H6	3.80 ± 1.00	5.59
G1922H1'	U1923H5	4.50 ± 1.50	6.63
U1923H1'	C1924H6	3.40 ± 0.60	4.65

a. An error bar of NOE distance of 0.5 Å is added on the range shown below.

Table 4.7 Violations of H69ΨΨΨ NMR structure to the NOEs derived from the

NMR experiments on H69UUU

Proton 1	Proton 2	NOE distance (Å)[a]	Distance in H69ΨΨΨ (Å)
G1910H2'	U1911H1'	unNOE	3.33
G1910H3'	U1911H5	3.50 ± 0.80	4.81
U1911H1'	A1919H2	4.50 ± 1.50	8.18
A1912H8	A1912H2'	3.00 ± 0.60	4.20
A1912H2	A1913H8	5.00 ± 1.00	10.12
A1912H2	A1913H2	3.90 ± 1.00	5.80
A1912H2	A1918H1'	5.00 ± 1.00	7.46
A1912H2	A1919H2	5.00 ± 1.00	3.41
A1912H2'	A1913H8	2.50 ± 0.50	3.55
A1913H2	A1916H1'	5.00 ± 1.00	7.14
C1914H4'	A1916H8	4.00 ± 1.00	5.57
C1914H4'	A1916H1'	4.00 ± 1.00	7.80
U1915H6	U1915H2'	2.30 ± 0.50	3.56
U1915H4'	A1916H8	4.10 ± 1.00	5.71
A1916H8	A1916H2'	4.50 ± 1.50	2.05
A1916H1'	U1917H6	5.00 ± 1.00	3.13
A1918H2'	A1919H8	unNOE	2.73
A1919H1'	C1920H6	3.70 ± 0.80	5.12
C1924H5	C1924H3'	3.10 ± 0.60	4.96

a. An error bar of NOE distance of 0.5 Å is added on the range shown below.

Table 4.8 Violations of H69UUU NMR structure to the NOEs derived from the

NMR experiments on H69ΨΨΨ

Proton 1	Proton 2	NOE distance (Å)[a]	Distance in H69UUU (Å)
A1912H8	A1913H8	3.50 ± 0.80	4.89
A1912H2	A1918H2	5.00 ± 1.00	3.46
A1913H8	C1914H1'	5.00 ± 1.00	7.04
A1913H2	C1914H1'	3.30 ± 0.60	5.08
A1913H1'	A1919H2	5.00 ± 1.00	7.46
A1913H3'	C1914H5	5.00 ± 1.00	2.82
C1914H3'	Ψ1915H6	3.10 ± 0.60	4.38
Ψ1915H6	Ψ1915H2'	3.50 ± 0.60	2.24
Ψ1915H6	A1916H2	5.00 ± 1.00	7.10
A1916H8	A1916H2'	2.60 ± 0.50	4.04
A1916H2	Ψ1917H1'	3.00 ± 0.50	4.74
A1916H2	Ψ1917H2'	5.00 ± 1.00	7.74
A1916H2	Ψ1917H4'	5.00 ± 1.00	6.81
A1916H2	A1918H1'	unNOE	3.45
A1916H1'	Ψ1917H6	4.20 ± 1.00	6.11
A1916H1'	Ψ1917H2'	5.00 ± 1.00	7.40
A1916H1'	A1918H8	5.00 ± 1.00	7.79
A1916H3'	Ψ1917H6	4.10 ± 1.00	2.60
Ψ1917H1'	A1918H8	5.00 ± 1.00	3.01
Ψ1917H1'	A1918H2	5.00 ± 1.00	7.42
A1918H8	A1918H2'	3.30 ± 0.60	4.57
A1918H8	A1918H3'	2.80 ± 0.50	3.86
A1918H2'	A1919H8	2.70 ± 0.50	4.00

Proton 1	Proton 2	NOE distance (Å)[a]	Distance in H69UUU (Å)
C1920H3'	G1921H8	2.70 ± 0.50	3.81
G1921H1'	G1922H8	3.60 ± 0.80	5.01
U1923H5	C1924H6	5.00 ± 1.00	7.35
U1923H1'	C1924H6	3.40 ± 0.60	5.01

a. An error bar of NOE distance of 0.5 Å is added on the range shown below.

compared (Al-Hashimi, Pitt *et al.* 2003). The key NOEs and NOE violations are shown in Table 4.5, 4.6, 4.7, and 4.8 for comparison of the NMR structures of H69UUU and H69ΨΨΨ, and the crystal structure of H69ΨΨΨ of 23S rRNA from *D. radiodurans*.

4.4 Discussion

In this study, the structures of unmodified (H69UUU) and pseudouridylated (H69ΨΨΨ) H69 were solved by NMR spectroscopy. The global folding of the two constructs are very similar in that 1) a hairpin structure is observed in both cases, with a stem region of five base pair long; 2) U/Ψ1911 and A1919 do not form a Watson-Crick base pair, but bridge stacking interactions between the neighboring loop residues and stem capping base pair G1910•C1920; 3) a continuous stacking system consisting residues U/Ψ1911, A1912, A1913, and C1914 is observed, with subtle different backbone contours in H69UUU and H69ΨΨΨ. The most significant structural difference is revealed in the loop region from U/Ψ1915 to A1918, where four base rings form a stacking "ladder" in H69ΨΨΨ, while the "ladder" is interrupted in H69UUU at U1915-A1916 and U1917-A1918 steps. This structural difference is accompanied by different sugar puckers of these loop residues. Another significant difference is the dynamics of residue C1914, where conformational change is observed in the H69ΨΨΨ 2D NOESY spectrum. It is unambiguously shown in the comparison of the two structures that pseduouridylation does have structural effects on H69 folding.

The first pseudouridylation site in H69 is residue 1911. No base pairing formation (U/Ψ1911XA1919) is observed in either structure, and U/Ψ1911 is sandwiched between G1910 and A1912 in a "parallel-displaced" configuration, though the NMR spectra and structure of H69ΨΨΨ suggests that at least one hydrogen bond involving a bridging water molecule is formed between the N1H of Ψ1911 and its backbone phosphate oxygen. It has been reported that formation of such a water-mediated hydrogen bond in pseudouridine contributes ≈ - 0.7 kcal/mol to stabilize the RNA folding (Newby and Greenbaum 2002). This observation agrees with earlier findings in thermodynamic studies on H69 pseudouridylations (Meroueh, Grohar *et al.* 2000; Sumita, Jiang *et al.* 2012). UV-melting experiments on a duplex construct (residues G1906 – A1912 and A1919 – C1924 of H69) with either U1911 or Ψ1911 shows that pseudouridylation at 1911 exerts a stabilizing effect of ≈ -1.0 kcal/mol. A similar thermodynamic contribution was also revealed in comparison of thermodynamic stabilities of H69UUU and H69ΨUU (1911-selective pseudouridylation) hairpin structures. The stabilizing effect of pseudouridylation at 1911 in H69 appears to be localized within a context extending to A1912, and completion of the hairpin structure with all the loop residues (no pseudouridylation at 1915 and 1917) does not appear to influence the stabilizing effect of Ψ1911. In addition, formation of a water-mediated hydrogen bond involving N1H and enhanced stacking of Ψ1911 with neighboring residues dominate the stabilizing contribution of pseudouridylation at this site, and the extra cross-strand interactions appear to provide insignificant stability bonus, if a minor difference of -0.3 kcal/mol (-1.0 – (-0.7)) kcal/mol) is

considered. This suggestion agrees with a possible hydrogen bond established between 1H6 of A1919 and O2 of Ψ1911, in an unfavorable geometry. As a result, mutation of Ψ1911C showed negligible effects on the subunit association profile and activity of ribosomes, when pseudouridylation at 1915 and 1917 by RluD was not affected (Liiv, Karitkina *et al.* 2005; Leppik, Peil *et al.* 2007). It was also reported in the same paper that mutation A1919G showed severe slow growth phenotype with existence of the pseudouridylations in H69. This dramatic growth defect could not be attributed completely to the loss of the hydrogen bond between 1H6 of A1919 and O2 of Ψ1911 due to substitution of a hydrogen bond donor (1H6) in A1919 to an acceptor (O6) in G1919, if G1919 assumes the same conformation as A1919 in H69ΨΨΨ. Base stacking may play a role in this case, and G often contributes a slightly less stabilizing effect as the capping residue of an RNA duplex compared to A (Freier, Burger *et al.* 1983; Freier, Alkema *et al.* 1985; O'Toole, Miller *et al.* 2005). Based upon the structures reported here, we hypothesize that a less favorable stacking involving G1919 and a loss of a hydrogen bond worked synergistically to alter the structural property of H69ΨΨΨ A1919G, leading to compromised ribosomal activity.

More profound structural effect is found with pseudouridylation at 1915 and 1917, where Ψ1915, A1916, Ψ1917, and A1918 form a continuous stacking system in H69ΨΨΨ, and the integrity of this stacking system is abolished in H69UUU, which has been observed in chemical probing experiments of large ribosomal subunits (Sakakibara and Chow 2012). Enhancement of stacking interactions by pseudouridylation is also suggested by other studies (Davis 1995;

Durant and Davis 1995; Kim, Theimer *et al.* 2010). While there is an essential difference that significant destabilizing effect was observed in this study, and stabilization of RNA folding were reported elsewhere. Loss of well-defined doublet contours of C1914 H6-H5 crosspeak in H69ΨΨΨ 2D NOESY spectrum indicates a highly dynamic local structure, with intermediate time-regime exchange between the different conformers. This destabilizing effect was revealed in thermodynamic studies on both H69 duplex construct and selective pseudouridylation of H69 hairpins (Meroueh, Grohar *et al.* 2000; Sumita, Jiang *et al.* 2012). Pseudouridylation at residues 1915 and 1917 confers a destabilizing effect of 0.9 kcal/mol. When they were pseudouridylated individually, Ψ1915 contributes a destabilizing effect of 0.7 kcal/mol and Ψ1917 0.3 kcal/mol. In the crystal structures containing large ribosomal subunit, a stacking system extended from m^3U/Ψ1915 to A1918 appears to be conserved in ribosomes from bacteria and at leat one eukaryote (yeast) (Harms, Schluenzen *et al.* 2001; Yusupov, Yusupova *et al.* 2001; Berk, Zhang *et al.* 2006; Weixlbaumer, Petry *et al.* 2007; Weixlbaumer, Jin *et al.* 2008; Schmeing, Voorhees *et al.* 2009; Jenner, Demeshkina *et al.* 2010; Korostelev, Zhu *et al.* 2010; Ratje, Loerke *et al.* 2010; Ben-Shem, Garreau de Loubresse *et al.* 2011). Mutations disrupting the stacking gave rise to suboptimal ribosomal subunit association and activity (Liiv, Karitkina *et al.* 2005; Leppik, Peil *et al.* 2007). Comparison of H69 structures in 1NKW and 2I2T shows that most dramatic conformational changes undergo from residue A1912 and Ψ1915 (Figure 4.11) (Harms, Schluenzen *et al.* 2001; Berk, Zhang *et al.* 2006; Saro and SantaLucia 2007). Thus, it is evident that enhanced stacking
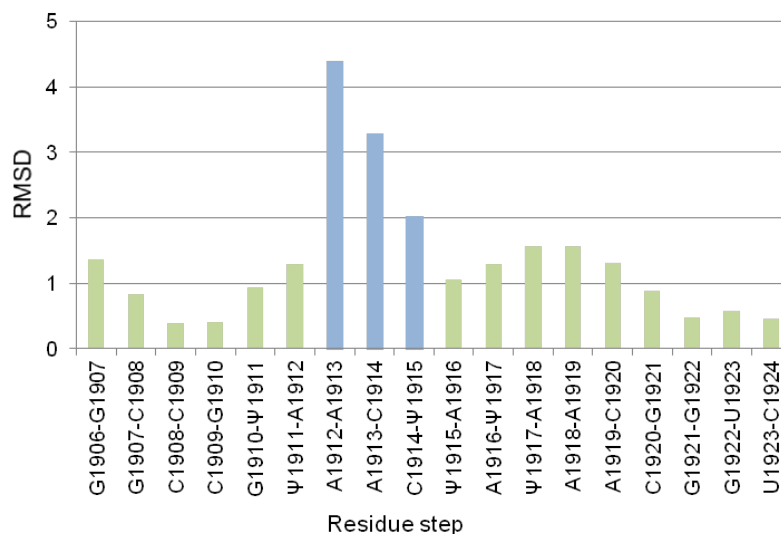
Figure 4.11 Conformational changes in each dinucleotide step upon ribosome association. Coordinates of the two neighboring residues in 1NKW and 2I2T were aligned and the RMSD for the two residues were calculated with all atoms (Saro and SantaLucia 2007; Sijenyi, Saro *et al.* 2011). Steps highlighted in light blue show most dramatic conformational changes on ribosome subunit association. Dinucleotide steps were chosen to illustrate the conformational changes in each residue and the relative conformational changes between the two residues, while an propagating effect of coordinates translation resulted from other residues, e.g. stem duplex residues, is minimized.

by pseudouridylation at 1915 and 1917, together with Ψ1911 plays a role in pre-organizing the folding of H69 loop region by destabilizing the upstream loop residues. The fact that bases of residues Ψ1915 - A1918 assume a continuous stacking in the post-association state suggests that pseudouridylation at 1915 and 1917 may contribute to the stabilization of b2a intersubunit bridge by reinforcing the stacking in the downstream loop region of H69.

4.5 Conclusion

H69 is a stem-loop structure, consisting of 19 nucleotides, in bacterial 23S ribosomal RNA, and its primary/secondary structure and modification pattern are highly conserved through all the three domains (Ofengand and Bakin 1997; Yusupov, Yusupova *et al.* 2001; Cannone, Subramanian *et al.* 2002). Recent X-ray crystallography studies show that H69 directly participates in establishment of B2a intersubunit bridge in both bacterial and eukaryotic ribosomes, and have direct contact to multiple translational factors during translation process (Yusupov, Yusupova *et al.* 2001; Weixlbaumer, Petry *et al.* 2007; Weixlbaumer, Jin *et al.* 2008; Schmeing, Voorhees *et al.* 2009; Jenner, Demeshkina *et al.* 2010; Korostelev, Zhu *et al.* 2010; Ratje, Loerke *et al.* 2010; Ben-Shem, Garreau de Loubresse *et al.* 2011). Comparison of the different X-ray crystal structures shows that nucleotide A1913 base is flipped out upon association of ribosomal subunits. In addition, the whole H69 hairpin structure undergoes subtle conformational changes during different stages in the translation (Figure 4.12). Since these crystal structures are representative of states when stable interactions between H69 and translational factors are still intact and the energy
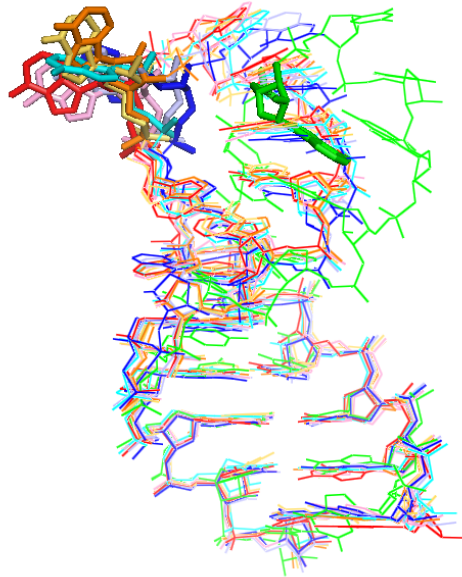
Figure 4.12 Comparison of H69 conformations at different stages of translation (Del Campo, Recinos *et al.* 2005; Berk, Zhang *et al.* 2006; Weixlbaumer, Petry *et al.* 2007; Weixlbaumer, Jin *et al.* 2008; Schmeing, Voorhees *et al.* 2009; Jenner, Demeshkina *et al.* 2010; Korostelev, Zhu *et al.* 2010; Ratje, Loerke *et al.* 2010). Crystal structures are shown in "line" format and A1913 in each structure is highlighted by "stick" format. The coloring scheme is as following: green, 1NKW, pre-initiation; cyan, 2I2T, post-initiation; red, 2WRO, pre-accommodation of A-site tRNA; pink, 3I8F, post-accommodation of A-site, tRNA; blue, 2XTG, pre-translocation; lightblue, 2XUX, post-translocation; orange-yellow, 3MRZ, bound with RF2; orange, 2WH2, bound with RRF. On association of the complete ribosomes, A1913 is flipped out from loop to form B2a intersubunit bridge with h44 of 16S rRNA. Subtle differences of H69 are shown between each stage of the translation process.

of the whole translation machinery is at a local minimum, it is likely that H69 experiences more dramatic conformational changes during the ratchet movements of the translation process (Fischer, Konevega *et al.* 2010). Thus, structural stability and flexibility are both important to the function of H69.

H69 oligonucleotide constructs (H69UUU and H69ΨΨΨ) adopt a hairpin structure, consisting of a 5 base pair stem region and a 9 residue loop region. Comparison of the two NMR structures reveals that pseudouridylation modifies local intraresidue and interresidue interactions. Ψ1911 N1H has been shown to participate in a water-mediated hydrogen bond with backbone oxygen atom and one additional hydrogen bond could be established between the 1H6 of A1919 and O2 of Ψ1911. Pseudouridylation at 1911 has a stabilizing effect to the folding of H69 by enhanced stacking combined with a water-mediated hydrogen bond (Davis 1995; Kim, Theimer *et al.* 2010). Unique to H69, pseudouridylation at 1915 and 1917 is destabilizing, as shown in the thermodynamic studies and NMR spectra of H69ΨΨΨ. Structure calculation shows that pseudouridylation at these two positions promote the stacking interactions from Ψ1915 to A1918, in a fashion similar to that in the crystal structure 1NKW. In contrast, in the NMR structure of H69UUU, base stacking is missing at the U1915-A1916 and U1917-A1918 steps. We hypothesize that the enhanced stacking in the lower half loop region (from Ψ1915 to A1918) destabilizes the C1914, which is positioned next to A1913. This destabilization enables structural flexibility of the region containing A1913 and prepares A1913 to flip out upon association of the ribosomeal subunits. In the post-initiation stages, pseudouridylation at 1915 and 1917 could

potentially provide extra stability to the B2a intersubunit bridge, by enhanced stacking possibly combined with water-mediated hydrogen bonds.

CHAPTER 5

INTERACTIONS BETWEEN 16S RIBOSOMAL RNA HELIX 9 AND S20 RIBOSOMAL PROTEIN IN THE 30S RIBOSOMAL SUBUNITS OF ESCHERICHIA COLI AND PSEUDOMONAS AERUGINOSA

5.1 Introduction

Ribosomes are universally conserved protein biosynthesis machinery. Each ribosome is composed of two discrete subunits, named small ribosomal subunit (ssu) and large ribosome subunit (lsu). In a bacterial ribosome, one copy of ribosomal RNA (rRNA) and 21 ribosomal proteins (rproteins) were identified in a ssu, and two copies of rRNAs and about 30 rproteins in a lsu (Berk, Zhang *et al*. 2006; Dunkle, Wang *et al*. 2011). The composition of a eukaryotic ribosome is even more complicated, with 10 additional rproteins in a ssu, and one extra copy of rRNA and 20 extra rproteins in a lsu (Ben-Shem, Garreau de Loubresse *et al*. 2011). The assembly of each subunit is a highly hierarchical and coordinated process (Mizushima and Nomura 1970; Held, Mizushima *et al*. 1973; Held, Ballou *et al*. 1974; Herold and Nierhaus 1987; Grondek and Culver 2004; Shajani, Sykes *et al*. 2011). The rRNA molecules play a role as the scaffold for the folding of the subunits, and constitute major catalysis activity (Serdyuk, Agalarov *et al*. 1983; Nitta, Ueda *et al*. 1998). Ribosomal proteins are incorporated into ribosomal subunits in an ordered fashion co-transcriptionally to complement the subunit morphology and function (McCarthy, Britten *et al*. 1962; Mangiarotti, Apirion *et al*. 1968; Osawa, Otaka *et al*. 1969; Homann and Nierhaus

1971; Lindahl 1973; Nierhaus, Bordasch *et al*. 1973; de Narvaez and Schaup 1979; Jakel and Gorlich 1998).

S20 is one ribosomal protein in the small subunit (Berk, Zhang *et al*. 2006; Dunkle, Wang *et al*. 2011). According to the Nomura 30S ribosomal subunit assembly map, S20 is a primary binding protein bound dominantly in the 5' domain of 16S rRNA, and works synergistically with S4, another small subunit primary binding protein, to promote the binding of S16, a secondary binding protein (Mizushima and Nomura 1970). In turn, S16 could affect the incorporation of S5 and S12, tertiary proteins in the small subunit. Any effect on the S20 protein could potentially ripple down the cascade to affect the assembly of the small subunit. S20 could potentially exert its effect through h44 of 16S rRNA, since nucleotide residues in the penultimate region of h44 complete the binding pocket of S20 in the small subunit (Figure 5.1) (Berk, Zhang *et al*. 2006; Dutca and Culver 2008) .

The S20 rprotein is structurally and functionally important, and deletion of S20 was shown to result in slow growth phenotype due to inefficient translation initiation (Tobin, Mandava *et al*. 2010). *In vitro* studies on 16S rRNA post-transcriptional modifications revealed that deletion of S20 could substantially decrease the modification level of several nucleotide residues located close to the tRNA binding sites on the 16S rRNA (Ryden-Aulin, Shaoping *et al*. 1993; Chow, Lamichhane *et al*. 2007). Most of the methyltransferases catalyzing modifications on these sites show substrate preference to the assembled small subunit, except that RsmB (Fmu) methylates 16S rRNA more effectively instead
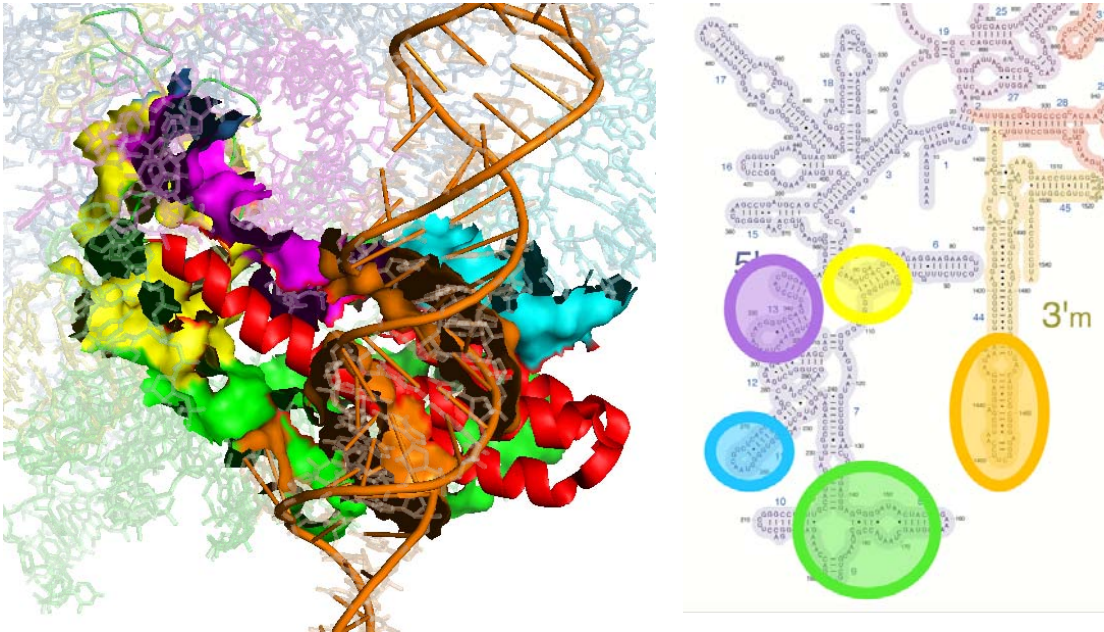
Figure 5.1 Binding pocket of ribosomal protein S20 in a small ribosomal subunit from bacteria (Berk, Zhang *et al*. 2006). 5' domain of the 16S rRNA is shown in "stick" and the 3' minor domain in cartoon. Chemical moiety of the 16S rRNA within 6 Å from the S20 is shown in "surface". The coloring scheme for each helix region of 16S rRNA constituting the S20 binding pocket is: helix 5 (h5), yellow; h8, h9, and h10, green; h11, blue; h13 and h14, purple; h44, brown. Composition of the binding pocket agree with the probing experiments very well (Dutca and Culver 2008). The picture of the 16S rRNA secondary structure of *E. coli* is from the Noller's lab website. http://rna.ucsc.edu/rnacenter/ribosome_images.html

of the small subunit (Tscherne, Nurse *et al*. 1999; O'Farrell, Pulicherla *et al*. 2006; Basturea and Deutscher 2007; Demirci, Larsen *et al*. 2010). These observations suggest that S20 has a long rang effect on the assembly of the small ribosomal subunit, possibly transferred by the h44 in a relatively late stage of the subunit assembly.

Extensive studies on the interactions of S20 protein and helix 9 (h9) (or 16S rRNA) from *E. coli* and *Pseudomonas aeruginosa* have been carried out in Dr. Cunningham's lab with a specially designed genetic system (Figure 5.2) (Lee, Holland-Staley *et al*. 1996; Lee, Varma *et al*. 1997). Using the system, Marny Waddington in the Cunningham lab found that, 16S rRNA from *P. aeruginosa* (Pa16S) was able to be assembled into ribosomes with all other rRNAs and rproteins from *E. coli*, and the activity of ribosomes was ≈25% compared to *E. coli* wild type (unpublished data). If the sequence of helix 9 on 16S rRNA from *P. aeruginosa* was replaced by that from *E. coli*, the activity of ribosomes containing this chimeric 16S rRNA with all other components from the *E. coli* was recovered to ≈96%. Alternatively, if S20 rprotein from *P. aeruginosa* (PaS20) was overexpressed from a plasmid by induction, the activity of ribosomes containing 16S rRNA from *P. aeruginosa* was increased to ≈96% by complementation of PaS20. The aforementioned observations help to narrow down the critical interaction partners to h9 of the 16S rRNA and S20.

Previous work done by Ami Lamichhane of Dr. Cunningham's lab suggested that, incorporation of non-cognate pairs of h9 and S20 could result in deficient assembly of the small ribosomal subunit and association of the
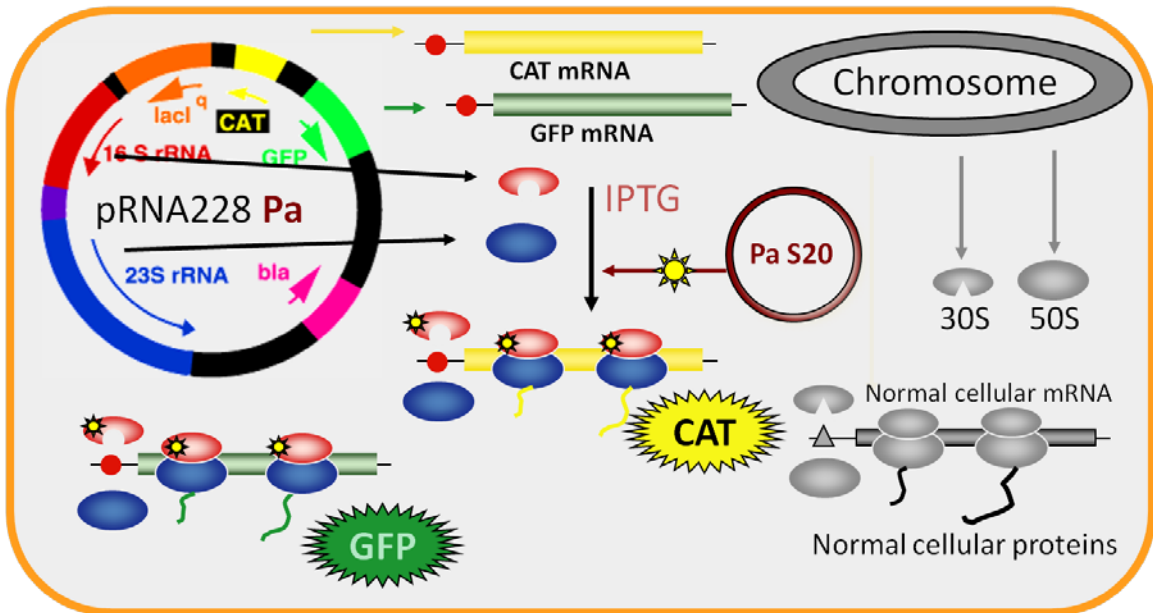
Figure 5.2 A genetic system for *in vivo* study on ribosomal functions (Lee, Holland-Staley *et al*. 1996; Lee, Varma *et al*. 1997). In an *E. coli* host cell, a plasmid encoding 5S, 16S, and 23S rRNA, under control of a *lacI* operator is introduced to make ribosomal RNAs independent from the host chromosomal RNA synthesis. The anti-Shine-Dalgarno sequence located close to the 3' terminus of the 16S rRNA encoded by the plasmid was engineered to be 5'-GGGAU-3', and the Shine-Dalgarno sequences of the Chlorophenicol acetyl transferase (CAT) and Green Fluorescence Protein (GFP) encoded in the same plasmid were mutated to 5'-AUCCC-3', accordingly. As the result, CAT and GFP can ONLY be translated by ribosomes assembled with 16S rRNA encoded by the plasmid. CAT and GFP are used as reporter genes, and the translation activity of the ribosomes containing the plasmid-encoded 16S rRNA can be quantified by the Minimal Inhibitory Concentration (MIC) determined by CAT and the intensity of the fluorescence from GFP. This system enables researchers to test the

activity of ribosomes containing modified ribosomal RNA encoded by the plasmid without disrupting the normal cellular function maintained by ribosomes produced from chromosome. One additional plasmid carrying the gene of a protein, e.g. ribosomal protein S20, can be introduced into the system. Effect of the protein by inducible overexpression on activity of ribosomes can be determined. The specially engineered anti-Shine-Dalgarno sequence on the 16S rRNA is useful in separation of the small ribosomal subunit containing plasmid-encoded 16S rRNA from those containing chromosomal 16S rRNA by affinity chromatography.

ribosomes, and compromised ribosomal activity. While the mechanism was still unclear, as if all the effects are originated from the different binding affinities between the cognate pairs and non-cognate pairs, or events downstream from the binding also contribute to the observed difference in behavior of the cognate and non-cognate pairs. A model system consisting of h9 RNA oligonucleotides and S20 rproteins were employed in the research described in this chapter, to determine the affinities of the cognate and non-cognate pairs, and show the conformational changes following the binding of pairs of h9 RNA and S20 rprotein.

5.2 Materials and Methods

5.2.1 Materials

Biotinylated h9 RNA oligonucleotides of *E. coli* and *P. aeruginosa* were purchased from Sigma-Aldrich®. T7 RNA polymerase was prepared as described (Davanloo, Rosenberg *et al*. 1984). Template and primer DNA sequences for *in vitro* transcription of mutated *E. coli* and *P. aeruginosa* h9 were purchased from IDT®. Biacore™ CM4 SPR chips, *N*-hydroxsuccinimide, and 1-1-(3-Dimethylaminopropyl)-3-Ethylcarbodiimide Hydrochloride (EDC) were purchased from GE healthcare. Neutravidin and ethylenediamine were purchased from thermo Fisher Scientific.

5.2.2 Preparation of the S20 rproteins from *E. coli* (EcS20) and *P. aeruginosa* (PaS20).

A pET15b plasmid carrying either the gene of EcS20 or PaS20, with a his[6] tag at the N-terminus of the encoded protein, was transformed into BL21 (DE3)

plysS host cells. A single colony from the LB agar plate containing 100 µg/mL ampicillin was used for inoculation of 25 mL culture in LB medium (100 µg/mL ampicillin) at 37 °C overnight. The miniculture was diluted into a large volume of pre-warmed LB medium (100 µg/mL ampicillin) by a ratio of 1:100 (v/v). When the OD_{600} reading reached 0.2, 1 mM IPTG (final concentration) was added into the culture medium and the bacteria were grown for another 90 min at 37 °C. The culture was chilled in ice-water bath once the time point was reached. Bacteria were pelleted by centrifugation (7,000 rpm, 15 min), and a denaturing lysis buffer (50 mM potassium phosphate, 300 mM KCl, 5mM imidazole, 6M urea, pH 8.0) was used to resuspend the bacterial cells. The suspension was passed through a French® press to disrupt the cell membrane and release the cellular content in the cytoplasm. Soluble components from the cell lysis were separated from insoluble debris by centrifugation (15,000 rpm, 20 min, twice) and filtration (0.22 µm) through a syringe tip filter. The supernatant was loaded onto a Ni-IMAC column pre-equilibrated with the lysis buffer. A constant flow peristaltic pump was used to control the flow rate at 2 mL/min. Following the loading of sample supernatant, lysis buffer and wash buffer (50 mM potassium phosphate, 300 mM KCl, 10mM imidazole, 6M urea, pH 8.0) were used to remove the impurities from the resin. The volume of each buffer was 10× of the column volume. In the next step, a buffer with urea gradient (6 – 0.6 M) was used to re-nature the S20 protein on-column. The gradient was created by diluting wash buffer with renaturing buffer (50 mM potassium phosphate, 300 mM KCl, 10mM imidazole, pH 8.0) in a 50 mL tube, where 2 mL of renaturing buffer was added into the 50

mL repertoire every minute (the flow rate was 2 mL/min). The far UV absorbance of eluents at several renaturation time points was measured for estimation of residual urea concentration. S20 protein was eluted from the column in elution buffer (50 mM potassium phosphate, 300 mM KCl, 250mM imidazole, pH 8.0) by gravity, after 1 hr of on-column renaturation. Fractions containing S20 protein were identified by protein SDS-PAGE, and pooled into a dialysis bag for buffer exchange with HBS-EP buffer (10 mM HEPES, 150 mM KCl, 3.4 mM EDTA, 0.005% Surfactant P20, pH 7.4) at a volume ratio of ≤ 1:250 (v/v) overnight. The S20 protein was concentrated by centrifugal filtration in a Millipore centrifugal filter unit with a NMWL of 3kDa. The S20 proteins from *E. coli* and *P. aeruginosa* were purified separately, and characterized with MALDI-tof mass spectrometry.

5.2.3 Preparation of the h9 RNAs from *E. coli* and *P. aeruginosa*

The h9 sequence starts at A179 (*E. coli* numbering), and additional nucleic acid residues were added at the 5' termini, either as a spacer (5'-CCGC-3') when the 5' terminus of the oligonucleotides were biotinylated for SPR chip immobilization, or as a transcription starting sequence (5'-GCA-3') to promote the yield without disrupting the secondary structure of h9s (Milligan, Groebe *et al*. 1987). The h9 RNA samples with 5'-GCA-3' at the 5' terminus (Figure 5.3 B and C) were prepared by *in vitro* transcription with T7 RNA polymerase (Wyatt, Chastain *et al*. 1991). The full length transcripts were purified by 20% (w/v) denaturing PAGE, recovered from gel slices by electroelution, desalted by Sap-pal® column, and characterized by MALDI-tof mass spectrometry.
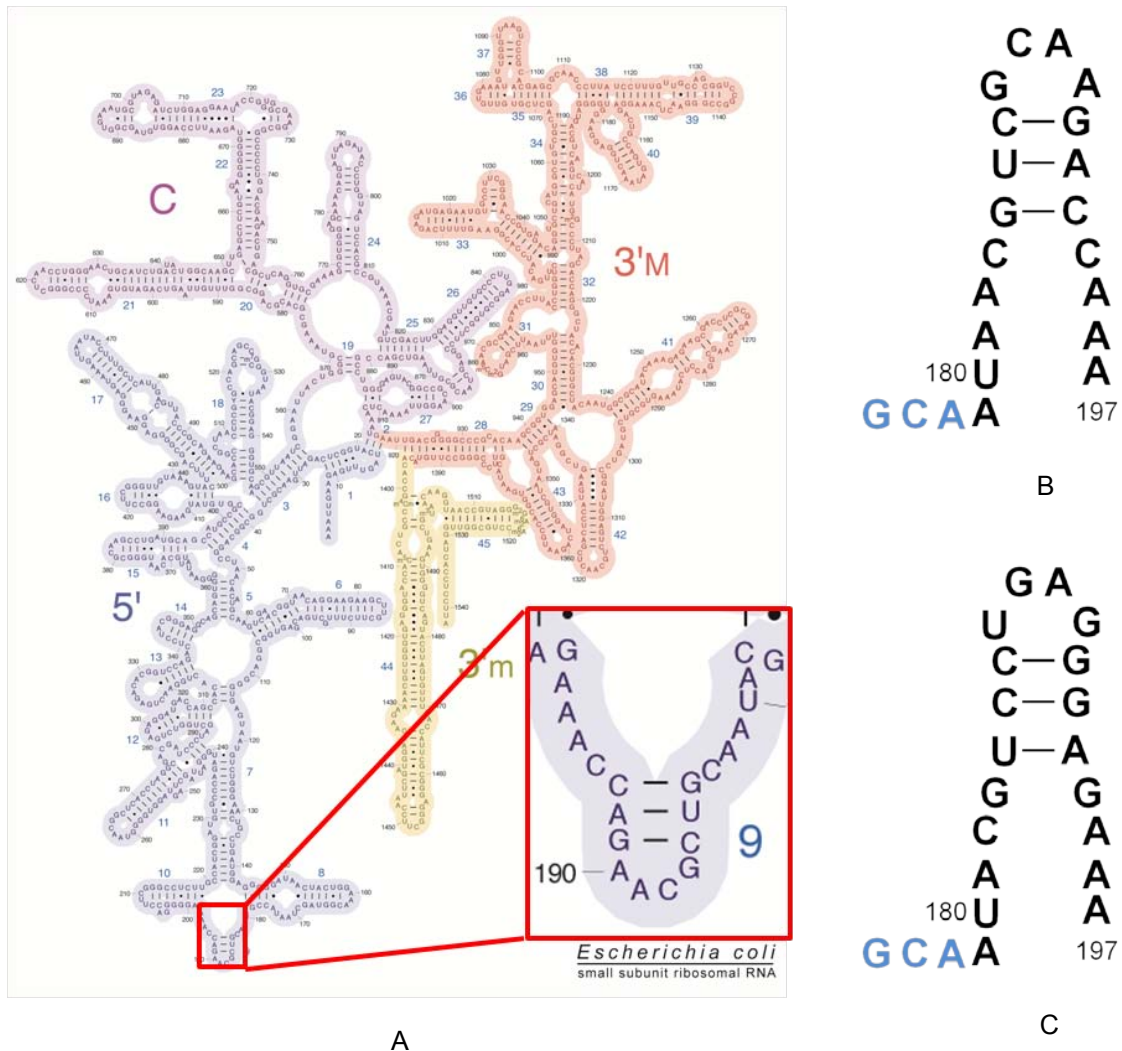
Figure 5.3 Secondary structures of *E. coli* 16S rRNA (A) and helix 9 (black letters, B), compared to the secondary structure of helix 9 from *P. aeruginosa* (black letters, C). The *E. coli* numbering system is used. The blue letters in figures on the right indicate the starting sequence used in *in vitro* transcription of the two RNA oligonucleotides to minimize the dimerization. The picture of the 16S rRNA secondary structure of *E. coli* is from the Noller's lab website. http://rna.ucsc.edu/rnacenter/ribosome_images.html

5.2.4 UV melting experiments on h9 RNAs

An Aviv 14DS UV-vis spectrometer was used in the melting experiments. The absorbance at 260 nm was recorded in a temperature range of 0 - 85 °C provided by a built-in thermocoupler. Five RNA samples dissolved in a melting buffer (15 mM NaCl, 20 mM sodium cacodylate, and 0.5 mM Na_2EDTA, pH 7.0) were transferred into five microcuvettes of four different light path lengths (0.1, 0.2, 0.5, 0.8 cm). The concentration of RNA in each sample was calculated from the absorbance at 85 °C. MeltWin 3.5 was employed in data processing.

5.2.5 NMR spectroscopy

All the NMR experiments were done on a Bruker Avance 700 MHz NMR spectrometer with a HCN cryoprobe. A sample of each h9 RNA (≈ 0.2 mM final concentration) dissolved in an NMR buffer (10 mM potassium phosphate, 50 mM KCl, pH 7.3, and 0.1 mM EDTA-Na_2 to a volume of 300 μL in H_2O/D_2O (90%/10%)) was used for 1D ^1H experiments carried out at 5, 15, 25, and 37 °C, and 2D NOESY experiments at 5 °C. 1D ^1H experiments was also carried out on a sample of cognate pair of h9 RNA and S20 rprotein from *E. coli* at 10 and 25 °C. The sparky 3.114 was used to process the NMR spectra.

5.2.6 Gel shift assay

Gel shift assay was done on a native 15% PAGE in 1× TBE buffer with a Biorad Mini-PROTEAN tetra cell. A 4% PAGE was used as the stacking gel. The voltage of the electrophoresis was set at 200 V (inner buffer chamber was connected to the anode, and the outer buffer chamber was connected to the cathode), and the experiment finished in ≈ 1hr. Pure h9 RNA samples, S20

rprotein samples, and mixtures of cognate and non-cognate pairs of h9 RNA with S20 rprotein samples were prepared in the SPR buffer (10 mM HEPES, 150 mM KCl, 3.4 mM EDTA, 0.005% Surfactant P20, pH 7.4) to a final concentration of 4 μM for h9 RNA and/or S20 rprotein. The gel was stained with GelCode® Blue to visualize the proteins.

5.2.7 SPR experiments

A Biacore CM4 SPR chip was used in the SPR experiments. After activation by EDC and NHS, 210 μL of a neutravidin solution (20 μg/mL final concentration) dissolved in NaOAc buffer (10mM, pH 4.5) was injected, followed by injection of 75 μL of an ethylenediamine solution (1M, pH 8.5 by HCl). Biotinylated-h9 RNAs (B-h9) were dissolved in the SPR buffer (10 mM HEPES, 150 mM KCl, 3.4 mM EDTA, 0.005% Surfactant P20, pH 7.4) and further diluted to ≈ 25 nM (nominal) with the same buffer. The samples were reannealed by being heated to 90 °C for 1 min and cooled down in ice-water bath immediately. The B-h9 samples were injected at a flow rate of 5 μL/min in a portion of 5 μL into the desired cell, respectively, until the desired immobilization level was reached (≈ 81 RU). The S20 rprotein samples were diluted with the SPR buffer to make a series of solutions of 10, 50, 100, 150, 200, 500 nM for each S20 rprotein. The flow rate of the experimental condition was 35 μL/min, and a "Kinject" method was used to inject each sample in triplets. The curves were processed with a "1:1 Langmuir" model in "Kinetics simultaneous ka/kd" module of BIAevaluation 4.1.1.

5.2.8 Circular Dichroism experiments

All the circular dichroism experiments were carried out at RT ($\approx$ 25 °C) in a HEPES-buffered potassium (HBK, 10 mM HEPES, 150 mM KCl, pH 7.4) solution comparable to the SPR buffer. The h9 RNA stock solutions and S20 rprotein stock solutions were diluted into a HBK buffer containing 1M (final concentration) NaCl to make running solutions, then the samples of pure h9 RNA and S20 rprotein, and mixtures of h9 RNA and S20 protein were prepared by mixing the h9 RNA and S20 rprotein running solutions with either the dilution buffer or with cognate/non-cognate pair partner solutions. The final concentration of the h9 RNA and/or S20 rprotein was $\approx$ 17 µM. All the samples (four pure samples: Ech9, Pah9, EcS20, and PaS20; four mixtures: Ech9+EcS20, Pah9+PaS20, Ech9+PaS20, and Pah9+EcS20) were subjected to buffer exchange with the HBK buffer overnight at 4°C to remove the excessive amount of NaCl.

5.3 Results and Discussion

5.3.1 S20 rproteins from *E.coli* and *P. aeruginosa*

Pure S20 rproteins from *E. coli* and *P. aeruginosa* were prepared according to the procedure described above (Figure 5.4 A and B). Molecular weights of both S20 rproteins are $\approx$ 134 Da smaller than those predicted from the amino acid sequences (Figure 5.4 C and D). It has been reported that methionine aminopeptidase (MAP) in the *E. coli* host cells can remove the first methionine residue from a protein/polypeptide substrate, and the first three residues in both recombinant S20 rproteins are N-MGG-C, satisfying the substrate specificity of the MAP (Ben-Bassat, Bauer *et al*. 1987). The difference in the measured and predicted molecular weights of both recombinant S20 rproteins is corresponding

A

B

C

D
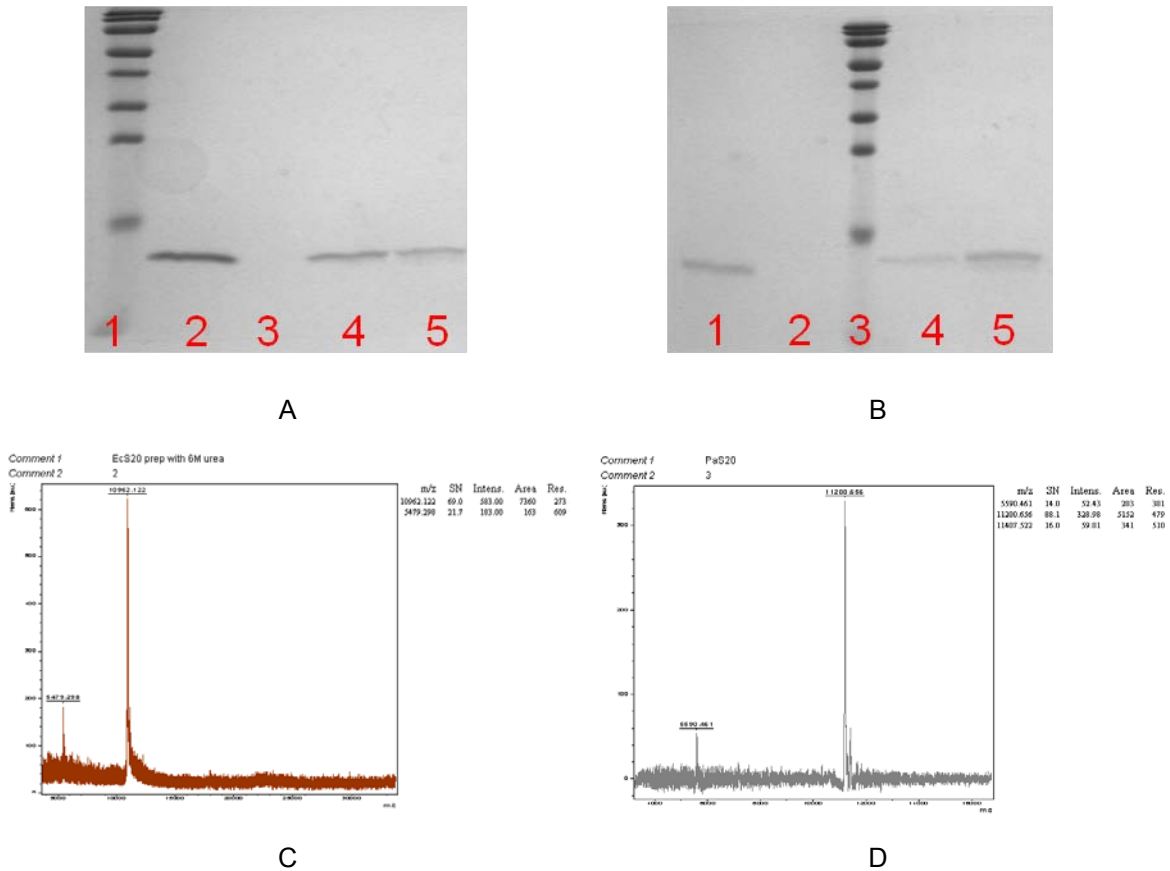
Figure 5.4 Characterization of S20 rproteins of *E. coli* (A and C) and *P. aeruginosa* (B and D) from purification by SDS-PAGE (A and B) and MALDI-tof (C and D). In the SDS-PAGE of *E. coli* S20 rprotein (A), lane 1 to 5 are MW ladder, before renaturation, and fractions 1 to 3 after renaturation, respectively. In the SDS-PAGE of *P. aeruginosa* S20 rprotein (B), lane 1 to 5 are before renaturation, fraction 1 after renaturation, MW ladder, fractions 2 and 3 after renaturation, respectively.

to the loss of a methionine residue.

5.3.2 h9 RNAs are folded into hairpin structure and bind S20 rproteins

The yield of h9 RNAs with 5'-GCA-3' at the 5'- terminus was ≈ 1.5 OD/mL from the T7 polymerase *in vitro* transcription reaction (Milligan, Groebe *et al*. 1987; Wyatt, Chastain *et al*. 1991). The molecular weights of both h9 RNA oligonucleotides were confirmed by MALDI-tof (Figure 5.5 A and B). The assumption that both h9 RNA (GCA) oligonucleotides are folded into a monomer hairpin structure was supported by UV-melting experiments (Figure 5.5 C and D) and NMR experiments (Figure 5.5 E and F). No concentration-dependence in the melting temperature of either h9 RNA was detected within a concentration range from ≈ 3 to 50 µM. The 1D ^{1}H NMR experiments further expanded this range to ≈ 0.1 mM, with indications that hairpin structures dominate the possible conformers. Results from preliminary binding experiments with gel shift assay showed that complexes of cognate and non-cognate pairs of h9 RNA and S20 rprotein could form (Figure 5.6).

5.3.3 Determination of dissociation constants by SPR experiments

In the 99 amino acid residues of EcS20 rprotein, 21 residues are positively charged at neutral pH, and 5 residues are negatively charged. The numbers for PaS20 are 103, 21, and 6, respectively. These positive charges carried by the analytes (S20 rproteins) promote non-specific binding greatly, since a SA chip (CM5 chip immobilized with streptavidin) surface usually carries negative charges from unreacted carboxylic group at neutral pH. A response of about 800 RU was observed with ≈ 100nM of S20 rprotein analyte flowing through the blank

A

B

C

D

E

F
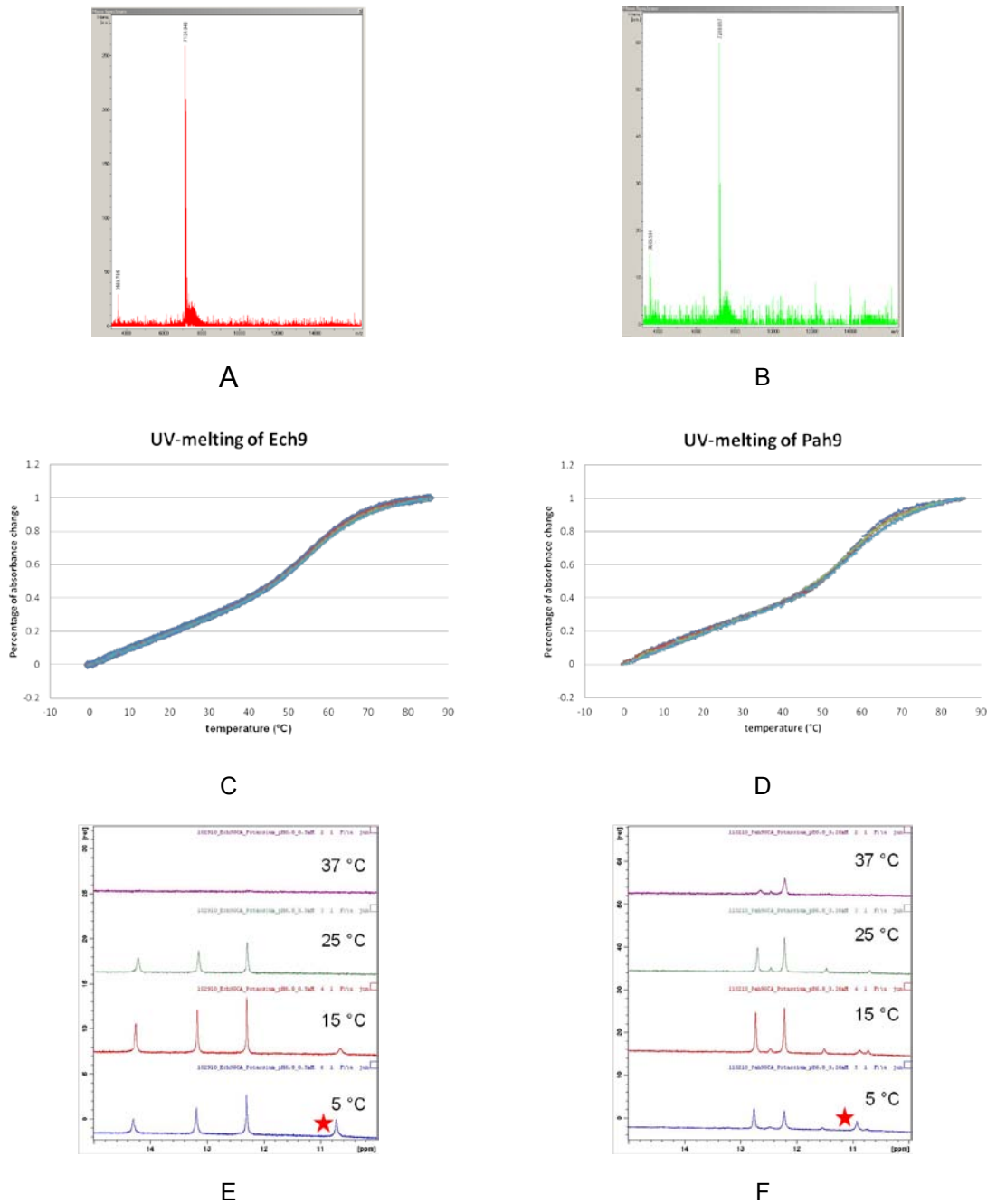
Figure 5.5 Characterization of 5'-GCA-3' h9 RNAs from *E. coli* (left) and *P. aeruginosa* (right). The MALDI-tof spectra (upper) show that the molecular weights of Ech9 and Pah9 are 7135 and 7208 Da, respectively, less than 1 Da away from the predicted molecular weights. It is revealed in the UV-melting

spectra (middle) of Ech9 (five samples from 3.7 to 53.2 µM) and Pah9 (five samples from 2.5 to 39.8 µM) that the normalized melting curves of each h9 RNA overlap nicely, and the difference between melting temperatures for each h9 RNA is within 1 °C. Three and two major peaks were observed between 12 and 15 ppm in the 1D ^{1}H NMR spectra of Ech9 (lower left) and Pah9 (lower right), respectively. Even though in both h9 RNAs, three Watson-Crick base pairs are suggested by phylogenetic studies (Figure 5.3 right), an A•U base pair is in the stem closing position in Pah9, which could be much less stable than a G•C base pair, or the folding of the Pah9 dangling ends region is different from that of the Ech9, which could also contribute to the loss of a N3H resonance in the spectrum. A resonance close to 11 ppm was observed in both cases (red star), which could be from a non-Watson-Crick base pair in the dangling ends region.
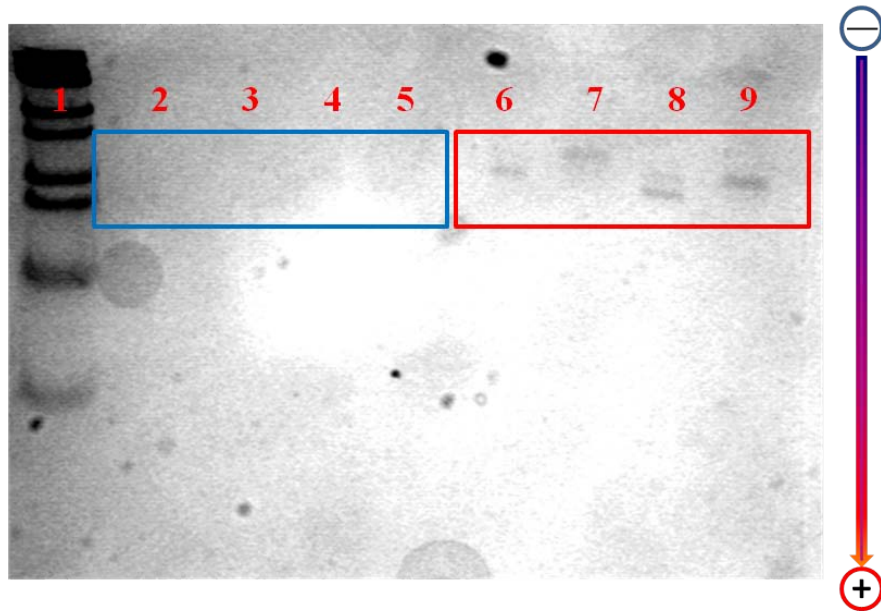
Figure 5.6 Gel shift assay on binding of cognate and non-cognate pairs of h9 RNA and S20 rproteins. Lanes 1 -9 are MW ladder, Ech9, Pah9, EcS20, PaS20, mixture of Ech9 + EcS20, mixture of Pah9 + EcS20, mixture of Ech9 + PaS20, mixture of Pah9 + PaS20. The gel was stained with GelCode® Blue. No band was observed in lanes 2 and 3, since only h9 RNAs were loaded into these two lanes. No band was observed in lanes 4 and 5, either, since the pI of both S20 rproteins are close to 11, and did not migrate into the gel with a net positive charge at experimental condition. Only when S20 rproteins formed complex with h9 RNAs, and the positive charges (EcS20 has 16 net positive charges, and PaS20 has 15 net positive charges) of the S20 rproteins were neutralized by the negative charges of the h9 RNAs (22 phosphate in the backbone in each h9 RNA), S20 rprotein could co-migrate into the gel and be visualized by protein staining reagent.

cell, where no bio-h9 RNA was immobilized, and the target maximal response was expected to be around 100 RU for kinetic studies. To minimize the non-specific binding, a CM4 chip, which has only 1/3 of the carboxylic acid modification capacity of CM5 chip, was used, and ethylenediamine, which has an additional free amino group carrying a positive charge, was applied to deactivate the activated carboxylic group after immobilization of the neutravidin, whose specificity is higher than that of the streptavidin (Figure 5.7). Smaller carboxylic acid capacity of the CM4 chip helped to reduce the volume of residual carboxylic group after immobilization, and the positively charged amino group of the ethylenediamine helped to neutralize the negative charges from residual carboxylic acid group. This new regime helped to reduce the non-specific binding in the blank cell to ≈ 20 RU, compared to 800 RU at the same condition.

Each group consisting of one of the triplet binding curves of all analyte concentrations was fit to a 1:1 Langmuir model with simultaneous k_a/k_d kinetics (Figure 5.8). One set of k_a/k_d and K_d were derived from each fitting, and the results were reported in the format of "the average ± standard deviation" derived from the triplet repetition (Table 5.1). It was reported that, the dissociation constant between 16S rRNA and S20 rprotein from *E. coli* was ≈ 83 nM, as determined by dialysis (Donly and Mackie 1988). This observation is supportive to the assumption that, interaction between h9 oligonucleotide and S20 rprotein (K_d = 77 nM) resembles that between the binding of 16S rRNA and S20 rprotein (K_d = 83 nM), from *E. coli*. The dissociation constants (K_d) of all the four pairs are within a factor of 6, ranging from 51 nM (Pah9 + EcS20) to 312 nM (Ech9 +

PaS20). This could be used to explain the phenomena observed in the gel shift assay. When the concentration of each h9 RNA and S20 was ≈4 μM (final concentration) in the mixture, the concentrations of the pairs containing S20 rproteins (stained with GelCode® blue) were from ≈ 3 to 3.6 μM, a difference indistinguishable from protein staining, and all the bands showed similar intensity.

It was revealed in the comparison of dissociation constants (K_d) that EcS20 binds stronger to both h9 RNAs than PaS20 (Table 5.1). The rate constants of association and dissociation processes favor binding of EcS20 to both h9 RNAs (Table 5.1). This observation results partially, if not completely, from the fact that there is potentially one more net positive charge on EcS20 (21 – 5 = 16) than PaS20 (21 – 6 =15) at experimental condition (pH 7.4). The his[6] tag and linker region has no contribution to the charges, since all of the charged residues are included in wild type S20 rprotein sequences. When the two h9 RNAs are compared, it was clearly shown that Pah9 binds S20 rproteins stronger than Ech9 (Table 5.1), while this effect seems to be less significant than the difference between the S20 rproteins. It could be postulated that Ech9 is a weaker binder compared to Pah9, and the EcS20 is a strong binder compared to PaS20, so the cognate pairs show intermediate binding affinity, while the non-cognate pairs show either too high or too low a binding affinity.

5.3.4 No significant conformational change was observed in CD spectra
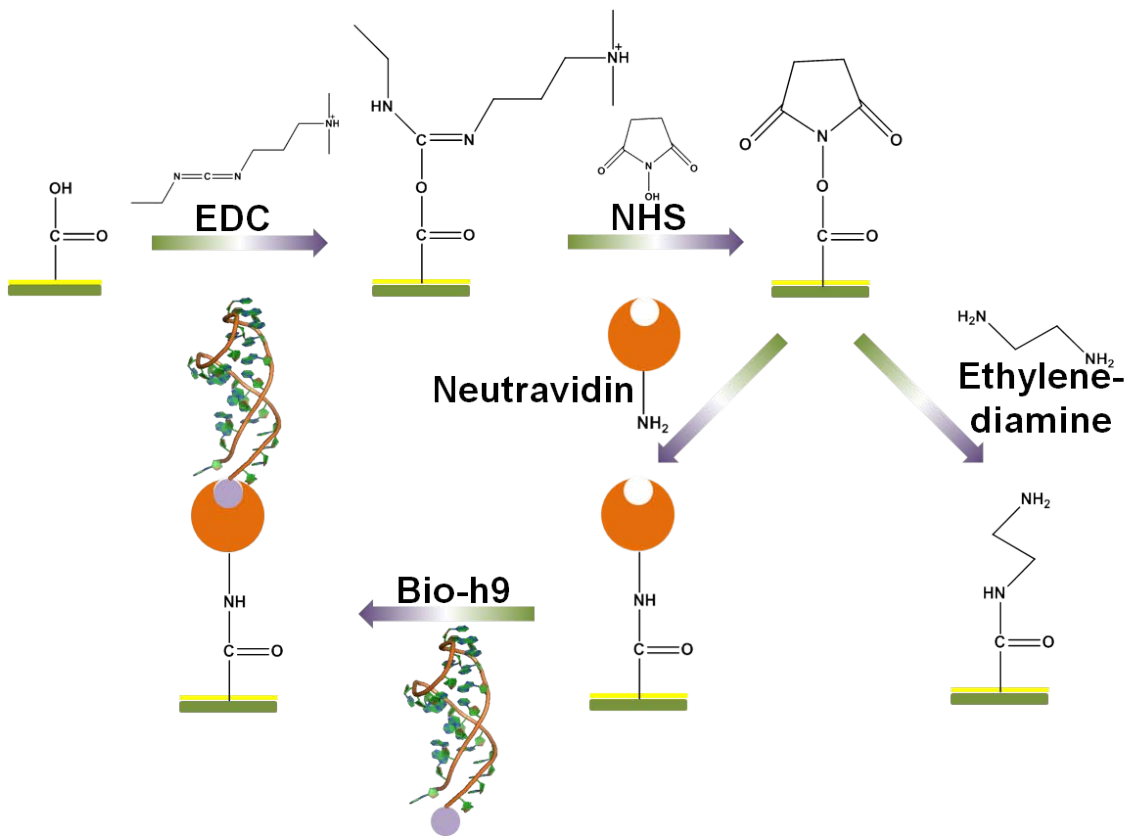
Figure 5.7 Immobilization of bio-h9 RNA on a SPR chip. After being activated by EDC and NHS, the chip surface was immobilized with neutravidin, and the residual activated carboxylic acid groups were deactivated with ethylenediamine. Lastly, Bio-h9 RNA was immobilized by interaction with the neutravidin.
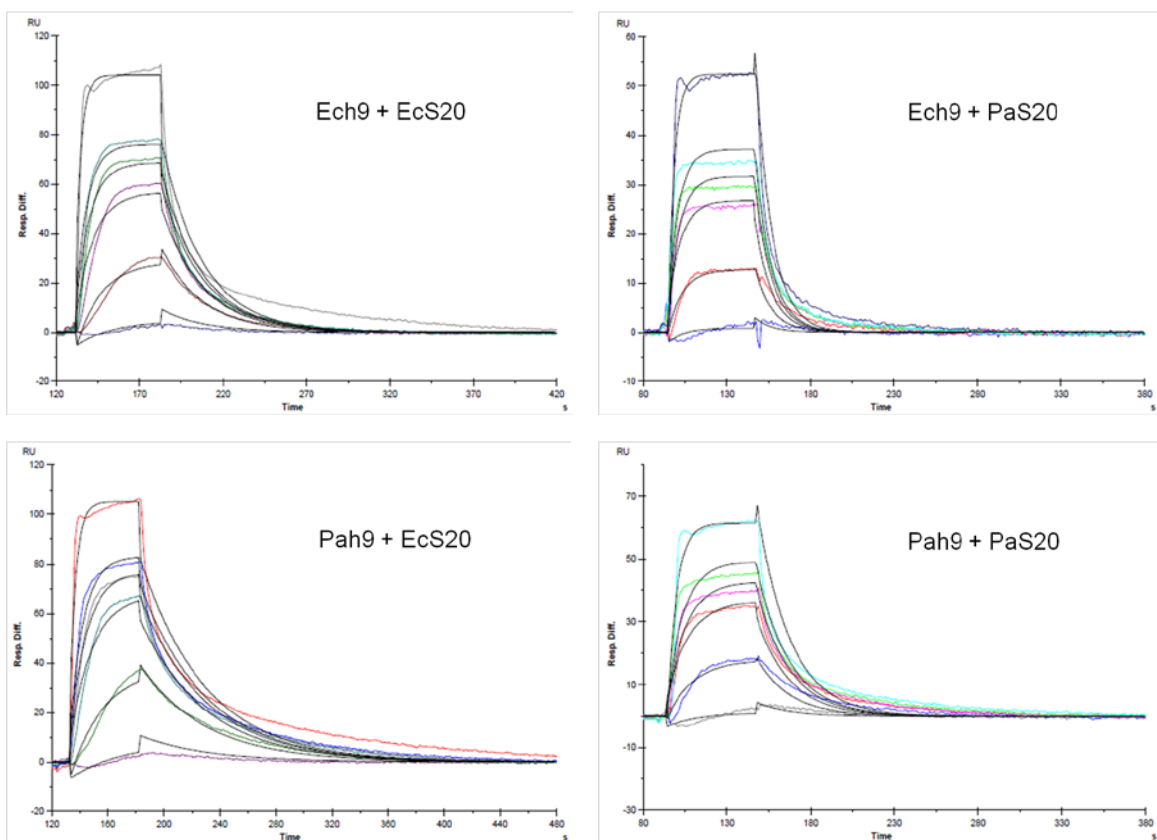
Figure 5.8 Binding curves of Surface Plasmon Resonance experiments. The curves of each group (pair of h9 RNA and S20 rprotein) were fit to a 1:1 Langmuir model simultaneously (global fitting) with both the association and dissociation phases.

Table 5.1 Kinetic and dissociation constants from the SPR experiments.

h9-S20 pair	k_a (10^5/Ms)	k_d (10^{-2}/s)	K_d (nM)
Ech9 + EcS20	5.3 ± 0.13	4.1 ± 0.04	77 ± 2
Ech9 + PaS20	3.2 ± 0.20	10 ± 0.26	312 ± 13
Pah9 + EcS20	4.1 ± 0.09	2.0 ± 0.08	51 ± 1
Pah9 + PaS20	2.6 ± 0.02	5.2 ± 0.18	198 ± 7

A buffer comparable to that of the SPR experiments was employed to provide consistency. Spectra of each pure h9 RNA and S20 rprotein, and each mixture of h9 RNA and S20 rprotein pair were acquired at RT. Since no strong chromophore is in either S20 protein, any absorbance in the 260 – 280 nm region is from the aromatic base rings of the h9 RNAs. No net CD signal was observed in the 260 – 280 nm region when the h9 RNA spectrum was subtracted from the corresponding S20 rprotein + h9 RNA spectrum (Figure 5.9). This suggests that there may be no significant conformational changes of the h9 RNAs on binding to the S20 rproteins, or the conformational changes do not contribute to the CD changes. No chemical shift change was observed in the 1D ^{1}H NMR spectra of the RNA exchangeable proton region on binding of h9 RNA to S20 rprotein, which indicates that the conformation of the duplex region may not need to undergo a significant readjustment. CD spectra in the far-UV region show that conformational changes of both S20 rproteins occur on binding of the S20 rprotein to the h9 RNAs, while the identity of the h9 RNA does not affect the conformational change by much. This trend was also observed in the comparison of dissociation constants involving the same S20 rprotein. For the same S20 rprotein, the difference in thermodynamic stability contribution ($\Delta G°_{37}$) of the association of S20 + Pah9 and S20 + Ech9 is only ≈ 0.25 kcal/mol (equivalent to a ratio of dissociation constant of 0.65), one tenth of the energy contribution from a single hydrogen bond (Hao 2006).

5.3.5 Binding affinity is correlated with the subunit assembly and ribosome association
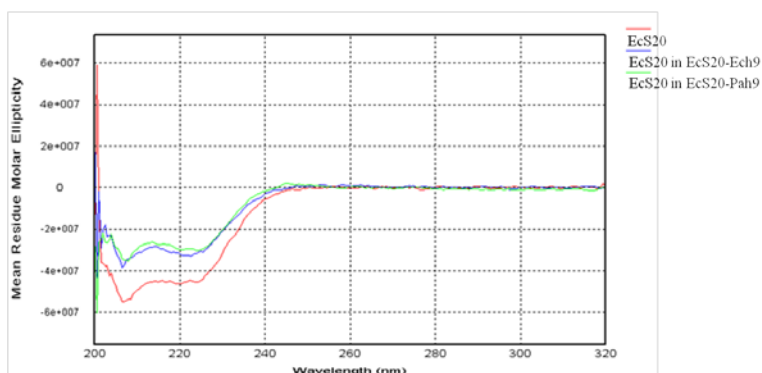
A
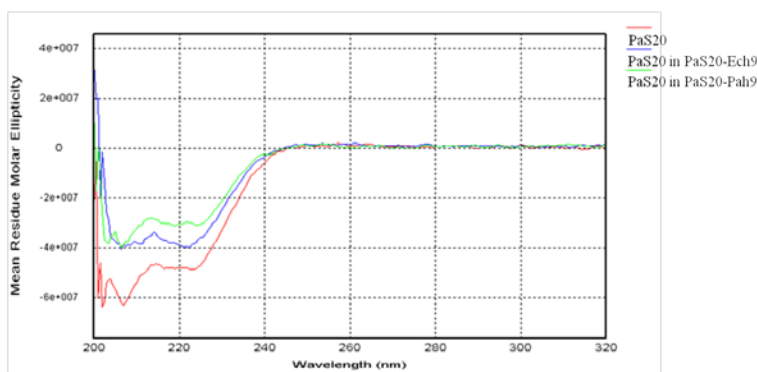
B

Figure 5.9 Difference spectra of EcS20 (A) and PaS20 (B) binding to cognate and non-cognate h9 RNAs. The red curves show the CD spectra of pure S20 proteins. Two troughs at 208 and 222 nm were observed in each curve, typical for proteins containing α-helix component. No net CD was observed in the 260 – 280 nm region in the S20 rprotein CD spectra, possibily due to the lack of a strong chromophore in either S20 rprotein. The green curves and blue curves show the mixture of S20 rprotein binding to Ech9 and Pah9, respectively. Change in trough intensity on binding of h9 RNA indicates that conformational change may happen in S20 rproteins on binding to h9 RNA, while the binding

event would not necessarily cause significant conformational change in the h9 RNAs or the CD signal from the conformational changes of the h9 RNAs would cancel off each other to give null net CD. No significant difference was observed when the two spectra of an S20 rprotein binding to two h9 RNAs were compared. This observation agrees with the small difference between the binding affinity of each S20 rprotein to the two h9 RNAs (Table 5.1).

Binding affinity of the h9 RNA and S20 rprotein pair is directly correlated to the incorporation of S20 rprotein into the small ribosomal subunit, which can associate with the large subunit. When the fluorescence intensities of the cognate pairs or non-cognate pairs of h9 RNA and S20 rprotein are compared respectively, stronger fluorescence intensity is always seen with the stronger binding affinity (Table 5.2, column "Flu/OD^a") (Lamichhane 2009). It is most significant when both pairs are non-cognate pair, stronger interactions between the Pa16S rRNA and EcS20 rprotein enabled more efficient incorporation of the EcS20-GFP fusion protein into the small subunit, than that of the Ec16S rRNA + PaS20 pair. While the much less significant difference seen between the two cognate pairs could result from either experimental error or suboptimal assembly due to incorporation of all *E. coli* rproteins, except PaS20, into a small subunit with the Pa16S rRNA. An alternative possibility is that the host cell constitutively expressed EcS20, which competed against PaS20-GFP to bind with Pa16S rRNA, and a stronger affinity between the Pa16S rRNA and EcS20 than between PaS16 rRNA + PaS20 made the incorporation of PaS20-GFP into a small subunit with Pa16S rRNA less efficient. While, since the expression level of the chromosomal EcS20 should be much lower than that of the over expressed PaS20-GFP, this difference was not dramatic.

Even though the correlation between binding affinity and assembly of the small subunit was revealed by the aforementioned comparisons, the binding affinity itself does not determines the assembly efficiency, since cognate pairs of 16S rRNA and S20 rprotein showed higher fluorescence intensities, compared to

Table 5.2 Results from biological experiments on h9-S20 interactions.

	K_d (nM)	Flu/OD[a]	16S rRNA %[b]	MIC (µg/mL)[c]	GFP (% WT)[d]
Ech9 + EcS20	77 ± 2	161	44 (30S) 42 (70S)	600	100
Ech9 + PaS20	312 ± 13	60	38 (30S) 36 (70S)	650	96
Pah9 + EcS20	51 ± 1	114	31 (30S) 36 (70S)	200	44
Pah9 + PaS20	198 ± 7	154	36 (30S) 43 (70S)	500	96

a: Fluorescence intensity from S20-GFP over 16S rRNA absorbance ratios were quantified from preparations of plasmid-encoded 16S rRNA containing small ribosomal subunit .

b: The percentages of plasmid-encoded 16S RNA in the total 16S rRNA (plasmid + chromosomal) from both small subunit preparations and whole ribosome preparations were quantified by RT-PCR with radio labeling.

c: Chloramphenicol was used to test the cellular expression level of chloramphenicol actyltransferase synthesized in the plasmid-encoded 16S rRNA containing ribosome.

d: Fluorescence intensity was used to quantify the cellular expression level of green fluorescence protein synthesized in the plasmid-encoded 16S rRNA containing ribosome.

those of the non-cognate pairs. This trend was also observed in other *in vitro* and *in vivo* experiments (Table 5.2). Compared to the wild type pair control (Ec16S rRNA + EcS20 rprotein), assembly of small subunits with Pa16S rRNA was compromised to different extent, depending on the cognate or non-cognate S20 rprotein was co-expressed, and association of the complete ribosomes alleviated the effect of incorporation of a Pa16S rRNA into a small subunit with all *E. coli* ribosomal proteins (the S20 was from *P. aeruginosa* in Pa16S rRNA + PaS20). In the case of Ec16S rRNA + PaS20, since there was the chromosomal EcS20 expressed, and the affinity between cognate pair of 16S rRNA and S20 rprotein from *E. coli* is stronger than that of Ec16S rRNA + PaS20 pair, the endogenous EcS20 complemented the structure and function of the ribosomes, where the 16S rRNA incorporation level is higher than the other non-cognate pair, and the ribosomal activity was close to wild type. Even though Pa16S rRNA + EcS20 form a non-cognate pair, the EcS20-GFP was able to be incorporated with Pa16S rRNA into the small ribosomal subunit, and the ribosomal function was shown to be suboptimal, instead of completely lost.

5.4 Conclusion

The assembly of ribosomal subunits is a highly ordered, co-transcriptional process that also affects the downstream association of ribosomal subunits (McCarthy, Britten *et al*. 1962; Mangiarotti, Apirion *et al*. 1968; Osawa, Otaka *et al*. 1969; Mizushima and Nomura 1970; Homann and Nierhaus 1971; Lindahl 1973; Nierhaus, Bordasch *et al*. 1973; Held, Ballou *et al*. 1974; de Narvaez and Schaup 1979; Herold and Nierhaus 1987; Jakel and Gorlich 1998; Grondek and

Culver 2004; Dutca and Culver 2008; Shajani, Sykes *et al*. 2011). S20 rprotein, a primary ribosome binding protein in the assembly of the small subunit, binds into the 5' region of the 16S rRNA first, possibly stabilized primarily by interactions with the h9 region. A direct contact would be formed at a much later stage between h44 of the 16S rRNA and S20 rprotein, when the transcription of the 16S rRNA 3' minor region is finished. Interactions between the S20 rprotein and h44 may play an important role in quality control, possibly by gauging the minor conformational differences of S20 rproteins in a cognate or non-cognate S20 + h9 (in the 16S rRNA context) complex. The conformational differences in h44 resulted from interactions of cognate and non-cogante S20 rprotein + h9 RNA pairs may potentially exert effects on the ribosomal association level, and the association of ribosomes, on the other hand, helped increase the level of incorporation of Pa16S rRNA into the small subunit, together with both EcS20 and PaS20. Even though no direct evidence shows that h9 RNA undergoes a significant conformational change on binding to the S20 protein, and the two S20 rproteins behaved very similar to each other as seen in the CD spectra, the possibility should not be excluded that conformational changes of either the h9 RNAs or S20 rproteins could happen on binding of the S20 rproteins to the h44 at a late stage of subunit assembly.

Binding affinity between the h9 RNA and S20 rprotein is not the decisive factor for assembly of subunit with cognate and non-cognate pair of h9 (16S rRNA) + S20 rprotein, and the downstream events, including binding of the h44 to S20, association of ribosomes, and translation process, are able to readjust

the structure and activity of the ribosomes (Figure 5.10). The methods established in the project could be used for further antibiotic study. For instance, should an antibiotic candidate targeting h9 be discovered, the SPR experiment regime could be used to study the binding kinetics and thermodynamics of interactions between the antibiotic molecule and the h9 RNA(s) by direct binding of the antibiotic to the h9 rRNA or competition of the binding between the antibiotic and S20 rprotein.
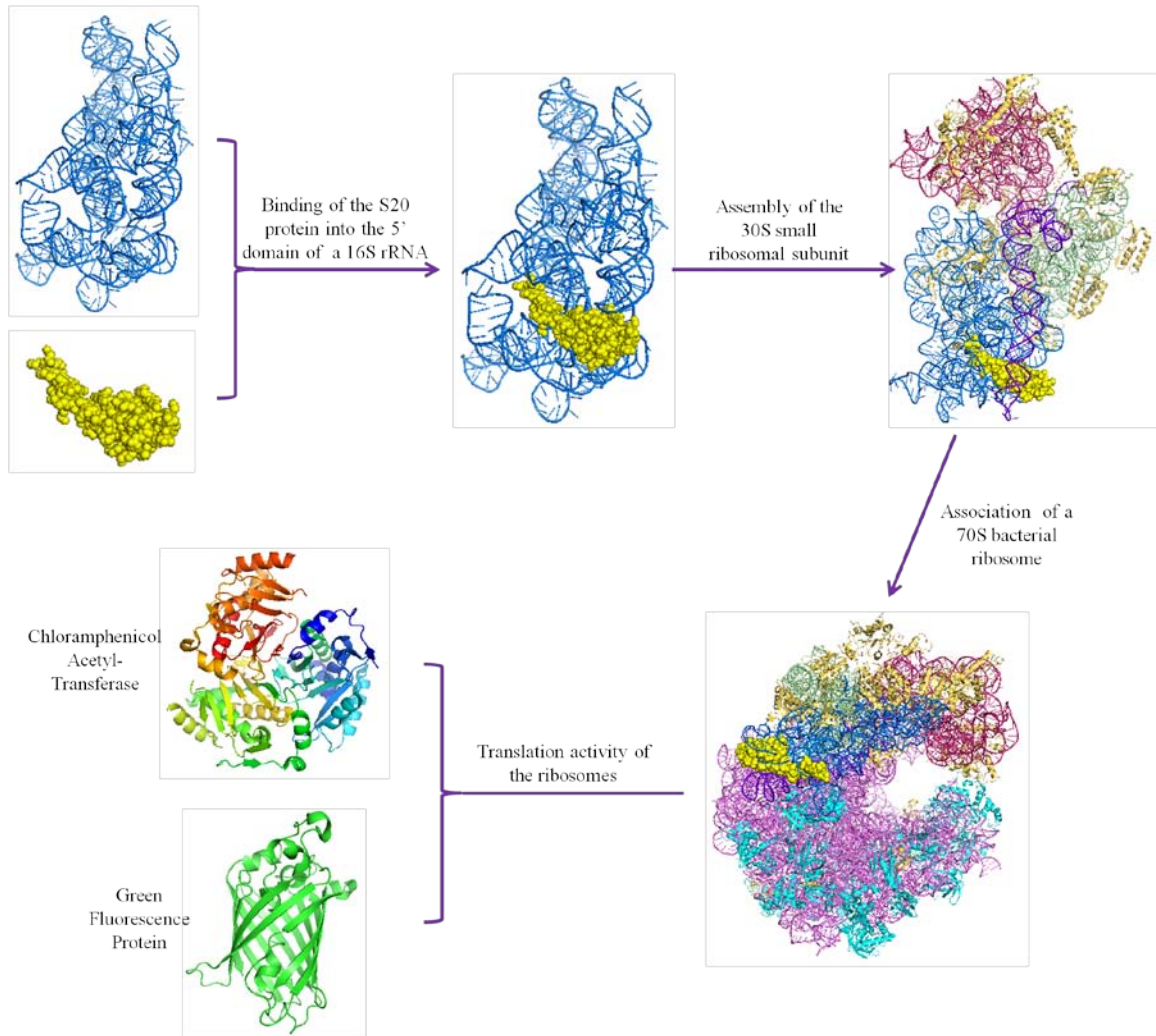
Figure 5.10. The cascade of effects starting with the binding of h9 of 16S rRNA and S20 rprotein.

APPENDIX 1

Structural Restraints for the H69UUU

Distance restraints for the H69UUU

! Base flipping control in Global fold

assign (residue 1 and name O6) (residue 2 and name O6) 4.6 0.5 0.5

assign (residue 1 and name N2) (residue 2 and name N2) 4.8 0.5 0.5

assign (residue 2 and name O6) (residue 3 and name N4) 4.7 0.5 0.5

assign (residue 2 and name N2) (residue 3 and name O2) 3.5 0.5 0.5

assign (residue 3 and name N4) (residue 4 and name N4) 3.7 0.5 0.5

assign (residue 3 and name O2) (residue 4 and name O2) 4.8 0.5 0.5

assign (residue 4 and name N4) (residue 5 and name O6) 5.4 0.5 0.5

assign (residue 4 and name O2) (residue 5 and name N2) 6.5 0.5 0.5

assign (residue 19 and name O2) (residue 18 and name O2) 5.0 0.5 0.5

assign (residue 19 and name N4) (residue 18 and name O4) 4.9 0.5 0.5

assign (residue 18 and name O2) (residue 17 and name N2) 3.2 0.5 0.5

assign (residue 18 and name O4) (residue 17 and name O6) 3.9 0.5 0.5

assign (residue 17 and name O6) (residue 16 and name O6) 4.9 0.5 0.5

assign (residue 17 and name N2) (residue 16 and name N2) 4.6 0.5 0.5

assign (residue 16 and name O6) (residue 15 and name N4) 5.3 0.5 0.5

assign (residue 16 and name N2) (residue 15 and name O2) 6.4 0.5 0.5

!Stem stacking distances

assign (residue 1 and name H1) (residue 2 and name H1) 4.5 0.5 0.5

assign (residue 1 and name H1) (residue 18 and name H3) 5.3 0.5 0.5

assign (residue 17 and name H1) (residue 18 and name H3) 3.6 0.5 0.5

assign (residue 2 and name H1) (residue 17 and name H1) 4.5 0.5 0.5

assign (residue 17 and name H1) (residue 16 and name H1) 4.5 0.5 0.5

assign (residue 16 and name H1) (residue 5 and name H1) 4.4 0.5 0.5

! NOEs from H2O NOESY

```
assign (residue 1 and name H1) (residue 19 and name H42) 2.4 0.5 0.5

assign (residue 2 and name H1) (residue 19 and name H42) 4.8 0.5 0.5

assign (residue 18 and name H3) (residue 19 and name H42) 4.6 0.5 0.5

assign (residue 2 and name H1) (residue 17 and name H22) 4.5 0.5 0.5

assign (residue 2 and name H1) (residue 19 and name H42) 4.2 0.5 0.5

assign (residue 18 and name H3) (residue 17 and name H22) 4.4 0.5 0.5

assign (residue 2 and name H1) (residue 18 and name H3) 2.1 0.5 0.5

assign (residue 3 and name H42) (residue 17 and name H1) 2.5 0.5 0.5

assign (residue 4 and name H42) (residue 17 and name H1) 3.3 0.5 0.5

assign (residue 3 and name H42) (residue 16 and name H1) 5.8 0.5 0.5

assign (residue 4 and name H42) (residue 16 and name H1) 2.8 0.5 0.5

assign (residue 5 and name H1) (residue 15 and name H42) 2.6 0.5 0.5

assign (residue 4 and name H42) (residue 15 and name H42) 6.8 0.5 0.5

! Carbon distances for the Base pairs

assign (residue 1 and name C1') (residue 19 and name C1') 10.7 0.2 0.2

assign (residue 1 and name C8) (residue 19 and name C6) 9.7 0.2 0.2

assign (residue 2 and name C1') (residue 18 and name C1') 10.7 0.2 0.2

assign (residue 2 and name C8) (residue 18 and name C6) 9.7 0.2 0.2

assign (residue 3 and name C1') (residue 17 and name C1') 10.7 0.2 0.2

assign (residue 3 and name C6) (residue 17 and name C8) 9.7 0.2 0.2

assign (residue 4 and name C1') (residue 16 and name C1') 10.7 0.2 0.2

assign (residue 4 and name C6) (residue 16 and name C8) 9.7 0.2 0.2

assign (residue 5 and name C1') (residue 15 and name C1') 10.7 0.2 0.2

assign (residue 5 and name C8) (residue 15 and name C6) 9.7 0.2 0.2

!Base pair dists

! for G1/ C19  base pair

assign (residue 1 and name N1) (residue 19 and name N3) 2.7 0.2 0.2

assign (residue 1 and name O6) (residue 19 and name N4) 2.7 0.2 0.2

assign (residue 1 and name N2) (residue 19 and name O2) 2.8 0.2 0.2
```

```
! for G2/U18 base pair

assign (residue 2 and name N1) (residue 18 and name O2) 3.1 0.2 0.2

assign (residue 2 and name O6) (residue 18 and name N3) 2.9 0.2 0.2

! for C3/ G17  base pair

assign (residue 3 and name N3) (residue 17 and name N1) 2.7 0.2 0.2

assign (residue 3 and name N4) (residue 17 and name O6) 2.7 0.2 0.2

assign (residue 3 and name O2) (residue 17 and name N2) 2.8 0.2 0.2

! for C4/ G16  base pair

assign (residue 4 and name N3) (residue 16 and name N1) 2.7 0.2 0.2

assign (residue 4 and name N4) (residue 16 and name O6) 2.7 0.2 0.2

assign (residue 4 and name O2) (residue 16 and name N2) 2.8 0.2 0.2

! for G5/ C15  base pair

assign (residue 5 and name N1) (residue 15 and name N3) 2.7 0.2 0.2

assign (residue 5 and name O6) (residue 15 and name N4) 2.7 0.2 0.2

assign (residue 5 and name N2) (residue 15 and name O2) 2.8 0.2 0.2

! Base to Base constraints, seen only in 400ms 2D NOESY.

assign (residue 1 and name H8) (residue 2 and name H8)   4.5 1.5 1.5

assign (residue 2 and name H8) (residue 3 and name H6)   4.5 1.5 1.5

assign (residue 3 and name H6) (residue 4 and name H6)   4.5 1.5 1.5

assign (residue 4 and name H6) (residue 5 and name H8)   4.5 1.5 1.5

assign (residue 5 and name H8) (residue 6 and name H6)   4.5 1.5 1.5

assign (residue 6 and name H6) (residue 7 and name H8)   4.5 1.5 1.5

assign (residue 7 and name H8) (residue 8 and name H8)   4.5 1.5 1.5

assign (residue 8 and name H8) (residue 9 and name H6)   4.7 1.0 1.0

assign (residue 9 and name H6) (residue 10 and name H6)  4.0 0.0 100.0

assign (residue 10 and name H6) (residue 11 and name H8)  4.0 0.0 100.0

assign (residue 11 and name H8) (residue 12 and name H6)  4.0 0.0 100.0

assign (residue 12 and name H6) (residue 13 and name H8)  4.0 0.0 100.0

assign (residue 13 and name H8) (residue 14 and name H8) 4.2 1.0 1.0
```

```
assign (residue 14 and name H8) (residue 15 and name H6) 5.0 1.0 1.0

assign (residue 15 and name H6) (residue 16 and name H8) 4.5 1.5 1.5

assign (residue 16 and name H8) (residue 17 and name H8) 4.5 1.5 1.5

assign (residue 17 and name H8) (residue 18 and name H6) 4.5 1.5 1.5

! residue 1 GUA (1906)

! Intra residue

assign (residue 1 and name H8) (residue 1 and name H1') 3.2 0.7 0.6

assign (residue 1 and name H8) (residue 1 and name H2') 3.4 0.9 0.4

assign (residue 1 and name H8) (residue 1 and name H3') 2.6 0.8 0.3

assign (residue 1 and name H8) (residue 1 and name H4') 4.1 0.9 1.1

assign (residue 1 and name H8) (residue 1 and name H5') 3.0 0.5 0.8

assign (residue 1 and name H8) (residue 1 and name H5'') 3.3 0.8 0.5

! residue 2 GUA (1907)

! Intra residue

assign (residue 2 and name H8) (residue 2 and name H1') 3.8 0.8 1.2

assign (residue 2 and name H8) (residue 2 and name H2') 3.6 0.6 1.4

assign (residue 2 and name H8) (residue 2 and name H3') 2.4 0.6 0.5

assign (residue 2 and name H8) (residue 2 and name H4') 4.0 1.0 1.0

assign (residue 2 and name H8) (residue 2 and name H5'') 2.8 0.3 1.0

assign (residue 2 and name H8) (residue 2 and name H5') 3.6 0.6 1.4

! Inter residue

assign (residue 1 and name H1') (residue 2 and name H8) 3.6 0.6 1.4

assign (residue 1 and name H2') (residue 2 and name H8) 2.5 0.7 0.4

assign (residue 1 and name H3') (residue 2 and name H8) 4.1 1.1 0.9

! residue 3 CYT (1908)

! Intra residue

assign (residue 3 and name H6) (residue 3 and name H1') 3.7 0.7 1.3

assign (residue 3 and name H6) (residue 3 and name H2') 4.0 1.0 1.0

assign (residue 3 and name H6) (residue 3 and name H3') 2.6 0.8 0.3
```

```
assign (residue 3 and name H6) (residue 3 and name H4') 4.0 1.0 1.0

assign (residue 3 and name H6) (residue 3 and name H5') 3.1 0.6 0.7

assign (residue 3 and name H6) (residue 3 and name H5'') 4.0 1.0 1.0

assign (residue 3 and name H5) (residue 3 and name H2') 4.5 1.5 1.5

assign (residue 3 and name H5) (residue 3 and name H3') 4.5 1.5 1.5

! Inter residue

assign (residue 3 and name H6) (residue 2 and name H1') 4.5 1.5 1.5

assign (residue 3 and name H6) (residue 2 and name H2') 2.5 0.7 0.4

assign (residue 3 and name H6) (residue 2 and name H3') 3.0 0.5 0.8

assign (residue 3 and name H6) (residue 2 and name H5'') 4.5 1.5 1.5

assign (residue 3 and name H5) (residue 2 and name H2') 3.8 0.8 1.2

assign (residue 3 and name H5) (residue 2 and name H3') 3.5 1.0 0.3

assign (residue 3 and name H5) (residue 2 and name H8) 4.0 1.0 1.0

assign (residue 3 and name H1') (residue 2 and name H2') 4.0 1.0 1.0

! residue 4 CYT (1909)

! Intra residue

assign (residue 4 and name H6) (residue 4 and name H1') 3.2 0.7 0.6

assign (residue 4 and name H6) (residue 4 and name H2') 3.5 0.5 1.5

assign (residue 4 and name H6) (residue 4 and name H3') 2.6 0.8 0.3

assign (residue 4 and name H6) (residue 4 and name H4') 4.2 1.2 0.8

assign (residue 4 and name H6) (residue 4 and name H5'') 2.7 0.2 1.1

assign (residue 4 and name H6) (residue 4 and name H5') 3.5 0.5 1.5

! Inter residue

assign (residue 4 and name H6) (residue 3 and name H1') 4.0 1.0 2.0

assign (residue 4 and name H6) (residue 3 and name H2') 2.5 0.7 0.4

assign (residue 4 and name H5) (residue 3 and name H1') 4.0 0.5 1.0

assign (residue 4 and name H5) (residue 3 and name H2') 3.5 0.5 1.5

assign (residue 4 and name H5) (residue 3 and name H3') 3.4 0.6 0.6

assign (residue 4 and name H1') (residue 3 and name H2') 3.7 0.7 1.3
```

! residue 5 GUA (1910)

! Intra residue

assign (residue 5 and name H8) (residue 5 and name H1') 3.8 0.8 1.2

assign (residue 5 and name H8) (residue 5 and name H2') 3.6 0.6 1.4

assign (residue 5 and name H8) (residue 5 and name H3') 2.5 0.0 1.3

! Inter residue

assign (residue 5 and name H8) (residue 4 and name H1') 3.6 0.6 1.4

assign (residue 5 and name H8) (residue 4 and name H2') 2.5 0.7 0.4

assign (residue 5 and name H8) (residue 4 and name H3') 3.2 0.7 0.6

assign (residue 5 and name H1') (residue 4 and name H2') 3.5 0.6 0.6

! residue 6 URI (1911)

! Intra residue

assign (residue 6 and name H6) (residue 6 and name H1') 3.6 0.8 0.8

assign (residue 6 and name H6) (residue 6 and name H2') 3.5 0.7 0.7

assign (residue 6 and name H6) (residue 6 and name H3') 3.0 0.6 0.6

assign (residue 6 and name H6) (residue 6 and name H4') 4.1 0.8 0.8

assign (residue 6 and name H5) (residue 6 and name H2') 4.0 0.0 50.0

assign (residue 6 and name H5) (residue 6 and name H3') 4.0 1.0 1.0

! Inter residue

assign (residue 5 and name H3') (residue 6 and name H6) 3.2 0.7 0.6

assign (residue 6 and name H6) (residue 5 and name H1') 3.6 0.8 0.8

assign (residue 6 and name H6) (residue 5 and name H2') 2.4 0.5 0.5

assign (residue 6 and name H5) (residue 5 and name H2') 3.5 0.8 0.8

assign (residue 6 and name H5) (residue 5 and name H3') 3.5 0.8 0.8

! residue 7 ADE (1912)

! Intra residue

assign (residue 7 and name H8) (residue 7 and name H1') 3.8 0.8 0.8

assign (residue 7 and name H8) (residue 7 and name H2') 3.0 0.6 0.6

assign (residue 7 and name H8) (residue 7 and name H3') 3.3 0.6 0.6

```
assign (residue 7 and name H8) (residue 7 and name H4') 4.0 1.0 1.0

assign (residue 7 and name H8) (residue 7 and name H5'') 3.2 0.6 0.6

assign (residue 7 and name H8) (residue 7 and name H5') 3.6 0.8 0.8

! Inter residue

assign (residue 7 and name H8) (residue 6 and name H1') 4.5 1.5 1.5

assign (residue 7 and name H8) (residue 6 and name H2') 2.9 0.6 0.6

assign (residue 7 and name H8) (residue 6 and name H3') 3.1 0.6 0.6

assign (residue 7 and name H8) (residue 6 and name H4') 4.5 1.5 1.5

! residue 8 ADE (1913)

! Intra residue

assign (residue 8 and name H8) (residue 8 and name H1') 3.5 0.8 0.8

assign (residue 8 and name H8) (residue 8 and name H2') 3.2 0.6 0.6

assign (residue 8 and name H8) (residue 8 and name H3') 2.8 0.5 0.5

assign (residue 8 and name H8) (residue 8 and name H4') 5.0 1.0 1.0

assign (residue 8 and name H8) (residue 8 and name H5') 3.7 0.8 0.8

assign (residue 8 and name H8) (residue 8 and name H5'') 3.7 0.8 0.8

! Inter residue

assign (residue 8 and name H8) (residue 7 and name H1') 4.8 1.0 1.0

assign (residue 8 and name H8) (residue 7 and name H2') 2.5 0.5 0.5

assign (residue 8 and name H8) (residue 7 and name H3') 3.0 0.5 0.8

assign (residue 7 and name H5'') (residue 8 and name H8) 4.5 1.5 1.5

assign (residue 8 and name H2) (residue 7 and name H2) 3.9 1.0 1.0

assign (residue 8 and name H1') (residue 7 and name H2) 4.2 1.0 1.0

assign (residue 8 and name H2') (residue 7 and name H2) 5.0 1.0 1.0

! residue 9 CYT (1914)

! Intra residue

assign (residue 9 and name H6) (residue 9 and name H1') 3.4 0.8 0.8

assign (residue 9 and name H6) (residue 9 and name H2') 2.5 0.5 0.5

assign (residue 9 and name H6) (residue 9 and name H3') 2.7 0.5 0.5
```

```
assign (residue 9 and name H6) (residue 9 and name H4') 4.5 1.5 1.5

assign (residue 9 and name H6) (residue 9 and name H5') 3.6 0.8 0.8

assign (residue 9 and name H6) (residue 9 and name H5'') 3.4 0.8 0.8

assign (residue 9 and name H5) (residue 9 and name H2') 3.9 1.0 1.0

assign (residue 9 and name H5) (residue 9 and name H3') 4.5 1.5 1.5

! Inter residue

assign (residue 9 and name H6) (residue 8 and name H2') 2.8 0.5 0.5

assign (residue 9 and name H6) (residue 8 and name H3') 3.1 0.6 0.6

assign (residue 9 and name H6) (residue 8 and name H4') 4.5 1.5 1.5

assign (residue 9 and name H5) (residue 8 and name H8) 3.9 1.0 1.0

assign (residue 9 and name H5) (residue 8 and name H2') 4.3 1.5 1.5

assign (residue 9 and name H5) (residue 8 and name H3') 3.8 1.0 1.0

assign (residue 9 and name H1') (residue 8 and name H2) 3.9 1.0 1.0

assign (residue 9 and name H1') (residue 8 and name H2') 4.5 1.5 1.5

! residue 10 URI (1915)

! Intra residue

assign (residue 10 and name H6) (residue 10 and name H1') 3.3 0.6 0.6

assign (residue 10 and name H6) (residue 10 and name H2') 2.3 0.5 0.5

assign (residue 10 and name H6) (residue 10 and name H3') 3.2 0.6 0.6

assign (residue 10 and name H6) (residue 10 and name H4') 4.5 1.5 1.5

assign (residue 10 and name H6) (residue 10 and name H5') 3.4 0.8 0.8

assign (residue 10 and name H6) (residue 10 and name H5'') 3.4 0.8 0.8

assign (residue 10 and name H5) (residue 10 and name H2') 4.5 1.5 1.5

! Inter residue

assign (residue 10 and name H6) (residue 9 and name H1') 4.0 1.0 1.0

assign (residue 10 and name H6) (residue 9 and name H2') 3.7 0.8 0.8

assign (residue 10 and name H6) (residue 9 and name H3') 4.2 1.0 1.0

assign (residue 10 and name H6) (residue 9 and name H4') 3.8 1.0 1.0

assign (residue 10 and name H5) (residue 8 and name H2) 4.0 0.0 50.0
```

assign (residue 10 and name H5) (residue 9 and name H1') 4.5 1.5 1.5

assign (residue 10 and name H5) (residue 9 and name H2') 5.0 1.0 1.0

assign (residue 10 and name H1') (residue 8 and name H2) 4.0 0.0 50.0

! residue 11 ADE (1916)

! Intra residue

assign (residue 11 and name H8) (residue 11 and name H1') 3.4 0.8 0.8

assign (residue 11 and name H8) (residue 11 and name H2') 4.5 1.5 1.5

assign (residue 11 and name H8) (residue 11 and name H3') 4.5 1.5 1.5

assign (residue 11 and name H8) (residue 11 and name H4') 4.5 1.5 1.5

assign (residue 11 and name H8) (residue 11 and name H5'') 4.5 1.5 1.5

assign (residue 11 and name H8) (residue 11 and name H5'') 4.5 1.5 1.5

assign (residue 11 and name H2) (residue 11 and name H1') 4.5 1.5 1.5

! Inter residue

assign (residue 11 and name H8) (residue 9 and name H4') 4.0 1.0 1.0

assign (residue 11 and name H8) (residue 10 and name H1') 4.5 1.5 1.5

assign (residue 11 and name H8) (residue 10 and name H2') 4.1 1.0 1.0

assign (residue 11 and name H8) (residue 10 and name H3') 3.8 1.0 1.0

assign (residue 11 and name H8) (residue 10 and name H4') 4.1 1.0 1.0

assign (residue 11 and name H8) (residue 10 and name H5') 4.5 1.5 1.5

assign (residue 11 and name H8) (residue 10 and name H5'') 4.0 1.0 1.0

assign (residue 11 and name H1') (residue 8 and name H2) 5.0 1.0 1.0

assign (residue 11 and name H1') (residue 8 and name H2) 4.0 0.0 50.0

assign (residue 11 and name H1') (residue 9 and name H4') 4.0 1.0 1.0

assign (residue 11 and name H1') (residue 9 and name H5') 3.8 1.0 1.0

assign (residue 11 and name H1') (residue 9 and name H5'') 4.5 1.5 1.5

! residue 12 URI (1917)

! Intra residue

assign (residue 12 and name H6) (residue 12 and name H1') 3.2 0.6 0.6

assign (residue 12 and name H6) (residue 12 and name H2') 2.4 0.5 0.5

```
assign (residue 12 and name H6) (residue 12 and name H3') 3.3 0.6 0.6

assign (residue 12 and name H6) (residue 12 and name H4') 4.5 1.5 1.5

assign (residue 12 and name H6) (residue 12 and name H5'') 3.4 0.6 0.6

assign (residue 12 and name H6) (residue 12 and name H5') 4.1 1.0 1.0

assign (residue 12 and name H5) (residue 12 and name H2') 4.5 1.5 1.5

! Inter residue

assign (residue 12 and name H6) (residue 11 and name H1') 5.0 1.0 1.0

assign (residue 12 and name H6) (residue 11 and name H4') 4.5 1.5 1.5

assign (residue 12 and name H5) (residue 8 and name H2) 4.5 1.5 1.5

assign (residue 12 and name H5) (residue 11 and name H2') 4.0 0.0 50.0

assign (residue 12 and name H1') (residue 11 and name H2) 4.5 1.5 1.5

! residue 13 ADE (1918)

! Intra residue

assign (residue 13 and name H8) (residue 13 and name H1') 3.5 0.6 0.6

assign (residue 13 and name H8) (residue 13 and name H2') 4.5 1.5 1.5

assign (residue 13 and name H8) (residue 13 and name H3') 3.2 0.6 0.6

assign (residue 13 and name H8) (residue 13 and name H4') 4.2 1.0 1.0

assign (residue 13 and name H8) (residue 13 and name H5') 3.7 1.0 1.0

assign (residue 13 and name H8) (residue 13 and name H5'') 3.6 0.8 0.8

! Inter residue

assign (residue 13 and name H8) (residue 7 and name H2) 5.0 1.0 1.0

assign (residue 13 and name H8) (residue 12 and name H1') 4.5 1.5 1.5

assign (residue 13 and name H8) (residue 12 and name H2') 4.5 1.5 1.5

assign (residue 13 and name H8) (residue 12 and name H3') 3.9 1.0 1.0

assign (residue 13 and name H8) (residue 12 and name H4') 3.5 0.8 0.8

assign (residue 13 and name H2) (residue 8 and name H1') 4.5 1.5 1.5

assign (residue 13 and name H1') (residue 7 and name H2) 5.0 1.0 1.0

assign (residue 13 and name H1') (residue 8 and name H2) 4.5 1.5 1.5

assign (residue 13 and name H1') (residue 11 and name H2) 4.5 1.5 1.5
```

! residue 14 ADE (1919)

! Intra residue

assign (residue 14 and name H8) (residue 14 and name H1') 3.7 0.8 0.8

assign (residue 14 and name H8) (residue 14 and name H2') 4.0 1.0 1.0

assign (residue 14 and name H8) (residue 14 and name H3') 3.0 0.6 0.6

assign (residue 14 and name H8) (residue 14 and name H4') 4.0 1.0 1.0

assign (residue 14 and name H8) (residue 14 and name H5') 4.0 1.0 1.0

assign (residue 14 and name H8) (residue 14 and name H5'') 3.3 0.6 0.6

assign (residue 14 and name H2) (residue 14 and name H1') 4.5 1.5 1.5

assign (residue 14 and name H2) (residue 14 and name H2') 4.0 0.0 50.0

assign (residue 14 and name H2) (residue 14 and name H3') 4.0 0.0 50.0

assign (residue 14 and name H2) (residue 14 and name H4') 4.0 0.0 50.0

! Inter residue

assign (residue 14 and name H8) (residue 13 and name H1') 4.5 1.5 1.5

assign (residue 14 and name H8) (residue 13 and name H2') 4.0 0.0 50.0

assign (residue 14 and name H8) (residue 13 and name H3') 3.1 0.6 0.6

assign (residue 14 and name H8) (residue 13 and name H4') 4.5 1.5 1.5

assign (residue 14 and name H8) (residue 13 and name H5') 5.0 1.0 1.0

assign (residue 14 and name H8) (residue 13 and name H5'') 5.0 1.0 1.0

assign (residue 14 and name H2) (residue 6 and name H1') 4.5 1.5 1.5

assign (residue 14 and name H2) (residue 6 and name H2') 4.0 0.0 50.0

assign (residue 14 and name H2) (residue 6 and name H3') 4.0 0.0 50.0

assign (residue 14 and name H2) (residue 7 and name H2) 5.0 1.0 1.0

assign (residue 14 and name H2) (residue 7 and name H1') 3.1 0.6 0.6

assign (residue 14 and name H2) (residue 13 and name H2) 4.5 1.5 1.5

assign (residue 14 and name H2) (residue 8 and name H1') 4.0 0.0 50.0

assign (residue 14 and name H1') (residue 7 and name H2) 4.9 1.0 1.0

assign (residue 14 and name H1') (residue 13 and name H2) 3.3 0.6 0.6

assign (residue 14 and name H5'') (residue 13 and name H3') 4.5 1.5 1.5

! residue 15 CYT (1920)

! Intra residue

assign (residue 15 and name H6) (residue 15 and name H1') 4.0 1.0 1.0

assign (residue 15 and name H6) (residue 15 and name H2') 3.5 0.5 1.5

assign (residue 15 and name H6) (residue 15 and name H3') 2.5 0.0 1.3

assign (residue 15 and name H5) (residue 15 and name H1') 4.5 1.5 1.5

assign (residue 15 and name H5) (residue 15 and name H2') 4.5 1.5 1.5

assign (residue 15 and name H5) (residue 15 and name H3') 4.5 1.5 1.5

! Inter residue

assign (residue 15 and name H6) (residue 14 and name H1') 3.7 0.8 0.8

assign (residue 15 and name H6) (residue 14 and name H2') 2.6 0.5 0.5

assign (residue 15 and name H6) (residue 14 and name H3') 3.5 0.8 0.8

assign (residue 15 and name H6) (residue 14 and name H4') 4.5 1.5 1.5

assign (residue 15 and name H5) (residue 14 and name H8) 4.5 1.5 1.5

assign (residue 15 and name H5) (residue 14 and name H1') 4.5 1.5 1.5

assign (residue 15 and name H5) (residue 14 and name H2') 4.0 0.0 50.0

assign (residue 15 and name H5) (residue 14 and name H3') 4.5 1.5 1.5

assign (residue 15 and name H1') (residue 14 and name H2) 3.2 0.6 0.6

! residue 16 GUA (1921)

! Intra residue

assign (residue 16 and name H8) (residue 16 and name H1') 3.5 0.5 1.5

assign (residue 16 and name H8) (residue 16 and name H2') 4.2 1.2 0.8

assign (residue 16 and name H8) (residue 16 and name H3') 3.0 0.5 0.8

assign (residue 16 and name H8) (residue 16 and name H4') 3.5 0.5 1.5

! Inter residue

assign (residue 16 and name H8) (residue 15 and name H1') 3.9 0.9 1.1

assign (residue 16 and name H8) (residue 15 and name H2') 2.5 0.7 0.4

assign (residue 16 and name H8) (residue 15 and name H3') 3.1 0.6 0.7

assign (residue 16 and name H1') (residue 15 and name H2') 3.0 0.0 2.0

! residue 17 GUA (1922)

! Intra residue

assign (residue 17 and name H8) (residue 17 and name H1') 3.8 0.8 1.2

assign (residue 17 and name H8) (residue 17 and name H2') 4.2 1.2 0.8

assign (residue 17 and name H8) (residue 17 and name H3') 2.9 0.4 0.9

assign (residue 17 and name H8) (residue 17 and name H4') 4.5 1.5 1.5

! Inter residue

assign (residue 17 and name H8) (residue 16 and name H1') 3.8 0.8 1.2

assign (residue 17 and name H8) (residue 16 and name H2') 2.5 0.7 0.4

assign (residue 17 and name H8) (residue 16 and name H3') 3.1 0.6 0.7

assign (residue 17 and name H1') (residue 16 and name H2') 4.5 1.5 1.5

! residue 18 URI (1923)

! Intra residue

assign (residue 18 and name H6) (residue 18 and name H1') 3.4 0.9 0.4

assign (residue 18 and name H6) (residue 18 and name H2') 3.6 0.6 1.5

assign (residue 18 and name H6) (residue 18 and name H3') 3.5 1.0 0.3

assign (residue 18 and name H6) (residue 18 and name H4') 4.5 1.5 1.5

! Inter residue

assign (residue 18 and name H6) (residue 17 and name H1') 4.0 0.0 2.0

assign (residue 18 and name H6) (residue 17 and name H2') 2.5 0.7 0.4

assign (residue 18 and name H6) (residue 17 and name H3') 3.0 0.5 0.8

assign (residue 18 and name H5) (residue 17 and name H1') 4.5 1.5 1.5

assign (residue 18 and name H5) (residue 17 and name H2') 3.7 0.7 1.3

assign (residue 18 and name H5) (residue 17 and name H3') 3.5 0.5 1.5

assign (residue 18 and name H5) (residue 17 and name H8) 3.9 0.9 1.1

assign (residue 18 and name H1') (residue 17 and name H1') 4.5 1.5 1.5

assign (residue 18 and name H1') (residue 17 and name H2') 3.8 0.8 1.2

! residue 19 CYT (1924)

! Intra residue

```
assign (residue 19 and name H6) (residue 19 and name H1') 3.4 0.4 1.6

assign (residue 19 and name H6) (residue 19 and name H2') 2.8 0.3 1.0

assign (residue 19 and name H6) (residue 19 and name H3') 2.5 0.7 0.4

assign (residue 19 and name H6) (residue 19 and name H4') 3.5 0.5 1.5

assign (residue 19 and name H5) (residue 19 and name H2') 3.5 0.5 1.5

assign (residue 19 and name H5) (residue 19 and name H3') 3.1 0.6 0.7

! Inter residue

assign (residue 19 and name H5) (residue 18 and name H2') 3.1 0.6 0.7

assign (residue 19 and name H6) (residue 18 and name H1') 3.5 0.5 1.5

assign (residue 19 and name H6) (residue 18 and name H2') 2.5 0.7 0.4

! Unnoe

assign (residue 4 and name H2') (residue 5 and name H5') 4.00 0.0 50.0

assign (residue 5 and name H1') (residue 6 and name H5) 4.00 0.0 50.0

assign (residue 5 and name H2') (residue 6 and name H1') 4.00 0.0 50.0

assign (residue 5 and name H2') (residue 6 and name H2') 4.00 0.0 50.0

assign (residue 6 and name H1') (residue 13 and name H5') 4.00 0.0 50.0

assign (residue 6 and name H1') (residue 13 and name H8) 4.00 0.0 50.0

assign (residue 6 and name H6) (residue 13 and name H5') 4.00 0.0 50.0

assign (residue 6 and name H6) (residue 7 and name H5'') 4.00 0.0 50.0

assign (residue 6 and name H1') (residue 16 and name H1') 4.00 0.0 50.0

assign (residue 6 and name H5'') (residue 6 and name H6) 3.00 0.0 50.0

assign (residue 6 and name H2') (residue 7 and name H1') 4.00 0.0 50.0

assign (residue 6 and name H2') (residue 7 and name H5') 4.00 0.0 50.0

assign (residue 6 and name H3') (residue 7 and name H5'') 4.00 0.0 50.0

assign (residue 6 and name H2') (residue 7 and name H3') 3.00 0.0 50.0

assign (residue 6 and name H1') (residue 7 and name H2') 3.00 0.0 50.0

assign (residue 7 and name H1') (residue 8 and name H1') 4.00 0.0 50.0

assign (residue 7 and name H2') (residue 8 and name H1') 4.00 0.0 50.0

assign (residue 7 and name H4') (residue 8 and name H5') 3.00 0.0 50.0
```

```
assign (residue 7 and name H5') (residue 8 and name H8) 3.00 0.0 50.0

assign (residue 7 and name H2) (residue 11 and name H2) 3.00 0.0 50.0

assign (residue 7 and name H2) (residue 11 and name H1') 4.00 0.0 50.0

assign (residue 7 and name H2) (residue 11 and name H2') 4.00 0.0 50.0

assign (residue 7 and name H2) (residue 13 and name H2) 3.00 0.0 50.0

assign (residue 7 and name H2') (residue 13 and name H1') 4.00 0.0 50.0

assign (residue 7 and name H3') (residue 13 and name H1') 4.00 0.0 50.0

assign (residue 7 and name H1') (residue 14 and name H1') 4.00 0.0 50.0

assign (residue 7 and name H2') (residue 14 and name H2) 4.00 0.0 50.0

assign (residue 7 and name H1') (residue 15 and name H1') 4.00 0.0 50.0

assign (residue 7 and name H2') (residue 15 and name H1') 4.00 0.0 50.0

assign (residue 7 and name H4') (residue 15 and name H1') 4.00 0.0 50.0

assign (residue 7 and name H5') (residue 16 and name H1') 4.00 0.0 50.0

assign (residue 7 and name H5'') (residue 16 and name H1') 4.00 0.0 50.0

assign (residue 7 and name H1') (residue 12 and name H2') 3.00 0.0 50.0

assign (residue 7 and name H1') (residue 13 and name H8) 4.00 0.0 50.0

assign (residue 7 and name H8) (residue 14 and name H2) 4.00 0.0 50.0

assign (residue 8 and name H1') (residue 9 and name H2') 3.00 0.0 50.0

assign (residue 8 and name H2') (residue 9 and name H5') 3.00 0.0 50.0

assign (residue 8 and name H2') (residue 9 and name H5'') 3.00 0.0 50.0

assign (residue 8 and name H4') (residue 9 and name H5) 4.00 0.0 50.0

assign (residue 8 and name H5'') (residue 9 and name H5) 4.00 0.0 50.0

assign (residue 8 and name H2') (residue 9 and name H4') 3.00 0.0 50.0

assign (residue 8 and name H1') (residue 10 and name H2') 4.00 0.0 50.0

assign (residue 8 and name H1') (residue 11 and name H1') 4.00 0.0 50.0

assign (residue 8 and name H1') (residue 11 and name H2') 4.00 0.0 50.0

assign (residue 8 and name H2') (residue 11 and name H8) 4.00 0.0 50.0

assign (residue 8 and name H4') (residue 11 and name H1') 3.00 0.0 50.0

assign (residue 8 and name H2) (residue 12 and name H5) 4.00 0.0 50.0
```

```
assign (residue 8 and name H2) (residue 12 and name H5') 3.00 0.0 50.0

assign (residue 8 and name H8) (residue 12 and name H5) 4.00 0.0 50.0

assign (residue 8 and name H8) (residue 12 and name H4') 3.00 0.0 50.0

assign (residue 8 and name H2') (residue 10 and name H5) 4.00 0.0 50.0

assign (residue 8 and name H2') (residue 12 and name H5) 4.00 0.0 50.0

assign (residue 8 and name H8) (residue 13 and name H1') 3.00 0.0 50.0

assign (residue 8 and name H5') (residue 14 and name H2) 4.00 0.0 50.0

assign (residue 8 and name H2) (residue 14 and name H1') 4.00 0.0 50.0

assign (residue 8 and name H2) (residue 14 and name H4') 4.00 0.0 50.0

assign (residue 8 and name H2) (residue 15 and name H1') 4.00 0.0 50.0

assign (residue 8 and name H2) (residue 15 and name H4') 4.00 0.0 50.0

assign (residue 8 and name H1') (residue 15 and name H4') 4.00 0.0 50.0

assign (residue 8 and name H4') (residue 15 and name H4') 4.00 0.0 50.0

assign (residue 8 and name H5') (residue 16 and name H5') 4.00 0.0 50.0

assign (residue 8 and name H5') (residue 16 and name H5'') 4.00 0.0 50.0

assign (residue 9 and name H5') (residue 9 and name H2') 4.00 0.0 50.0

assign (residue 9 and name H5'') (residue 9 and name H2') 4.00 0.0 50.0

assign (residue 9 and name H1') (residue 10 and name H5'') 4.00 0.0 50.0

assign (residue 9 and name H1') (residue 11 and name H1') 4.00 0.0 50.0

assign (residue 9 and name H1') (residue 11 and name H8) 4.00 0.0 50.0

assign (residue 9 and name H2') (residue 10 and name H2') 4.00 0.0 50.0

assign (residue 9 and name H3') (residue 10 and name H2') 3.00 0.0 50.0

assign (residue 9 and name H3') (residue 10 and name H5'') 3.00 0.0 50.0

assign (residue 9 and name H4') (residue 10 and name H3') 4.00 0.0 50.0

assign (residue 9 and name H4') (residue 10 and name H5'') 4.00 0.0 50.0

assign (residue 9 and name H5') (residue 10 and name H5) 4.00 0.0 50.0

assign (residue 9 and name H5') (residue 10 and name H6) 4.00 0.0 50.0

assign (residue 9 and name H5'') (residue 10 and name H5) 4.00 0.0 50.0

assign (residue 9 and name H3') (residue 10 and name H5) 4.00 0.0 50.0
```

```
assign (residue 9 and name H5') (residue 10 and name H5') 4.00 0.0 50.0

assign (residue 9 and name H5') (residue 10 and name H1') 4.00 0.0 50.0

assign  (residue 9 and name H4') (residue 10 and name H5') 4.00 0.0 50.0

assign (residue 9 and name H3') (residue 11 and name H8) 4.00 0.0 50.0

assign (residue 9 and name H4') (residue 11 and name H2') 4.00 0.0 50.0

assign (residue 9 and name H4') (residue 11 and name H4') 4.00 0.0 50.0

assign (residue 9 and name H5') (residue 11 and name H2) 3.00 0.0 50.0

assign (residue 9 and name H5') (residue 11 and name H8) 4.00 0.0 50.0

assign (residue 9 and name H5'') (residue 11 and name H8) 3.00 0.0 50.0

assign (residue 9 and name H5'') (residue 11 and name H2') 4.00 0.0 50.0

assign (residue 9 and name H5'') (residue 10 and name H4') 3.00 0.0 50.0

assign (residue 9 and name H4') (residue 10 and name H4') 4.00 0.0 50.0

assign (residue 9 and name H4') (residue 10 and name H1') 4.00 0.0 50.0

assign (residue 10 and name H6) (residue 11 and name H1') 4.00 0.0 50.0

assign (residue 10 and name H5) (residue 11 and name H8) 4.00 0.0 50.0

assign (residue 10 and name H1') (residue 11 and name H5') 3.00 0.0 50.0

assign (residue 10 and name H5') (residue 10 and name H2') 4.00 0.0 50.0

assign (residue 10 and name H5'') (residue 10 and name H2') 3.00 0.0 50.0

assign (residue 10 and name H1') (residue 11 and name H1') 4.00 0.0 50.0

assign (residue 10 and name H1') (residue 11 and name H5'') 3.00 0.0 50.0

assign (residue 10 and name H2') (residue 11 and name H5') 4.00 0.0 50.0

assign (residue 10 and name H3') (residue 11 and name H2') 4.00 0.0 50.0

assign (residue 10 and name H3') (residue 11 and name H3') 4.00 0.0 50.0

assign (residue 10 and name H3') (residue 11 and name H5') 4.00 0.0 50.0

assign (residue 10 and name H5') (residue 11 and name H2) 4.00 0.0 50.0

assign (residue 10 and name H5') (residue 11 and name H1') 4.00 0.0 50.0

assign (residue 10 and name H5'') (residue 11 and name H1') 4.00 0.0 50.0

assign (residue 10 and name H5) (residue 13 and name H4') 4.00 0.0 50.0

assign (residue 11 and name H5') (residue 11 and name H8) 3.00 0.0 50.0
```

```
assign (residue 10 and name H4') (residue 11 and name H5') 4.00 0.0 50.0

assign (residue 11 and name H2) (residue 12 and name H6) 3.00 0.0 50.0

assign (residue 11 and name H2) (residue 12 and name H5) 4.00 0.0 50.0

assign (residue 11 and name H2) (residue 12 and name H3') 4.00 0.0 50.0

assign (residue 11 and name H2) (residue 12 and name H5') 4.00 0.0 50.0

assign (residue 11 and name H2) (residue 13 and name H2') 4.00 0.0 50.0

assign (residue 11 and name H2) (residue 13 and name H3') 4.00 0.0 50.0

assign (residue 11 and name H8) (residue 12 and name H2') 4.00 0.0 50.0

assign (residue 11 and name H1') (residue 12 and name H5) 4.00 0.0 50.0

assign (residue 11 and name H2') (residue 12 and name H6) 3.00 0.0 50.0

assign (residue 11 and name H2') (residue 12 and name H2') 4.00 0.0 50.0

assign (residue 11 and name H2') (residue 12 and name H3') 4.00 0.0 50.0

assign (residue 11 and name H2') (residue 12 and name H5') 4.00 0.0 50.0

assign (residue 11 and name H2') (residue 12 and name H5'') 4.00 0.0 50.0

assign (residue 11 and name H3') (residue 12 and name H2') 4.00 0.0 50.0

assign (residue 11 and name H3') (residue 12 and name H5') 4.00 0.0 50.0

assign (residue 11 and name H3') (residue 12 and name H5'') 4.00 0.0 50.0

assign (residue 11 and name H4') (residue 12 and name H5') 4.00 0.0 50.0

assign (residue 11 and name H5') (residue 12 and name H6) 3.00 0.0 50.0

assign (residue 11 and name H5'') (residue 12 and name H5) 4.00 0.0 50.0

assign (residue 11 and name H2') (residue 13 and name H2) 4.00 0.0 50.0

assign (residue 11 and name H2') (residue 13 and name H1') 3.00 0.0 50.0

assign (residue 12 and name H5) (residue 13 and name H2) 4.00 0.0 50.0

assign (residue 12 and name H5) (residue 13 and name H1') 4.00 0.0 50.0

assign (residue 12 and name H5) (residue 13 and name H8) 3.00 0.0 50.0

assign (residue 12 and name H1') (residue 13 and name H4') 4.00 0.0 50.0

assign (residue 12 and name H3') (residue 13 and name H4') 4.00 0.0 50.0

assign (residue 12 and name H4') (residue 13 and name H1') 4.00 0.0 50.0

assign (residue 12 and name H4') (residue 13 and name H4') 4.00 0.0 50.0
```

```
assign (residue 12 and name H4') (residue 13 and name H5'') 4.00 0.0 50.0

assign (residue 12 and name H5'') (residue 13 and name H8) 4.00 0.0 50.0

assign (residue 12 and name H5'') (residue 13 and name H1') 3.00 0.0 50.0

assign (residue 12 and name H5') (residue 13 and name H2) 4.00 0.0 50.0

assign (residue 12 and name H5') (residue 13 and name H1') 4.00 0.0 50.0

assign (residue 12 and name H5'') (residue 12 and name H2') 4.00 0.0 50.0

assign (residue 12 and name H5'') (residue 12 and name H5'') 4.00 0.0 50.0

assign (residue 12 and name H3') (residue 13 and name H3') 4.00 0.0 50.0

assign (residue 12 and name H5') (residue 14 and name H1') 4.00 0.0 50.0

assign (residue 12 and name H2') (residue 14 and name H4') 4.00 0.0 50.0

assign (residue 12 and name H2') (residue 14 and name H1') 3.00 0.0 50.0

assign (residue 12 and name H3') (residue 14 and name H5'') 4.00 0.0 50.0

assign (residue 12 and name H3') (residue 14 and name H4') 4.00 0.0 50.0

assign (residue 12 and name H3') (residue 14 and name H1') 4.00 0.0 50.0

assign (residue 12 and name H4') (residue 14 and name H1') 4.00 0.0 50.0

assign (residue 12 and name H4') (residue 13 and name H2') 4.00 0.0 50.0

assign (residue 12 and name H2') (residue 13 and name H2') 4.00 0.0 50.0

assign (residue 12 and name H2') (residue 13 and name H3') 4.00 0.0 50.0

assign (residue 12 and name H1') (residue 13 and name H2) 3.00 0.0 50.0

assign (residue 12 and name H1') (residue 13 and name H2') 4.00 0.0 50.0

assign (residue 12 and name H5') (residue 14 and name H2) 4.00 0.0 50.0

assign (residue 12 and name H5'') (residue 14 and name H1') 4.00 0.0 50.0

assign (residue 12 and name H3') (residue 14 and name H5') 4.00 0.0 50.0

assign (residue 12 and name H4') (residue 13 and name H2) 4.00 0.0 50.0

assign (residue 12 and name H3') (residue 14 and name H2) 4.00 0.0 50.0

assign (residue 12 and name H2') (residue 14 and name H2) 4.00 0.0 50.0

assign (residue 12 and name H5) (residue 15 and name H1') 4.00 0.0 50.0

assign (residue 12 and name H6) (residue 14 and name H2) 4.00 0.0 50.0

assign (residue 13 and name H2) (residue 14 and name H8) 4.00 0.0 50.0
```

```
assign (residue 13 and name H2) (residue 14 and name H2') 3.00 0.0 50.0

assign (residue 13 and name H2) (residue 14 and name H3') 3.00 0.0 50.0

assign (residue 13 and name H1') (residue 14 and name H2) 3.00 0.0 50.0

assign (residue 13 and name H1') (residue 14 and name H1') 4.00 0.0 50.0

assign (residue 13 and name H2') (residue 14 and name H8) 3.00 0.0 50.0

assign (residue 13 and name H2') (residue 14 and name H1') 4.00 0.0 50.0

assign (residue 13 and name H2') (residue 14 and name H5') 3.00 0.0 50.0

assign (residue 13 and name H2') (residue 14 and name H5'') 4.00 0.0 50.0

assign (residue 13 and name H2) (residue 15 and name H1') 4.00 0.0 50.0

assign (residue 13 and name H2) (residue 15 and name H4') 4.00 0.0 50.0

assign (residue 13 and name H2) (residue 15 and name H5') 4.00 0.0 50.0

assign (residue 14 and name H1') (residue 15 and name H1') 4.00 0.0 50.0

assign (residue 14 and name H1') (residue 15 and name H5') 4.00 0.0 50.0

assign (residue 14 and name H2') (residue 15 and name H1') 3.00 0.0 50.0
```

Dihedral angle restraints for the H69UUU

```
!Delta

assign (resid 1 and name c5') (resid 1 and name c4')

        (resid 1 and name c3') (resid 1 and name o3') 1 84.00 20.0  2

assign (resid 2 and name c5') (resid 2 and name c4')

        (resid 2 and name c3') (resid 2 and name o3') 1 84.00 20.0  2

assign (resid 3 and name c5') (resid 3 and name c4')

        (resid 3 and name c3') (resid 3 and name o3') 1 84.00 20.0  2

assign (resid 4 and name c5') (resid 4 and name c4')

        (resid 4 and name c3') (resid 4 and name o3') 1 84.00 20.0  2

assign (resid 5 and name c5') (resid 5 and name c4')

        (resid 5 and name c3') (resid 5 and name o3') 1 84.00 20.0  2

assign (resid 6 and name c5') (resid 6 and name c4')

        (resid 6 and name c3') (resid 6 and name o3') 1 84.00 20.0  2
```

```
assign (resid 7 and name c5') (resid 7 and name c4')

      (resid 7 and name c3') (resid 7 and name o3') 1 84.00 20.0  2

assign (resid 8 and name c5') (resid 8 and name c4')

      (resid 8 and name c3') (resid 8 and name o3') 1 84.00 20.0  2

assign (resid 9 and name c5') (resid 9 and name c4')

      (resid 9 and name c3') (resid 9 and name o3') 1 157.00 40.0  2

assign (resid 10 and name c5') (resid 10 and name c4')

      (resid 10 and name c3') (resid 10 and name o3') 1 157.00 40.0  2

assign (resid 11 and name c5') (resid 11 and name c4')

      (resid 11 and name c3') (resid 11 and name o3') 1 84.00 20.0  2

assign (resid 12 and name c5') (resid 12 and name c4')

      (resid 12 and name c3') (resid 12 and name o3') 1 157.00 40.0  2

assign (resid 13 and name c5') (resid 13 and name c4')

      (resid 13 and name c3') (resid 13 and name o3') 1 84.00 20.0  2

assign (resid 14 and name c5') (resid 14 and name c4')

      (resid 14 and name c3') (resid 14 and name o3') 1 84.00 20.0  2

assign (resid 15 and name c5') (resid 15 and name c4')

      (resid 15 and name c3') (resid 15 and name o3') 1 84.00 20.0  2

assign (resid 16 and name c5') (resid 16 and name c4')

      (resid 16 and name c3') (resid 16 and name o3') 1 84.00 20.0  2

assign (resid 17 and name c5') (resid 17 and name c4')

      (resid 17 and name c3') (resid 17 and name o3') 1 84.00 20.0  2

assign (resid 18 and name c5') (resid 18 and name c4')

      (resid 18 and name c3') (resid 18 and name o3') 1 84.00 20.0  2

assign (resid 19 and name c5') (resid 19 and name c4')

      (resid 19 and name c3') (resid 19 and name o3') 1 84.00 20.0  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 20.0 2
```

```
assign (resid 1 and name o3') (resid 2 and name p)

      (resid 2 and name o5') (resid 2 and name c5') 1 -62 20.0 2

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

      (resid 3 and name p) (resid 3 and name o5') 1 -71 20.0 2

assign (resid 3 and name c3') (resid 3 and name o3')

      (resid 4 and name p) (resid 4 and name o5') 1 -71 20.0 2

assign (resid 3 and name o3') (resid 4 and name p)

      (resid 4 and name o5') (resid 4 and name c5') 1 -62 20.0 2

assign (resid 4 and name c3') (resid 4 and name o3')

      (resid 5 and name p) (resid 5 and name o5') 1 -71 20.0 2

assign (resid 4 and name o3') (resid 5 and name p)

      (resid 5 and name o5') (resid 5 and name c5') 1 -62 20.0 2

assign (resid 5 and name c3') (resid 5 and name o3')

      (resid 6 and name p) (resid 6 and name o5') 1 -71 20.0 2

assign (resid 5 and name o3') (resid 6 and name p)

      (resid 6 and name o5') (resid 6 and name c5') 1 -62 20.0 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 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 -71 20.0 2
assign (resid 14 and name o3') (resid 15 and name p)

        (resid 15 and name o5') (resid 15 and name c5') 1 -62 20.0 2
assign (resid 15 and name c3') (resid 15 and name o3')

        (resid 16 and name p) (resid 16 and name o5') 1 -71 20.0 2
assign (resid 15 and name o3') (resid 16 and name p)

        (resid 16 and name o5') (resid 16 and name c5') 1 -62 20.0 2
assign (resid 16 and name c3') (resid 16 and name o3')

        (resid 17 and name p) (resid 17 and name o5') 1 -71 20.0 2
```

```
assign (resid 16 and name o3') (resid 17 and name p)

       (resid 17 and name o5') (resid 17 and name c5') 1 -62 20.0 2

assign (resid 17 and name c3') (resid 17 and name o3')

       (resid 18 and name p) (resid 18 and name o5') 1 -71 20.0 2

assign (resid 17 and name o3') (resid 18 and name p)

       (resid 18 and name o5') (resid 18 and name c5') 1 -62 20.0 2

assign (resid 18 and name c3') (resid 18 and name o3')

       (resid 19 and name p) (resid 19 and name o5') 1 -71 20.0 2

assign (resid 18 and name o3') (resid 19 and name p)

       (resid 19 and name o5') (resid 19 and name c5') 1 -62 20.0 2

!Beta

assign (resid 2 and name p) (resid 2 and name o5')

       (resid 2 and name c5') (resid 2 and name c4') 1 172 20.0 2

assign (resid 3 and name p) (resid 3 and name o5')

       (resid 3 and name c5') (resid 3 and name c4') 1 172 20.0 2

assign (resid 4 and name p) (resid 4 and name o5')

       (resid 4 and name c5') (resid 4 and name c4') 1 172 20.0 2

assign (resid 5 and name p) (resid 5 and name o5')

       (resid 5 and name c5') (resid 5 and name c4') 1 172 20.0 2

assign (resid 6 and name p) (resid 6 and name o5')

       (resid 6 and name c5') (resid 6 and name c4') 1 -180 90.0 2

assign (resid 7 and name p) (resid 7 and name o5')

       (resid 7 and name c5') (resid 7 and name c4') 1 -180 90.0 2

assign (resid 8 and name p) (resid 8 and name o5')

       (resid 8 and name c5') (resid 8 and name c4') 1 -180 90.0 2

assign (resid 9 and name p) (resid 9 and name o5')

       (resid 9 and name c5') (resid 9 and name c4') 1 -180 90.0 2

assign (resid 10 and name p) (resid 10 and name o5')

       (resid 10 and name c5') (resid 10 and name c4') 1 -180 90.0 2
```

```
assign (resid 11 and name p) (resid 11 and name o5')

       (resid 11 and name c5') (resid 11 and name c4') 1 -180 90.0 2

assign (resid 12 and name p) (resid 12 and name o5')

       (resid 12 and name c5') (resid 12 and name c4') 1 -180 90.0 2

assign (resid 13 and name p) (resid 13 and name o5')

       (resid 13 and name c5') (resid 13 and name c4') 1 -180 90.0 2

assign (resid 14 and name p) (resid 14 and name o5')

       (resid 14 and name c5') (resid 14 and name c4') 1 -180 90.0 2

assign (resid 15 and name p) (resid 15 and name o5')

       (resid 15 and name c5') (resid 15 and name c4') 1 172 20.0 2

assign (resid 16 and name p) (resid 16 and name o5')

       (resid 16 and name c5') (resid 16 and name c4') 1 172 20.0 2

assign (resid 17 and name p) (resid 17 and name o5')

       (resid 17 and name c5') (resid 17 and name c4') 1 172 20.0 2

assign (resid 18 and name p) (resid 18 and name o5')

       (resid 18 and name c5') (resid 18 and name c4') 1 172 20.0 2

assign (resid 19 and name p) (resid 19 and name o5')

       (resid 19 and name c5') (resid 19 and name c4') 1 172 20.0 2

!Gamma

assign (resid 1 and name o5') (resid 1 and name c5')

       (resid 1 and name c4') (resid 1 and name c3') 1 60 20.0 2

assign (resid 2 and name o5') (resid 2 and name c5')

       (resid 2 and name c4') (resid 2 and name c3') 1 60 20.0 2

assign (resid 3 and name o5') (resid 3 and name c5')

       (resid 3 and name c4') (resid 3 and name c3') 1 60 20.0 2

assign (resid 4 and name o5') (resid 4 and name c5')

       (resid 4 and name c4') (resid 4 and name c3') 1 60 20.0 2

assign (resid 5 and name o5') (resid 5 and name c5')

       (resid 5 and name c4') (resid 5 and name c3') 1 60 20.0 2
```

```
assign (resid 6 and name o5') (resid 6 and name c5')

       (resid 6 and name c4') (resid 6 and name c3') 1 60 60.0 2

assign (resid 7 and name o5') (resid 7 and name c5')

       (resid 7 and name c4') (resid 7 and name c3') 1 60 60.0 2

assign (resid 8 and name o5') (resid 8 and name c5')

       (resid 8 and name c4') (resid 8 and name c3') 1 60 60.0 2

assign (resid 9 and name o5') (resid 9 and name c5')

       (resid 9 and name c4') (resid 9 and name c3') 1 120 120.0 2

assign (resid 10 and name o5') (resid 10 and name c5')

       (resid 10 and name c4') (resid 10 and name c3') 1 60 60.0 2

assign (resid 11 and name o5') (resid 11 and name c5')

       (resid 11 and name c4') (resid 11 and name c3') 1 120 120.0 2

assign (resid 12 and name o5') (resid 12 and name c5')

       (resid 12 and name c4') (resid 12 and name c3') 1 60 60.0 2

assign (resid 13 and name o5') (resid 13 and name c5')

       (resid 13 and name c4') (resid 13 and name c3') 1 120 120.0 2

assign (resid 14 and name o5') (resid 14 and name c5')

       (resid 14 and name c4') (resid 14 and name c3') 1 120 120  2

assign (resid 15 and name o5') (resid 15 and name c5')

       (resid 15 and name c4') (resid 15 and name c3') 1 58 20.0 2

assign (resid 16 and name o5') (resid 16 and name c5')

       (resid 16 and name c4') (resid 16 and name c3') 1 58 20.0 2

assign (resid 17 and name o5') (resid 17 and name c5')

       (resid 17 and name c4') (resid 17 and name c3') 1 58 20.0 2

assign (resid 18 and name o5') (resid 18 and name c5')

       (resid 18 and name c4') (resid 18 and name c3') 1 58 20.0 2

assign (resid 19 and name o5') (resid 19 and name c5')

       (resid 19 and name c4') (resid 19 and name c3') 1 58 20.0 2

!Epsilon
```

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

        (resid 1 and name o3') (resid 2 and name p) 1 -160 20.0 2

assign (resid 2 and name c4') (resid 2 and name c3')

        (resid 2 and name o3') (resid 3 and name p) 1 -160 20.0 2

assign (resid 3 and name c4') (resid 3 and name c3')

        (resid 3 and name o3') (resid 4 and name p) 1 -160 20.0 2

assign (resid 4 and name c4') (resid 4 and name c3')

        (resid 4 and name o3') (resid 5 and name p) 1 -160 20.0 2

assign (resid 5 and name c4') (resid 5 and name c3')

        (resid 5 and name o3') (resid 6 and name p) 1 -120 120.0 2

assign (resid 6 and name c4') (resid 6 and name c3')

        (resid 6 and name o3') (resid 7 and name p) 1 -120 120.0 2

assign (resid 7 and name c4') (resid 7 and name c3')

        (resid 7 and name o3') (resid 8 and name p) 1 -120 120.0 2

assign (resid 8 and name c4') (resid 8 and name c3')

        (resid 8 and name o3') (resid 9 and name p) 1 -120 120.0 2

assign (resid 9 and name c4') (resid 9 and name c3')

        (resid 9 and name o3') (resid 10 and name p) 1 -120 120.0 2

assign (resid 10 and name c4') (resid 10 and name c3')

        (resid 10 and name o3') (resid 11 and name p) 1 -120 120.0 2

assign (resid 11 and name c4') (resid 11 and name c3')

        (resid 11 and name o3') (resid 12 and name p) 1 -120 120.0 2

assign (resid 12 and name c4') (resid 12 and name c3')

        (resid 12 and name o3') (resid 13 and name p) 1 -120 120.0 2

assign (resid 13 and name c4') (resid 13 and name c3')

        (resid 13 and name o3') (resid 14 and name p) 1 -120 120.0 2

assign (resid 14 and name c4') (resid 14 and name c3')

        (resid 14 and name o3') (resid 15 and name p) 1 -120 120.0 2

assign (resid 15 and name c4') (resid 15 and name c3')
```

```
        (resid 15 and name o3') (resid 16 and name p) 1 -160 20.0 2
assign (resid 16 and name c4') (resid 16 and name c3')
        (resid 16 and name o3') (resid 17 and name p) 1 -160 20.0 2
assign (resid 17 and name c4') (resid 17 and name c3')
        (resid 17 and name o3') (resid 18 and name p) 1 -160 20.0 2
assign (resid 18 and name c4') (resid 18 and name c3')
        (resid 18 and name o3') (resid 19 and name p) 1 -160 20.0 2
!Chi
assign (resid 1 and name O4') (resid 1 and name C1')
        (resid 1 and name N9 ) (resid 1 and name C4 )  1.0 -160 20.0 2
assign (resid 2 and name O4') (resid 2 and name C1')
        (resid 2 and name N9 ) (resid 2 and name C4 )  1.0 -160 20.0 2
assign (resid 3 and name O4') (resid 3 and name C1')
        (resid 3 and name N1 ) (resid 3 and name C2 )   1.0 -160 20.0 2
assign (resid 4 and name O4') (resid 4 and name C1')
        (resid 4 and name N1 ) (resid 4 and name C2 )   1.0 -160 20.0 2
assign (resid 5 and name O4') (resid 5 and name C1')
        (resid 5 and name N9 ) (resid 5 and name C4 )  1.0 -160 20.0 2
assign (resid 6 and name O4') (resid 6 and name C1')
        (resid 6 and name C5 ) (resid 6 and name C4 )  1 -110 110 2
assign (resid 7 and name O4') (resid 7 and name C1')
        (resid 7 and name N9 ) (resid 7 and name C4 )  1 -110 110 2
assign (resid 8 and name O4') (resid 8 and name C1')
        (resid 8 and name N9 ) (resid 8 and name C4 )  1 -110 110 2
assign (resid 9 and name O4') (resid 9 and name C1')
        (resid 9 and name N1 ) (resid 9 and name C2 )   1 -110 110 2
assign (resid 10 and name O4') (resid 10 and name C1')
        (resid 10 and name C5 ) (resid 10 and name C4 )  1 -110 110 2
assign (resid 11 and name O4') (resid 11 and name C1')
```

```
        (resid 11 and name N9 ) (resid 11 and name C4 )  1 -110 110 2
assign (resid 12 and name O4') (resid 12 and name C1')
        (resid 12 and name C5 ) (resid 12 and name C4 )  1 -110 110 2
assign (resid 13 and name O4') (resid 13 and name C1')
        (resid 13 and name N9 ) (resid 13 and name C4 )  1 -110 110 2
assign (resid 14 and name O4') (resid 14 and name C1')
        (resid 14 and name N9 ) (resid 14 and name C4 )  1 -110 110 2
assign (resid 15 and name O4') (resid 15 and name C1')
        (resid 15 and name N1 ) (resid 15 and name C2 )   1.0 -160 20.0 2
assign (resid 16 and name O4') (resid 16 and name C1')
        (resid 16 and name N9 ) (resid 16 and name C4 )  1.0 -160 20.0 2
assign (resid 17 and name O4') (resid 17 and name C1')
        (resid 17 and name N9 ) (resid 17 and name C4 )  1.0 -160 20.0 2
assign (resid 18 and name O4') (resid 18 and name C1')
        (resid 18 and name N1 ) (resid 18 and name C2 ) 1 -160 20.0 2
assign (resid 19 and name O4') (resid 19 and name C1')
        (resid 19 and name N1 ) (resid 19 and name C2 ) 1 -160 20.0 2
```

APPENDIX 2

Structural Restraints for the H69ΨΨΨ

Distance restraints for the H69ΨΨΨ

! Base flipping control in Global fold

assign (residue 1 and name O6) (residue 2 and name O6) 4.6 0.5 0.5

assign (residue 1 and name N2) (residue 2 and name N2) 4.8 0.5 0.5

assign (residue 2 and name O6) (residue 3 and name N4) 4.7 0.5 0.5

assign (residue 2 and name N2) (residue 3 and name O2) 3.5 0.5 0.5

assign (residue 3 and name N4) (residue 4 and name N4) 3.7 0.5 0.5

assign (residue 3 and name O2) (residue 4 and name O2) 4.8 0.5 0.5

assign (residue 4 and name N4) (residue 5 and name O6) 5.4 0.5 0.5

assign (residue 4 and name O2) (residue 5 and name N2) 6.5 0.5 0.5

assign (residue 19 and name O2) (residue 18 and name O2) 5.0 0.5 0.5

assign (residue 19 and name N4) (residue 18 and name O4) 4.9 0.5 0.5

assign (residue 18 and name O2) (residue 17 and name N2) 3.2 0.5 0.5

assign (residue 18 and name O4) (residue 17 and name O6) 3.9 0.5 0.5

assign (residue 17 and name O6) (residue 16 and name O6) 4.9 0.5 0.5

assign (residue 17 and name N2) (residue 16 and name N2) 4.6 0.5 0.5

assign (residue 16 and name O6) (residue 15 and name N4) 5.3 0.5 0.5

assign (residue 16 and name N2) (residue 15 and name O2) 6.4 0.5 0.5

!Stem stacking distances

assign (residue 1 and name H1) (residue 2 and name H1) 4.5 0.5 0.5

assign (residue 1 and name H1) (residue 18 and name H3) 5.3 0.5 0.5

assign (residue 17 and name H1) (residue 18 and name H3) 3.6 0.5 0.5

assign (residue 2 and name H1) (residue 17 and name H1) 4.5 0.5 0.5

assign (residue 17 and name H1) (residue 16 and name H1) 4.5 0.5 0.5

assign (residue 16 and name H1) (residue 5 and name H1) 4.4 0.5 0.5

! NOEs from H2O NOESY

```
assign (residue 1 and name H1) (residue 19 and name H42) 2.4 0.5 0.5

assign (residue 2 and name H1) (residue 19 and name H42) 4.8 0.5 0.5

assign (residue 18 and name H3) (residue 19 and name H42) 4.6 0.5 0.5

assign (residue 2 and name H1) (residue 17 and name H22) 4.5 0.5 0.5

assign (residue 2 and name H1) (residue 19 and name H42) 4.2 0.5 0.5

assign (residue 18 and name H3) (residue 17 and name H22) 4.4 0.5 0.5

assign (residue 2 and name H1) (residue 18 and name H3) 2.1 0.5 0.5

assign (residue 3 and name H42) (residue 17 and name H1) 2.5 0.5 0.5

assign (residue 4 and name H42) (residue 17 and name H1) 3.3 0.5 0.5

assign (residue 3 and name H42) (residue 16 and name H1) 5.8 0.5 0.5

assign (residue 4 and name H42) (residue 16 and name H1) 2.8 0.5 0.5

assign (residue 5 and name H1) (residue 15 and name H42) 2.6 0.5 0.5

assign (residue 4 and name H42) (residue 15 and name H42) 6.8 0.5 0.5

! Carbon distances for the Base pairs

assign (residue 1 and name C1') (residue 19 and name C1') 10.7 0.2 0.2

assign (residue 1 and name C8) (residue 19 and name C6) 9.7 0.2 0.2

assign (residue 2 and name C1') (residue 18 and name C1') 10.7 0.2 0.2

assign (residue 2 and name C8) (residue 18 and name C6) 9.7 0.2 0.2

assign (residue 3 and name C1') (residue 17 and name C1') 10.7 0.2 0.2

assign (residue 3 and name C6) (residue 17 and name C8) 9.7 0.2 0.2

assign (residue 4 and name C1') (residue 16 and name C1') 10.7 0.2 0.2

assign (residue 4 and name C6) (residue 16 and name C8) 9.7 0.2 0.2

assign (residue 5 and name C1') (residue 15 and name C1') 10.7 0.2 0.2

assign (residue 5 and name C8) (residue 15 and name C6) 9.7 0.2 0.2

!Base pair dists

! for G1/ C19  base pair

assign (residue 1 and name N1) (residue 19 and name N3) 2.7 0.2 0.2

assign (residue 1 and name O6) (residue 19 and name N4) 2.7 0.2 0.2

assign (residue 1 and name N2) (residue 19 and name O2) 2.8 0.2 0.2
```

```
! for G2/U18 base pair

assign (residue 2 and name N1) (residue 18 and name O2) 3.1 0.2 0.2

assign (residue 2 and name O6) (residue 18 and name N3) 2.9 0.2 0.2

! for C3/ G17  base pair

assign (residue 3 and name N3) (residue 17 and name N1) 2.7 0.2 0.2

assign (residue 3 and name N4) (residue 17 and name O6) 2.7 0.2 0.2

assign (residue 3 and name O2) (residue 17 and name N2) 2.8 0.2 0.2

! for C4/ G16  base pair

assign (residue 4 and name N3) (residue 16 and name N1) 2.7 0.2 0.2

assign (residue 4 and name N4) (residue 16 and name O6) 2.7 0.2 0.2

assign (residue 4 and name O2) (residue 16 and name N2) 2.8 0.2 0.2

! for G5/ C15  base pair

assign (residue 5 and name N1) (residue 15 and name N3) 2.7 0.2 0.2

assign (residue 5 and name O6) (residue 15 and name N4) 2.7 0.2 0.2

assign (residue 5 and name N2) (residue 15 and name O2) 2.8 0.2 0.2

! Base to Base constraints.

assign (residue 1 and name H8) (residue 2 and name H8) 4.5 1.5 1.5

assign (residue 2 and name H8) (residue 3 and name H6) 4.5 1.5 1.5

assign (residue 3 and name H6) (residue 4 and name H6) 5.0 1.0 1.0

assign (residue 4 and name H6) (residue 5 and name H8) 4.5 1.5 1.5

assign (residue 5 and name H8) (residue 6 and name H6) 5.0 1.0 1.0

assign (residue 6 and name H6) (residue 7 and name H8) 5.0 1.0 1.0

assign (residue 7 and name H8) (residue 8 and name H8) 3.5 0.8 0.8

assign (residue 8 and name H8) (residue 9 and name H6) 5.0 1.0 1.0

assign (residue 9 and name H6) (residue 10 and name H6) 4.0 0.0 50.0

assign (residue 10 and name H6) (residue 11 and name H8) 4.0 1.0 1.0

assign (residue 11 and name H8) (residue 12 and name H6) 5.0 1.0 1.0

assign (residue 12 and name H6) (residue 13 and name H8) 5.0 1.0 1.0

assign (residue 13 and name H8) (residue 14 and name H8) 5.0 1.0 1.0
```

assign (residue 14 and name H8) (residue 15 and name H6) 5.0 1.0 1.0

assign (residue 15 and name H6) (residue 16 and name H8) 4.5 1.5 1.5

assign (residue 16 and name H8) (residue 17 and name H8) 4.5 1.5 1.5

assign (residue 17 and name H8) (residue 18 and name H6) 4.5 1.5 1.5

! Base to Base constraints, H1'/H5.

assign (residue 1 and name H1') (residue 2 and name H1') 4.5 1.5 1.5

assign (residue 2 and name H1') (residue 3 and name H1') 4.5 1.5 1.5

assign (residue 2 and name H1') (residue 3 and name H5) 4.5 1.5 1.5

assign (residue 4 and name H1') (residue 5 and name H1') 4.5 1.5 1.5

assign (residue 7 and name H1') (residue 8 and name H1') 4.0 0.0 50.0

assign (residue 8 and name H1') (residue 9 and name H1') 4.0 0.0 50.0

assign (residue 8 and name H1') (residue 9 and name H5) 4.0 0.0 50.0

assign (residue 8 and name H1') (residue 11 and name H1') 4.0 0.0 50.0

assign (residue 9 and name H1') (residue 10 and name H1') 4.0 0.0 50.0

assign (residue 10 and name H1') (residue 11 and name H1') 4.0 0.0 50.0

assign (residue 11 and name H1') (residue 12 and name H1') 4.6 1.0 1.0

assign (residue 14 and name H1') (residue 15 and name H1') 5.0 1.0 1.0

assign (residue 14 and name H1') (residue 15 and name H5) 5.0 1.0 1.0

assign (residue 15 and name H1') (residue 16 and name H1') 4.5 1.5 1.5

assign (residue 15 and name H1') (residue 15 and name H5) 4.5 1.5 1.5

assign (residue 17 and name H1') (residue 18 and name H5) 4.5 1.5 1.5

assign (residue 18 and name H1') (residue 19 and name H1') 4.5 1.5 1.5

! residue 1 GUA (1906)

! Intra residue

assign (residue 1 and name H8) (residue 1 and name H1') 3.5 0.8 0.8

assign (residue 1 and name H8) (residue 1 and name H2') 3.5 0.8 0.8

assign (residue 1 and name H8) (residue 1 and name H3') 2.7 0.5 0.5

assign (residue 1 and name H8) (residue 1 and name H4') 4.1 1.0 1.0

! residue 2 GUA (1907)

```
! Intra residue

assign (residue 2 and name H8) (residue 2 and name H1') 3.6 0.8 0.8

assign (residue 2 and name H8) (residue 2 and name H2') 3.6 0.8 0.8

assign (residue 2 and name H8) (residue 2 and name H3') 2.4 0.5 0.5

assign (residue 2 and name H8) (residue 2 and name H4') 4.0 1.0 1.0

! Inter residue

assign (residue 1 and name H1') (residue 2 and name H8) 3.8 0.8 0.8

assign (residue 1 and name H2') (residue 2 and name H8) 2.4 0.5 0.5

assign (residue 1 and name H3') (residue 2 and name H8) 4.0 1.0 1.0

assign (residue 1 and name H4') (residue 2 and name H8) 4.0 0.0 50.0

! residue 3 CYT (1908)

! Intra residue

assign (residue 3 and name H6) (residue 3 and name H1') 3.3 0.5 0.5

assign (residue 3 and name H6) (residue 3 and name H2') 3.9 1.0 1.0

assign (residue 3 and name H6) (residue 3 and name H3') 2.4 0.5 0.5

assign (residue 3 and name H6) (residue 3 and name H4') 3.9 1.0 1.0

assign (residue 3 and name H5) (residue 3 and name H2') 4.0 0.0 50.0

assign (residue 3 and name H5) (residue 3 and name H3') 3.6 0.8 0.8

! Inter residue

assign (residue 3 and name H6) (residue 2 and name H1') 3.9 1.0 1.0

assign (residue 3 and name H6) (residue 2 and name H2') 2.4 0.5 0.5

assign (residue 3 and name H6) (residue 2 and name H3') 3.0 0.6 0.6

assign (residue 3 and name H5) (residue 2 and name H1') 5.0 1.0 1.0

assign (residue 3 and name H5) (residue 2 and name H2') 3.6 0.8 0.8

assign (residue 3 and name H5) (residue 2 and name H3') 3.6 0.8 0.8

assign (residue 3 and name H5) (residue 2 and name H8) 4.0 1.0 1.0

! residue 4 CYT (1909)

! Intra residue

assign (residue 4 and name H6) (residue 4 and name H1') 3.5 0.8 0.8
```

```
assign (residue 4 and name H6) (residue 4 and name H2') 3.6 0.8 0.8

assign (residue 4 and name H6) (residue 4 and name H3') 2.5 0.5 0.5

assign (residue 4 and name H6) (residue 4 and name H4') 3.4 0.6 0.6

! Inter residue

assign (residue 4 and name H6) (residue 3 and name H1') 3.5 0.8 0.8

assign (residue 4 and name H6) (residue 3 and name H2') 2.4 0.5 0.5

assign (residue 4 and name H5) (residue 3 and name H2') 3.7 0.8 0.8

! residue 5 GUA (1910)

! Intra residue

assign (residue 5 and name H8) (residue 5 and name H1') 3.8 0.8 0.8

assign (residue 5 and name H8) (residue 5 and name H2') 3.8 0.8 0.8

assign (residue 5 and name H8) (residue 5 and name H3') 2.5 0.5 0.5

! Inter residue

assign (residue 5 and name H8) (residue 4 and name H1') 3.8 1.0 1.0

assign (residue 5 and name H8) (residue 4 and name H2') 2.5 0.5 0.5

assign (residue 5 and name H8) (residue 4 and name H3') 3.6 0.8 0.8

! residue 6 PSU (1911)

! Intra residue

assign (residue 6 and name H6) (residue 6 and name H2') 3.6 0.8 0.8

assign (residue 6 and name H6) (residue 6 and name H3') 2.8 0.5 0.5

assign (residue 6 and name H6) (residue 6 and name H4') 4.5 1.0 1.0

! Inter residue

assign (residue 6 and name H6) (residue 5 and name H1') 3.8 1.0 1.0

assign (residue 6 and name H6) (residue 5 and name H2') 2.6 0.5 0.5

assign (residue 6 and name H6) (residue 5 and name H3') 3.1 0.6 0.6

! residue 7 ADE (1912)

! Intra residue

assign (residue 7 and name H8) (residue 7 and name H1') 4.1 1.0 1.0

assign (residue 7 and name H8) (residue 7 and name H2') 4.0 1.0 1.0
```

assign (residue 7 and name H8) (residue 7 and name H3') 2.8 0.5 0.5

! Inter residue

assign (residue 7 and name H8) (residue 6 and name H2') 2.8 0.5 0.5

assign (residue 7 and name H8) (residue 6 and name H3') 2.7 0.5 0.5

assign (residue 7 and name H8) (residue 6 and name H4') 5.0 1.0 1.0

assign (residue 7 and name H2) (residue 8 and name H1') 3.6 0.8 0.8

assign (residue 7 and name H2) (residue 8 and name H8) 5.0 1.0 1.0

assign (residue 7 and name H2) (residue 13 and name H2) 5.0 1.0 1.0

! residue 8 ADE (1913)

! Intra residue

assign (residue 8 and name H8) (residue 8 and name H1') 3.6 0.8 0.8

assign (residue 8 and name H8) (residue 8 and name H2') 3.2 0.6 0.6

assign (residue 8 and name H8) (residue 8 and name H3') 2.8 0.5 0.5

assign (residue 8 and name H8) (residue 8 and name H4') 4.5 1.5 1.5

! Inter residue

assign (residue 8 and name H8) (residue 7 and name H1') 4.5 1.5 1.5

assign (residue 8 and name H8) (residue 7 and name H2') 2.9 0.6 0.6

assign (residue 8 and name H8) (residue 7 and name H3') 2.9 0.6 0.6

assign (residue 8 and name H2) (residue 13 and name H1') 4.5 1.5 1.5

assign (residue 8 and name H2) (residue 9 and name H1') 3.3 0.6 0.6

assign (residue 8 and name H2) (residue 12 and name H1') 5.0 1.0 1.0

assign (residue 8 and name H1') (residue 7 and name H3') 5.0 1.0 1.0

! residue 9 CYT (1914)

! Intra residue

assign (residue 9 and name H6) (residue 9 and name H1') 3.4 0.6 0.6

assign (residue 9 and name H6) (residue 9 and name H2') 2.9 0.5 0.5

assign (residue 9 and name H6) (residue 9 and name H3') 2.5 0.5 0.5

assign (residue 9 and name H6) (residue 9 and name H4') 3.7 0.8 0.8

assign (residue 9 and name H5) (residue 9 and name H3') 5.0 1.0 1.0

assign (residue 9 and name H5) (residue 9 and name H2') 5.0 1.0 1.0

! Inter residue

assign (residue 9 and name H6) (residue 8 and name H1') 5.0 1.0 1.0

assign (residue 9 and name H6) (residue 8 and name H2') 2.9 0.6 0.6

assign (residue 9 and name H6) (residue 8 and name H3') 3.0 0.6 0.6

assign (residue 9 and name H5) (residue 8 and name H2') 4.5 1.5 1.5

assign (residue 9 and name H5) (residue 8 and name H3') 5.0 1.0 1.0

assign (residue 9 and name H5) (residue 8 and name H8) 5.0 1.0 1.0

assign (residue 9 and name H1') (residue 8 and name H8) 5.0 1.0 1.0

assign (residue 9 and name H1') (residue 8 and name H2') 4.5 1.5 1.5

assign (residue 9 and name H1') (residue 11 and name H2') 4.5 1.5 1.5

assign (residue 9 and name H1') (residue 8 and name H3') 5.0 1.0 1.0

! residue 10 PSU (1915)

! Intra residue

assign (residue 10 and name H6) (residue 10 and name H1') 3.3 0.6 0.6

assign (residue 10 and name H6) (residue 10 and name H2') 3.5 0.6 0.6

assign (residue 10 and name H6) (residue 10 and name H3') 3.0 0.6 0.6

assign (residue 10 and name H6) (residue 10 and name H4') 3.9 1.0 1.0

! Inter residue

assign (residue 10 and name H6) (residue 9 and name H1') 5.0 1.0 1.0

assign (residue 10 and name H6) (residue 9 and name H2') 3.6 0.8 0.8

assign (residue 10 and name H6) (residue 9 and name H3') 3.1 0.6 0.6

! residue 11 ADE (1916)

! Intra residue

assign (residue 11 and name H8) (residue 11 and name H1') 3.5 0.6 0.6

assign (residue 11 and name H8) (residue 11 and name H2') 2.6 0.5 0.5

assign (residue 11 and name H8) (residue 11 and name H3') 4.5 1.0 1.0

assign (residue 11 and name H8) (residue 11 and name H4') 4.3 1.0 1.0

assign (residue 11 and name H2) (residue 11 and name H1') 5.0 1.0 1.0

! Inter residue

assign (residue 11 and name H8) (residue 10 and name H1') 5.0 1.0 1.0

assign (residue 11 and name H8) (residue 10 and name H2') 3.0 0.5 0.5

assign (residue 11 and name H8) (residue 10 and name H3') 2.8 0.5 0.5

assign (residue 11 and name H8) (residue 10 and name H4') 5.0 1.0 1.0

assign (residue 11 and name H8) (residue 9 and name H1') 5.0 1.0 1.0

assign (residue 11 and name H2) (residue 10 and name H6) 5.0 1.0 1.0

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 H2') 5.0 1.0 1.0

assign (residue 11 and name H2) (residue 12 and name H4') 5.0 1.0 1.0

assign (residue 11 and name H1') (residue 10 and name H2') 4.2 1.0 1.0

assign (residue 11 and name H1') (residue 12 and name H2') 5.0 1.0 1.0

! residue 12 PSU (1917)

! Intra residue

assign (residue 12 and name H6) (residue 12 and name H1') 3.3 0.6 0.6

assign (residue 12 and name H6) (residue 12 and name H2') 3.3 0.6 0.6

assign (residue 12 and name H6) (residue 12 and name H3') 3.2 0.6 0.6

assign (residue 12 and name H6) (residue 12 and name H4') 4.1 1.0 1.0

! Inter residue

assign (residue 12 and name H6) (residue 11 and name H1') 4.2 1.0 1.0

assign (residue 12 and name H6) (residue 11 and name H2') 2.6 0.5 0.5

assign (residue 12 and name H6) (residue 11 and name H3') 4.1 1.0 1.0

! residue 13 ADE (1918)

! Intra residue

assign (residue 13 and name H8) (residue 13 and name H1') 3.8 0.8 0.8

assign (residue 13 and name H8) (residue 13 and name H2') 3.3 0.6 0.6

assign (residue 13 and name H8) (residue 13 and name H3') 2.8 0.5 0.5

assign (residue 13 and name H2) (residue 13 and name H1') 5.0 1.0 1.0

! Inter residue

assign (residue 13 and name H8) (residue 11 and name H1') 5.0 1.0 1.0

assign (residue 13 and name H8) (residue 12 and name H1') 5.0 1.0 1.0

assign (residue 13 and name H8) (residue 12 and name H2') 3.2 0.6 0.6

assign (residue 13 and name H8) (residue 12 and name H3') 3.0 0.6 0.6

assign (residue 13 and name H8) (residue 12 and name H4') 3.5 0.8 0.8

assign (residue 13 and name H2) (residue 8 and name H1') 3.8 1.0 1.0

assign (residue 13 and name H2) (residue 12 and name H1') 5.0 1.0 1.0

assign (residue 13 and name H2) (residue 14 and name H1') 3.9 1.0 1.0

assign (residue 13 and name H1') (residue 12 and name H2') 5.0 1.0 1.0

! residue 14 ADE (1919)

! Intra residue

assign (residue 14 and name H8) (residue 14 and name H1') 4.5 1.0 1.0

assign (residue 14 and name H8) (residue 14 and name H2') 4.5 1.0 1.0

assign (residue 14 and name H8) (residue 14 and name H3') 2.7 0.6 0.6

assign (residue 14 and name H8) (residue 14 and name H4') 3.9 1.0 1.0

! Inter residue

assign (residue 14 and name H8) (residue 13 and name H1') 5.0 1.0 1.0

assign (residue 14 and name H8) (residue 13 and name H2') 2.7 0.5 0.5

assign (residue 14 and name H8) (residue 13 and name H3') 3.3 0.8 0.8

assign (residue 14 and name H2) (residue 7 and name H1') 3.4 0.6 0.6

assign (residue 14 and name H2) (residue 15 and name H1') 3.6 0.8 0.8

assign (residue 14 and name H2) (residue 8 and name H1') 5.0 1.0 1.0

! residue 15 CYT (1920)

! Intra residue

assign (residue 15 and name H6) (residue 15 and name H1') 3.7 0.8 0.8

assign (residue 15 and name H6) (residue 15 and name H2') 3.0 0.6 0.6

assign (residue 15 and name H6) (residue 15 and name H3') 2.7 0.5 0.5

! Inter residue

assign (residue 15 and name H6) (residue 14 and name H2') 2.8 0.5 0.5

```
assign (residue 15 and name H6) (residue 14 and name H3') 3.1 0.6 0.6

assign (residue 15 and name H6) (residue 14 and name H1') 5.0 1.0 1.0

assign (residue 15 and name H5) (residue 14 and name H2') 4.5 1.5 1.5

assign (residue 15 and name H5) (residue 14 and name H3') 4.5 1.5 1.5

assign (residue 14 and name H2') (residue 15 and name H1') 4.5 1.5 1.5

! residue 16 GUA (1921)

! Intra residue

assign (residue 16 and name H8) (residue 16 and name H1') 3.4 0.6 0.6

assign (residue 16 and name H8) (residue 16 and name H2') 4.0 1.0 1.0

assign (residue 16 and name H8) (residue 16 and name H3') 3.2 0.6 0.6

assign (residue 16 and name H8) (residue 16 and name H4') 4.5 1.5 1.5

! Inter residue

assign (residue 16 and name H8) (residue 15 and name H1') 3.9 1.0 1.0

assign (residue 16 and name H8) (residue 15 and name H2') 2.4 0.5 0.5

assign (residue 16 and name H8) (residue 15 and name H3') 2.7 0.5 0.5

! residue 17 GUA (1922)

! Intra residue

assign (residue 17 and name H8) (residue 17 and name H1') 3.6 0.8 0.8

assign (residue 17 and name H8) (residue 17 and name H2') 4.5 1.5 1.5

assign (residue 17 and name H8) (residue 17 and name H3') 2.9 0.5 0.5

assign (residue 17 and name H8) (residue 17 and name H4') 3.8 1.0 1.0

! Inter residue

assign (residue 17 and name H8) (residue 16 and name H1') 3.6 0.8 0.8

assign (residue 17 and name H8) (residue 16 and name H2') 2.6 0.5 0.5

assign (residue 17 and name H8) (residue 16 and name H3') 3.1 0.6 0.6

assign (residue 17 and name H8) (residue 18 and name H5) 3.9 1.0 1.0

! residue 18 URI (1923)

! Intra residue

assign (residue 18 and name H6) (residue 18 and name H1') 3.1 0.6 0.6
```

```
assign (residue 18 and name H6) (residue 18 and name H2') 3.3 0.6 0.6

assign (residue 18 and name H6) (residue 18 and name H3') 2.5 0.5 0.5

assign (residue 18 and name H6) (residue 18 and name H4') 3.5 0.8 0.8

! Inter residue

assign (residue 18 and name H6) (residue 17 and name H1') 3.8 1.0 1.0

assign (residue 18 and name H6) (residue 17 and name H2') 2.4 0.5 0.5

assign (residue 18 and name H6) (residue 17 and name H3') 2.8 0.5 0.5

assign (residue 18 and name H6) (residue 19 and name H5) 5.0 1.0 1.0

! residue 19 CYT (1924)

! Intra residue

assign (residue 19 and name H6) (residue 19 and name H1') 3.2 0.6 0.6

assign (residue 19 and name H6) (residue 19 and name H2') 4.0 1.0 1.0

assign (residue 19 and name H6) (residue 19 and name H3') 3.0 0.6 0.6

! Inter residue

assign (residue 19 and name H6) (residue 18 and name H5) 5.0 1.0 1.0

assign (residue 19 and name H6) (residue 18 and name H1') 3.4 0.6 0.6

assign (residue 19 and name H6) (residue 18 and name H2') 2.2 0.5 0.5

assign (residue 19 and name H6) (residue 18 and name H3') 3.2 0.6 0.6

!unnoe

assign (resid 5 and name h2') (resid 7 and name h8) 4.00 0.0 50.0

assign (resid 5 and name h4') (resid 6 and name h6) 4.00 0.0 50.0

assign (residue 7 and name H2) (residue 7 and name H1') 4.0 0.0 50.0

assign (resid 7 and name h2) (resid 12 and name h3') 4.00 0.0 50.0

assign (resid 7 and name h2) (resid 13 and name h1') 4.00 0.0 50.0

assign (resid 7 and name h2) (resid 13 and name h8) 4.00 0.0 50.0

assign (resid 7 and name h2') (resid 8 and name h1') 4.00 0.0 50.0

assign (resid 7 and name h1') (resid 13 and name h2) 4.00 0.0 50.0

assign (resid 8 and name h2) (resid 11 and name h4') 4.00 0.0 50.0

assign (resid 8 and name h2) (resid 11 and name h2) 4.00 0.0 50.0
```

```
assign (resid 8 and name h2) (resid 11 and name h1') 4.00 0.0 50.0

assign (resid 8 and name h2) (resid 12 and name h6) 4.00 0.0 50.0

assign (resid 8 and name h2) (resid 13 and name h8) 4.00 0.0 50.0

assign (resid 8 and name h2) (resid 13 and name h2') 4.00 0.0 50.0

assign (resid 8 and name h5') (resid 14 and name h2) 4.00 0.0 50.0

assign (resid 8 and name h1') (resid 11 and name h2) 4.00 0.0 50.0

assign (resid 8 and name h4') (resid 13 and name h2) 4.00 0.0 50.0

assign (resid 9 and name h1') (resid 10 and name h3') 4.00 0.0 50.0

assign (resid 9 and name h1') (resid 11 and name h3') 4.00 0.0 50.0

assign (resid 9 and name h1') (resid 11 and name h2) 4.00 0.0 50.0

assign (resid 9 and name h2') (resid 11 and name h2) 4.00 0.0 50.0

assign (resid 9 and name h4') (resid 11 and name h8) 3.00 0.0 50.0

assign (residue 9 and name H1') (residue 10 and name H4') 4.0 0.0 50.0

assign (resid 11 and name h3') (resid 13 and name h8) 4.00 0.0 50.0

assign (resid 11 and name h2') (resid 13 and name h1') 3.00 0.0 50.0

assign (resid 11 and name h1') (resid 13 and name h1') 4.00 0.0 50.0

assign (resid 11 and name h2) (resid 13 and name h1') 4.00 0.0 50.0

assign (resid 12 and name h6) (resid 13 and name h1') 4.00 0.0 50.0

assign (resid 12 and name h3') (resid 13 and name h1') 4.00 0.0 50.0

assign (resid 12 and name h4') (resid 13 and name h1') 4.00 0.0 50.0

assign (resid 13 and name h8) (resid 15 and name h4') 4.00 0.0 50.0

assign (resid 13 and name h2) (resid 14 and name h2) 4.00 0.0 50.0

assign (resid 13 and name h2) (resid 15 and name h1') 4.00 0.0 50.0

assign (resid 13 and name h1') (resid 14 and name h1') 4.00 0.0 50.0

assign (resid 13 and name h1') (resid 15 and name h4') 4.00 0.0 50.0

assign (resid 13 and name h2') (resid 14 and name h1') 4.00 0.0 50.0

assign (residue 14 and name H2) (residue 13 and name H1') 4.0 0.0 50.0

assign (resid 14 and name h8) (resid 15 and name h2') 4.00 0.0 50.0

assign (resid 14 and name h8) (resid 15 and name h3') 4.00 0.0 50.0
```

```
assign (resid 14 and name h8) (resid 16 and name h8) 4.00 0.0 50.0
```

Dihedral angle restraints for the H69ΨΨΨ

```
!Delta
assign (resid 1 and name c5') (resid 1 and name c4')
       (resid 1 and name c3') (resid 1 and name o3') 1 84.00 20.0  2
assign (resid 2 and name c5') (resid 2 and name c4')
       (resid 2 and name c3') (resid 2 and name o3') 1 84.00 20.0  2
assign (resid 3 and name c5') (resid 3 and name c4')
       (resid 3 and name c3') (resid 3 and name o3') 1 84.00 20.0  2
assign (resid 4 and name c5') (resid 4 and name c4')
       (resid 4 and name c3') (resid 4 and name o3') 1 84.00 20.0  2
assign (resid 5 and name c5') (resid 5 and name c4')
       (resid 5 and name c3') (resid 5 and name o3') 1 84.00 20.0  2
assign (resid 6 and name c5') (resid 6 and name c4')
       (resid 6 and name c3') (resid 6 and name o3') 1 84.00 20.0  2
assign (resid 7 and name c5') (resid 7 and name c4')
       (resid 7 and name c3') (resid 7 and name o3') 1 84.00 20.0  2
assign (resid 8 and name c5') (resid 8 and name c4')
       (resid 8 and name c3') (resid 8 and name o3') 1 84.00 20.0  2
assign (resid 9 and name c5') (resid 9 and name c4')
       (resid 9 and name c3') (resid 9 and name o3') 1 84.00 20.0  2
assign (resid 13 and name c5') (resid 13 and name c4')
       (resid 13 and name c3') (resid 13 and name o3') 1 84.00 20.0  2
assign (resid 14 and name c5') (resid 14 and name c4')
       (resid 14 and name c3') (resid 14 and name o3') 1 84.00 20.0  2
assign (resid 15 and name c5') (resid 15 and name c4')
       (resid 15 and name c3') (resid 15 and name o3') 1 84.00 20.0  2
assign (resid 16 and name c5') (resid 16 and name c4')
```

```
        (resid 16 and name c3') (resid 16 and name o3') 1 84.00 20.0  2
assign (resid 17 and name c5') (resid 17 and name c4')
        (resid 17 and name c3') (resid 17 and name o3') 1 84.00 20.0  2
assign (resid 18 and name c5') (resid 18 and name c4')
        (resid 18 and name c3') (resid 18 and name o3') 1 84.00 20.0  2
assign (resid 19 and name c5') (resid 19 and name c4')
        (resid 19 and name c3') (resid 19 and name o3') 1 84.00 20.0  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 20.0 2
assign (resid 1 and name o3') (resid 2 and name p)
        (resid 2 and name o5') (resid 2 and name c5') 1 -62 20.0 2
assign (resid 2 and name c3') (resid 2 and name o3')
        (resid 3 and name p) (resid 3 and name o5') 1 -71 20.0 2
assign (resid 2 and name o3') (resid 3 and name p)
        (resid 3 and name o5') (resid 3 and name c5') 1 -62 20.0 2
assign (resid 3 and name c3') (resid 3 and name o3')
        (resid 4 and name p) (resid 4 and name o5') 1 -71 20.0 2
assign (resid 3 and name o3') (resid 4 and name p)
        (resid 4 and name o5') (resid 4 and name c5') 1 -62 20.0 2
assign (resid 4 and name c3') (resid 4 and name o3')
        (resid 5 and name p) (resid 5 and name o5') 1 -71 20.0 2
assign (resid 4 and name o3') (resid 5 and name p)
        (resid 5 and name o5') (resid 5 and name c5') 1 -62 20.0 2
assign (resid 5 and name c3') (resid 5 and name o3')
        (resid 6 and name p) (resid 6 and name o5') 1 0 120.0 2
assign (resid 5 and name o3') (resid 6 and name p)
        (resid 6 and name o5') (resid 6 and name c5') 1 0 120.0 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 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 -71 20.0 2

assign (resid 15 and name o3') (resid 16 and name p)

        (resid 16 and name o5') (resid 16 and name c5') 1 -62 20.0 2

assign (resid 16 and name c3') (resid 16 and name o3')

        (resid 17 and name p) (resid 17 and name o5') 1 -71 20.0 2

assign (resid 16 and name o3') (resid 17 and name p)

        (resid 17 and name o5') (resid 17 and name c5') 1 -62 20.0 2

assign (resid 17 and name c3') (resid 17 and name o3')

        (resid 18 and name p) (resid 18 and name o5') 1 -71 20.0 2

assign (resid 17 and name o3') (resid 18 and name p)

        (resid 18 and name o5') (resid 18 and name c5') 1 -62 20.0 2

assign (resid 18 and name c3') (resid 18 and name o3')

        (resid 19 and name p) (resid 19 and name o5') 1 -71 20.0 2

assign (resid 18 and name o3') (resid 19 and name p)

        (resid 19 and name o5') (resid 19 and name c5') 1 -62 20.0 2

!Beta

assign (resid 2 and name p) (resid 2 and name o5')

        (resid 2 and name c5') (resid 2 and name c4') 1 172 20.0 2

assign (resid 3 and name p) (resid 3 and name o5')

        (resid 3 and name c5') (resid 3 and name c4') 1 172 20.0 2

assign (resid 4 and name p) (resid 4 and name o5')

        (resid 4 and name c5') (resid 4 and name c4') 1 172 20.0 2
```

```
assign (resid 5 and name p) (resid 5 and name o5')

        (resid 5 and name c5') (resid 5 and name c4') 1 172 20.0 2
assign (resid 6 and name p) (resid 6 and name o5')

        (resid 6 and name c5') (resid 6 and name c4') 1 -180 90.0 2
assign (resid 7 and name p) (resid 7 and name o5')

        (resid 7 and name c5') (resid 7 and name c4') 1 -180 90.0 2
assign (resid 8 and name p) (resid 8 and name o5')

        (resid 8 and name c5') (resid 8 and name c4') 1 -180 90.0 2
assign (resid 9 and name p) (resid 9 and name o5')

        (resid 9 and name c5') (resid 9 and name c4') 1 -180 90.0 2
assign (resid 10 and name p) (resid 10 and name o5')

        (resid 10 and name c5') (resid 10 and name c4') 1 -180 90.0 2
assign (resid 11 and name p) (resid 11 and name o5')

        (resid 11 and name c5') (resid 11 and name c4') 1 -180 90.0 2
assign (resid 12 and name p) (resid 12 and name o5')

        (resid 12 and name c5') (resid 12 and name c4') 1 -180 90.0 2
assign (resid 13 and name p) (resid 13 and name o5')

        (resid 13 and name c5') (resid 13 and name c4') 1 -180 90.0 2
assign (resid 14 and name p) (resid 14 and name o5')

        (resid 14 and name c5') (resid 14 and name c4') 1 -180 90.0 2
assign (resid 15 and name p) (resid 15 and name o5')

        (resid 15 and name c5') (resid 15 and name c4') 1 172 20.0 2
assign (resid 16 and name p) (resid 16 and name o5')

        (resid 16 and name c5') (resid 16 and name c4') 1 172 20.0 2
assign (resid 17 and name p) (resid 17 and name o5')

        (resid 17 and name c5') (resid 17 and name c4') 1 172 20.0 2
assign (resid 18 and name p) (resid 18 and name o5')

        (resid 18 and name c5') (resid 18 and name c4') 1 172 20.0 2
assign (resid 19 and name p) (resid 19 and name o5')
```

```
        (resid 19 and name c5') (resid 19 and name c4') 1 172 20.0 2
!Gamma
assign (resid 1 and name o5') (resid 1 and name c5')
        (resid 1 and name c4') (resid 1 and name c3') 1 60 20.0 2
assign (resid 2 and name o5') (resid 2 and name c5')
        (resid 2 and name c4') (resid 2 and name c3') 1 60 20.0 2
assign (resid 3 and name o5') (resid 3 and name c5')
        (resid 3 and name c4') (resid 3 and name c3') 1 60 20.0 2
assign (resid 4 and name o5') (resid 4 and name c5')
        (resid 4 and name c4') (resid 4 and name c3') 1 60 20.0 2
assign (resid 5 and name o5') (resid 5 and name c5')
        (resid 5 and name c4') (resid 5 and name c3') 1 60 20.0 2
assign (resid 6 and name o5') (resid 6 and name c5')
        (resid 6 and name c4') (resid 6 and name c3') 1 120 120 2
assign (resid 7 and name o5') (resid 7 and name c5')
        (resid 7 and name c4') (resid 7 and name c3') 1 120 120 2
assign (resid 8 and name o5') (resid 8 and name c5')
        (resid 8 and name c4') (resid 8 and name c3') 1 120 120 2
assign (resid 9 and name o5') (resid 9 and name c5')
        (resid 9 and name c4') (resid 9 and name c3') 1 120 120 2
assign (resid 10 and name o5') (resid 10 and name c5')
        (resid 10 and name c4') (resid 10 and name c3') 1 120 120 2
assign (resid 11 and name o5') (resid 11 and name c5')
        (resid 11 and name c4') (resid 11 and name c3') 1 120 120 2
assign (resid 12 and name o5') (resid 12 and name c5')
        (resid 12 and name c4') (resid 12 and name c3') 1 120 120 2
assign (resid 13 and name o5') (resid 13 and name c5')
        (resid 13 and name c4') (resid 13 and name c3') 1 120 120 2
assign (resid 14 and name o5') (resid 14 and name c5')
```

```
        (resid 14 and name c4') (resid 14 and name c3') 1 120 120 2
assign (resid 15 and name o5') (resid 15 and name c5')
        (resid 15 and name c4') (resid 15 and name c3') 1 60 20.0 2
assign (resid 16 and name o5') (resid 16 and name c5')
        (resid 16 and name c4') (resid 16 and name c3') 1 60 20.0 2
assign (resid 17 and name o5') (resid 17 and name c5')
        (resid 17 and name c4') (resid 17 and name c3') 1 60 20.0 2
assign (resid 18 and name o5') (resid 18 and name c5')
        (resid 18 and name c4') (resid 18 and name c3') 1 60 20.0 2
assign (resid 19 and name o5') (resid 19 and name c5')
        (resid 19 and name c4') (resid 19 and name c3') 1 60 20.0 2
!Epsilon
assign (resid 1 and name c4') (resid 1 and name c3')
        (resid 1 and name o3') (resid 2 and name p) 1 -160 20.0 2
assign (resid 2 and name c4') (resid 2 and name c3')
        (resid 2 and name o3') (resid 3 and name p) 1 -160 20.0 2
assign (resid 3 and name c4') (resid 3 and name c3')
        (resid 3 and name o3') (resid 4 and name p) 1 -160 20.0 2
assign (resid 4 and name c4') (resid 4 and name c3')
        (resid 4 and name o3') (resid 5 and name p) 1 -160 20.0 2
assign (resid 5 and name c4') (resid 5 and name c3')
        (resid 5 and name o3') (resid 6 and name p) 1 -120 120.0 2
assign (resid 6 and name c4') (resid 6 and name c3')
        (resid 6 and name o3') (resid 7 and name p) 1 -120 120.0 2
assign (resid 7 and name c4') (resid 7 and name c3')
        (resid 7 and name o3') (resid 8 and name p) 1 -120 120.0 2
assign (resid 8 and name c4') (resid 8 and name c3')
        (resid 8 and name o3') (resid 9 and name p) 1 -120 120.0 2
assign (resid 9 and name c4') (resid 9 and name c3')
```

```
        (resid 9 and name o3') (resid 10 and name p) 1 -120 120.0 2
assign (resid 10 and name c4') (resid 10 and name c3')
        (resid 10 and name o3') (resid 11 and name p) 1 -120 120.0 2
assign (resid 11 and name c4') (resid 11 and name c3')
        (resid 11 and name o3') (resid 12 and name p) 1 -120 120.0 2
assign (resid 12 and name c4') (resid 12 and name c3')
        (resid 12 and name o3') (resid 13 and name p) 1 -120 120.0 2
assign (resid 13 and name c4') (resid 13 and name c3')
        (resid 13 and name o3') (resid 14 and name p) 1 -120 120.0 2
assign (resid 14 and name c4') (resid 14 and name c3')
        (resid 14 and name o3') (resid 15 and name p) 1 -120 120.0 2
assign (resid 15 and name c4') (resid 15 and name c3')
        (resid 15 and name o3') (resid 16 and name p) 1 -160 20.0 2
assign (resid 16 and name c4') (resid 16 and name c3')
        (resid 16 and name o3') (resid 17 and name p) 1 -160 20.0 2
assign (resid 17 and name c4') (resid 17 and name c3')
        (resid 17 and name o3') (resid 18 and name p) 1 -160 20.0 2
assign (resid 18 and name c4') (resid 18 and name c3')
        (resid 18 and name o3') (resid 19 and name p) 1 -160 20.0 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 20 2
assign (resid 2 and name O4') (resid 2 and name C1')
        (resid 2 and name N9 ) (resid 2 and name C4 )  1 -160 20 2
assign (resid 3 and name O4') (resid 3 and name C1')
        (resid 3 and name N1 ) (resid 3 and name C2 )  1 -160 20 2
assign (resid 4 and name O4') (resid 4 and name C1')
        (resid 4 and name N1 ) (resid 4 and name C2 )  1 -160 20 2
assign (resid 5 and name O4') (resid 5 and name C1')
```

```
        (resid 5 and name N9 ) (resid 5 and name C4 )  1 -160 20 2
assign (resid 6 and name O4') (resid 6 and name C1')
        (resid 6 and name N1 ) (resid 6 and name C2 )  1 -110 110 2
assign (resid 7 and name O4') (resid 7 and name C1')
        (resid 7 and name N9 ) (resid 7 and name C4 )  1 -110 110 2
assign (resid 8 and name O4') (resid 8 and name C1')
        (resid 8 and name N9 ) (resid 8 and name C4 )  1 -110 110 2
assign (resid 9 and name O4') (resid 9 and name C1')
        (resid 9 and name N1 ) (resid 9 and name C2 )  1 -110 110 2
assign (resid 10 and name O4') (resid 10 and name C1')
        (resid 10 and name N1 ) (resid 10 and name C2 )  1 -110 110 2
assign (resid 10 and name H1') (resid 10 and name C1')
        (resid 10 and name C5 ) (resid 10 and name C6 )  1 180 45 2
assign (resid 11 and name O4') (resid 11 and name C1')
        (resid 11 and name N9 ) (resid 11 and name C4 )  1 -110 110 2
assign (resid 12 and name O4') (resid 12 and name C1')
        (resid 12 and name N1 ) (resid 12 and name C2 )  1 -110 110 2
assign (resid 13 and name O4') (resid 13 and name C1')
        (resid 13 and name N9 ) (resid 13 and name C4 )  1 -110 110 2
assign (resid 14 and name O4') (resid 14 and name C1')
        (resid 14 and name N9 ) (resid 14 and name C4 )  1 -110 110 2
assign (resid 15 and name O4') (resid 15 and name C1')
        (resid 15 and name N1 ) (resid 15 and name C2 )  1 -160 20 2
assign (resid 16 and name O4') (resid 16 and name C1')
        (resid 16 and name N9 ) (resid 16 and name C4 )  1 -160 20 2
assign (resid 17 and name O4') (resid 17 and name C1')
        (resid 17 and name N9 ) (resid 17 and name C4 )  1 -160 20 2
assign (resid 18 and name O4') (resid 18 and name C1')
        (resid 18 and name N1 ) (resid 18 and name C2 )  1 -160 20 2
```

```
assign (resid 19 and name O4') (resid 19 and name C1')

       (resid 19 and name N1 ) (resid 19 and name C2 )  1 -160 20 2
```

REFERENCES

Abeysirigunawardena, S. (2008). Biophysical investigations of structural and stability changes in Helix 69 of Eschrichia coli 23S rRNA. Department of Chemistry. Detroit, Wayne State University. **Ph. D.**

Al-Hashimi, H. M., S. W. Pitt, A. Majumdar, W. Xu and D. J. Patel (2003). "Mg2+-induced variations in the conformation and dynamics of HIV-1 TAR RNA probed using NMR residual dipolar couplings." Journal of Molecular Biology **329**(5): 867-873.

Albergo, D. D., L. A. Marky, K. J. Breslauer and D. H. Turner (1981). "Thermodynamics of (dG--dC)3 double-helix formation in water and deuterium oxide." Biochemistry **20**(6): 1409-1413.

Ali, I. K., L. Lancaster, J. Feinberg, S. Joseph and H. F. Noller (2006). "Deletion of a conserved, central ribosomal intersubunit RNA bridge." Mol Cell **23**(6): 865-874.

Altman, S. (1989). "Ribonuclease P: an enzyme with a catalytic RNA subunit." Adv Enzymol Relat Areas Mol Biol **62**: 1-36.

Altona, C. (1982). "Conformational analysis of nucleic acids. Determination of backbone geometry of single-helical RNA and DNA in aqueous solution." Recl.: J. R. Neth. Chem. Soc. FIELD Full Journal Title:Recueil: Journal of the Royal Netherlands Chemical Society **101**(12): 413-433.

Armache, J. P., A. Jarasch, A. M. Anger, E. Villa, T. Becker, S. Bhushan, F. Jossinet, M. Habeck, G. Dindar, S. Franckenberg, V. Marquez, T. Mielke, M. Thomm, O. Berninghausen, B. Beatrix, J. Soding, E. Westhof, D. N. Wilson and R. Beckmann (2010). "Cryo-EM structure and rRNA model of a translating

eukaryotic 80S ribosome at 5.5-A resolution." <u>Proc Natl Acad Sci U S A</u> **107**(46): 19748-19753.

Auffinger, P. and E. Westhof (1997). "RNA hydration: three nanoseconds of multiple molecular dynamics simulations of the solvated tRNA(Asp) anticodon hairpin." <u>J Mol Biol</u> **269**(3): 326-341.

Badis, G., M. Fromont-Racine and A. Jacquier (2003). "A snoRNA that guides the two most conserved pseudouridine modifications within rRNA confers a growth advantage in yeast." <u>RNA</u> **9**(7): 771-779.

Bartel, D. P. (2009). "MicroRNAs: Target Recognition and Regulatory Functions." <u>Cell</u> **136**(2): 215-233.

Basturea, G. N. and M. P. Deutscher (2007). "Substrate specificity and properties of the Escherichia coli 16S rRNA methyltransferase, RsmE." <u>RNA</u> **13**(11): 1969-1976.

Baudin-Baillieu, A., C. Fabret, X. H. Liang, D. Piekna-Przybylska, M. J. Fournier and J. P. Rousset (2009). "Nucleotide modifications in three functionally important regions of the Saccharomyces cerevisiae ribosome affect translation accuracy." <u>Nucleic Acids Res</u> **37**(22): 7665-7677.

Bax, A., G. M. Clore and A. M. Gronenborn (1990). "H-1-H-1 Correlation Via Isotropic Mixing of C-13 Magnetization, a New 3-Dimensional Approach for Assigning H-1 and C-13 Spectra of C-13-Enriched Proteins." <u>Journal of Magnetic Resonance</u> **88**(2): 425-431.

Ben-Bassat, A., K. Bauer, S. Y. Chang, K. Myambo, A. Boosman and S. Chang (1987). "Processing of the initiation methionine from proteins: properties of the

Escherichia coli methionine aminopeptidase and its gene structure." J Bacteriol **169**(2): 751-757.

Ben-Shem, A., N. Garreau de Loubresse, S. Melnikov, L. Jenner, G. Yusupova and M. Yusupov (2011). "The structure of the eukaryotic ribosome at 3.0 A resolution." Science **334**(6062): 1524-1529.

Ben-Shem, A., L. Jenner, G. Yusupova and M. Yusupov (2010). "Crystal structure of the eukaryotic ribosome." Science **330**(6008): 1203-1209.

Berk, V., W. Zhang, R. D. Pai and J. H. Cate (2006). "Structural basis for mRNA and tRNA positioning on the ribosome." Proc Natl Acad Sci U S A **103**(43): 15830-15834.

Bevington, P. R. (1969). Data reduction and error analysis for the physical sciences. New York,, McGraw-Hill.

Blanchard, S. C. and J. D. Puglisi (2001). "Solution structure of the A loop of 23S ribosomal RNA." Proc Natl Acad Sci U S A **98**(7): 3720-3725.

Bommarito, S., N. Peyret and J. SantaLucia, Jr. (2000). "Thermodynamic parameters for DNA sequences with dangling ends." Nucleic Acids Res **28**(9): 1929-1934.

Borer, P. N., B. Dengler, I. Tinoco, Jr. and O. C. Uhlenbeck (1974). "Stability of ribonucleic acid double-stranded helices." J Mol Biol **86**(4): 843-853.

Borovinskaya, M. A., R. D. Pai, W. Zhang, B. S. Schuwirth, J. M. Holton, G. Hirokawa, H. Kaji, A. Kaji and J. H. Cate (2007). "Structural basis for aminoglycoside inhibition of bacterial ribosome recycling." Nat Struct Mol Biol **14**(8): 727-732.

Brahms, J. and W. F. Mommaerts (1964). "A Study of Conformation of Nucleic Acids in Solution by Means of Circular Dichroism." J Mol Biol **10**: 73-88.

Breslauer, K. J. (1995). "Extracting thermodynamic data from equilibrium melting curves for oligonucleotide order-disorder transitions." <u>Methods Enzymol</u> **259**: 221-242.

Brunger, A. T. (2007). "Version 1.2 of the Crystallography and NMR system." <u>Nat Protoc</u> **2**(11): 2728-2733.

Brunger, A. T., P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson and G. L. Warren (1998). "Crystallography & NMR system: A new software suite for macromolecular structure determination." <u>Acta Crystallogr D Biol Crystallogr</u> **54**(Pt 5): 905-921.

Busch, H., R. Reddy, L. Rothblum and Y. C. Choi (1982). "Snrnas, Snrnps, and Rna Processing." <u>Annu Rev Biochem</u> **51**: 617-654.

Cabello-Villegas, J. and E. P. Nikonowicz (2005). "Solution structure of psi32-modified anticodon stem-loop of Escherichia coli tRNAPhe." <u>Nucleic Acids Research</u> **33**(22): 6961-6971.

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

Cantor, C. R. and P. R. Schimmel (1980). <u>Techniques for the study of biological structure and function</u>. San Francisco, W. H. Freeman.

Carter, A. P., W. M. Clemons, D. E. Brodersen, R. J. Morgan-Warren, B. T. Wimberly and V. Ramakrishnan (2000). "Functional insights from the structure of the 30S

ribosomal subunit and its interactions with antibiotics." <u>Nature</u> **407**(6802): 340-348.

Carter, A. P., W. M. Clemons, Jr., D. E. Brodersen, R. J. Morgan-Warren, T. Hartsch, B. T. Wimberly and V. Ramakrishnan (2001). "Crystal structure of an initiation factor bound to the 30S ribosomal subunit." <u>Science</u> **291**(5503): 498-501.

Cate, J. H., A. R. Gooding, E. Podell, K. Zhou, B. L. Golden, C. E. Kundrot, T. R. Cech and J. A. Doudna (1996). "Crystal structure of a group I ribozyme domain: principles of RNA packing." <u>Science</u> **273**(5282): 1678-1685.

Cate, J. H., M. M. Yusupov, G. Z. Yusupova, T. N. Earnest and H. F. Noller (1999). "X-ray crystal structures of 70S ribosome functional complexes." <u>Science</u> **285**(5436): 2095-2104.

Cech, T. R. (1990). "Self-splicing of group I introns." <u>Annu Rev Biochem</u> **59**: 543-568.

Cerny, J. and P. Hobza (2007). "Non-covalent interactions in biomacromolecules." <u>Phys Chem Chem Phys</u> **9**(39): 5291-5303.

Charette, M. and M. W. Gray (2000). "Pseudouridine in RNA: what, where, how, and why." <u>IUBMB Life</u> **49**(5): 341-351.

Charollais, J., D. Pflieger, J. Vinh, M. Dreyfus and I. Iost (2003). "The DEAD-box RNA helicase SrmB is involved in the assembly of 50S ribosomal subunits in Escherichia coli." <u>Mol Microbiol</u> **48**(5): 1253-1265.

Chattopadhyay, A., C. Thibaudeau and P. Acharaya (1999). <u>Stereoelectronic Effects in Nucleosides and Nucleotides and their Structural Implications</u>, Uppsala University Press.

Chow, C. S. and F. M. Bogdan (1997). "A Structural Basis for RNAminus signLigand Interactions." <u>Chem Rev</u> **97**(5): 1489-1514.

Chow, C. S., T. N. Larnichhane and S. K. Mahto (2007). "Expanding the nucleotide repertoire of the ribosome with post-transcriptional modifications." <u>Acs Chemical Biology</u> **2**(9): 610-619.

Chow, P. K., T. H. Ng, M. Chew, I. C. Song, H. C. Kee and P. O. Mack (1999). "A practical technique of colour image analysis: applications in experimental research." <u>Ann Acad Med Singapore</u> **28**(1): 155-158.

Chui, H. M., J. P. Desaulniers, S. A. Scaringe and C. S. Chow (2002). "Synthesis of helix 69 of Escherichia coli 23S rRNA containing its natural modified nucleosides, m(3)Psi and Psi." <u>J Org Chem</u> **67**(25): 8847-8854.

Claridge, T. D. W. (1999). <u>High-resolution NMR techniques in organic chemistry</u>. Amsterdam ; New York, Pergamon.

Clore, G. M. and A. M. Gronenborn (1989). "Determination of three-dimensional structures of proteins and nucleic acids in solution by nuclear magnetic resonance spectroscopy." <u>Crit Rev Biochem Mol Biol</u> **24**(5): 479-564.

Clore, G. M. and A. M. Gronenborn (1998). "New methods of structure refinement for macromolecular structure determination by NMR." <u>Proceedings of the National Academy of Sciences of the United States of America</u> **95**(11): 5891-5898.

Crick, F. (1956). "Ideas on protein synthesis."

Crick, F. H. (1966). "Codon--anticodon pairing: the wobble hypothesis." <u>J Mol Biol</u> **19**(2): 548-555.

Culver, G. M. and H. F. Noller (2000). "In vitro reconstitution of 30S ribosomal subunits using complete set of recombinant proteins." <u>Methods Enzymol</u> **318**: 446-460.

D'Souza, V., A. Dey, D. Habib and M. F. Summers (2004). "NMR structure of the 101-nucleotide core encapsidation signal of the Moloney murine leukemia virus." <u>J Mol Biol</u> **337**(2): 427-442.

Davanloo, P., A. H. Rosenberg, J. J. Dunn and F. W. Studier (1984). "Cloning and expression of the gene for bacteriophage T7 RNA polymerase." <u>Proc Natl Acad Sci U S A</u> **81**(7): 2035-2039.

Davies, J. and B. D. Davis (1968). "Misreading of ribonucleic acid code words induced by aminoglycoside antibiotics. The effect of drug concentration." <u>Journal of Biological Chemistry</u> **243**(12): 3312-3316.

Davis, D. R. (1995). "Stabilization of RNA stacking by pseudouridine." <u>Nucleic Acids Research</u> **23**(24): 5020-5026.

Davis, J. H., M. Tonelli, L. G. Scott, L. Jaeger, J. R. Williamson and S. E. Butcher (2005). "RNA helical packing in solution: NMR structure of a 30 kDa GAAA tetraloop-receptor complex." <u>J Mol Biol</u> **351**(2): 371-382.

de Narvaez, C. C. and H. W. Schaup (1979). "In vivo transcriptionally coupled assembly of Escherichia coli ribosomal subunits." <u>J Mol Biol</u> **134**(1): 1-22.

Decatur, W. A. and M. J. Fournier (2002). "rRNA modifications and ribosome function." <u>Trends Biochem Sci</u> **27**(7): 344-351.

Del Campo, M., C. Recinos, G. Yanez, S. C. Pomerantz, R. Guymon, P. F. Crain, J. A. McCloskey and J. Ofengand (2005). "Number, position, and significance of the

pseudouridines in the large subunit ribosomal RNA of Haloarcula marismortui and Deinococcus radiodurans." <u>RNA</u> **11**(2): 210-219.

Demirci, H., L. H. G. Larsen, T. Hansen, A. Rasmussen, A. Cadambi, S. T. Gregory, F. Kirpekar and G. Jogl (2010). "Multi-site-specific 16S rRNA methyltransferase RsmF from Thermus thermophilus." <u>RNA (New York, N Y)</u> **16**(8): 1584-1596.

Denmon, A. P., J. Wang and E. P. Nikonowicz (2011). "Conformation effects of base modification on the anticodon stem-loop of Bacillus subtilis tRNA(Tyr)." <u>J Mol Biol</u> **412**(2): 285-303.

Desaulniers, J. P., Y. C. Chang, R. Aduri, S. C. Abeysirigunawardena, J. SantaLucia, Jr. and C. S. Chow (2008). "Pseudouridines in rRNA helix 69 play a role in loop stacking interactions." <u>Org Biomol Chem</u> **6**(21): 3892-3895.

Desaulniers, J. P., H. M. Chui and C. S. Chow (2005). "Solution conformations of two naturally occurring RNA nucleosides: 3-methyluridine and 3-methylpseudouridine." <u>Bioorg Med Chem</u> **13**(24): 6777-6781.

Devoe, H. and I. Tinoco, Jr. (1962). "The hypochromism of helical polynucleotides." <u>J Mol Biol</u> **4**: 518-527.

Devoe, H. and I. Tinoco, Jr. (1962). "The stability of helical polynucleotides: base contributions." <u>J Mol Biol</u> **4**: 500-517.

Dieckmann, T. and J. Feigon (1997). "Assignment methodology for larger RNA oligonucleotides: application to an ATP-binding RNA aptamer." <u>J Biomol NMR</u> **9**(3): 259-272.

Diener, J. L. and P. B. Moore (1998). "Solution structure of a substrate for the archaeal pre-tRNA splicing endonucleases: The bulge-helix-bulge motif." Molecular Cell **1**(6): 883-894.

Dingley, A. J. and S. Grzesiek (1998). "Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide (2)J(NN) couplings." Journal of the American Chemical Society **120**(33): 8293-8297.

Dodds, J. A., T. J. Morris and R. L. Jordan (1984). "Plant Viral Double-Stranded-Rna." Annual Review of Phytopathology **22**: 151-168.

Donly, B. C. and G. A. Mackie (1988). "Affinities of ribosomal protein S20 and C-terminal deletion mutants for 16S rRNA and S20 mRNA." Nucleic Acids Res **16**(3): 997-1010.

Doty, P., H. Boedtker, J. R. Fresco, R. Haselkorn and M. Litt (1959). "Secondary Structure in Ribonucleic Acids." Proc Natl Acad Sci U S A **45**(4): 482-499.

Duc, A.-C. (2009). TOWARDS THE DEVELOPMENT OF NEW ANTI-INFECTIVES: TARGETING BACTERIAL RNAS AND THE RIBOSOME WITH ANTIBIOTIC LEADS. Department of Chemistry. Detroit, Wayne State University. **Ph. D**.

Dunkle, J. A., L. Wang, M. B. Feldman, A. Pulk, V. B. Chen, G. J. Kapral, J. Noeske, J. S. Richardson, S. C. Blanchard and J. H. Cate (2011). "Structures of the bacterial ribosome in classical and hybrid states of tRNA binding." Science **332**(6032): 981-984.

Durant, P. C. and D. R. Davis (1995). "Stabilization of RNA stacking by pseudouridine." Nucleic Acids Res **23**(24): 5020-5026.

Durant, P. C. and D. R. Davis (1999). "Stabilization of the anticodon stem-loop of tRNALys,3 by an A+-C base-pair and by pseudouridine." J Mol Biol **285**(1): 115-131.

Dutca, L. M. and G. M. Culver (2008). "Assembly of the 5' and 3' minor domains of 16S ribosomal RNA as monitored by tethered probing from ribosomal protein S20." Journal of Molecular Biology **376**(1): 92-108.

Ejby, M., M. A. Sorensen and S. Pedersen (2007). "Pseudouridylation of helix 69 of 23S rRNA is necessary for an effective translation termination." Proc Natl Acad Sci U S A **104**(49): 19410-19415.

El Hage, A., M. Sbai and J. H. Alix (2001). "The chaperonin GroEL and other heat-shock proteins, besides DnaK, participate in ribosome biogenesis in Escherichia coli." Mol Gen Genet **264**(6): 796-808.

Ero, R., L. Peil, A. Liiv and J. Remme (2008). "Identification of pseudouridine methyltransferase in Escherichia coli." RNA **14**(10): 2223-2233.

Feig, A. and O. Ulenbeck (1999). The Role of Metal Ions in RNA Biochemistry. The RNA World. **37**.

Ferat, J. L. and F. Michel (1993). "Group-Ii Self-Splicing Introns in Bacteria." Nature **364**(6435): 358-361.

Fischer, N., A. L. Konevega, W. Wintermeyer, M. V. Rodnina and H. Stark (2010). "Ribosome dynamics and tRNA movement by time-resolved electron cryomicroscopy." Nature **466**(7304): 329-333.

Flanagan, M. T. and R. H. Pantell (1984). "Surface-Plasmon Resonance and Immunosensors." Electronics Letters **20**(23): 968-970.

Francois, B., R. J. Russell, J. B. Murray, F. Aboul-ela, B. Masquida, Q. Vicens and E. Westhof (2005). "Crystal structures of complexes between aminoglycosides and decoding A site oligonucleotides: role of the number of rings and positive charges in the specific binding leading to miscoding." Nucleic Acids Res **33**(17): 5677-5690.

Frank, J., A. Verschoor, Y. Li, J. Zhu, R. K. Lata, M. Radermacher, P. Penczek, R. Grassucci, R. K. Agrawal and S. Srivastava (1995). "A model of the translational apparatus based on a three-dimensional reconstruction of the Escherichia coli ribosome." Biochem Cell Biol **73**(11-12): 757-765.

Frank, J., J. Zhu, P. Penczek, Y. Li, S. Srivastava, A. Verschoor, M. Radermacher, R. Grassucci, R. K. Lata and R. K. Agrawal (1995). "A model of protein synthesis based on cryo-electron microscopy of the E. coli ribosome." Nature **376**(6539): 441-444.

Freier, S. M., D. Alkema, A. Sinclair, T. Neilson and D. H. Turner (1985). "Contributions of dangling end stacking and terminal base-pair formation to the stabilities of XGGCCp, XCCGGp, XGGCCYp, and XCCGGYp helixes." Biochemistry **24**(17): 4533-4539.

Freier, S. M., B. J. Burger, D. Alkema, T. Neilson and D. H. Turner (1983). "Effects of 3' Dangling End Stacking on the Stability of Ggcc and Ccgg Double Helices." Biochemistry **22**(26): 6198-6206.

Freier, S. M., R. Kierzek, J. A. Jaeger, N. Sugimoto, M. H. Caruthers, T. Neilson and D. H. Turner (1986). "Improved free-energy parameters for predictions of RNA duplex stability." Proc Natl Acad Sci U S A **83**(24): 9373-9377.

Freistroffer, D. V., M. Y. Pavlov, J. MacDougall, R. H. Buckingham and M. Ehrenberg (1997). "Release factor RF3 in E.coli accelerates the dissociation of release factors RF1 and RF2 from the ribosome in a GTP-dependent manner." EMBO J **16**(13): 4126-4133.

Frolova, L. Y., R. Y. Tsivkovskii, G. F. Sivolobova, N. Y. Oparina, O. I. Serpinsky, V. M. Blinov, S. I. Tatkov and L. L. Kisselev (1999). "Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis." RNA **5**(8): 1014-1020.

Fürtig, B. and C. R. J. W. H. Schwalbe (2003). "NMR Spectroscopy of RNA." ChemBioChem **4**(10): 936-962.

Gabashvili, I. S., R. K. Agrawal, C. M. Spahn, R. A. Grassucci, D. I. Svergun, J. Frank and P. Penczek (2000). "Solution structure of the E. coli 70S ribosome at 11.5 A resolution." Cell **100**(5): 537-549.

Gautheret, D. and R. R. Gutell (1997). "Inferring the conformation of RNA base pairs and triples from patterns of sequence variation." Nucleic Acids Res **25**(8): 1559-1564.

Getz, M., X. Sun, A. Casiano-Negroni, Q. Zhang and H. M. Al-Hashimi (2007). "NMR studies of RNA dynamics and structural plasticity using NMR residual dipolar couplings." Biopolymers **86**(5-6): 384-402.

Gilbert, W. (1986). "Origin of life: The RNA world." Nature **319**(6055): 618-618.

Ginsburg, D. and J. A. Steitz (1975). "The 30 S ribosomal precursor RNA from Escherichia coli. A primary transcript containing 23 S, 16 S, and 5 S sequences." Journal of Biological Chemistry **250**(14): 5647-5654.

Gorenstein, D. G. (1981). "Nucleotide conformational analysis by 31P nuclear magnetic resonance spectroscopy." <u>Annu Rev Biophys Bioeng</u> **10**: 355-386.

Gotz, F., E. R. Dabbs and C. O. Gualerzi (1990). "Escherichia coli 30S mutants lacking protein S20 are defective in translation initiation." <u>Biochim Biophys Acta</u> **1050**(1-3): 93-97.

Gratzer, W. B. and R. Mendelsohn (1978). <u>Optical methods for studying protein conformation</u>, Elsevier.

Green, G. A. and D. S. Jones (1986). "The nucleotide sequence of a cytoplasmic tRNAPhe from Scenedesmus obliquus and comparison with a tRNATyr species." <u>Biochem J</u> **236**(2): 601-603.

Greenfield, N. and G. D. Fasman (1969). "Computed circular dichroism spectra for the evaluation of protein conformation." <u>Biochemistry</u> **8**(10): 4108-4116.

Greider, C. W. and E. H. Blackburn (1989). "A Telomeric Sequence in the Rna of Tetrahymena Telomerase Required for Telomere Repeat Synthesis." <u>Nature</u> **337**(6205): 331-337.

Griffey, R. H., C. D. Poulter, A. Bax, B. L. Hawkins, Z. Yamaizumi and S. Nishimura (1983). "Multiple quantum two-dimensional 1H--15N nuclear magnetic resonance spectroscopy: chemical shift correlation maps for exchangeable imino protons of Escherichia coli tRNAMetf in water." <u>Proc Natl Acad Sci U S A</u> **80**(19): 5895-5897.

Grondek, J. F. and G. M. Culver (2004). "Assembly of the 30S ribosomal subunit: positioning ribosomal protein S13 in the S7 assembly branch." <u>RNA</u> **10**(12): 1861-1866.

Gutell, R. R., J. C. Lee and J. J. Cannone (2002). "The accuracy of ribosomal RNA comparative structure models." Curr Opin Struct Biol **12**(3): 301-310.

Hall, K. B. (1995). "Uses of 13C- and 15N-labeled RNA in NMR of RNA-protein complexes." Methods Enzymol **261**: 542-559.

Harms, J., F. Schluenzen, R. Zarivach, A. Bashan, S. Gat, I. Agmon, H. Bartels, F. Franceschi and A. Yonath (2001). "High resolution structure of the large ribosomal subunit from a mesophilic eubacterium." Cell **107**(5): 679-688.

Hartz, D., J. Binkley, T. Hollingsworth and L. Gold (1990). "Domains of initiator tRNA and initiation codon crucial for initiator tRNA selection by Escherichia coli IF3." Genes Dev **4**(10): 1790-1800.

Held, W. A., B. Ballou, S. Mizushima and M. Nomura (1974). "Assembly mapping of 30 S ribosomal proteins from Escherichia coli. Further studies." Journal of Biological Chemistry **249**(10): 3103-3111.

Held, W. A., S. Mizushima and M. Nomura (1973). "Reconstitution of Escherichia coli 30 S ribosomal subunits from purified molecular components." Journal of Biological Chemistry **248**(16): 5720-5730.

Herold, M. and K. H. Nierhaus (1987). "Incorporation of six additional proteins to complete the assembly map of the 50 S subunit from Escherichia coli ribosomes." Journal of Biological Chemistry **262**(18): 8826-8833.

Hiller, D. A., V. Singh, M. Zhong and S. A. Strobel (2011). "A two-step chemical mechanism for ribosome-catalysed peptide bond formation." Nature **476**(7359): 236-239.

Hirashima, A. and A. Kaji (1972). "Purification and properties of ribosome-releasing factor." Biochemistry **11**(22): 4037-4044.

Holley, R. W., G. A. Everett, J. T. Madison and A. Zamir (1965). "Nucleotide Sequences in Yeast Alanine Transfer Ribonucleic Acid." Journal of Biological Chemistry **240**(5): 2122-&.

Holzwarth, G. and P. Doty (1965). "The Ultraviolet Circular Dichroism of Polypeptides." J Am Chem Soc **87**: 218-228.

Homann, H. E. and K. H. Nierhaus (1971). "Ribosomal proteins. Protein compositions of biosynthetic precursors and artifical subparticles from ribosomal subunits in Escherichia coli K 12." Eur J Biochem **20**(2): 249-257.

Ibba, M. and D. Soll (1999). "Quality control mechanisms during translation." Science (New York, N Y) **286**(5446): 1893-1897.

Inoue, K., J. Alsina, J. Chen and M. Inouye (2003). "Suppression of defective ribosome assembly in a rbfA deletion mutant by overexpression of Era, an essential GTPase in Escherichia coli." Mol Microbiol **48**(4): 1005-1016.

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

Jakel, S. and D. Gorlich (1998). "Importin beta, transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells." EMBO J **17**(15): 4491-4502.

Jenner, L. B., N. Demeshkina, G. Yusupova and M. Yusupov (2010). "Structural aspects of messenger RNA reading frame maintenance by the ribosome." Nat Struct Mol Biol **17**(5): 555-560.

Jin, H., A. C. Kelley and V. Ramakrishnan (2011). "Crystal structure of the hybrid state of ribosome in complex with the guanosine triphosphatase release factor 3." Proc Natl Acad Sci U S A **108**(38): 15798-15803.

John SantaLucia, J. (2000). The use of spectroscopic techniques in the study of DNA stability. Oxford ; New York, Oxford University Press.

Jones, P. G., M. Mitta, Y. Kim, W. Jiang and M. Inouye (1996). "Cold shock induces a major ribosomal-associated protein that unwinds double-stranded RNA in Escherichia coli." Proc Natl Acad Sci U S A **93**(1): 76-80.

Jorgensen, W. L. and D. L. Severance (1990). "Aromatic Aromatic Interactions - Free-Energy Profiles for the Benzene Dimer in Water, Chloroform, and Liquid Benzene." Journal of the American Chemical Society **112**(12): 4768-4774.

Kaczanowska, M. and M. Ryden-Aulin (2007). "Ribosome biogenesis and the translation process in Escherichia coli." Microbiology and molecular biology reviews : MMBR **71**(3): 477-494.

Kamita, M., Y. Kimura, Y. Ino, R. M. Kamp, B. Polevoda, F. Sherman and H. Hirano (2011). "N(alpha)-Acetylation of yeast ribosomal proteins and its effect on protein synthesis." Journal of Proteomics **74**(4): 431-441.

Kay, L. E., M. Ikura, R. Tschudin and A. Bax (2011). "Three-dimensional triple-resonance NMR Spectroscopy of isotopically enriched proteins. 1990." Journal of magnetic resonance (San Diego, Calif : 1997) **213**(2): 423-441.

Kay, L. E., G. Y. Xu, A. U. Singer, D. R. Muhandiram and J. D. Formankay (1993). "A Gradient-Enhanced Hcch Tocsy Experiment for Recording Side-Chain H-1 and

C-13 Correlations in H2o Samples of Proteins." <u>Journal of Magnetic Resonance Series B</u> **101**(3): 333-337.

Khaitovich, P., T. Tenson, P. Kloss and A. S. Mankin (1999). "Reconstitution of functionally active Thermus aquaticus large ribosomal subunits with in vitro-transcribed rRNA." <u>Biochemistry</u> **38**(6): 1780-1788.

Kim, N.-K., C. A. Theimer, J. R. Mitchell, K. Collins and J. Feigon (2010). "Effect of pseudouridylation on the structure and activity of the catalytically essential P6.1 hairpin in human telomerase RNA." <u>Nucleic Acids Research</u> **38**(19): 6746-6756.

Kim, N. K., Q. Zhang, J. Zhou, C. A. Theimer, R. D. Peterson and J. Feigon (2008). "Solution structure and dynamics of the wild-type pseudoknot of human telomerase RNA." <u>J Mol Biol</u> **384**(5): 1249-1261.

Kipper, K., C. Hetenyi, S. Sild, J. Remme and A. Liiv (2009). "Ribosomal intersubunit bridge B2a is involved in factor-dependent translation initiation and translational processivity." <u>J Mol Biol</u> **385**(2): 405-422.

Klein, D. J., P. B. Moore and T. A. Steitz (2004). "The contribution of metal ions to the structural stability of the large ribosomal subunit." <u>RNA</u> **10**(9): 1366-1379.

Klein, D. J., T. M. Schmeing, P. B. Moore and T. A. Steitz (2001). "The kink-turn: a new RNA secondary structure motif." <u>Embo J</u> **20**(15): 4214-4221.

Klinge, S., F. Voigts-Hoffmann, M. Leibundgut, S. Arpagaus and N. Ban (2011). "Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6." <u>Science</u> **334**(6058): 941-948.

Kooyman, R. P. H., H. Kolkman, J. Vangent and J. Greve (1988). "Surface-Plasmon Resonance Immunosensors - Sensitivity Considerations." <u>Analytica Chimica Acta</u> **213**(1-2): 35-45.

Korostelev, A., J. Zhu, H. Asahara and H. F. Noller (2010). "Recognition of the amber UAG stop codon by release factor RF1." <u>EMBO J</u> **29**(15): 2577-2585.

Kowalak, J. A., E. Bruenger and J. A. Mccloskey (1995). "Posttranscriptional Modification of the Central Loop of Domain-V in Escherichia-Coli 23-S Ribosomal-Rna." <u>Journal of Biological Chemistry</u> **270**(30): 17758-17764.

Lamichhane, A. (2009). Identification of drug targets and drug leads in *Pseudomonas aeruginosa*. <u>Department of Biological Sciences</u>. Detroit, Wayne State University. **M. S.**

Lane, D. J., B. Pace, G. J. Olsen, D. A. Stahl, M. L. Sogin and N. R. Pace (1985). "Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses." <u>Proceedings of the National Academy of Sciences of the United States of America</u> **82**(20): 6955-6959.

Latham, M. P., D. J. Brown, S. A. McCallum and A. Pardi (2005). "NMR methods for studying the structure and dynamics of RNA." <u>Chembiochem : a European journal of chemical biology</u> **6**(9): 1492-1505.

Lecompte, O., R. Ripp, J. C. Thierry, D. Moras and O. Poch (2002). "Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale." <u>Nucleic Acids Res</u> **30**(24): 5382-5390.

Lee, K., C. A. Holland-Staley and P. R. Cunningham (1996). "Genetic analysis of the Shine-Dalgarno interaction: selection of alternative functional mRNA-rRNA combinations." RNA (New York, N Y) **2**(12): 1270-1285.

Lee, K., S. Varma, J. SantaLucia, Jr. and P. R. Cunningham (1997). "In vivo determination of RNA structure-function relationships: analysis of the 790 loop in ribosomal RNA." Journal of Molecular Biology **269**(5): 732-743.

Leontis, N. B. and E. Westhof (1998). "The 5S rRNA loop E: chemical probing and phylogenetic data versus crystal structure." Rna **4**(9): 1134-1153.

Leppik, M., L. Peil, K. Kipper, A. Liiv and J. Remme (2007). "Substrate specificity of the pseudouridine synthase RluD in Escherichia coli." FEBS J **274**(21): 5759-5766.

Li, Z., S. Pandit and M. P. Deutscher (1999). "Maturation of 23S ribosomal RNA requires the exoribonuclease RNase T." RNA **5**(1): 139-146.

Li, Z., S. Pandit and M. P. Deutscher (1999). "RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA." EMBO J **18**(10): 2878-2885.

Liang, X. H., Q. Liu and M. J. Fournier (2007). "rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity." Mol Cell **28**(6): 965-977.

Liedberg, B., C. Nylander and I. Lundstrom (1995). "Biosensing with surface plasmon resonance--how it all started." Biosens Bioelectron **10**(8): i-ix.

Liiv, A., D. Karitkina, U. Maivali and J. Remme (2005). "Analysis of the function of E. coli 23S rRNA helix-loop 69 by mutagenesis." BMC Mol Biol **6**: 18.

Lindahl, L. (1973). "Two new ribosomal precursor particles in E. coli." <u>Nat New Biol</u> **243**(127): 170-172.

Lovgren, J. M., G. O. Bylund, M. K. Srivastava, L. A. Lundberg, O. P. Persson, G. Wingsle and P. M. Wikstrom (2004). "The PRC-barrel domain of the ribosome maturation protein RimM mediates binding to ribosomal protein S19 in the 30S ribosomal subunits." <u>RNA</u> **10**(11): 1798-1812.

Maivali, U. and J. Remme (2004). "Definition of bases in 23S rRNA essential for ribosomal subunit association." <u>RNA</u> **10**(4): 600-604.

Mangiarotti, G., D. Apirion, D. Schlessinger and L. Silengo (1968). "Biosynthetic precursors of 30S and 50S ribosomal particles in Escherichia coli." <u>Biochemistry</u> **7**(1): 456-472.

Marchler-Bauer, A., S. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, F. Lu, G. H. Marchler, M. Mullokandov, M. V. Omelchenko, C. L. Robertson, J. S. Song, N. Thanki, R. A. Yamashita, D. Zhang, N. Zhang, C. Zheng and S. H. Bryant (2011). "CDD: a Conserved Domain Database for the functional annotation of proteins." <u>Nucleic Acids Res</u> **39**(Database issue): D225-229.

Marquez, V., T. Frohlich, J.-P. Armache, D. Sohmen, A. Donhofer, A. Mikolajka, O. Berninghausen, M. Thomm, R. Beckmann, G. J. Arnold and D. N. Wilson (2011). "Proteomic characterization of archaeal ribosomes reveals the presence of novel archaeal-specific ribosomal proteins." <u>Journal of Molecular Biology</u> **405**(5): 1215-1232.

Martino, L., S. Pennell, G. Kelly, T. T. Bui, O. Kotik-Kogan, S. J. Smerdon, A. F. Drake, S. Curry and M. R. Conte (2011). "Analysis of the interaction with the hepatitis C virus mRNA reveals an alternative mode of RNA recognition by the human La protein." Nucleic Acids Res.

McCarthy, B. J., R. J. Britten and R. B. Roberts (1962). "The Synthesis of Ribosomes in E. coli: III. Synthesis of Ribosomal RNA." Biophys J **2**(1): 57-82.

McDowell, J. A. and D. H. Turner (1996). "Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: solution structure of (rGAGGUCUC)2 by two-dimensional NMR and simulated annealing." Biochemistry **35**(45): 14077-14089.

Melzer, M. J., D. M. Sether, A. V. Karasev, W. Borth and J. S. Hu (2008). "Complete nucleotide sequence and genome organization of pineapple mealybug wilt-associated virus-1." Arch Virol **153**(4): 707-714.

Mengel-Jorgensen, J., S. S. Jensen, A. Rasmussen, J. Poehlsgaard, J. J. L. Iversen and F. Kirpekar (2006). "Modifications in Thermus thermophilus 23 S ribosomal RNA are centered in regions of RNA-RNA contact." Journal of Biological Chemistry **281**(31): 22108-22117.

Meroueh, M., P. J. Grohar, J. Qiu, J. SantaLucia, Jr., S. A. Scaringe and C. S. Chow (2000). "Unique structural and stabilizing roles for the individual pseudouridine residues in the 1920 region of Escherichia coli 23S rRNA." Nucleic Acids Res **28**(10): 2075-2083.

Miller, P. A., Z. Shajani, G. A. Meints, D. Caplow, G. Goobes, G. Varani and G. P. Drobny (2006). "Contrasting views of the internal dynamics of the HhaI

methyltransferase target DNA reported by solution and solid-state NMR spectroscopy." J Am Chem Soc **128**(50): 15970-15971.

Milligan, J. F., D. R. Groebe, G. W. Witherell and O. C. Uhlenbeck (1987). "Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates." Nucleic Acids Res **15**(21): 8783-8798.

Mizushima, S. and M. Nomura (1970). "Assembly mapping of 30S ribosomal proteins from E. coli." Nature **226**(5252): 1214.

Moazed, D. and H. F. Noller (1989). "Intermediate states in the movement of transfer RNA in the ribosome." Nature **342**(6246): 142-148.

Moazed, D., S. Stern and H. F. Noller (1986). "Rapid chemical probing of conformation in 16 S ribosomal RNA and 30 S ribosomal subunits using primer extension." J Mol Biol **187**(3): 399-416.

Mora, L., V. Heurgue-Hamard, S. Champ, M. Ehrenberg, L. L. Kisselev and R. H. Buckingham (2003). "The essential role of the invariant GGQ motif in the function and stability in vivo of bacterial release factors RF1 and RF2." Mol Microbiol **47**(1): 267-275.

Morosyuk, S. V., P. R. Cunningham and J. SantaLucia, 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**(1): 197-211.

Motorin, Y. and M. Helm (2011). "RNA nucleotide methylation." Advanced review **2**(September/October): 611-631.

Neuhaus, D. (2011). Nuclear Overhauser Effect. Encyclopedia of Magnetic Resonance.

Neuhaus, D. and M. P. Williamson (1989). The nuclear Overhauser effect in structural and conformational analysis. New York City, VCH.

Newby, M. I. and N. L. Greenbaum (2002). "Investigation of Overhauser effects between pseudouridine and water protons in RNA helices." Proc Natl Acad Sci U S A **99**(20): 12697-12702.

Newby, M. I. and N. L. Greenbaum (2002). "Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine." Nat Struct Biol **9**(12): 958-965.

Nierhaus, K. H., K. Bordasch and H. E. Homann (1973). "Ribosomal proteins. 43. In vivo assembly of Escherichia coli ribosomal proteins." J Mol Biol **74**(4): 587-597.

Nierhaus, K. H. and D. N. Wilson (2004). Protein Synthesis and Ribosome Structure, WILEY-VCH Verlag GmbH & Co. KGaA.

Nikonowicz, E. P. and A. Pardi (1992). "Three-dimensional heteronuclear NMR studies of RNA." Nature **355**(6356): 184-186.

Nikonowicz, E. P. and A. Pardi (1993). "An efficient procedure for assignment of the proton, carbon and nitrogen resonances in 13C/15N labeled nucleic acids." J Mol Biol **232**(4): 1141-1156.

Nissen, P., J. A. Ippolito, N. Ban, P. B. Moore and T. A. Steitz (2001). "RNA tertiary interactions in the large ribosomal subunit: the A-minor motif." Proc Natl Acad Sci U S A **98**(9): 4899-4903.

Nitta, I., T. Ueda and K. Watanabe (1998). "Possible involvement of Escherichia coli 23S ribosomal RNA in peptide bond formation." RNA **4**(3): 257-267.

Noller, H. F., V. Hoffarth and L. Zimniak (1992). "Unusual resistance of peptidyl transferase to protein extraction procedures." Science **256**(5062): 1416-1419.

Nord, S., G. O. Bylund, J. M. Lovgren and P. M. Wikstrom (2009). "The RimP protein is important for maturation of the 30S ribosomal subunit." J Mol Biol **386**(3): 742-753.

O'Farrell, H. C., N. Pulicherla, P. M. Desai and J. P. Rife (2006). "Recognition of a complex substrate by the KsgA/Dim1 family of enzymes has been conserved throughout evolution." RNA (New York, N Y) **12**(5): 725-733.

O'Toole, A. S., S. Miller and M. J. Serra (2005). "Stability of 3' double nucleotide overhangs that model the 3' ends of siRNA." RNA **11**(4): 512-516.

Ofengand, J. (2002). "Ribosomal RNA pseudouridines and pseudouridine synthases." FEBS Lett **514**(1): 17-25.

Ofengand, J. and A. Bakin (1997). "Mapping to nucleotide resolution of pseudouridine residues in large subunit ribosomal RNAs from representative eukaryotes, prokaryotes, archaebacteria, mitochondria and chloroplasts." J Mol Biol **266**(2): 246-268.

Olsen, G. J. and C. R. Woese (1993). "Ribosomal RNA: a key to phylogeny." FASEB journal : official publication of the Federation of American Societies for Experimental Biology **7**(1): 113-123.

Ordal, M. A., L. L. Long, R. J. Bell, S. E. Bell, R. R. Bell, R. W. Alexander, Jr. and C. A. Ward (1983). "Optical properties of the metals Al, Co, Cu, Au, Fe, Pb, Ni, Pd, Pt, Ag, Ti, and W in the infrared and far infrared." Appl Opt **22**(7): 1099-1020.

Ortiz-Meoz, R. F. and R. Green (2011). "Helix 69 is key for uniformity during substrate selection on the ribosome." Journal of Biological Chemistry **286**(29): 25604-25610.

Osawa, S., E. Otaka, T. Itoh and T. Fukui (1969). "Biosynthesis of 50 s ribosomal subunit in Escherichia coli." J Mol Biol **40**(3): 321-351.

Pardi, A. (1995). "Multidimensional heteronuclear NMR experiments for structure determination of isotopically labeled RNA." Methods Enzymol **261**: 350-380.

Pardi, A. and E. P. Nikonowicz (1992). "Simple procedure for resonance assignment of the sugar protons in carbon-13 labeled RNAs." J. Am. Chem. Soc. **114**(23): 9202-9203.

Paul, M. S. and B. L. Bass (1998). "Inosine exists in mRNA at tissue-specific levels and is most abundant in brain mRNA." Embo Journal **17**(4): 1120-1127.

Pauling, L. (1982). "Lewis,G.N. And the Nature of the Chemical-Bond." Abstracts of Papers of the American Chemical Society **183**(Mar): 66-Ched.

Petersheim, M. and D. H. Turner (1983). "Nuclear overhauser studies of CCGGAp, ACCGGp, and ACCGGUp." Biochemistry **22**(2): 264-268.

Polacek, N. and A. S. Mankin (2005). "The ribosomal peptidyl transferase center: structure, function, evolution, inhibition." Crit Rev Biochem Mol Biol **40**(5): 285-311.

Puglisi, J. D. and I. Tinoco, Jr. (1989). "Absorbance melting curves of RNA." Methods Enzymol **180**: 304-325.

Rabl, J., M. Leibundgut, S. F. Ataide, A. Haag and N. Ban (2011). "Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1." Science **331**(6018): 730-736.

Raether, H. (1988). "Surface-Plasmons on Smooth and Rough Surfaces and on Gratings." Springer Tracts in Modern Physics **111**: 1-133.

Ramakrishnan, V. (2002). "Ribosome Structure and the Mechanism of Translation." Cell **108**(4): 557-572.

Ratje, A. H., J. Loerke, A. Mikolajka, M. Brunner, P. W. Hildebrand, A. L. Starosta, A. Donhofer, S. R. Connell, P. Fucini, T. Mielke, P. C. Whitford, J. N. Onuchic, Y. Yu, K. Y. Sanbonmatsu, R. K. Hartmann, P. A. Penczek, D. N. Wilson and C. M. Spahn (2010). "Head swivel on the ribosome facilitates translocation by means of intra-subunit tRNA hybrid sites." Nature **468**(7324): 713-716.

Raychaudhuri, S., J. Conrad, B. G. Hall and J. Ofengand (1998). "A pseudouridine synthase required for the formation of two universally conserved pseudouridines in ribosomal RNA is essential for normal growth of Escherichia coli." RNA **4**(11): 1407-1417.

Razga, F., J. Koca, J. Sponer and N. B. Leontis (2005). "Hinge-like motions in RNA kink-turns: the role of the second a-minor motif and nominally unpaired bases." Biophys J **88**(5): 3466-3485.

Rentzeperis, D., J. Ho and L. A. Marky (1993). "Contribution of loops and nicks to the formation of DNA dumbbells: melting behavior and ligand binding." Biochemistry **32**(10): 2564-2572.

Riazance, J. H., W. A. Baase, W. C. Johnson, Jr., K. Hall, P. Cruz and I. Tinoco, Jr. (1985). "Evidence for Z-form RNA by vacuum UV circular dichroism." Nucleic Acids Res **13**(13): 4983-4989.

Rife, J. P. and P. B. Moore (1998). "The structure of a methylated tetraloop in 16S ribosomal RNA." Structure **6**(6): 747-756.

Rodnina, M. V., A. Savelsbergh, V. I. Katunin and W. Wintermeyer (1997). "Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome." Nature **385**(6611): 37-41.

Ryden-Aulin, M., Z. Shaoping, P. Kylsten and L. A. Isaksson (1993). "Ribosome activity and modification of 16S RNA are influenced by deletion of ribosomal protein S20." Mol Microbiol **7**(6): 983-992.

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

Sakakibara, Y. and C. S. Chow (2011). "Probing conformational states of modified helix 69 in 50S ribosomes." J Am Chem Soc **133**(22): 8396-8399.

Sakakibara, Y. and C. S. Chow (2012). "Role of Pseudouridine in Structural Rearrangements of Helix 69 During Bacterial Ribosome Assembly." ACS Chem Biol.

SantaLucia, J., Jr. and D. H. Turner (1997). "Measuring the thermodynamics of RNA secondary structure formation." Biopolymers **44**(3): 309-319.

Santalucia, J., R. Kierzek and D. H. Turner (1991). "Functional-Group Substitutions as Probes of Hydrogen-Bonding between Ga Mismatches in Rna Internal Loops." Journal of the American Chemical Society **113**(11): 4313-4322.

Saponara, A. G. and M. D. Enger (1974). "The isolation from ribonucleic acid of substituted uridines containing alpha-aminobutyrate moieties derived from methionine." Biochim Biophys Acta **349**(1): 61-77.

Saro, P. and J. J. SantaLucia (2012). RNA 1-2-3. Ann Arbor, DNA software

Wayne State University**:** Software for RNA, DNA, and protein sequence alignment, secondary and tertiary structure prediction and structure analysis tools.

Schirmer, R. E. and J. H. Noggle (1972). "Quantitative Application of Nuclear Overhauser Effect to Determination of Molecular Structure." Journal of the American Chemical Society **94**(9): 2947-&.

Schmeing, T. M., A. C. Seila, J. L. Hansen, B. Freeborn, J. K. Soukup, S. A. Scaringe, S. A. Strobel, P. B. Moore and T. A. Steitz (2002). "A pre-translocational intermediate in protein synthesis observed in crystals of enzymatically active 50S subunits." Nat Struct Biol **9**(3): 225-230.

Schmeing, T. M., R. M. Voorhees, A. C. Kelley, Y. G. Gao, F. V. t. Murphy, J. R. Weir and V. Ramakrishnan (2009). "The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA." Science **326**(5953): 688-694.

Schuwirth, B. S., M. A. Borovinskaya, C. W. Hau, W. Zhang, A. Vila-Sanjurjo, J. M. Holton and J. H. Cate (2005). "Structures of the bacterial ribosome at 3.5 A resolution." Science **310**(5749): 827-834.

Scolnick, E., R. Tompkins, T. Caskey and M. Nirenberg (1968). "Release factors differing in specificity for terminator codons." Proc Natl Acad Sci U S A **61**(2): 768-774.

Serdyuk, I. N., S. C. Agalarov, S. E. Sedelnikova, A. S. Spirin and R. P. May (1983). "Shape and compactness of the isolated ribosomal 16 S RNA and its complexes with ribosomal proteins." J Mol Biol **169**(2): 409-425.

Shajani, Z., M. T. Sykes and J. R. Williamson (2011). "Assembly of Bacterial Ribosomes." Annual Review of Biochemistry, Vol 80 **80**: 501-526.

Sharpe Elles, L. M., M. T. Sykes, J. R. Williamson and O. C. Uhlenbeck (2009). "A dominant negative mutant of the E. coli RNA helicase DbpA blocks assembly of the 50S ribosomal subunit." Nucleic Acids Res **37**(19): 6503-6514.

Sigel, R. K. O., D. G. Sashital, D. L. Abramovitz, A. G. Palmer, S. E. Butcher and A. M. Pyle (2004). "Solution structure of domain 5 of a group II intron ribozyme reveals a new RNA motif." Nat Struct Mol Biol **11**(2): 187-192.

Sijenyi, F. (2008). Structure and dynamics of E. coli 16S rRNA helix 23 and human 18S rRNA helix 24 by NMR spectroscopy. Chemistry. Detroit, Wayne State University. **Ph. D.**

Sijenyi, F., P. Saro, Z. Ouyang, K. Damm-Ganamet, M. Wood, J. Jiang and J. SantaLucia, Jr. (2011). The RNA Folding Problems: Different levels of RNA Structure Prediction. RNA 3D Structure Analysis and Prediction. N. B. Leontis and E. Westhof, Springer.

Singer, C. E. and G. R. Smith (1972). "Histidine Regulation in Salmonella-Typhimurium .13. Nucleotide Sequence of Histidine Transfer Ribonucleic-Acid." Journal of Biological Chemistry **247**(10): 2989-&.

Sklenar, V., H. Miyashiro, G. Zon, H. T. Miles and A. Bax (1986). "Assignment of the 31P and 1H resonances in oligonucleotides by two-dimensional NMR spectroscopy." FEBS Lett **208**(1): 94-98.

Smith, J. S. and E. P. Nikonowicz (1998). "NMR structure and dynamics of an RNA motif common to the spliceosome branch-point helix and the RNA-binding site for phage GA coat protein." Biochemistry **37**(39): 13486-13498.

Spahn, C. M., E. Jan, A. Mulder, R. A. Grassucci, P. Sarnow and J. Frank (2004). "Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor." Cell **118**(4): 465-475.

Sprinzl, M., C. Horn, M. Brown, A. Ioudovitch and S. Steinberg (1998). "Compilation of tRNA sequences and sequences of tRNA genes." Nucleic Acids Research **26**(1): 148-153.

Steitz, T. A. (2008). "A structural understanding of the dynamic ribosome machine." Nat Rev Mol Cell Biol **9**(3): 242-253.

Strickland, E. H. (1974). "Aromatic contributions to circular dichroism spectra of proteins." CRC Crit Rev Biochem **2**(1): 113-175.

Sumita, M., J. P. Desaulniers, Y. C. Chang, H. M. Chui, L. Clos, 2nd and C. S. Chow (2005). "Effects of nucleotide substitution and modification on the stability and structure of helix 69 from 28S rRNA." RNA **11**(9): 1420-1429.

Sumita, M., J. Jiang, J. Santalucia, Jr. and C. S. Chow (2012). "Comparison of solution conformations and stabilities of modified helix 69 rRNA analogs from bacteria and human." Biopolymers **97**(2): 94-106.

Sun, S. and E. R. Bernstein (1996). "Aromatic van der Waals clusters: Structure and nonrigidity." Journal of Physical Chemistry **100**(32): 13348-13366.

Sykes, M. T., Z. Shajani, E. Sperling, A. H. Beck and J. R. Williamson (2010). "Quantitative proteomic analysis of ribosome assembly and turnover in vivo." J Mol Biol **403**(3): 331-345.

Teichner, A., P. Londei and P. Cammarano (1986). "Intralineage Diversity of Archaebacterial Ribosomes - a Dichotomy of Ribosome Features Separates Sulfur-Dependent Archaebacteria and Methanococcaceae from Other Archaebacterial Taxa." Journal of Molecular Evolution **23**(4): 343-353.

Thompson, J., D. F. Kim, M. O'Connor, K. R. Lieberman, M. A. Bayfield, S. T. Gregory, R. Green, H. F. Noller and A. E. Dahlberg (2001). "Analysis of mutations at residues A2451 and G2447 of 23S rRNA in the peptidyltransferase active site of the 50S ribosomal subunit." Proc Natl Acad Sci U S A **98**(16): 9002-9007.

Tinoco, I., Jr. (1959). "The Optical Rotation of Oriented Helices. I. Electrical Orientation of Poly-r-Benzyl-L-Glutamate in Ethylene Dichloride." J. Am. Chem. Soc. **81**: 1540-1544.

Tjandra, N. and A. Bax (1997). "Measurement of dipolar contributions to 1JCH splittings from magnetic-field dependence of J modulation in two-dimensional NMR spectra." J Magn Reson **124**(2): 512-515.

Tobin, C., C. S. Mandava, M. Ehrenberg, D. I. Andersson and S. Sanyal (2010). "Ribosomes lacking protein S20 are defective in mRNA binding and subunit association." J Mol Biol **397**(3): 767-776.

Tomsic, J., L. A. Vitali, T. Daviter, A. Savelsbergh, R. Spurio, P. Striebeck, W. Wintermeyer, M. V. Rodnina and C. O. Gualerzi (2000). "Late events of translation initiation in bacteria: a kinetic analysis." EMBO J **19**(9): 2127-2136.

Trobro, S. and J. Aqvist (2008). "Role of ribosomal protein L27 in peptidyl transfer." Biochemistry **47**(17): 4898-4906.

Tsai, A. G., A. E. Engelhart, M. M. Hatmal, S. I. Houston, N. V. Hud, I. S. Haworth and M. R. Lieber (2009). "Conformational variants of duplex DNA correlated with cytosine-rich chromosomal fragile sites." Journal of Biological Chemistry 284(11): 7157-7164.

Tscherne, J. S., K. Nurse, P. Popienick, H. Michel, M. Sochacki and J. Ofengand (1999). "Purification, cloning, and characterization of the 16S RNA m5C967 methyltransferase from Escherichia coli." Biochemistry 38(6): 1884-1892.

Valadkhan, S. and J. L. Manley (2001). "Splicing-related catalysis by protein-free snRNAs." Nature 413(6857): 701-707.

Vallurupalli, P. and P. B. Moore (2003). "The solution structure of the loop E region of the 5S rRNA from spinach chloroplasts." J Mol Biol 325(5): 843-856.

van Dijk, M. and A. M. J. J. Bonvin (2009). "3D-DART: a DNA structure modelling server." Nucleic Acids Research 37(Web Server issue): W235-239.

Van Ryk, D. I. and S. Venkatesan (1999). "Real-time kinetics of HIV-1 Rev-Rev response element interactions. Definition of minimal binding sites on RNA and protein and stoichiometric analysis." Journal of Biological Chemistry 274(25): 17452-17463.

Varani, G., F. Aboul-ela and F. H. T. Allain (1996). "NMR investigation of RNA structure." Progress in Nuclear Magnetic Resonance Spectroscopy 29(1-2): 51-127.

Varani, G., C. Cheong and I. Tinoco, Jr. (1991). "Structure of an unusually stable RNA hairpin." Biochemistry 30(13): 3280-3289.

Varani, G. and W. H. McClain (2000). "The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems." EMBO Rep **1**(1): 18-23.

Varani, G. and I. Tinoco, Jr. (1991). "RNA structure and NMR spectroscopy." Q Rev Biophys **24**(4): 479-532.

Vicens, Q. and E. Westhof (2003). "Molecular recognition of aminoglycoside antibiotics by ribosomal RNA and resistance enzymes: an analysis of x-ray crystal structures." Biopolymers **70**(1): 42-57.

Videler, H., L. L. Ilag, A. R. McKay, C. L. Hanson and C. V. Robinson (2005). "Mass spectrometry of intact ribosomes." FEBS Lett **579**(4): 943-947.

Vuister, G. W., R. Boelens, A. Padilla and R. Kaptein (1991). "Statistical analysis of double NOE transfer pathways in proteins as measured in 3D NOE-NOE spectroscopy." J Biomol NMR **1**(4): 421-438.

Wallace, B. A. and R. W. Janes (2009). Modern techniques for circular dichroism and synchrotron radiation circular dichroism spectroscopy.

Wang, A. H., G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel and A. Rich (1979). "Molecular structure of a left-handed double helical DNA fragment at atomic resolution." Nature **282**(5740): 680-686.

Wang, B. and E. V. Anslyn (2011). Chemosensors : principles, strategies, and applications, John Wiley & Sons, Inc.

Wang, S., P. W. Huber, M. Cui, A. W. Czarnik and H. Y. Mei (1998). "Binding of neomycin to the TAR element of HIV-1 RNA induces dissociation of Tat protein by an allosteric mechanism." Biochemistry **37**(16): 5549-5557.

Warrant, R. W. and S. H. Kim (1978). "alpha-Helix-double helix interaction shown in the structure of a protamine-transfer RNA complex and a nucleoprotamine model." Nature **271**(5641): 130-135.

Watkins, N. E., Jr. and J. SantaLucia, Jr. (2005). "Nearest-neighbor thermodynamics of deoxyinosine pairs in DNA duplexes." Nucleic Acids Res **33**(19): 6258-6267.

Weixlbaumer, A., H. Jin, C. Neubauer, R. M. Voorhees, S. Petry, A. C. Kelley and V. Ramakrishnan (2008). "Insights into translational termination from the structure of RF2 bound to the ribosome." Science **322**(5903): 953-956.

Weixlbaumer, A., S. Petry, C. M. Dunham, M. Selmer, A. C. Kelley and V. Ramakrishnan (2007). "Crystal structure of the ribosome recycling factor bound to the ribosome." Nat Struct Mol Biol **14**(8): 733-737.

White, S. A., M. Nilges, A. Huang, A. T. Brunger and P. B. Moore (1992). "NMR analysis of helix I from the 5S RNA of Escherichia coli." Biochemistry **31**(6): 1610-1621.

Whitmore, L. and B. A. Wallace (2008). "Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases." Biopolymers **89**(5): 392-400.

Wijmenga, S. S., H. A. Heus, B. Werten, G. A. van der Marel, J. H. van Boom and C. W. Hilbers (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**(2): 134-141.

Williamson, J. R. (2005). "Assembly of the 30S ribosomal subunit." Q Rev Biophys **38**(4): 397-403.

Wohnert, J., A. J. Dingley, M. Stoldt, M. Gorlach, S. Grzesiek and L. R. Brown (1999). "Direct identification of NH...N hydrogen bonds in non-canonical base pairs of RNA by NMR spectroscopy." Nucleic Acids Res **27**(15): 3104-3110.

Wower, I. K., J. Wower and R. A. Zimmermann (1998). "Ribosomal protein L27 participates in both 50 S subunit assembly and the peptidyl transferase reaction." Journal of Biological Chemistry **273**(31): 19847-19852.

Wuthrich, K. (1986). NMR of Proteins and Nucleic Acids. New York, Wiley-Interscience.

Wyatt, J. R., M. Chastain and J. D. Puglisi (1991). "Synthesis and purification of large amounts of RNA oligonucleotides." Biotechniques **11**(6): 764-769.

Yusupov, M. M., G. Z. Yusupova, A. Baucom, K. Lieberman, T. N. Earnest, J. H. Cate and H. F. Noller (2001). "Crystal structure of the ribosome at 5.5 A resolution." Science **292**(5518): 883-896.

Zaher, H. S., J. J. Shaw, S. A. Strobel and R. Green (2011). "The 2'-OH group of the peptidyl-tRNA stabilizes an active conformation of the ribosomal PTC." EMBO J **30**(12): 2445-2453.

Zhang, W., J. A. Dunkle and J. H. Cate (2009). "Structures of the ribosome in intermediate states of ratcheting." Science **325**(5943): 1014-1017.

Zucker, F. H. and J. W. Hershey (1986). "Binding of Escherichia coli protein synthesis initiation factor IF1 to 30S ribosomal subunits measured by fluorescence polarization." Biochemistry **25**(12): 3682-3690.

ABSTRACT

**"FINE-TUNING" OF RIBOSOME STRUCTURE AND FUNCTIONS BY
PSEUDOURIDYLATION AND RNA-PROTEIN INTERACTIONS**

by

JUN JIANG

August 2012

Advisor: Prof. John SantaLucia Jr.

Major: Chemistry (Biochemistry)

Degree: Doctor of Philosophy

Ribosomal structure and functions appear to be "fine-tuned" by pseudouridylation and RNA-protein interactions. Pseudouridylation may promote base stacking interactions by mediating the base stacking between residues on both sides. In the RNA duplex region, this enhanced stacking interaction contributes to stabilization of duplex folding. In the loop region, enhanced stacking in one structural motif may destabilize the conformation of adjacent structural residues. This hypothesis is supported by UV-melting experiments, where pseudouridylation significantly stabilized H69 stem duplex folding, while destabilize the loop conformation. In addition, NMR also supports this hypothesis. The NMR structure of H69UUU shows that there are interruptions of base stacking between steps U1915 to A1916, and U1917 to A1918. The NMR structure and crystal structures of H69ΨΨΨ shows that Ψ1915 and Ψ1917 (*E. coli* numbering) mediate the extensive stacking. The dynamic property was also revealed in the NMR spectrum of H69ΨΨΨ (C1914), which suggests that the enhanced stacking interactions from Ψ1915 to C1924 may contribute to the

destabilization of the loop conformation of H69ΨΨΨ and pre-organization of the stem-loop structure of H69ΨΨΨ for ribosomal subunits association. RNA-protein interactions are another method to regulate ribosome biogenesis and activity. An ambient dissociation constant between the ribosomal RNA and ribosomal protein is required for the optimal ribosome biogenesis and activity (77nM for the *E. coli* cognate pair and 198 nM for the *P. aeruginosa* cognate pair). Either a too weak (dissociation constants of 77 nM vs. 312 nM for Ech9 – EcS20/PaS20) or a too strong interaction (dissociation constants of 198 nM vs. 51 nM for Pah9 – PaS20/EcS20) could affect the ability of S20 ribosomal protein to coordinate the correct folding of 16S rRNA.

AUTOBIOGRAPHICAL STATEMENT
JUN JIANG

EDUCATIONAL BACKGROUND

Doctor of Philosophy (Aug 2006 – Aug 2012): Advisor: Prof. John SantaLucia, Department of Chemistry, Wayne State University, Detroit, Michigan. Dissertation Title ""Fine-tuning" of ribosomal structure and functions by pseudouridylation and RNA-protein interactions".

Master of Science (Sep 2002 – Sep 2005): Advisor: Prof. Xiaoda Yang, Department of Chemical Biology, School of Pharmaceutical Sciences, Peking University, Beijing, P.R. China

Bachelor of Pharmacy (Sep 1998 – Jul 2002): Advisor: Prof. Xiangtao Liu School of Pharmaceutical Sciences, Peking University, Beijing, P.R. China.

PUBLICATIONS

1. **Jiang, J.**, Chow, C. and SantaLucia, J., Jr. "Effects of pseudouridylations on the solution structure of Helix 69 of 23S ribosomal RNA from *Escherichia coli*" (Manuscript in preparation).

2. Sumita, M, **Jiang, J.**, SantaLucia, J., Jr, Chow, C. (2012) "Comparison of solution conformations and stabilities of modified helix 69 rRNA analogues from bacteria and human." Biopolymer **97**(2): 94-106.

3. Sijenyi, F., Saro, P., Ouyang, Z., Damm-Ganamet, K., Wood, M., **Jiang, J.**, and SantaLucia, J., Jr. (2011) "The RNA Folding Problems: Different levels of RNA Structure Prediction", in RNA 3D Structure Analysis and Prediction. Leontis, N. and Westhof, E. (Eds.) Series Nucleic Acids and Molecular Biology, Springer.

4. **Jiang, J.**, Yang, X., and Wang, K. (2007) "Inhibition of cysteine protease papain by metal ions and polysulfide complexes, especially mercuric ion." (Journal of Chinese Pharmaceutical Sciences **2007** (16) 1-8.