COMBINING *IN SILICO* AND *IN VITRO* STUDIES ON TRNA DYNAMICS AND PSEUDOURIDINE FORMATION BY TRUB

JENNA FRIEDT
B.Sc. University of Lethbridge, 2011

A Thesis
Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfilment of the
Requirements for the Degree

MASTER OF SCIENCE

Department of Chemistry & Biochemistry
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

Abstract

RNA is an important molecule in the cell capable of carrying sequence information for protein production and acting as a functional component in essential cellular processes. In order to perform its many roles, RNA must often adopt specific three-dimensional structures that can be stabilized by the introduction of chemical modifications. Pseudouridine (Ψ) is the most abundant post-transcriptional RNA modification and is formed by enzymes known as pseudouridine synthases. These enzymes share a common core fold and active site structure, but the exact mechanism of pseudouridine formation is still undetermined. Here, a combination of *in silico* and *in vitro* studies were used to identify residues involved in catalysis to better understand the mechanism of pseudouridine formation. Molecular dynamics (MD) simulations of the *Escherichia coli* pseudouridine synthase TruB, responsible for the formation of the conserved Ψ55 in tRNA, were performed that revealed an interaction network of three active site residues. *In vitro* mutational studies showed that these residues are not required for substrate binding, but are highly important for catalysis. Additionally, MD simulations of unmodified tRNA were performed to establish a foundation for nucleic acid simulation by monitoring the effect of Mg^{2+} ions on the stability of tRNA *in silico*. These simulations will allow for a complex of tRNA and TruB to be simulated in order to complement the experimental studies and better understand the role of protein-RNA interactions. Together, computational and experimental work will assist in determining the mechanism of pseudouridine formation that may be common to all pseudouridine synthases.

Acknowledgements

I would like to start by thanking my supervisor, Dr. Ute Wieden-Kothe, for all the support and guidance she has provided me throughout my undergraduate and M.Sc. studies. Since the very first undergraduate course you taught me, you have been such an inspiration and motivational force in my life. Your enthusiasm and dedication to science and students has made an extremely significant impact on me. Thank you for challenging me in everyday work, encouraging my independence, and providing me with every opportunity to succeed in my academic career.

Thank you to my committee members, Dr. Marc Roussel and Dr. Steven Mosimann, for sharing your knowledge and advice. Your encouragement and support have been greatly appreciated. Thank you also to Dr. Juli Feigon for accepting to act as external examiner.

I have been fortunate enough to work with many extremely talented and intelligent people. The members of the Kothe, Wieden, and iGEM labs have made my years at the U of L more than memorable. I am so grateful for the constructive feedback, friendships, and necessary distractions you have provided. Thank you specifically to Fern Leavens, Laura Keffer-Wilkes, Raja Kamalampeta, Evan Caton, and Evan Mercier for their support and contributions to the work I have done.

Finally, I would like to thank my family and friends for supporting my academic investments. Thank you to my parents, Kent and Cheryl, for their love and encouragement of my ambitions, even when it may take me far away from them. I will always owe a part of any of my successes to your unending support.

Table of Contents

Abstract iii
Acknowledgements iv
Table of Contents v
List of Tables vii
List of Figures viii
List of Abbreviations ix
Chapter 1 – Introduction 1
  1.1 RNA: A Diverse and Versatile Biomolecule 1
  1.2 RNA Modifications 2
    1.2.1 Pseudouridine 2
    1.2.2 Pseudouridine in Noncoding RNAs 4
  1.3 Pseudouridine Synthases 6
    1.3.1 Catalytic Mechanism of Pseudouridine Formation 9
    1.3.2 TruB 12
  1.4 Molecular Dynamics Simulations 16
  1.5 tRNA and Mg^{2+} 19
  1.6 Molecular Dynamics Simulations of Nucleic Acids 23
Chapter 2 – The Role of Active Site Residues in Pseudouridylation by TruB 26
  2.1 Materials and Methods 26
    2.1.1 Buffers and Reagents 26
    2.1.2 Protein Expression and Purification 27
    2.1.3 [^{3}H]-Labeled tRNA Preparation 28
    2.1.4 Nitrocellulose Filtration Assays 30
    2.1.5 Tritium Release Assay 31
    2.1.6 2-Aminopurine tRNA Preparation 31
    2.1.7 Fluorescence Spectroscopy and Stopped-Flow Experiments 33
    2.1.8 TruB Molecular Dynamics Simulations 34
  2.2 Results 37
    2.2.1 Molecular Dynamics Simulations Reveal Active Site Interactions 37
    2.2.2 Role of D90 and R181 in tRNA Binding by TruB 44
    2.2.3 Catalytic Activity of TruB D90 and R181 Variants 46

2.2.4 Kinetic Analysis of TruB Variants' Interaction with tRNA 50
2.2.5 Molecular Dynamics Simulations of TruB D90N 53
2.3 Discussion 55
2.3.1 TruB Wild Type MD Simulations 56
2.3.2 Binding and Catalysis by TruB D90 and R181 Variants 59
2.3.3 Pre-Steady-State Kinetics of TruB Variants 61
2.3.4 TruB D90N MD Simulations 63
2.3.5 The Role of R181 and D90 in Pseudouridine Formation by TruB 64
Chapter 3 – Effect of Mg^{2+} Ions on tRNA Dynamics *In Silico* 67
3.1 Methods 68
3.1.1 Generating tRNA Models 68
3.1.2 tRNA Molecular Dynamics Simulations 70
3.1.3 MD Simulation Analysis 71
3.1.4 Cluster Analysis of tRNA MD Simulations 72
3.2 Results 73
3.2.1 General Stability of tRNA Simulations 73
3.2.2 The Role of Mg^{2+} in tRNA Simulations 79
3.2.3 High Occupancy Mg^{2+} Sites 83
3.2.4 Cluster Analysis of tRNA Simulations 87
3.3 Discussion 92
3.3.1 tRNA Simulations are Remarkably Stable 93
3.3.2 Mg^{2+} Is Not Required for Maintaining Overall tRNA L-Shaped Structure 94
3.3.3 Non-Specific High Occupancy Mg^{2+} Sites 96
3.3.4 Mg^{2+} Effects on Local Structure Properties 97
3.3.5 Application to Other RNA Simulations 99
3.3.6 Simulation of tRNA-Protein Complexes 100
Chapter 4 – Conclusions and Future Directions 102
4.1 TruB Catalytic Mechanism 102
4.2 tRNA Dynamics 107
References 113
Appendix 126

List of Tables

Table 2.1 Thermocycler conditions for PCR amplification of tRNA^{Phe} gene from pCF0 plasmid. 29
Table 2.2 Primers used for generating DNA template of 5’-half tRNA 32
Table 2.3 Thermocycler conditions for amplification of template DNA of 5’-half tRNA 32
Table 2.4 Dissociation constants for TruB-tRNA interaction determined from nitrocellulose filtration 46
Table 2.5 Apparent rates of single-turnover pseudouridine formation by TruB variants. 49
Table 2.6 Pre-steady-state kinetic rates for 2AP-tRNA interacting with TruB 52
Table 3.1 tRNA crystal structures used to find potential Mg^{2+} binding sites 69
Table 3.2 tRNA models for simulation. 70
Table 3.3 Number of simulation frames in each cluster based on Mg^{2+} occupancy at a reference position observed in the X-18Mg^{2+} simulation. 73

List of Figures

Figure 1.1 Isomerization of uridine to pseudouridine 3
Figure 1.2 Secondary and tertiary structure of tRNA^{Phe}. 6
Figure 1.3 Proposed mechanisms for pseudouridine formation 11
Figure 2.1 Global fluctuations of TruB in MD simulations... 39
Figure 2.2 Conserved TruB active site residues 40
Figure 2.3 Active site residue interactions during MD simulations 42
Figure 2.4 Binding of tRNA by TruB variants 45
Figure 2.5 Pseudouridine formation by TruB variants under multiple-turnover conditions 47
Figure 2.6 Pseudouridine formation by TruB variants under single-turnover conditions 48
Figure 2.7 tRNA interaction with TruB observed by 2AP fluorescence 51
Figure 2.8 Global fluctuations of TruB D90N MD simulations...................................... 53
Figure 2.9 Active site residue interactions during TruB D90N MD simulations 55
Figure 3.1 tRNA elbow angle and torsion angle.. 72
Figure 3.2 Cluster analysis reference Mg^{2+} position 73
Figure 3.3 Initial placement of Mg^{2+} for tRNA simulations 75
Figure 3.4 RMSD and RMSF of tRNA simulations 78
Figure 3.5 Number of Watson-Crick base pairs during tRNA simulations 80
Figure 3.6 Stability of the G18-U55 tertiary interaction.. 82
Figure 3.7 Inner and outer sphere Mg^{2+} ions 84
Figure 3.8 High occupancy Mg^{2+} sites 86
Figure 3.9 High occupancy Mg^{2+} sites of occupied and unoccupied simulation clusters 89
Figure 3.10 tRNA elbow angle in the occupied and unoccupied clusters 91
Figure A.1 Acceptor arm and anticodon arm contribution to global flexibility 126
Figure A.2 D-arm and T-arm contribution to global flexibility...................................... 127
Figure A.3 Stability of the G15-C48 tertiary interaction 128
Figure A.4 Stability of the G19-C56 tertiary interaction 129
Figure A.5 Number of Watson-Crick base pairs in the occupied and unoccupied clusters 130
Figure A.6 Stability of G18-U55 base pair in the occupied and unoccupied clusters 131
Figure A.7 Stability of G15-C48 base pair in the occupied and unoccupied clusters 132
Figure A.8 Stability of G19-C56 base pair in the occupied and unoccupied clusters 133
Figure A.9 Mobility of Mg^{2+} ions during simulation 134

List of Abbreviations

Ψ	Pseudouridine
2AP	2-aminopurine
ATP	Adenosine triphosphate
CHARMM	Chemistry at Harvard Molecular Mechanics
CTP	Cytidine triphosphate
D	Aspartate
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleotide triphosphates
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
GMP	Guanosine monophosphate
GTP	Guanosine triphosphate
iPPase	Inorganic pyrophosphatase
IPTG	β-D-1-thiogalactopyranoside
LB	Luria-Bertani broth
MD	Molecular dynamics
miRNA	Micro RNA
mRNA	Messenger RNA
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDBID	Protein data bank identification code
PMSF	Phenylmethylsulfonylfluoride
PTC	Peptidyltransferase center
R	Arginine
RMSD	Root-mean-square deviation
RMSF	Root-mean-square fluctuation
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
SDS	Sodium dodecyl sulfate
siRNA	Small interfering RNA
snoRNP	Small nucleolar ribonucleoprotein
snRNP	Small nuclear ribonucleoprotein
sRNP	Small ribonucleoprotein
tRNA	Transfer RNA
U	Uridine
UTP	Uridine triphosphate
UV	Ultraviolet
VMD	Visual Molecular Dynamics software

Chapter 1 – Introduction

1.1 RNA: A Diverse and Versatile Biomolecule

Ribonucleic acids (RNA) are one of the major molecules of life and play important roles in cellular processes such as gene regulation, splicing, telomere maintenance, and protein synthesis. RNA is unique in that it can carry sequence information as deoxyribonucleic acid (DNA) does, but can also fold into complex three-dimensional structures that allow it to catalyze cellular reactions similar to proteins. There are a number of types of coding and non-coding RNAs found in the cell, each with a specific function. Protein-coding genes are transcribed from DNA into messenger RNA (mRNA) that is translated into a protein by the ribosome. Ribosomal RNA (rRNA) makes up a majority of the ribosome and constitutes the catalytic component necessary for protein synthesis (Nissen et al., 2000). Also involved in protein synthesis is transfer RNA (tRNA), which is a highly modified and structured RNA charged with amino acids to be incorporated into a growing polypeptide chain. Besides these RNAs involved in protein synthesis, there are also small nuclear RNAs (snRNAs) in eukaryotes that are a major component of the spliceosome complex required for removing introns from pre-mRNAs. In archaea and eukaryotes, small nucleolar RNAs (snoRNAs) function in complexes with proteins to guide the chemical modification of a variety of other RNA molecules in the cell. Additionally in many eukaryotes, micro RNAs (miRNAs) and small interfering RNAs (siRNAs) are important for gene regulation and silencing, and have potential applications in gene therapy (Pereira and Lopes-Cendes, 2012; Wang et al., 2011). In order for RNA to efficiently contribute to cellular processes, it often has to adopt a highly defined three-dimensional structure. To promote the stability of the native folded state, chemical modifications to the RNA can be made to strengthen intramolecular interactions and stabilize local secondary and tertiary structures.

1.2 RNA Modifications

With only four canonical nucleotides (adenosine, guanosine, cytidine, and uridine) to generate diversity in RNA, mechanisms have evolved to increase the complexity of these molecules by expanding the number of potential nucleotides. This is done through post-transcriptional modifications, of which there are over 100 currently found in RNA (Dunin-Horkawicz et al., 2006). The most common of these are methylation and pseudouridylation, which are found in RNAs in all three domains of life (Jackman and Alfonzo, 2013). One of the main functions of modified bases is to stabilize RNA structure, though modifications are also important for avoiding misfolding, enhanced recognition by cellular targets, and efficient participation in metabolic processes (Grosjean, 2005).

1.2.1 Pseudouridine

In the 1950s, the first modified RNA nucleotide, pseudouridine, was discovered (Davis and Allen, 1957). Pseudouridine is the 5-ribosyl isomer of uridine and is found in noncoding RNAs such as tRNA, rRNA, snRNA, and snoRNA (Charette and Gray, 2000). The N-C glycosidic bond between the base and the sugar molecule in uridine is converted to a C-C glycosidic bond in pseudouridine (Fig. 1.1). Since the Watson-Crick face of pseudouridine is unchanged from uridine, it is still able to base-pair with adenine. The C-C glycosidic bond in pseudouridine, unique in RNA or DNA, creates an extra hydrogen bond donor at N1 of the base, which allows for the formation of novel interactions within the RNA. For instance, in the crystal structure of tRNA^{Gln} bound to glutaminyl-tRNA synthetase and ATP, a water molecule was found coordinated between the N1-H of pseudouridine and the backbone phosphate of the preceding nucleotide, which may contribute to the stabilizing effect of pseudouridine (Arnez and Steitz, 1994). Pseudouridine has also been found to promote the 3'-endo sugar conformation in A-RNA helices, improving base-pairing and increasing the base stacking of nearby nucleotides (Davis, 1995; Ge and Yu, 2013). Together these properties result in a more rigid backbone of pseudouridine-containing RNA with enhanced local stability.

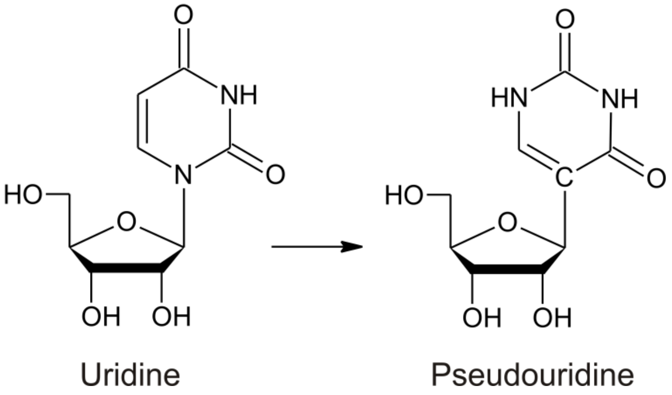

Figure 1.1 Isomerization of uridine to pseudouridine. Uridine is converted to pseudouridine through the breakage of the N-C glycosidic bond, rotation of the base, and reattachment of the base to the ribose through a C-C glycosidic bond.

The role of pseudouridine is still unclear, but changes to pseudouridylation patterns have interesting implications in certain diseases and stress responses. For example, elevated levels of oxidized pseudouridine have been found in the urine of some Alzheimer's and Parkinson's disease patients (Lee et al., 2007). As well, the bone marrow failure disease, dyskeratosis congenita, has been associated – among others – with mutations in protein-coding genes for the H/ACA small nucleolar ribonucleoprotein (snoRNP) complex responsible for pseudouridine formation in humans, though the exact cause of this disease is not yet clear (Grozdanov et al., 2009; Heiss et al., 1998).

Pseudouridylation can also be regulated as additional pseudouridines in U2 snRNA are formed in yeast under heat stress or nutrient deprivation (Wu et al., 2011). Though there are currently no known pseudouridines found in mRNA, artificial H/ACA guide RNAs have been used to successfully introduce pseudouridines into mRNA in yeast and *Xenopus* (Chen et al., 2010; Karijolich and Yu, 2011). Substituting uridine in stop codons for pseudouridine can cause the read-through of mRNA transcripts in both bacteria and eukaryotes (Fernandez et al., 2013; Karijolich and Yu, 2011). Additionally, pseudouridines found in the anticodon of tRNA can give rise to alternative codon usage (Tomita et al., 1999). This raises interesting questions such as whether there are intrinsic pseudouridine sites in mRNA that are in too low abundance to be detected, and if pseudouridines in other RNA molecules can be used for regulation of protein expression.

1.2.2 Pseudouridine in Noncoding RNAs

RNA modifications are commonly found near functionally important regions, and pseudouridine is no exception. In snRNA, pseudouridines are found in the 5’-end region of U1, the loop region of U5, and near the branch site recognition sequence of U2 where it is important for stabilizing the branch site and bulged adenosine (Newby and Greenbaum, 2001; Yu et al., 2011). Pseudouridines in rRNA help to improve folding, stability, and assembly of the ribosome, and are often positioned far from protein binding sites, but close to the functional centers of the ribosome (Cunningham et al., 1991; Ofengand and Fournier, 1998). For example, a number of pseudouridines are clustered in the decoding center, the A-site finger region just above the ribosomal A-site, the peptidyltransferase center (PTC), and helix 69 where they help to form an inter-subunit bridge between the large and small ribosomal subunits. The removal of single pseudouridines from rRNA shows little to no growth phenotype, but for example removing all pseudouridines from the PTC causes defects in rRNA processing, cell growth, and translation (King et al., 2003).

tRNA is one of the most highly modified RNA molecules in the cell (Fig. 1.2) (Lane et al., 1995). In bacteria, there are seven potential sites of pseudouridylation in tRNA, including the universally conserved Ψ55 in the T-arm of all elongator tRNAs (Sprinzl et al., 1998; Zamir et al., 1965). While pseudouridines do not seem to influence the overall L-shaped structure of tRNA, the presence of pseudouridines can stabilize local areas within the molecule (Byrne et al., 2010; Jackman and Alfonzo, 2013; Shi and Moore, 2000). For example, pseudouridines found at the base of the D- or anticodon-stem provide a strong stem-closing base pair, and pseudouridine in the anticodon loop can assist in binding of tRNA to the ribosome and can therefore affect translation fidelity and reading frame maintenance (Lecointe et al., 2002). Urbonavičius and coworkers showed that the loss of pseudouridines and other modifications in tRNA increased the frequency of frameshifting, potentially due to increased susceptibility to slippage of the tRNA in the decoding center or from poor recruitment of aminoacyl-tRNA to the ribosomal A-site (Urbonavičius et al., 2001). Although some of the pseudouridines in tRNA are highly conserved, many of them are not required for cell viability (Gutgsell et al., 2000; Gutgsell et al., 2001; Kinghorn et al., 2002; Raychaudhuri et al., 1999). The presence of Ψ55 is not required for cell growth under normal conditions, but has been shown to enhance viability at high growth temperatures in *Escherichia coli*, which indicates that Ψ55 increases tRNA thermal stability (Kinghorn et al., 2002). Additionally, the absence of Ψ55 causes the upregulation of the tRNA modifications Gm18, m^5s^2U54 and m^1A58 in the hyperthermophile *Thermus thermophilus*, suggesting that this modification may help to regulate the modification of other sites in tRNA (Ishida et al., 2011). These types of studies help to shed light on potential cellular roles of this abundant modification, though a full understanding of pseudouridine's cellular function(s) is still lacking.

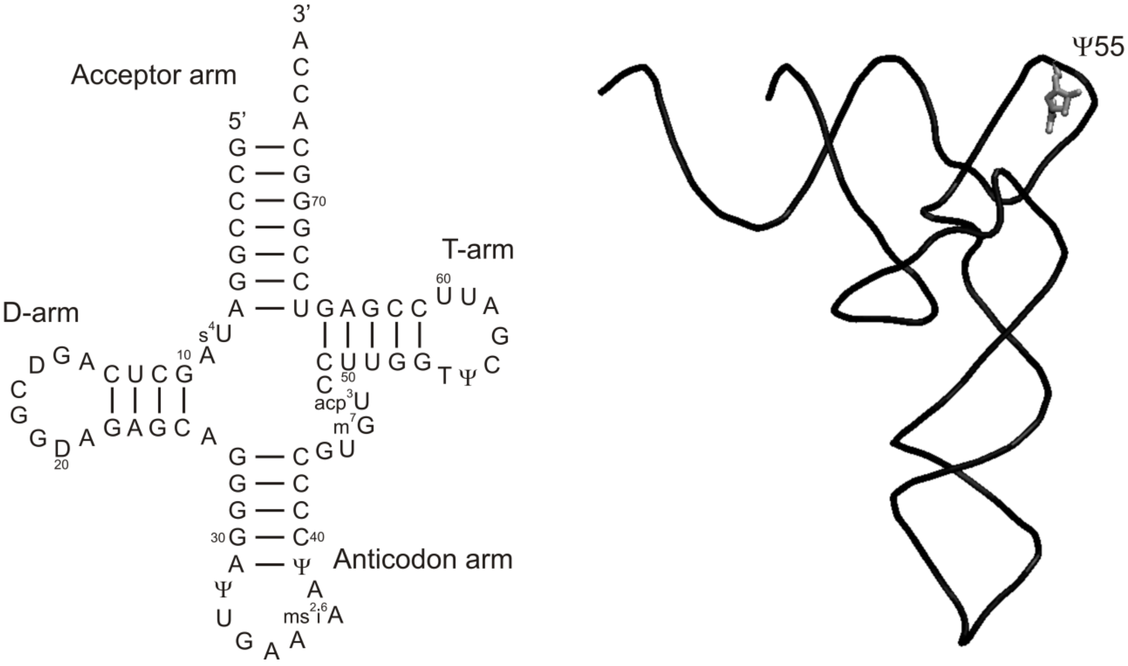

Figure 1.2 Secondary and tertiary structure of tRNA^{Phe}. Predicted secondary structure of tRNA^{Phe} from *E. coli* (left), including the modified nucleotides 4-thiouridine (s^4U) at position 8, dihydrouridine (D) at positions 16 and 20, Ψ at positions 32, 39 and 55, 2-methylthio-*N*^{6}-isopentenyladenosine (ms^2i^6A) at position 37, 7-methylguanosine (m^7G) at position 46, 3-(3-amino-3-carboxypropyl) uridine (acp^3U) at position 47 and 5-methyluridine (or ribothymidine, T) at position 54. Three-dimensional structure of tRNA^{Phe} from *S. cerevisiae* (PDBID 1EHZ), including the stick-representation of Ψ55 in the T-arm, is shown on the right.

1.3 Pseudouridine Synthases

Rather than being incorporated by the transcriptional machinery like the four canonical nucleotides, pseudouridines are formed as a post- or possibly co-transcriptional modification through the isomerization of uridines that are already present in the RNA (Ciampi et al., 1977; Johnson and Söll, 1970). The enzymes responsible for this isomerization reaction are called pseudouridine synthases. They are classified into one of six families based on sequence similarity: RluA, RsuA, TruA, TruB, TruD, and Pus10 (Hamma and Ferré-D'Amaré, 2006; Koonin, 1996; McCleverty et al., 2007). The families are named for their representative *E. coli* protein, except for Pus10 which is not found in bacteria. Bacteria use stand-alone enzymes that typically modify one or a few uridine targets and do not require additional cofactors or input energy. In addition to stand-alone enzymes, archaea and eukaryotes utilize H/ACA sRNPs to form pseudouridines. These complexes are composed of four protein subunits: Cbf5 (the catalytic component, dyskerin in humans), Nop10, Gar1, and L7Ae (Nhp2 in humans). H/ACA sRNPs utilize an H/ACA guide RNA that base-pairs with the substrate RNA to properly position the target uridine in the active site of Cbf5 (Wu and Feigon, 2007). Since the guide RNA specifies the substrate RNA, a number of pseudouridines can be formed by H/ACA sRNPs through the use of different guide RNAs bound to the active complex.

Crystal structures of enzymes from all six families of pseudouridine synthases have been solved and revealed that they share a common catalytic core fold and active-site cleft despite very low sequence identity (Hamma and Ferré-D'Amaré, 2006; McCleverty et al., 2007). Because of the diversity in the number and type of modified RNAs, pseudouridine synthases often recognize their substrates using unique inserts, N- or C-terminal extensions, or peripheral domains. For instance, TruB and RluA have a small thumb loop insert that undergoes a large conformational change upon RNA binding (Hoang et al., 2006; Pan et al., 2003). TruD also contains a thumb region, however it is a large domain that moves like a hinge toward the active site when RNA is bound (Hoang and Ferré-D'Amaré, 2004). RluA has a forefinger loop that works with the thumb loop to pinch the loop of its RNA substrate (Hoang et al., 2006). Forefinger loops are also found in enzymes of the RsuA and TruA families. TruA is the only known pseudouridine synthase that appears to function as a dimer (Foster et al., 2000). Enzymes in the TruB family contain a C-terminal PUA domain, named for its presence in some *p*seudo*u*ridine synthases and *a*rchaeosine-transglycosylases. It is thought that this domain is important for recognizing the 3'-CCA end of tRNA by TruB or the 3'-ACA of the H/ACA guide RNA by Cbf5 (Li and Ye, 2006; Zhou et al., 2011). Pus10 is characterized by having thumb and forefinger loops as well as a large N-terminal THUMP domain named after its presence in 4-thiouridine synthases, RNA methyltransferases and pseudouridine synthases (McCleverty et al., 2007).

The only absolutely conserved residue among the six families of pseudouridine synthases is an aspartate found in the active site, which has been shown to be essential for catalysis (Behm-Ansmant et al., 2003; Chan and Huang, 2009; Conrad et al., 1999; Del Campo et al., 2001; Huang et al., 1998; Kaya and Ofengand, 2003; Ramamurthy et al., 1999; Raychaudhuri et al., 1999; Zebarjadian et al., 1999). Site-directed mutagenesis of this residue abolishes catalysis, but does not affect substrate recognition or binding (Ramamurthy et al., 1999; Wright et al., 2011). In addition to this residue, pseudouridine synthases contain a basic residue, either arginine or lysine, in the active site that is capable of forming a salt bridge interaction with the catalytic aspartate. As well, an aromatic residue in the active site is found to pack tightly to the ribose of the target uridine and likely stabilizes the conformation of the base within the active site (Phannachet et al., 2005). All pseudouridine synthases use a tyrosine for this role, except for TruD where this residue is a phenylalanine. The tyrosine has been proposed to act as a general base to abstract a proton from C5 to complete the isomerization process (Phannachet et al., 2005). Although these are the only residues common to all six families, other residues are conserved within each individual family. For example, RluA, RsuA, TruA, and Pus10 employ an additional arginine residue in close proximity to the catalytic aspartate that moves to occupy the gap in the RNA helix left when the target uridine gets flipped out into the active site of the enzyme (Hoang et al., 2006; Kamalampeta et al., 2013). A similar role is fulfilled by a conserved histidine in the active site of enzymes from the TruB family (Duan et al., 2009; Hoang and Ferré-D’Amaré, 2001). Despite the variety of RNA substrates modified by these enzymes, the conservation of these active site residues across all pseudouridine synthase families has led to the suggestion of a shared mechanism for pseudouridine formation.

1.3.1 Catalytic Mechanism of Pseudouridine Formation

Because of the common catalytic domain and active site aspartate residue found in all six families, it is thought that pseudouridine synthases also share a universal catalytic mechanism (Mueller and Ferré-D’Amaré, 2009). This idea is also supported by a uniformly slow catalytic step seen for the pseudouridine synthase families RluA, TruA, and TruB (Wright et al., 2011). However, the actual mechanism of pseudouridine formation is still undetermined. The minimal process for the formation of pseudouridine would be breaking of the N-C glycosidic bond, flipping or rotation of the uracil base, and reformation of a C-C glycosidic bond. To do this, the conserved aspartate residue is proposed to act as a nucleophile to attack the uridine, thus initiating the first step in pseudouridine formation.

Originally, two mechanisms for pseudouridine formation were put forward, where the aspartate attacked either C6 of the uracil base or C1’ of the ribose ring (Fig 1.3 A-B) (Gu et al., 1999; Huang et al., 1998; Kammen et al., 1988). Attacking at C6 would result in the formation of a Michael adduct, where the uracil base would be attached to the aspartate via an ester bond. This would provide an axis for rotation to position the base for reattachment at C5, followed by the leaving of aspartate and deprotonation of C5 to give the final product. Alternatively, attack on the ribose at C1’ would cause the release of the uracil base and generation of an acylal intermediate. The base would then be free to flip or rotate into position for C-C bond formation using aspartate as a leaving group, with deprotonation of C5 giving the final pseudouridine product. Attack on the ribose sugar at C1’ is similar to the mechanism employed by DNA glycosylases, which have the same first step of catalysis to perform, i.e. breaking of the N-glycosidic bond (Labahn et al., 1996). For many years, the formation of a covalent adduct by Michael addition was an attractive idea, since pseudouridine synthases were not found to release the uracil base into the bulk solvent during catalysis (Arena et al., 1978). However, extensive NMR studies by the Mueller group using a 5-fluorouridine-substituted RNA substrate and TruB have produced strong evidence against the Michael addition mechanism (Miracco and Mueller, 2011). They showed the formation of two isomeric products that differed in stereochemistry at C2’, which is unlikely to have been the result of attack on C6. Instead, they proposed a modified mechanism for acylal formation as well as an alternative mechanism where the catalytic aspartate abstracts a proton from C2’ of the sugar, releasing the base and forming a glycal intermediate that will again require deprotonation of C5 after C-C bond formation (Fig. 1.3C).

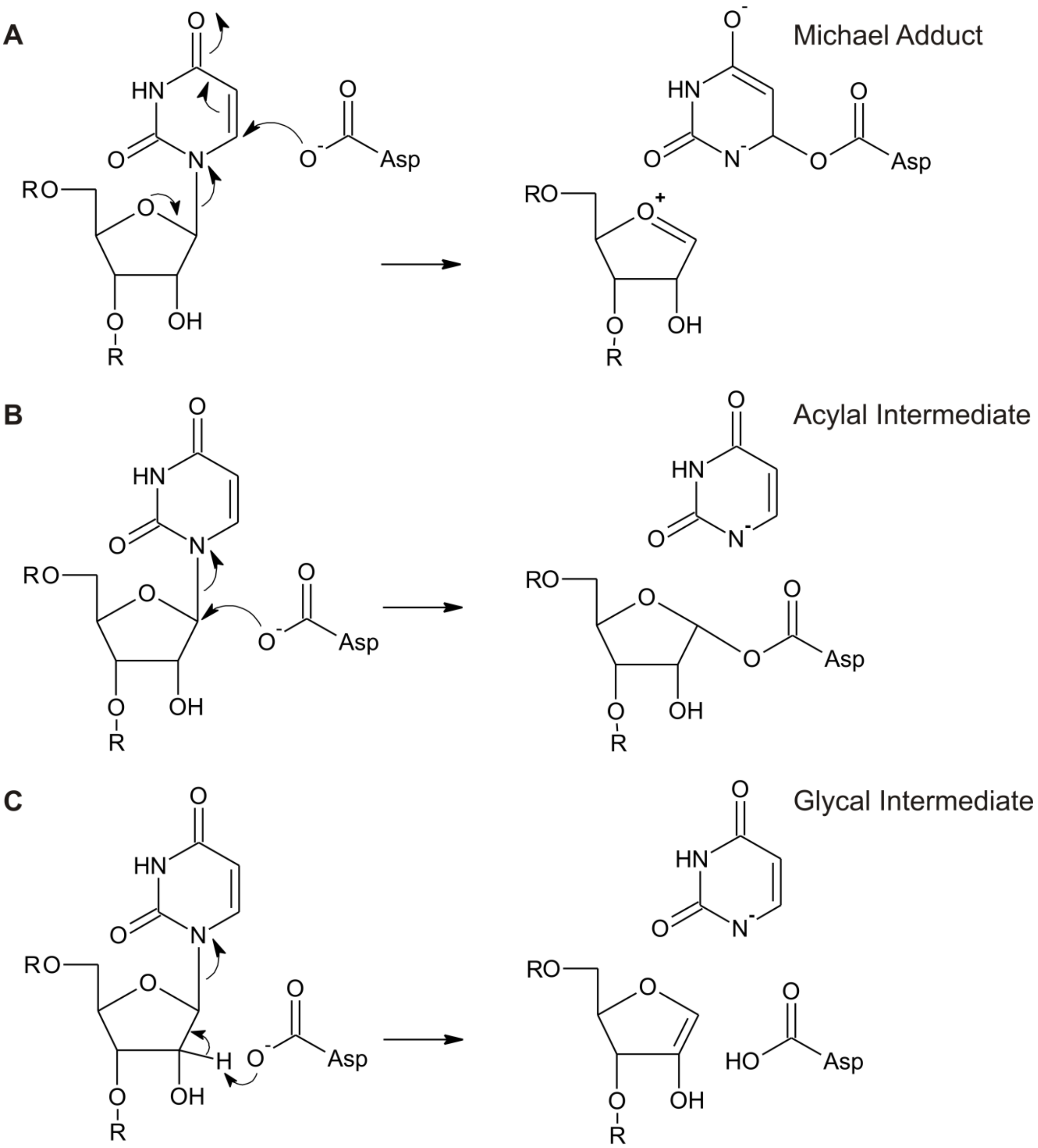

Figure 1.3 Proposed mechanisms for pseudouridine formation. (A) Attack by the catalytic aspartate on C6 of the base will result in a Michael adduct where the uracil base is covalently linked to the aspartate. (B) Formation of an acylal intermediate will occur if the aspartate attacks at position C1’ of the ribose. (C) Recently proposed mechanism of glycal intermediate formation following attack by the catalytic aspartate on C2’ of the ribose. (Figure adapted from Gu et al., 1999 and Miracco and Mueller, 2011.)

Determining the role of other conserved active site residues may identify additional residues that work together with the catalytic aspartate and contribute to pseudouridine formation. The catalytic role of the basic residue in pseudouridine synthases may be to keep the aspartate deprotonated so as to make it more nucleophilic, or it may also serve as a source of protons (Hamma and Ferré-D'Amaré, 2006). As mentioned above, the conserved aromatic residue could act as a general base, but this mechanism would have to be modified for the TruD family that has a phenylalanine at this position (Phannachet et al., 2005). Further studies still need to be performed to determine which mechanism is most likely to occur and the exact role of each of these residues in catalysis. Additionally, it will be important to show that the same C2’-isomeric products are formed by the other pseudouridine synthase families, thereby providing further evidence for whether or not a universal mechanism is used.

1.3.2 TruB

TruB was first identified and characterized in 1995 as the bacterial pseudouridine synthase responsible for the universally conserved Ψ55 in the T-arm of tRNA (Nurse et al., 1995). Since then, a number of biochemical and crystallographic studies have been performed to gain more information about this enzyme. TruB is composed of a large catalytic domain that resembles that of the other pseudouridine synthases families, and a small globular C-terminal PUA domain (Hoang and Ferré-D’Amaré, 2001). The catalytic domain contains the conserved active site residues D48, R181, and Y76. R181 can form a solvent-inaccessible salt bridge with the catalytic D48 residue, which likely helps to alleviate the energetically disfavored buried charges while keeping D48 activated in a deprotonated state (Hoang and Ferré-D’Amaré, 2001). Besides the PUA domain, TruB enzymes that pseudouridylate tRNAs, as opposed to those involved in H/ACA sRNPs, also contain two unique inserts (Hoang and Ferré-D'Amaré, 2001). Insert 1 is a small loop at the top of the active site cleft and insert 2, also called the thumb loop, is a highly conserved and mobile element involved in tRNA binding. Cbf5, a homolog of TruB and part of the H/ACA sRNP machinery, also contains a thumb loop with a similar sequence to TruB, but it is shorter in length (Duan et al., 2009). Though the apo and bound conformations of TruB are strikingly similar throughout the catalytic domain, there are noticeable differences in the conformation of insert 1, the thumb loop, and the PUA domain upon RNA binding (Pan et al., 2003). Non-specific interactions between insert 1 and the thumb loop are broken such that insert 1 can move closer to the RNA and the thumb loop can undergo a large conformational change to pinch down on the RNA in the active site (Phannachet and Huang, 2004). The PUA domain also moves as a rigid body by approximately 6 Å in the direction of the newly bound tRNA in order to form non-specific interactions with the backbone of the acceptor stem (Hoang et al., 2005; Pan et al., 2003). These regions form some of the interactions important for substrate recognition and specificity.

The first crystal structure of a pseudouridine synthase in complex with RNA was that of TruB bound to a truncated tRNA substrate consisting of just the T-arm nucleotides (Hoang and Ferré-D'Amaré, 2001). This and other crystallographic studies have helped reveal the mechanism of RNA binding to TruB, which is an interesting problem since TruB must bind and recognize all elongator tRNAs in the cell. It was found that the minimal substrate for TruB must contain the canonical stem-loop structure of the T-arm, as well as the invariant nucleotides U54, U55, C56, and A58 (tRNA numbering) (Gu et al., 1998; Hoang and Ferré-D’Amaré, 2001; Pan et al., 2003; Phannachet and Huang, 2004). The binding pocket of TruB is decorated with positively-charged residues and is highly complementary to the stem-loop shape of the tRNA T-arm, which retains a similar shape after binding to TruB as in the context of full-length tRNA. Upon RNA binding, insert 1, the thumb loop, and the PUA domain move towards the RNA and make direct interactions with it. This suggests a two-step binding process, where the tRNA rigidly docks into the binding cleft through electrostatic interactions and surface complementarity, followed by induced fit of protein elements to maximize interactions with the RNA (Pan et al., 2003; Phannachet and Huang, 2004). Since the U55 target is buried within the tertiary structure of tRNA and makes long-distance contacts with G18 in the D-loop, TruB disrupts this interaction and flips U55 and the two subsequent bases out of the helical stack and into the active site (Hoang and Ferré-D’Amaré, 2001). This mechanism is similarly employed by RluA, RsuA, and TruA, as well as the methyltransferase TrmA responsible for the modification at U54 of tRNA (Alian et al., 2008; Hamma and Ferré-D'Amaré, 2006). In TruB, these flipped-out nucleotides are stabilized by the insertion of a conserved histidine residue (H43) below the reverse Hoogsteen base pair formed between U54 and A58 (Hoang and Ferré-D’Amaré, 2001; Pan et al., 2003). One of the flipped-out nucleotides, C56, is specifically recognized by forming a hydrogen bond with a conserved aspartate in insert 1 (D90) and through a water-mediated hydrogen bond with R141 in the thumb loop (Phannachet and Huang, 2004). This nucleotide is very important for substrate recognition, as mutations to any other nucleotide have been shown to abolish TruB activity (Becker et al., 1997; Gu et al., 1998). Upon its large conformational change, the thumb loop forms direct contacts with the RNA, pinching the major groove of the RNA loop and anchoring the RNA in the active site (Pan et al., 2003). U55 gets positioned within the active site by stacking between Y76 and Y179 (Hoang and Ferré-D'Amaré, 2001; Pan et al., 2003). Catalysis is further facilitated by the proper orientation of D48 through a salt bridge interaction with R181 that points D48 toward U55 and closes up the active site. After pseudouridine formation, small perturbations in the active site are thought to propagate to the thumb loop, potentially through transient interactions with R181 or residues in insert 1, causing the thumb loop to release the RNA to allow for product dissociation (Hoang et al., 2005; Phannachet and Huang, 2004).

Biochemical studies on TruB have analyzed the role of certain residues in pseudouridine formation. For example, when the catalytic aspartate is substituted with alanine, cysteine, or asparagine, there is no detectable activity despite similar ability to bind tRNA (Hoang et al., 2005; Ramamurthy et al., 1999). This and other studies that mutated the analogous aspartate residue in the remaining pseudouridine synthase families have led to the consensus that this is the catalytic residue in these enzymes. Site-directed mutagenesis of the conserved H43 showed that only a phenylalanine substitution at this position retained activity similar to wild-type, while glutamine, asparagine, alanine, or glycine substitutions greatly impaired catalysis (Hamilton et al., 2005). This study confirmed the importance of this residue in stabilizing the flipped-out conformation of nucleotides 55-57 in the active site. It has been suggested that the role of the conserved aromatic residue, Y76 in TruB, is to play a part in the final proton abstraction from C5 of the base (Phannachet et al., 2005). Interestingly, no biochemical studies have been performed to test the role of the third active site residue common to all families, that of the basic residue (either arginine or lysine). Kinetic studies have been performed with TruB, revealing a two-step binding process with an overall rate of 6 s^{-1} and a rate-limiting, single-round rate of pseudouridine formation of 0.5 s^{-1} (Wright et al., 2011). This relatively slow rate of catalysis may be due to evolutionary selection or the complexity of the isomerization preventing any faster catalysis.

The exact role of TruB in the cell has been under debate for the last 10 years since an interesting finding by the Ofengand lab. This group as well as the Stansfield lab found that TruB knockout strains of *E. coli* showed no growth phenotype on their own, but could not compete with wild type *E. coli* when grown in co-culture (Gutgsell et al., 2000; Kinghorn et al., 2002). Interestingly, the cells could be rescued by being supplemented with an inactive TruB gene (TruB D48C) that is incapable of forming Ψ55. This suggests that TruB may have more than one role in the cell, and that pseudouridine formation may not be its primary function. Since TruB must disrupt the tertiary structure of its tRNA substrate in order for catalysis to occur, it could be that TruB acts as an RNA chaperone by assisting in the folding of tRNA molecules, though this hypothesis has yet to be tested.

1.4 Molecular Dynamics Simulations

Crystallographic studies are useful for examining interactions within biological molecules by providing static snapshots of a protein or nucleic acid. Using crystallographic data as a starting point, molecular dynamics (MD) simulations can be used to generate and analyze series of snapshots representative of general motions that occur within the molecule. MD simulations are essentially atomistic hypothetical single-molecule experiments that examine the thermal fluctuations of a particular biomolecule in an aqueous solution (Šponer et al., 2012). In molecular mechanics simulations, biomolecules are represented as simple ball-and-spring models with bond forces described using Newtonian physics and potential-energy functions (Adcock and McCammon, 2006). Due to advances in computational power and optimized simulation theory, MD simulations can generate information on conformational changes, ligand docking, protein folding, communication networks, and kinetic and thermodynamic properties on time scales near biological relevance (nanoseconds to microseconds). MD protocols are well established for protein systems, and as such have been used on a number of different systems. One of the more recent and impressive examples of MD simulations was performed on an all-atoms model of the 70S ribosome for 1.3 μs, indicating how MD simulations can provide information on conformational freedom and sampling even in larger macromolecules or macromolecular complexes (Whitford et al., 2013).

The first MD simulation of a protein was done on the bovine pancreatic trypsin inhibitor in 1976, lasting for 9.2 ps (McCammon, 1976). This simplistic model, only 500 atoms simulated in a vacuum, was the stepping stone for development of more complex simulation parameters and MD theory allowing for longer simulation times (up to microseconds) and inclusion of explicit solvent and ions. A critical component to ensure validity and stability of an MD simulation is choosing the appropriate potential-energy function and force field parameters. Parameters for simulations have been obtained from experimental or quantum mechanical studies of small molecules or fragments, and are assumed to reliably represent these fragments in the context of a larger molecule (Adcock and McCammon, 2006). Force fields typically have parameters describing equilibrium bond lengths, bond angles, dihedral angles, improper dihedral angles, and nonbonded interactions for biological building blocks. A number of different force fields have been developed for proteins and subsequently optimized to describe lipids (CHARMM22), nucleic acids (BMS force field, AMBER ff99), and carbohydrates (CHARMM CSFF) (Cornell et al., 1995; Feller and MacKerell, 2000; Kuttel et al., 2002; Langley, 1998; MacKerell et al., 1998; Reiling et al., 1996). One limiting feature of these force fields in molecular mechanics is that covalent bond breakage or formation is not permitted, and thus can introduce artificial strain in certain molecules (Šponer et al., 2012).

MD simulations are highly modular and can be optimized for particular molecular systems and to answer specific questions. They involve iterative calculations of the instantaneous forces in the system and track the resultant movements of the atoms over a specified time step (Adcock and McCammon, 2006). The simulations can be performed with implicit or explicit solvent and ion molecules, and *in vitro* experimental conditions can be closely mimicked by defining pressure, temperature, volume, or energy limits. However, since MD simulations are generated based on empirical models and generalized force field parameters, the accuracy is limited by the quality of the initial model, force fields, simulation constraints, and the length of the simulation. For this reason, MD simulations are most informative when combined with and validated by experimental data.

As computing power continues to grow and simulation parameters become more refined, there is an expected increase in the utility of MD simulations for studying biomolecules (Adcock and McCammon, 2006). Not only can large macromolecules like the ribosome be reliably simulated for long time scales, but MD simulations can also be used for ligand docking studies, monitoring conformational changes and transient protein sub-states, and rational design of inhibitors or antibiotics (ElHefnawi et al., 2012). These types of studies are currently limited mostly by simulation lengths, since these types of motions can occur on time scales of microseconds to several minutes (Adcock and McCammon, 2006; Šponer et al., 2012). Already, an increasing number of simulations are being performed on a variety of protein-RNA complexes, such as an miRNA-mRNA-AGO2-FXR1 complex thought to control translation upregulation, the HIV-1 trans-activator protein Tat with its RNA response element and potential inhibitory peptides, and non-canonical rRNA elements interacting with ribosomal protein L1 (Krepl et al., 2013; Li et al., 2013; Ye et al., 2012). With the growing capabilities and knowledge of the limitations of MD simulations, this computational technique can be used to gain unique insights into structural, dynamical, and functional characteristics of biomolecular systems.

1.5 tRNA and Mg^{2+}

As mentioned above, tRNA is a highly structured biomolecule involved in important cellular processes. It is best known for its adaptor role in protein synthesis, where it delivers amino acids to the growing polypeptide chain by decoding the mRNA through correct codon-anticodon recognition in the ribosome. The *E. coli* K12 genome contains 86 tRNA genes that represent 45 different tRNA species (Withers et al., 2006). These include at least one tRNA for each of the 20 standard amino acids found in proteins, as well as other tRNA molecules like the specific initiator tRNA ($tRNA^{fMet}$) that carries a formylated methionine (fMet) used to initiate translation in bacteria. Most tRNAs consist of approximately 75 nucleotides with a small variable loop region (class I), while some others contain a larger variable region that can form an additional hairpin (class II) (Giegé et al., 2012; Sprinzl et al., 1998). Besides cytosolic tRNAs, eukaryotes often contain an additional set of mitochondrial tRNAs that share some sequence features with cytosolic tRNAs but are often less modified and may have odd structural features such as missing D- or T-arms (Ernsting et al., 2009; Yuan et al., 2010). Since all tRNAs must be accommodated in the ribosome and recognized by other translational initiation and elongation machinery (e.g. EF-Tu), this narrowly defines the structural diversity of these molecules. However, there are many other interaction partners of tRNA that in some cases work on a very particular tRNA (e.g. aminoacyl-tRNA synthetases), so sufficient structural and sequence diversity must be present for this specific recognition. In bacteria, tRNAs are typically expressed in a long transcript that is enzymatically processed at the 5’ and 3’ ends by endonucleases such as RNase P, specifically modified by a variety of modification enzymes, and aminoacylated by the respective aminoacyl-tRNA synthetase. Some eukaryotic tRNAs contain introns that must be removed by splicing endonucleases followed by ligation with a tRNA ligase (Hopper et al., 2010). Eukaryotic tRNAs must also interact with export proteins (e.g. Exportin-1, RanGTP) that transport the processed tRNAs to the cytoplasm and in some cases to the mitochondria (Duchêne et al., 2009). Turnover and rapid tRNA decay pathway enzymes also interact with tRNAs to recycle damaged tRNA or as a stress response to nutrient deprivation and oxidative stress (Lee and Collins, 2005; Thompson et al., 2008). tRNAs have also been found to participate in more obscure roles, such as cell wall and amino acid biosynthesis (Sheppard et al., 2008; Stewart et al., 1971). This variety of cellular partners makes tRNA a unique and diverse molecule that can be fine-tuned to facilitate its participation in essential processes.

The secondary structure of tRNA has a cloverleaf shape, with three stem-loops and one stem. These are named the D-arm, anticodon arm, T-arm, and acceptor stem, based on modifications or functional properties. When folded into the native tertiary structure, the four stems are coaxially-stacked into an L-shape that is stabilized through numerous long-range interactions (Holbrook et al., 1977; Marck and Grosjean, 2002). Many of these interactions form between D- and T-loop nucleotides in the elbow region of the tRNA. The L-shaped structure tightly compacts the negatively-charged phosphate backbone in what would generally be thought of as an unfavourable conformation. Because of this, an extensive amount of research has analyzed the stability of the three-dimensional tRNA structure and what stabilizing effect, if any, can be achieved with cations.

Unlike most proteins, RNA very often requires the presence of mono- or divalent cations or polycations such as polyamines to stabilize its native conformation (Serebrov et al., 2001). Mg^{2+} and other metal ions can also influence the folding process of RNAs and facilitate the function of catalytic ribozymes (Fedor, 2002; Misra and Draper, 2002). A number of properties of RNA can influence the interactions it has with Mg^{2+}, including chain flexibility, phosphate positioning, and charge density (Bowman et al., 2012). There are four generally accepted classes of ions associated with nucleic acids: inner sphere (or chelated/coordinated), outer sphere (or glassy), diffuse (or condensed), and free (Bowman et al., 2012; Hayes et al., 2012). These classes are based on properties such as degree of coordination, rates of diffusion, and proximity to the nucleic acid. Inner sphere ions are those bound directly to RNA, usually through shedding of part of their hydration shell, and have limited diffusion rates. Outer sphere ions are closely coordinated with RNA and often interact through a single hydration layer. In MD simulations of the SAM-1 riboswitch, it was found that a majority of Mg^{2+} ions were found in the outer sphere and that these ions may be as influential as inner sphere ions on RNA structure (Hayes et al., 2012). Diffuse ions have multiple hydration shells and higher rates of diffusion, while free ions are found in the bulk solution, are not coordinated by the RNA, and have the fastest diffusion rates.

The first solved crystal structure of a tRNA molecule indicated the binding of four Mg^{2+} ions within the tertiary structure (Holbrook et al., 1977; Quigley et al., 1978). Early studies on tRNA folding using fluorescent probes revealed that Mg^{2+}, spermidine, and temperature all affected tRNA conformations, with high concentrations of Mg^{2+} and/or spermidine stabilizing a slow-tumbling conformation likely equivalent to that seen in the crystal structure (Ehrenberg et al., 1979; Nilsson et al., 1983). NMR studies on unmodified tRNA transcripts have showed that the presence of 5 mM Mg^{2+} can assist the tRNA in adopting the native fold by compensating for the lack of stabilizing interactions usually provided by the various nucleotide modifications (Hall et al., 1989). Circular dichroism, UV, and fluorescence techniques have also been used to identify the presence of at least two strong Mg^{2+} binding sites in modified and unmodified tRNA that were filled at Mg^{2+} concentrations below 0.1 mM, in addition to up to 30 non-specific, weak binding sites (Serebrov et al., 1998). The binding of Mg^{2+} has also been seen to influence the conformation of tRNA, based on fluorescence studies showing a series of transitions as Mg^{2+} concentration was increased, corresponding to high affinity Mg^{2+} binding sites (<20 µM) and lower affinity binding sites (2-20 mM) (Serebrov et al., 2001). However, the number of binding sites can also be affected by the overall concentration of ions in solution, with low ionic strength yielding 4-6 strong metal ion binding sites and higher monovalent cation concentrations reducing this number to only one or two sites (Leroy and Gueron, 1977; Romer and Hach, 1975; Stein and Crothers, 1976).

While crystal structures can provide direct information about potential metal ion binding sites, the position of those ions could be affected by crystal packing or may only be artefacts of the crystallization process. A basic search of tRNA crystal structures alone shows variable positioning of Mg^{2+} ions, with some structures having no metal ions identified within the three-dimensional structure at all (Basavappa and Sigler, 1991). Therefore, despite the previously cited work on Mg^{2+} binding sites in tRNA, it is still not entirely clear where the metal ion binding sites are, how many there are, or what type of ion best binds in that position. In the last 15 years, MD simulations of RNA and DNA have been able to predict and in some cases propose new metal ion binding sites (Auffinger and Westhof, 2000; Cheatham and Kollman, 1997; Young et al., 1997). Computational methods used in combination with experimental work should be useful in learning more about the specificity and function of certain metal ion binding sites with respect to nucleic acid structure.

1.6 Molecular Dynamics Simulations of Nucleic Acids

Despite the major advances in protein MD simulations mentioned above, the simulation of nucleic acids has lagged behind due in part to the highly charged nature of nucleic acids and the long range electrostatic interactions that can occur between and within nucleic acid strands (Norberg and Nilsson, 2002). It was not until the 1990s when simulation protocols had been optimized for appropriate treatment of long range electrostatics that simulations of nucleic acids could be reliably performed (Šponer et al., 2012). Much research has been invested in accurately defining the force field parameters that best represent the structure and covalent connectivity of nucleic acids to remain accurate over long simulation times (e.g. 100s of nanoseconds) (Cheatham and Kollman, 2000).

Two of the most commonly used force fields in MD simulations are the AMBER and CHARMM force fields (Šponer et al., 2012). AMBER is based on the second generation force field developed by Cornell et al., and optimization of these parameters have better defined the glycosidic bond angles and relieved some of the backbone tension from earlier versions (Cheatham et al., 1999; Cornell et al., 1995). Additionally, further parameters have been released for the AMBER force field that have been optimized specifically for RNA simulations (Yildirim et al., 2010). The CHARMM force field has been extensively tested on B-DNA, but not yet on RNA or non-canonical DNA structures (Ditzler et al., 2010). Recent changes to this force field include revision of the dihedral angle parameters of the 2’-hydroxyl group to better simulate A-RNA systems (Denning et al., 2011). Studies done on RNA-only simulations of a small riboswitch and an A-RNA duplex showed that variations of the AMBER force field produced simulations that maintained the original shape of the RNA helix better than those that had used the CHARMM force field (Banáš et al., 2012; Beššeová et al., 2012). However, both AMBER and CHARMM force fields have been found to perform well with protein and protein-nucleic acid systems (Ditzler et al., 2010).

While solvation and counter-ions are often taken into consideration to mimic *in vivo* conditions for biomolecules *in silico*, these considerations are especially important when simulating nucleic acids due to their highly charged phosphate backbone and potential for complex tertiary folds (Auffinger et al., 1999). In particular, placement of water and cations such as Mg^{2+} are important for electrostatic shielding and preventing unfolding during simulations (Alexander et al., 2010; Auffinger et al., 1999). Simulations of the bacterial 5S rRNA loop E showed that monovalent ions were also important for RNA stability and that Mg^{2+} ions placed within the RNA remained close to where they were found in crystal structures (Auffinger et al., 2004). As mentioned above (Section 1.5), crystallographic data and *in vitro* folding studies on tRNA have been used to identify 4-6 potential Mg^{2+} binding sites that likely contribute to promoting and stabilizing the native structure (Quigley et al., 1978; Serebrov et al., 2001). Due to this stabilizing effect, the concentration of ions used in simulation can affect RNA structure, where too low of a concentration will not be sufficient for maintaining tertiary interactions but too high of a concentration can perturb the native structure (Eargle et al., 2008). In addition, these considerations may change when simulating free RNA or RNA bound to protein, as well as modified versus unmodified RNA (Eargle et al., 2008). There is more to be learned about the stability of RNA with respect to modification and ion coordination, and improved treatment of RNA and cations in MD simulations combined with experimental techniques should help to clarify the importance of RNA-Mg^{2+} interactions.

Chapter 2 – The Role of Active Site Residues in Pseudouridylation by TruB

Biochemical studies have shown that an absolutely conserved aspartate residue is essential for catalysis by pseudouridine synthases (Behm-Ansmant et al., 2003; Conrad et al., 1999; Del Campo et al., 2001; Huang et al., 1998; Kaya and Ofengand, 2003; Ramamurthy et al., 1999; Raychaudhuri et al., 1999; Zebarjadian et al., 1999). However, the actual mechanism of pseudouridine formation and the roles of some other conserved active site residues have not been determined. Here, I employed a strategy combining computational with biochemical investigations to describe the function of two more residues in the active site of the *E. coli* pseudouridine synthase TruB. First, *in silico* MD simulations were used to analyze the intramolecular dynamics of TruB and identify important interactions occurring in the active site that may contribute to catalysis and that were not visible in crystal structures. Three residues were found to closely interact through a majority of the simulations: D48, R181, and D90. Crystallographic studies have previously shown that the catalytic residue D48 interacts with R181. However, here an additional residue, D90 was found to also interact with R181. This triad of residues was further analyzed biochemically for their potential to contribute to the function of TruB. Their functional roles were characterized through the use of mutational studies, binding and catalysis assays, rapid-kinetic stopped-flow experiments, and MD simulation analysis of a TruB variant.

2.1 Materials and Methods

2.1.1 Buffers and Reagents

TAKEM$_4$ Buffer: 50 mM Tris-HCl pH 7.5, 70 mM NH_4Cl, 30 mM KCl, 1 mM EDTA, 4 mM $MgCl_2$. [C5-^{3}H]-UTP for *in vitro* transcriptions was purchased from

Moravek. T4 RNA ligase was purchased from New England Biolabs. All other enzymes were purchased from Fermentas. 2-aminopurine-labeled 3'-half tRNA was purchased from Dharmacon.

2.1.2 Protein Expression and Purification

The TruB gene in the pET28a-EcTruB plasmid (Wright et al., 2011) was mutated using QuikChange® site-directed mutagenesis (Strategene) to encode the amino acid substitutions D90E, D90N, D90A, R181K, R181M, and R181A. TruB wild type and variants were expressed in *E. coli* BL21(DE3) cells in the presence of 50 µg/mL kanamycin. Cultures were grown in LB media to an optical density at 600 nm (OD_{600}) of 0.6, after which protein expression was induced with 1 mM β-D-1-thiogalactopyranoside (IPTG). The equivalent of one OD_{600} of cells was collected for SDS-PAGE analysis at 30 min, 1, 2, and 3 hours after induction to monitor protein expression. After 3 h, the cells were harvested by centrifugation at 5000 × *g* for 15 min at 4°C and flash frozen with liquid nitrogen.

All TruB variants were purified as previously described (Wright et al., 2011) using affinity and size-exclusion chromatography. Briefly, frozen cell pellets were thawed and resuspended in 5 mL/g of cells of Buffer A (20 mM Tris-HCl pH 8.1, 400 mM KCl, 5% glycerol (v/v), 1 mM β-mercaptoethanol, 0.5 mM phenylmethylsulfonylfluoride (PMSF), 30 mM imidazole). Lysozyme (1 mg/mL), sodium deoxycholate (12.5 mg/mL), and sonication were used to lyse the cells, followed by centrifugation at 30 000 × *g* for 30 min at 4°C to pellet the cell debris. The cleared lysate was added to Ni^{2+}-sepharose (Fast Flow, GE Healthcare) in Buffer A and incubated for 60 min to allow for binding of the histidine-tagged protein to the resin. The resin was washed 6 times with Buffer A before eluting the protein with Buffer B (20 mM Tris-HCl pH 8.1, 400 mM KCl, 5% (v/v) glycerol, 1 mM β-mercaptoethanol, 500 mM imidazole). Samples of each wash and elution step were analyzed by SDS-PAGE. Elution samples containing the protein were pooled and concentrated at 4500 × *g* and 4°C in a Vivaspin ultrafiltration device (molecular weight cut-off of 10 000 Da).

The TruB variants were further purified and re-buffered using Superdex 75 (XK26/100, GE Healthcare) size exclusion chromatography. After applying the concentrated protein sample to the column, it was eluted in Buffer C (20 mM HEPES-KOH pH 7.5, 150 mM KCl, 1mM β-mercaptoethanol, 0.5 mM EDTA, 5 mM $MgCl_2$, 20% glycerol (v/v)) while monitoring the absorbance at 280 nm. Peak fractions were analyzed by SDS-PAGE, and fractions containing the protein were pooled and concentrated as before. Final protein concentration was determined using A_{280} (molar extinction coefficient of 20 985 M^{-1} cm^{-1}, calculated using ProtParam (Gill and von Hippel, 1989) and by SDS-PAGE comparison with a TruB preparation of known concentration. A_{260} and urea-PAGE analysis showed that protein preparations were not detectably contaminated with nucleic acid, and final protein purity was greater than 95% as judged by SDS-PAGE.

2.1.3 [^{3}H]-Labeled tRNA Preparation

Template DNA for *E. coli* $tRNA^{Phe}$ was amplified with a polymerase chain reaction (PCR) from the pCF0 plasmid containing the $tRNA^{Phe}$ gene (Sampson et al., 1989) using the primers T7 promoter sense (5’-GCTGCAGTAATACGACTCACTATAG-3’) and $EcotRNA^{Phe}$ antisense (5’-mUmGGTGCCCGGACTCG-3’). The reaction components included 1X *Pfu* buffer with

$MgSO_4$, 0.4 mM dNTPs, 0.4 µM of forward and reverse primer, 0.02 U/µL *Pfu* DNA polymerase, and 50 ng pCF0 plasmid. A PCR thermocycler was used to perform the reaction, with cycle conditions shown in Table 2.1. PCR products were analyzed by 1% agarose gel electrophoresis to verify product size.

Table 2.1 Thermocycler conditions for PCR amplification of tRNA^{Phe} gene from pCF0 plasmid.

Step	Temperature (°C)	Time	No. of cycles
1	95	5 min	1
2	95	30 s	6
	45 (1°C increase/cycle)	30 s	
	72	20 s	
3	95	30 s	29
	50	30 s	
	72	20 s	
4	72	11 min	1

The PCR-amplified template was used for *in vitro* transcription of tRNA^{Phe}. Reaction components included 1X transcription buffer (40 mM Tris-HCl pH 7.5, 15 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl); 10 mM DTT; 3 mM each ATP, GTP, and CTP; 96 µM UTP; 4 µM [C5-^{3}H]-UTP; 5 mM GMP; 0.01 U/µL iPPase; 0.3 µM T7 RNA polymerase; 0.12 U/µL RNase inhibitor; and 10% DNA template (v/v). The reaction was incubated at 37°C for 4 h, after which DNase (0.002 U/µL) was added to the reaction followed by incubation for an additional hour at 37°C. The reaction was then quenched and diluted to a final KCl concentration of 0.2 M by adding R0’ buffer (100 mM Tris-acetate pH 6.3, 10 mM $MgCl_2$, 15% ethanol) and R3’ buffer (100 mM Tris-acetate pH 6.3, 10 mM $MgCl_2$, 15% ethanol, 1150 mM KCl).

A Nucleobond PC100 column (Macherey & Nagel) was used to purify the [^{3}H]-tRNA. The column was equilibrated with R0' buffer before applying the tRNA *in vitro* transcription solution. The column was washed twice with R1'a buffer (100 mM Tris-acetate pH 6.3, 10 mM $MgCl_2$, 15% ethanol, 300 mM KCl) followed by elution of the tRNA with R3' buffer. The elution was precipitated with 0.75 volumes of cold isopropanol at -20°C. The tRNA was pelleted by centrifugation at 4500 × *g* for 45 min at 4°C and the pellet was washed with 70% ethanol before dissolving in H_2O. Scintillation counting and absorbance measurements at 260 nm (molar extinction coefficient of 5×10^5 M^{-1} cm^{-1} (Peterson and Uhlenbeck, 1992)) were used to quantify the tRNA concentration and specific activity.

2.1.4 Nitrocellulose Filtration Assays

[^{3}H]-tRNA was refolded in 1X $TAKEM_4$ by incubation at 65°C for 5 min followed by cooling to room temperature for 10 min. tRNA (6.25 µL, 100 nM final concentration) was incubated with increasing concentrations of protein (18.75 µL, 0-30 µM final concentration) in 1X $TAKEM_4$. After 10 min incubation at room temperature, the reaction mixture was transferred to a nitrocellulose membrane (0.2 µm, Whatman) pre-soaked in $TAKEM_4$ buffer and then filtered by vacuum. The membrane was immediately washed with 1 mL cold 1X $TAKEM_4$ and dissolved in scintillation cocktail (EcoLite(+), MP Biomedicals) for 30 min. Scintillation counting was used to determine the amount of radioactivity on the membrane, corresponding to the amount of tRNA bound to protein. The amount of tRNA bound was plotted against the protein concentration and was fit with the following hyperbolic function to calculate the dissociation constant (K_D) for tRNA binding:

$$Y = B_{max} \times [\text{protein}] / (K_D + [\text{protein}])$$

2.1.5 Tritium Release Assay

[^{3}H]-tRNA was refolded as for nitrocellulose filtration assays. TruB was diluted in 1X TAKEM$_4$ and pre-warmed at 37°C before adding the folded tRNA and incubating at 37°C. Samples were taken from the reaction at various time points, and quenched by adding to 1 mL of 5% Norit A charcoal (w/v) in 0.1 M HCl. After incubation for 15-20 min, the solution was centrifuged at 13 000 × *g* for 2 min, and the cleared supernatant was transferred to 0.5 mL of 5% Norit A charcoal (w/v) in 0.1 M HCl. The solution was centrifuged again before filtering 1 mL of the supernatant through glass wool to remove any residual charcoal and collecting the filtrate for scintillation counting. Thereby, the amount of [^{3}H] was determined that was released during the conversion of uridine to pseudouridine by TruB. Single-turnover experiments were analyzed by 1-exponential fitting to determine an apparent rate of catalysis (k_{app}), using the equation:

$$Y = Y_\infty + A \times \exp(-k_{app} \times t)$$

2.1.6 2-Aminopurine tRNA Preparation

2-aminopurine (2AP)-tRNA was prepared by annealing two halves of the tRNA and ligating to give full-length tRNA. The following 3’-half was purchased from Dharmacon and contained a 2AP label at position 57 (full-length numbering):

5’-AAAUCCCCGUGUCCUUGGUUCG-2AP-UUCCGAGUCCGCGCACCA-3’

The 5’-half was generated by assembly and amplification of two DNA primers (sequences found in Table 2.2) followed by *in vitro* transcription. To assemble the primers, a reaction mixture of 1X *Pfu* buffer with MgSO$_4$, 200 µM dNTPs, 0.5 µM 5’-half sense primer, 0.5 µM 5’-half antisense primer, and 0.02 U/µL *Pfu* DNA polymerase was heated to 94°C for 1 min for denaturation, 44°C for 1 min to anneal the primers to each other, followed by primer extension at 72°C for 30 s for a total of 20 cycles. The product of this assembly was then amplified using the same reaction mixture, but using the T7 promoter and 5’-half antisense methyl primers (Table 2.2) in place of the 5’-half sense and 5’-half antisense primers, respectively, and adding 5% assembly product (v/v). The cycling conditions for the amplification are listed in Table 2.3.

Table 2.2 Primers used for generating DNA template of 5’-half tRNA. Underlined sequence indicates region of overlap between 5’-half sense and antisense primers.

Primer Name	Sequence
5’-half sense	5’-GCGTAATACGACTCACTATAGCGCGGATAGCTCAGTCG-3’
5’-half antisense	5’-TCAATCCCCTGCTCTACCGACTGAGCTATCCG-3’
T7 promoter	5’-GCGTAATACGACTCACTATAG-3’
5’-half antisense methyl	5’-mUmCAATCCCCTGCTCTACCGAC-3’

Table 2.3 Thermocycler conditions for amplification of template DNA of 5’-half tRNA.

Step	Temperature (°C)	Time	No. of cycles
1	94	4 min	1
2	94	1 min	10
	48	2 min	
	72	45 s	

The amplified 5’-half tRNA template was then used for *in vitro* transcription using the same procedure as outlined for [^{3}H]-tRNA but using 3 mM non-radioactive UTP in place of the mixture of non-radioactive and [^{3}H]-UTP. The *in vitro* transcription was incubated at 37°C overnight, and DNase was added as before to digest the template. The RNA was purified using the BioRad Model 491 Prep Cell continuous electrophoresis purification system. 6X RNA loading dye (8 M urea, 20 mM EDTA, 0.025% bromophenol blue, 1X TBE (90 mM Tris base pH 8.3, 90 mM boric acid, 1 mM EDTA)) was added to the transcribed RNA to a 1X final concentration followed by heating to 100°C for 10 min. The RNA was loaded on a 5 cm 12% urea-PAGE (28 mm diameter) and a constant voltage of 250 V was applied while pumping with 1X TBE at a flow rate of 0.75 mL/min to capture the RNA eluting from the bottom of the gel. The absorbance at 260 nm was monitored and 1.2 mL fractions were collected. Peak fractions were analyzed by 12% urea-PAGE, and fractions containing 5'-half tRNA were pooled and precipitated with 0.05 µg/µL glycogen, 0.3 M sodium acetate, and 3 volumes of cold ethanol at -20°C.

To generate full-length 2AP-tRNA, 15 µM 5'-half tRNA was annealed in 1.5-fold excess to 2AP-3'-half tRNA in 1X T4 RNA ligase buffer (New England Biolabs) by heating at 95°C for 3 min, 65°C for 10 min, and then slowly cooling from 65°C to 20°C over 45 min (1°C decrease per min). T4 RNA ligase (0.02 U/mL) was then added to the mixture followed by incubation at 37°C for 4 h. Ligated tRNA was purified from unligated tRNA halves by separation on 12% urea-PAGE, followed by UV-shadowing and gel excision. The tRNA was eluted from the gel using an EluTrap electroelution system (Whatman) and ethanol precipitated to concentrate the final sample. Absorbance measurements at 260 nm were used to determine the concentration of ligated 2AP-tRNA (molar extinction coefficient of 5×10^{5} M^{-1} cm^{-1}).

2.1.7 Fluorescence Spectroscopy and Stopped-Flow Experiments

Fluorescence spectroscopy was used to analyze binding between TruB and 2AP-tRNA using a Quanta Master 60 fluorescence spectrometer (Photon Technology International). A sample of 0.05 µM 2AP-tRNA diluted in 1X $TAKEM_4$ was excited at 325 nm and the emission was measured from 340-400 nm. 2AP-tRNA was titrated with

TruB wild type (0.03-4 µM) and R181K (0.03-5 µM) and fluorescence was monitored. Fluorescence changes at 364 nm were plotted against protein concentration. The data were fit with a hyperbolic function (see above) to determine the K_D for TruB binding to 2AP-tRNA.

Pre-steady-state fluorescence experiments were performed using a KinTek SF-2004 stopped-flow apparatus. A final concentration of 0.1 µM 2AP-tRNA was rapidly mixed with 3 µM TruB (wild type, D90N, and R181K) and excited at 325 nm. Emission was monitored at wavelengths longer than 350 nm. For TruB wild type, 30 s time courses were recorded and the data was fit with the following 3-exponential function:

$$F = F_{\infty} + A \times \exp(-k_{app1} \times t) + B \times \exp(-k_{app2} \times t) + C \times \exp(-k_{app3} \times t)$$

where F_{∞} is the fluorescence end-level; A, B, and C are the respective amplitudes of fluorescence change for each phase; t is time; and k_{app} indicates the apparent rate of the respective phase. Experiments with TruB D90N and R181K were best fit with 2-exponential functions, since they were lacking the final phase of fluorescence decrease that was observed in the TruB wild type experiments.

2.1.8 TruB Molecular Dynamics Simulations

TruB wild type was simulated in two different conformations as observed in the presence (bound) and absence (apo) of RNA. In crystal structures of *E. coli* TruB in the absence of RNA, insufficient resolution was obtained to place the thumb loop structure (residues 124-152) (Pan et al., 2003). Therefore, to simulate TruB-apo, first a homology model of *E. coli* TruB-apo was generated using the Swiss-Model server (Arnold et al., 2006; Kiefer et al., 2009) and the *Thermatoga maritima* TruB-apo crystal structure (PDBID: 1ZE1 (Phannachet and Huang, 2004)) as a template such that the ordered thumb loop could be added to the crystal structure of *E. coli* TruB-apo (PDBID: 1R3F). To do this, the homology model of *T. maritima* TruB was superimposed with the crystal structure of *E. coli* TruB-apo with respect to the catalytic domain. The coordinates of residues 124-152 corresponding to the thumb loop of *T. maritima* TruB were added to the coordinates of *E. coli* TruB-apo.

To generate a starting structure for TruB-bound, the crystal structure for *E. coli* TruB bound to a short RNA (PDBID: 1K8W (Hoang and Ferré-D'Amaré, 2001)) and the TruB-apo model were used. Since the crystal structure for TruB-bound was missing amino acids at both the N- and C- termini, *E. coli* TruB-bound was superimposed with TruB-apo with respect to residue 10, and the coordinates of residues 8-9 from TruB-apo were added to the coordinates of *E. coli* TruB-bound. The structures were then superimposed with respect to residues 310-311, and the coordinates of residues 312-314 from TruB-apo were added to the coordinates of *E. coli* TruB-bound. Water molecules within 10 Å of the protein as observed in the X-ray structure (PDBID: 1K8W) were also included in the TruB-bound model. This was not done for the TruB-apo model since there were no water molecules identified in the crystal structure. The model for the TruB D90N simulation was generated from the TruB-bound model, modifying D90 to N90 with the Mutate Residue function in VMD (Humphrey et al., 1996).

All models were solvated using the solvate package in VMD to create a water box that extended at least 10 Å around the protein. The models were then minimized in five stages using a 0.5 fs time step, in which the protein molecules were fixed for the first 10 000 steps, followed by fixing the water molecules for the next 10 000 steps, repetition of the first two stages, and releasing all constraints for 100 000 steps or until the energy of the system had stabilized. The final minimized structures were ionized and neutralized using the autoionize package in VMD. Since wet-lab work with TruB was conducted using TAKEM$_4$ buffer, the structures were ionized and subsequently neutralized by placing 100 mM K$^+$ and Cl$^-$ ions and 3 mM Mg^{2+} no closer than 5 Å to each other or to the protein. The models were further minimized for 100 000 steps at 300 K without any atom constraints. The final model for TruB-apo contained 4811 protein atoms, 45 321 water atoms, 37 K$^+$ ions, 30 Cl$^-$ ions, and 1 Mg^{2+} ion (50 200 atoms total). The final model for TruB-bound contained 4811 protein atoms, 44 634 water atoms, 37 K$^+$ ions, 30 Cl$^-$ ions, and 1 Mg^{2+} ion (49 513 atoms total). The final model for TruB D90N contained 4813 protein atoms, 44 826 water atoms, 36 K$^+$ ions, 30 Cl$^-$ ions, and 5 Mg^{2+} ions (49 710 atoms total).

Before starting the simulations, the models were equilibrated at 310 K and 360 K for 150 ps with a step size of 0.5 fs at constant pressure (1 atm) and variable volume (NPT ensemble) while invoking periodic boundary conditions. A Nosé-Hoover Langevin piston was used for pressure control, and Langevin dynamics were used to maintain a constant temperature. Production phase simulations were performed at 310 K using the final atomic velocities from the equilibration at 310 K and the atomic coordinates from the equilibration at 360 K. The particle mesh Ewald method was used to calculate full-system electrostatics. Simulations were generated in NAMD2 with CHARMM27 parameters (Brooks et al., 2009; Phillips et al., 2005) for 40 ns using periodic boundary conditions, NPT ensemble conditions with Nosé-Hoover Langevin piston for pressure control and Langevin dynamics for temperature control. The simulations were analyzed using scripts written in house, evoked in VMD. Data from the trajectories was sampled every 1-2 ps. Root-mean-square deviation (RMSD) was calculated by comparing the coordinates of the backbone atoms in each frame to the coordinates of the first frame of simulation. Root-mean-square fluctuation (RMSF) was calculated for the backbone atoms of each residue for the whole simulation excluding the first 5 ns where the simulation was still stabilizing, giving a time-averaged deviation for each residue. Active site interactions were analyzed by measuring the distance between the terminal carbon atom of the carboxylate side chain group of aspartate and the terminal carbon of the guanidinium side chain group of arginine.

2.2 Results

2.2.1 Molecular Dynamics Simulations Reveal Active Site Interactions

In order to analyze functional dynamics and changes in intramolecular interactions of TruB through MD simulations, TruB was modeled in two functionally different conformations: TruB in the absence of RNA (TruB-apo) and TruB bound to RNA, but with the RNA removed from the model (TruB-bound). The TruB-apo model was generated using a crystal structure of *E. coli* TruB (PDBID: 1R3F) and a homology model of the conformation of the thumb loop from *T. maritima* TruB (PDBID: 1ZE1) that was not resolved in the *E. coli* TruB structure. Starting from a crystal structure of *E. coli* TruB bound to a short RNA (PDBID: 1K8W), the RNA was first removed and then the N- and C-terminal residues from the TruB-apo model were added to generate the TruB-bound model with the same amino acid length as the TruB-apo model. Therefore, both models contained residues 8-314 from the *E. coli* TruB sequence. The models were solvated and ionized with the addition of 37 K^+, 30 Cl^-, and 1 Mg^{2+} ion. TruB-apo and TruB-bound were simulated for 40 ns each at 310 K.

The RMSD of the backbone atoms for both simulations was calculated for all residues (Fig. 2.1A) and for all residues excluding the thumb loop residues 124-152 (Fig. 2.1B). In the first 5 ns of both simulations, the backbone atoms moved significantly from the starting structure before stabilizing around an RMSD of 3.5 Å. When the thumb loop residues were omitted, the fluctuations in the RMSD were not as large as for the whole protein, and the average RMSD was slightly lower (approximately 3 Å). This was most apparent in the first 10 ns of the TruB-bound simulation. The average RMSD was not significantly different between the two simulations, suggesting that the remainder of the protein in both models behaved similar *in silico* and was equally flexible. Since the RMSD began to stabilize after 5 ns of simulation, further analyses were performed on 5-40 ns of simulation only.

RMSF values were calculated for all TruB residues from 5-40 ns (Fig. 2.1C) to assess which regions of TruB were the most flexible during the simulations. The RMSF for amino acids in the TruB-apo and TruB-bound simulations were generally similar, with insert 1 (residues 85-100), the thumb loop, residues in a periphery helix (α6, residues 215-230), and the C-terminal PUA domain showing the highest RMSF values. The flexibility of the thumb loop was expected, as the RMSD calculation after removing the thumb loop residues already indicated that this was a dynamic region, as well as the fact that the thumb loop is known to change conformation upon the binding of RNA to TruB (Pan et al., 2003). Likewise, crystal structures have shown that insert 1 changes its interaction partners upon RNA binding (Pan et al., 2003), so movement in this small loop was not surprising. Insert 1 was more flexible in the TruB-bound simulation, with RMSF values for these amino acids being two-fold higher than in the TruB-apo simulation. The

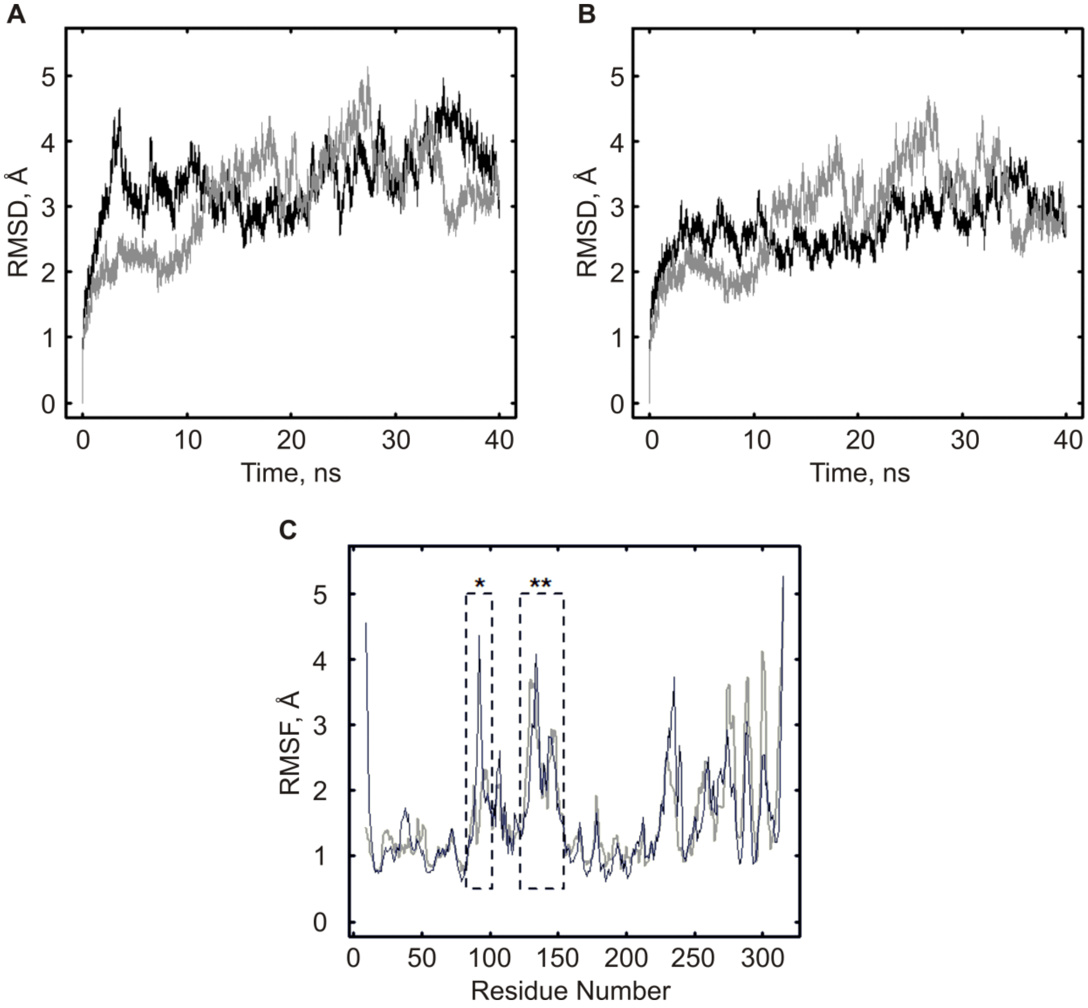

Figure 2.1 Global fluctuations of TruB in MD simulations. TruB was simulated in the apo (grey) and bound (black) conformation at 310 K for 40 ns. (A) RMSD calculation for the backbone atoms of all residues. (B) RMSD calculation for backbone atoms of all residues excluding thumb loop residues 124-152. (C) RMSF values for simulation from 5-40 ns, with insert 1 (*) and thumb loop (**) residues indicated with dashed boxes.

PUA domain is attached to the catalytic domain of TruB by a short linker that could have provided enough flexibility for this small domain to exhibit more motion than that of the catalytic domain, which was the most stable region of TruB in both simulations. In general, the flexibility seen for the TruB-apo and TruB-bound simulations were considered suitable for further analysis of the dynamics of the two models.

In order to identify and analyze interactions between relevant active site residues, a multiple sequence alignment of 100 bacterial TruB sequences was used to determine highly conserved residues that may be functionally significant. Figure 2.2A shows a coloured representation of the results of the alignment, with highly conserved residues coloured blue and poorly conserved residues coloured red, as mapped onto the TruB-apo conformation. In general, the catalytic domain was found to be more highly conserved than the PUA domain. Since pseudouridine synthases are classified based on sequence similarity, it was not surprising to see high conservation within the active site and the RNA-interaction face of the protein. Some of the 100% conserved residues include the three conserved active site residues: the catalytic aspartate (D48), Y76, which is thought

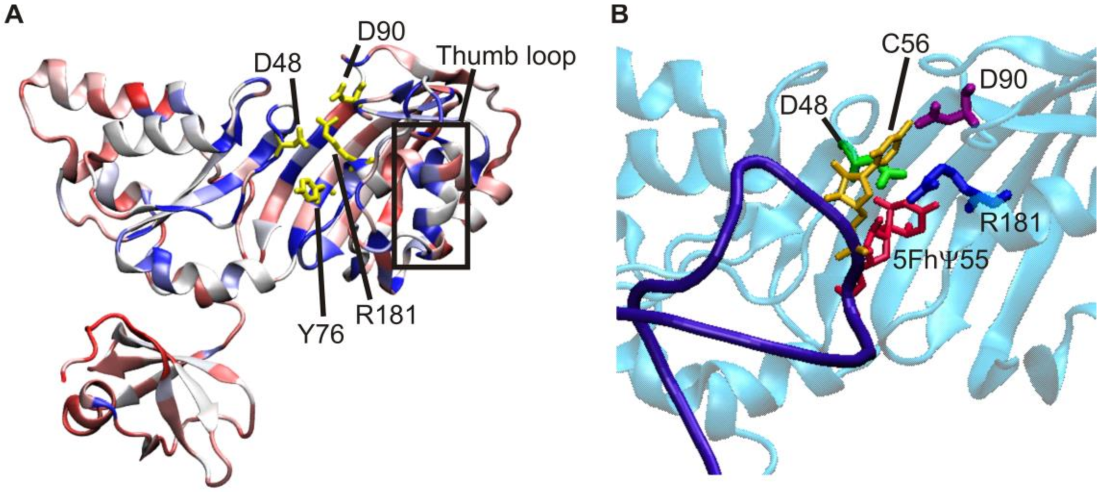

Figure 2.2 Conserved TruB active site residues. (A) TruB sequence conservation shown by a colour gradient of red to white to blue, indicating 0-100% conserved, respectively. The conservation of each residue was determined from a multiple sequence alignment of 100 bacterial TruB sequences. Stick models of four conserved residues are shown in yellow, including the catalytic D48, Y76, R181, and D90. The thumb loop, indicated by the black box, is shown in the TruB-apo conformation. (B) Zoomed-in view of TruB active site (cyan), including 22-mer T-loop RNA (dark blue). Stick models show proximity of D48 (green) and R181 (blue) of TruB to a 5-fluoro-6-hydroxy-pseudouridine (5FhΨ55) of RNA (red), as well as proximity of D90 of TruB (purple) to C56 of RNA (yellow). Image was created using PDBID: 1K8W and VMD.

to help stabilize the conformation of the flipped-out RNA bases after binding (Phannachet et al., 2005), and R181, which has been seen from crystallographic data to interact with D48 (Chaudhuri et al., 2004; Hoang and Ferré-D'Amaré, 2001; Hoang et al., 2005; Pan et al., 2003). Also conserved on the β-sheet behind the active site are the residues L200 and R202. Interestingly, a second-shell residue located in insert 1, D90, is also 100% conserved and makes contacts with C56 of the RNA substrate (Pan et al., 2003; Phannachet and Huang, 2004). R181 is located at the base of the helix below insert 1 (Fig. 2.2B) and was found to be positioned in close proximity to both D48 and D90 throughout the two TruB simulations. The side chains of these three residues were within hydrogen-bonding distance of each other. To analyze whether interactions occurred between D48, R181, and D90, the distances between the side chains were measured throughout the simulations.

Changes in the interaction patterns between D48, R181, and D90 were analysed to determine potential roles for the residues in terms of substrate binding and catalysis. The side chain of aspartate has two oxygen atoms that can interact with the hydrogen atoms of three nitrogens in the guanidinium group of the arginine side chain. In addition to salt bridge interactions, this creates a complex network of potential hydrogen-bonding interactions that can form between D48, R181, and D90. To best represent the distance between these side chains in a single measurement, the distance between the carboxylate carbon atom of the aspartate side chain and the guanidinium carbon atom of the arginine side chain was measured for D48-R181 and D90-R181 in all simulation frames of both the apo and bound conformation (Fig. 2.3).

In the TruB-apo simulation, the distance between the terminal carbons of D48 and

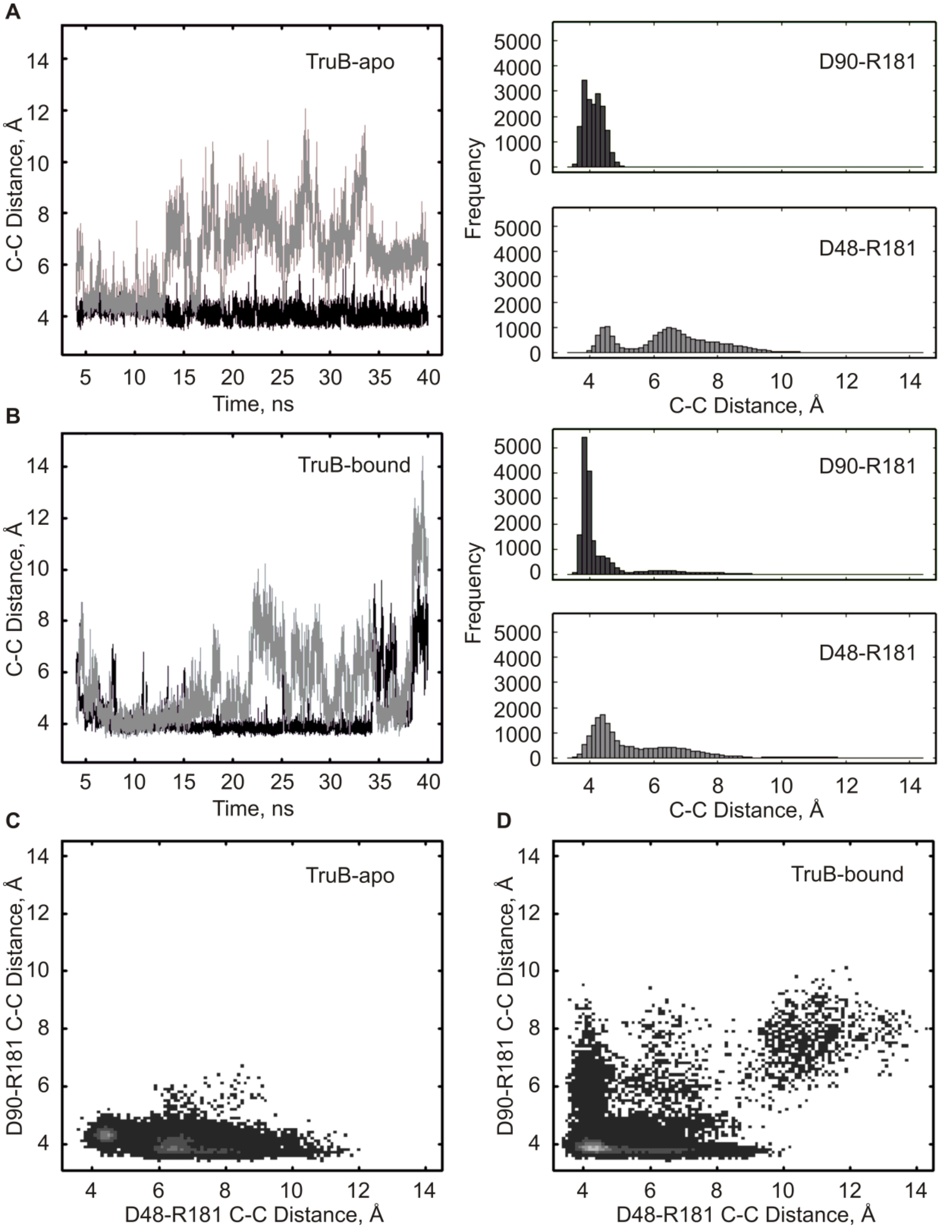

Figure 2.3 Active site residue interactions during MD simulations. Distances between center of mass of guanidinium carbon of arginine and carboxylate carbon of aspartate determined every 2 ps of simulation time: D48-R181 interaction (grey); D90-R181 interaction (black). Time-resolved changes and histogram of C-C distances for TruB-apo simulation (A) and TruB-bound simulation (B). Two-dimensional histogram of C-C distances of D90-R181 interaction plotted against C-C distances of D48-R181 for the TruB-apo simulation (C) and TruB-bound simulation (D). The occupancy of each bin is

colored from black to white, indicating 0-500 counts, respectively. Histogram bin size is 0.15 Å for (A) and (B) and 0.1 Å for (C) and (D).

R181K was relatively short (less than 5 Å) for the first 12 ns. Afterwards, large fluctuations up to 10 Å were seen with only transient sampling of the short-distance conformation. This explains the two populations seen in the histogram, with one small population centered near 5 Å and another population centered around 6.5 Å that was widely distributed up to 10 Å. In contrast, the C-C distance between D90 and R181 was quite short throughout the simulation, with one population in the histogram centered around 5 Å. This indicated that a more stable interaction was formed between D90 and R181 than between D48 and R181. In the TruB-bound simulation (Fig. 2.3B), D48 and R181 were again found in close proximity for the first 15 ns of simulation before exhibiting larger fluctuations in the measured distance for the remainder of the simulation. However, D48 and R181 more frequently sampled the short distance conformation than was seen in the TruB-apo simulation, which was best represented in the histogram showing a more highly populated maximum near 4.5 Å and a less populated but extended distribution up to 12 Å. D90 and R181 in the TruB-bound simulation were found to be separated by an even shorter distance (approximately 0.5 Å less) than seen in the TruB-apo simulation. The only fluctuations away from this short distance conformation were seen from 5-7 ns and 35-40 ns, but even these fluctuations were at a shorter distance than that observed for the maximum distance between D48 and R181. The fluctuation from 35-40 ns was due to a rather large displacement in the thumb loop that caused R181 and D90 to move further away from each other. Overall, both TruB-apo and TruB-bound had a stable interaction that formed between D90 and R181, and the interaction between D48 and R181 was in closer proximity in the TruB-bound simulation than in the TruB-apo simulation. This can be seen in the two-dimensional histograms shown in Figure 2.3C-D. The highest density population from the TruB-apo simulation data was seen when D90 and R181 were in closer proximity than D48 and R181. The TruB-bound simulation showed again the highest density of frames occurring when D90 and R181 were in close proximity, but that both residue pairs sampled conformations at further distances from each other for at least part of the simulation.

2.2.2 Role of D90 and R181 in tRNA Binding by TruB

To elucidate the role of the active site residues D90 and R181 in pseudouridine formation, substitutions were made at each position and the resultant TruB variants were analyzed for changes in function. Site-directed mutagenesis was used to modify the coding sequence of TruB to generate substitutions at D90 (to glutamate, asparagine, or alanine) and R181 (to lysine, methionine, or alanine). The substitutions were chosen to disrupt both the charge and the size of the original residue at each position. The proteins were overexpressed in *E. coli* BL21(DE3) cells with an N-terminal hexa-histidine tag and purified by Ni^{2+}-sepharose affinity chromatography and Superdex 75 size exclusion chromatography. SDS-PAGE analysis showed the proteins were purified to greater than 95% purity, and A_{260}/A_{280} ratios and urea-PAGE analysis indicated that there was no detectable RNA contamination in any of the final protein samples.

The ability of the TruB variants to bind to tRNA was tested by performing nitrocellulose filtration assays with $[^3H]$-labeled $tRNA^{Phe}$. tRNA bound to TruB was retained on the nitrocellulose filter and quantified by scintillation counting, allowing for the determination of a dissociation constant (K_D) of tRNA binding to each TruB variant.

Prior to filtering the reaction mixture, tRNA and protein were incubated for 10 min at room temperature. Under these conditions, TruB wild type would be able to bind substrate tRNA, perform catalysis, and release product tRNA, and therefore the measured K_D would indicate binding of product tRNA to TruB. It has been previously determined in the Kothe lab that TruB binds substrate and product tRNA with a similar affinity (Keffer-Wilkes, 2012), so a heterogeneous population of substrate and product tRNA would still be expected to produce a similar K_D.

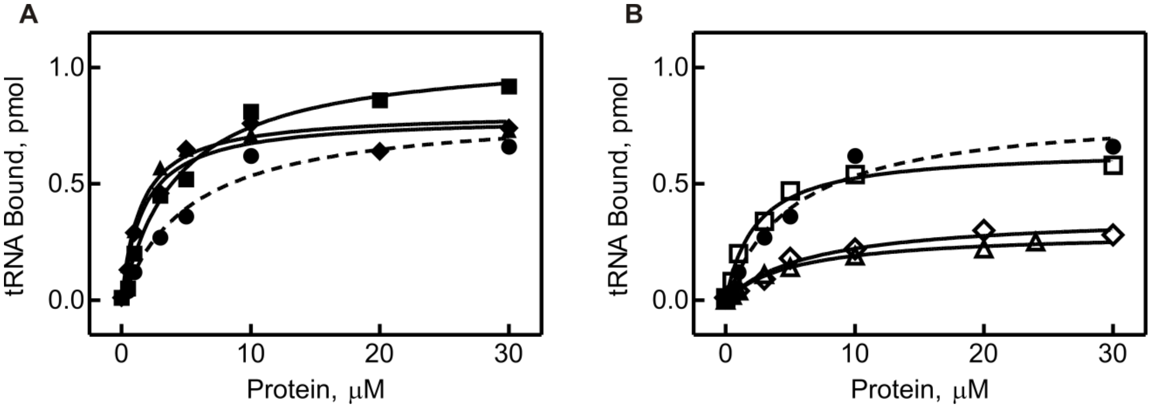

Figure 2.4 Binding of tRNA by TruB variants. Nitrocellulose filtration was used to determine the dissociation constant of [^{3}H]-labeled tRNA^{Phe} binding to TruB variants with substitutions at (A) position 90 and (B) position 181. 0.1 µM tRNA was incubated with increasing concentrations of TruB (0-30 µM). A hyperbolic equation was fit to the data to determine the K_D (smooth lines for TruB variants). (A) TruB wild type (circles, dashed line), TruB D90E (squares), TruB D90N (triangles), TruB D90A (diamonds). (B) TruB wild type (circles, dashed line), TruB R181K (open squares), TruB R181M (open triangles), TruB R181A (open diamonds).

For each TruB variant, 100 nM [^{3}H]-tRNA^{Phe} was incubated with increasing concentrations of protein (0-30 µM). Plotting the amount of tRNA bound at increasing protein concentration for the D90 variants (Fig. 2.4A) and the R181 variants (Fig. 2.4B) showed an increase in the amount of tRNA bound as protein concentration increased. TruB wild type and the D90 variants reached a maximum of 1 pmol tRNA bound, whereas the R181 variants reached a maximum 0.75 pmol tRNA bound for R181K, and

0.4 pmol for R181M and R181A. Binding was not expected to saturate at 100% tRNA bound since the experiment represents non-equilibrium conditions. The subsequent washing of the nitrocellulose membrane may result in dissociation of some of the tRNA-protein complexes, which can also cause a slight overestimation of the K_D. In particular, a low saturation point can be indicative of a fast rate of dissociation. A hyperbolic function was fit to the binding data and the resultant K_D values are summarized in Table 2.4. All TruB D90 and R181 variants appear to bind tRNA with a similar affinity as TruB wild type, with K_D values in the low micromolar range. TruB D90A had the lowest K_D of 2.0 ± 0.5 µM, and TruB R181A had the highest K_D of 9 ± 2 µM. Taking the precision of the measurements into account, no or only very small differences were seen between the affinity of TruB wild type for tRNA and the affinity of the TruB variants for tRNA.

Table 2.4 Dissociation constants for TruB-tRNA interaction determined from nitrocellulose filtration (Fig. 2.4). K_D values represent mean of three replicates ± standard deviation.

TruB	**K_D, µM**
Wild type	5 ± 2
D90E	4 ± 1
D90N	4 ± 2
D90A	2.0 ± 0.5
R181K	2.3 ± 0.5
R181M	8 ± 3
R181A	9 ± 2

2.2.3 Catalytic Activity of TruB D90 and R181 Variants

Next, the catalytic activity of the TruB variants was assessed to further characterize the role of D90 and R181. Tritium release assays were performed using [^{3}H]-tRNA labeled at position C5 of the uridine base. When uridine is converted to pseudouridine and the new glycosidic bond is formed between C5 of the base and C1’ of the ribose, the tritium is released into solution and can be quantified by scintillation counting to determine the amount of pseudouridine formed in the reaction. Multiple-turnover and single-turnover experiments were performed with enzyme concentrations of 20 nM or 200 nM, and 3 µM, respectively, and a tRNA concentration of 950 nM. The ability of the TruB variants to form pseudouridines under multiple-turnover conditions would suggest that the variants retain the ability to modify the tRNA and release the product to regenerate an active pool of enzyme. Surprisingly, initial multiple-turnover experiments using an enzyme concentration of 20 nM showed very little to no detectable product formation within 60 min for all D90 and R181 variants (data not shown). Under the same conditions, TruB wild type rapidly converted nearly all substrate into product within the first 5 min. When enzyme concentration was increased to 200 nM, complete conversion of substrate to product was seen for TruB wild type within 1 min, whereas only TruB D90E and D90N showed significant product formation under these conditions

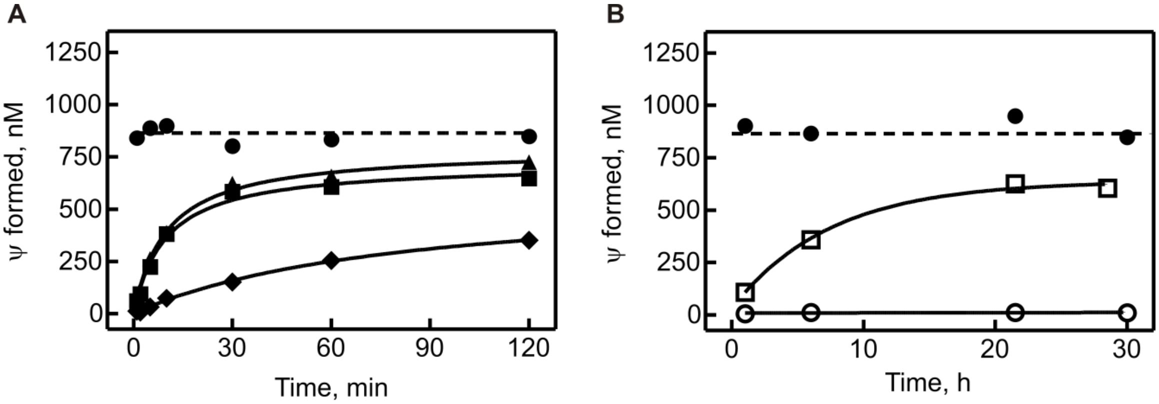

Figure 2.5 Pseudouridine formation by TruB variants under multiple-turnover conditions. [^{3}H]-labeled tRNA^{Phe} was incubated with TruB at 37°C and pseudouridine formation was quantified by tritium release assay. (A) Multiple-turnover conditions with 950 nM tRNA and 200 nM TruB wild type (closed circles), TruB D90E (closed squares), TruB D90N (closed triangles), and TruB D90A (closed diamonds). (B) Multiple-turnover conditions with 950 nM tRNA and 200 nM TruB wild type (closed circles) and TruB R181K (open squares) performed for longer incubation times to confirm activity above baseline (no enzyme added, open circles). Data was fit with a hyperbolic function (smooth lines).

(Fig. 2.5A). TruB D90A was also able to form pseudouridines, but had only reached 30% product formation after 120 min. Much longer incubation times (up to 30 h) were needed to see product formation for the R181 variants, with the TruB R181K being the only R181 variant to give readings above baseline levels (Fig. 2.5B). Even so, TruB R181K still required several hours to catalyze just one round of pseudouridine formation under

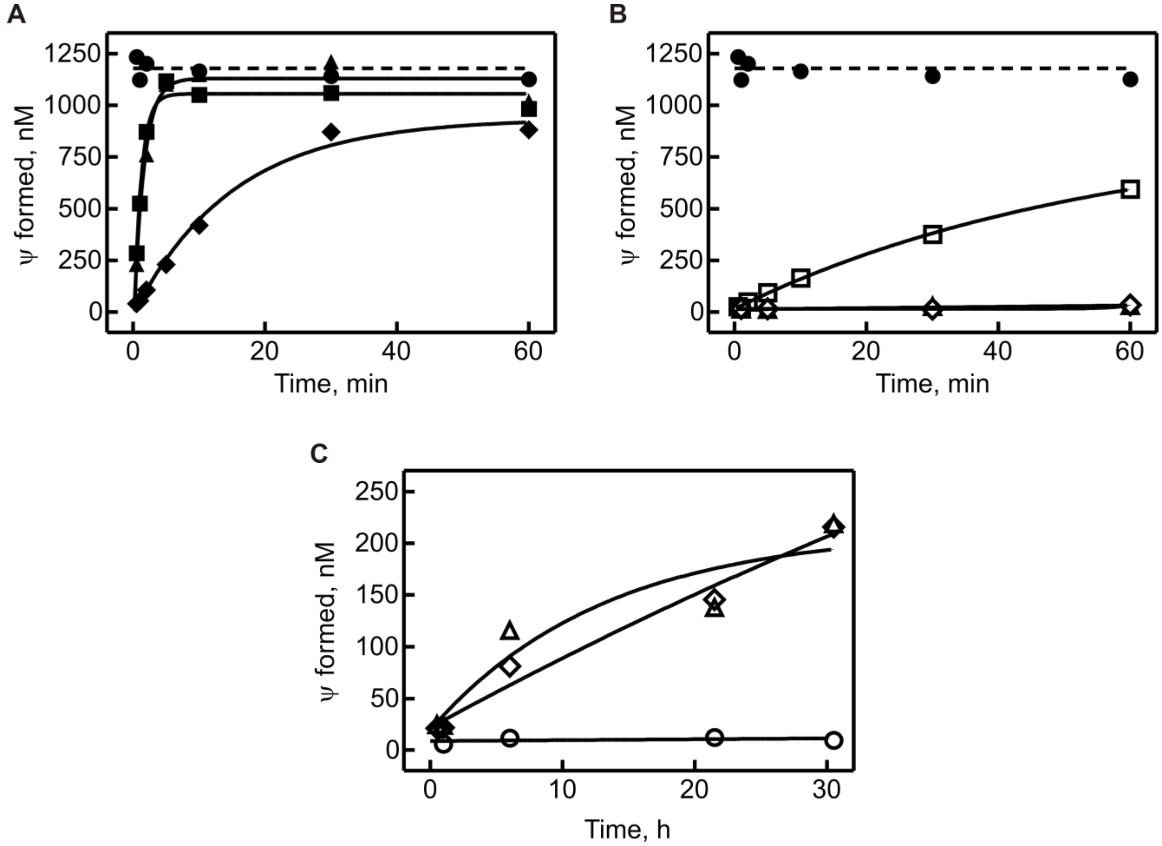

Figure 2.6 Pseudouridine formation by TruB variants under single-turnover conditions. [^{3}H]-labeled tRNA^{Phe} was incubated with TruB at 37°C and pseudouridine formation was quantified by tritium release assay. (A) Analysis of TruB variants with substitutions to D90 by single-turnover reactions with an excess of enzyme (3 µM) over tRNA (950 nM): TruB wild type (closed circles), TruB D90E (closed squares), TruB D90N (closed triangles), and TruB D90A (closed diamonds). (B) Effect of substitutions to R181 in TruB on single-turnover pseudouridine formation with 3 µM enzyme and 950 nM tRNA: TruB wild type (closed circles), TruB R181K (open squares), TruB R181M (open triangles), and TruB R181A (open diamonds). (C) Reaction with 3 µM TruB R181M (open triangles) and TruB R181A (open diamonds) was incubated for longer time points to confirm activity above baseline (no enzyme added to reaction, open circles). (A-C) A single-exponential equation was fit to the data to determine the apparent rate of catalysis, k_{app} (smooth lines).

multiple-turnover conditions. Since such a drastic effect on pseudouridylation by all TruB variants was not expected, especially the conservative R181K substitution, single-turnover experiments were performed to determine if the other R181 variants retained catalytic activity and to measure the rate of catalysis for all variants. Using an enzyme concentration of 3 µM, 3-fold higher than tRNA concentration, product formation was seen for all D90 and R181 variants (Fig. 2.6A, B), albeit at much slower rates than by TruB wild type. In fact, TruB R181M and R181A were required to be incubated with tRNA for up to 30 h before activity above baseline levels were seen (Fig. 2.6C). The single-turnover time courses were fit with a single-exponential function to obtain the rate of pseudouridine formation (Table 2.5). Compared to the single-turnover rate by TruB wild type (Wright et al., 2011), a drastic reduction in catalytic rate was seen for all variants. The D90 variants were generally more active than the R181 variants, showing a reduction in catalytic rate of approximately 30-, 50-, and 500-fold for D90E, D90N, and D90A, respectively. R181K displayed over 2000-fold reduction in catalytic activity compared to wild type, and R181M and R181A were the most impaired at pseudouridine formation with a reduction in catalytic rate of more than 20 000-fold.

Table 2.5 Apparent rates of single-turnover pseudouridine formation by TruB variants. Single-turnover rates determined from tritium release assay (Fig. 2.6) and the fold decrease in catalytic rate as compared to TruB wild type (Wright et al., 2011). k_{app} represents average rate from duplicate data ± standard deviation.

TruB	**k_{app}, s^{-1}**	**Fold decrease**
Wild type	0.5 ± 0.2	--
D90E	0.017 ± 0.002	30
D90N	0.010 ± 0.002	52
D90A	0.0010 ± 0.0002	526
R181K	0.00024 ± 0.00008	2056
R181M	$< 3 \times 10^{-5}$	> 20 000
R181A	$< 3 \times 10^{-5}$	> 20 000

2.2.4 Kinetic Analysis of TruB Variants' Interaction with tRNA

To convert uridine to pseudouridine, a number of steps have to take place including binding of tRNA, flipping out of the target uridine and the two subsequent bases into TruB's active site, breaking of the N-C glycosidic bond, rotation or flipping of the detached uracil base, formation of a new C-C glycosidic bond, and release of the modified tRNA. To better understand at which step in catalysis the TruB variants were impaired, rapid-kinetic stopped-flow analysis was performed using a 2AP-labeled tRNA. 2AP is a fluorescent nucleobase analog that is quenched when involved in base-stacking interactions and is often used to study RNA folding (Souliere et al., 2011). Upon binding to TruB, nucleotides 55-57 of the tRNA are flipped out into the active site of the enzyme in order for catalysis to proceed. By placing a 2AP label at position 57, changes in fluorescence can be monitored that are associated with disturbing the base-stacking interactions of 2AP within the native tertiary structure of the tRNA. To ensure that 2AP-tRNA was still capable of binding to TruB, fluorescence titrations were used to determine the affinity of TruB for 2AP-tRNA (Fig. 2.7A). These experiments were performed by monitoring the fluorescence of 50 nM 2AP-tRNA after excitation at 325 nm and then titrating with TruB wild type or TruB R181K. The emission maximum for 2AP was monitored at 364 nm and the relative change in fluorescence was plotted against protein concentration. A hyperbolic function was fit to the data and gave a K_D of 1.7 ± 0.2 µM for wild type and 2.7 ± 0.3 µM for R181K, which is similar to that observed for the nitrocellulose filtration assay performed with [^{3}H]-tRNA. This indicated that 2AP-tRNA was capable of binding to TruB despite the non-natural base analog and allowed for further testing with this substrate.

Pre-steady-state stopped-flow experiments were performed by rapidly mixing 0.1 µM 2AP-tRNA with an excess of TruB wild type, D90N, and R181K (3 µM). 2AP-tRNA was excited at a wavelength of 325 nm and fluorescence was monitored at wavelengths longer than 350 nm (Fig 2.7B). A biphasic rapid increase in fluorescence followed by a decrease in fluorescence was detected when 2AP-tRNA was mixed with TruB wild type. The data was best fit with a 3-exponential function, producing three apparent rates. When 2AP-tRNA was mixed with TruB D90N and R181K, the fluorescence change followed a similar biphasic increase in fluorescence as was seen for TruB wild type, but differed at the end of the time course as the fluorescence remained constant without decreasing up to 30 s. Therefore, a 2-exponential function was fit to the data for TruB D90N and R181K, resulting in two apparent rates. The first apparent rate for TruB D90N and R181K was the same, within error, as k_{app1} measured for TruB wild type, and k_{app2} was only two-fold

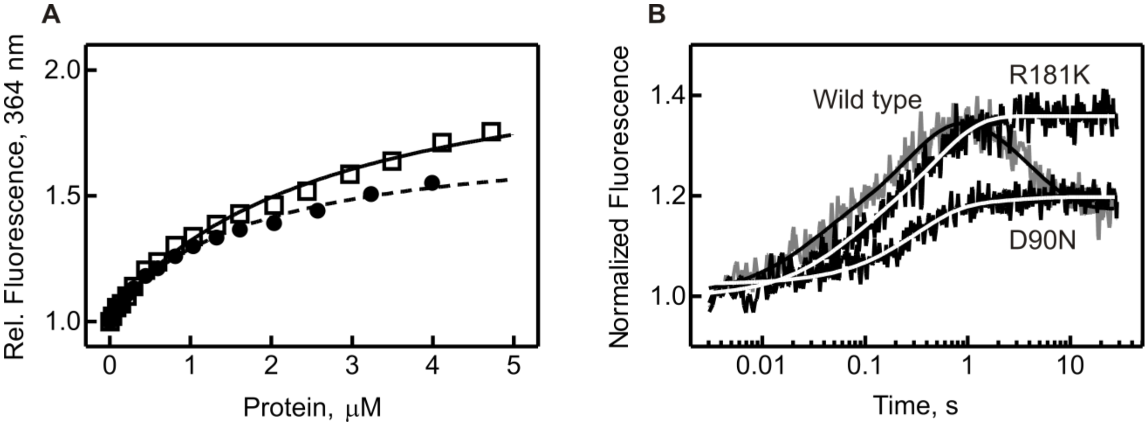

Figure 2.7 tRNA interaction with TruB observed by 2AP fluorescence. (A) 50 nM 2AP-tRNA was excited at 325 nm and fluorescence was monitored from 340-400 nm in the presence of increasing concentrations of TruB. The relative fluorescence change at 364 nm is shown for wild type (closed circles) and R181K (open squares). Fitting a hyperbolic function to the data gave a K_D of 1.7 ± 0.2 µM for wild type and 2.7 ± 0.3 µM for R181K. (B) Pre-steady-state kinetics of 2AP-tRNA interacting with TruB. 0.1 µM tRNA was rapidly mixed with 3 µM TruB wild type (grey), D90N, and R181K in a stopped-flow apparatus. The time courses were fit with a 3-exponential function for TruB wild type (smooth black line), and a 2-exponential function for D90N and R181K (smooth white lines).

slower than that for TruB wild type (Table 2.6). The third apparent rate seen only in the TruB wild type time course was similar to the previously determined rate of catalysis for TruB (0.5 ± 0.2 s^{-1} (Wright et al., 2011)) and likely reflects catalysis and product release by TruB. This phase was not seen for TruB D90N and R181K, which could be explained by the very slow rate of catalysis by these variants. The first two apparent rates could then correspond to a rapid two-step binding event including a conformational change such as, perhaps, flipping of the bases into TruB's active site. Although a relatively large error was associated with k_{app1}, this was likely due to the time points being recorded on a logarithmic scale, allowing only a small number of data points to be fit. Modifying the analysis equation to ignore this first phase resulted in a much poorer fit; this phase was therefore considered valid in each of the traces. Despite the two-fold reduction in the rates for k_{app1} and k_{app2} for tRNA interacting with TruB D90N and R181K as compared to TruB wild type, this difference is not large enough to have caused the 50-fold and 2000-fold reduction in catalytic rate. Therefore, binding and most likely also base-flipping were not strongly affected by substitutions of D90 and R181, and the observed slow rates of pseudouridine formation by the TruB variants seen in the tritium release assay can be attributed to one of the steps of catalysis being impaired.

Table 2.6 Pre-steady-state kinetic rates for 2AP-tRNA interacting with TruB. Apparent rate constants determined from exponential fitting of stopped-flow analysis of 2AP-tRNA interacting with TruB wild type, D90N, R181K. The wild type time course was fit with a 3-exponential function, and the D90N and R181K time courses were fit with 2-exponential functions.

TruB	k_{app1}, s^{-1}	k_{app2}, s^{-1}	k_{app3}, s^{-1}
Wild type	82 ± 61	4.2 ± 0.5	0.26 ± 0.02
D90N	39 ± 40	1.9 ± 0.4	--
R181K	23 ± 19	1.7 ± 0.7	--

2.2.5 Molecular Dynamics Simulations of TruB D90N

To better understand the roles of D90 and R181 in pseudouridine formation, further MD simulations were performed on a variant of TruB to complement the previously described *in vitro* experiments. Starting from the TruB-bound model to best represent an active conformation of TruB, D90 was substituted by asparagine and simulated under the same conditions as TruB wild type. RMSD and RMSF calculations were performed to measure global fluctuations and to identify flexible regions in the protein (Fig. 2.8). In the first few nanoseconds of simulation, the RMSD of all backbone residues increased to 2.5 Å where it remained until 20 ns, followed by a slow increase in the RMSD up to 40 ns. When the thumb loop residues were omitted from the calculation, the RMSD was significantly lower for the last 20 ns, showing only a slight increase to approximately 3 Å in the final 10 ns. The RMSF calculation revealed that the thumb loop

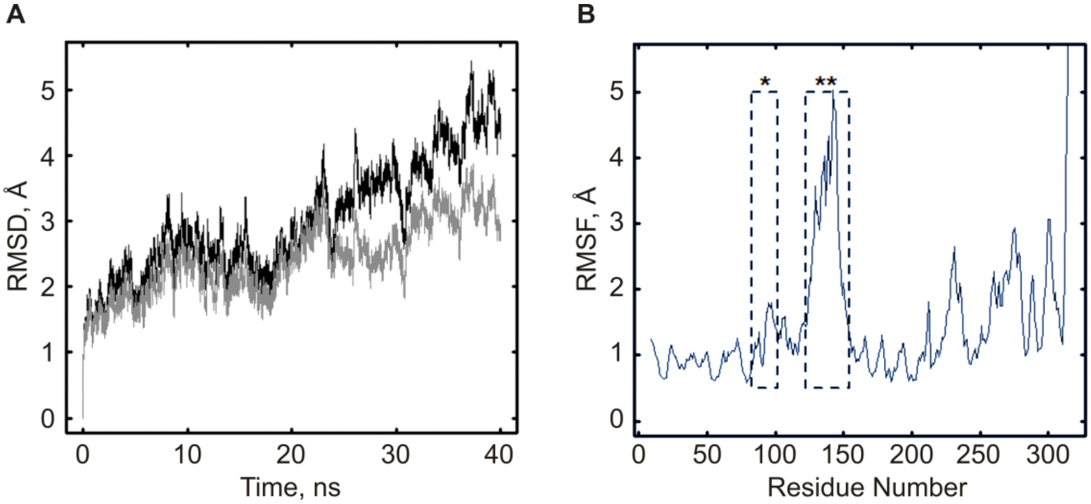

Figure 2.8 Global fluctuations of TruB D90N MD simulations. TruB D90N was simulated at 310 K for 40 ns, starting from the TruB-bound conformation. (A) RMSD calculation for the backbone atoms of all residues (black) and for backbone atoms of all residues excluding thumb loop residues 124-152 (grey). (B) RMSF values for TruB D90N residues from 5-40 ns of simulation with insert 1 (*) and thumb loop (**) residues indicated with dashed boxes.

residues were highly flexible, as was seen in the TruB-apo and TruB-bound simulations. This reflects the decrease in the RMSD seen when the thumb loop residues were omitted from the calculation. Helix α6 and the PUA domain, which had high RMSF values in the TruB-apo and TruB-bound simulations, also had high flexibility in the TruB D90N simulation. Interestingly, insert 1 that contains the substituted residue 90 was not very flexible, in contrast to what was seen in the TruB-bound simulations. This could indicate that the presence of asparagine at position 90 restricts the motion of insert 1 in the bound conformation more than the presence of an aspartate at the same position.

Distances between R181 and either D48 or N90 were also calculated in the same way as for the TruB wild type simulations (Fig 2.9). Similar to the D90-R181 interaction in the TruB-apo and TruB-bound simulations, the distance between N90 and R181 is relatively short for the duration of the TruB D90N simulation. The histogram showed one population centered around 4.5 Å, extending to approximately 7 Å. In contrast to the TruB-apo and TruB-bound simulations, the distance between D48 and R181 was noticeably shorter than the distance between N90 and R181 after the first 12 ns of simulation (approximately 3.8 Å). Although a second population at 6 Å was seen in the histogram, removing the first 12 ns of simulation data revealed only one population in the histogram at <4 Å distance (data not shown). The shorter distance suggests that a more stable interaction was formed between D48 and R181 than for N90 and R181 in TruB D90N. Also, the D48-R181 interaction was more stable in the TruB D90N simulation than it was for either of the TruB wild type simulations. The two-dimensional histogram (Fig. 2.9C) revealed a dense population at shorter D48-R181 distances, indicating that R181 was more frequently in closer proximity to D48 than to N90. This population was slightly shifted to shorter D48-R181 distances in comparison to the TruB-apo and TruB-bound simulations that showed R181 more often in closer proximity to D90 than to D48.

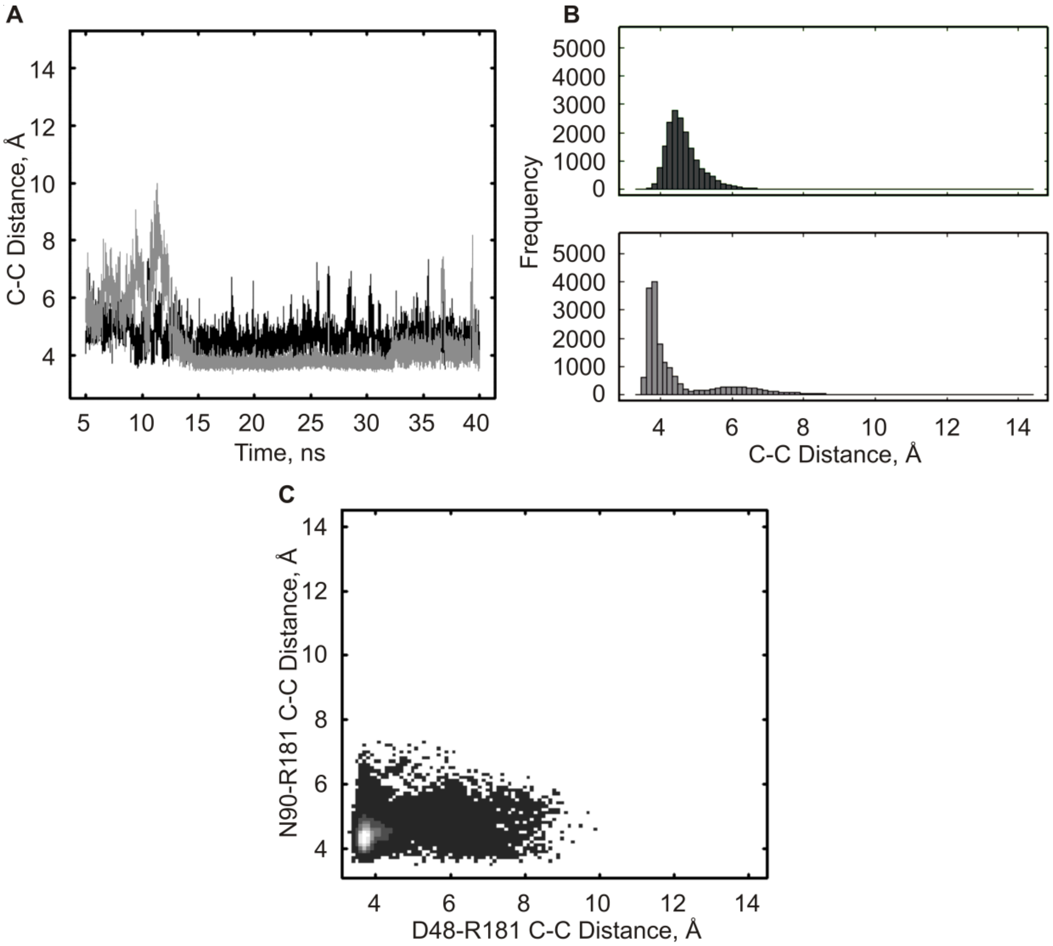

Figure 2.9 Active site residue interactions during TruB D90N MD simulations. Distances between center of mass of guanidinium carbon of arginine and carboxylate carbon of aspartate determined every 2 ps of simulation time: D48-R181 interaction (grey); N90-R181 interaction (black). (A) Time-resolved changes in C-C distances measured for the TruB D90N simulation. (B) Histogram of C-C distances during the TruB-D90N simulation, with a bin size of 0.15 Å. (C) Two-dimensional histogram of C-C distances of N90-R181 interaction plotted against C-C distances of D48-R181. The occupancy of each bin is colored from black to white, indicating 0-500 counts, respectively, with a bin size of 0.1 Å.

2.3 Discussion

The results presented here showed the importance of two active site residues, besides the already known catalytic D48, in pseudouridine formation by TruB. First, MD simulations of TruB wild type revealed the ability of the active site R181 to interact with both D48 and D90, a residue located in an insert unique to TruB. While transient interactions were formed between R181 and D48, a more stable interaction was seen between R181 and D90. Second, nitrocellulose filtration assays were used to show that substitutions to D90 and R181 did not affect TruB's affinity for tRNA. However, catalysis was greatly affected in these TruB variants. Substituting D90 for glutamate, asparagine, or alanine caused a reduction in the rate of catalysis by 30-, 50-, and 500-fold, respectively. Likewise, substituting R181 for lysine caused a 2000-fold reduction in catalytic rate, while the methionine- and alanine-substituted variants were over 20 000-fold slower at catalysis. Third, pre-steady-state kinetics of TruB's interaction with 2AP-tRNA indicated that binding was not affected and occurred in a two-step event as tRNA was mixed with TruB wild type, D90N, or R181K. Differences between the time courses were seen in a final decrease in fluorescence for TruB wild type that was not observed for D90N or R181K and likely reflects catalysis and product release. Finally, MD simulations of TruB D90N indicated that R181 and N90 were still capable of interacting, but in contrast to the TruB wild type simulations, the transient interaction between R181 and D48 was stabilized in the TruB D90N simulation.

2.3.1 TruB Wild Type MD Simulations

TruB was simulated in two functionally different conformations in order to analyze intramolecular dynamics that occur at the atomic level. In accordance with observations from crystal structures, RMSF calculations for the TruB-apo and TruB-bound simulations indicated that the thumb loop, insert 1, and PUA domain were highly flexible, which may facilitate the conformational changes that are known to occur upon

RNA binding. The thumb loop residues were seen to greatly contribute to the RMSD calculations in both simulations. This clearly reflects the inherent flexibility needed for this region to undergo such a large conformational change from the apo to the bound conformation. Additionally, when tRNA is bound to TruB, interactions between the tRNA and the thumb loop likely stabilize the bound conformation of the thumb loop (Phannachet and Huang, 2004). Since the simulations were performed in the absence of tRNA, these potentially stabilizing interactions were removed and therefore it was not surprising to see flexibility in the thumb loop. Were tRNA to be added to the simulation, I would expect the thumb loop flexibility to be significantly reduced.

Protein MD simulations performed using CHARMM parameters typically give RMSD values from 1-2 Å (Li et al., 2005; Wieden et al., 2010). Although the RMSD values of the TruB simulations were consistently higher than this (closer to 3 Å), the increased flexibility and global motion were not considered to render the simulations unusable for the following reasons. The relatively high RMSD could be due to the overall shape of TruB or the presence of inherently flexible regions such as the thumb loop, insert 1, and PUA domain. Though the catalytic domain and PUA domain are both globular in shape, the linker region between them could allow for increased structural fluctuations that contribute to a higher RMSD. There were no significant drifts or trends in the RMSD that would indicate a breakdown in the parameters that were used for simulating TruB, nor were there any apparent distortions in the structure as the simulations progressed. Additional simulation time can be generated to ensure that no trends in the RMSD develop and that sufficient time was given to explore conformational space. As these were the first MD simulations performed with TruB, our work has laid the foundation for further analyses and can be used to direct changes in simulation conditions that can be explored to answer more specific questions about the dynamics of TruB.

Crystal structures of TruB have indicated the presence of a salt bridge interaction between the catalytic D48 residue and R181 at the base of a nearby helix (Chaudhuri et al., 2004; Hoang and Ferré-D'Amaré, 2001; Hoang et al., 2005; Pan et al., 2003). Both of these residues are 100% conserved in TruB. D90 in insert 1 can interact with the nucleotide adjacent to the target uridine in the tRNA substrate (C56) and is also 100% conserved. The MD simulations of TruB performed here revealed that R181 was capable of forming a salt bridge with D90 in addition to D48, which has not been reported in crystallographic studies. Since the salt bridge between D48 and R181 has been proposed to keep D48 in a deprotonated state in order for it to act as a nucleophile in pseudouridine formation (Hoang and Ferré-D'Amaré, 2001), the presence of an additional interaction partner for R181 has interesting implications for the role of these three residues in catalysis. In particular, R181 was more frequently in closer proximity to D90 than to D48 in both the TruB-apo and TruB-bound simulations. This could indicate that the interaction between R181 and D48 is more transient than suggested based on crystal structures and that R181 interacting with D90 may be important for disrupting the interaction with D48 in order to make the catalytic residue available for pseudouridine formation. The interaction between R181 and D48 may be more important for the positioning of D48 for catalysis than for activating it for catalysis. *In vitro* mutational studies were used to better dissect the role of these three residues in regards to the function of TruB.

2.3.2 Binding and Catalysis by TruB D90 and R181 Variants

A series of substitutions were made to D90 and R181 to test the importance of the charge, size, and identity of these residues for tRNA binding and pseudouridine formation by TruB. Since D90 has been seen in crystal structures to hydrogen bond with C56 in the tRNA (Pan et al., 2003; Phannachet and Huang, 2004), it was expected that substitutions to D90 would have a greater effect on tRNA binding than on catalysis. However, my results show that the TruB D90 variants all had a similar affinity for tRNA as TruB wild type, indicating that D90 is not essential for substrate binding. Surprisingly though, these variants were all severely impaired at pseudouridine formation, with even the most active variant (TruB D90E) showing a 30-fold decrease in catalytic rate. This was quite unexpected, as it was not apparent from crystallographic data how this residue could be involved in catalysis. Even substituting D90 for asparagine caused over 50-fold reduction in pseudouridine formation. This suggests that although the size of the residue is important, since TruB D90A was the least active of the D90 variants, the negative charge of the residue may be even more important. Since TruB D90E was the only D90 variant that retained the negative charge found in TruB wild type, the salt bridge interaction with R181 discovered through MD simulations is likely more important than hydrogen bond interactions with C56 in the RNA that does not require a charged residue and should be retained in the TruB D90N variant. Crystal structures of TruB bound to RNA have shown that C56 can also be recognized by residues in the thumb loop (Phannachet and Huang, 2004), and therefore substitutions at D90 could potentially be compensated for by these additional interactions, which may explain why tRNA binding was not affected.

Although R181 has been discussed in a number of structural papers to interact with the catalytic aspartate (Chaudhuri et al., 2004; Hoang and Ferré-D’Amaré, 2001; Hoang et al., 2005; Pan et al., 2003), to our knowledge no biochemical studies have been reported on the role of this residue in the function of TruB or other pseudouridine synthase families. Nitrocellulose filtration assays showed that all TruB R181 variants bound tRNA with a similar affinity as TruB wild type. Although within error, TruB R181M and TruB R181A had a K_D 1.5-fold higher than that of TruB wild type. The slight increase in the measured affinity for these two variants could be due to a faster rate of dissociation, causing more of the tRNA-TruB complexes to dissociate in the subsequent wash step. The binding curves for TruB R181M and R181A had the lowest saturation point, which is further evidence that a faster rate of dissociation may be the cause of the slightly higher affinity for tRNA. Therefore, the main role of R181 is unlikely to be assisting with tRNA binding, as replacing the positively charged side chain of arginine with methionine and alanine did not greatly affect TruB’s ability to bind tRNA.

When the activity of the TruB R181 variants was tested, high concentrations of enzyme and long incubation times were needed to measure pseudouridine formation above baseline levels. Such a drastic effect was unexpected, especially for the conservative substitution of arginine for lysine that was over 2000-fold slower at catalysis. In fact, methionine and alanine substitutions produced enzymes that were first thought to be inactive as incubation times of up to 30 hours were required to detect modest levels of pseudouridine formation. This suggests that the identity of R181 is highly critical for pseudouridine formation and that a basic residue is needed in the catalytic pocket. One of the major differences between arginine and lysine are the number of potential hydrogen bond donors. It could be that R181 is needed to interact with multiple residues simultaneously and therefore more than the one amino group found in lysine is required. Since it was seen through MD simulations that R181 can interact with D48 and D90, the guanidinium group may be the central hub of these two interactions and therefore cannot be substituted for an amino acid with fewer hydrogen bond donors. Although the side chain of lysine is shorter by one bond than the side chain of arginine, visual inspection of substituting R181 for lysine in the crystal structure of TruB in the bound conformation revealed that lysine should still be long enough to effectively interact with either D48 or D90. Therefore, the multiple hydrogen bond donors on the guanidinium group of arginine and the resultant network of interaction partners is most likely the critical property of R181 that is required for fast catalysis.

2.3.3 Pre-Steady-State Kinetics of TruB Variants

The rate of catalysis determined from tritium release assays reflects both the rate of substrate binding and pseudouridine formation (but not product release). However, these steps can be further divided into sub-steps such as initial substrate binding, base-flipping, and catalysis. Since the TruB variants were not significantly affected at binding tRNA, a 2AP-labeled tRNA substrate was used to test whether there was a step between binding and catalysis that was causing the reduced catalytic rates. With this fluorophore, rearrangements in the helical stacking of the tRNA were expected to produce fluorescence changes that could be monitored by pre-steady-state stopped-flow experiments. Fluorescence titrations showed that the incorporation of this modified base at position 57 in the tRNA did not affect the affinity of TruB for tRNA, as was expected based on biochemical studies showing that base substitutions to G57 were well tolerated by TruB (Gu et al., 1998).

When 2AP-tRNA was rapidly mixed with TruB wild type, a triphasic time course was seen with two rapid increases in fluorescence followed by a slower decrease in fluorescence. The fluorescence change upon mixing with TruB D90N and R181K displayed a similar biphasic increase, but not a decrease in fluorescence at the end of the time course. The rates of the two fluorescence increases were similar in magnitude for all three proteins and probably correspond to two separate binding events, such as initial binding of tRNA and flipping out of the bases into TruB's active site. Previous studies in the Kothe lab observing changes in tRNA absorbance have shown that TruB binds tRNA in two steps with an overall rate of 6.0 ± 1.8 s^{-1} (Wright et al., 2011), which is very similar to the k_{app2} measured here. Interestingly, the highly sensitive 2AP label allows detection of an additional rapid phase of fluorescence increase. Due to the nature of this fluorophore, the signal change could be associated with changes in base stacking, such as flipping-out of the nucleotide into the enzyme active site that is not detectable with unlabeled tRNA. However, it cannot be ruled out that this rapid fluorescence change represents another step such as the bimolecular initial binding of TruB and tRNA.

The fluorescence decrease seen during the TruB wild type trace had a rate comparable to TruB's rate of catalysis. Under single-turnover conditions, TruB catalyzes pseudouridine formation at a rate of 0.5 ± 0.2 s^{-1} (Wright et al., 2011). Therefore, k_{app3} likely represents product release, which is expected to be limited by the rate of catalysis and could produce a fluorescence decrease as the 2AP-label gets repositioned back in the helical stack of the T-arm upon tRNA dissociation and the fluorescence of 2AP is quenched. Since TruB D90N and R181K were significantly slower at catalysis, it was not surprising that this phase was absent in their respective traces. These experiments suggest that TruB D90N and R181K were not kinetically impaired at binding tRNA or base flipping and were most likely exclusively impaired at one of the steps of catalysis.

2.3.4 TruB D90N MD Simulations

MD simulations were also performed on TruB D90N in the bound conformation to analyze any differences in the interaction pattern of D48, R181, and N90 in comparison to TruB wild type. RMSD and RMSF calculations indicated that TruB D90N had similar flexibility in simulation as TruB wild type. The only noticeable difference was seen in the mobility of insert 1 that contains the D90N substitution, which became much more stable in this simulation than it was for the TruB-bound simulation. As the substitution was the only difference between the starting structure of the TruB-bound and TruB D90N simulations, the increased stability of insert 1 could be attributed to the presence of asparagine over aspartate. Interestingly, the TruB D90N RMSF values for insert 1 were more similar to TruB-apo than to TruB-bound. In fact, residue 90 in TruB D90N and TruB-apo had the second lowest RMSF value of all insert 1 residues. Therefore, although asparagine substitutions may make insert 1 more stable in the bound conformation, it is likely not the only factor that can contribute to stability (or flexibility) in this region.

When the distances between D48-R181 and N90-R181 were calculated for the TruB D90N simulation, an interesting trend was observed that was different from that seen for TruB-apo and TruB-bound. While the TruB wild type simulations showed that R181 was more frequently in closer proximity to D90, in the TruB D90N simulation

R181 made a closer contact to D48 than to N90. In fact, of all three simulations, the closest proximity contact (3.8 Å) between R181 and D48 was seen in the TruB D90N simulation. However, the distance between R181 and N90 was similar to the short distance seen between R181 and D90, indicating that this interaction was maintained and not significantly affected. Therefore, changing D90 to asparagine stabilized the interaction between R181 and D48, but did not drastically alter the interaction between R181 and N90. Combining this data with the observed 50-fold reduction in catalytic rate by TruB D90N suggests that forming a stable interaction between R181 and D48 is detrimental to catalysis. Though D90 does not directly interact with D48, a network of interactions including R181 seems to connect these two residues and changes to the interaction pattern of D90 (by substitution with asparagine) can greatly affect TruB's ability to form pseudouridine.

2.3.5 The Role of R181 and D90 in Pseudouridine Formation by TruB

We propose the following model for the role of the network of D48, R181, and D90 in catalysis. An interaction between R181 and D48 is necessary for catalysis, but the stability of this interaction must be mediated by a secondary interaction between R181 and D90. If D48 is not able to interact with R181, as was the case for the TruB R181M and R181A variants, catalysis is almost completely abolished. However, if the interaction between R181 and D48 is too stable, as was seen in the TruB D90N MD simulation, the rate of catalysis is also reduced. Therefore, R181 must be allowed to interact with D48, but the strength of the interaction needs to be mediated by additional contacts to allow for efficient catalysis by TruB. This role can be filled by D90, as a stable R181-D90 interaction was seen during the TruB-apo and TruB-bound simulations. Since a greater reduction in catalytic rate was seen for substitutions to R181, the R181-D48 interaction is more critical for catalysis than the R181-D90 interaction. D90 may be important for fine-tuning the position of R181 and the stability of the R181-D48 interaction in order to best facilitate pseudouridine formation. A moderate interaction between R181 and D48 is likely necessary for properly positioning D48 in the active site of TruB when tRNA binds, but not too strong of an interaction which would prevent D48 from participating in catalysis.

Though a basic active site residue is conserved in all pseudouridine synthase families, this residue is sometimes a lysine (as in RsuA, TruD, and Pus10) (Hoang and Ferré-D'Amaré, 2004; McCleverty et al., 2007; Sivaraman et al., 2002). The multiple hydrogen bond donors found in arginine is likely the critical property of this residue in contributing to catalysis by TruB, but is not essential in those enzymes that contain lysine, as Pus10 was found to have a rate constant for catalysis similar to arginine-containing enzymes, including TruB (Kamalampeta et al., 2013; Wright et al., 2011). Similarly, D90 is found in an insert unique to TruB and therefore the corresponding residue is not found in other pseudouridine synthases. If a universal mechanism for catalysis is used by all pseudouridine synthase families, the role of D90 in mediating the salt bridge interaction between the catalytic aspartate and the active site basic residue would either not be necessary or must be filled by a completely different residue in pseudouridine synthases other than TruB. The different families of pseudouridine synthases are often defined by the presence of unique domains or inserts, so the other families have likely evolved a slightly different network of active site interactions with which to mediate the positioning of the catalytic aspartate for pseudouridine formation.

This has already been seen when comparing TruB to RluA, RsuA, and TruA; a histidine residue is needed to stabilize the flipped-out nucleotides in TruB, but RluA, RsuA, and TruA utilize an arginine at a different position for this role. The families that contain lysine in the active site could use this residue mostly for shielding the negative charge of the catalytic aspartate and then rely on another residue for positioning the aspartate for catalysis. Alternatively, in these pseudouridine synthases, the active site lysine could both position and activate the catalytic aspartate if its interaction with the catalytic aspartate is weaker than the R181-D48 interaction in TruB wild type, but stronger than in the TruB R181K variant. In both cases, an analogous residue for TruB’s D90 would not be needed. Lastly, another type of residue or an aspartate at a different position could be used by the other pseudouridine synthase families to mediate the important interaction between the conserved catalytic residue and basic residue that has been shown here to be critical for pseudouridine formation.

Chapter 3 – Effect of Mg^{2+} Ions on tRNA Dynamics *In Silico*

Though TruB has been crystallized with a truncated RNA substrate (Hoang and Ferré-D'Amaré, 2001; Hoang et al., 2005), no structures have been solved with a full-length tRNA. Using the truncated substrate as a guide, a full-length tRNA can be modeled into the active site of TruB to gain insight into what additional contacts can be made between the protein and the RNA. To get a better idea of the dynamics of this interaction, an MD simulation of the modeled complex can be used to better understand substrate recognition by TruB, protein-induced conformational changes to the tRNA, or changes to interactions between active site residues (such as D48, R181, and D90 discussed in Chapter 2) when bound to tRNA. In preparation for this type of simulation, I performed MD simulations of unmodified $tRNA^{Phe}$ from *E. coli* in order to optimize conditions for nucleic acid simulation. While protocols for MD simulations of proteins are well established, there is still much to be learned about the treatment of nucleic acids *in silico*. For example, minimization and equilibration steps are often more stringent for nucleic acids to account for their highly charged phosphate backbone and to allow for water and ions to diffuse into the RNA structure (Auffinger and Westhof, 1997; Eargle et al., 2008). The placement and concentration of cations such as Mg^{2+}, reported to be important for RNA folding, must also be carefully considered for nucleic acid simulations (Auffinger et al., 1999; Serebrov et al., 2001). For this reason, I simulated tRNA under a variety of Mg^{2+} ion conditions and analyzed the effect of Mg^{2+} ion concentration and initial placement on global and local stability of the tRNA. The simulations were also analyzed after being sorted into two clusters based on Mg^{2+} occupancy at a specific position within the tRNA. The knowledge gained from these simulations will direct the development of simulation procedures for a tRNA-TruB complex.

3.1 Methods

3.1.1 Generating tRNA Models

Unmodified *E. coli* $tRNA^{Phe}$ was modelled using the respective crystal structure (PDBID: 3L0U (Byrne et al., 2010)) and the coordinates for the 3'-CCA nucleotides from *Saccharomyces cerevisiae* $tRNA^{Phe}$ (PDBID: 1EHZ (Shi and Moore, 2000)). Before starting the simulations, potential Mg^{2+} binding sites in tRNA were identified using the Database of Metal Ion Binding Sites in RNA (MeRNA, http://merna.lbl.gov) and the Protein Data Bank (http://www.rcsb.org). In total, 28 tRNA structures containing Mg^{2+} (Table 3.1) were superimposed with respect to the phosphodiester backbone, and the probability of an Mg^{2+} ion being found within a three-dimensional grid (3x3x3 Å grid spacing) was calculated. Alternatively, Mg^{2+} ions were placed using the cionize program in VMD (Humphrey et al., 1996), which places ions at points of minimal energy based on Coulombic potential and recalculates the potential after each new ion is added. In total, seven tRNA models were generated (Table 3.2) with Mg^{2+} being positioned based on highest probability in the 3x3x3 Å grid (named with "X"), by cionize in VMD (named "C"), or by random placement no closer than 5 Å to each other or the tRNA using the autoionize program in VMD (named with "R").

Each structure was solvated by placing the tRNA in a water box that extended at least 10 Å from the tRNA using the solvate package in VMD. The models were then minimized in five stages using a 0.5 fs time step, in which the tRNA and Mg^{2+} ions were fixed for the first 10 000 steps, followed by fixing only the water molecules for the next

Table 3.1 tRNA crystal structures used to find potential Mg^{2+} binding sites. All models were superimposed on the PDBID 3L0U structure to determine high probability locations of Mg^{2+} occupancy from crystal structures (used in the "X"-named models).

PDBID	Molecule	Organism	No. of Mg^{2+} ions
3L0U	Unmodified $tRNA^{Phe}$	*E. coli*	2
1EHZ	Mature $tRNA^{Phe}$	*S. cerevisiae*	6
1TN2	Mature $tRNA^{Phe}$	*S. cerevisiae*	5
1TRA	Mature $tRNA^{Phe}$	*S. cerevisiae*	5
2TRA	Mature $tRNA^{Asp}$	*S. cerevisiae*	1
3TRA	Mature $tRNA^{Asp}$	*S. cerevisiae*	1
4TNA	Mature $tRNA^{Phe}$	*S. cerevisiae*	4
4TRA	Mature $tRNA^{Phe}$	*S. cerevisiae*	4
6TNA	Mature $tRNA^{Phe}$	*S. cerevisiae*	4
1EVV	Mature $tRNA^{Phe}$	*S. cerevisiae*	10
1I9V	Neomycin:$tRNA^{Phe}$	*S. cerevisiae*	3
1B23	EF-Tu:GTP:$tRNA^{Cys}$	*T. aquaticus*	3
1OB5	EF-Tu:GDPNP:Phe-$tRNA^{Phe}$	*S. cerevisiae*	3
1TTT	EF-Tu:GDPNP:Phe-$tRNA^{Phe}$	*S. cerevisiae*	6
1FFY	IleRS:Mupirocin:$tRNA^{Ile}$	*S. aureus*	10
1FIR	Reverse transcription primer $tRNA^{Lys,3}$	HIV	1
1J1U	TyrRS:$tRNA^{Tyr}$:L-Tyr	*M. jannaschii*	4
2AZX	TrpRS:Trp-$tRNA^{Trp}$	*H. sapiens*	1
1N77	GluRS:ATP:$tRNA^{Glu}$	*T. thermophilus*	1
2IY5	PheRS:PheOH-AMP:$tRNA^{Phe}$	*T. thermophilus*	1
2DXI	GluRS:$tRNA^{Glu}$	*T. thermophilus*	2
2CV1	GluRS:ATP:Glutamol:$tRNA^{Glu}$	*T. thermophilus*	4
2ZUE	ArgRS:ANP:$tRNA^{Arg}$	*P. horikoshii*	1
2DLC	TyrRS:Tyr-AMP:$tRNA^{Tyr}$	*S. cerevisiae*	1
3EPH	DMATase:$tRNA^{Cys}$	*S. cerevisiae*	10
3EPJ	DMATase:DMAT:$tRNA^{Cys}$	*S. cerevisiae*	16
2ZXU	MiaA:$tRNA^{Phe}$	*E. coli*	10
2ZM5	MiaA:$tRNA^{Phe}$	*E. coli*	7

10 000 steps, repetition of the first two stages, and releasing all atom constraints for 100 000 steps or until the energy of the system had stabilized. After minimization, the system was neutralized with 100 mM K^+ and Cl^- ions placed no closer than 5 Å to each other or the tRNA (all models contained 30 Cl^- ions, see Table 3.2 for number of K^+ ions). A final minimization step was carried out for 100 000 steps at 300 K with no atoms constraints.

Table 3.2 tRNA models for simulation.

Model Name	Number of Mg^{2+} ions	Number of K^+ ions	Number of tRNA atoms	Number of water atoms	Total number of atoms
$0Mg^{2+}$	**0**	106	2445	48 297	50 878
$R\text{-}4Mg^{2+}$	**4** (randomly placed)	98	2445	48 285	50 862
$R\text{-}18Mg^{2+}$	**18** (randomly placed)	70	2445	48 243	50 806
$X\text{-}4Mg^{2+}$	**4** (from 3x3x3Å grid)	98	2445	48 276	50 853
$X\text{-}18Mg^{2+}$	**18** (4 from 3x3x3Å grid and 14 randomly placed)	70	2445	48 234	50 797
$C\text{-}4Mg^{2+}$	**4** (from cionize)	98	2445	48 294	50 871
$C\text{-}18Mg^{2+}$	**18** (4 from cionize and 14 randomly placed)	70	2445	48 252	50 815

3.1.2 tRNA Molecular Dynamics Simulations

The tRNA models were equilibrated using the procedure outlined by Eargle, *et al.* 2008. Briefly, harmonic constraints (1 kcal mol^{-1} $Å^{-2}$) were applied to the tRNA atoms, ions, and heavy atoms for 25 ps at 100 K using a 0.5 fs time step. The constraints were then applied to only the tRNA backbone atoms for an additional 25 ps at 100 K, followed by 3.9 ns equilibration at 298 K with all constraints removed. All models were stable through the equilibration stage, and simulations were continued for the $R\text{-}18Mg^{2+}$, $X\text{-}18Mg^{2+}$, and $C\text{-}18Mg^{2+}$ models since these models had the lowest energy after equilibration. The $0Mg^{2+}$ model was also simulated, though it had the highest energy after equilibration, to serve as a control simulation. Simulations were performed at 310 K in NAMD2 using the CHARMM27 parameters (Brooks et al., 2009; Phillips et al., 2005) for 40 ns using periodic boundary conditions, NPT ensemble methods, and a Nosé-Hoover Langevin piston for pressure control and Langevin dynamics for temperature control. The particle mesh Ewald method was used to calculate full-system electrostatics. VMD was used to visualize and analyze the simulations using scripts written in-house.

3.1.3 MD Simulation Analysis

A number of analyses were performed to describe the global and local dynamics of the tRNA. RMSD and RMSF were calculated using scripts evoked in VMD. The number of Watson-Crick base pairs was determined by counting the number of hydrogen bond donor and acceptor atoms in A-U and G-C base pairs that were within a 2.7 Å cut-off. For A-U base pairs, the distance was measured between N1 of adenine and H3 of uridine, and between H61 of adenine and O4 of uridine. Likewise, the following distances were measured for G-C base pairs: O6 of guanine and H41 of cytidine, H1 of guanine and N3 of cytidine, and H21 of guanine and O2 of cytidine. Each of these atom pair distances were calculated individually and any nucleotides with these atom pairs within 2.7 Å in a particular frame of simulation, regardless of nucleotide position or presence of the other hydrogen bond(s), were counted toward the total number of Watson-Crick base pairs formed. The same method was used to rate the stability of the G15-C48, G18-U55, and G19-C56 tertiary interactions, by analyzing the distance between the respective hydrogen bonding atoms in each interaction. For example, G18-U55 forms a *trans* Watson-Crick sugar-edge base pair (Byrne et al., 2010), so the distance between H1 of guanine and O2 of uridine, and H21 of guanine and O2 of uridine was calculated. Changes in the tRNA elbow angle were determined by measuring the angle formed between the phosphate atoms of the nucleotides C72, C56, and A35 (Fig 3.1A). Likewise, changes in the torsion angle of the tRNA were found by measuring the dihedral angle formed between the phosphate atoms of the nucleotides C72, C56, C41, and G29 (Fig 3.1B).

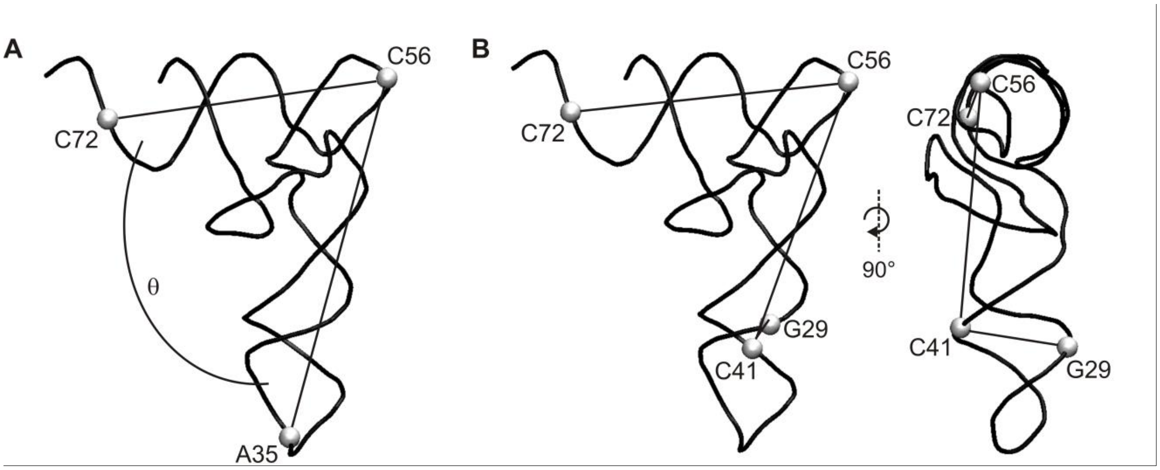

Figure 3.1 tRNA elbow angle and torsion angle. (A) To measure the elbow angle of tRNA during the simulations, the angle formed between the phosphate atom of C72, C56, and A35 was measured. (B) The torsion angle of the tRNA was determined by measuring the dihedral angle formed by the phosphate atoms of C72, C56, C41, and G29.

3.1.4 Cluster Analysis of tRNA MD Simulations

Cluster analyses were performed on the R-18Mg^{2+}, X-18Mg^{2+}, and C-18Mg^{2+} simulations by clustering the simulation frames based on whether or not Mg^{2+} was present at a specific location within the tRNA structure. This position was determined by analyzing which positions were highly populated by Mg^{2+} ions in each frame of the simulations, similar to the generation of the Mg^{2+} positions for the X-18Mg^{2+} model. A 4x4x4 Å grid was used to locate coordinates of high Mg^{2+} residency throughout the three simulations. The highest probability position from each of the three tRNA simulations containing Mg^{2+} was used to cluster all frames in each simulation. Two clusters were defined as those frames containing a Mg^{2+} ion within 3.5 Å of the reference position and those that did not (occupied and unoccupied, respectively). This criterion gave a significant number of frames in each cluster when using the reference Mg^{2+} position from

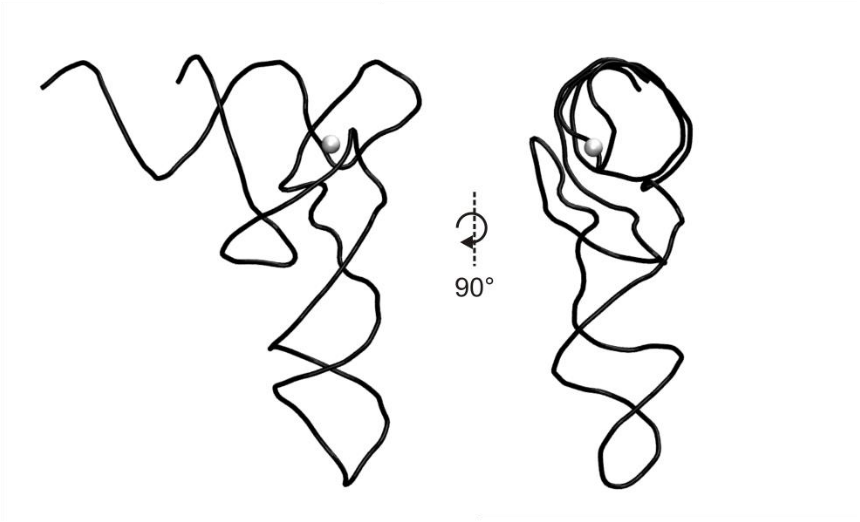

Figure 3.2 Cluster analysis reference Mg^{2+} position. High probability Mg^{2+} ion position (white sphere) identified from the X-18Mg^{2+} simulation and used to cluster the R-18Mg^{2+}, X-18Mg^{2+}, and C-18Mg^{2+} simulations based on Mg^{2+} occupancy within 3.5 Å of this reference position.

the X-18Mg^{2+} simulation (Table 3.3 and Figure 3.2); therefore, the clusters generated using this reference position were used for further analysis of the same parameters as used for the whole simulation (e.g. number of hydrogen bonds) to determine whether the presence of Mg^{2+} influenced any of those characteristics.

Table 3.3 Number of simulation frames in each cluster based on Mg^{2+} occupancy at a reference position observed in the X-18Mg^{2+} simulation.

Simulation	Occupied Frames	Unoccupied Frames	% Occupied
R-18Mg^{2+}	2252	49 761	4
X-18Mg^{2+}	30 386	21 619	58
C-18Mg^{2+}	25 985	26 025	50

3.2 Results

3.2.1 General Stability of tRNA Simulations

To optimize tRNA simulation conditions, seven tRNA models (starting from PDBID: 3L0U (Byrne et al., 2010)) that had different initial placements and concentrations of Mg^{2+} ions were originally minimized and equilibrated. Previous MD studies performed on tRNA in complex with EF-Tu have also tested the effect of various ion concentrations on tRNA stability using 11, 18, or 26 Mg^{2+} ions in the simulations (Eargle et al., 2008). The following models generated for this thesis project contained 0, 4, or 18 Mg^{2+} ions in order to maintain the minimum number of Mg^{2+} binding sites in tRNA (i.e. 4), as well as to test the effect of additional Mg^{2+} ions on tRNA stability. Some of the models contained 4 Mg^{2+} ions placed directly within the 3D tRNA structure in order to mimic the strong Mg^{2+} binding sites that have previously been documented, while others contained Mg^{2+} ions randomly placed 5 Å away from the tRNA. The approximate concentration of Mg^{2+} in the models containing 4 or 18 ions was 10 mM or 50 mM, respectively. This is higher than cellular conditions but is similar to the increased concentration of divalent cations expected to be found near nucleic acid (Cowan, 1995; Eargle et al., 2008).

First, the MeRNA database and the Protein Data Bank were used to find 28 structures of tRNA free in solution or bound to protein that contained Mg^{2+} ions and therefore potential binding sites. All structures were superimposed and placed in a three-dimensional grid with 3x3x3 Å grid spacings. The probability of an Mg^{2+} ion being found in each grid box in or around the tRNA was calculated. Figure 3.3 shows the four highest probability positions of Mg^{2+} ions located within this grid (red spheres), with each position being occupied in more than 10% of the aligned structures. The coordinates for these positions were used for the X-4Mg^{2+} and X-18Mg^{2+} models. Second, the cionize package in VMD was also used to identify potential Mg^{2+} ion binding sites by placing Mg^{2+} ions at positions of highest negative charge density. The first four ions placed by this method (Fig. 3.3, grey spheres) were used for the C-$4Mg^{2+}$ and C-$18Mg^{2+}$ models. The potential Mg^{2+} binding sites identified by crystal structures and by negative charge density were all located in the elbow region, however, the two placement methods did not generate any redundant positions. Finally, the autoionize package in VMD was used to randomly place four Mg^{2+} ions for the R-$4Mg^{2+}$ model and 18 Mg^{2+} ions for the R-$18Mg^{2+}$ model. tRNA was also modeled without the addition of Mg^{2+} ($0Mg^{2+}$).

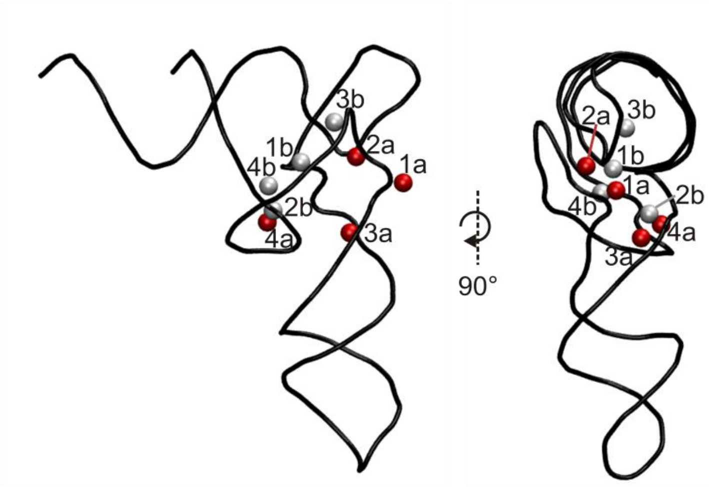

Figure 3.3 Initial placement of Mg^{2+} for tRNA simulations. Four highest probability Mg^{2+} ion binding sites were identified based on tRNA crystal structures from the Protein Data Bank in a 3x3x3 Å grid (red spheres). Numbers 1a-4a indicate highest to lowest probability positions. Grey spheres indicate the first four Mg^{2+} ions placed by cionize in VMD based on negative charge density, labeled 1b-4b in the order of placement.

The relative energy of the seven models after minimization and equilibration was lower for the three models that contained 18 Mg^{2+} ions than the models that contained 0 or 4 Mg^{2+} ions, regardless of ion placement method. For this reason, the models containing 4 Mg^{2+} ions were not used for simulation. Additionally, previous studies have found that less than 4 Mg^{2+} ions led to unwinding of the anticodon stem of tRNA within 20 ns of simulation and that little difference in tRNA stability was seen when either 18 or

26 Mg^{2+} ions were used in the model (Eargle et al., 2008). The following four models were selected for simulation up to 40 ns: 0Mg^{2+}, R-18Mg^{2+}, X-18Mg^{2+}, and C-18Mg^{2+}. Thereby, the effect of the three Mg^{2+} ion placement strategies could be tested in terms of stability of the tRNA, perturbations to the structure, and differences in secondary structural elements and long-range tertiary contacts.

Visual inspection of all four simulations revealed that the L-shaped tRNA structure was generally maintained for all four simulations. However, a slight distortion in the acceptor stem was seen in the last 15 ns of the 0Mg^{2+} and R-18Mg^{2+} simulations and in last 5 ns of the C-18Mg^{2+} simulation. The helical pitch of the stem unwound into an extended ladder-like conformation, though this distortion was localized to the acceptor stem only and did not cause changes to the rest of the tRNA structure. The 0Mg^{2+} and R-18Mg^{2+} simulations showed transient unwinding of the acceptor stem as early as 10 ns into the simulation, however the unwound conformation was not long-lived until 25 ns into the simulation. Recently it has been reported that a modified force field was necessary to preserve helical properties of the preQ$_1$ riboswitch in simulations longer than 100 ns (Banáš et al., 2012). It could be that the force field used for our simulations was not as accurate at maintaining the A-form helix in tRNA for long simulation times, and that the unwinding could be due to biases in the force field that cause the formation of ladder-like conformations. However, since helix unwinding was not seen for the X-18Mg^{2+} simulation, the force field can still be considered suitable for simulations of this length and unwinding may be attributed to differences between the four models. In any case, the effect of different force fields should be examined for longer simulations of tRNA.

The RMSD of backbone atoms was calculated for each tRNA simulation as a measure of the global flexibility of the tRNA (Fig. 3.4A). A sharp increase in the RMSD for all simulations occurred in the first nanosecond, followed by stabilization after 5 ns. All simulations were similar in their deviation from the starting structure, with an average RMSD around 4.5 Å. This value was higher than what was seen for the TruB simulations (Chapter 2), but this could in part be due to the more extended form of tRNA as compared to the globular form of TruB. The acceptor arm and anticodon arm would be expected to exhibit higher flexibility than the globular domains of TruB, and therefore the higher RMSD was not surprising. Additionally, simulations of an EF-Tu-tRNA complex with the addition of 18 Mg^{2+} ions gave RMSDs for the tRNA of 3.5-4 Å (Eargle et al., 2008), deviations of a similar size to those seen here. Interestingly, the $0Mg^{2+}$ simulation behaved much like the other simulations, despite the lack of Mg^{2+} ions. All simulations sampled similar minimum and maximum RMSD conformations, with the C-$18Mg^{2+}$ simulation showing the largest range between minimum and maximum RMSD values. Backbone deviations in the $0Mg^{2+}$, R-$18Mg^{2+}$, and C-$18Mg^{2+}$ simulations were not different from the X-$18Mg^{2+}$ simulation, despite the previously mentioned helix unwinding. This unwinding therefore did not significantly affect the global stability of the different tRNA models. Since the simulations appeared to stabilize after 5 ns, further analysis was only performed on the data from 5-40 ns.

RMSF calculations were used to reveal which regions of the tRNA were the most flexible. As expected, the most flexible regions of the tRNA were the 5’- and 3’-ends and the nucleotides in the anticodon loop (Fig. 3.4B). These nucleotides had RMSF values greater than 4 Å in all simulations. The nucleotides in the D-arm and T-arm showed the

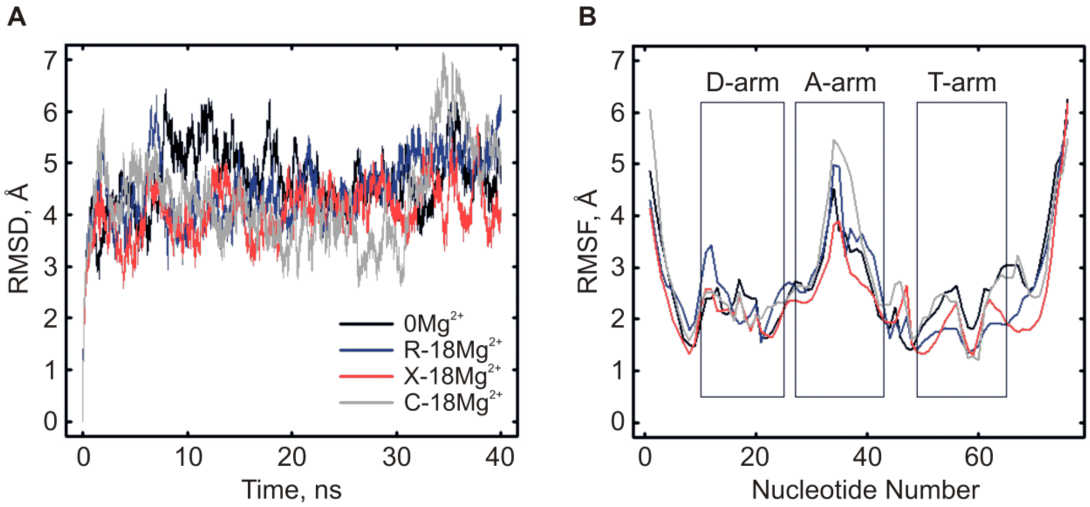

Figure 3.4 RMSD and RMSF of tRNA simulations. Four different tRNA models were simulated at 310 K for 40 ns, including $0Mg^{2+}$ (black), $R\text{-}18Mg^{2+}$ (blue), $X\text{-}Mg^{2+}$ (red), and $C\text{-}18Mg^{2+}$ (grey). (A) RMSD for the tRNA backbone atoms. (B) RMSF of tRNA nucleotides from 5-40 ns of simulations. Boxed areas indicate the nucleotides in the D-arm, anticodon arm (A-Arm), and T-arm.

lowest RMSF values, with almost all nucleotides having an RMSF less than 3 Å. In fact, when the anticodon stem and accepter stem were omitted from the RMSD calculation, the RMSD was more frequently less than 4 Å (Appendix Fig. A.1), which is more comparable to previously reported RMSD values for tRNA and those calculated for TruB. The D-arm and T-arm nucleotides were the most stable during the simulation and most of the fluctuations contributing to the high RMSD can be attributed to motions in the anticodon stem and acceptor stem (Appendix Fig. A.2). Since at least three long-range tertiary interactions between the D- and T-loops in the elbow region have been identified through crystal structures (Byrne et al., 2010), stability within this compact region of the tRNA was expected. Again, the $0Mg^{2+}$ simulation showed similar RMSF values when compared to the simulations that contained 18 Mg^{2+} ions. For this reason, it was unlikely that the Mg^{2+} ions contributed significantly to the overall stability of the tRNA molecule *in silico*. These calculations indicate that tRNA was stable during these simulations, regardless of the presence or absence of Mg^{2+} ions.

3.2.2 The Role of Mg^{2+} in tRNA Simulations

To further examine structural effects due to Mg^{2+} ions during tRNA simulation, specific secondary and tertiary interactions were analyzed for each of the four simulations. In particular, the number of A-U and G-C Watson-Crick hydrogen bonds present and the presence of three long-range tertiary interactions between the D- and T-arms (G15-C48, G18-U55, and G19-C56) were monitored throughout the simulations as a measure of how well the 3D tRNA structure was maintained.

In the predicted secondary structure of *E. coli* $tRNA^{Phe}$, there are 4 potential Watson-Crick A-U base pairs and 16 potential G-C base pairs, all of which are seen in the crystal structure of unmodified $tRNA^{Phe}$ (PDBID: 3L0U). The distance between each of the hydrogen and hydrogen bond acceptor atoms was calculated to determine the number of these atom pairs within 2.7 Å for each frame of the simulation. A cut-off of 2.7 Å was chosen as it is the sum of the van der Waals radii of oxygen and hydrogen (Klein, 2006), and is consistent with the predicted distance for a moderate to weak hydrogen bond (Jeffrey, 1997). Both of the hydrogen bonds in an A-U base pair and all three of the hydrogen bonds in a G-C base pair were measured separately. If at least one of the hydrogen bonds making up the base pair was found to be within 2.7 Å, the base pair was considered present during that simulation frame, regardless of the presence of the other hydrogen bond(s). Shown in Fig. 3.5 is only the N1-H3 bond of adenine and uridine, respectively, and the O6-H41 bond between guanine and cytidine, respectively, since the other hydrogen bonds from each base pair gave similar distributions. The average number of A-U base pairs in the four simulations was approximately 3-4, which

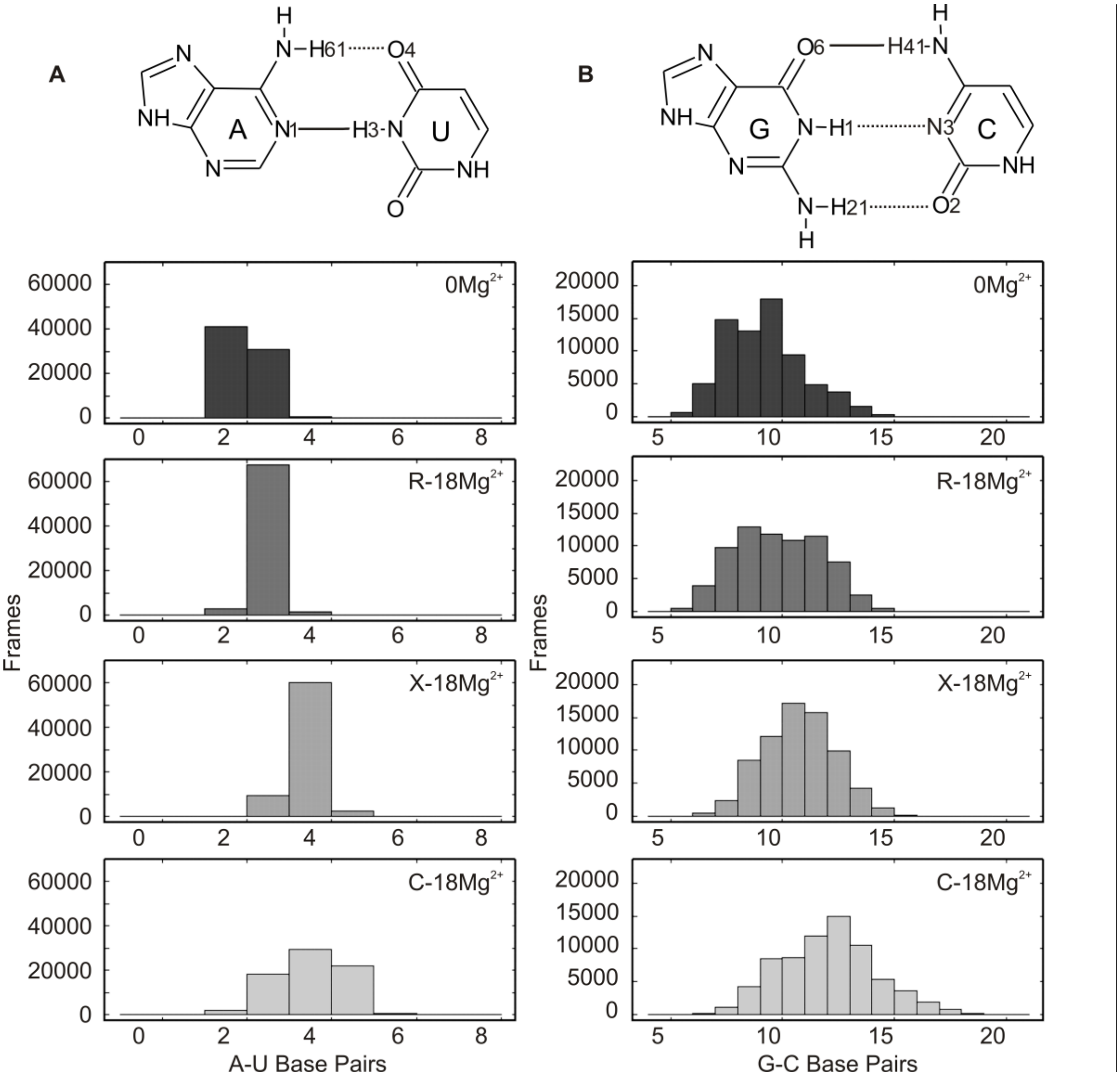

Figure 3.5 Number of Watson-Crick base pairs during tRNA simulations. Each hydrogen bond distance was individually measured for each potential A-U and G-C base pair throughout the simulation. If the respective hydrogen bond donor and acceptor atom were within 2.7 Å, the bond was counted as being base-paired. The script used to measure the distances treated each hydrogen bond of the base pair separately and the data represented in the histograms corresponds to the hydrogen bond denoted by a solid line in the top panel. (A) A-U and (B) G-C Watson-Crick base pairs are shown for $0Mg^{2+}$ (black), $R\text{-}18Mg^{2+}$ (dark grey), $X\text{-}18Mg^{2+}$ (medium grey), and $C\text{-}18Mg^{2+}$ (light grey).

indicates that at least 75% of the predicted A-U base pairs were present through a majority of the simulation time. Similarly, approximately 10-12 G-C base pairs (60-75% of the predicted number) were commonly found in all four of the simulations.

In some frames of the X-18Mg^{2+} and C-18Mg^{2+} simulations, more than the predicted 4 A-U and 16 G-C base pairs were present (up to 6 or 19 base pairs, respectively). This was not surprising since the coaxial stacking of the L-shape of tRNA brings in close proximity nucleotides that would not be predicted by secondary structure to be capable of base pairing in three dimensions (e.g. D- and T-arm nucleotides); however, the calculation performed here was unable to distinguish these kinds of base pairs. Alternatively, the hydrogen bond angle could also be considered in order to decrease the contribution of non-optimal hydrogen bonds that may have been recognized using this method. In general, the X-18Mg^{2+} and C-18Mg^{2+} simulations were found to have a higher number of A-U and G-C base pairs than the 0Mg^{2+} and R-18Mg^{2+} simulations. Though this difference constituted only 1-2 base pairs, these could represent base pairs that were slightly weaker in the 0Mg^{2+} and R-18Mg^{2+} simulations. Hence, placing Mg^{2+} ions within the tRNA helped to stabilize some of these weaker base pairs in the tRNA. However, this stabilizing effect was not seen in the RMSD or RMSF calculations and may therefore reflect local stabilization rather than global stabilization. Interestingly, having 70% of the predicted base pairs with a moderately strong hydrogen bond was sufficient for maintaining the tRNA structure *in silico*.

There are three tertiary interactions identified in the crystal structure of unmodified *E. coli* $tRNA^{Phe}$ between D- and T-arm residues in $tRNA^{Phe}$: G15-C48, G18-U55, and G19-C56. Base pairs between these nucleotides were monitored in a similar way as for A-U and G-C base pairs; the distance was measured between the hydrogen bond donor and acceptor atom in each base pair interaction. A cut-off of 2.7 Å was again used to define a moderately strong hydrogen bond. Figure 3.6 shows the distance between H21 of guanine and O2 of uridine for the interaction between G18 and U55. All four simulations showed a single population centred around 2.0 Å, indicating the presence of a strong hydrogen bond between G18 and U55 throughout a majority of each simulation. A similar result was obtained for the G15-C48 and G19-C56 base pairs (Appendix Figs. A.3, A.4). Since no difference in the stability of the three examined tertiary interactions in the four simulations was observed, the initial concentration and

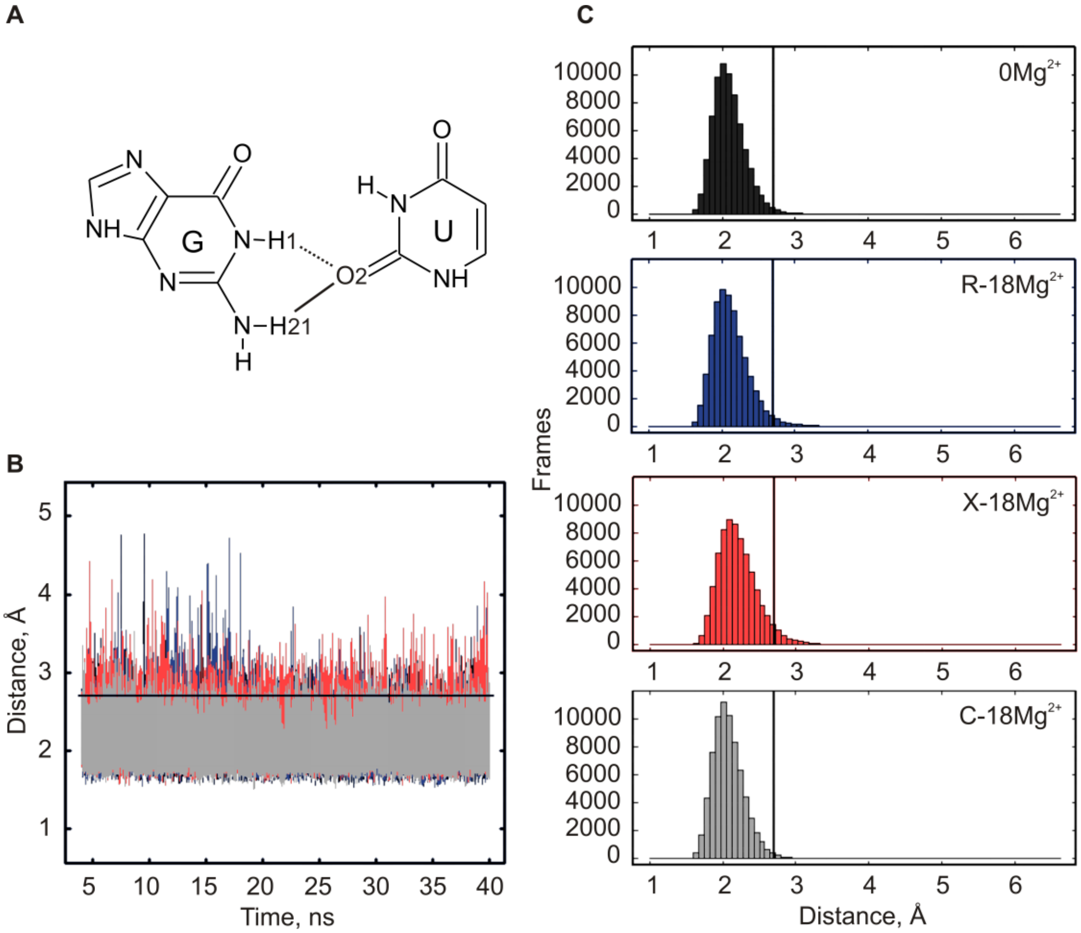

Figure 3.6 Stability of the G18-U55 tertiary interaction. (A) The potential hydrogen bonds in the *trans* Watson-Crick sugar-edge G18-U55 base pair. Data in (B) and (C) represent the hydrogen bond denoted by the solid line. (B) Time-resolved changes and (C) histogram of the distance between G18 and U55, as defined by the solid line shown in (A) for the simulation of $0Mg^{2+}$ (black), R-$18Mg^{2+}$ (blue), X-Mg^{2+} (red), and C-$18Mg^{2+}$ (grey). The black bar indicates a distance of 2.7 Å, the cut-off used to define a hydrogen bond.

placement of Mg^{2+} ions did not significantly affect these long-range interactions. Together with the data on the number of base pairs present during the simulations, Mg^{2+} placement and concentration was not important for preserving the overall shape and tertiary interactions of tRNA during simulation.

3.2.3 High Occupancy Mg^{2+} Sites

Previous RNA MD studies have defined four different classes of Mg^{2+} ions based on the measured distance from the RNA. Of particular interest are inner sphere ions that make direct contacts with the RNA and are found within 3 Å of the RNA, and outer sphere ions that are found within 5 Å of the RNA (Hayes et al., 2012). Here, the number of inner and outer sphere Mg^{2+} ions was monitored throughout each simulation (Fig. 3.7). As previously seen, the outer sphere was always highly populated, often containing 30-50% of all Mg^{2+} ions in the simulations (6-9 Mg^{2+}). Although most of the Mg^{2+} ions were not initially placed closer than 5 Å to the tRNA, a number of ions very quickly found residence within the structure in both the outer and inner sphere. The R-18Mg^{2+} simulation did not have a highly populated inner sphere, with only a few simulation frames having one inner sphere Mg^{2+} ion. However, a significant number of Mg^{2+} ions populated the outer sphere as early as 5 ns into the simulation. The X-18Mg^{2+} simulation showed at least two inner sphere ions and an average of 8-9 outer sphere ions through the majority of the simulation. The inner sphere ions corresponded to 2a and 3a (as numbered in Fig. 3.3) that were initially placed within the model. Likewise, the four inner sphere ions seen in the C-18Mg^{2+} simulation corresponded to the four Mg^{2+} ions originally placed in the model. This speaks toward the validity of placing Mg^{2+} ions based on negative charge density, as these ions were apparently highly coordinated with the tRNA.

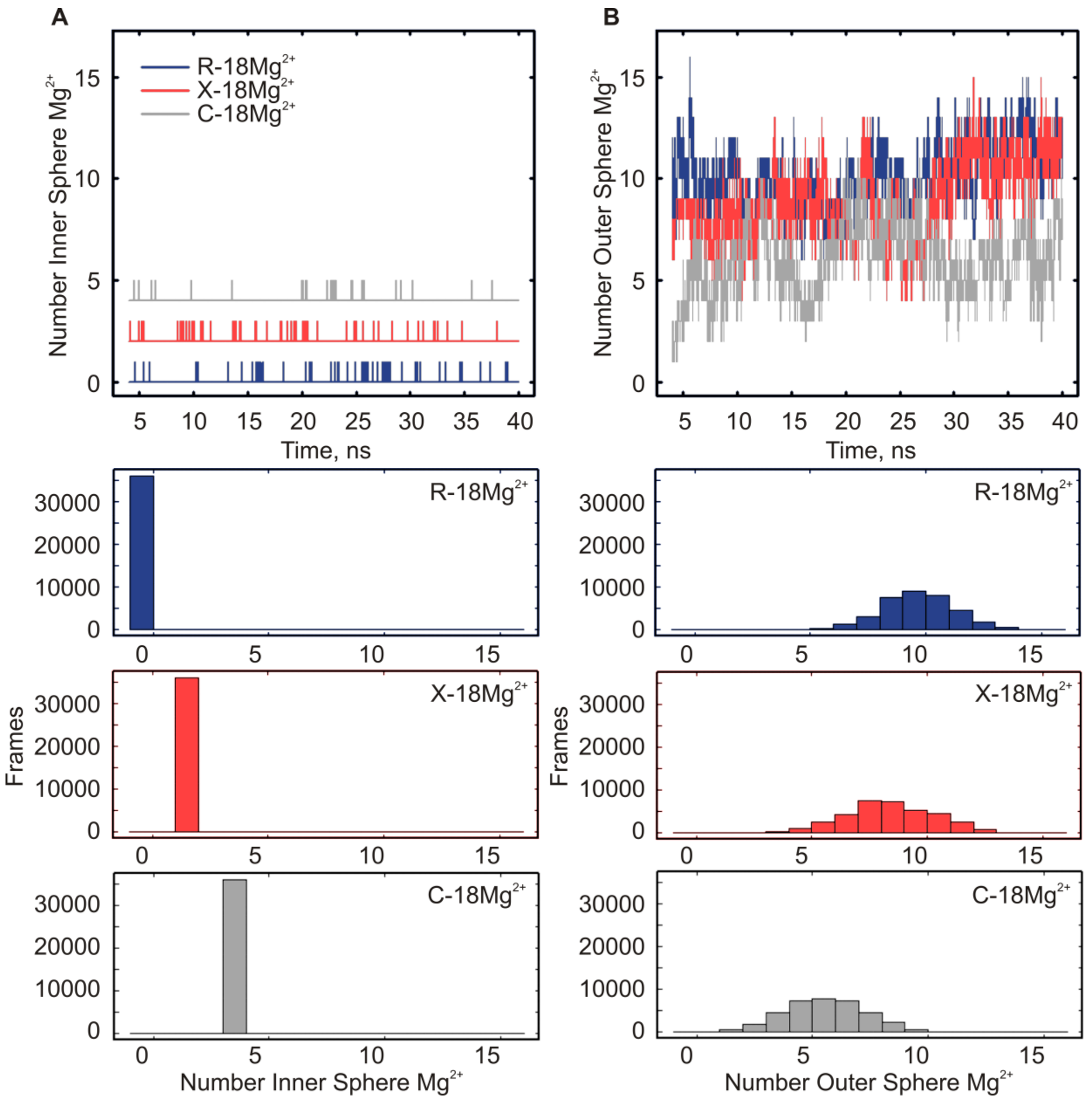

Figure 3.7 Inner and outer sphere Mg^{2+} ions. The number of Mg^{2+} ions found in the (A) inner sphere (within 3 Å) and (B) outer sphere (between 3-5 Å) of the tRNA. Time resolved count of inner and outer sphere ions are shown in the top panels for the R-$18Mg^{2+}$ (blue), X-$18Mg^{2+}$ (red), and C-$18Mg^{2+}$ (grey) simulations, and individual histograms of the data are shown in the bottom panels (same color scheme).

Interestingly, the C-$18Mg^{2+}$ simulation showed the lowest population of outer sphere Mg^{2+} ions, which could be due to the presence of a more highly populated inner sphere that did not require a highly populated outer sphere to balance the large negative charge density of the tRNA. In fact, the sum of inner and outer sphere ions was essentially the same for all three simulations (9-10 Mg^{2+}). Therefore, inner and outer sphere Mg^{2+} ions appeared to be important, though the lack of inner sphere ions in the R-18Mg^{2+} simulation suggest that Mg^{2+} ions don't necessarily need to be bound within the structure as long as there are enough ions in the outer sphere.

The presence of inner and outer sphere ions in all three simulations was further examined by generating occupancy maps to indicate where in the tRNA structure an Mg^{2+} ion was likely to be found. Using the VolMap plugin in VMD, the models were analyzed for areas of high Mg^{2+} occupancy. Figure 3.8 shows the regions of at least 10% Mg^{2+} occupancy through all frames of each simulation. All simulations show regions of high Mg^{2+} occupancy in the elbow region of the tRNA, as was expected due to the high negative charge density in this region. However, the pattern of high Mg^{2+} occupancy was distinct for each of the three simulations. In agreement with the counting of inner and outer sphere ions, Mg^{2+} ions initially placed not closer than 5 Å to the tRNA in the R-18Mg^{2+} simulation were able to find residence within the 3D structure. Four main regions of high occupancy within the elbow region were found for this simulation, three of which extended toward the acceptor stem with one found near the base of the elbow closer to the anticodon stem. In the X-18Mg^{2+} simulation, a number of high occupancy positions were also found in the elbow region, but were positioned closer to the D-loop. Although the regions of high Mg^{2+} occupancy were found to be near the original ion placement positions for this model, none of the positions directly overlapped. Three of the four Mg^{2+} ions originally placed in this model rearranged to the positions indicated by the occupancy maps. However, one of the four original Mg^{2+} ions (number 1a from Fig. 3.3) moved entirely out of the elbow region after approximately 20 ns and diffused away into the water box of the model before returning to the outer sphere after a few nanoseconds.

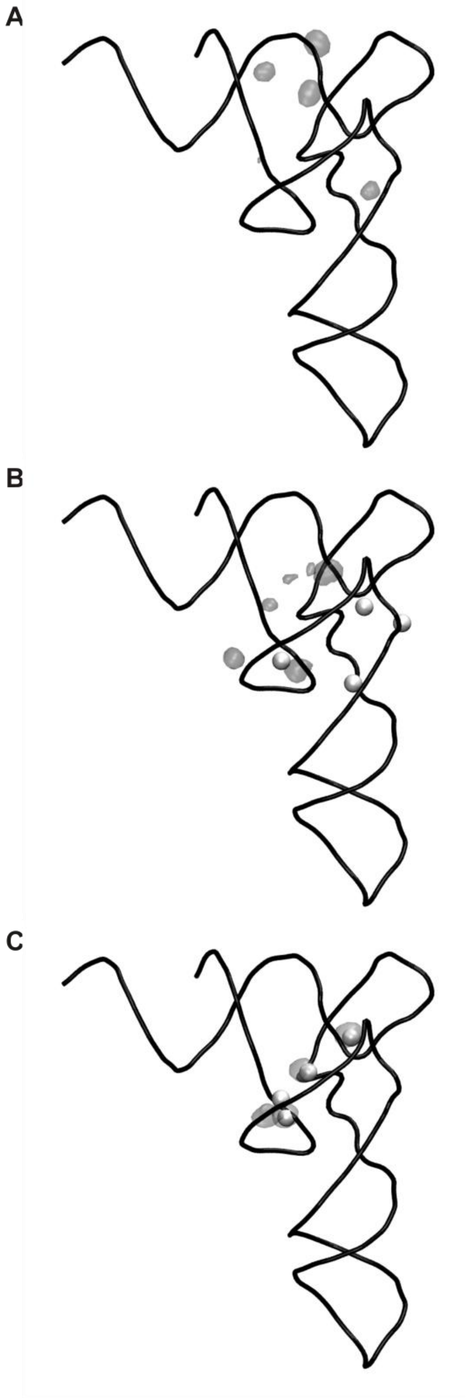

Figure 3.8 High occupancy Mg^{2+} sites. Regions of Mg^{2+} occupancy for more than 10% of the simulation are shown by grey clouds in the tRNA structure for the (A) R-18Mg^{2+}, (B) X-18Mg^{2+}, and (C) C-18Mg^{2+} simulations. The original four Mg^{2+} ions placed for the models are shown as white spheres (B and C).

The high occupancy positions in the C-18Mg^{2+} simulation closely resembled those found in the X-18Mg^{2+} simulation, but were even more tightly clustered in the D-loop region. All four of the original Mg^{2+} ions placed in this model closely corresponded to the regions of high Mg^{2+} occupancy, indicating that these ions did not move significantly from their starting positions and were likely coordinated with the tRNA, which was further evidence that the placement strategy used for this model was valid. In general, the occupancy maps calculated from these three simulations indicate that there are a variety of potential Mg^{2+} binding sites within the elbow region of the tRNA. This also suggests that the binding of Mg^{2+} within tRNA is predominantly non-specific and allows for general stabilization through the shielding of negative charges.

3.2.4 Cluster Analysis of tRNA Simulations

Due to the presence of high Mg^{2+} occupancy sites in the simulations, it was tested whether the presence or absence of Mg^{2+} at a particular position corresponded to local stabilizing effects that could not be seen when the simulation was analyzed as a whole. To do this, the models were placed in a three-dimensional grid with 4x4x4 Å grid spacings to locate the Cartesian coordinates of highest Mg^{2+} residency in each of the three simulations. Simulation frames that contained an Mg^{2+} ion within 3.5 Å of this determined reference position were designated “occupied” and those that did not were considered “unoccupied”. This criterion resulted in a significant number of frames in each cluster for the R-18Mg^{2+}, X-18Mg^{2+}, and C-18Mg^{2+} simulations when the reference position from the X-18Mg^{2+} simulation was used (see Methods, Table 3.2 and Fig. 3.2). Since the lowest number of frames in a cluster was 2252 (R-18Mg^{2+} occupied cluster), a random selection of 2252 frames from each of the other clusters was created for further analysis. Because of the uneven distribution of the R-18Mg^{2+} clusters (2252 frames occupied vs. 49 761 frames unoccupied), the unoccupied cluster inherently resembled the analysis done on the entire simulation. The X-18Mg^{2+} and C-18Mg^{2+} simulations had more evenly distributed occupied and unoccupied clusters and therefore would more reliably reflect differences due to the presence or absence of Mg^{2+}.

New occupancy maps were generated for the occupied and unoccupied clusters (Fig. 3.9) for comparison to the occupancy map from the full simulation. The R-18Mg^{2+} occupied cluster showed an extended distribution of Mg^{2+} occupancy positions that were located in the elbow region and acceptor stem. Similar positions of high Mg^{2+} occupancy were seen in the unoccupied cluster (comparable to the overall occupancy map), but showed much less diffuse regions of high occupancy than for the occupied cluster. The occupied and unoccupied clusters from the X-18Mg^{2+} showed little difference in the high Mg^{2+} occupancy regions and also resembled the occupancy seen for the whole simulation. Likewise, the occupied and unoccupied clusters from the C-18Mg^{2+} simulation showed very few differences between them and did not significantly differ from the occupancy map for the whole simulation. These results indicate that the presence or absence of Mg^{2+} did not greatly influence the position of other Mg^{2+} ions in the simulation.

The occupied and unoccupied clusters were further analyzed for local effects on tRNA structure. As was seen for the whole simulation analysis, no significant difference was observed between the occupied and unoccupied cluster analysis for the three simulations when the number of base pairs (Appendix Fig. A.5) and tertiary interactions were examined (Appendix Figs. A.6-A.8). However, Mg^{2+} occupancy did appear to have

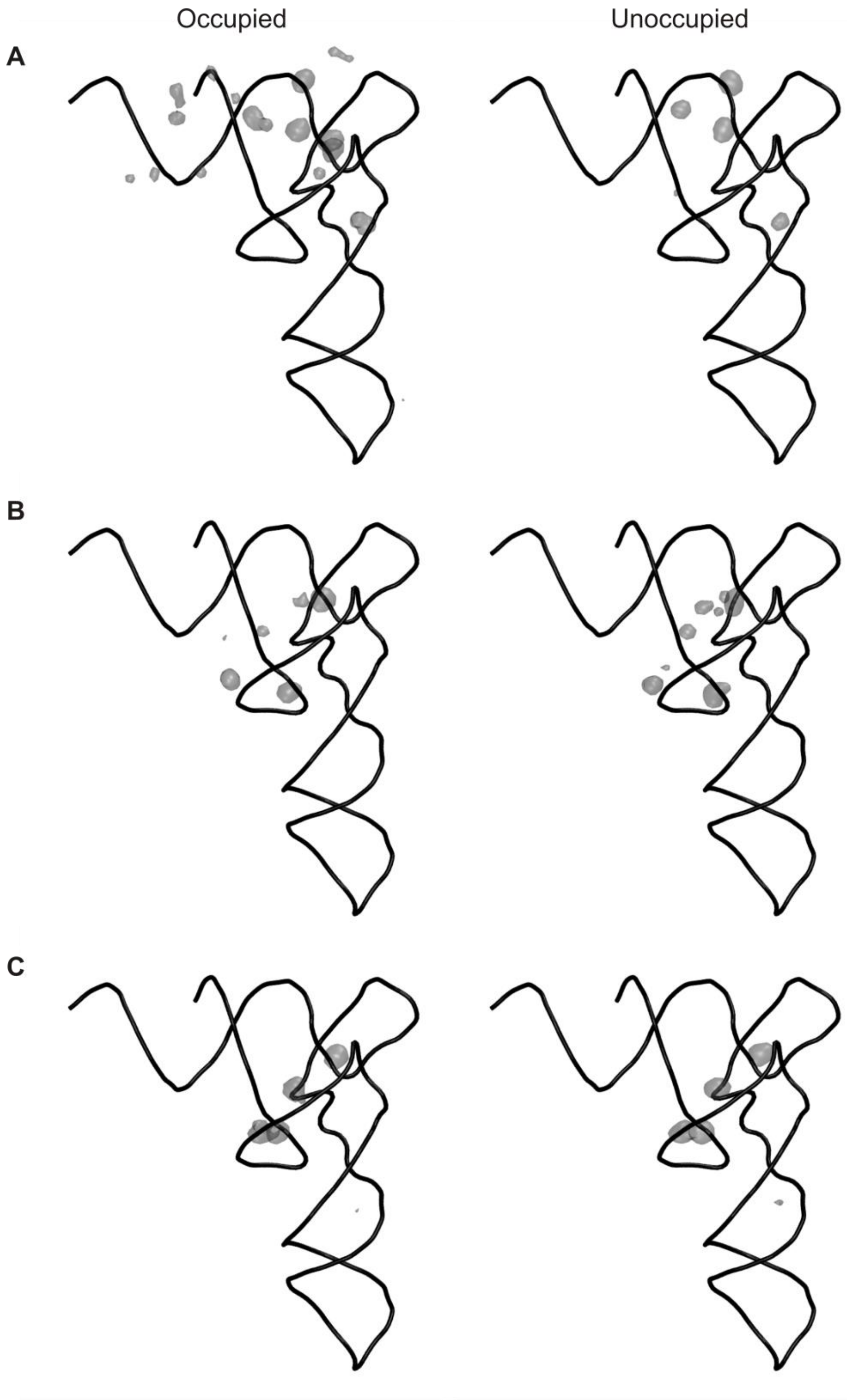

Figure 3.9 High occupancy Mg^{2+} sites of occupied and unoccupied simulation clusters. Regions of Mg^{2+} occupancy for more than 10% of the occupied (left) and unoccupied (right) simulation clusters shown as grey clouds for the (A) R-18Mg^{2+}, (B) X-18Mg^{2+}, and (C) C-18Mg^{2+} simulations. 2252 frames were analyzed in each cluster.

a slight effect on the angle between the tRNA acceptor stem and anticodon stem (elbow angle) when the occupied frames were compared to the unoccupied frames for each of the three simulations. This analysis was done by measuring the angle formed between the phosphate atom of C72 in the acceptor arm, C56 in the elbow, and A35 in the anticodon arm (Byrne et al., 2010). The occupied cluster from the R-18Mg^{2+} simulation showed a relatively narrow distribution of the elbow angle from 58-74° with two populations centered at 64° and 68° (Fig. 3.10A). Conversely, the unoccupied cluster showed a much wider distribution of the elbow angle (55-81°) with only one population centered around 67°. The distribution of the elbow angle for the occupied and unoccupied clusters of the X-18Mg^{2+} simulation spanned a similar range relative to each other, but showed a marked difference in the distribution centres (Fig. 3.10B). The occupied cluster had an average elbow angle near 70° whereas the unoccupied cluster was centered around 76°. In both of these simulations, the occupied cluster appeared to maintain a more compact tRNA structure, in that the elbow angle was slightly smaller than that seen for the unoccupied cluster. This was different from the trend seen for the C-18Mg^{2+} simulation clusters (Fig. 3.10C). In this simulation, while the range of the elbow angle was similar for the occupied and unoccupied clusters, the occupied cluster had a larger average angle of 68° compared to the unoccupied cluster that had an average elbow angle of 66°, indicating that the tRNA was similar or slightly less compact in the occupied cluster than in the unoccupied cluster. The differences seen in the elbow angle of the occupied and unoccupied clusters for the three simulations could indicate that the presence or absence of Mg^{2+} within the elbow region can result in propagation of structural differences that can travel toward the acceptor stem and anticodon stem of the tRNA. Since the ends of

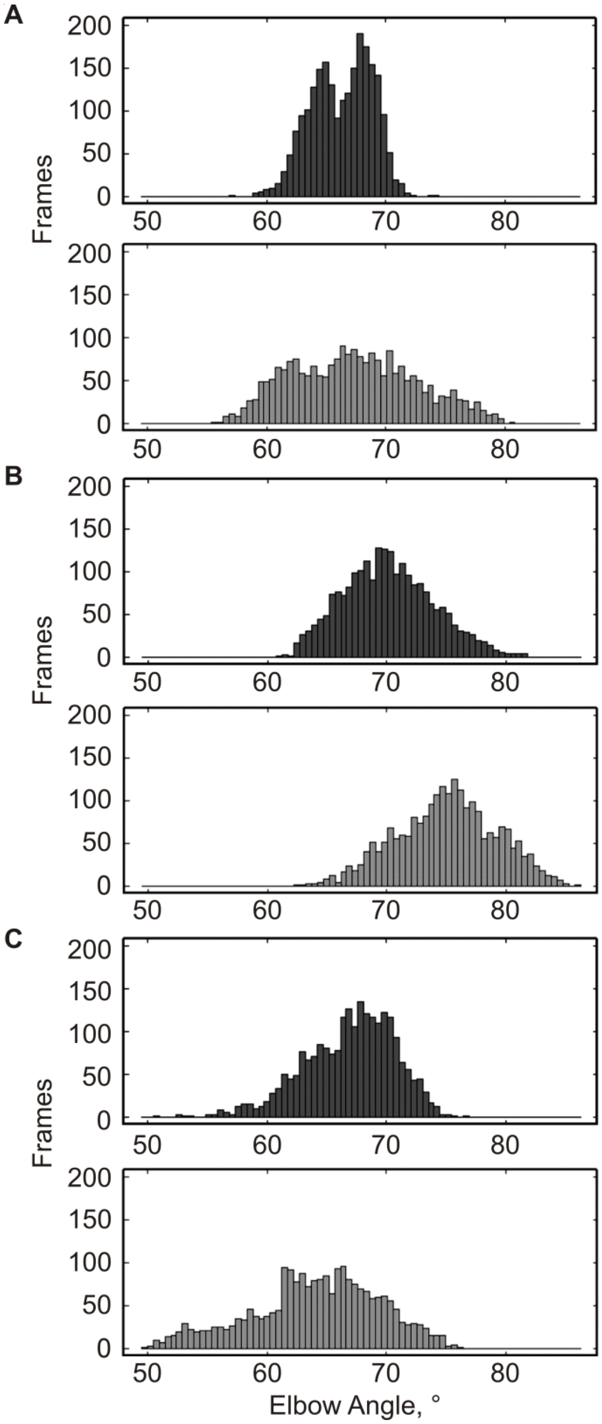

Figure 3.10 tRNA elbow angle in the occupied and unoccupied clusters. The angle between the acceptor arm and anticodon arm of the tRNA (elbow angle) was measured between the phosphate atom of C72, C56, and A35 for both the occupied (black) and unoccupied (grey) cluster for the (A) R-18Mg^{2+}, (B) X-18Mg^{2+}, and (C) C-18Mg^{2+} simulations.

the L-shape are inherently more flexible (as seen by the RMSF calculation, Fig. 3.4B), Mg^{2+} ions in the elbow region may serve to better constrain large movements at the termini of the tRNA by stabilizing the elbow region.

3.3 Discussion

The results presented here examined the role of Mg^{2+} ions in simulations of unmodified tRNA. First, the deviation from the starting tRNA structure through the simulations in the presence of 0 or 18 Mg^{2+} ions was remarkably similar. All simulations stabilized at an RMSD of 4.5 Å, to which fluctuations in the acceptor arm and anticodon arm contributed the most. The number of Watson-Crick base pairs maintained throughout the simulations constituted nearly 70% of the predicted number of base pairs, with the models containing Mg^{2+} placed directly within the tRNA structure showing 1-2 more base pairs than the models that did not have Mg^{2+} placed within the structure. Three tertiary interactions in the elbow region were highly stable in all simulations. Second, up to 50% of the Mg^{2+} ions were found to occupy the inner and outer sphere in each simulation, and the common region of high Mg^{2+} occupancy was the elbow region. Mg^{2+} ions placed within the structure to balance the negative charge density essentially did not move from their starting locations, whereas randomly placed Mg^{2+} ions and those placed based on crystal structure positions showed higher mobility. Finally, the frames from each simulation were sorted into two clusters based on the presence or absence of Mg^{2+} within 3.5 Å of a reference position within the tRNA structure. A slight difference between the clusters was seen in the tRNA elbow angle, where the R-18Mg^{2+} and X-18Mg^{2+} simulations had a smaller elbow angle in the occupied frames, and the C-18Mg^{2+} simulation had a larger elbow angle in the occupied frames. However, there was no apparent difference between the occupied and unoccupied clusters in terms of the number of base pairs, stability of tertiary interactions, or Mg^{2+} occupancy.

3.3.1 tRNA Simulations are Remarkably Stable

All-atoms models of unmodified $tRNA^{Phe}$ were simulated in the absence or presence of Mg^{2+} with different placement strategies. After minimization and equilibration, the three models containing 18 Mg^{2+} ions had the lowest energy, as was expected due to the ability of the divalent cations to shield the high negative charge of the tRNA backbone. Although the energy of the $0Mg^{2+}$ model was the highest after minimization and equilibration, it was the same order of magnitude as the energy of the other models (~3% higher in energy than the models containing 18 Mg^{2+} ions) and visual inspection of the minimization and equilibration steps did not reveal any noticeable changes to the tRNA structure. Therefore, these four models were selected for simulation and further analysis.

The global stability of the models was assessed by computing the RMSD and RMSF after 40 ns of simulation. As expected, the major contribution to the RMSD was from motion in the 5'- and 3'-ends and by the nucleotides in the anticodon loop. This trend has been reported for short simulations of tRNA, as well as tRNA simulated in complex with EF-Tu, where the greatest flexibility was seen in the acceptor stem and anticodon stem and little flexibility was seen in the core of the structure (Auffinger et al., 1999; Eargle et al., 2008). Very little flexibility was seen in the D- and T-arms, likely due to strong interactions between these regions that serve to rigidify the structure. Interestingly, despite the difference in Mg^{2+} concentration and placement strategy, all models exhibited similar global flexibility profiles. This indicates that the three Mg^{2+}

placement strategies were all suitable for generating models that were stable during simulation. Surprisingly, the absence of Mg^{2+} also generated a stable simulation, indicating that tRNA likely adopts an inherently stable structure for *in silico* studies. tRNA *in vitro* is able to adopt the native fold and be recognized by aminoacyl-tRNA synthetase in 2-10 mM Mg^{2+} for at least 60-65% of a tRNA population (Bhaskaran et al., 2012). While Mg^{2+} may help tRNA fold from a denatured state, it may not be necessary for stabilizing tRNA that is already folded. Single-molecule experiments with fluorescently-labeled tRNA have shown that at Mg^{2+} concentrations higher than 0.1 mM, where any strong Mg^{2+} binding sites are likely filled, have no additional stabilizing effects as seen by the distance between the acceptor and anticodon arms (Daher and Rueda, 2012). Further evidence for Mg^{2+}-independence was shown here by the presence of a majority of the predicted secondary structure Watson-Crick base pairs and three tertiary interactions forming strong hydrogen bonds through most of the simulation for all models. Since it was expected that the presence of Mg^{2+} would confer additional stability to the tRNA, closer examinations of the simulations were performed to assess if Mg^{2+} ions had a more specific, rather than global, effect on stabilizing tRNA structural elements.

3.3.2 Mg^{2+} Is Not Required for Maintaining Overall tRNA L-Shaped Structure

The requirement for Mg^{2+} in stabilizing tRNA structure *in silico* was examined by analyzing certain secondary and tertiary structural elements. It was found that when Mg^{2+} ions were initially placed within the tRNA structure, 1-2 additional Watson-Crick base pairs were stabilized during the simulation. However, regardless of Mg^{2+} concentration and position, approximately 70% of the predicted number of base pairs were maintained at a moderately strong hydrogen bond distance of 2.7 Å. This indicates that the helical stability of the four models was quite similar, especially since no fraying of base pairs was seen in any of the simulations. In early MD simulations of DNA, the terminal base pair of a duplex was found to be susceptible to breaking, but improvements to force field parameters and the treatment of electrostatic interactions have made base pair fraying less common (Cheatham and Young, 2000; Cieplak et al., 1997; Young et al., 1997). Still, some MD studies of nucleic acid duplexes impose artificial restraints on the terminal base pair to prevent fraying or limit the analysis to the dynamics of the interior nucleotides and not the terminal nucleotides (Dornberger et al., 2001; Kolaskar and Joshi, 1998). In the simulations performed here, all base pairs in the acceptor stem maintained hydrogen bond interactions through all 40 ns of simulation, including the terminal base pair that would be most likely to break. Additionally, the presence of stable tertiary interactions (G18-U55, G15-C48, and G19-C56) between the D-arm and T-arm were seen in all four simulations. These interactions were very strong, as a majority of the simulation time for all four models showed the hydrogen bond atom pairs to be well within the 2.7 Å cut-off. This suggests that the concentration and position of Mg^{2+} was not important for maintaining the L-shaped structure that brings these nucleotides from the D- and T-arms in close proximity. In general, stable secondary and tertiary structures in the tRNA are formed for the duration of 40 ns simulation time, regardless of Mg^{2+} concentration and initial placement. Surprisingly, despite the extent of literature documenting the importance of Mg^{2+} for RNA folding and stability (Bowman et al., 2012; Misra and Draper, 2002; Serebrov et al., 2001), the $0Mg^{2+}$ simulation did not exhibit any drastic structural defects when compared to the simulations containing Mg^{2+}. Since larger structural rearrangements are more likely to occur on a microsecond time scale, it could be that effects of Mg^{2+} on tRNA structure may only be observed in longer simulations and the importance of Mg^{2+} was not yet apparent in these simulations. Hence, longer simulation times will be necessary to gain a more complete understanding of potential structural effects Mg^{2+} has on tRNA *in silico*. Alternatively, the different Mg^{2+} requirements in our *in silico* studies and in *in vitro* tRNA folding studies could be explained by a role of Mg^{2+} ions in the folding process, but not for the stability of the final, folded product.

3.3.3 Non-Specific High Occupancy Mg^{2+} Sites

Different layers of Mg^{2+} interaction with RNA have been seen in MD simulations of the SAM-1 riboswitch (Hayes et al., 2012). Similar layers of inner and outer sphere ions, defined by the distance from the tRNA, were also populated in the tRNA simulations presented here. Interestingly, in the R-18Mg^{2+} simulation that did not start with any Mg^{2+} closer than 5 Å to the tRNA, 50% of the Mg^{2+} ions moved within 5 Å to the tRNA in just a few nanoseconds and remained in close proximity to the tRNA for a majority of the simulation. In fact, all three simulations containing Mg^{2+} revealed that about half the Mg^{2+} ions were found in the inner or outer sphere. When the specific occupancy of these ions were mapped onto the tRNA structure for each simulation, it was found that the elbow region in general was most frequently occupied by Mg^{2+} ions, but that the specific locations of occupancy varied between each simulation (Fig. 3.8). This suggests that while shielding of negative charge density in the elbow region is important, the positioning of Mg^{2+} in the inner and outer sphere is not specifically defined.

The areas of high Mg^{2+} occupancy in the X-18Mg^{2+} and C-18Mg^{2+} simulations closely correlated to the original positions of the Mg^{2+} ions used for those models. It was not expected that these ions would show highly dynamic motions or diffusion away from the tRNA due to their tight placement within the 3D structure of the tRNA. Accordingly, all four Mg^{2+} ions specifically placed in the C-18Mg^{2+} simulation essentially did not move from their starting positions, suggesting that their placement was highly coordinated by the tRNA and important for balancing the negative charge density in that region. However, the Mg^{2+} ions placed based on crystal structure position (X-18Mg^{2+} model) showed more movement from their starting positions, with three of the four ions shifting to a nearby position (Fig. 3.8 and Mg^{2+} RMSF < 2 Å, Appendix Fig. A.9), and one of the Mg^{2+} ions moving away from its starting position within the tRNA into the surrounding water box where it then diffused in and out of the outer sphere. Therefore, the placement of these ions was perhaps not as optimal as it was in the C-18Mg^{2+} simulation. In the R-18Mg^{2+} simulation, a number of Mg^{2+} ions placed more than 5 Å from the tRNA moved into the core of the tRNA structure in the first few nanoseconds of simulation. Interestingly, all three simulations showed that Mg^{2+} ions were capable of diffusing in and out of the tRNA structure. This suggests that some Mg^{2+} ions associate loosely with the tRNA and can exchange with other ions in solution. In terms of *in vivo* interactions between Mg^{2+} and RNA, the ability for Mg^{2+} to enter the tRNA structure suggests that the effect of Mg^{2+} on tRNA folding can be conferred during folding, but also after the tRNA has adopted its L-shaped conformation.

3.3.4 Mg^{2+} Effects on Local Structure Properties

To better analyze the potential effect Mg^{2+} ions had on the stability and structure of tRNA, the simulations were sorted into clusters based on Mg^{2+} occupancy at a particular position. The reference position was taken from the X-18Mg^{2+} simulation that was occupied in 58% of all frames and was similar in position to Mg^{2+} 2a (Fig. 3.3). Using this position to cluster the remaining simulations allowed us to analyze a significant number of frames in the occupied and unoccupied clusters for all three simulations. Analysis of these two clusters revealed very few differences between the occupied and unoccupied frames. The occupancy maps, number of base pairs, and stability of the tertiary interactions were similar for the occupied and unoccupied cluster from the same simulation (Appendix Figs. A.5-A.8). Additionally, no significant difference was seen when comparing the occupied and unoccupied clusters from different simulations. This is further evidence that the presence or absence of Mg^{2+} did not greatly influence tRNA structure *in silico* and that all three Mg^{2+} placement methods used for generating the models were suitable for producing stable tRNA simulations, in particular because three different placement methods resulted in high Mg^{2+} occupancy near the same coordinates. Since this analysis was only based on one Mg^{2+} position identified from one of the simulations, we cannot rule out that further analysis on frames clustered using a different reference position may indicate other parameters that may be affected by the presence of a specific Mg^{2+} position within the tRNA.

Though high Mg^{2+} occupancy sites were found for the three Mg^{2+}-containing simulations, it was not clear how these sites could affect tRNA structure or stability. One structural parameter that showed a slight difference between the occupied and unoccupied cluster was the elbow angle of the tRNA. The change in the elbow angle may reflect the helix unwinding seen in the R-18Mg^{2+} and C-18Mg^{2+} simulations. Additionally, the differences in the high Mg^{2+} occupancy of the acceptor stem and elbow regions between the three simulations could also have influenced the change in the elbow angle (and helix unwinding) of the tRNA. The more rapid unwinding of the acceptor stem seen in the $0Mg^{2+}$ simulation could indicate that Mg^{2+} is important for maintaining the helical structure. Without Mg^{2+} ions, the acceptor stem may become too flexible and lead to unwinding of the helix. On the other hand, the occupancy map for the R-$18Mg^{2+}$ simulation showed higher occupancy of Mg^{2+} that extended into the acceptor stem. The presence of too many Mg^{2+} ions could have caused distortion of the acceptor stem that made it more likely to adopt an unwound conformation than if the Mg^{2+} ions were located closer to the elbow region, as seen in the X-$18Mg^{2+}$ and C-$18Mg^{2+}$ simulations. A combination of adjusting the force field parameters and fine-tuning Mg^{2+} ion positions may be useful in preventing the acceptor stem from unwinding in longer simulations.

3.3.5 Application to Other RNA Simulations

The simulations performed here indicate that tRNA is a remarkably stable molecule that maintains its three-dimensional shape with little requirement for the stabilizing effects of Mg^{2+}. Even though the force field used for these simulations was optimized more for protein simulation than nucleic acid simulation, drastic rearrangements or distortions in the tRNA were not observed. Initial placement of Mg^{2+} ions did not greatly influence the global or local stability of the tRNA. While the Mg^{2+} ions placed based on crystallographic data showed little movement likely due to coordination by the tRNA, the Mg^{2+} ions positioned by charge density were even less mobile and more highly coordinated. This suggests that although all of the tested placement methods for Mg^{2+} generated stable tRNA simulations, balancing the negative charge density of an RNA backbone may be the most important factor to consider when placing ions in a nucleic acid simulation. However, the ability of Mg^{2+} ions to artificially perturb nucleic acid structure must also be considered, and should be analyzed by comparison to crystal structures. Since there were no apparent distortions to the structure due to the placement of Mg^{2+} in the C-18Mg^{2+} simulation, or in any of the other simulations, it can be assumed that the resultant Mg^{2+} locations and the number of Mg^{2+} ions added to the elbow region did not cause any unnatural perturbations during our tRNA simulations.

The apparent lack of a requirement for Mg^{2+} was quite surprising and may reflect the need for longer simulation times. Whether this Mg^{2+}-independence holds true for other RNA molecules is an interesting question. Mg^{2+} binding has been found to affect the width of the major and minor grooves of DNA, but these effects are only seen locally and do not propagate throughout the structure (Hud and Polak, 2001). On the other hand, Mg^{2+} ions were found to be considerably more important for simulations of the internal loop E fragment of bacterial 5S rRNA, where the absence of the 5 Mg^{2+} ions identified in crystal structures caused larger RMSDs, fluctuations in the sugar pucker, and transient unstacking events (Auffinger et al., 2003). Therefore, the findings presented here may not be universal for nucleic acid simulations, especially for the simulation of nucleic acids with highly defined Mg^{2+} binding sites. However, for those molecules that have undetermined binding sites for Mg^{2+}, the placement of Mg^{2+} ions to balance the negative charge can be used to give relatively stable simulations.

3.3.6 Simulation of tRNA-Protein Complexes

The work reported here presents a significant step forward in our understanding of how to model, simulate, and analyze nucleic acids in MD simulations. Since an unmodified tRNA was used here, it would be interesting to analyze differences in structure or stability that occur when the tRNA is partially or fully modified. Additionally, this information can be used in conjunction with protein MD simulations to generate a model of TruB bound to tRNA for simulation. Since simulations have now been performed for both TruB and tRNA alone, the dynamics of a tRNA-TruB complex can be easily compared to the data that has already been generated. In designing the complex, the crystal structure of TruB bound to a short RNA piece mimicking the T-arm (PDBID: 1K8W) can be used to position a full-length tRNA structure within the active site by superimposition of the two RNAs. Different concentrations of Mg^{2+} can also be tested in this simulation, however studies have shown that TruB can still perform pseudouridylation in the absence of Mg^{2+} (Nagamatsu and Miyazawa, 1983; Nurse et al., 1995). Relevant dynamics that could be studied with a tRNA-TruB simulation include the base-flipping of the target uridine and the two subsequent bases into TruB's active site, interactions between the thumb loop of TruB with the tRNA, or potential interactions between the PUA domain of TruB with the acceptor stem of the tRNA. These simulations could provide information on how TruB recognizes its full-length tRNA substrate, which has not yet been possible through crystallographic studies, or about the potential ability of TruB to rearrange the tRNA structure according to a proposed role of TruB to act as an RNA chaperone (Gutgsell et al., 2000).

Chapter 4 – Conclusions and Future Directions

4.1 TruB Catalytic Mechanism

Although pseudouridine was first discovered over 50 years ago, it has only been in the last decade or more that substantial leaps in our knowledge about the function and formation of this modified nucleotide have taken place. A great part of this was due to crystallographic studies that have provided us with structural information for all pseudouridine synthase families. Alignment of these structures have revealed a common catalytic core domain with three conserved residues in the active site: an aspartate, an aromatic residue (tyrosine or phenylalanine), and a basic residue (arginine or lysine). So far, studies have focused on the function of the conserved aspartate, revealing that it is essential for catalysis. The aromatic residue is also thought to contribute to catalysis by helping to position the target uridine or by acting as a general base in the final steps of catalysis to deprotonate the isomerized product (Phannachet et al., 2005). However, no biochemical studies have explored the role of the basic residue and how it might function in pseudouridine formation. Here, I presented work that showed the importance of R181 in the pseudouridine synthase TruB, as well as interactions with another conserved residue, D90, that likely work together with the catalytic aspartate (D48) to facilitate pseudouridine formation.

Crystal structures have shown that R181 and D48 can form a salt bridge interaction and that D90 makes contacts with C56 of the tRNA (Hoang and Ferré-D'Amaré, 2001; Pan et al., 2003; Phannachet and Huang, 2004). MD simulations of TruB wild type performed here in the absence of RNA showed that R181 and D90 were also able to interact. However, TruB D90N simulations showed that R181 formed a more stable interaction with D48 than with N90, which was not seen in the wild type simulations. This suggests that a careful network of interactions stems from R181 that may be important for catalysis, a finding that was further supported by *in vitro* studies. Though substitutions at positions 90 and 181 did not affect substrate binding or binding kinetics, all TruB variants were severely impaired at catalysis. The chosen substitutions to R181 and D90 would all have disrupted the interaction network seen in the MD simulations, and led to the proposal of a model where interactions between R181 and D90 are important for mediating interactions between R181 and D48. The strength of the R181-D48 interaction must be fine-tuned in order for D48 to be properly positioned by R181, but not prevented from participating in catalysis. Interestingly, the R181-D90 salt bridge interaction seen in the MD simulations was more important for TruB's function than the hydrogen bond interaction between D90 and C56 of the RNA, since TruB D90N that loses the ability to form the salt bridge while maintaining hydrogen bond capabilities was less active than the most active TruB variant, D90E, that was still capable of forming a salt bridge with R181.

Knowing the importance of R181 and D90 for pseudouridine formation by TruB, we can make use of the variants that have been generated to further dissect the steps in catalysis. The TruB variants tested here showed similar rates for k_{app2} compared to TruB wild type, suggesting that base-flipping was likely not affected by the amino acid substitutions. To confirm that k_{app2} corresponds to base-flipping, similar experiments with 2AP-tRNA can be performed with RluA, which also flips its target uridine and two other bases into the active site (Hoang et al., 2006), to see if an additional fluorescence change is also observed. A rapid initial rate has also not been seen in experiments performed with a 3’ fluorescein-labeled tRNA, which would not be expected to give a signal change upon base-flipping (unpublished data, Kothe lab), and therefore further suggests that base-flipping is being monitored with the 2AP-tRNA. Another way to confirm this hypothesis would be to impair base-flipping, perhaps by the substitution of H43, as this residue is thought to stabilize the conformation of the flipped-out nucleotides (Hoang and Ferré-D’Amaré, 2001). TruB H43 variants would therefore not be expected to produce a fluorescence change due to base flipping with a 2AP-tRNA substrate. Since pre-steady-state tRNA binding kinetics have only been analyzed for TruB D90N and R181K, stopped-flow experiments with 2AP-tRNA can be performed for the remaining D90 and R181 variants. It will be interesting to see if three fluorescence changes, as for TruB wild type, are seen for TruB D90E, the most active of the TruB variants. If the third, slow rate does in fact correspond to catalysis and product release as proposed for the TruB wild type time course, then TruB D90E is the most likely variant to also produce this fluorescence decrease and its rate should reflect the measured rate of catalysis. Additionally, titrations with each TruB variant will reveal whether one of the observed fluorescence changes is concentration dependent. This would indicate that the corresponding fluorescence change reflects the bimolecular binding event of tRNA encountering TruB. Potentially, this would allow us to distinguish between the bimolecular binding event and the flipping of bases into TruB’s active site. Thereby, it could be revealed which of these two steps is faster. These studies will help to complete the kinetic mechanism of pseudouridine formation by TruB, and similar experiments can be used to study other pseudouridine synthases that are also known to use a base-flipping mechanism (e.g. RluA).

As outlined in the introduction, there are currently three proposed mechanisms for pseudouridine formation. The two most probable of these mechanisms result in either the formation of an acylal intermediate, where the catalytic aspartate is covalently linked to the ribose of the target uridine, or the formation of a glycal intermediate that has no covalent connections between the RNA and the protein (Miracco and Mueller, 2011). Using the TruB R181 variants that displayed very slow rates of catalysis, assays can be performed to try to trap intermediates that may build up during pseudouridine formation. For example, the acylal and glycal intermediate mechanisms would both generate a free uracil base, but the Michael addition pathway would form a covalent complex between the enzyme and the base. After incubating tRNA radio-labeled on the uracil base with TruB R181M or R181A and separating the protein from the tRNA, scintillation counting of each fraction would indicate whether the protein was linked to the radioactive uracil base. This would suggest that pseudouridine formation follows the Michael addition mechanism. Alternatively, if radioactivity was not found in the protein fraction, indicating that the base does not form a covalent complex with the protein, an electrophoretic gel-shift assay could be used to monitor if a covalent complex forms instead between the protein and the tRNA. This should only occur if the acylal intermediate is formed. These types of analyses have not been possible with other TruB variants that still retained appreciable rates of catalysis. The use of the impaired TruB R181 variants in these two assays can provide evidence for or against each of the proposed mechanisms of pseudouridylation. If a universal pseudouridylation mechanism is utilized, it is expected that similar intermediates can be trapped for the other pseudouridine synthase families. Since a basic active site residue is conserved in all families, corresponding mutations in these other enzymes should also sufficiently impair catalysis to test for the presence of similar covalent complexes as further evidence for a common mechanism of pseudouridine formation.

An interesting question surrounding RNA modifications is the specific function of each modification. A number of genes coding for RNA modification enzymes are not essential, but removal of some of these genes can in some cases decrease an organism's fitness or have lethal effects when two or more of these genes are removed (Gutgsell et al., 2000; Gutgsell et al., 2001; King et al., 2003; Kinghorn et al., 2002; Phizicky and Hopper, 2010). Interestingly, the absence of certain modifications in rRNA have been shown to increase susceptibility to antibiotics that target the PTC (Toh and Mankin, 2008). This allows for speculation that certain modification enzymes evolved to confer intrinsic antibiotic resistance. Modification enzymes may then be an attractive new target for antibiotics as a way to increase the efficacy of treatments that are already in use. Developing inhibitors that target specific pseudouridine synthases may be an effective way to treat bacterial infections. However, specificity for bacterial enzymes is necessary since the human homolog of TruB, dyskerin, is important for the maturation and stability of telomerase RNA in addition to pseudouridylation of ribosomal and spliceosomal RNAs in the context of the H/ACA snoRNP (Meier, 2003). Because of this role, mutations in dyskerin have been implicated with aberrant telomerase function and dyskeratosis congenita, and the loss of function of dyskerin in humans (and the subsequent loss of pseudouridines in the ribosome) have been shown to result in lowered affinity of tRNA binding to the ribosome and a decrease in translational fidelity, and may potentially impair ribosome biogenesis (Jack et al., 2011). The regulation and sequential order of the incorporation of RNA modifications may also give us an idea of when during the cell cycle pseudouridine modifications are most beneficial or what processes are most affected by the presence or absence of modifications, for example under certain stress conditions. Therefore, RNA modifications should be thought of in the global context of an RNA molecule and how these modifications may work together to provide a particular functional advantage.

4.2 tRNA Dynamics

The folding, stability, and catalytic function of RNAs often rely on the presence of Mg^{2+} or other cations (Fedor, 2002; Misra and Draper, 2002; Serebrov et al., 2001). Crystal structures of RNA molecules have indicated the presence of Mg^{2+} binding sites that may contribute to the various proposed roles of Mg^{2+}. To better understand the dynamics of Mg^{2+} binding sites and RNA-protein interactions that are not apparent from crystallographic data, MD simulations can be used to analyze general motions and critical molecular contacts at the atomic level. In fact, MD simulations of nucleic acids have been used to correctly predict or propose new binding sites for metal ions (Auffinger and Westhof, 2000; Cheatham and Kollman, 1997; Young et al., 1997). An advantage to using MD simulations is that any crystal-packing effects that may influence Mg^{2+} binding are removed, however other limitations are manifested in the selection of appropriate force fields to best represent the nucleic acid and the ionic solution (Cheatham and Kollman, 2000). Here, I performed MD simulations on tRNA in the presence and absence of Mg^{2+} that indicated there is no absolute requirement for Mg^{2+} ions for tRNA stability *in silico*, and that the initial placement of Mg^{2+} ions may have only limited local structural effects.

Placing Mg^{2+} ions in high probability locations identified from tRNA crystal structures, positions of high negative charge within the tRNA structure, or random placement around the tRNA did not exhibit any differences in the global stability of the tRNA. Even the absence of Mg^{2+} ions produced a similarly stable simulation of tRNA that did not drastically distort the tRNA for up to 40 ns. Long-range tertiary interactions were remarkably stable in all four simulations, though the presence of Mg^{2+} within the tRNA structure (for the X-18Mg^{2+} and C-18Mg^{2+} simulations) may have stabilized 1-2 extra Watson-Crick hydrogen bonds compared to the simulations that did not have Mg^{2+} within the structure. The similarity in the RMSD, RMSF, and tertiary interaction analyses for all four simulations indicate that Mg^{2+} is not essential for global stabilization or maintenance of the L-shaped tRNA structure on a nanosecond time scale *in silico*. However, some local stabilizing effects of Mg^{2+} may be conferred to Watson-Crick base pairing and the tRNA elbow angle. While the R-18Mg^{2+} and X-18Mg^{2+} simulations showed a slightly more compact elbow angle in simulation frames that had Mg^{2+} occupying a specific position within the structure, the C-18Mg^{2+} simulation was slightly more compact in the frames that did not have Mg^{2+} at this same position. Although Mg^{2+} spent the most amount of time in the elbow region of the tRNA in all simulations, there were small differences in the specific Mg^{2+} occupancy sites. These differences may account for the variable fluctuations measured in the elbow angle between the three simulations. Therefore, the presence of Mg^{2+} may have small effects on local structure, but the overall stability of tRNA is not greatly affected by Mg^{2+} ions.

Though it appears that Mg^{2+} is not important for general tRNA stability *in silico*, MD simulations should be performed under conditions that best represent the *in vivo* or *in* *vitro* state of a molecular so as to maintain validity in comparison to experimental results. For this reason and despite the findings presented in this thesis, Mg^{2+} should still be included in nucleic acid simulations. Three different placement methods were tested here, and all methods can be considered valid approaches for the inclusion of Mg^{2+} ions. This is due to similar regions of high Mg^{2+} occupancy, the sum of inner and outer sphere Mg^{2+} ions, and the ability to cluster the three simulations based on Mg^{2+} occupancy of a common reference position. However, placing Mg^{2+} in positions to balance the highly negative charge of the RNA backbone may be the best approach when performing these simulations. Based on the Mg^{2+} occupancy maps, the four Mg^{2+} ions placed in the C-18Mg^{2+} simulation were the most stable (i.e. least mobile) of all Mg^{2+} ions placed in any of the simulations. This suggests that these ions were highly coordinated by the tRNA and therefore positively contributed to the stability of the simulation. Placement based on positions identified by crystal structures are likely more important for DNA or RNA that has highly defined metal binding sites, which was not the case for tRNA as only four Mg^{2+} binding sites were identified in more than 10% of 28 aligned tRNA structures having over 120 Mg^{2+} ions in total. In fact, even the most populated Mg^{2+} binding site was seen in less than 30% of the aligned structures, further supporting the idea that tRNA does not contain specific Mg^{2+} binding sites.

There appears to be no minimum number of Mg^{2+} ions required for tRNA stability *in silico*, but further studies can test for whether there is a maximum number of Mg^{2+} ions that can be added before artificial perturbations are seen in the tRNA structure. This can be done using any or all of the three placement methods used in this study. However, placement of additional Mg^{2+} within the structure would be expected to have a greater effect than random placement of Mg^{2+} in the surrounding water box since ions will more easily move from unfavorable placement in the water box than from a coordinated position within the structure. Longer simulations may need to be performed before an effect on tRNA stability due to Mg^{2+} is manifested. This might be seen in the eventual fraying of secondary or tertiary base pairs, or in distortions to the L-shaped structure. Less of a requirement for Mg^{2+} would be expected in simulations of modified tRNA, since a number of modifications (e.g. pseudouridine) are thought to stabilize RNA (Davis, 1995; Jackman and Alfonzo, 2013) and may alleviate the need for Mg^{2+} that may be more apparent in unmodified RNA molecules. For example, the additional base pairs that were stabilized in the X-18Mg^{2+} and C-18Mg^{2+} simulations may be present in simulations of fully modified tRNA in the absence of Mg^{2+}. Other effects of modifications on the dynamics of tRNA can be assessed by systematically adding modifications to the tRNA for simulation and comparing the stability and intramolecular interactions to simulations with different sets of modifications. In particular, pseudouridine has been shown to coordinate a water molecule with the backbone of the preceding nucleotide (Arnez and Steitz, 1994); analysis of this interaction may confirm the stabilizing effect of pseudouridine on RNA structure.

Ultimately, this work provides the first steps toward the simulation of a complex of tRNA and TruB. Since a crystal structure of TruB bound to a full-length tRNA substrate is not available, simulation of a modeled complex will provide insight into additional interactions that has so far not been possible to study. For instance, these simulations may reveal dynamic interactions between active site residues of TruB and the tRNA, changes in the protein or tRNA structure due to binding (e.g. base-flipping of the tRNA), or additional interactions that may contribute to substrate recognition by TruB, such as interactions between the PUA domain of TruB with the tRNA acceptor stem. Additionally, steered MD simulations can be performed where constraints are added to the target uridine to gradually move it from the native conformation to the flipped-out position as seen in crystal structures of TruB bound to a short RNA substrate (Hoang and Ferré-D'Amaré, 2001; Pan et al., 2003; Phannachet and Huang, 2004). This type of simulation will better indicate what residues are involved in base-flipping and what kind of active site rearrangements are needed to accommodate the flipped-out bases, which may help to better define the role of residues like R181 and D90 in pseudouridine formation (see above). Since TruB binds substrate and product tRNA with similar affinities (Keffer-Wilkes, 2012; Ramamurthy et al., 1999), simulations performed on complexes of TruB bound to substrate and TruB bound to product may indicate differences in recognition patterns between uridine and pseudouridine that explain when or how TruB senses that catalysis is complete. In relation to the Mg^{2+} studies presented here, TruB does not require Mg^{2+} to form pseudouridine in tRNA (Nagamatsu and Miyazawa, 1983; Nurse et al., 1995), though the effect (if any) Mg^{2+} has on tRNA binding to TruB is not known since it is not the rate-limiting step of pseudouridine formation. Therefore, Mg^{2+} can still be included in the simulations to ensure the stability of the *in silico* complex is not compromised, as well as to mimic the conditions used for the TruB simulations and the experimental binding and activity assays. Simulating a tRNA-TruB complex and analyzing the dynamics of this interaction may be a direct way of examining the hypothesis that TruB has an additional and more important role in the cell as an RNA chaperone (Gutgsell et al., 2000). This may be seen in fine-tuning of the overall tRNA structure when bound to TruB (such as elbow angle changes through interactions with the PUA domain) or increased flexibility of tRNA regions that do not interact with TruB residues (such as the D-arm nucleotides that may have less contact to the T-arm). The conformation of the tRNA in these types of simulations can be complemented by single-molecule studies using fluorescently-labeled tRNA that may indicate the presence of intermediate tRNA structures when interacting with TruB.

RNA has a variety of interaction partners in the cell that in some cases contribute to RNA structure, function, stability, or degradation. Due to improvements in nucleic acid force fields for MD simulations (Denning et al., 2011; Yildirim et al., 2010), complex systems involving DNA and RNA can be more reliably simulated to give information about intramolecular interactions and dynamics. Further advances in the parameterization of non-canonical motifs (e.g. tetraloops or quadruple helices) will allow for an even greater number of systems to be simulated for longer times. Simulations of nucleic acids in complex with modification enzymes, transcription or translation machinery, and spliceosome components can provide information about critical interactions involved in essential cellular processes. These types of studies can also be used to aid in the identification of new drug targets or the design of novel inhibitors. In combination with experimental techniques, MD simulations of nucleic acids can be a powerful way to closely examine influences from ions, solvents, proteins, and other biomolecules on the structure and function of essential RNA molecules.

References

Adcock, S.A., and McCammon, J.A. (2006). Molecular dynamics: survey of methods for simulating the activity of proteins. Chem Rev *106*, 1589-1615.

Alexander, R.W., Eargle, J., and Luthey-Schulten, Z. (2010). Experimental and computational determination of tRNA dynamics. FEBS Lett *584*, 376-386.

Alian, A., Lee, T.T., Griner, S.L., Stroud, R.M., and Finer-Moore, J. (2008). Structure of a TrmA-RNA complex: A consensus RNA fold contributes to substrate selectivity and catalysis in m5U methyltransferases. Proc Natl Acad Sci U S A *105*, 6876-6881.

Arena, F., Ciliberto, G., Ciampi, S., and Cortese, R. (1978). Purification of pseudouridylate synthetase I from *Salmonella typhimurium*. Nucleic Acids Res *5*, 4523-4536.

Arnez, J.G., and Steitz, T.A. (1994). Crystal structure of unmodified tRNA(Gln) complexed with glutaminyl-tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure. Biochemistry *33*, 7560-7567.

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics *22*, 195-201.

Auffinger, P., Bielecki, L., and Westhof, E. (2003). The Mg^{2+} binding sites of the 5S rRNA loop E motif as investigated by molecular dynamics simulations. Chem Biol *10*, 551-561.

Auffinger, P., Bielecki, L., and Westhof, E. (2004). Symmetric K^+ and Mg^{2+} ion-binding sites in the 5S rRNA loop E inferred from molecular dynamics simulations. J Mol Biol *335*, 555-571.

Auffinger, P., Louise-May, S., and Westhof, E. (1999). Molecular dynamics simulations of solvated yeast tRNA(Asp). Biophys J *76*, 50-64.

Auffinger, P., and Westhof, E. (1997). RNA hydration: three nanoseconds of multiple molecular dynamics simulations of the solvated tRNA(Asp) anticodon hairpin. J Mol Biol *269*, 326-341.

Auffinger, P., and Westhof, E. (2000). Water and ion binding around RNA and DNA (C,G) oligomers. J Mol Biol *300*, 1113-1131.

Banáš, P., Sklenovsky, P., Wedekind, J.E., Šponer, J., and Otyepka, M. (2012). Molecular mechanism of preQ1 riboswitch action: a molecular dynamics study. J Phys Chem B *116*, 12721-12734.

Basavappa, R., and Sigler, P.B. (1991). The 3 Å crystal structure of yeast initiator tRNA: functional implications in initiator/elongator discrimination. EMBO J *10*, 3105-3111.

Becker, H.F., Motorin, Y., Sissler, M., Florentz, C., and Grosjean, H. (1997). Major identity determinants for enzymatic formation of ribothymidine and pseudouridine in the T psi-loop of yeast tRNAs. J Mol Biol *274*, 505-518.

Behm-Ansmant, I., Urban, A., Ma, X., Yu, Y.T., Motorin, Y., and Branlant, C. (2003). The *Saccharomyces cerevisiae* U2 snRNA:pseudouridine-synthase Pus7p is a novel multisite-multisubstrate RNA:Psi-synthase also acting on tRNAs. RNA *9*, 1371-1382.

Beššeová, I., Banáš, P., Kührová, P., Košinová, P., Otyepka, M., and Šponer, J. (2012). Simulations of A-RNA duplexes. The effect of sequence, solute force field, water model, and salt concentration. J Phys Chem B *116*, 9899-9916.

Bhaskaran, H., Rodriguez-Hernandez, A., and Perona, J.J. (2012). Kinetics of tRNA folding monitored by aminoacylation. RNA *18*, 569-580.

Bowman, J.C., Lenz, T.K., Hud, N.V., and Williams, L.D. (2012). Cations in charge: magnesium ions in RNA folding and catalysis. Curr Opin Struct Biol *22*, 262-272.

Brooks, B.R., Brooks, C.L., 3rd, Mackerell, A.D., Jr., Nilsson, L., Petrella, R.J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., *et al.* (2009). CHARMM: the biomolecular simulation program. J Comput Chem *30*, 1545-1614.

Byrne, R.T., Konevega, A.L., Rodnina, M.V., and Antson, A.A. (2010). The crystal structure of unmodified tRNA$^{\mathrm{Phe}}$ from *Escherichia coli*. Nucleic Acids Res *38*, 4154-4162.

Chan, C.M., and Huang, R.H. (2009). Enzymatic characterization and mutational studies of TruD - the fifth family of pseudouridine synthases. Arch Biochem Biophys *489*, 15-19.

Charette, M., and Gray, M.W. (2000). Pseudouridine in RNA: What, where, how, and why. IUBMB Life *49*, 341-351.

Chaudhuri, B.N., Chan, S., Perry, L.J., and Yeates, T.O. (2004). Crystal structure of the apo forms of psi 55 tRNA pseudouridine synthase from *Mycobacterium tuberculosis*: a hinge at the base of the catalytic cleft. J Biol Chem *279*, 24585-24591.

Cheatham, T.E., 3rd, Cieplak, P., and Kollman, P.A. (1999). A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. J Biomol Struct Dyn *16*, 845-862.

Cheatham, T.E., 3rd, and Kollman, P.A. (2000). Molecular dynamics simulation of nucleic acids. Annu Rev Phys Chem *51*, 435-471.

Cheatham, T.E., 3rd, and Young, M.A. (2000). Molecular dynamics simulation of nucleic acids: successes, limitations, and promise. Biopolymers *56*, 232-256.

Cheatham, T.E., and Kollman, P.A. (1997). Molecular Dynamics Simulations Highlight the Structural Differences among DNA:DNA, RNA:RNA, and DNA:RNA Hybrid Duplexes. J Am Chem Soc *119*, 4805-4825.

Chen, C., Zhao, X., Kierzek, R., and Yu, Y.T. (2010). A flexible RNA backbone within the polypyrimidine tract is required for U2AF65 binding and pre-mRNA splicing *in vivo*. Mol Cell Biol *30*, 4108-4119.

Ciampi, M.S., Arena, F., and Cortese, R. (1977). Biosynthesis of pseudouridine in the *in vitro* transcribed tRNA^{Tyr} precursor. FEBS Lett *77*, 75-82.

Cieplak, P., Cheatham, T.E., and Kollman, P.A. (1997). Molecular Dynamics Simulations Find That 3' Phosphoramidate Modified DNA Duplexes Undergo a B to A Transition and Normal DNA Duplexes an A to B Transition. J Am Chem Soc *119*, 6722-6730.

Conrad, J., Niu, L., Rudd, K., Lane, B.G., and Ofengand, J. (1999). 16S ribosomal RNA pseudouridine synthase RsuA of *Escherichia coli*: deletion, mutation of the conserved Asp102 residue, and sequence comparison among all other pseudouridine synthases. RNA *5*, 751-763.

Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., and Kollman, P.A. (1995). A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. J Am Chem Soc *117*, 5179-5197.

Cowan, J.A. (1995). The biological chemistry of magnesium (VCH Publishers).

Cunningham, P.R., Richard, R.B., Weitzmann, C.J., Nurse, K., and Ofengand, J. (1991). The absence of modified nucleotides affects both *in vitro* assembly and *in vitro* function of the 30S ribosomal subunit of *Escherichia coli*. Biochimie *73*, 789-796.

Daher, M., and Rueda, D. (2012). Fluorescence Characterization of the Transfer RNA-like Domain of Transfer Messenger RNA in Complex with Small Binding Protein B. Biochemistry *51*, 3531-3538.

Davis, D.R. (1995). Stabilization of RNA stacking by pseudouridine. Nucleic Acids Res *23*, 5020-5026.

Davis, F.F., and Allen, F.W. (1957). Ribonucleic acids from yeast which contain a fifth nucleotide. J Biol Chem *227*, 907-915.

Del Campo, M., Kaya, Y., and Ofengand, J. (2001). Identification and site of action of the remaining four putative pseudouridine synthases in *Escherichia coli*. RNA *7*, 1603-1615.

Denning, E.J., Priyakumar, U.D., Nilsson, L., and Mackerell, A.D., Jr. (2011). Impact of 2'-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. J Comput Chem *32*, 1929-1943.

Ditzler, M.A., Otyepka, M., Sponer, J., and Walter, N.G. (2010). Molecular dynamics and quantum mechanics of RNA: conformational and chemical change we can believe in. Accounts of chemical research *43*, 40-47.

Dornberger, U., Spackovj, N., Walter, A., Gollmick, F.A., Sponer, J., and Fritzsche, H. (2001). Solution structure of the dodecamer d-(CATGGGCC-CATG)2 is B-DNA. Experimental and molecular dynamics study. J Biomol Struct Dyn *19*, 159-174.

Duan, J., Li, L., Lu, J., Wang, W., and Ye, K. (2009). Structural Mechanism of Substrate RNA Recruitment in H/ACA RNA-Guided Pseudouridine Synthase. Mol Cell *34*, 427-439.

Duchêne, A.-M., Pujol, C., and Maréchal-Drouard, L. (2009). Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria. Curr Genet *55*, 1-18.

Dunin-Horkawicz, S., Czerwoniec, A., Gajda, M.J., Feder, M., Grosjean, H., and Bujnicki, J.M. (2006). MODOMICS: a database of RNA modification pathways. Nucleic Acids Res *34*, D145-149.

Eargle, J., Black, A.A., Sethi, A., Trabuco, L.G., and Luthey-Schulten, Z. (2008). Dynamics of Recognition between tRNA and Elongation Factor Tu. J Mol Biol *377*, 1382-1405.

Ehrenberg, M., Rigler, R., and Wintermeyer, W. (1979). On the structure and conformational dynamics of yeast phenylalanine-accepting transfer ribonucleic acid in solution. Biochemistry *18*, 4588-4599.

ElHefnawi, M., ElGamacy, M., and Fares, M. (2012). Multiple virtual screening approaches for finding new hepatitis C virus RNA-dependent RNA polymerase inhibitors: structure-based screens and molecular dynamics for the pursue of new poly pharmacological inhibitors. BMC Bioinformatics *13 Suppl 17*, S5.

Ernsting, B., Edwards, D., Aldred, K., Fites, J., and Neff, C. (2009). Mitochondrial genome sequence of *Unionicola foili* (Acari: Unionicolidae): a unique gene order with implications for phylogenetic inference. Exp Appl Acarol *49*, 305-316.

Fedor, M.J. (2002). The role of metal ions in RNA catalysis. Curr Opin Struct Biol *12*, 289-295.

Feller, S.E., and MacKerell, A.D. (2000). An Improved Empirical Potential Energy Function for Molecular Simulations of Phospholipids. J Phys Chem B *104*, 7510-7515.

Fernandez, I.S., Ng, C.L., Kelley, A.C., Wu, G., Yu, Y.T., and Ramakrishnan, V. (2013). Unusual base pairing during the decoding of a stop codon by the ribosome. Nature *500*, 107-110.

Foster, P.G., Huang, L., Santi, D.V., and Stroud, R.M. (2000). The structural basis for tRNA recognition and pseudouridine formation by pseudouridine synthase I. Nat Struct Biol *7*, 23-27.

Ge, J., and Yu, Y.T. (2013). RNA pseudouridylation: new insights into an old modification. Trends Biochem Sci *38*, 210-218.

Giegé, R., Jühling, F., Pütz, J., Stadler, P., Sauter, C., and Florentz, C. (2012). Structure of transfer RNAs: similarity and variability. Wiley Interdiscip Rev RNA *3*, 37-61.

Gill, S.C., and von Hippel, P.H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem *182*, 319-326.

Grosjean, H. (2005). Modification and editing of RNA: historical overview and important facts to remember. In Fine-Tuning of RNA Functions by Modification and Editing, H. Grosjean, ed. (Springer Berlin Heidelberg), pp. 1-22.

Grozdanov, P.N., Fernandez-Fuentes, N., Fiser, A., and Meier, U.T. (2009). Pathogenic NAP57 mutations decrease ribonucleoprotein assembly in dyskeratosis congenita. Hum Mol Genet *18*, 4546-4551.

Gu, X., Liu, Y., and Santi, D.V. (1999). The mechanism of pseudouridine synthase I as deduced from its interaction with 5-fluorouracil-tRNA. Proc Natl Acad Sci U S A *96*, 14270-14275.

Gu, X., Yu, M., Ivanetich, K.M., and Santi, D.V. (1998). Molecular recognition of tRNA by tRNA pseudouridine 55 synthase. Biochemistry *37*, 339-343.

Gutgsell, N., Englund, N., Niu, L.H., Kaya, Y., Lane, B.G., and Ofengand, J. (2000). Deletion of the *Escherichia coli* pseudouridine synthase gene truB blocks formation of pseudouridine 55 in tRNA *in vivo*, does not affect exponential growth, but confers a strong selective disadvantage in competition with wild-type cells. RNA *6*, 1870-1881.

Gutgsell, N.S., Del Campo, M., Raychaudhuri, S., and Ofengand, J. (2001). A second function for pseudouridine synthases: A point mutant of RluD unable to form pseudouridines 1911, 1915, and 1917 in *Escherichia coli* 23S ribosomal RNA restores normal growth to an RluD-minus strain. RNA *7*, 990-998.

Hall, K.B., Sampson, J.R., Uhlenbeck, O.C., and Redfield, A.G. (1989). Structure of an unmodified tRNA molecule. Biochemistry *28*, 5794-5801.

Hamilton, C.S., Spedaliere, C.J., Ginter, J.M., Johnston, M.V., and Mueller, E.G. (2005). The roles of the essential Asp-48 and highly conserved His-43 elucidated by the pH dependence of the pseudouridine synthase TruB. Arch Biochem Biophys *433*, 322-334.

Hamma, T., and Ferré-D'Amaré, A.R. (2006). Pseudouridine Synthases. Chem Biol *13*, 1125-1135.

Hayes, R.L., Noel, J.K., Mohanty, U., Whitford, P.C., Hennelly, S.P., Onuchic, J.N., and Sanbonmatsu, K.Y. (2012). Magnesium fluctuations modulate RNA dynamics in the SAM-I riboswitch. J Am Chem Soc *134*, 12043-12053.

Heiss, N.S., Knight, S.W., Vulliamy, T.J., Klauck, S.M., Wiemann, S., Mason, P.J., Poustka, A., and Dokal, I. (1998). X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. Nat Genet *19*, 32-38.

Hoang, C., Chen, J., Vizthum, C.A., Kandel, J.M., Hamilton, C.S., Mueller, E.G., and Ferré-D'Amaré, A.R. (2006). Crystal structure of pseudouridine synthase RluA: indirect sequence readout through protein-induced RNA structure. Mol Cell *24*, 535-545.

Hoang, C., and Ferré-D'Amaré, A.R. (2004). Crystal structure of the highly divergent pseudouridine synthase TruD reveals a circular permutation of a conserved fold. RNA *10*, 1026-1033.

Hoang, C., and Ferré-D’Amaré, A.R. (2001). Cocrystal structure of a tRNA Psi 55 pseudouridine synthase: Nucleotide flipping by an RNA-modifying enzyme. Cell *107*, 929-939.

Hoang, C., Hamilton, C.S., Mueller, E.G., and Ferré-D'Amaré, A.R. (2005). Precursor complex structure of pseudouridine synthase TruB suggests coupling of active site perturbations to an RNA-sequestering peripheral protein domain. Protein Sci *14*, 2201-2206.

Holbrook, S.R., Sussman, J.L., Warrant, R.W., Church, G.M., and Kim, S.H. (1977). RNA-ligant interactions. (I) Magnesium binding sites in yeast tRNA^{Phe}. Nucleic Acids Res *4*, 2811-2820.

Hopper, A.K., Pai, D.A., and Engelke, D.R. (2010). Cellular dynamics of tRNAs and their genes. FEBS Lett *584*, 310-317.

Huang, L., Pookanjanatavip, M., Gu, X., and Santi, D.V. (1998). A conserved aspartate of tRNA pseudouridine synthase is essential for activity and a probable nucleophilic catalyst. Biochemistry *37*, 344-351.

Hud, N.V., and Polak, M. (2001). DNA-cation interactions: The major and minor grooves are flexible ionophores. Curr Opin Struct Biol *11*, 293-301.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J Mol Graph *14*, 33-38, 27-38.

Ishida, K., Kunibayashi, T., Tomikawa, C., Ochi, A., Kanai, T., Hirata, A., Iwashita, C., and Hori, H. (2011). Pseudouridine at position 55 in tRNA controls the contents of other modified nucleotides for low-temperature adaptation in the extreme-thermophilic eubacterium *Thermus thermophilus*. Nucleic Acids Res *39*, 2304-2318.

Jack, K., Bellodi, C., Landry, Dori M., Niederer, Rachel O., Meskauskas, A., Musalgaonkar, S., Kopmar, N., Krasnykh, O., Dean, Alison M., Thompson, Sunnie R., *et al.* (2011). rRNA Pseudouridylation Defects Affect Ribosomal Ligand Binding and Translational Fidelity from Yeast to Human Cells. Mol Cell *44*, 660-666.

Jackman, J.E., and Alfonzo, J.D. (2013). Transfer RNA modifications: nature's combinatorial chemistry playground. Wiley Interdiscip Rev RNA *4*, 35-48.

Jeffrey, G.A. (1997). An introduction to hydrogen bonding (New York, Oxford University Press).

Johnson, L., and Söll, D. (1970). *In vitro* biosynthesis of pseudouridine at the polynucleotide level by an enzyme extract from *Escherichia coli*. Proc Natl Acad Sci U S A *67*, 943-950.

Kamalampeta, R., Keffer-Wilkes, L.C., and Kothe, U. (2013). tRNA Binding, Positioning, and Modification by the Pseudouridine Synthase Pus10. J Mol Biol *425*, 3863-3874.

Kammen, H.O., Marvel, C.C., Hardy, L., and Penhoet, E.E. (1988). Purification, structure, and properties of *Escherichia coli* tRNA pseudouridine synthase I. J Biol Chem *263*, 2255-2263.

Karijolich, J., and Yu, Y.T. (2011). Converting nonsense codons into sense codons by targeted pseudouridylation. Nature *474*, 395-398.

Kaya, Y., and Ofengand, J. (2003). A novel unanticipated type of pseudouridine synthase with homologs in bacteria, archaea, and eukarya. RNA *9*, 711-721.

Keffer-Wilkes, L.C. (2012). Substrate binding and catalysis by the pseudouridine synthases RluA and TruB. In Department of Chemistry & Biochemistry (Lethbridge, University of Lethbridge).

Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L., and Schwede, T. (2009). The SWISS-MODEL Repository and associated resources. Nucleic Acids Res *37*, D387-392.

King, T.H., Liu, B., McCully, R.R., and Fournier, M.J. (2003). Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. Mol Cell *11*, 425-435.

Kinghorn, S.M., O'Byrne, C.P., Booth, I.R., and Stansfield, I. (2002). Physiological analysis of the role of truB in *Escherichia coli*: a role for tRNA modification in extreme temperature resistance. Microbiology *148*, 3511-3520.

Klein, R.A. (2006). Modified van der Waals atomic radii for hydrogen bonding based on electron density topology. Chemical Physics Letters *425*, 128-133.

Kolaskar, A.S., and Joshi, R.R. (1998). Molecular dynamics simulation of a 13-mer duplex DNA: a PvuII substrate. J Biomol Struct Dyn *15*, 1155-1165.

Koonin, E.V. (1996). Pseudouridine synthases: four families of enzymes containing a putative uridine-binding motif also conserved in dUTPases and dCTP deaminases. Nucleic Acids Res *24*, 2411-2415.

Krepl, M., Reblova, K., Koca, J., and Šponer, J. (2013). Bioinformatics and molecular dynamics simulation study of L1 stalk non-canonical rRNA elements: kink-turns, loops, and tetraloops. J Phys Chem B *117*, 5540-5555.

Kuttel, M., Brady, J.W., and Naidoo, K.J. (2002). Carbohydrate solution simulations: producing a force field with experimentally consistent primary alcohol rotational frequencies and populations. J Comput Chem *23*, 1236-1243.

Labahn, J., Scharer, O.D., Long, A., Ezaz-Nikpay, K., Verdine, G.L., and Ellenberger, T.E. (1996). Structural basis for the excision repair of alkylation-damaged DNA. Cell *86*, 321-329.

Lane, B.G., Ofengand, J., and Gray, M.W. (1995). Pseudouridine and O2'-methylated nucleosides. Significance of their selective occurrence in rRNA domains that function in ribosome-catalyzed synthesis of the peptide bonds in proteins. Biochimie *77*, 7-15.

Langley, D.R. (1998). Molecular Dynamic Simulations of Environment and Sequence Dependent DNA Conformations: The Development of the BMS Nucleic Acid Force Field and Comparison with Experimental Results. J Biomol Struct Dyn *16*, 487-509.

Lecointe, F., Namy, O., Hatin, I., Simos, G., Rousset, J.P., and Grosjean, H. (2002). Lack of pseudouridine 38/39 in the anticodon arm of yeast cytoplasmic tRNA decreases *in vivo* recoding efficiency. J Biol Chem *277*, 30445-30453.

Lee, S.H., Kim, I., and Chung, B.C. (2007). Increased urinary level of oxidized nucleosides in patients with mild-to-moderate Alzheimer's disease. Clin Biochem *40*, 936-938.

Lee, S.R., and Collins, K. (2005). Starvation-induced cleavage of the tRNA anticodon loop in *Tetrahymena thermophila*. J Biol Chem *280*, 42744-42749.

Leroy, J.L., and Gueron, M. (1977). Electrostatic effects in divalent ion binding to tRNA. Biopolymers *16*, 2429-2446.

Li, C.H., Zuo, Z.C., Su, J.G., Xu, X.J., and Wang, C.X. (2013). The interactions and recognition of cyclic peptide mimetics of Tat with HIV-1 TAR RNA: a molecular dynamics simulation study. J Biomol Struct Dyn *31*, 276-287.

Li, L., and Ye, K. (2006). Crystal structure of an H/ACA box ribonucleoprotein particle. Nature *443*, 302-307.

Li, X., Hassan, S.A., and Mehler, E.L. (2005). Long dynamics simulations of proteins using atomistic force fields and a continuum representation of solvent effects: calculation of structural and dynamic properties. Proteins *60*, 464-484.

MacKerell, A.D., Bashford, D., Bellott, Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., *et al.* (1998). All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. J Phys Chem B *102*, 3586-3616.

Marck, C., and Grosjean, H. (2002). tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. RNA *8*, 1189-1232.

McCammon, J.A., ed. (1976). Molecular dynamics study of the bovine pancreatic trypsin inhibitor (Orsay, France, CECAM, Universite de Paris).

McCleverty, C.J., Hornsby, M., Spraggon, G., and Kreusch, A. (2007). Crystal structure of human Pus10, a novel pseudouridine synthase. J Mol Biol *373*, 1243-1254.

Meier, U.T. (2003). Dissecting dyskeratosis. Nat Genet *33*, 116-117.

Miracco, E.J., and Mueller, E.G. (2011). The products of 5-fluorouridine by the action of the pseudouridine synthase TruB disfavor one mechanism and suggest another. J Am Chem Soc *133*, 11826-11829.

Misra, V.K., and Draper, D.E. (2002). The linkage between magnesium binding and RNA folding. J Mol Biol *317*, 507-521.

Mueller, E.G., and Ferré-D’Amaré, A.R., eds. (2009). Pseudouridine Formation, the Most Common Transglycosylation in RNA (Austin, Landes Bioscience).

Nagamatsu, K., and Miyazawa, Y. (1983). Thermal and Mg^{2+} dependent behavior of pseudouridines at 39th and 55th of yeast $tRNA^{Phe}$. Nucleic Acids Symp Ser, 133-136.

Newby, M.I., and Greenbaum, N.L. (2001). A conserved pseudouridine modification in eukaryotic U2 snRNA induces a change in branch-site architecture. RNA *7*, 833-845.

Nilsson, L., Rigler, R., and Wintermeyer, W. (1983). The influence of spermine on the structural dynamics of yeast $tRNA^{Phe}$. Biochim Biophys Acta *740*, 460-465.

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

Norberg, J., and Nilsson, L. (2002). Molecular dynamics applied to nucleic acids. Accounts of chemical research *35*, 465-472.

Nurse, K., Wrzesinski, J., Bakin, A., Lane, B.G., and Ofengand, J. (1995). Purification, cloning, and properties of the tRNA psi 55 synthase from *Escherichia coli*. RNA *1*, 102-112.

Ofengand, J., and Fournier, M.J., eds. (1998). The pseudouridine residues of rRNA: Number, location, biosynthesis, and function (New York, ASM Press).

Pan, H., Agarwalla, S., Moustakas, D.T., Finer-Moore, J., and Stroud, R.M. (2003). Structure of tRNA pseudouridine synthase TruB and its RNA complex: RNA recognition through a combination of rigid docking and induced fit. Proc Natl Acad Sci U S A *100*, 12648-12653.

Pereira, T.C., and Lopes-Cendes, I. (2012). Emerging RNA-based drugs: siRNAs, microRNAs and derivates. Cent Nerv Syst Agents Med Chem *12*, 217-232.

Peterson, E.T., and Uhlenbeck, O.C. (1992). Determination of recognition nucleotides for Escherichia coli phenylalanyl-tRNA synthetase. Biochemistry *31*, 10380-10389.

Phannachet, K., Elias, Y., and Huang, R.H. (2005). Dissecting the roles of a strictly conserved tyrosine in substrate recognition and catalysis by pseudouridine 55 synthase. Biochemistry *44*, 15488-15494.

Phannachet, K., and Huang, R.H. (2004). Conformational change of pseudouridine 55 synthase upon its association with RNA substrate. Nucleic Acids Res *32*, 1422-1429.

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. J Comput Chem *26*, 1781-1802.

Phizicky, E.M., and Hopper, A.K. (2010). tRNA biology charges to the front. Genes Dev *24*, 1832-1860.

Quigley, G.J., Teeter, M.M., and Rich, A. (1978). Structural analysis of spermine and magnesium ion binding to yeast phenylalanine transfer RNA. Proc Natl Acad Sci U S A *75*, 64-68.

Ramamurthy, V., Swann, S.L., Paulson, J.L., Spedaliere, C.J., and Mueller, E.G. (1999). Critical aspartic acid residues in pseudouridine synthases. J Biol Chem *274*, 22225-22230.

Raychaudhuri, S., Niu, L., Conrad, J., Lane, B.G., and Ofengand, J. (1999). Functional effect of deletion and mutation of the *Escherichia coli* ribosomal RNA and tRNA pseudouridine synthase RluA. J Biol Chem *274*, 18880-18886.

Reiling, S., Schlenkrich, M., and Brickmann, J. (1996). Force field parameters for carbohydrates. J Comput Chem *17*, 450-468.

Romer, R., and Hach, R. (1975). tRNA conformation and magnesium binding. A study of a yeast phenylalanine-specific tRNA by a fluorescent indicator and differential melting curves. Eur J Biochem *55*, 271-284.

Sampson, J.R., DiRenzo, A.B., Behlen, L.S., and Uhlenbeck, O.C. (1989). Nucleotides in yeast tRNA^{Phe} required for the specific recognition by its cognate synthetase. Science *243*, 1363-1366.

Serebrov, V., Clarke, R.J., Gross, H.J., and Kisselev, L. (2001). Mg^{2+}-induced tRNA folding. Biochemistry *40*, 6688-6698.

Serebrov, V., Vassilenko, K., Kholod, N., Gross, H.J., and Kisselev, L. (1998). Mg^{2+} binding and structural stability of mature and *in vitro* synthesized unmodified *Escherichia coli* tRNA^{Phe}. Nucleic Acids Res *26*, 2723-2728.

Sheppard, K., Yuan, J., Hohn, M.J., Jester, B., Devine, K.M., and Soll, D. (2008). From one amino acid to another: tRNA-dependent amino acid biosynthesis. Nucleic Acids Res *36*, 1813-1825.

Shi, H., and Moore, P.B. (2000). The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: a classic structure revisited. RNA *6*, 1091-1105.

Sivaraman, J., Sauve, V., Larocque, R., Stura, E.A., Schrag, J.D., Cygler, M., and Matte, A. (2002). Structure of the 16S rRNA pseudouridine synthase RsuA bound to uracil and UMP. Nat Struct Biol *9*, 353-358.

Souliere, M.F., Haller, A., Rieder, R., and Micura, R. (2011). A Powerful Approach for the Selection of 2-Aminopurine Substitution Sites to Investigate RNA Folding. J Am Chem Soc *133*, 16161-16167.

Šponer, J., Cang, X., and Cheatham, T.E., 3rd (2012). Molecular dynamics simulations of G-DNA and perspectives on the simulation of nucleic acid structures. Methods *57*, 25-39.

Sprinzl, M., Horn, C., Brown, M., Ioudovitch, A., and Steinberg, S. (1998). Compilation of tRNA sequences and sequences of tRNA genes. Nucleic Acids Res *26*, 148-153.

Stein, A., and Crothers, D.M. (1976). Equilibrium binding of magnesium(II) by *Escherichia coli* $\text{tRNA}^{\text{fMet}}$. Biochemistry *15*, 157-160.

Stewart, T.S., Roberts, R.J., and Strominger, J.L. (1971). Novel species of tRNA. Nature *230*, 36-38.

Thompson, D.M., Lu, C., Green, P.J., and Parker, R. (2008). tRNA cleavage is a conserved response to oxidative stress in eukaryotes. RNA *14*, 2095-2103.

Toh, S.-M., and Mankin, A.S. (2008). An Indigenous Posttranscriptional Modification in the Ribosomal Peptidyl Transferase Center Confers Resistance to an Array of Protein Synthesis Inhibitors. J Mol Biol *380*, 593-597.

Tomita, K., Ueda, T., and Watanabe, K. (1999). The presence of pseudouridine in the anticodon alters the genetic code: a possible mechanism for assignment of the AAA lysine codon as asparagine in echinoderm mitochondria. Nucleic Acids Res *27*, 1683-1689.

Urbonavičius, J., Qian, Q., Durand, J.M., Hagervall, T.G., and Bjork, G.R. (2001). Improvement of reading frame maintenance is a common function for several tRNA modifications. EMBO J *20*, 4863-4873.

Wang, Z., Rao, D.D., Senzer, N., and Nemunaitis, J. (2011). RNA interference and cancer therapy. Pharm Res *28*, 2983-2995.

Whitford, P.C., Blanchard, S.C., Cate, J.H., and Sanbonmatsu, K.Y. (2013). Connecting the kinetics and energy landscape of tRNA translocation on the ribosome. PLoS Comput Biol *9*, e1003003.

Wieden, H.J., Mercier, E., Gray, J., Steed, B., and Yawney, D. (2010). A combined molecular dynamics and rapid kinetics approach to identify conserved three-dimensional communication networks in elongation factor Tu. Biophys J *99*, 3735-3743.

Withers, M., Wernisch, L., and Reis, M.D. (2006). Archaeology and evolution of transfer RNA genes in the *Escherichia coli* genome. RNA *12*, 933-942.

Wright, J.R., Keffer-Wilkes, L.C., Dobing, S.R., and Kothe, U. (2011). Pre-steady-state kinetic analysis of the three *Escherichia coli* pseudouridine synthases TruB, TruA, and RluA reveals uniformly slow catalysis. RNA *17*, 2074-2084.

Wu, G., Xiao, M., Yang, C., and Yu, Y.T. (2011). U2 snRNA is inducibly pseudouridylated at novel sites by Pus7p and snR81 RNP. EMBO J *30*, 79-89.

Wu, H., and Feigon, J. (2007). H/ACA small nucleolar RNA pseudouridylation pockets bind substrate RNA to form three-way junctions that position the target U for modification. Proc Natl Acad Sci U S A *104*, 6655-6660.

Ye, W., Qin, F., Zhang, J., Luo, R., and Chen, H.F. (2012). Atomistic mechanism of microRNA translation upregulation via molecular dynamics simulations. PLoS One *7*, e43788.

Yildirim, I., Stern, H.A., Kennedy, S.D., Tubbs, J.D., and Turner, D.H. (2010). Reparameterization of RNA chi Torsion Parameters for the AMBER Force Field and Comparison to NMR Spectra for Cytidine and Uridine. J Chem Theory Comput *6*, 1520-1531.

Young, M.A., Ravishanker, G., and Beveridge, D.L. (1997). A 5-nanosecond molecular dynamics trajectory for B-DNA: analysis of structure, motions, and solvation. Biophys J *73*, 2313-2336.

Yu, A.T., Ge, J., and Yu, Y.T. (2011). Pseudouridines in spliceosomal snRNAs. Protein Cell *2*, 712-725.

Yuan, M.L., Wei, D.D., Wang, B.J., Dou, W., and Wang, J.J. (2010). The complete mitochondrial genome of the citrus red mite *Panonychus citri* (Acari: Tetranychidae): high genome rearrangement and extremely truncated tRNAs. BMC Genomics *11*, 597.

Zamir, A., Holley, R.W., and Marquisee, M. (1965). Evidence for the Occurrence of a Common Pentanucleotide Sequence in the Structures of Transfer Ribonucleic Acids. J Biol Chem *240*, 1267-1273.

Zebarjadian, Y., King, T., Fournier, M.J., Clarke, L., and Carbon, J. (1999). Point mutations in yeast CBF5 can abolish *in vivo* pseudouridylation of rRNA. Mol Cell Biol *19*, 7461-7472.

Zhou, J., Liang, B., and Li, H. (2011). Structural and functional evidence of high specificity of Cbf5 for ACA trinucleotide. RNA *17*, 244-250.

Appendix

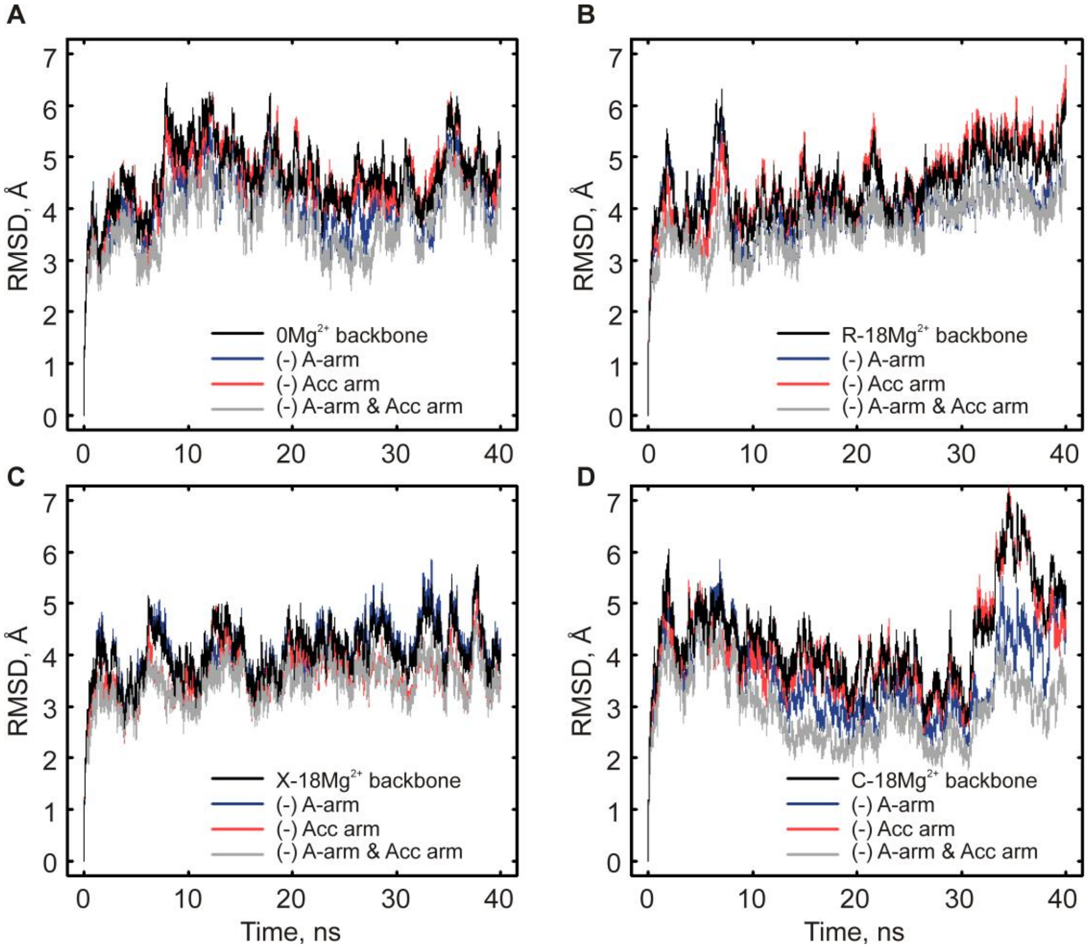

Figure A.1 Acceptor arm and anticodon arm contribution to global flexibility. RMSD was calculated for tRNA simulations at 310 K for 40 ns, including all backbone atoms (black) and backbone atoms excluding the anticodon arm ((-) A-arm, blue), the acceptor arm ((-) Acc arm, red), and both the anticodon and acceptor arms ((-) A-arm & Acc arm, grey). (A) 0Mg^{2+}, (B) R-18Mg^{2+}, (C) X-Mg^{2+}, and (D) C-18Mg^{2+}.

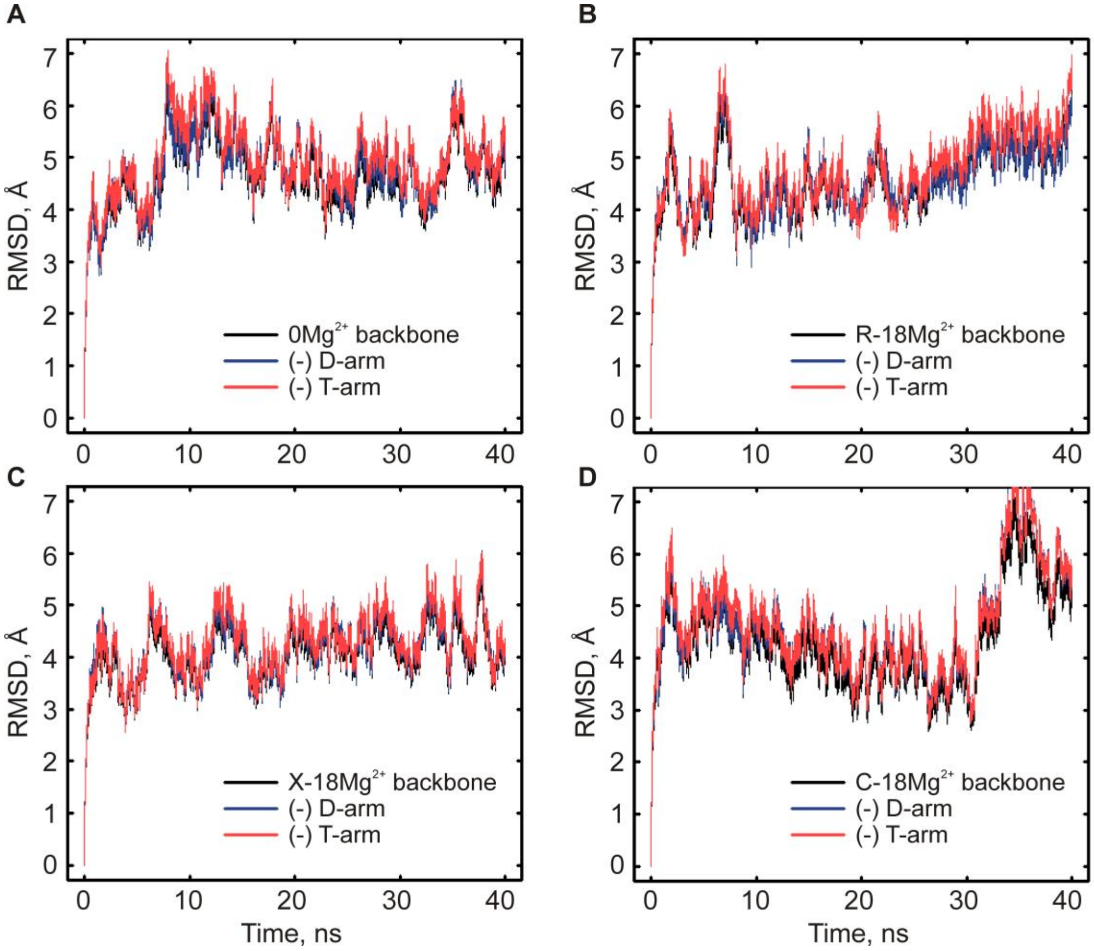

Figure A.2 D-arm and T-arm contribution to global flexibility. RMSD was calculated for tRNA simulations at 310 K for 40 ns, including all backbone atoms (black) and backbone atoms excluding the D-arm (blue) or the T-arm (red). (A) 0Mg^{2+}, (B) R-18Mg^{2+}, (C) X-Mg^{2+}, and (D) C-18Mg^{2+}.

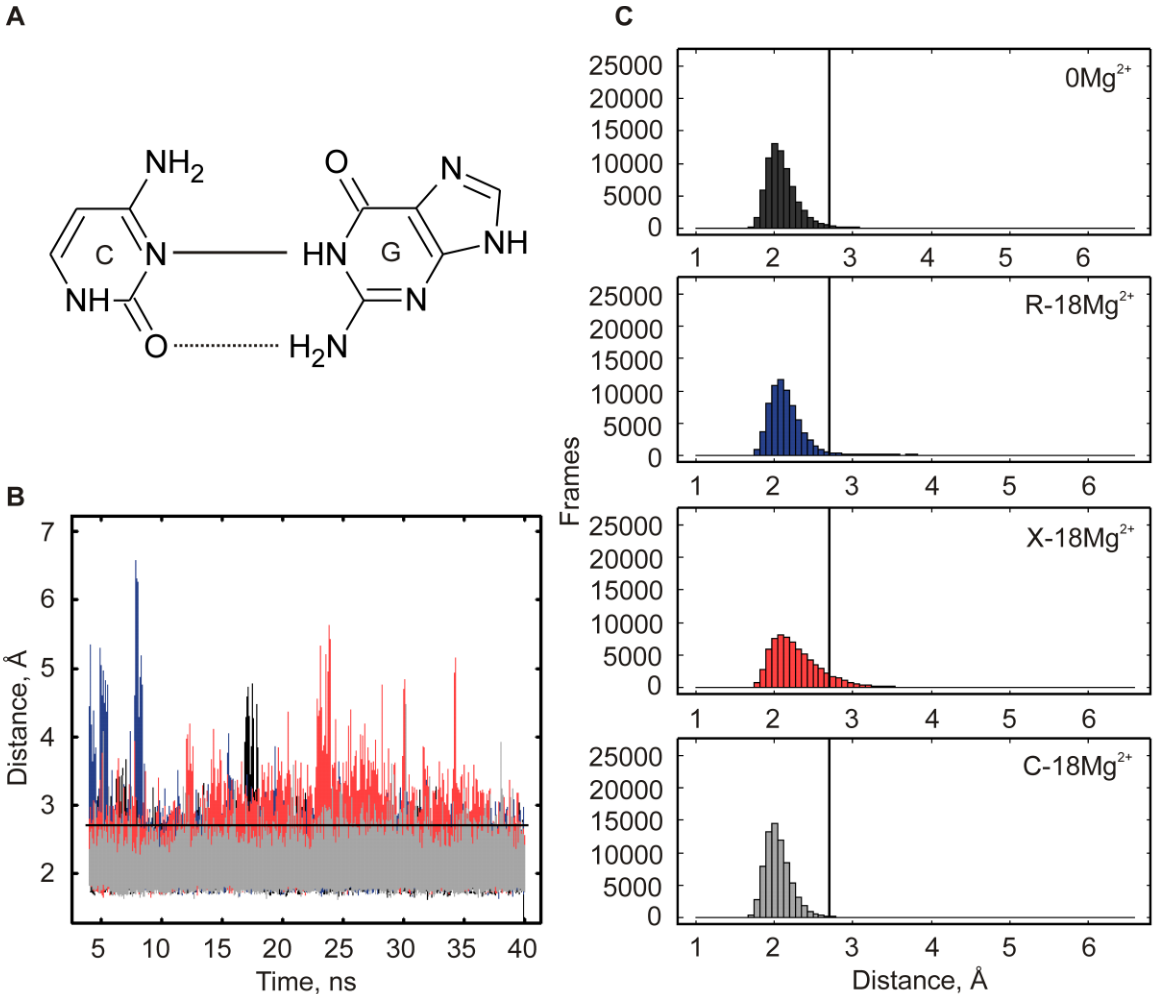

Figure A.3 Stability of the G15-C48 tertiary interaction. (A) The potential hydrogen bonds in the *trans* Watson-Crick G15-C48 base pair. Data in (B) and (C) represent the hydrogen bond denoted by the solid line. (B) Time-resolved changes and (C) histogram of the distance between G15 and C48 for the simulation of $0Mg^{2+}$ (black), $R\text{-}18Mg^{2+}$ (blue), $X\text{-}Mg^{2+}$ (red), and $C\text{-}18Mg^{2+}$ (grey). The black bar indicates a distance of 2.7 Å, the cut-off used to define a hydrogen bond.

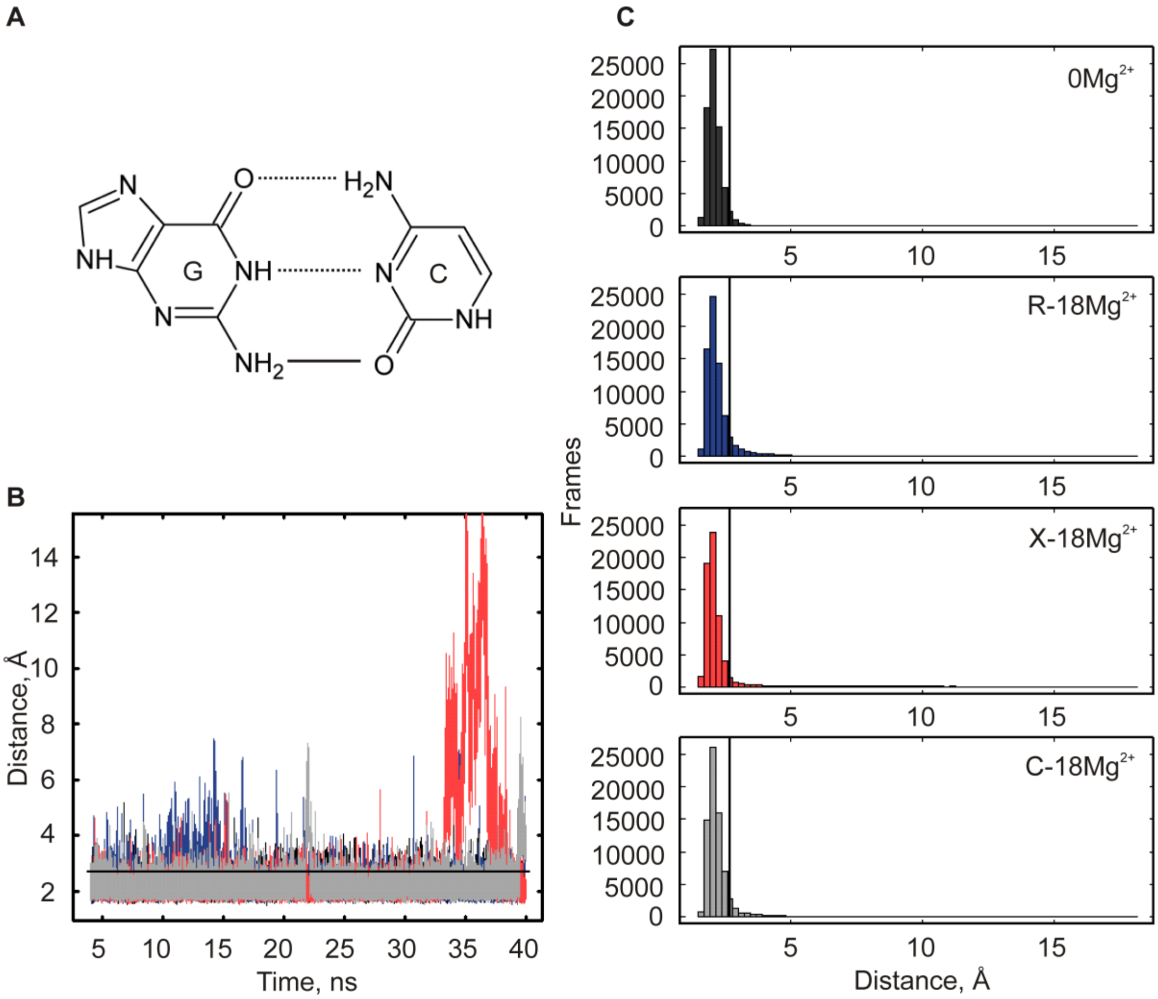

Figure A.4 Stability of the G19-C56 tertiary interaction. (A) The potential hydrogen bonds in the *cis* Watson-Crick G19-C56 base pair. Data in (B) and (C) represent the hydrogen bond denoted by the solid line. (B) Time-resolved changes and (C) histogram of the distance between G19 and C56 for the simulation of 0Mg^{2+} (black), R-18Mg^{2+} (blue), X-Mg^{2+} (red), and C-18Mg^{2+} (grey). The black bar indicates a distance of 2.7 Å, the cut-off used to define a hydrogen bond.

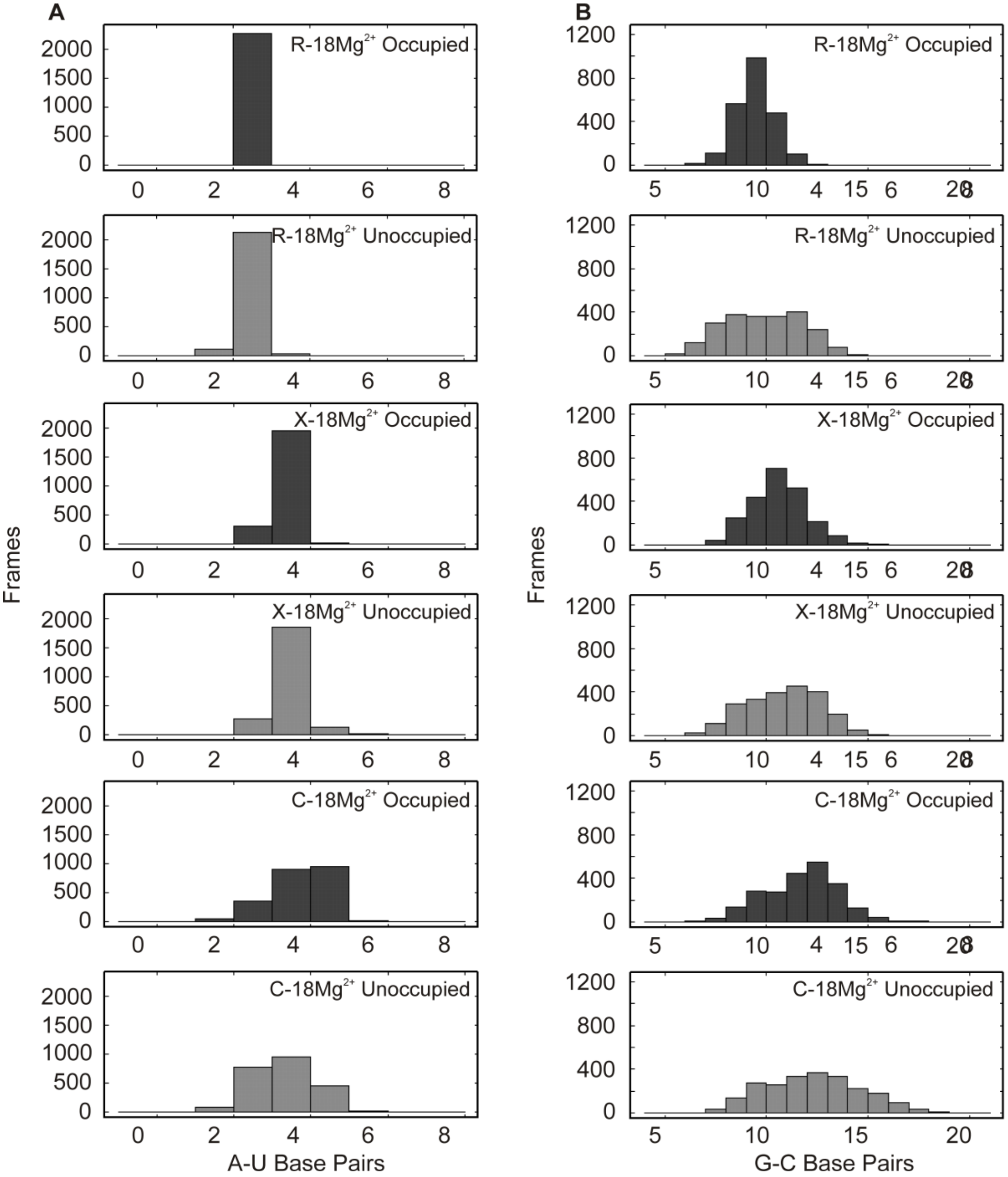

Figure A.5 Number of Watson-Crick base pairs in the occupied and unoccupied clusters. Hydrogen bond distances were calculated as in Fig. 3.4 to determine the number of A-U and G-C base pair through the simulations. Each bond that was within a 2.7 Å cut-off was counted as a base pair, regardless of the presence of the other bond(s) in the base pair (similar to Fig. 3.4). The data represented in the histograms corresponds to the same hydrogen bond as shown in Fig. A.3 and Fig. 3.4. Simulations were clustered based on Mg^{2+} occupancy at a specific coordinate in the three-dimensional model. (A) A-U and (B) G-C Watson-Crick base pairs are shown for the occupied (black) and unoccupied (grey) cluster for the R-18Mg^{2+}, X-18Mg^{2+}, and C-18Mg^{2+} simulations.

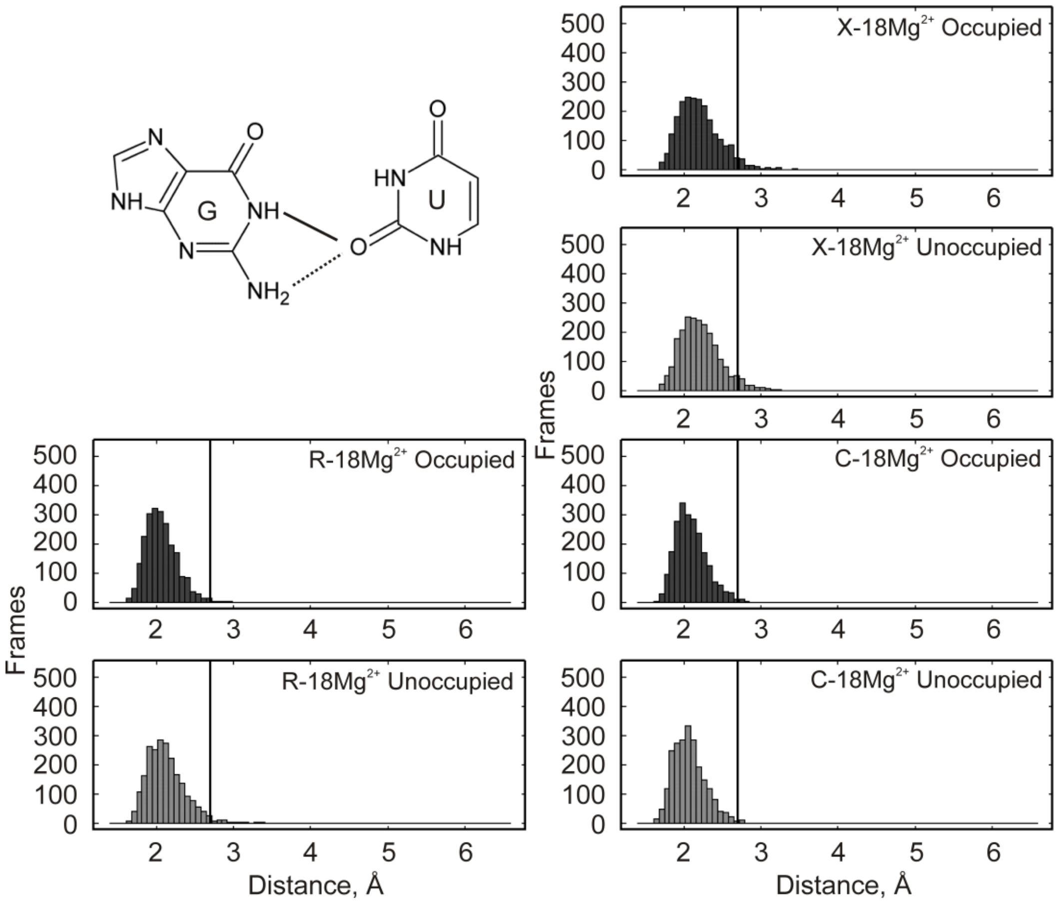

Figure A.6 Stability of G18-U55 base pair in the occupied and unoccupied clusters. Hydrogen bond distances were calculated for the *trans* Watson-Crick sugar-edge G18-U55 base pair as in Fig. 3.5 for the occupied (black) and unoccupied (grey) cluster for the R-18Mg^{2+}, X-18Mg^{2+}, and C-18Mg^{2+} simulations. The solid black line in the top schematic indicates the bond distance represented in the remaining panels. A black bar in each panel marks a distance of 2.7 Å used to define a moderately-strong hydrogen bond.

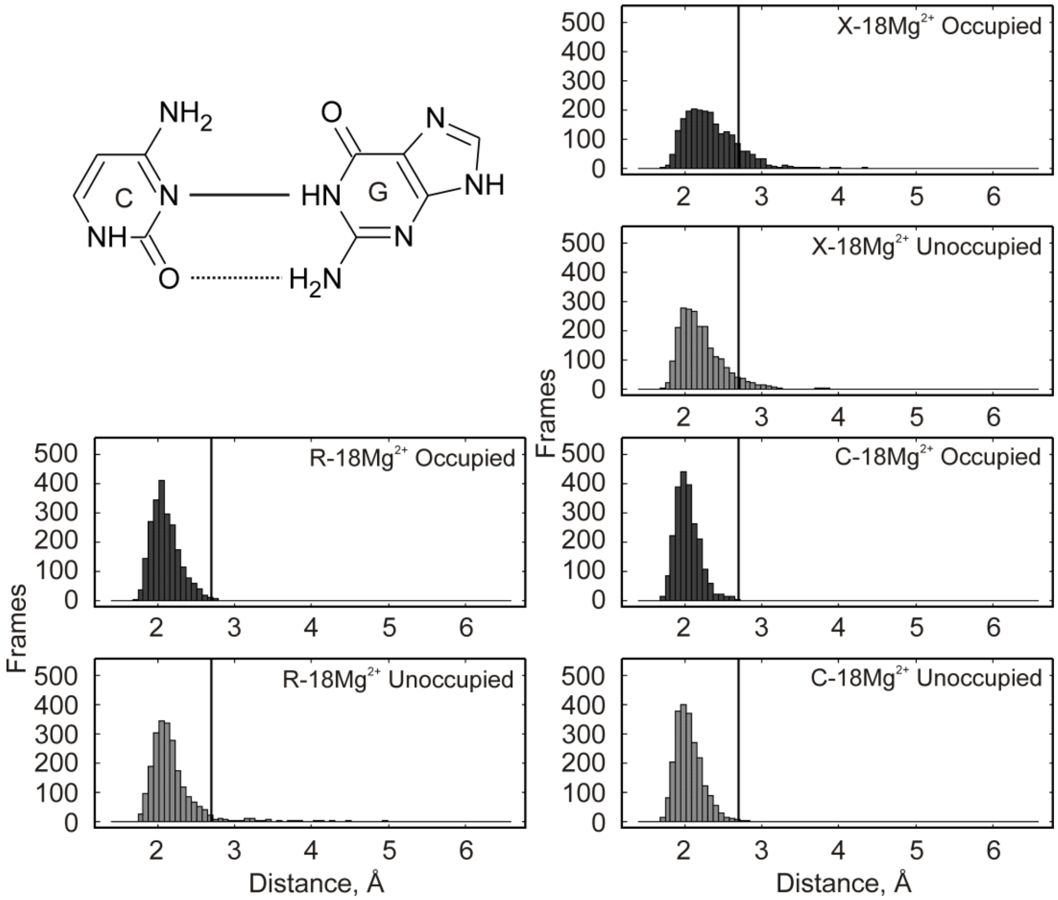

Figure A.7 Stability of G15-C48 base pair in the occupied and unoccupied clusters. Hydrogen bond distances were calculated for the *trans* Watson-Crick G15-C48 base pair as in Fig. A.3 for the occupied (black) and unoccupied (grey) cluster for the R-18Mg^{2+}, X-18Mg^{2+}, and C-18Mg^{2+} simulations. The solid black line in the top schematic indicates the bond distance represented in the remaining panels. A black bar in each panel marks a distance of 2.7 Å used to define a moderately-strong hydrogen bond.

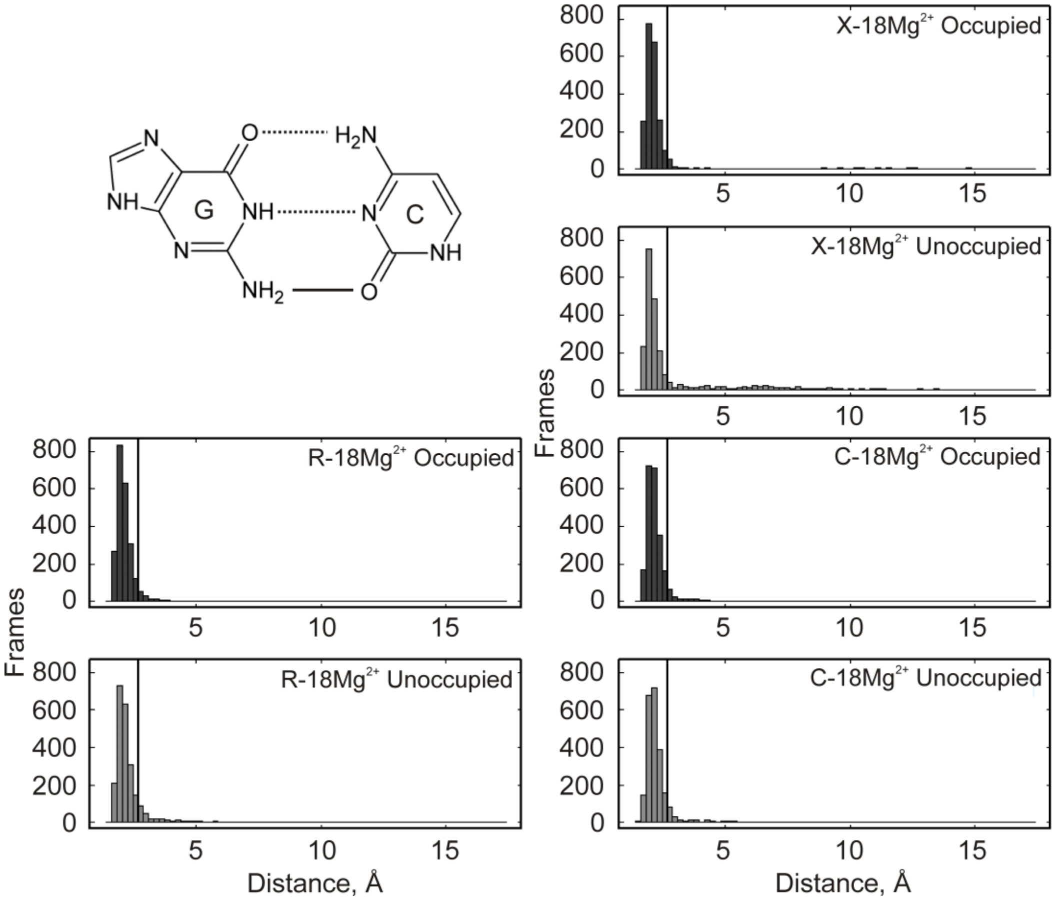

Figure A.8 Stability of G19-C56 base pair in the occupied and unoccupied clusters. Hydrogen bond distances were calculated for the *cis* Watson-Crick G19-C56 base pair as in Fig. A.4 for the occupied (black) and unoccupied (grey) cluster for the R-18Mg^{2+}, X-18Mg^{2+}, and C-18Mg^{2+} simulations. The solid black line in the top schematic indicates the bond distance represented in the remaining panels. A black bar in each panel marks a distance of 2.7 Å used to define a moderately-strong hydrogen bond.

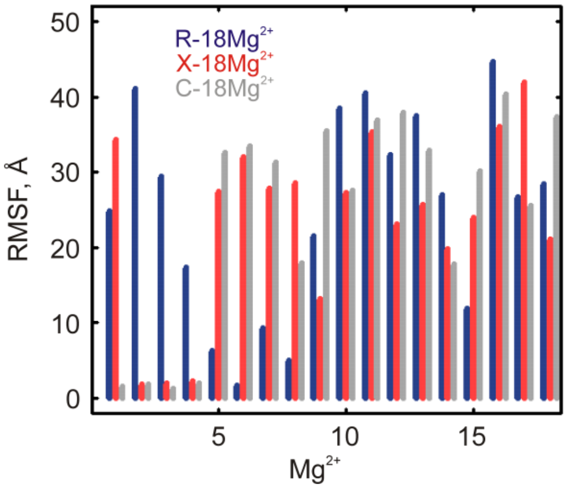

Figure A.9 Mobility of Mg^{2+} ions during simulation. The RMSF of the Mg^{2+} ions were calculated for the R-18Mg^{2+} (blue), X-18Mg^{2+} (red), and C-18Mg^{2+} (grey) simulations. For the X-18Mg^{2+} and C-18Mg^{2+} simulations, the first four ions were specifically placed within the tRNA structure. The remaining 14 ions for the X-18Mg^{2+} and C-18Mg^{2+} simulations and all ions from the R-18Mg^{2+} simulation were randomly placed in the surrounding water box no closer than 5 Å to the tRNA or other ions.