Single-Molecule Kinetic and Thermodynamic Studies of

Cation-Facilitated RNA Folding: Isolated Tertiary Interactions

by

Julie L. Fiore

B.S. and B.A., University of Pittsburgh, 2003

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Chemical Physics

2011

This thesis entitled:
Single-Molecule Kinetic and Thermodynamic Studies of Cation-Facilitated RNA Folding: Isolated Tertiary
Interactions
written by Julie L. Fiore
has been approved for the Chemical Physics program

_____
David J. Nesbitt

_____
Arthur Pardi

Date: _____

The final copy of this thesis has been examined by both the signatories, and we find that both the
content and the form meet acceptable presentation standards of scholarly work in the above
mentioned discipline.

Fiore, Julie L. (Ph.D. Chemical Physics)

Single-Molecule Kinetic and Thermodynamic Studies of Cation-Facilitated RNA Folding: Isolated

      Tertiary Interactions

Thesis directed by Professor David J. Nesbitt

RNA functional diversity is coupled with its ability to fold into unique structures, a process that is generally hierarchical—tertiary interactions occur between preformed secondary structural elements, e.g., loops and helices. For RNA to fold into compact, biochemically competent shapes, counterion neutralization of the negatively charged-phosphate backbone is required. The objective of this thesis is to investigate the physical principles that dictate how an RNA molecule achieves and maintains its tertiary structure. Toward this end, single-molecule fluorescence resonance energy transfer methods are combined with temperature control to probe the mediation of RNA folding landscapes by cation-facilitated tertiary interactions.

The primary focus of this thesis is kinetic and thermodynamic characterization of the ubiquitous GAAA tetraloop–receptor tertiary interaction using freely diffusing and immobilized single-molecule assays. The apparent first-order rate constants (k_{dock} and k_{undock}) for the intramolecular docking/undocking of the tetraloop and receptor are measured as function of monovalent, divalent, and trivalent cation concentration. We observe that the [cation] needed to promote folding is correlated with charge density of the ion, which we interpret in terms of counterion condensation on the RNA. The temperature dependence of k_{dock} and k_{undock} are also determined, which yield the standard state and transition state free energies, enthalpies, and entropies for docking and undocking. At physiological conditions, the transition state for tetraloop–receptor docking is early, with its formation rate-limited by an entropic barrier. The overall docking reaction is exothermic and entropically costly, consistent with the large number of hydrogen bonding and base-stacking interactions formed by the tertiary contact. Surprisingly, we reveal an entropic origin of Mg^{2+}-facilitated RNA folding, which conflicts with the common expectation that increasing [Mg^{2+}] aids folding by reducing electrostatic repulsions of the RNA

backbone. We propose instead that higher $[Mg^{2+}]$ promotes folding by decreasing the entropic penalty of counterion uptake in the folding transition state and by reducing disorder in the unfolded conformational ensemble. This work elucidates potential RNA folding paradigms, such as early transitions states and an entropic origin of tertiary cooperativity and cation-facilitated folding.

To my grandparents and

in loving memory of my mother

Acknowledgments

Without the tireless support of my doctoral advisor, David J. Nesbitt, this thesis would not be possible. David is a patient teacher, generous with both his time and knowledge. I appreciate the many hours he spent working with and advising me—poring over data, revising manuscripts, challenging me to both understand and communicate with the highest scientific standards, etc. My current goal is to implement what he has taught me in my new endeavors. I admire David's enthusiasm, intelligence, scientific expertise, and dedication to teaching. I am honored by his belief in my potential, and the encouragement he gave me to try to reach it. The many freedoms he gave me to pursue my curiosities and present my research have benefitted me tremendously.

This thesis has also come to fruition because of collaboration with Arthur Pardi, Christopher Downey, and Jose Hodak. Art and Chris designed the RNAs used in this work and offered insightful scientific advice throughout my graduate work. Chris taught me many things, including basic biochemistry techniques. Jose Hodak trained me in single-molecule methods and left me with a beautiful microscope system.

I would also like to thank my thesis committee—David Nesbitt, Carl Lineberger, Margaret Murnane, Arthur Pardi, and Thomas Perkins—for supporting me in this process and providing me with many useful comments and advice. I am especially grateful to Art and David for their comments on my thesis. I would also like to acknowledge my oral exam committee: David Nesbitt, Arthur Pardi, Eric Cornell, Ralph Jimenez, and Stephen Cundiff.

Many educators have encouraged me and helped keep my passion for science alive. It was in high school that I decided I would be a Chemistry major, inspired by a phenomenal teacher, Ronald Pollack. He told me that even if I thought I wanted to go to medical school, majoring in Chemistry would be wise and favorably viewed by admissions committees. He also told me that I would not be satisfied with a B.S. in Chemistry, so I should expect to pursue a Ph.D. As it turns out, he was right on both counts; I would not be satisfied until I got a Ph.D. and this trajectory did successfully lead me to medical school. It

Biochemistry for allowing me the opportunity to study here. I would especially like to thank the graduate secretary, Cori Fagan-Edminster, for staying so on top of graduation logistics. I thank Robert Parson, director of the Chemical Physics Program, and David Jonas for additional advising.

Being a part of the JILA community has been truly special to me. I would like to extend a special thanks to the JILA shops and staff for their hard work and dedication. In particular, I offer my gratitude to the following people for technical and/or emotional support: Computing—James R. Raith, Michael Paige, Joel Frahm, Alan Dunwell, James Mckown, and Peter Ruprecht; Instrument Specialists—Hans Green, Blaine Horner, and Ariel Paul; Electronics—James Fung-a-Fat, Carl Sauer, J.M. Whitmore, David Tegart, and Terry Brown; Adminstration—Maryly Dole, Pam Leland, Agnieszka Lynch, Loree Kaleth, and Krista Beck; Supply—Jeffrey Sauter, David Errickson, Brian Lynch, and Randall Holliness; and in the Scientific Reports Office—Julie Phillips and Lynn Hogan.

I would like to acknowledge all past and present Nesbitt laboratory members, a large group of intelligent people with whom I have had the pleasure of interacting. Their support, knowledge, and good nature have made this process enjoyable, even in the most stressful times. In this regard, I extend my gratitude to the current Nesbitt group: Mike Ziemkiewicz, Thomas Baker, Melanie Roberts, Larry Fiegland, Rob Roscioli, Erik Holmstrom, Andrew Gisler, Andrej Grubisic, Kevin Early, Nicholas Dupuis, Daniel Nelson, Mia Zutz, Grant Buckingham, and Chi-Hshuan Chang. Of these, I would like to give special acknowledgment to Mike for his many pep talks; I appreciate his refreshingly inquisitive personality and admire his pure love of learning. Tom's vast knowledge and scientific prowess have been an invaluable resource for me. Melanie has been there to offer encouraging words and humor on even the hardest days. I have particularly enjoyed seeing the RNA constituency of the lab grow. In that regard, I appreciate the new techniques that Larry introduced to the lab and have enjoyed working with and learning from him. Erik—sharp, confident, and creative—has been a pleasure to collaborate with on some of the projects in this thesis. Nick Dupuis has been a breath of fresh air in the laboratory. I am grateful for his enthusiasm, wisdom, and support, especially in the final weeks of writing and preparing for my defense. I would also like to acknowledge past Nesbitt group members: Erin Whitney, Oliver Monti, Jose Hodak,

ix

Michael Deskevich, Bradford Perkins, Alex Zolot, Vasiliy Fomenko, Feng Deng, Oliver Piestert, Jens-Peter Kneyemer, Nicole Marme, Chandra Savage, Richard Walters, Christian Pluetzer, Haemi Lee, Stephanie Meyer, Shervin Rahimpour, Kevin Dean, Jessica Rouge, Danny Bell, Kasia Kobeszko, Erin Sharp, Jeff Peterson, Volker Schweikhard, and Lora Nugent-Glandorf. I am particularly grateful to Brad for his friendship and advice. I extend my gratitude to Vasiliy Fomenko, who helped me in countless ways during graduate school.

I would also like to offer my gratitude to my colleagues in the fields of RNA and Biophysics, who encouraged and inspired me. I am particularly grateful to Samuel Butcher, Arthur Pardi, Daniel Herschlag, Nils Walter, Vincent Chu, D.M.J. Lilley, Robert Batey, David Rueda, Goran Pljevaljcic, Elvin Aleman, Ashok Deniz, Thomas Perkins, Kathleen Hall, and Michael Yarus for their support and/or interest in my work.

I cannot imagine how I could have done this without my friends and family. In particular, Martha Hosotani, Emily Steigerwalt, Stephanie Meyer, Haemi Lee, Amanda Carpenter, Lora Nugent-Glandorf, Juan Pino, Adam Pearlstein, Ashley Carter, Carrie DeFurio, James Fung-a-Fat, Carl Sauer, and the Wagner family have been an invaluable support group. It is with Martha that I share my favorite memories of graduate school. I appreciate her advice, scientific or otherwise. Stephanie, with her easy going personality, always managed to make me feel happy and proud. I am thankful to Amanda for her encouragement; I admire her strength and appreciate her point of view. I am also lucky to have a large and loving family. I would especially like to thank my father, Kevin Fiore; and brothers Thomas and Daniel Fiore; my aunt Maureen; and my grandparents, Stephen and Theresa Gribschaw, Marshall Fiore, and in memory, Ruth Fiore. My mother was my biggest supporter, believing in me and my every dream. Her work ethic, generosity, and love have been my inspiration.

Last and certainly not least, I would like to acknowledge Samuel Butcher—the love of my life, my crew, my colleague, my best friend, and my partner. I am grateful for his unwavering love, encouragement, and support, particularly during this challenging and selfish time. Most of all, I am thankful for his unfaltering belief in me.

Table of Contents

Chapter 6   The Role of Counterion Valence and Size in GAAA Tetraloop–Receptor Docking/Undocking Kinetics

Tables

Figures

filtered through a confocal pinhole, and detected by avalanche photodiodes. (*B*) Schematic of the microscope (courtesy of Larry Fiegland) and a depiction of the laser focused by an objective through cover glass into a solution containing fluorescently labeled molecules (not to scale). In this work donor and acceptor signals (*green* and *red*) are summed over horizontal/vertical polarization channels.

and K'_{Mg} are Mg^{2+}-dissociation constants and the rate constants reflect docking and undocking resolved by FRET. (*C*) Simplified parallel model to describe [Na$^+$] and [Mg^{2+}]-dependence for the observed fraction of docked molecules with Mg^{2+} and Na$^+$ dissociation constants.

Figure 3.6 E_{FRET} distributions and Gaussian fits (*black*) showing, donor-only (*leftmost peak*), undocked (*green*) and docked (*red*) populations at (*A*) 25 mM Na$^+$ and (*B*) 1.0 M Na$^+$, with shot-noise predictions for each peak (*blue dashed lines*). Note that the undocked peak shifts to higher center E_{FRET} value and broadens with increasing [Na$^+$] (see text for details).

Figure 3.7   (*A*) Least squares fits of fractional docked population ($N_{docked}/(N_{docked} + N_{undocked})$) versus [Na$^+$] to Eq. 3.7, resulting in f_{max} = 0.55 ± 0.05, a Hill coefficient 1.3 ± 0.3, and K_D = 180 ± 30 mM. The asymptotic value (f_{max}) is consistent with Mg^{2+} studies in Figure 3.5, suggesting a ≈ 32–34% non-docking RNA subpopulation. (*B*) f_{free} as a function of [Mg^{2+}] at low [Na$^+$] (25 mM) with a fit to Eq. 3.10 that also allows for a [Na$^+$] docking pathway (Figure 3.5 *C*), yielding f_{max} = 0.55 ± 0.04, n = 8 ± 2, $K_D^{Mg^{2+}}$ = 1.06 ± 0.03 mM, and demonstrating high cooperativity with respect to Mg^{2+} observed under minimal Na$^+$ conditions.

Figure 3.8   Evidence for a cation-induced increase in electrostatic compaction and conformational sampling of the undocked state tetraloop–receptor construct. (*A*) and (*B*) Systematic shift in mean E_{FRET} of the undocked peak with increasing [Na$^+$] and [Mg^{2+}], respectively, fit by a Hill-type model (Eq. 3.11) with $E_{FRET}^0(Na^+, Mg^{2+})$ = 0.227 ± 0.004, 0.227 ± 0.003; $E_{FRET}^0(Na^+, Mg^{2+})$ = 0.15 ± 0.02, 0.07 ± 0.02, n(Na$^+$, Mg^{2+}) = 2.1 ± 0.4, 2.6 ± 0.8 and K_D(Na$^+$, Mg^{2+}) = 180 ± 20 mM, 0.9 ± 0.2 mM. (*C*) and (*D*) Systematic shifts in undocked E_{FRET} peak widths as function of [Na$^+$] and [Mg^{2+}], respectively, yielding K_D(Na$^+$, Mg^{2+}) = 0.23 M ± 0.02, 1.2 ± 0.4 mM; n(Na$^+$, Mg^{2+}) = 3.6 ± 0.8, 2.7 ± 1.2; and $\Delta\sigma$(Na$^+$, Mg^{2+}) = 0.10 ± 0.01, 0.07 ± 0.03, respectively.

Figure 3.9   Calculated Debye shielding lengths in the presence of 50 mM hemisodium HEPES buffer with the addition of (*A*) [NaCl] in the absence MgCl$_2$ and (*B*) [MgCl$_2$] without and with 100 mM NaCl.

predictions (*colored lines*) and Gaussian fits (*black lines*). The shot-noise predicted width, σ_{SN}, is calculated from the standard-error propagation of finite photon counting statistic, as $\sigma_{SN} = (E_m(1-E_m)/T)^{1/2}$, where E_m is the center of the Gaussian peak and T is the minimum average count rate (photons/bin) for molecules included in the analysis (161). The undocked state is significantly broadened beyond shot noise ($\sigma_{undocked}/\sigma_{SN} = 1.4 \pm 0.1$) while the docked peaks are nearly shot-noise limited ($\sigma_{docked}/\sigma_{SN} = 1.06 \pm 0.08$). This observation is consistent with a larger conformational accessibility of the tetraloop in the undocked vs docked states, as was discussed in detail elsewhere (128).

129

Chapter 1   Introduction

The correlation between RNA structure and its biological function offers diverse opportunity for RNA-based pharmaceuticals and biotechnology (1-4). Central to such applications is prediction of RNA three-dimensional structures.   RNA folding is generally hierarchical—functional structures are achieved through tertiary interaction of preformed secondary elements, e.g., flexible junctions enable unpaired nucleotides to act as beacons for helix recognition and packing. Secondary structure prediction from the known thermodynamics is quite reliable (5), though prediction of tertiary structure remains a major challenge (6).   Moreover, static tertiary structures alone are not enough to predict function, as time-dependent structural dynamics occurring during biochemical processes can also be important (2,3). As a result, the kinetics and thermodynamics of RNA folding must be addressed, i.e., the energy landscape for folding must be characterized. Toward this end, it is necessary to explore the kinetic and thermody-namic properties of individual and combined tertiary interactions.   Furthermore, formation of a com-pact RNA structure requires counterions to screen the electrostatic repulstion of the negatively charged phosphate backbone (7-12). However, little is known about how cations affect folding landscapes. In this thesis, temperature-controlled single-molecule fluorescence energy transfer methods are used to probe the real-time docking and undocking of a ubiquitous tertiary interaction, the GAAA tetraloop–tetraloop receptor motif. We assess the energetic differences between unfolded, transition, and folded states as a function of cation concentration (e.g., [K^+] and [Mg^{2+}]). With such information, we offer in-sights into cation-mediated RNA folding landscapes.

1.1   RNA: More than a Messenger

According to the central dogma of molecular biology, genetic information is uni-directionally relayed from DNA to RNA by transcription and from RNA to proteins by translation, while the genome is preserved and propagated by DNA replication, Figure 1.1 (13) . Yet RNA is much more than the messenger of the genetic code; it is implicated in essentially all cellular processes, e.g. regulating transcription and translation (2,14-17). RNA is even the genomic entity in many viruses and can be used as a template for reverse transcription, i.e., retroviruses can transcribe DNA from RNA, Figure 1.1 (18,19). Astoundingly, as much as 93% of human DNA is transcribed into RNA (20,21), though only about 1.5% of the human
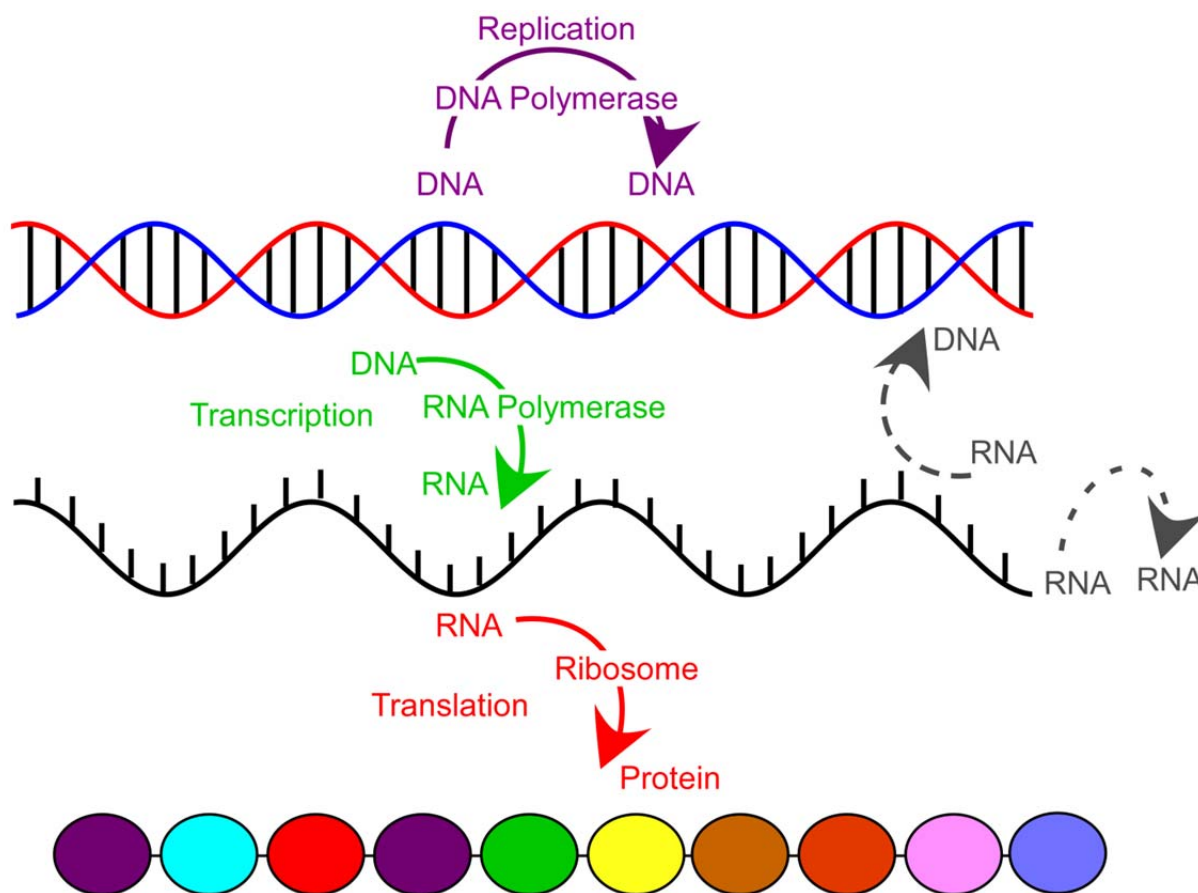
Figure 1.1   The central dogma of molecular biology. DNA is replicated by DNA polymerase and transcribed into RNA by RNA polymerase. RNA is translated into proteins by ribosomes. In special viral cases, RNA is replicated by RNA-dependent RNA polymerase. In retroviruses, DNA is transcribed from RNA by reverse transcriptase.

genome codes for protein, i.e., only a small fraction of human DNA represents "genes" that are phenotypically manifested (22). The fraction of non-coding DNA (or RNA) in a genome generally correlates with species complexity and may have an evolutionary advantage (23-26). The discovery that RNA can act as an enzyme, or ribozyme, supports the idea that the excess non-coding RNA may serve important biological functions (27,28). Indeed, even the ribosome, the site of protein synthesis, is an RNA catalyst (16,29). With the ability to carry genetic information and catalyze reactions, RNA may be a primordial molecule that supported pre-cellular life (17,30). According to this "RNA world" hypothesis, the first biotic world was made up of RNA, with DNA later evolving as a more robust harbor for the genetic code. Similarly, the evolutionary need for diversity and enzyme stability may have led to a larger set of molecules capable of catalysis, i.e., proteins (17,30).

To explore the functional repertoire of RNA beyond messaging, we need to look no further than the non-coding regions of the RNA transcript. In messenger RNA (mRNA), the long non-coding regions, known as introns, must be removed for proper translation of the gene by the ribosome. Similarly, to make other types of RNAs, e.g., transfer RNA (tRNA) or ribosomal RNA (rRNA), intron regions of the RNA transcript must be excised (Figure 1.2 *A*). Excision of the introns and subsequent splicing of the functional regions can be catalyzed by the intron itself, for which two types of mechanism are known, defining group I and group II introns (31). The splicing mechanism of group I introns is shown in Figure 1.2 *B*; a guanosine co-factor is utilized for two trans-esterification steps that lead to release of the intron and joining of the exons (32). To form the active site for guanosine attack, the self-splicing ribozyme must fold into an intricate shape, as depicted in the structure of the *Tetrahymena thermophila* ribozyme in Figure 1.2 *C* (33). Indeed, RNA versatility is coupled with its ability to assemble into complex three dimensional structures, which create unique sites for catalysis and molecular recognition. This synergistic connection between RNA structure and function offers diverse opportunities for RNA-based biotechnologies and pharmaceutical development, central to which is understanding how RNA acquires and maintains its shape, i.e., RNA folding (1-4).
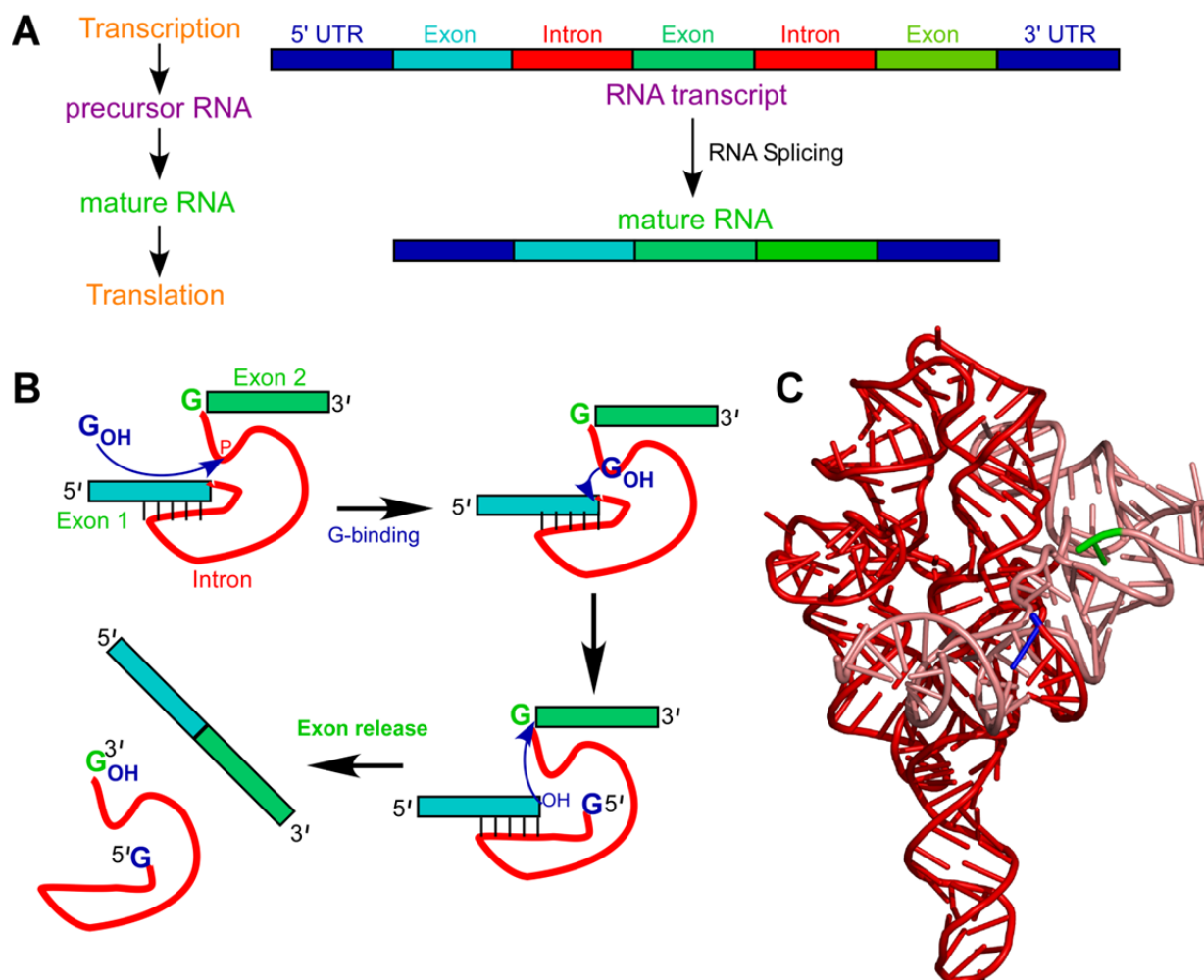
Figure 1.2   RNA splicing and self splicing by Group I introns. (*A*) RNA is transcribed from DNA into precursor RNA, which contains excess non-coding regions (introns) that are removed prior to nucleus export for translation into proteins or other functions. (*B*) Mechanism of group I intron self-splicing ribozymes: A free guanosine attacks the 5′ phosphate at the active site (labeled P), which allows for subsequent nucleophilic attack at exon 1, freeing the 5′ end of the intron while exon 1 maintains it position with respect to the intron through base-pairing interactions. A second OH attack occurs freeing the 3′ end of the intron and splicing together the two exons for mature mRNA. (*C*) The active splicing site is achieved by the complex 3D shape of the intron. Shown is the *Tetrahymena thermophila* ribozyme with the active site (5′G) shown in *blue* and the 3′ G shown in *green* to match *B* (PDB ID 1X8W). The RNA phosphate sugar backbone is drawn as a *tube*, and the bases are shown as *sticks*.

1.2   RNA Folding

RNA, like its DNA cousin, is a polymer composed of four nucleotide building blocks—each contains one

of four nucleobases (adenine (A), guanine (G), cytosine (C), or uracil (U)), a five-carbon sugar (ribose),

and a phosphate (Figure 1.3). The ribose 2′-hydroxyl that DNA lacks is crucial to RNA's more diverse

functionality and also makes RNA more prone to degradation by hydrolysis than DNA. The nucleotide

sequence of the RNA makes up its primary structure (Figure 1.4). Just as in DNA, nucleotides can hydro-

gen bond through the bases to form secondary helical structure akin to the famous DNA double helix

(Figure 1.3). A perfectly base-paired RNA helix is A-form—12 basepairs/turn. However, RNA, unlike

DNA, is generally single-stranded, i.e., it lacks a complementary strand such that it loops and folds onto

itself to make the base-pairing interactions of secondary structure (Figure 1.4). The optimization of
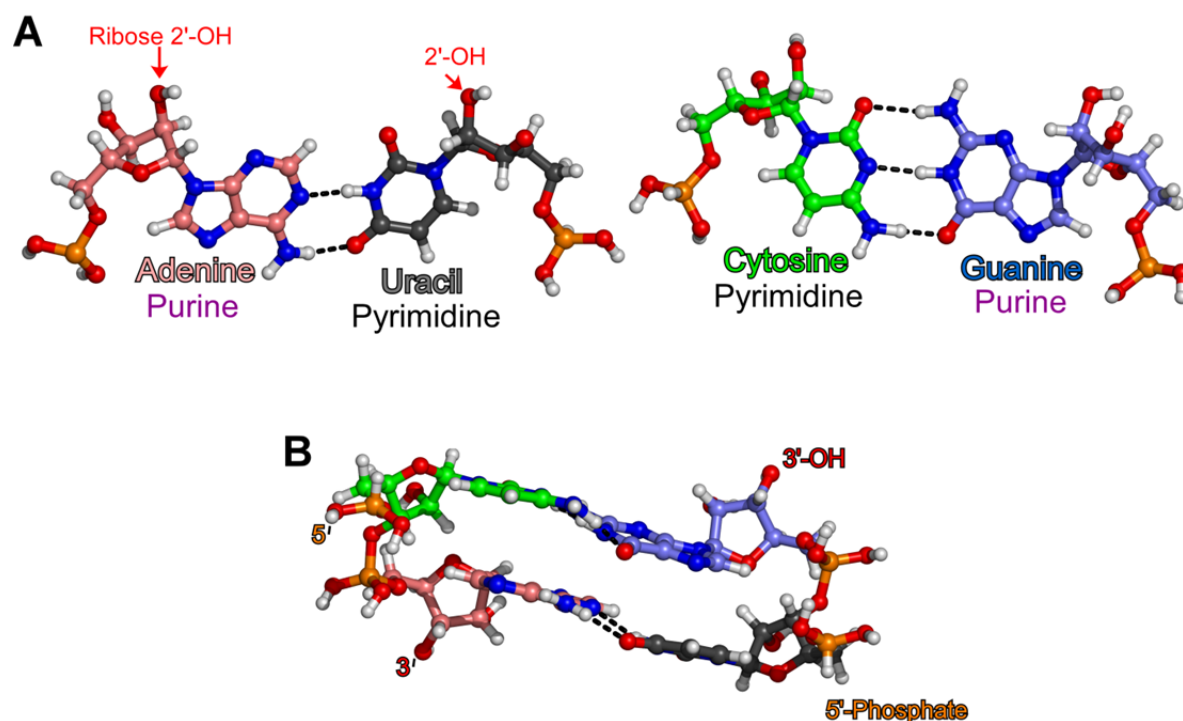
Figure 1.3   Base pairing of the four RNA nucleotides consisting of a ribose, phosphate, and a nucleobase: adenine (A), uracil (U), cytosine (C), and guanine (G). (*A*) Purines (A or G) Watson-Crick base pair with pyrimidines (U or C) through two or three hydrogen bonds, respectively (H-bonds shown as *black dashes*). The ribose 2′-OH is indicated and atoms are colored: C is the color of the residue, P is *orange*, N *blue*, O *red*, H *white*. (*B*) RNA has a 5′−3′ directionality with base pairing interactions enabling formation of a double helix.

A

5′ GCGGAUUUAGCUCAGUUGGGAGAGCGCCAGACUGAAGAUCUGGAGGUCCUGUGUUCGAUCCACAGAAUUCGCACCA 3′

Primary Structure

Secondary Structure

Tertiary Structure

B

Helix

Junction

Helix

Unfolded

Tertiary Folding

Folded
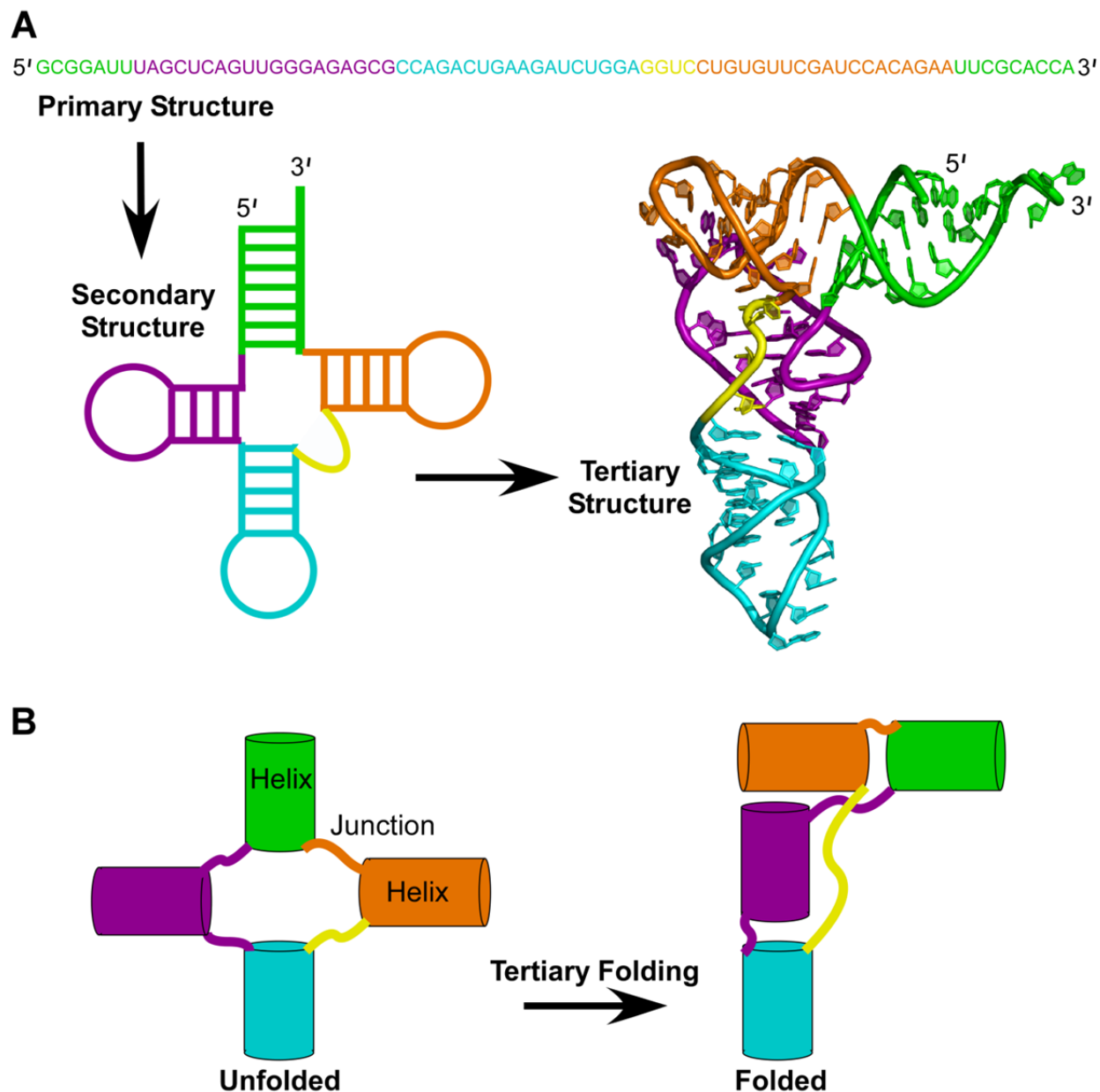
Figure 1.4   The hierarchy of RNA structure depicted for yeast phenylalanine tRNA (PDB ID 6TNA). (*A*) The nucleotide sequence makes up the RNA primary structure. The single-stranded RNA folds onto itself to optimize the number of base pairs (*short lines*), forming unpaired hairpin loops and bulges. Interactions between unpaired regions and/or 2′-hydroxyls lead to tertiary structure (crystal structure shown). The RNA phosphate sugar backbone is drawn as a *tube*, and the bases are shown as *rings*. (*B*) Schematic of tertiary folding in tRNA. Formation of secondary structure results in helical regions and connected by flexible junctions that allow the helices to interact and form the tertiary structure crucial to functionality.

base-pairing interactions results in the formation of relatively rigid helical regions capped by loops

connected by unpaired and thus much more flexible junctions. Interactions between the helical regions, i.e., tertiary interactions, enable the formation of compact ordered structures, as shown for the characteristic cloverleaf shape of secondary structure and L-shaped tertiary structure of tRNA (Figure 1.4) (34).

This process of RNA folding, i.e., formation of higher order structure, is generally hierarchical with tertiary structure occurring through interactions of preformed secondary structure elements (6,35-37). From the example of tRNA (Figure 1.4), this general theme of how RNA folds is revealed—flexible junctions enable unpaired nucleotides to act as beacons between helical regions for structural assembly—a helix–junction–helix secondary structure packs into a specific tertiary structure (Figure 1.4 *B*). Obtaining the correct 3D structure is critical to RNA function. For example, the functional fold of tRNA allows transport of amino acids through the ribosome for peptide linkage in a growing protein (34). In addition to requiring tertiary interactions to guide assembly (6,36,38,39), RNA must overcome a severe frustration—folding is opposed by the electrostatic repulsion of the negatively charged phosphate backbone. Therefore, the ability of RNA molecules to achieve compact, functional structures depends inherently on counterion neutralization and tertiary interactions (12,40). Thus, cations (e.g., K^+ and Mg^{2+}) can play a critical role in RNA folding, through both site-specific chelation and diffuse interactions through a hydration shell (40,41).

1.2.1 Tertiary Motifs: Packing Strategies of RNA

Tertiary interactions stabilize folded RNA structures (42) and can be categorized into relatively few structural motifs, as characterized by their involvement of secondary structure features: coaxial helical stacks (as seen in Figure 1.4 *A*), kissing hairpins, tetraloop–receptor interactions, A-minor motifs, pseudoknots, loop–loop interactions, and ribose zippers (37,43). Interactions involving loops are prevalent because hairpin loops are common secondary structure elements employed as caps to helical regions (Figure 1.4 *A*). The most common such loops are tetraloops of the sequence GNRA, where N is any nucleotide and R is a purine—comprising one-third of the tetraloops in ribosomal RNA and half of the te-

traloops in some catalytic RNAs (37,44-47). GNRA tetraloops are exceptionally stable, a property at-

tributable to a characteristic U-turn structure, i.e., a sharp bend in the backbone between the G and N

nucleotides (Figure 1.5 *A*), allowing for hydrogen bonding and base-stacking within the loop (48-51).

GNRA loops form long-range tertiary contacts, i.e., interactions with distal receptor regions of the RNA
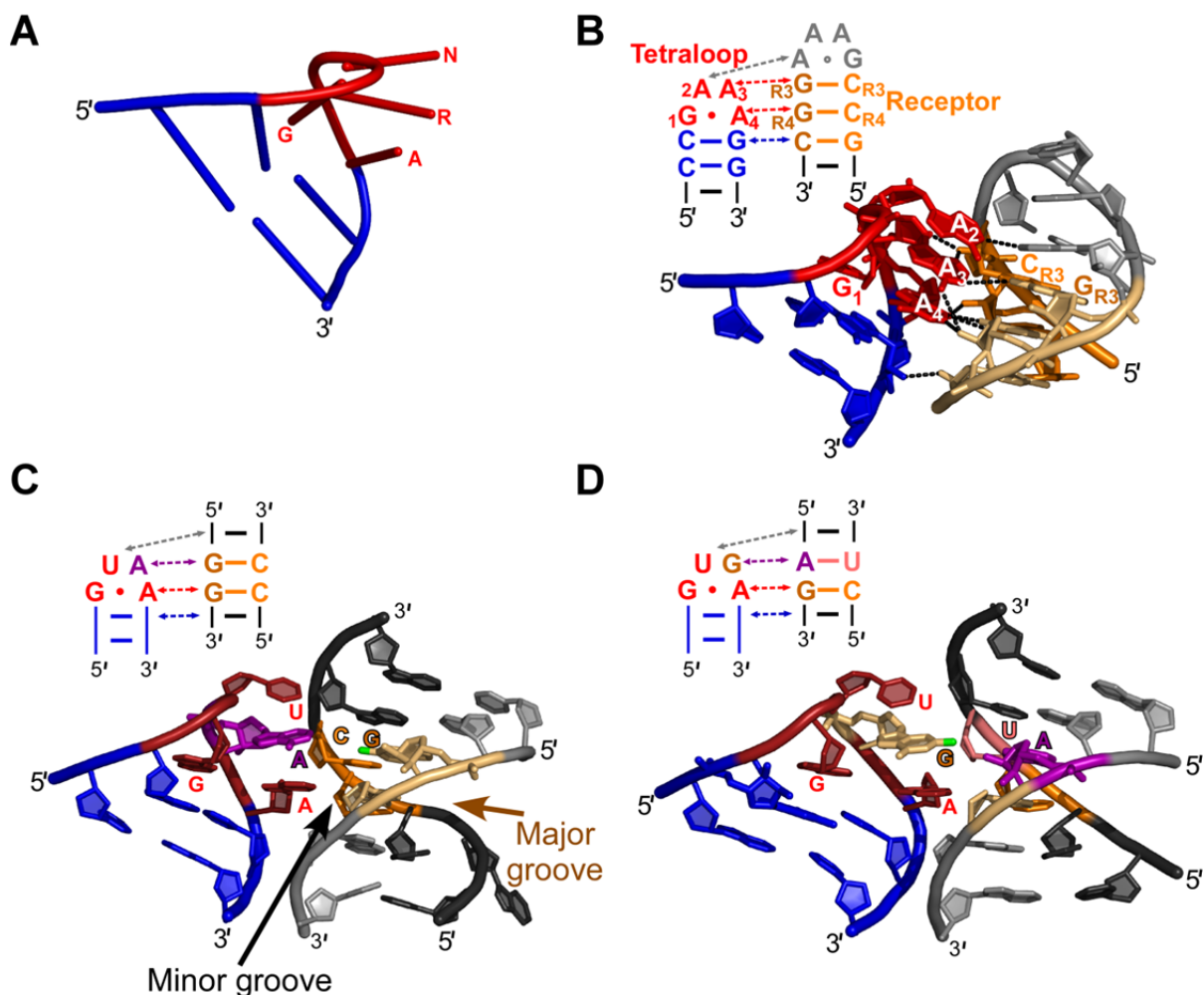
Figure 1.5   GNRA tetraloops engage in interactions with the helical minor grooves. (*A*) NMR structure of a GAAA tetraloop is representative of the GNRA U turn structure that exposes the Watson Crick edges of the last NRA nucleotides for binding (PDB ID 1ZIG). (*B*) GAAA tetraloop (*red*) interaction with tandem C:G basepair (*orange*) in a helix minor groove as observed in intermolecular crystal contacts of the hammerhead ribozyme (PDB ID 1HMH). The 9 possible intermolecular hydrogen bonds are shown as *dashed black lines*. (*C*)-(*D*) Other examples of GNRA tetraloop–helix interactions as observed in the crystal structure of RNase P(PDB ID 3Q1Q). The GUAA loop binds at a CC:GG basepairs and GAGA binds at CU:AG basepairs, a common phylogenetic preference that accommodates a bulky guanine C2 amino group (bright green). *Short lines* and *circles* indicate Watson-Crick and noncanonical basepairs, respectively.

(44,45,52). The structures of the GNRA loops share potential recognition features—the Watson–Crick base pairing edges of the last three bases of the loop are exposed, and the unique backbone contour can provide shape recognition, as shown in Figure 1.5 *A* (51,53-55).

The simplest engagement of a GNRA tetraloop in a tertiary interaction is with the minor groove of continuous helical "receptor." In this case, the tetraloop inserts into the wide and shallow minor groove, which is more accessible than the major groove in an RNA helix, Figure 1.5 *C* (45,56). The structural basis for such an interaction was revealed in the 2.6 Å hammerhead ribozyme crystal structure, in which intermolecular contacts were observed between a GAAA loop and a helix minor groove at the site of two consecutive purines, shown in Figure 1.5 *B* (57,58). This packing strategy is stabilized by a hydrogen bonding network (9 potential hydrogen bonds, though at least 3 are expected to be relatively weak, i.e., with bond lengths > 3 Å), as shown in Figure 1.5 *B*. Insertion of adenosines into distal minor grooves was later recognized as the highly utilized A-minor motif (59,60). A-minor motifs are the most abundant tertiary motifs known in RNA structures, e.g., 186 such interactions are identified in the large ribosomal RNA subunit (59,61). Four varieties of A-minor motifs have been identified, the most common type I (61) is observed in the GAAA–helix interaction in the A4 nucleotide (Figure 1.5 *B*). Type I is defined by the nestling geometry of the adenine into the minor groove, which optimizes hydrogen bond opportunities (57). Most of the interactions in the tetraloop–helix interaction involve the ribose 2′ hydroxyls, highlighting the importance of this characteristic difference between RNA from DNA.

The P4–P6 domain of the *Tetrahymena thermophila* ribozyme (shown in *red* in Figure 1.2 *C)* also reveals themes of helical packing through tertiary interactions, including the involvement of a GNRA tetraloop in a tertiary interaction. The domain folds independently of the full ribozyme, making a hinged structure stabilized by two major tertiary contacts, a GAAA tetraloop–11 nucleotide (nt) receptor and an A-rich bulge–helix interaction, Figure 1.6 (62).   Both of these interactions utilize the more broadly characterized tertiary motifs, the A-minor and ribose zipper motifs, and are therefore very interesting for exploring the general physical principles that govern RNA folding. Indeed, the GAAA tetraloop–11 nt receptor (T(GAAA)–R(11nt)) motif is the strongest, most specific, and widespread GNRA–receptor in-
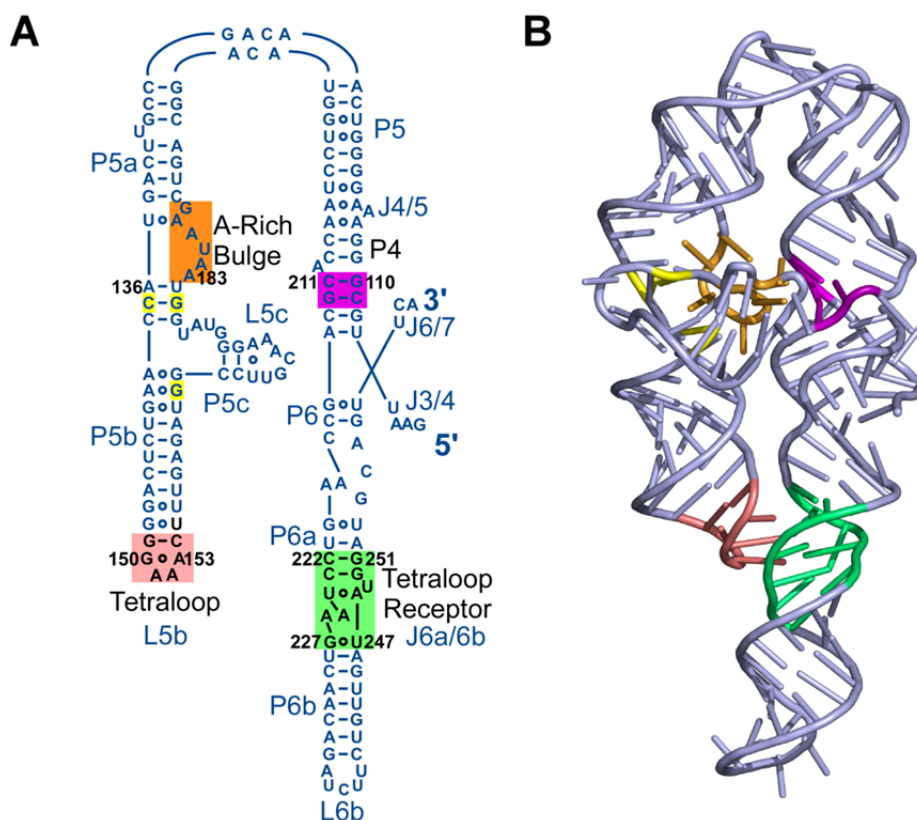
Figure 1.6   Secondary structure (*A*) and crystal structure (*B*) of the *Tetrayhmena thermophila* P4−P6 domain (PDB ID 1GID) highlighting the GAAA tetraloop−tetraloop receptor (*pink* and *green*) and A-rich bulge−P4 helix (*orange* and *purple*) interactions. Nucleotides that stabilize the A-rich bulge are highlighted in yellow. Junction regions are labeled J, helices are labeled P, and loops as L.

teraction, stabilizing the folded structures of group I and group II Introns and RNase P (62-68). The asymmetric A-rich bulge is also a highly utilized motif in group I introns (45,69). The extensive background of structural data on these interactions provides insights into the kinetic and thermodynamic studies of the cation-mediated RNA folding interactions explored in this thesis; thus I review the structural features of these tertiary motifs.

1.2.2   GAAA Tetraloop− 11 Nucleotide Receptor Motif

While the continuous helix is the smallest possible tetraloop receptor (Figure 1.5), additional stability and specificity can be achieved through a larger receptor motif. A natural such receptor is a highly conserved 11 nt asymmetric internal loop (5′-UAUGG-3′:5′-CCUAAG-3′), which was first identified through

phylogenetic (co-variation) sequence analysis of group I and II introns (63). Through chemical modifica-

tions, Murphy and Cech first identified the tertiary contact between a GAAA tetraloop and an internal

loop on the opposing helix that stabilizes the fold of the *Tetrahymena* ribozyme's P4–P6 domain (Figure

1.6). The stabilizing interaction showed a preference for GAAA tetraloops (70). Due to its high specifici-

ty, affinity, and abundance, Costa and Michel proposed that the T(GAAA)−R(11 nt) interaction must pro-

vide a common strategy for RNA helical packing (63).

The 2.8 Å crystal structure of the *Tetrahymena* Group I intron's 160 nt P4–P6 domain, which

contains the canonical GAAA −11 nt receptor interaction (Figure 1.6), reveals the remarkable features

that give rise to the specificity and affinity of this interaction (62). Each of the three consecutive ade-

nines of the tetraloop tightly pack into the minor groove of the receptor (P6) helix, which classifies them

more broadly as A-minor motifs (59,60). The T(GAAA)−R(11 nt) interaction is also characterized by a

specific hydrogen bond and base stacking pattern (Figure 1.7). The first **A** (A151) of the tetraloop makes

two H-bonds with the A of the U-A reverse Hoogsteen base pair (U224 and A48). The second **A** (A152) of

the loop hydrogen bonds with the receptor's G250 (1 H bond) below it and U224 (2 H bonds) above it

via the ribose zipper motif, i.e., inter-digitated 2′-OH interactions (43). The third **A** (A153) of the loop

interacts with the C:G of the receptor, such that the **G•A** pair of the tetraloop forms a base quadruplet

with the C:G pair in the receptor (**G•A•**C-G) making 4 H-bonds, identical to the type I A-minor motif of

A4 in the the GAAA-minor groove interaction from the hammerhead structure in Figure 1.5 *B* (57)*.* Also,

just as in the hammerhead structure, many of the hydrogen bonds of this tetraloop−receptor interaction

utilize ribose 2′-hydoxyls. Furthermore, two consecutives adenines in the receptor (nts 225 and 226)

are aligned side by side forming a pseudo-base pair, called the adenosine platform motif, which stacks

on the G of the G·U wobble pair (Figure 1.7 *A* and *B*), achieving near coaxial alignment of the helices de-

spite the asymmetry of the internal loop (62,71). The adenosine platform opens the minor groove of the

tetraloop receptor, allowing A151 to stack upon it (62,71), as seen in Figure 1.7 *A* and *B*. An additional

hydrogen bond is also made between the 2′−hydroxyls of top C:G basepair (G251) of the receptor and

the C:G pair at the base of the tetraloop (C154). In total, 10 hydrogen bonds are formed between the
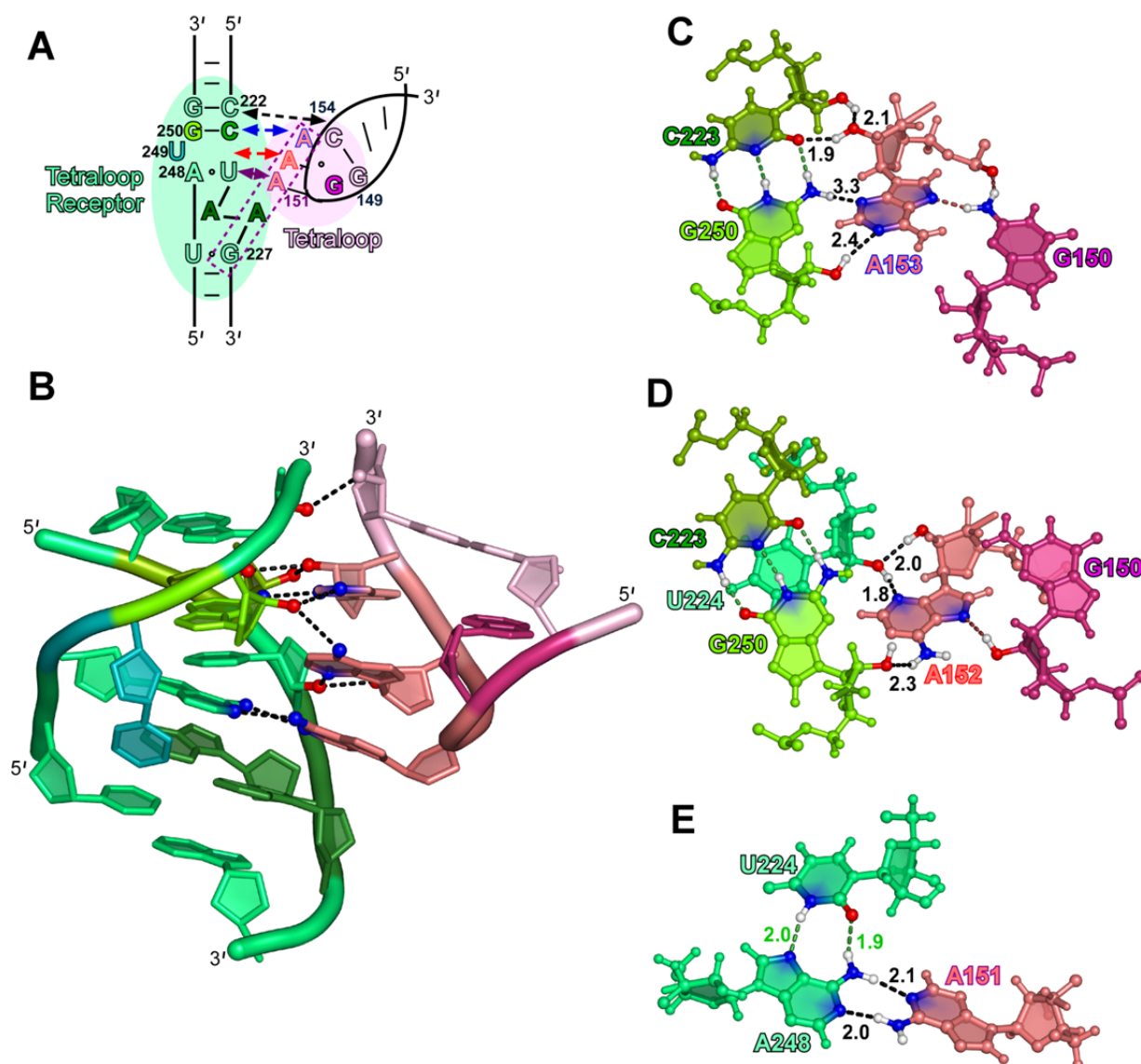
Figure 1.7   Structure of the GAAA tetraloop−11 nt tetraloop receptor motif from the P4−P6 domain. (*A*) Schematic of the interaction with hydrogen bond contacts indicated by arrows and base stacking of the tetraloop onto A226 of the adenosine platform in the receptor indicated with a *dashed* box. (*B*) 10 hydrogen bonds between the tetraloop and receptor shown as *black dotted lines*, *blue* = nitrogen, *red* = oxygen (hydrogens not shown) (*C*) A153, the top most adenosine, makes a base quadruplet with the receptor C223-G250 basepair. (*D*) A152 hydrogen bonds with the 2′-hydroxyls of U224 and G250, a ribose zipper motif. (*E*) A151 makes two H-bonds with the U-A reverse Hoogsten base pair (U224·A48). Hydrogen bond distances (Å) between the tetraloop and receptor are shown in *black* with a Watson-Crick pair (in *E*) labeled in *green* for comparison (PDB ID 1HR2).

tetraloop and the receptor, many of which are as short (i.e. strong) as Watson Crick base-pairing bonds.

These hydrogen bonds make the T(GAAA)−R(11 nt) motif a very specific and strong interaction, as mu-

tations to the tetraloop would disrupt the hydrogen bonding network and/or introduce steric clashes (60,62).

The same structure of the T(GAAA)−R(11 nt) interaction has been identified in many other natural and designed RNAs; therefore, the interaction is modular and can be isolated. For example, the crystal structures of RNase P, other Group I introns, and the NMR structure of a rationally designed dimer system contain the same T(GAAA)−R(11nt) motif (64,72,73). The structural robustness of the motif has allowed the interaction to be the subject of extensive biophysical characterization, and is an essential premise of this thesis.

Though the GAAA tetraloop is a rigid unit (indistinguishable whether free in solution or bound to a receptor) (50,51,62), the bound and unbound forms of its 11 nt tetraloop receptor are markedly different (62,74), as depicted in Figure 1.8. Specifically, the free form of the tetraloop receptor involves a high degree of base stacking; the central region is made up by 3 inter-digitated adenosines (*red*), a "base zipper" motif, while two uridines (*blue*) form a U·U mismatch pair stacked with the C:G base pairs (Figure 1.8 *A*). In the bound form, most of these stacking interactions are disrupted, with two of the A's aligned side by side, making up the adenosine platform, while one of the U's is unstacked and unpaired (Figure 1.8 *B*). These structural differences suggest that the tetraloop receptor must undergo conformational rearrangement upon tetraloop docking. However, this rearrangement is not achieved by cation concentration alone; even at very high (125 mM Mg^{2+}), the the tetraloop receptor is not detectable in a native (bound) form (75).

Metal ions are critical to proper RNA folding, and the formation of the GAAA tetraloop−11 nt receptor serves as model system for understanding RNA-metal interactions (76,77) and the role of metal ions in RNA folding, as is explored in this thesis. Metal cations associated with the bound tetraloop−receptor have been identified through NMR (Figure 1.9) and crystallography. In the crystal structure of *Tetrahymena* P4−P6 domain, a number of potential metal ion binding sites were identified near the GAAA tetraloop and tetraloop receptor (see Figure 1.6 for numbering): (i) a magnesium coordinated to G250 (analogous to G8 in Figure 1.9) phosphate oxygen of the receptor (62,64,78,79), (ii) a monovalent ion
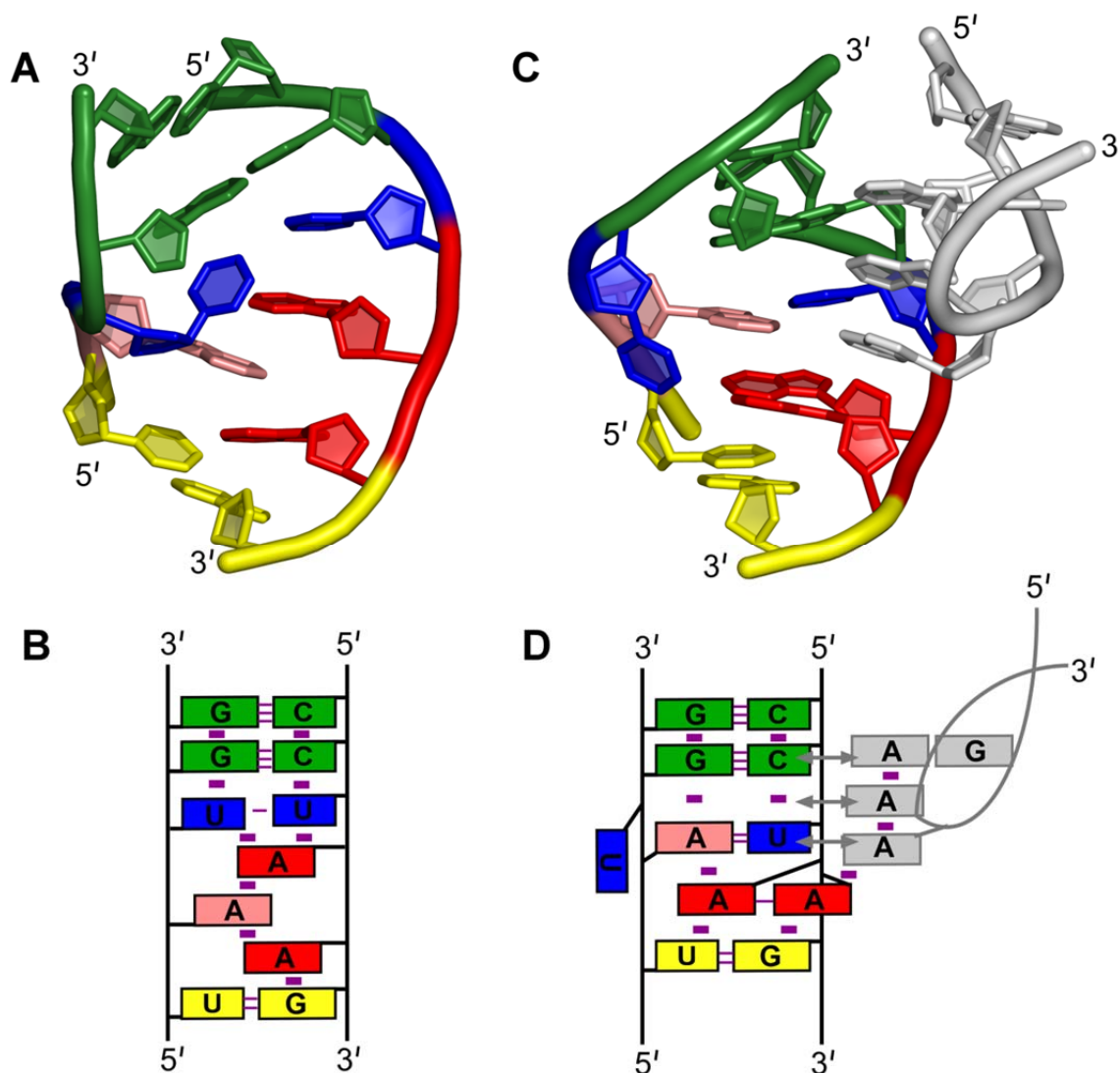
Figure 1.8 Free solution NMR structure of the 11 nt tetraloop receptor vs GAAA bound structure. (*A−B*) Lowest energy structure of the free structure 11 nt receptor determined by solution NMR spectroscopy and secondary structure schematic. The central region is made up by 3 interdigitated adenosines (*red*). Two uridines (*blue*) form a U·U mismatch pair stacked with the C:G base pairs (PDB ID 1TLR). Hydrogen bonds within the receptor are indicated as purple lines and base stacking as purple rectangles. (*C−D*) Crystal structure and secondary schematic of the GAAA bound tetraloop receptor from the P4-P6 domain (PDB ID 1HR2). GAAA–receptor hydrogen bonds are in shown in detail in Figure 1.7. In the bound form, two of the adenosines align side by side, making up the adenosine platform. One U (*blue*) is unstacked and unpaired.

(K+) coordination site below the adenosine platform nucleotides (A37 and A38 in **Figure 1.9** *B*) (64,71,79,80), and (iii) a cobalt hexammine binding site at the consecutive G·U wobble pairs (analogous to nts 18·25 and 17·26 in **Figure 1.9** *B*) in the major groove of the tetraloop helix (78,81). The NMR

structure of a dual tetraloop–receptor complex in solution showed similar metal binding sites (82).

Five manganese or 2 cobalt hexamine ions were also shown to localize on the complex using solution NMR (Figure 1.9) (82). Though the method used was unable to discern the hydration states of the ions, all of the position constraints could be satisfied by fully hydrated ions. The Mn^{2+} sites are shown in Figure 1.9 *A.* Sites 2, 3, and 5 are all in similar location to the 3 metal sites noted for the crystal structure. Site 5 (the G·U wobble pair) can be occupied by either manganese or cobalt hexamine ions and overlap well with the crystal structure (78,81). Site 2 (the AA platform) shows an associated manganese, which is ∼ 6.9 Å away from the K^+ site observed in the crystal structure (64,71,79,80), see Figure 1.9 *B*, potentially to accommodate a different coordination geometry. The authors determined that divalent ions can effectively compete with K^+ for this site, as spectra were compared with varying concentra-
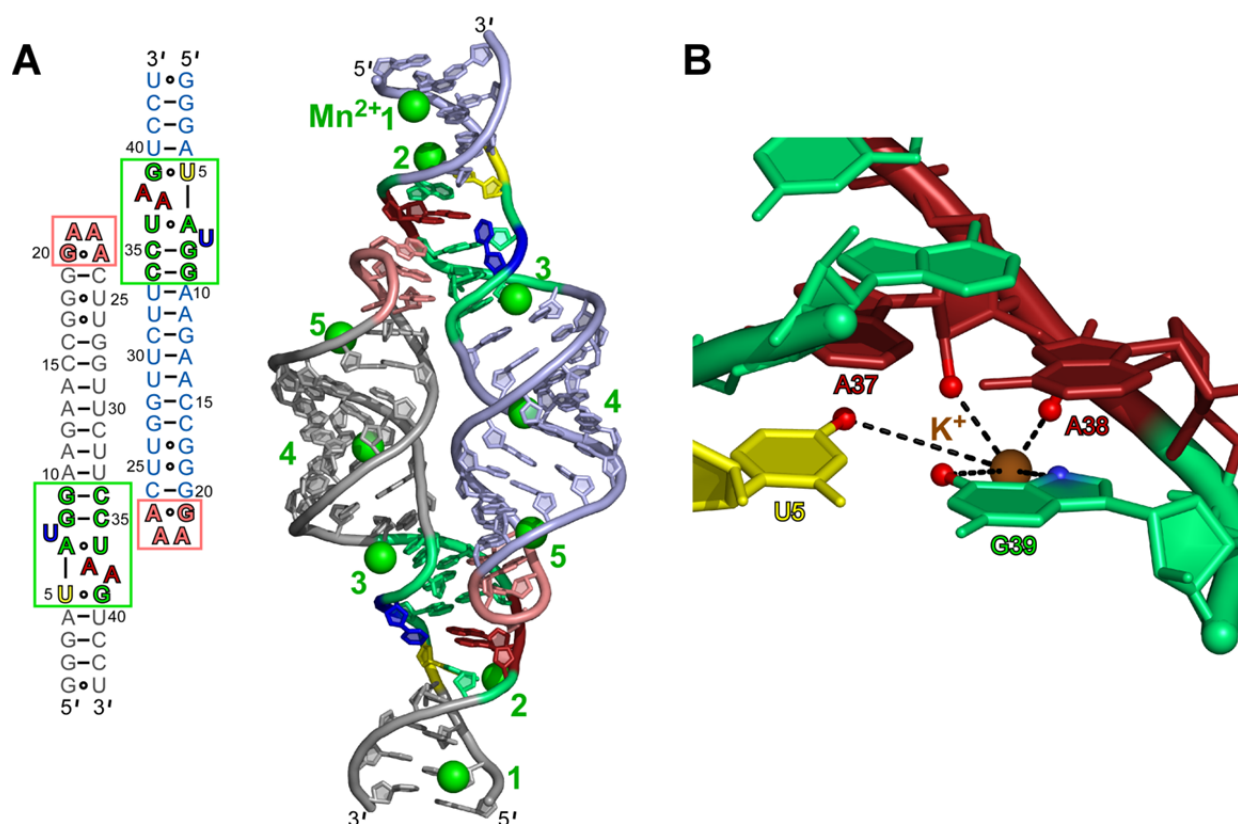
Figure 1.9   Metal ions and GAAA tetraloop–receptor structure. (*A*) Mn^{2+} (*green*) localized on the homodimer tetraloop receptor complex as determined by NMR. All position could be satisfied by hydrated ions (PDB ID 2I7Z). (*B*) The tetraloop receptor in the Azoarcus Group I intron crystal structure contains a K^+ chelation site below AA platform. The five proposed chelation sites are shown (PDB ID 1U6B).

tions of Mn^{2+} and K^+. However, K^+ has been shown to enhance the activity of the *Azoarcus* ribozyme and K^+ and $Na+$ have a higher affinity for this site over other monovalent ions (80,83). Both manganese and cobalt hexamine localize at the sequential C:G base pairs (site 3), indicating that the direct coordination of Mg^{2+} observed in the crystal structure is not required. Furthermore, Nonlinear Poisson–Boltzmann calculations of the complex's electrostatic surface have revealed that the metal ion localization sites overlap with high negative electrostatic potential, e.g. major grooves, as would be anticipated for diffusely bound ions (i.e., hydrated and localized, but not directly coordinated) (82,84,85). These calculations, combined with the observation that the tetraloop–receptor complex does not change structure over a range of ionic conditions, suggest that metal ions do not play a critical structural role in the tetraloop–receptor interaction, but rather are simply accumulated in the regions of negative electrostic potential within structure (82).

The solvent accessibility of the GAAA tetraloop–receptor interface was also assessed. It was determined that an ~730 $Å^2$ surface is buried in the interaction site (78), which suggests that water/cosolutes must be released upon tetraloop docking into the receptor. It was also observed that the increased hydrostatic pressure destabilizes the T(GAAA)–R(11 nt), to which a number of solvent interactions may contribute (86). The [Mg^{2+}]-dependence of the the bimolecular association of a GAAA tetraloop and receptor were also studied, yielding a dissociation constant (K_d) of 0.4 ± 0.05 mM at 125 mM $MgCl_2$ (76). This abundance of biophysical data provides an ideal platform on which explore kinetic and thermodynamic role of the tetraloop–receptor interaction in cation-mediated RNA folding.

1.2.3   A-rich Bulge–Helix Interaction

Nearly half of known group I introns possess a P5 extension, in which there is a highly conserved A-rich bulge (Figure 1.6) (45,69). The bulge motif is implicated in ribozyme activation (87-90) and is critical for proper folding of the P4–P6 domain, as demonstrated by deletion for mutation from A186 to U (70,91,92). The bulge itself has an interesting structure, as was seen in the crystal structure of P4–P6 domain (Figure 1.6) (62), making a corkscrew turn that flares the bases A183 and A184 out from the

helix (Figure 1.10). Two Mg^{2+} ions coordinate to the the the backbone of the bulge, which presumably eases repulsions between the closely positioned phosphates induced by it corkscrew turn. A186 also interacts with the G:C basepair below it (*yellow* in Figure 1.10) and the P5abc three-helix junction (Figure 1.6) (62). A183 and A184 make the tertiary contact with P4 (Figure 1.10), each making two hydrogen bonds sharing the 2′-hydroxyls, comprising the teeth of a ribose zipper, as shown with *dashed black lines* in Figure 1.10 *B* (62,93,94). Thus, only 4 hydrogen bonds zipper this tertiary interaction compared to 10 in the GAAA tetraloop−receptor motif. The *Tetrahymena* ribozyme is active without the te-traloop−receptor interaction (Figure 1.6), suggesting that (95) the tertiary interaction can form inde-
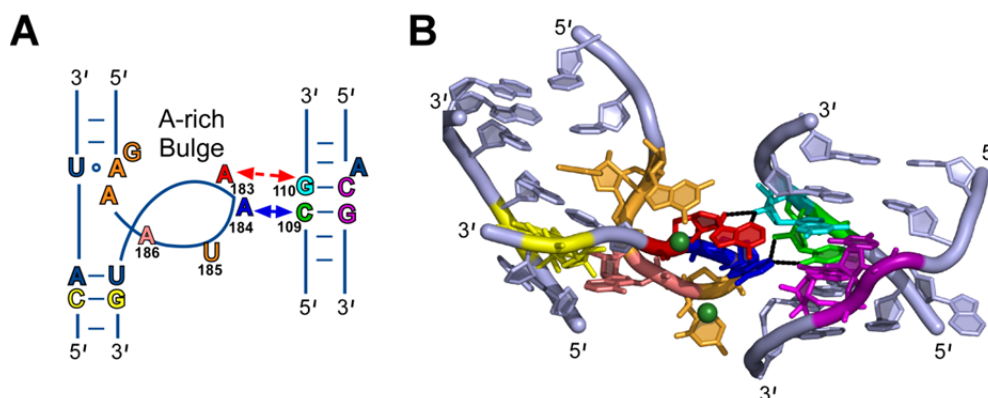
Figure 1.10   Structure of the adenine (A-rich) bulge interaction from the P4−P6 domain of the *Tetrahymena thermophila* ribozyme. (*A*) Secondary structure of the A-rich bulge showing the corkscrew turn of the A-rich bulge allow A183 and A184 to make a tertiary contacts with G110 and C109 on the opposing helix. (*B*) Crystal structure of the A-rich bulge interaction shows the 4 hydrogen bonds (*dashed black lines*) of the tertiary interaction along with the two directly coordinated Mg^{2+} ions (*green spheres*). (PDB 1GID).

pendently of the full domain, as supported by recent mutational studies of the P4−P6 domain (92). There is evidence that the A-rich bulge and tetraloop−receptor interaction work cooperatively to stabilize the P4−P6 domain (92), though a physical basis for this has not yet been identified. Kinetically and thermodynamically characterizing the tetraloop−receptor and A-rich bulge interactions individually may give insight into this question.

1.3   RNA Folding Landscapes

Thus far we have focused on the static yet sequential formation of tertiary structure from secondary elements. However, static secondary and tertiary structure data alone are not enough to explain functionality because structures can be highly dynamic and responsive to environment (e.g., temperature) (2,3). RNA, like proteins , encounter what it is known as Levinthal's paradox—random sampling of all possible conformations would take a tremendous amount of time (e.g., $\sim 10^{27}$ years for a 101 amino acid protein) (96,97). Yet RNAs can fold on the time scale of seconds and minutes, with secondary structure forming in microseconds. Therefore, there must be a funneling folding pathway to expedite this process by directing the RNA to the native state without random sampling (96,97). As a result, in order to understand how RNA folds, one needs to understand the energetics of the folding pathways, i.e., the full free energy, enthalpy and entropy landscape for folding. A major road block in achieving a predictive understanding of RNA folding landscapes, in particular the energy barriers for folding, is that they are often "rugged", i.e., with alternative conformations acting as kinetic traps resulting in slow and/or heterogeneous folding and/or unfolding rates (98-100). Moreover, the entropic and enthalpic challenge of folding a charged biopolymer (7-12) highlights the particularly critical role of Mg^{2+} and other counterions in thermodynamically influencing the overall folding free energy landscape. Such an understanding will also require characterization of the folding transition state, as well as the role of cations in stabilizing these transition states (7,8,10,101-103). A major aim of this thesis is to utilize simplified RNAs for studying the nature of folding landscapes.

1.4   A Single-Molecule Approach to Investigating RNA folding

As mentioned above, the concept of RNA structures as static, as portrayed in the crystal structures, can be misleading. RNA structures can be dynamic and responsive to cellular environment (temperature, $[Mg^{2+}]$, [metabolite], etc.), which can determine their functionality. Single-molecule methods have been instrumental in correlating the structural dynamics of RNA with function (3) and elucidating the path-

ways for RNA folding in real time (35,104). Single- molecule fluorescence resonance energy transfer (smFRET) methods have been widely applied to RNA folding (104,105). Such methods allow for isolation and direct characterization of RNA conformational dynamics, often with heterogeneous subpopulations that exhibit different rate constants and conformations and even rarely populated states readily distinguished under a wide range of experimental conditions, e.g., varying [Mg^{2+}] (3,106,107). Single-molecule FRET methods provide both folding and unfolding rate constants under equilibrium conditions, and can therefore offer insights into transition states and thermodynamics to reveal how the RNA folding landscape is perturbed by the environment (8,103,108)

1.5   Single-Molecule Fluorescence Resonance Energy Transfer (smFRET)

Fluorescence resonance energy transfer (FRET) is a very useful method for exploring conformational changes in biological molecules as it is a sensitive probe of distances between strategically positioned fluorescent labels, Figure 1.11 *A* (109). In particular, FRET is an extremely powerful probe of the real-time conformational changes in RNA , revealing folding and unfolding transitions on a 10−100 Å scale (105,110). The dipolar coupling of the transition moments of a donor and acceptor fluorophores result in a highly sensitive distance dependent efficiency for energy transfer (E_{FRET}) (111):

$$E_{\text{FRET}} = \frac{R_0^6}{R_0^6 + R^6},$$

1.1

where R is the distance between the fluorophores and R_0 is the Förster radius, or distance at 50% energy transfer. R_0 is dictated by the spectroscopic properties of the dyes as

$$R_0^6 = \frac{9000\,(\ln 10)\,Q_D \kappa^2 J}{128\pi^5 n^4 N_A},$$

1.2

where Q_D is the quantum yield of the donor, κ^2 is the orientational factor between the donor and acceptor, n is the refractive index of the medium, and N_A is Avogadro's number. *J* is the overlap integral for the donor emission and acceptor absorption.
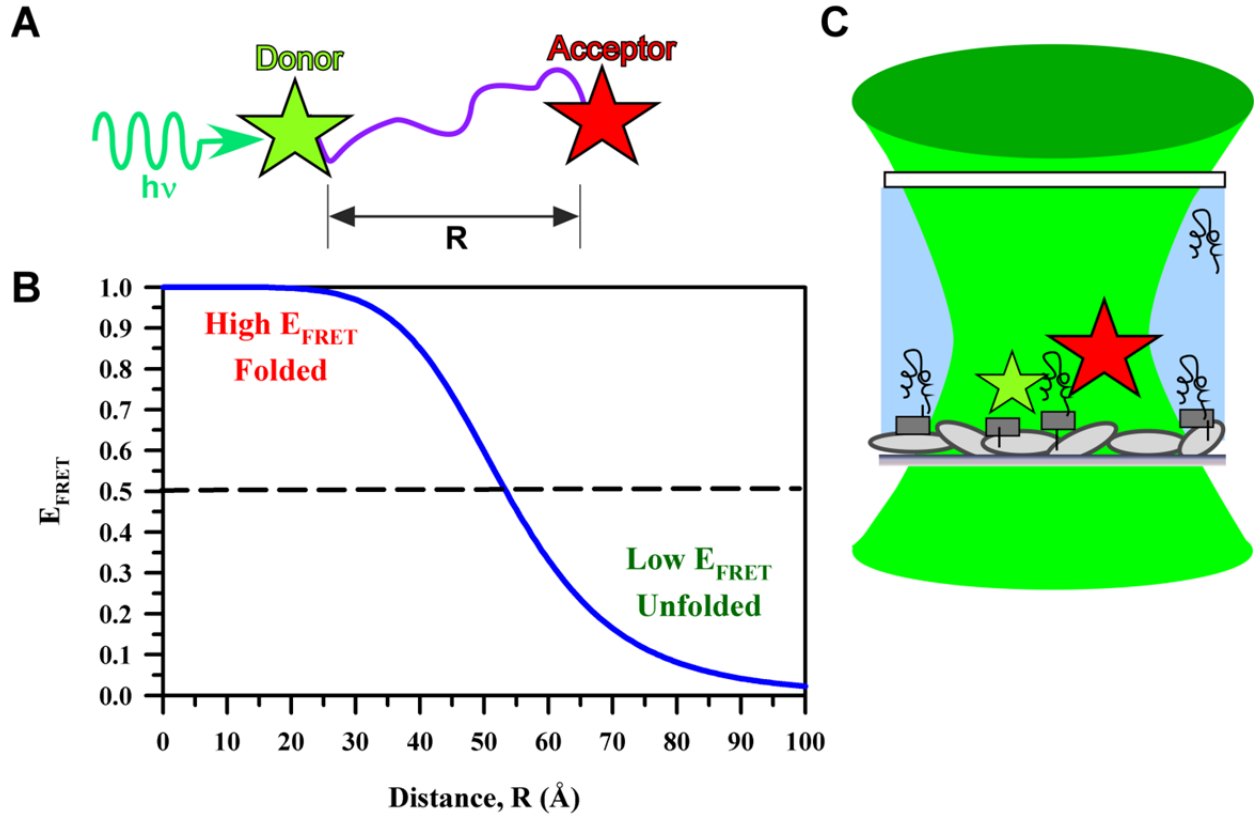
Figure 1.11   Single-molecule fluorescence resonance energy transfer (FRET) as distance ruler for conformational changes. (*A*) Schematic for FRET between a laser-excited donor fluorophore a distance (*R*) from an acceptor. (*B*) Distance dependence of the efficiency of energy transfer (E_{FRET}) for a Cy3-Cy5 dye pair with R_0 = 53.4 Å and the potential to resolve folded and unfolded molecule by E_{FRET} (Eq. 1.1). (*C*) Schematic of a FRET labeled molecule immobilized in a laser focus for single-molecule observation of donor and acceptor emission intensities.

$$J(\lambda) = \frac{\int_0^\infty F_D(\lambda)\,\varepsilon_A(\lambda)\,\lambda^4 \mathrm{d}\lambda}{\int_0^\infty F_D(\lambda)\,\mathrm{d}\lambda}, \qquad\qquad 1.3$$

with the fluorescence spectrum of the donor, $F_D(\lambda)$, and the absorption of the acceptor in terms of the

extinction coefficient, $\varepsilon_A(\lambda)$. κ^2 is a quantity that can vary between 0 and 4 depending on the dipole

alignment of the flourophores and is generally estimated (albeit often without basis) to be 2/3 for free

rotation of the fluorophores on a time scale much faster than the excited state lifetime (~ns)(112,113).

R_0 can be tuned by choice of fluorophore spectral properties to optimize sensitivity to distances of inter-

est. Cy3 and Cy5 are an ideal dye pair for the RNA folding explored in terms of their photostability, ab-

sorption cross sections, similarly good quantum yields, and R_0 (114,115). For the Cy3 and Cy5 labeled RNA used in this work, we determine $J = 8.25$ x 10^{-13} M^{-1} cm^3. With an estimated quantum yield of Cy3 labeled nucleic acids of ~0.15 and $n = 1.33$ in water, this J yields $R_0 = 53.4$ Å from Eq. 1.3. Cy3 quantum yield is sensitive to environment, e.g., temperature and attachment to biological molecules. The range for Q_D at room temperature is ~0.15–0.25 for Cy3 in single stranded vs. double stranded nucleic acids, thus the low end for R_0 is 53.4 Å and the high end is 58 Å, so R_0 for the Cy3–Cy5 pair on nucleic acids is within 50–60 Å (113,116,117).

For the Cy3–Cy5 pair we can view the distance dependence from Eq. 1.1, as shown in Figure 1.11 *B*. Cy3–Cy5 is an excellent probe pair for distance scales of RNA folding changes over a ~20–80 Å range, where one should note that the sensitivity to distance changes within a molecule is highest at distances around R_0. With a molecule strategically labeled to yield such distance changes upon folding and unfolding and single-molecule detection, one can observe the real time folding pathway of an RNA molecule at equilibrium. This requires the ability to observe the fluorescence of single molecules, thus demands a small detection volume, which can be achieved with laser-scanning confocal microscopy (118). The high numerical aperture microscope objective focuses the excitation laser (e.g., 532 nm) to a diffraction limited spot of ~270 nm, with the detection volume limited by a confocal pinhole. Working at low concentrations of the molecule of interest ensures that only a single molecule is observed in the focus at any given time (Figure 1.11 *C*). Longer time observation of molecules is afforded by immobilization (119).

From the fluorescence emission of the donor and acceptor labels, I_D and I_A, E_{FRET} is,

$$E_{FRET} = \frac{I_A}{I_A + \frac{Q_A}{Q_D} I_D},$$

1.4

where Q_A and Q_D are the acceptor and donor quantum yields, respectively. Thus, E_{FRET} can be used to probe the conformational state of the dual labeled RNA molecule by detection of the spectrally resolved donor and acceptor fluorescence emission. Determination of E_{FRET} from experimentally observed donor and acceptor emission rates will be discussed in Section 2.3

1.6   Single-Molecule FRET Strategy for Investigating Cation-Mediated Tertiary Interactions

RNA secondary structure is reasonably predictable from the known thermodynamic parameters (5,120), with software such as *mfold* freely available to do just that (121). Tertiary structure, however, is much less well understood and relies heavily on challenging methods such as crystallography and NMR (6). The modularity of RNA tertiary structure, i.e., the reconstitution of large RNAs from discrete components (122) has provided hope that one can predict RNA tertiary structures from knowledge of the individual tertiary interactions. For example, much like predicting the stability of an RNA helix from the known thermodynamics of Watson-Crick base pairs (120) , one could hope to predict the tertiary folded structure from the thermodynamics of tertiary interactions, which can also be categorized into relatively few motifs (37). Therefore, just as knowledge of the thermodynamic parameters was crucial for accurate secondary structure prediction, the same knowledge will be required for tertiary structure prediction. Proper RNA folding requires counterions to minimize repulsions of the negatively charged phosphate backbone. Therefore, structure prediction also relies on knowledge of how salt (e.g., NaCl) concentration affects tertiary structure stability. In addition, an understanding of the kinetics of tertiary folding is essential for determining RNA functionality (6,123). Toward this end, individual folding motifs and their dependence on cation environment must be characterized, both in isolation and in combination, for a unifying thermodynamic and kinetic description of RNA folding to emerge.

The well- studied P4-P6 domain (Figure 1.6) is an ideal place to begin characterizing tertiary interactions, as it folds independently and contains two distinct tertiary interactions (detailed above)—the ubiquitous GAAA tetraloop−11 nt receptor and the A-rich bulge−helix, both of which utilize exposed adenines to make tertiary contacts with an RNA minor groove, Figure 1.12 (62). In Figure 1.12, a strategy for breaking the P4–P6 domain (Figure 1.6) down into basic elements of tertiary structure is shown—a helix-junction-helix motif assembles with the help of a tertiary interaction. With strategic labeling with Cy3 and Cy5, this folding transition can be monitored by FRET. By studying the tertiary interactions indidually, one can hope to gain insight into how multiple tertiary interactions work together to guide RNAs to their functional structure. Thus, we aim to kinetically and thermodynamically char-
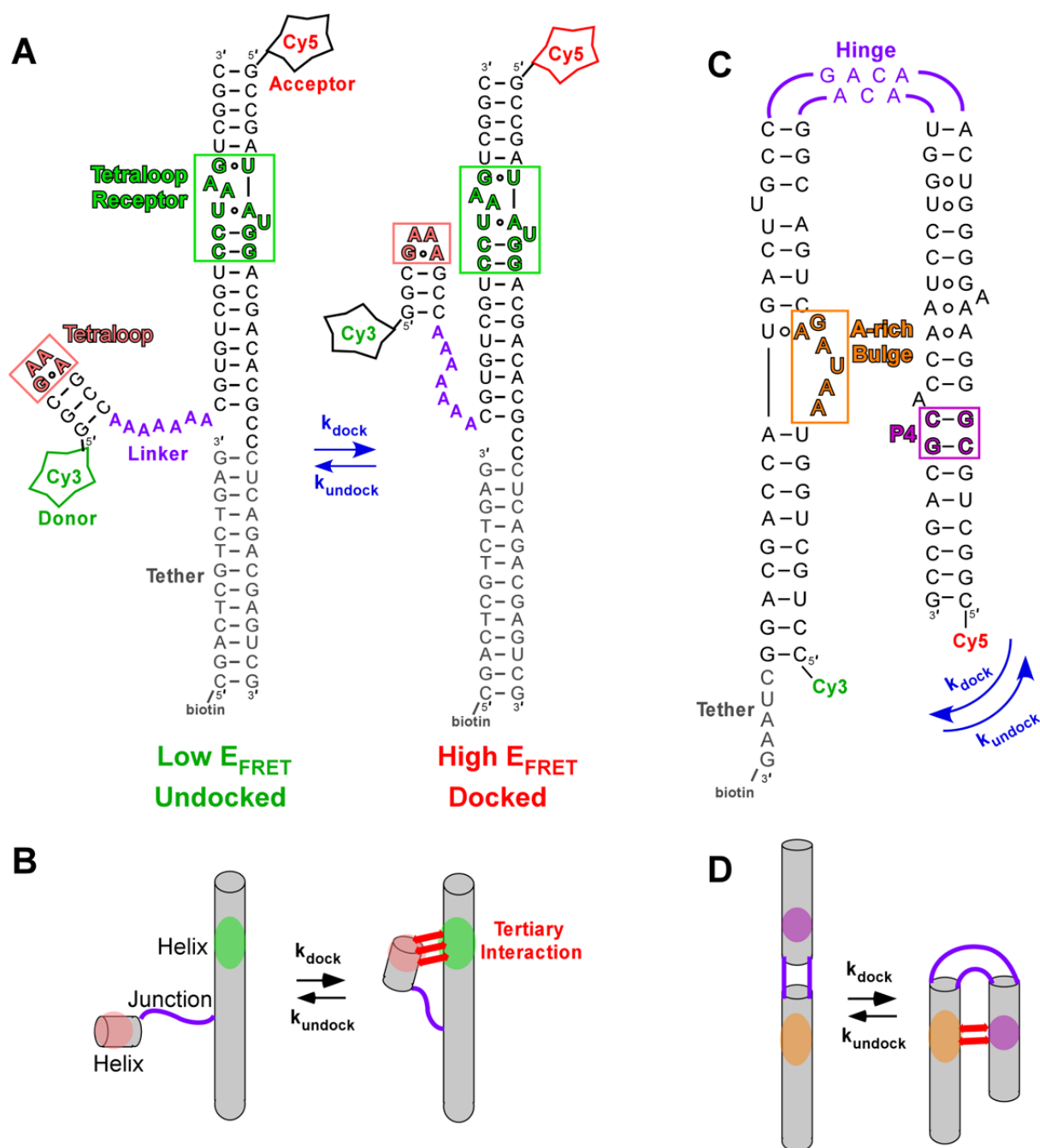
Figure 1.12 Isolation of tertiary interactions in FRET labeled constructs. (*A*) The GAAA tetraloop and 11 nt tetraloop receptor motif are connected by a single-stranded linker (*yellow*). Cy3 and Cy5 fluorophore labels allow for monitoring of docking and undocking using FRET methods. A biotinylated region (tether) is used for immobilization on strepatividin coated glass surfaces. (*C*) The A-rich bulge can be isolated in a reduction of the P4–P6 domain (Figure 1.6). (*B* and *D*) The tetraloop–receptor and A-rich bulge RNAs are model RNA systems for exploring the simplest folding motif—helix–junction–helix. The unpaired adenines in both constructs act as beacons between helical regions.

acterize isolated tertiary motifs in the context of simplified RNA folding systems, in particular the GAAA

tetraloop–receptor interaction (Figure 1.12 *A*). Using single-molecule techniques, specifically confocal FRET microscopy and time-correlated single-photon counting, we can observe both docking and un-docking transitions in the RNA at equilibrium and therefore gain insight into the folding mechanisms and the role of cations in the folding process. We can also explore the role of secondary elements, name-ly junctions, on the kinetics and thermodynamics of folding. We aim to elucidate the underlying physical principles that govern RNA folding and provide necessary information for tertiary structure prediction.

1.7   Overview of the Thesis

The major focus of this work is studying the single-molecule kinetics and thermodynamics of RNA fold-ing due to the ubiquitous GAAA tetraloop–11 nt receptor interaction (Figure 1.12*A*). As described above the interaction is modular (42,63), making it the subject of extensive   biophysical characterization (50,51,62,66,70,73,74,77,92,106,124-127), which allows for interpretation of our results in the context of other studies. The GAAA tetraloop–receptor interaction is isolated in a construct that allows for mon-itoring of intramolecular docking of the tetraloop into the receptor using single-molecule FRET methods (Figure 1.12*A* and *B*) (106). Connected by a flexible single-stranded junction (poly A), the GAAA te-traloop readily and specifically docks into its receptor, modulating the fluorescence resonance energy transfer (FRET) between the donor (Cy3) and acceptor (Cy5) fluorophores (106). The efficiency of en-ergy transfer (E_{FRET}) is monitored by single-molecule confocal microscopy—calculated ratiometrically from the donor and acceptor emission intensities (Eq. 1.4) (106,128). Such real-time E_{FRET} traces permit observation of the [Mg^{2+}]-dependent docking (k_{dock}) and undocking (k_{undock}) kinetics, observable by fluc-tuations between two well-resolved high (docked) and low (undocked) E_{FRET} states with FRET efficien-cies of ~0.3 and 0.7, respectively (Figure 1.13). The apparent first-order rate constants for docking and undocking (k_{dock} and k_{undock}) are determined from the dwell times (τ) of molecule in the undocked and docked states (Figure 1.13). This method allows for isolation of the tetraloop–receptor interaction, as well as serving as a simplified model system for studying the role of cations in RNA folding as it is a basic example of a helix-junction-helix motif (Figure 1.4 and Figure 1.12 *A* and *B*) (124).
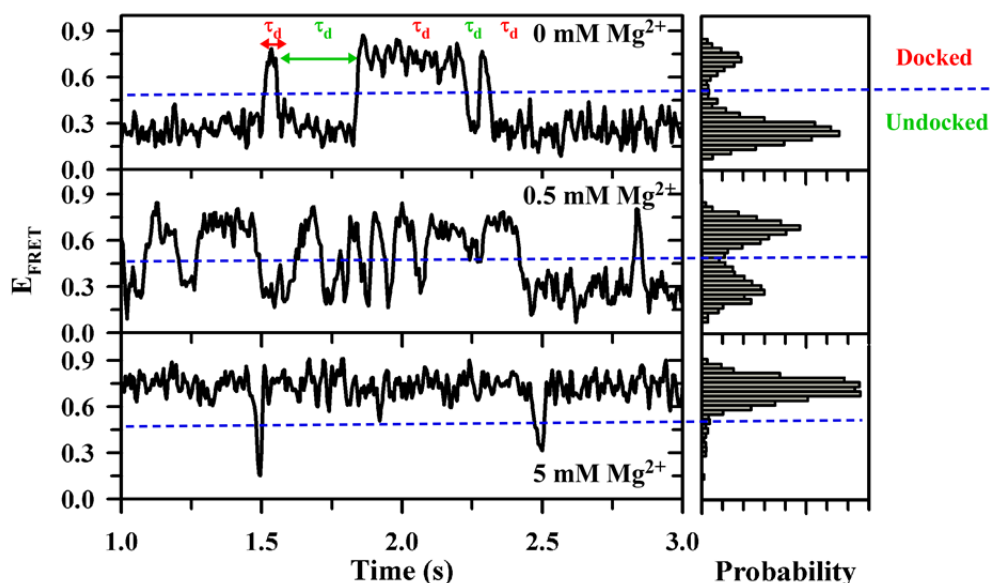
Figure 1.13   Single molecule Mg^{2+}-dependent kinetics of intramolecular tetraloop-receptor docking/undocking. Sample real-time single-molecule FRET efficiency (E_{FRET}) traces resolving docking and undocking transition of the tetraloop and receptor (Figure 1.12 *A* and *B*) at varying [Mg^{2+}]. Two E_{FRET} states, docked and undocked, are identified, as seen by the corresponding probability distributions.

In Chapter 2 the temperature-controlled single-molecule FRET microscope is described. Data acquistion, sample preparation, and analysis are also explained therein. In Chapter 3, a freely diffusing single-molecule FRET assay is developed to explore the effect of Na^+ and Mg^{2+} on the population distributions of the tetraloop−receptor construct (128). Chapter 4 describes measurements of the tetraloop−receptor interaction as a function of temperature, allowing for extraction of the enthalpy and entropy of docking (117). The overall docking reaction is exothermic and entropically costly, consistent with the large number of hydrogen bonding and base stacking interactions that occur within the tertiary contact. We interpret the GAAA-tetraloop−receptor tertiary interaction in the context of previous RNA folding thermodynamic studies to illuminate a possible enthalpy vs entropy-driven folding paradigm. In Chapter 5, the temperature-dependent docking and undocking kinetics of the tetraloop and receptor are explored as function of [Mg^{2+}], comparing the usual A_7 linked construct (Figure 1.12 *A*) with a U_7 alternative (unless otherwise noted, the A_7 construct is utilized). Quite surprisingly, our work reveals an entropic origin of Mg^{2+}-facilitated RNA folding, which contrasts with the common expectation that *increas-*

ing [Mg^{2+}] facilitates folding due to *reduced electrostatic repulsion* of opposing RNA helices. Instead, we propose that higher [Mg^{2+}] facilitates RNA folding by (i) decreasing the entropic penalty of counterion uptake in the tertiary folding transition state and (ii) by reducing disorder of the unfolded conformational ensemble. We also show the tetraloop–receptor folding transition state is "early" or unbound-like and is dominated by an entropic barrier, which we suggest may be a general feature of RNA folding. In Chapter 6, we investigate the role of cation valence and size in the kinetics and equilibrium of the tetraloop–receptor interaction. Specifically, we show that Na^+, K^+, Mg^{2+}, Ca^{2+}, and $Co(NH_3)_6^{3+}$ can equivalently accelerate docking and decelerate undocking, which we attribute to a four-state kinetic model. This model allows for extraction of the cation-binding affinities, which are dictated by the counterion valence. Cation charge density also affects the stoichiometry of cation uptake with folding. Spermidine^{3+} also facilitates docking, but to a lesser extent than the other cations. Chapter 7 summarizes the work of this thesis, revealing an unprecedented level of detail in characterization of a tertiary folding transition and insights into cation-mediated RNA folding landscapes. Future experiments are proposed, as these studies can be readily applied to other tertiary interactions, namely the A-rich bulge (Figure 1.12 *B*).

Chapter 2   Experiment

2.1   Single-Molecule Confocal FRET Microscopy[1]

2.1.1   Spectral Properties of the Cy3−Cy5 FRET Pair

The single-molecule FRET microscope system is optimized for use of Cy3 as a FRET donor, which can readily be excited by 532 nm laser sources (peak absorbance 550 nm) with its fluorescence well spectrally separated from the excitation source (peak emission at 570 nm), as shown in Figure 2.1. Cy5 is chosen as the FRET acceptor to resolve RNA conformational changes on a ∼30−70 Å scale. As mentioned in Chapter 1, the Förster radius for this pair is $R_0 \sim 53\text{--}55$ Å, which can be determined from the spectral overlap of the donor emission and acceptor absorbance (Eq. 1.4). Cy3 and Cy5 are also a good
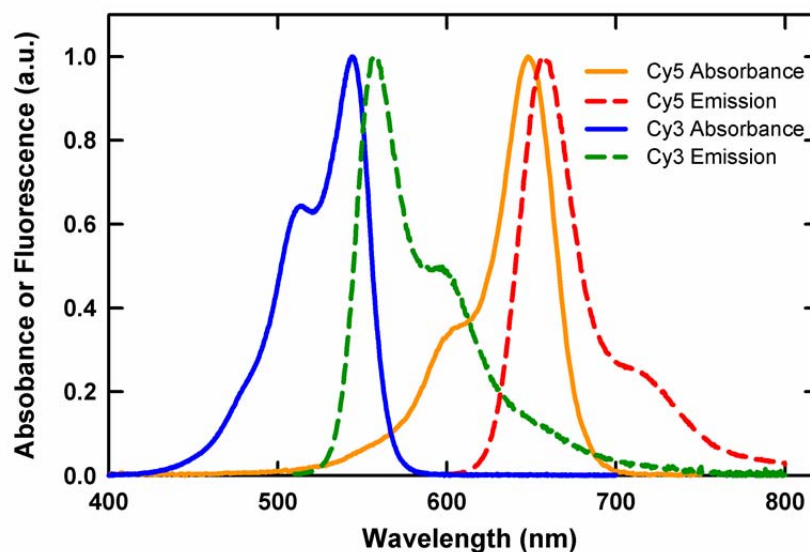
Figure 2.1   Absorption and emission spectra of Cy3 and Cy5. The extinction coefficients of Cy3 and Cy5 at peak absorbance are 150,000 and 250,000 M^{-1} cm^{-1}, respectively.

FRET pair for microscopy because the fluorescence can be well separated with a dichroic beam splitter at 645 nm, for example (Figure 2.1). The cyanine (Cy) dyes are ideal for single-molecule experiments because they have large absorptions cross sections at their peak absorbance (>10^{-16} cm^2), are bright (quantum yields > 0.15), and are photo-stable (>10^6 absorption/emission events before photobleaching) (114). The quantum yield for Cy3 increases considerably when bound to nucleic acids, in which case it is 0.15−0.25 at room temperature (116). No significant changes of Cy5 have been noted upon coupling to nucleic acids, and its quantum yield is ~0.28 according the dye manufacturer (Amersham Biosciences). The quantum yield decreases as a function of temperature for Cy3 and Cy5 dyes, whether free or coupled to nucleic acids (116,117). The similar quantum yields of donor and acceptor is a feature that greatly aids FRET experiments, such that nearly constant total fluorescence signal levels are maintained, irrespective of the conformational state of the molecule. This also simplifies determination of the energy transfer efficiency (E_{FRET}), as will be seen in section 2.3.1. The Cy dyes can also be made more photo-stable while minimizing undesirable photophysics (blinking and dark states) by addition of an oxygen scavenging cocktail to the sample buffer, which generates long-lived, high signal single-molecule trajectories (129,130).

2.1.2 Confocal Microscope for Single-Molecule FRET[2]

The single-molecule FRET apparatus is described with limited details in the following Chapters and our publications (106,117,124). Here I elaborate on the instrument design and characterization. The optical system for single-molecule FRET detection was built on an inverted microscope base (Olympus IX-70) using epi-illumination with the addition of a confocal pinhole, and two-color time-resolved detection. A mode-locked 82 MHz (~150 ps pulses) frequency doubled Nd:YAG laser at 532 nm (Model 3800, Spectra Physics, Mountain View, CA,) is used as an excitation source for fluorescent molecules (Figure 2.2). The laser beam is spatially filtered through a single-mode fiber and linearly polarized, as ensured by a polarizing beam splitter cube. The beam can be attenuated with neutral density filters to achieve

---

[2] The original microscope system (sections 2.1.2−2.1.5) was designed and built by Dr. Jose H. Hodak. Parts of these sections are adapted from his unpublished documentation of the instrument.
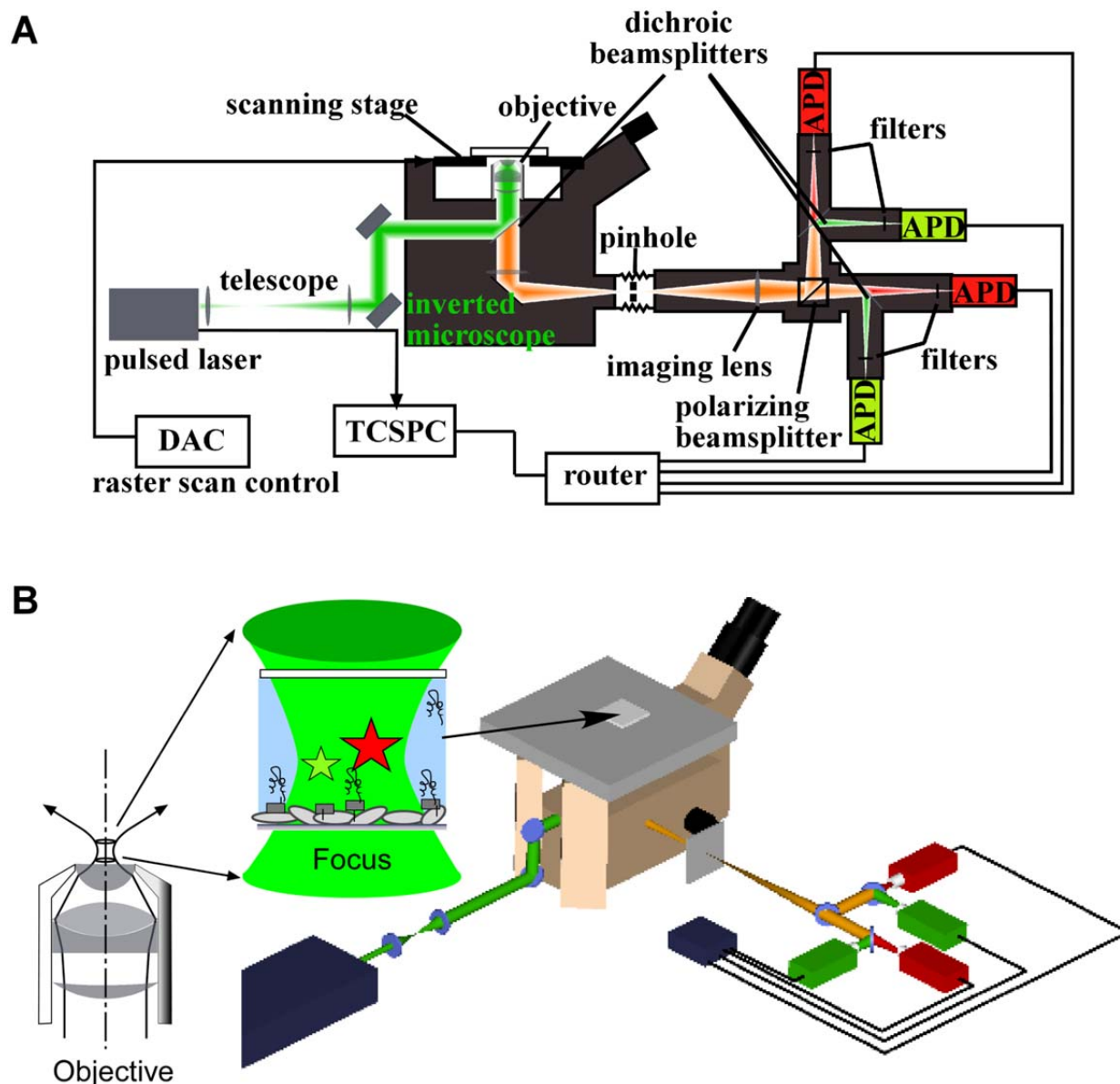
Figure 2.2   Schematic of scanning confocal microscope setup. (*A*) A pulsed 532 nm laser excitation source is focused into an inverted microscope. Fluorescence emission is collected by the same objective, filtered through a confocal pinhole, and detected by avalanche photodiodes. (*B*) Schematic of the microscope (courtesy of Larry Fiegland) and a depiction of the laser focused by an objective through cover glass into a solution containing fluorescently labeled molecules (not to scale). In this work donor and acceptor signals (*green* and *red*) are summed over horizontal/vertical polarization channels.

the desired power ranges of 0.05–100 μW for experiments. The beam is expanded to a $1/e^2$ intensity

diameter of 12 mm, to overfill the back aperture of the microscope objective (9.25 mm) by ~30%. Trace

infrared light from the Nd:YAG 1064 nm fundamental is removed by a bandpass filter (EX 530/10,

Chroma Technology, Rockingham, VT) upon entrance of the excitation laser light into the back port of the microscope. The beam is directed into a water immersion objective (Olympus (Center Valley, PA), UPlanApo 60X, 1.2 numerical aperture) by a dichroic mirror (550 DRLP, Chroma Technology), and is focused to a diffraction-limited spot. The objective is equipped with a correction collar, set to 0.17 for focusing through the standard cover glass thickness (Corning No 1 ½). By way of the coarse and fine adjustment knobs of the microscope base, the beam focus can be positioned at the top surface of the cover glass or in solution to excite molecules of interest.

Fluorescence emission from a laser-excited molecule is collected by the objective and passed through the dichroic onto a mirror, which directs the light through a tube lens (f=180 mm) and out of the microscope's side port (Figure 2.2). A pinhole wheel (Olympus OSP-TUR) is located at the image plane of the microscope (78 mm away from the microscope wall) to spatially filter out-of-plane fluorescence, confocally limiting background and detection volume. To select the size for the confocal aperture, we assume that the full numerical aperture (NA) of the objective is illuminated such that the laser is focused to the diffraction limit. In this limit the radius to the first diffraction minimum (r_d) of the Airy pattern is 270 nm $[r_d = 1.22 \lambda / (2\ NA)]$, where NA = 1.2 and the wavelength (λ) = 532 nm. This corresponds to a 540 nm-diameter central-airy disk at the focus of the objective. Based on a 60X magnification system, the image diameter of this central portion of the Airy pattern should be ~32.4 μm at the image of the tube lens. Thus, a 32.5 μm diameter pinhole would transmit the central disk of the Airy pattern (or 84% of the total intensity). The confocal pinhole was chosen to be 50 μm as a compromise between out of focus light rejection and high signal throughput. The fluorescence focused through a 50 μm pinhole is imaged by a biconvex lens (f=100 mm) onto the active area of four avalanche photodiode detectors. Before reaching the detectors, the fluorescence is separated by polarization with a broad-band polarizer cube (CVI PBSH-450-1300-100), dividing the fluorescence into vertical and horizontal polarizations with respect to cube surface. Each polarization is further separated into donor (Cy3) and acceptor (Cy5) channels by a dichroic beam splitter (645DCXR, Chroma Technology). This dichroic is chosen to optimize reflection of Cy3 and transmission of Cy5 emission. Photon color is further discrimminated

by transmission through a bandpass filters positioned 40 mm before the detectors, HQ585/70M and HQ700/75M (Chroma Technology), for donor and acceptor channels, respectively. (The HQ700/75M used originally was eventually replaced by a 665LP to allow for observation of further red-shifted fluorophores with no effect on the detection of Cy5 used in this work). All optical surfaces except for the photodiode window have a broad band anti-reflection coating.

Photon detection is performed by four single photon counting avalanche photodiodes (APDs, SPCM-AQR-14, Perkin-Elmer Optoelectronics, Fremont, CA) placed at the focus of the imaging lens. APDs are useful because their spectral response is flat over the investigated range, they have a high detection efficiency ($\sim$65%), fast time response, can tolerate high count rates (up to 10 MHz), low dark counts, large active area (175 μm in diameter), and short time to reset after a photon arrival event (dead time = 50 ns). The APDs are mounted on xyz translations stages to facilitate alignment at the lens focus for optimal signal collection. The 4-channel detection provides flexibility for additional investigations of fluorophore polarization anisotropy decay and rotational diffusion times. For the present studies, however, we focus on unpolarized signals, summing over horizontal/vertical channels for donor and acceptor.

The output of the avalanche photodiodes is connected to a de-multiplexing unit (HRT-82 Becker & Hickl) and fed into a time-correlated single-photon counting (TCSPC) module (SPC-134 Becker & Hickl). A reverse biased silicon photodiode (MRD510 Motorola) is illuminated with residual 1064 nm light from the mode-locked Nd:YAG laser to generate a synchronization signal. The time to amplitude converter (TAC) is started by a pulse from any one of the four avalanche photodiodes, beginning a linear voltage ramp to track the time until the next laser pulse arrives. The voltage ramp is stopped by the synchronization signal from the next laser pulse. The start–stop time is then used to infer the the arrival time of the photon after is excitation pulse based on the known interval of the laser pulse train. This reversed start/stop clocking method is much more efficient than clocking every laser pulse, i.e., the time –reversed start-stop cycles occur with orders of magnitude less frequency than the laser repetition rate ($\sim$10 kHz typical photon count rates vs. the 82 MHz laser repetition rate). Not only is time-reversed sin-

gle-photon countain more efficient, it is also necessary because there is a finite "dead time" to reset the voltage ramp after each counting start-stop cycle event. In the SPC-134 module this dead time is 100 ns. Information would be lost if a photon arrived before the reset—less resets means less chance of missing a photon. Data acquisition is achieved with software developed in the Nesbitt lab using Lab Windows CVI.

The laser excitation was later upgraded to a more user-friendly source with shorter pulse to decrease the instrument response time and greater intensity stability than the mode-locked Nd:YAG— a Time Bandwidth Model Fulmineo, SESAM patented modelocked 532 nm laser, including the power supply providing 50 mW maximum output at 20.55 MHz in 10 ps pulses. This laser is equipped with a sync output that can be directly fed as a synchronization signal for TCSPC, precluding the need for an external photodiode signal. The microscope system is also equipped with an alternative excitation source, a 635 nm Picoquant, GmbH pulsed laser (LDH-P-635) coupled to a single mode fiber, with pulse width of ~70 ps and driver (PDL 800-B), allowing for variable repetition rates of 5 to 80MHz repetition rate. The red laser is useful for studies utilizing direct excitation of Cy5.

2.1.3 Alignment

The system is aligned by removing the confocal pinhole and finding the position of maximum count rate for all 4 detectors using tetramethylrhodamine (TMR) for the donor channels and Atto 655 for the acceptor channels in water at ~100 nM and 400 nM concentrations, respectively. The fluorescence lifetimes and count rates are monitored in the SPCM Becker and Hickl software. This software was also used to ensure that the signals from all of the detectors are overlapped in time by adjusting the cable length from the APD into the router. Caution is taken when setting the laser power to avoid damaging the APDs with high photon count-rates (<10 MHz count rates). After the detectors are positioned, the pinhole spun in from the wheel, and the laser power is set to 1μW at the microscope objective (placing power meter head directly onto the objective with immersion water). The pinhole position is optimized by an xyz scanning stage. The z (z is the axis in line with the beam path) position is obtained by plotting

the signal level as a function of z position, (optimizing xy at each step). The z position is then set to the maximum transmission position and not altered unless references signals are below normal. The detectors are iteratively optimized with the pinhole in their xy positioning (their z positioning is optimized as it is fairly insensitive to alignment). Daily alignment consists of the xy positioning of the pinhole and detectors with reference solutions of TMR and Atto 655 designed to yield ~100 kHz/μW with 100 nM, and 400 nM, respectively. All cover glasses used in this work were 22 by 22 mm corning brand No 1-1/2, which has thickness of ~170 μm.

2.1.4   Time-Correlated Single Photon Counting (TCSPC)

TCSPC allows for every photon detected to be stored in a "time stamped" detection mode, which records (i) the APD channel on which the photon landed, i.e., donor or acceptor and polarization, (ii) time delay (τ_{micro}) after the laser excitation pulse (<200 ps resolution), and (iii) wall clock time (τ_{macro}) after the start of data acquisition (50 ns resolution). This information can be used to determine the wall clock time, polarization anisotropy and fluorescence lifetime for any time window (bin) of photons desired. A full width half maximum (FWHM) total instrument response function of 570 ± 5 ps was measured by prompt Raman scattering signal from H_2O, which is shorter now with the incorporation of the new laser with shorter pulses (current IRF = 450 ps FWHM). The SCPM-134 module is a time-reversed counting method (arrival of a photon triggers counting, which is stopped by the arrival of the next laser pulse), used in the "FIFO" mode, in which every detected photon is stored in a record containing the channel information (as passed by the HRT-82 router). A crucial element of TCSPC is that the count rates must be low enough such that it is negligibly probable for multiple photon arrivals from the same laser pulse, otherwise the information on the second photon is lost. Fluorescence lifetime data is calculated by making a histogram of the photon arrival times over a series of laser pulses.

2.1.5   Raster Scanned Imaging of Single-Molecule Fluorescence

A microscope sample stage was machined in the JILA instrument shop to accommodate a 3D nano-piezo

stage scanning stage (PI 517-3CL Physik Instrumente) and the microscope objective. The scanning stage

is driven by an amplifier with closed-loop control (E503 and E509C3A Physik Instrumente) and con-
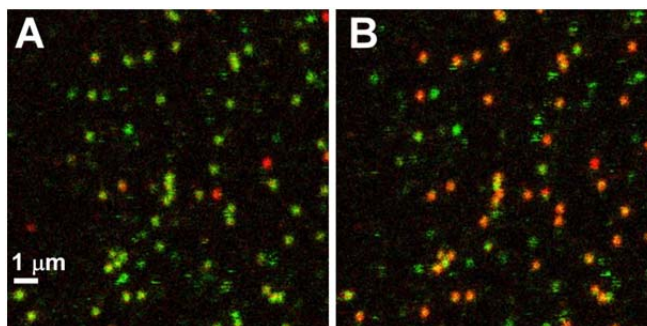
Figure 2.3   Raster-scanned images of the same RNA molecules with (*A*) no and (*B*) high [Mg^{2+}]. Each pixel depicts a false color representation of donor/acceptor emission with intensity proportional to number of donor (*green*) vs. acceptor (*red*) fluorescence photons. Each image is 12.5 × 12.5 μm; the intensity scale is 0–10 kcounts/s for an incident power of 1.1 μW, pixel integration time of 2 ms/pixel. At increased [Mg^{2+}] the molecules spends more time folded, so the FRET efficiency is high and more acceptor (red) photons are observed.

trolled by a computer card (PCI 6733 National Instruments Austin, TX) with digital to analog converters

(DAC), Figure 2.2. The stage raster scans the sample with respect to the objective at 2 ms per pixel to

collect 12.5 × 12.5 μm (256 × 256 pixel), line by line (trace and retrace provide separate images) to ob-

tain surface images (Figure 2.3). The stage scanning range is 100 μm and pixel integration times and im-

age size are variable parameters. The aforementioned conditions optimize the desired number of mole-

cules per image with good signal levels while minimizing photobleaching and the experimenter's time.

The size of the fluorescent molecule are diffraction limited as indicated by the full width half maximum

of molecule intensity profiles in images (∼270 nm, see section 2.1.6, Eq. 2.1), as seen in Figure 2.3.

The large time constant of the feedback loop together with the inertial mass of the scanning

stage causes hysteresis, evident in the trace-retrace images. The use of a closed loop scanning stage

ensures that a location will be reproduced after a trace-retrace cycle, or after the completion of an

image. Four images are collected simultaneously in this setup (trace and retrace, donor and acceptor),

each of them is set to a gray intensity scale linearly proportional to the number of counts per pixel. Pairs

of these images (gray planes) are added in an RGB color scale using a linear combination of red and green where the coefficients in the linear combination are given by the gray intensity of each component plane.

As a daily diagnostic for the scanning mode, samples are spin-coated using one droplet (~25 µL) of a ~300 pm solution of TMR (0.1 mg/ml of polyvinyl alcohol in water) onto a clean cover glass at 4000 rpm, 3 s. This measurement provides a signal reference of the line scans of 60 kHz/µW (or 60 x 10^3 counts/s/µW).

2.1.6 Instrument Characterization: Collection Efficiency

The microscope system is characterized by the collection efficiencies of Cy3 and Cy5 fluorescence emission on the channels designed for donor and acceptor detection, 1 and 2, respectively (with channels 1 and 2 summed over polarizations), i.e., $\beta_A{}^1$, $\beta_A{}^2$, $\beta_D{}^1$, and $\beta_D{}^2$ define the unitless collection efficiency of the acceptor on channel 1 or channel 2 and donor on channels 1 and 2, respectively. Crosstalk between donor and acceptor channels is characterized by $\beta_A{}^1$ and $\beta_D{}^2$. To determine these values we need to calculate the predicted photon emission rate from the fluorophores and compare it to the observed count rate. First we need to determine the laser intensity at the focus and thus must know the illumination area. The illuminated area at the focus is obtained by fitting a 2D Gaussian function to the point spread function of an imaged molecule (Tom Baker's image analysis program). For a diffraction-limited focal spot, the full width half maximum (r_{FWHM}) is approximately,

$$r_{FWHM} = \frac{1.22\,\lambda}{2\,NA}\,,$$

2.1

where NA is the numerical aperture (1.2 for our objective) and λ is the wavelength of the laser (532 nm). Thereby, r_{FWHM} in our system should be 270 nm, which is confirmed from fitting the intensity profiles of single molecule in images. The collection efficiency of the system is assessed by comparing the measured rates for photon emission with the predicted emission rate (Φ_{emit}), which is

$$\phi_{emit} = \sigma_{dye}\,\frac{I_0}{E_{photon}}\,Q_{dye}\,,$$

2.2

where σ_{dye} is the absorption cross section of the dye molecule at the excitation wavelength, Φ_{laser} is the photon flux of the laser source (photons/s), I_0 is peak intensity of the diffraction limited focus (J/s/cm^2) , E_{photon} is the photon energy or hc/λ (for 532 nm E_{photon} = 3.74 x 10^{-19} J), and Q_{dye} is the dye quantum yield. The cross section is simply determined from the absorbance of the dye molecule, which is directly proportional to the extinction coefficient (ε) according to Beer's law. Thus with ε known for Cy3 and Cy5 at the absorbance maximum (ε_{max} =250,000 and 150,000 M^{-1} cm^{-1}, respectively), the absorbance spectra can be converted into extinction coefficient spectra, from which the cross section (σ_{dye}) can be calculated at any wavelength by 2303/$N_A\varepsilon = \sigma$, where N_A is Avogadro's number. From absorbance spectra of the Cy3 and Cy5, we determine σ_{Cy3} = 3.49 x 10^{-16} cm^2 and σ_{Cy5} = 2.702 x 10^{-17} cm^2 at 532 nm. The quantum yield (Q_{dye}) of Cy3 is estimated to be 0.25 on the RNA construct and the Cy5 quantum yield is 0.3 (116). The laser power is measured at the microscope objective (1 µW). To calculate the peak intensity of the laser at the objective, we consider the Gaussian intensity profile of the beam,

$$I(r) = I_0\ e^{-4(\ln 2)r^2 / r_{FWHM}^2} \ ,$$
2.3

where I_0 is the peak intensity, r is the radial distance, r_{FWHM} is the full width half maximum, and $I(r)$ is related to the laser power P, by the integral,

$$P = \int_0^\infty r\ dr \int_0^{2\pi} d\theta\ I_0\ e^{-4(\ln 2)r^2 / r_{FWHM}^2} \ ,$$
2.4

yielding the laser peak intensity at the focus,

$$I_0 = \frac{4\ (\ln 2)\ P}{\pi\ r_{FWHM}^2} \ .$$
2.5

Thus, for 1 µW of input power and a r_{FWHM} of 270 nm (Eq. 2.1), I_0 = 1.2 kW/cm^2 (or kJ s^{-1} cm^{-2}). Thus we predict a 283 emission rate (Φ_{emit}) for Cy3 in the absence of FRET (acceptor bleached), and only 26.3 kHz for direct excitation of Cy5 at 532 nm (donor bleached) according to Eq.   2.2. Measuring the detected emission count rates on 20 doubly labeled RNA molecule (Figure 2.5 A) , subtracting background and dividing by the predicted count rate, we determine the collection efficiency of Cy3 and Cy5

on the donor (*1*) and acceptor (*2*) detectors to be $\beta_{A}{}^{1}$ = 0.000(3)%, $\beta_{A}{}^{2}$ = 2.42(18) % and Cy5 to be $\beta_{D}{}^{1}$ =

2.69(24) %, $\beta_{D}{}^{2}$ = 0.211(18)%. Thus, there is no cross talk of Cy5 onto the channel optimized for Cy3, but

there is a minor bleedthrough of Cy3 onto the acceptor channels. In this work we only use Cy3 and Cy5

labeled RNAs. Collection efficiencies would need to be reassessed for alternate dye pairs. Also one

should take note that throughout this worked the detected count rate of the nominally donor and accep-

tor signals is referred to as emission intensity (in units of kcounts/s or kHz).

2.1.7   Instrument Characterization: Detection Volume

Fluorescence correlation spectroscopy is used to assess dimensions of the confocal (detection) volume

(Figure 2.4 *A*). Mean cross correlations are measured for tetramethyrhodamine (TMR), which has a

well-known diffusion coefficient (D = 280 μm^2/s). The cross correlations,

$$G(\tau) = \frac{\delta F_{i}(t)\delta F_{j}(t+\tau)}{\langle F_{i}(t)\rangle\langle F_{j}(t)\rangle},$$

2.6

are calculated by a software analysis of 60 s time-correlated single-photon counting data traces; two

such cross correlations are averaged from the same sample to calculate the mean cross correlation and

standard deviations (σ). $G(\tau)$ is fit to the equation for 3D diffusion,

2.7

$$G(\tau) = \frac{1}{N}\left(1+\frac{4D\tau}{r_{0}{}^{2}}\right)^{-1}\left(1+\frac{4D\tau}{z_{0}{}^{2}}\right)^{-\frac{1}{2}},$$

with $1/\sigma^2$ weighting. N is the mean occupancy of molecules in the detection volume, D is the diffusion

coefficient, and r_0 and z_0 are the $1/e^2$ intensity radii for the lateral and axial dimesion of the 3D-Gaussian

profile, as depicted in Figure 2.4 (131,132). These values yield the confocal "detection" volume of 0.36

fL, calculated as V_{eff} =$\pi^{3/2}$ $r_0{}^2$ z_0 (132). To ensure single occupancy of the confocal volume, one only need

be sure that the average occupancy of this tiny volume is less than one using Poisson statistics (section

2.2.4).
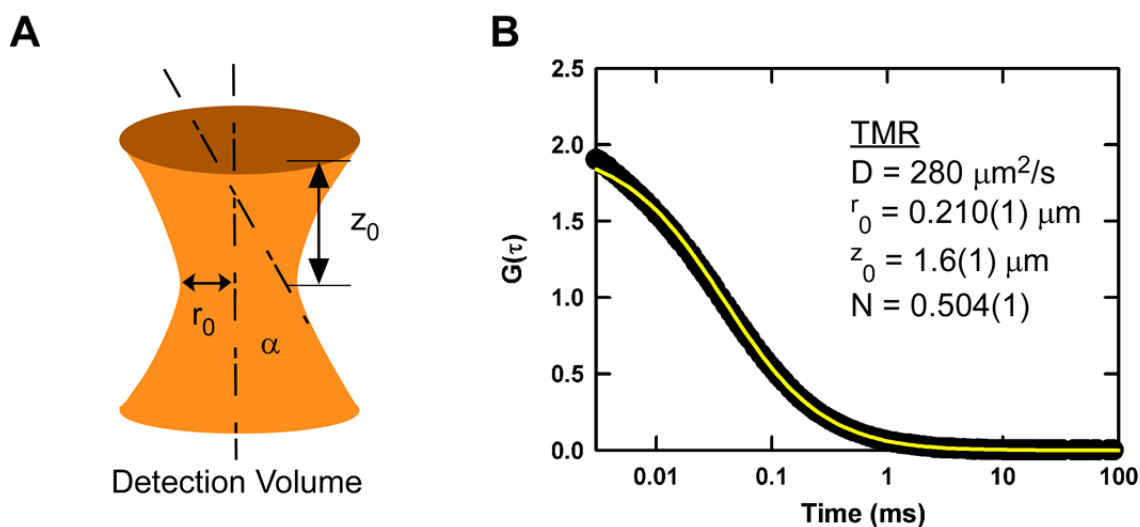
Figure 2.4   (*A*) Gaussian profile of the confocal detection volume with $1/e^2$ radius intensity drop off in the lateral (r_0) and axial (z_0) dimensions. (*B*) Cross correlation of donor channels for tetramethylrhodamine (TMR) with a known diffusion coefficient (*D*) and fit to Eq. 2.7 to yield the r_0, z_0, and *N* (the mean occupancy of focus).

2.1.8   Single-Molecule Fluorescence Time Trajectories

Fluorescence trajectories are acquired by locating individual RNA molecules in the focus of the laser via an intensity search algorithm during line scans. In the event that the intensity surpasses the threshold, the pixel with the highest intensity is identified and the stage is driven to this target pixel for acquisition of a fluorescence trajectory (Figure 2.5).   Because data analysis is time consuming, the contents of the FIFO memory in the SPCM are stored in a binary file for later processing. See Chapter 3 for discussion of acquisition of time trajectories as molecules freely diffuse through the focus. With a count rate of >6 kHz at 5 ms binning, the shot-noise limited signal/noise is >5/1. Single molecules are identifiable by photobleaching, donor and acceptor can be seen to bleach successively. For example, after the acceptor photobleaches in Figure 2.5, the donor continues to emit, until it, too, photobleaches. For measurements, powers of ~1 µW (measured on the microscope objective) are used for immobilized RNA and ~100 µW for freely diffusing RNA are typically used. Powers are chosen as a compromise between sig-

nal/noise and fluorophore photobleaching rates.
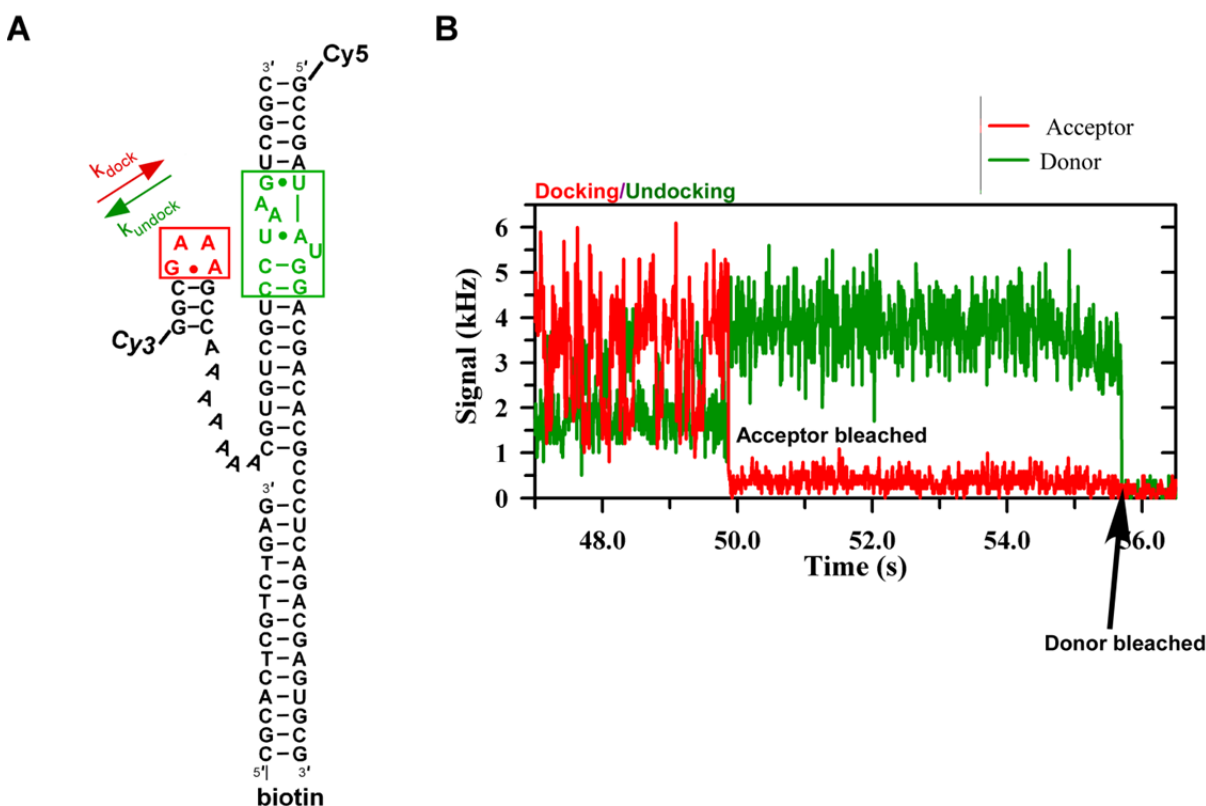
Figure 2.5   Tetraloop–receptor docking/undocking monitored by FRET.   (*A)* Schematic of the te-traloop–receptor RNA construct characterized.. Cy3 (donor) and Cy5(acceptor) labels allow monitoring of the GAAA tetraloop docking into its receptor by changes in FRET efficiency (E_{FRET}). The RNA is immo-bilized on glass surfaces with biotin-streptavidin binding for long time observation. (*B*) The donor and acceptor fluorescence emission from a single molecule are monitored in real time, fluctuating as the te-traloop and receptor dock and undock (100 mM NaCl, 50 mM HEPES, 100μM EDTA). Single molecules are identifiable by photobleaching; in this case the acceptor photobleaches, then the donor, resulting in background signals.

2.1.9   Temperature Control

Temperature control (± 0.2 °C) of the sample is achieved by placing the microfluidic flow cell (section

2.2.2) (128) into an enclosed heated stage (HSC60, Instec, Boulder, CO), which is placed on the micro-

scope stage with sticky, pliable plastic supports. A seal is created between the cover glass of the sample
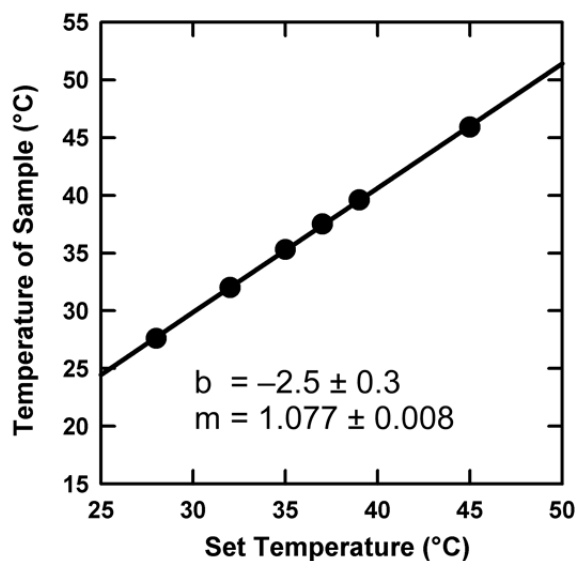
Figure 2.6   Calibration of the assembled sample heating system using the Bioptechs objective heater and Instec HSC60 stage. The objective heater controller is set to 1 °C cooler than the stage heater. The actual temperature is measured by a thermocouple inserted into a buffer-filled flow cell. The set and actual temperatures are in good agreement. The calibration curve (with slope *m* and intercept *b)* is used to correct the set temperatures for data analysis.

holder and the heated stage by Dow corning grease to ensure good thermal contact. The objective is

thermally isolated from the microscope turret with a threaded plastic spacer. The objective is in ther-

mal contact with the sample through the immersion water, so it is also temperature controlled with a a

resistively heated collars (Bioptechs, Butler, PA). To account for the thermal characteristics of the objec-

tive is set 1 degree cooler than the stage per suggestion from Bioptechs. A thermocouple inserted into

the buffer-filled flow cell is used to calibrate the temperature of the fully assembled sample-heating sys-

tem under data acquisition conditions, i.e, with immersion water and proper focusing. The calibration is

performed by raising and lowering the temperature to confirm that there is no hysteresis. The sample is

allowed to equilibrate at the temperature for 5 min before making the measurement.

2.2 Sample Preparation

2.2.1 Preparation of the GAAA Tetraloop–Receptor RNA Construct

The GAAA tetraloop–receptor tertiary interaction is isolated in an RNA construct (Figure 2.5), where docking and undocking of the tetraloop and receptor domains are enabled by a single-stranded A_7 linker. The A_7 linker sequence is one of several constructs used to study the effect of the sequence flexibility and length on the dynamics of the tetraloop–receptor interaction (106,124). A negative control "non-docking" system was also verified by replacement of the GAAA loop with a UUCG loop (106).

Tetraloop–receptor RNA constructs are prepared as previously described (106,124). Synthetic 5′ amino-modified (with a three carbon linker) RNA oligomers (Dharmacon, Lafayette, CO) are labeled with Cy3 and Cy5 *N*-succinimidyl esters (NHS) according to manufacturer protocols (Amersham Biosciences, Piscataway, NJ). Unreacted dyes are removed by microfiltration, followed by C-18 reverse-phase high performance liquid chromatography to separate the labeled and unlabeled RNAs. Synthetic RNA samples are desalted and deprotected. Hybridization of the Cy3- and Cy5-labeled RNA oligomers forms the tetraloop and receptor domains. The sequences of the RNA oligonucleotides are 5′-Cy5-GCC GAU AUG GAC GAC ACG CCC UCA GAC GAG UGC G-3′ and 5′-Cy3-GGC GAA AGC CAA AAA AAC GUG UCG UCC UAA GUC GGC-3′. The complete construct (Figure 2.5) is formed by annealing the Cy3 (1 µM) and Cy5 (1.5 µM) RNA oligomers with 2 µM biotinylated DNA oligomer (5′-biotin-CGC ACT CGT CTG AG-3′, Integrated DNA Technologies, Coralville, IA) by heating to 70 °C and cooling slowly to room temperature in an annealing buffer of 50 mM HEPES, 100 mM NaCl, 100 µM EDTA, pH 7.5. The 14mer DNA/RNA hybrid tether is biotinylated for immobilization on passivated glass surfaces via biotin-streptavidin chemistry. Such tethering capabilities are not necessary to perform smFRET experiments on freely diffusing RNA, though it is always included to allow for explicit comparison between studies (106,124). All sample solution contain 50 mM hemisodium HEPES (pH 7.5, Sigma Aldrich, St. Louis, MO), 100 mM NaCl and 0.1 mM EDTA unless otherwise noted. All buffers are 0.2 µm sterile filtered prior to preparing RNA solutions, and are prepared in sterile water (either autoclaved doubly distilled or LC-MS grade Chromsolv, or

NANOPure water). Chemicals are also purchased in RNAse free forms if available. All pipette tips and vials are purchased RNAse free.

2.2.2 Sample Holder Preparation for Single-Molecule Experiments

RNA Samples for single-molecule observation are prepared in a flow cell that enables easy changing of experimental conditions ([salt]) and can kept be air tight to reduce photobleaching of the dyes by an influx of oxygen (Figure 2.7). RNA samples are always aqueous solution. The flow-cell sample holder is assembled from a fluorinated polymer block (polychlorotrifluoroethylene, PCTFE) by milling a "microchannel" (0.2 mm deep, 10 mm long, 3.0 mm wide) on one of the block faces (Figure 2.7). Two 0.5 mm diameter holes are drilled in from the opposite side to connect at a 60° angle to the channel ends. A cover glass is attached to the holder with a thin layer of silicone rubber adhesive to form an ~10 μL volume flow cell. A slot is milled above the channel, leaving only 0.5 mm of polymer above the illumination area to limit back scattering of the excitation source. For a detail schematic, see Appendix A. After final solutaion are flushed into the cell, data are collected under static conditions, with entrance and exit holes covered by tape. All cover glass (22 x 22 mm Corning (Corning, NY) No. 1-1/2) is cleaned by soaking overnight in concentrated nitric acid followed by thorough rinsing with NANOPure water (Barnstead International, Dubuque, IA) and a 45 min treatment with ozone/UV light (UVO cleaner model 42, Jelight
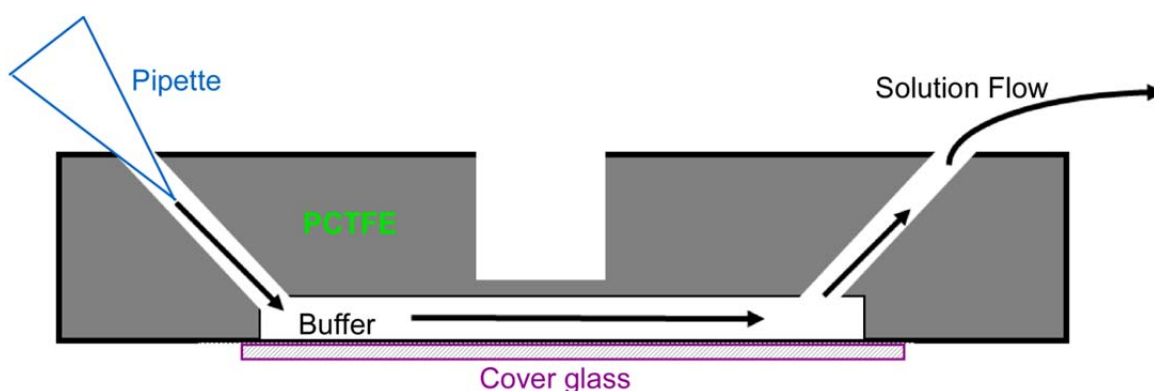
Figure 2.7   Schematic of microfluidic sample flow cell (not to scale). The buffer chamber volume created by the coverglass and plastic (PCTFE) block is ~10 μL and can be closed off from the environment by tape. Solutions are flowed through the cell with a micropipettor to prepare the sample or change buffer conditions. Technical drawings for flow cells are shown in Appendix A.

Company, Irvine, CA). An alternative version of the flow cell was also designed to allow for mounting on the microscope stage and to accommodate tubing for solution flow (Appendix A), such that the same molecules can be imaged under different solvent conditions. After each use, the sample holders are cleaned with acetone, methanol, NANOpure water, ethanol, and ozone cleaned for 45 min and glued with fresh coverglass on the day of experiments. Solvents must be flushed through the inlet holes of the cell when cleaning them. Glue should be allowed to cure for 2 hours before the onset of sample preparation. The sample holder can also be cleaned with piranha solution (50% concentrated sulfuric acid and 50% hydrogen peroxide (30% concentration)), rinsed thoroughly with NANOpure water, then ozone cleaned. LC-MS grade Chromsolv has also given the best results.

2.2.3   Single-Molecule Sample Preparation: Immobilized

The glass immobilization surfaces are prepared in the assembled sample holders (Figure 2.7) by first flushing the cell with 1 mL of water and 1 mL of working buffer (50 mM hemisodium HEPES, 0.1 mM EDTA, 100 mM NaCl) by inserting the micropipette tip into the hole of the flow cell. The holder is, then flushed with 120 µL of bovine serum albumin (10 mg/mL bovine serum albumin in working buffer), followed by 10 min. incubation, 1 mL buffer rinse, and then a 120 µL strepatividin (0.2 mg/mL) flush/10 min. incubation. Then ~100 µL of the desired RNA solution are flowed in (~50 pM) to achieve the desired surface coverage. A final 1 mL contains the desired conditions for sample observation (i.e., appropriate salt concentration) and an oxygen scavenging cocktail. Unless otherwise specified, experiments are performed in 50 mM hemisodium HEPES buffer (pH 7.5 at 25 °C) with 100 mM NaCl, 0.1 mM EDTA. An oxygen scavenger of 60 nM protocatechuic acid, 5 mM protocatechuate-3,4-dioxygenase, and 2 mM Trolox (130) is added to the buffer to reduce fluorophore photobleaching and photophysics. An alternative enzymatic oxygen scavenging solution of glucose (9 mg/mL), glucose oxidase (0.43 mg/mL), catalase (0.072 mg/mL), and 2 mM Trolox was initially used and noted accordingly in such instances (129).

2.2.4   Single-Molecule Sample Preparation: Freely Diffusing

For freely diffusing studies (118,133) , the laser focus is positioned 15 μm above the cover glass in solution to observe freely diffusing molecule at powers of 50-100 μW. At such powers, photobleaching of freely diffusing molecules is minimal in our experimental apparatus, eliminating the need for an enzymatic oxygen scavenging system (118,134). The coverglass is prepared by first rinsing with 1 mL of buffer, flushing with 10 mg/mL of BSA, 10 min. incubation, 1 mL buffer flush, then flushing in the desired RNA sample (~100 pM).   Labeled RNA diffusing through the confocal excitation region is readily monitored via isolated bursts of fluorescence photons, resulting in a mean occupancy of about $\lambda = 0.03$ molecules in the ~0.5 fL confocal volume, and the molecule is only detectable for ~ 1 ms as it quickly diffuses out of the focal spot. The probability of m molecules in the detection volume is $P(m) = \dfrac{\lambda^m e^{-\lambda}}{m!}$, translating into < 0.044% probability for events with $m \geq 1$. Because the freely diffusing molecule spends so little time in the focus, the higher laser powers are needed to obtain a sufficient number of photons for E_{FRET} analysis. See Chapter 3 for more details on performing freely diffusing data acquisition and analysis.

2.3   FRET Analysis

2.3.1 E_{FRET} from Donor and Acceptor Emission Intensities

Immobilized single-molecule trajectories are analyzed with 3−10 ms data binning, which clearly resolves the undocked and docked events of RNAs studied in this thesis, Figure 2.5 *B*, (106,128). From these time trajectories each bin can be converted into a FRET efficiency, E_{FRET}, from the donor and acceptor mission intensities. Ideally, if collection of the donor and acceptor emission (I_D and I_A) is:

$$E_{FRET} = \frac{I_A}{I_A + \dfrac{Q_A}{Q_D} I_D} = \frac{R_0^6}{R_0^6 + R^6},$$
<div align="right">2.8</div>

where Q_A and Q_D are the acceptor and donor quantum yields, respectively.   Thus, E_{FRET} can be used probe the conformational state of the dual-labeled RNA molecule by its sensitivity to fluorophore sepa-

ration (R) (see also Chapter 1). In an experiment, we do not detect the true donor and acceptor emission intensities, but rather some fraction of the signals determined by the collection efficiencies of the instrument. The intensity-based E_{FRET} (Eq. 2.3) is calculated from the background subtracted signals on the two channels, ΔI_1 and ΔI_2, designed primarily for donor and acceptor detection, respectfully. Corrections are implemented for (i) collection efficiencies and crosstalk of the donor and acceptor emission on channels 1 and 2 (β_1^A, β_2^A, β_1^D, β_2^D) (ii) differential quantum yields of the donor and acceptor (Q_D and Q_A), and (iii) fractional direct excitation of the acceptor vs. donor (α_A, where $1-\alpha_D = \alpha_A$), yielding,

$$E_{FRET} == \frac{\beta_1^D \Delta I_2 - \beta_2^D \Delta I_1 - \dfrac{\alpha_A}{\alpha_D}\dfrac{Q_A}{Q_D}(\beta_2^A \Delta I_1 - \beta_1^A \Delta I_2)}{\beta_1^D \Delta I_2 - \beta_2^D \Delta I_1 + \dfrac{Q_A}{Q_D}(\beta_2^A \Delta I_1 - \beta_1^A \Delta I_2)} \qquad 2.9$$

Quantum yield ratios and collection efficiencies are determined in independent measurements of singly labeled constructs, (Q_A/Q_D) = 1.2 ± 0.3, β_1^A (0.00000 ± 0.00003), β_2^A (0.0242 ± 0.0018), β_1^D (0.0269 ± 0.0024), β_2^D = 0.00211 ± 0.00018 (see section 2.1.6). For full derivation of the correction equation see Chapter 3 (128). Fractional direct laser excitation of the acceptor and donor is calculated from the extinction coefficients of Cy3 and Cy5 at 532 nm, α_A = 0.07 ± 0.01, α_D = 0.93 ± 0.01 (128). For each time bin in a trajcetory, E_{FRET} is calculated with the correction factors (Figure 2.8).
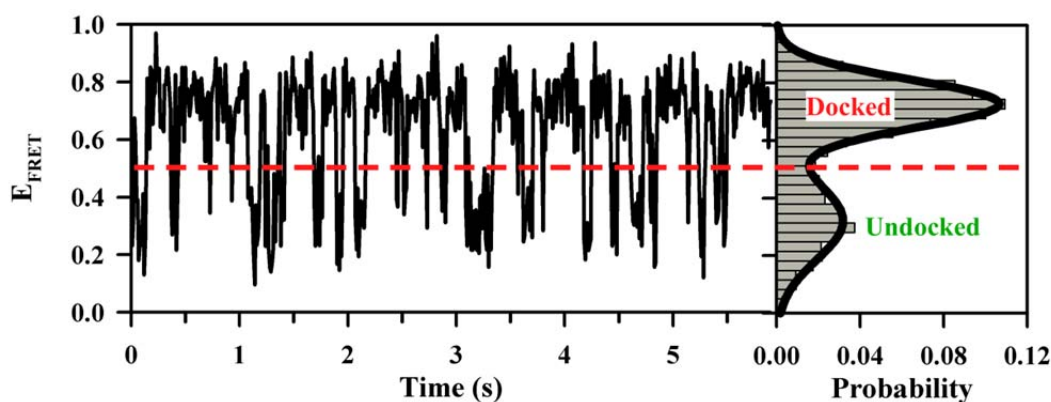
Figure 2.8 E_{FRET} trajectory for tetraloop–receptor docking calculated using intensity correction for the molecule in Figure 2.5 (1 mM MgCl$_2$, 100 mM NaCl, 100 μM EDTA, 50 mM HEPES). Docked and undocked states are well resolved with mean E_{FRET} of 0. 29(2) and 0.7(2) as determined from many molecule trajectories.

Analysis of either immobilized molecules or freely diffusing yield population distribution of the tetraloop−receptor RNA with E_{FRET} centers for the docked and undocked states determined by fitting to a sum of two Gaussian distributions, with the mean E_{FRET} given by the Gaussian center (Figure 2.8). Immobilized trajectories yield more information than freely diffusing trajecotires because the molecule is observed long enough to detect events of docking and undocking, such that these dwell times can be used to calculate the rate constants for docking and undocking. All single-molecule analysis is performed in Lab Window CVI programs developed in the Nesbitt lab and shared by multiple users.

2.3.2 E_{FRET} from Donor Fluorescence Lifetime

An alternative method for calculating E_{FRET} in our systems is by utilizing the microtime data, i.e., the fluorescence lifetime of the donor. This method has an advantage over the intensity-based method in the previous section because it does not require any of the corrections (quantum yield, collection efficiencies, and direct excitation of the acceptor). The only requirement for this method is that there is no cross talk of acceptor onto the donor channel, such that the lifetime observed originates from purely donor photons, which is achieved with our optical filter selections. The main disadvantage of this method is that more photons are needed to determine the accurate lifetime need to calculate E_{FRET}, as compared to the intensity method.

Determining E_{FRET} from donor lifetimes is based on the kinetic scheme for the possible relaxation pathways of a donor fluorophore (*D*) excited by a laser photon (*hυ*) (Figure 2.9 *A*). In this scheme, the efficiency of energy transfer is the fractional excitation rate of an acceptor molecule (*A*) by the donor,

$$E_{\text{FRET}} = \frac{k_T}{k_T + k_{\text{Drad}} + k_{\text{Dnonrad}}}, \qquad\qquad 2.10$$

where k_T is the rate constant for energy transfer from the donor to acceptor, and k_{Drad} and k_{Dnonrad} are the radiative and nonradiative decay rate constants for the donor (Figure 2.9 *A*). The fluorescence life-

time of the donor (τ_D) in the absence of energy transfer is the reciprocal of the sum of the radiative and nonradiative rates,

$$\tau_D = \frac{1}{k_{\text{Drad}} + k_{\text{Dnonrad}}} \qquad\qquad 2.11$$

In the presence of an acceptor, there is additional opportunity of donor relaxation by energy transfer to the acceptor and thus the donor lifetime in the presence of an acceptor (τ_{DA}) is

$$\tau_{DA} = \frac{1}{k_T + k_{\text{Drad}} + k_{\text{Dnonrad}}}, \qquad\qquad 2.12$$

where k_T depends on the distance (R) between donor and acceptor as $1/R^6$ (111). Rearrangement and substitution of Eqs, 2.11 and 2.12 into Eq. 2.10 yields an expression for E_{FRET} in terms of only donor life-times

$$E_{\text{FRET}} = 1 - \frac{\tau_{\text{DA}}}{\tau_{\text{D}}}. \qquad\qquad 2.13$$

τ_D can be measured for an RNA lacking an acceptor label, or from photons arriving after the acceptor photobleaches (Figure 2.5). τ_{DA} is dependent on the conformational state of the RNA and thus must be assessed for the donor in the presence of active acceptor in either the undocked or docked state. There are insufficient donor photons to determine a fluorescence lifetime on a per bin (typically 5 ms) level of a single-molecule trajectory. When the molecule is undocked, E_{FRET} is low and therefore donor emission is especially weak. Thus, though it is challenging to generate fluorescence lifetimes on a bin-wise level, we can group all photons in a trajectory that originate from the docked state or undocked states by set-ting a threshold between the well-resolved states of high and low donor emission signal. The donor-only lifetime can be assessed by grouping the photons after the acceptor photobleaches (Figure 2.5).

By grouping donor photons in a single-molecule trajectory as donor only, docked, or undocked, we obtain enough photons to determine the fluorescence lifetimes, τ_D, τ_{DA}(docked) and τ_{DA}(docked), as shown in Figure 2.9 *B*. Each of the Cy3 lifetimes clearly deviates from mono-exponential behavior pre-dicted by the scheme in Figure 2.9. This observation is consistent with measurements by Lilley and coworkers that have attributed the presence of multiple lifetimes to alternative conformations of the
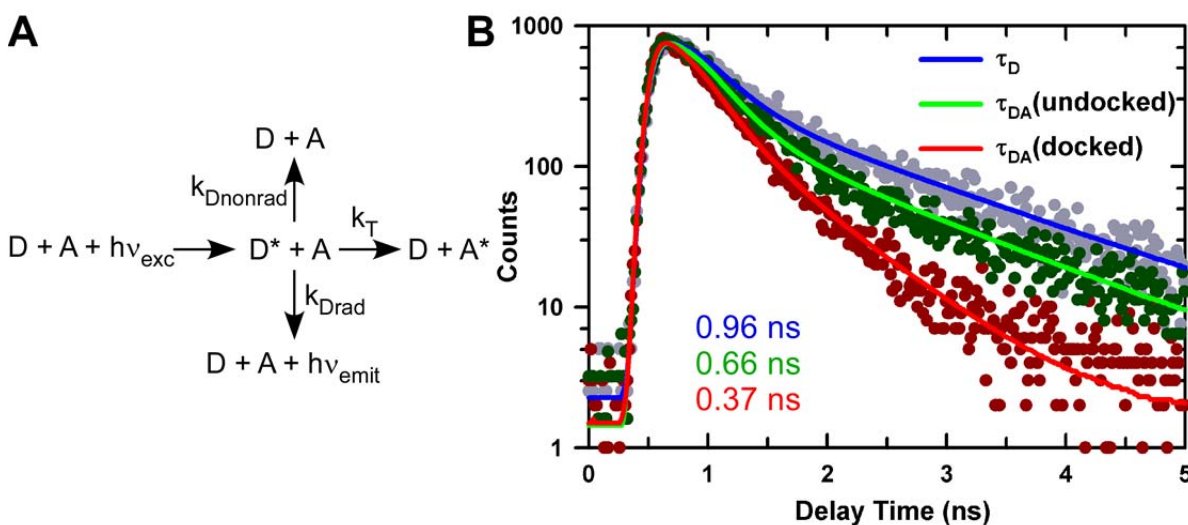
Figure 2.9   Dependence of donor fluorescence lifetime on FRET. (*A*) Kinetic scheme for relaxation of a donor molecule (D) excited to a state D* by a photon ($h\nu_{exc}$), from which it can relax with a nonradiative rate ($k_{Dnonrad}$), radiative rate (k_{Drad}), or transfer energy to an acceptor molecule by FRET (k_T). (*B*) The fluorescence lifetimes of the donor (Cy3) in the RNA FRET construct (Figure 2.5) in the absence of Cy5 (τ_D), in the presence of the Cy5 while the molecule is undocked ($\tau_{DA}(undocked)$) or docked (($\tau_{DA}(undocked)$)). The scheme in *A* predicts a monoexponential decay rate, yet the Cy3 decays are clearly multi-exponenial. There is a shortening of the lifetimes due to FRET, with energy transfer most efficient in the docked conformation when the donor is closer to the acceptor. Data are fit with a bi-exponential decay convoluted with the instrument response function using a maximum likelihood estimation (PicoQuant Symphotime software). The average lifetime for each donor state is shown on the plot (colored coded). Measurements were made at PicoQuant, GmbH on Microtime 200 with IRF < 100 ps FWHM. Similar measurments can be made in the Nesbittt lab (see Figure 6.10).

Cy3, which interconvert on a time scale much slower than the fluorescence lifetime (135) (113,116). For example, when the dye is stacked on nucleic acid bases it possesses a high quatum yield (long lifetime). When the Cy3 is unstacked, the quantum yield is low corresponding to a short fluorescence lifetime (113). In either scenario, Cy3 can undergo energy transfer. The intensity (or quantum yield) observed for a donor over the course of time trace is summed over many fluorescence lilfetimes (>5 ms). Therefore, the intensity-based is reflective of an average lifetime or a population weighted of the multiple donor lifetimes. Thus, to determin an E_{FRET} from the lifetime method that would be comparable to the intensity method, we need to asses the average donor lifetime, i.e., ($\tau_{average} = P_1\tau_1 + P_2\tau_2 + ...$), where P_i, the probability for a given lifetime, is

$$P_i = \frac{\int_0^\infty A_i e^{-\tau_i t}}{\int_0^\infty (A_1 e^{-\tau_1 t} + A_2 e^{-\tau_2 t} + ...)} = \frac{A_i \tau_i}{A_1 \tau_1 + A_2 \tau_2 + ...}, \qquad 2.14$$

and A_i is the amplitude for a lifetime of τ_i and t is the delay time. Whether donor only, docked, or un-

docked, the Cy3 flourescence decay can be well described by a double exponential fit (Figure 2.5). The

lifetime components and average lifetimes are summarized for the single molecule in Figure 2.5 and for

several molecules in Table 2-1. Single-molecule measurements of the donor-only lifetime are in good

agreement with bulk measurements that are presented in section 6.8.1 ($\tau_{\text{average}} = 0.9 \pm 0.1$ ns).

Table 2-1   Cy3 lifetimes for donor only, undocked, and docked A$_7$ tetraloop−receptor RNA

	P_1	τ_1 (ns)	P_2	τ_2 (ns)	τ_{average}
Donor Only (τ_D)	0.41 ± 0.01	0.214 ± 0.004	0.59 ± 0.01	1.31 ± 0.14	0.86 ± 0.09
Undocked (τ_{DA})	0.49 ± 0.03	0.18 ± 0.03	0.51± 0.03	1.07 ± 0.06	0.59 ± 0.03
Docked (τ_{DA})	0.58 ± 0.06	0.13 ± 0.01	0.42 ± 0.06	0.60 ± 0.04	0.32 ± 0.02

Values are the average from multiple single molecules and uncertainties are the standard deviation of
the mean.

The donor lifetime is longest in the absence of acceptor and shortest when the molecule is

docked because energy transfer is more efficient. With the average lifetimes summarized in Table 2-1,

we can calclulate the average E_{FRET} for the tetraloop−receptor RNA according to Eq. 2.13. These results

are compared to the intensity-based method in Table 2-2. We can see that the two methods yield E_{FRET}

within experimental uncertainties, offering an independent verification that the microscope system is

well characterized for accurate extraction of E_{FRET} from the intensity-based method. E_{FRET} values are

also in agreement with prediction for the Cy3–Cy5 distances in the RNA construct (Table 2-2) from

structural modeling of the RNA construct as helical domains (Appendix B), we estimate the fluorophore

distances are ~40 and 60 Å for the docked and undocked states, respectively. From $E_{\text{FRET}}(R) = R_0^6/(R_0^6$

$+ R^6)$, and $R_0 \sim 53$ Å the corresponding E_{FRET} prediction can be calculated, as shown in Table 2-2.

Table 2-2   Measurement of average E_{FRET} from immobilized single-molecules by the lifetime or intensity method for the A_7 tetraloop−receptor RNA (Figure 2.5)

	E_{FRET} (undocked)	E_{FRET} (docked)
Lifetime method	0.32 ± 0.06	0.63 ± 0.04
Intensity method	0.29 ± 0.02	0.69 ± 0.02
Prediction	0.3	0.8

The lifetime E_{FRET} method is challenging because a large number of photons are required to fit an exponential decay, an even stricter requirement in the case of Cy3, which deviates from monoexponential behavior. Furthermore, the relative populations of these states can be affected by the attachment scheme of the donor to the nucleic acid (e.g., phosphoramidite or NHS ester) and the structure of the molecule (DNA vs RNA, double stranded vs. single stranded), with evidence of even a third longer lifetime in other systems (116,135). For the RNA construct, there is no apparent systematic shift in the populations of the lifetime states with docking/undocking (in the donor only condition the molecule is mostly docked, Figure 2.5). Accurate extraction of the very short lifetime component is also challenging because of convolution with the instrument response function. Multi-parameter fluorescence measurement (e.g. donor/acceptor emission intensity and fluorescence lifetimes) are a valuable tool for single-molecule spectroscopy, and further implementation of such tools is warranted (113,133). Analysis of the polarization information available in the described experimental apparatus may yield additional insight into the multi-exponential character of the Cy3 lifetimes, though is not discussed in this thesis.

Chapter 3   Monovalent and Divalent Promoted GAAA-Tetraloop–Receptor Tertiary Interactions from Freely Diffusing Single-Molecule Studies

This chapter is published and reprinted with permission from *Biophysical Journal* (2008), 95:3892-3905, Fiore *et al.*, © 2008 Biophysical Society. [3]

3.1   Abstract

Proper assembly of RNA into catalytically active 3D structures requires multiple tertiary binding interactions, individual characterization of which is crucial to a detailed understanding of global RNA folding. This work focuses on single-molecule fluorescence studies of freely diffusing RNA constructs that isolate the GAAA tetraloop–receptor tertiary interaction. Freely diffusing conformational dynamics are explored as a function of Mg^{2+} and Na^+ concentration, both of which promote facile docking, but with 500-fold different affinities. Systematic shifts in mean E_{FRET} values and linewidths with increasing $[Na^+]$ are observed for the undocked species and can be interpreted with a Debye model in terms of electrostatic relaxation and increased flexibility in the RNA. Furthermore, we identify a 34 ± 2% fraction of freely diffusing RNA constructs remaining undocked even at saturating $[Mg^{2+}]$ levels, which agrees quantitatively with the 32 ± 1% fraction previously reported for immobilized constructs. This verifies that the kinetic heterogeneity observed in the docking rates is not the result of surface tethering. Finally, the K_D value and Hill coefficient for $[Mg^{2+}]$-dependent docking decrease significantly for $[Na^+]$ = 25 mM vs. 125 mM, indicating Mg^{2+} and Na^+ synergy in the RNA folding process.

3.2   Introduction

RNA catalytic and biological functionality, such as translation and self-splicing, require that RNA molecules fold into specific three-dimensional forms (17). Central to achieving and maintaining a correctly folded RNA structure are tertiary contacts (6,42), although the contributions of *individual* tertiary interactions to the RNA folding pathway and conformational dynamics are not well understood (103,126). Proper RNA folding also requires counterions to minimize repulsions of the negatively charged phosphate backbone through site-specific coordination to the RNA and/or nonspecific (delocalized) electrostatic screening (6,12,40,41,58,62,81,136-139). However, tertiary interactions and counterions do not always lead to the desired result; they can stabilize misfolded or kinetically trapped states, thereby preventing RNA from rapidly achieving a catalytically competent structure (6,9,140-142). These kinetically trapped states often result in long-lived intermediate conformations and heterogeneous folding rates (143-145). This complicated folding behavior is not limited to large ribozymes with multiple domains, with similarly complex behavior noted even for simpler RNA systems such as the hairpin ribozyme (3,146). Clearly any predictive understanding of RNA folding dynamics will require addressing not only global issues of topological structure, but also *individual* tertiary interactions and their dependence on cationic environment.

In response to this need, we have been systematically investigating RNA tertiary binding motifs at the single-molecule level to characterize *isolated* interactions as a function of solution environment. The present work focuses on cation-dependent RNA folding due to the ubiquitous GNRA-hairpin tetraloop–receptor interaction (63), whereby a GAAA tetraloop docks into an 11-nucleotide internal receptor loop (42,62). These GAAA tetraloop and receptor structures, both free and receptor-bound, have been extensively studied by NMR and x-ray crystallography (50,62,74,80). Thermodynamic and kinetic contributions of the tetraloop–receptor interaction to RNA folding have been investigated at the ensemble level in a variety of cationic environments in isolation (77,124), in P4–P6 domains (70,126) and in group II introns (127). In bulk studies, however, there is no synchronization between folding/unfolding

events for different molecules. Therefore, one must either perturbatively induce a short-lived synchronization (e.g., laser-induced temperature or pH jump methods and rapid mixing) to initiate observable folding dynamics, or be restricted to observing static equilibrium properties of the system. Conformational heterogeneity, for example, due to misfolding events, is largely obscured because of the limited resolution of different states by traditional ensemble methods such as gel electrophoresis.

   Studies at the single-molecule level allow direct observation of conformational populations, providing a powerful tool for investigating structural fluctuations in nucleic acids (3,8,105,119,146-149). In particular, time-resolved–single-molecule fluorescence resonance energy transfer (smFRET) methods allow thorough characterization of RNA-folding dynamics under both equilibrium and nonequilibrium conditions. Folding and unfolding rate constants are directly measurable, intermediate or misfolded conformations are explicitly identifiable, and subpopulations that exhibit different rate constants and conformations (i.e., kinetic and static heterogeneity) are readily distinguishable (3,8,105,107,119,146-148,150-152). Such measurements often profit from the longer observation times afforded by immobilization of the molecule, e.g., tethering to a cover glass. Immobilization raises concerns that surface proximity and/or tethering methods may influence the folding dynamics and/or of functionality of biomolecules (153,154). Seminal studies of the *Tetrahymena thermophila* and hairpin ribozymes showed that catalysis was unaffected by immobilization (3,119), which suggests that RNA functionality is uninfluenced by surface tethering. Furthermore, Ha and coworkers showed that the kinetic heterogeneity identified for surface-immobilized hairpin ribozymes is also present when the molecules are encapsulated in liposome vesicles (155), where interactions with vesicular walls are thought to be minimal (156,157). Single-molecule studies of freely diffusing species complement tethered-molecule investigations, by providing the capability to interrogate and resolve conformational populations in solution without potential for surface interaction. By way of example, Deniz and coworkers have revealed many insights into the [Mg^{2+}]-dependent folding of the hairpin ribozyme through a systematic conformational study utilizing freely diffusing smFRET (149). Additionally, the Mg^{2+}-binding parameters ob-

served for immobilized hairpin ribozymes (K_D and Hill coefficient) were found to be quantitatively replicated in freely diffusing RNA (149).
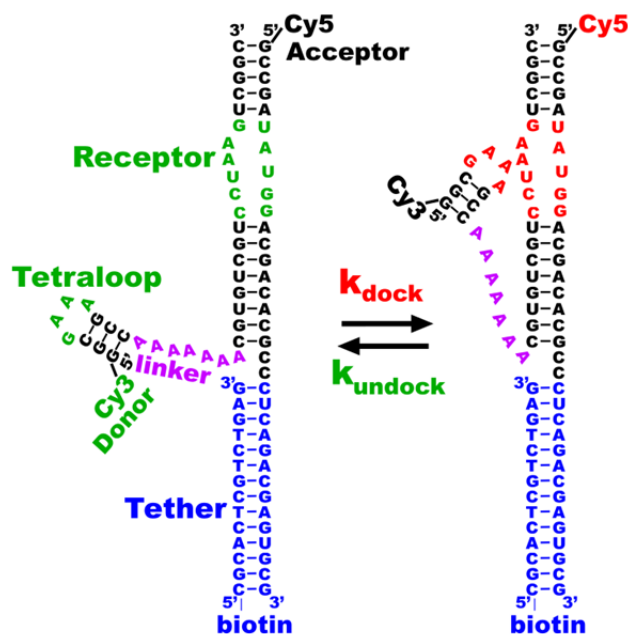
Figure 3.1   RNA construct for Cy3–Cy5 FRET-monitoring of GAAA tetraloop–receptor docking/undocking. The GAAA tetraloop and receptor are connected by a flexible A_7 linker (*purple*) and highlighted in the undocked (*green*) and docked (*red*) states. A biotinylated region (*blue*) is also retained for quantitative comparison with previous tethered results.

The focus of this work is smFRET studies of freely diffusing RNA containing a single tertiary binding motif, specifically the GAAA tetraloop–receptor interaction (Figure 3.1).   Folding of this RNA construct occurs when the tetraloop docks into the receptor domain via a flexible, single-stranded A_7 arm (Figure 3.1) (106,124). In particular, we explore the equilibrium effect of divalent (Mg^{2+}) and monovalent (Na^+) ions on the tetraloop–receptor docking/undocking in the absence of surface immobilization. Even for this simple, isolated RNA tertiary interaction, the folding dynamics prove quite rich. Both Mg^{2+} and Na^+ lead to proper formation of the tetraloop–receptor contact, though Na^+ permits additional electrostatic relaxation of the undocked structure over the concentration range studied. We also find that Mg^{2+} and Na^+ interact *synergistically* to enable tetraloop–receptor docking. Specifically, at low [Na^+]

($\approx$ 25 mM), Mg^{2+} displays highly cooperative Hill plots as a function of Mg^{2+} concentration. Mg^{2+}-induced folding becomes much more efficient (and correspondingly less cooperative) at higher Na^+ concentrations (125 mM Na^+), emphasizing the influence of electrostatic screening in the tetraloop–receptor construct. These results highlight the role of electrostatic shielding and conformational relaxation in the *undocked* species, as further demonstrated by a dramatic [Na^+]-dependent *increase* in both peak locations and widths in the FRET distributions.

As a secondary thrust, we quantitatively compare our studies of freely diffusing RNA with previous work on the surface-tethered GAAA tetraloop–receptor construct. The previous studies on the immobilized tetraloop–receptor construct revealed significant kinetic heterogeneity, i.e., two populations with distinctly different docking kinetics. These two populations could be characterized as either (i) an actively *docking* population exhibiting multiple docking/undocking transitions during a typical 30 second trajectory prior to photobleaching and (ii) a *non-docking* population with no transitions evident on a several minute time scale (106,124). Given the simplicity of this RNA system, such heterogeneity is particularly surprising. The freely diffusing results confirm that the fraction of actively docking vs. nondocking populations observed for tethered constructs is native to the RNA. We also find these populations to be independent of salt concentration and therefore not arising from improperly formed secondary structures (i.e., hairpin opening and closing.) Furthermore, quantitative comparison between (i) freely diffusing and (ii) immobilized conformational populations demonstrate GAAA-tetraloop–receptor folding dynamics to be unaffected by immobilization over a wide range of Mg^{2+} and Na^+ concentrations and cationic environments.

3.3   Materials and Methods

3.3.1   Preparation of RNA Construct

The GAAA tetraloop–receptor tertiary interaction is isolated in an RNA construct (depicted in Figure 3.1), where docking and undocking of the tetraloop and receptor domains are coupled by a single-stranded A_7 linker. The A_7 linker sequence is one of several constructs used to study the effect of the

sequence flexibility and length on the dynamics of the tetraloop–receptor interaction (106,124), for which the fraction of actively docking versus non-docking heterogeneity was quantitatively similar. Tetraloop–receptor RNA constructs are prepared as previously described (106,124). Briefly, synthetic amino-modified RNA oligomers (Dharmacon, Lafayette, CO) are labeled with Cy3 and Cy5 *N*-succinimidyl esters (Amersham Biosciences, Piscataway, NJ) and high performance liquid chromatography purified. (Mention of commercial products is for information only; it does not imply National Institute of Standards and Technology recommendation or endorsement, nor does it imply that products mentioned are necessarily the best available for the purpose.) Annealing of the Cy3- and Cy5-labeled RNA oligomers forms the tetraloop and receptor domains. A DNA oligomer (Integrated DNA Techonologies, Coralville, IA) is also hybridized to the complementary extension of one of the RNA oligomers generating the tether domain (Figure 3.1). The 14mer DNA/RNA hybrid tether is biotinylated for immobilization on passivated glass surfaces via biotin-streptavidin chemistry. Though such tethering capabilities are not necessary to perform smFRET experiments on freely diffusing RNA, it enables explicit comparison with our previous studies (106,124). The distances of the Cy3 and Cy5 in the undocked and docked states (Figure 3.1) are estimated to be < 70 Å and > 35 Å, respectively, from helical constraints (106). All solutions for freely diffusing studies of Mg^{2+}-mediated docking are diluted to ~100 pM RNA in donor (Cy3) strand in a standard buffer containing 50 mM hemisodium HEPES (pH 7.5, Sigma Aldrich, St. Louis, MO), 100 mM NaCl and 0.1 mM EDTA with varying [$MgCl_2$]. Detailed [Na^+]-dependent and monovalent/divalent-synergy studies of the tetraloop–receptor motif are performed in 50 mM hemisodium HEPES (pH 7.5) and 0.1 mM EDTA, with [NaCl] and [$MgCl_2$] varied. Note that 50 mM hemisodium HEPES contains 25 mM Na^+ even prior to added NaCl. All buffers are 0.2 μm sterile filtered and autoclaved prior to preparing RNA solutions.

3.3.2   Experimental Apparatus

To enable smFRET studies of freely diffusing RNA, we have adapted the previously described scanning confocal microscope system (106,124) in a similar manner to Deniz *et al*. and Rothwell *et al*. (118,133).

Briefly, the fluorescence excitation source is a mode-locked 82 MHz frequency doubled Nd:YAG laser (Model 3800, Spectra Physics, Mountain View, CA), spatially filtered through a single-mode fiber and linear polarizer. Trace infrared light from the Nd:YAG 1064 nm fundamental is removed by a bandpass filter (EX 530/10, Chroma Technology, Rockingham, VT) upon entrance of the excitation laser light into the back port of the microscope. The excitation light is focused through a water immersion objective (Olympus (Center Valley, PA), UPlanApo 60X, 1.2 numerical aperture) to a diffraction-limited spot. Fluorescence emission collected by the microscope objective is isolated from the excitation source with a dichroic beam splitter (550DRLP, Chroma Technology, Rockingham, VT) and spatially filtered with a 50 μm pinhole, limiting detection to the confocal volume. A broad-band-polarizing beam splitter cube divides the fluorescence into vertical and horizontal polarizations with respect to the linearly polarized laser excitation axis. Each polarization is further separated into donor (Cy3) and acceptor (Cy5) channels by a dichroic beam splitter (645DCXR, Chroma Technology), with photon color further defined by transmission through bandpass filters HQ585/70M and HQ700/75M (Chroma Technology), respectively. Photon detection is performed by four avalanche photodiodes (SPCM-AQR-14, Perkin-Elmer Optoelectronics, Fremont, CA). A full width half maximum instrument time response of 570 ± 5 ps is measured by prompt Raman scattering signal from H_2O. The 4-channel detection provides flexibility for additional investigations of fluorophore polarization anisotropy decay and rotational diffusion times. For the present studies, however, we focus on unpolarized FRET signals, summing over horizontal/vertical channels. Fluorescence trajectories are acquired for freely diffusing fluorescently labeled RNA constructs using a time-correlated single-photon counting (TCSPC) module (SPC-134 Becker & Hickl, Berlin).

Sample solutions are studied in a flow-cell sample holder assembled from a fluorinated polymer block (polychlorotrifluoroethylene, PCTFE) by milling a "microchannel" (0.2 mm deep, 10 mm long, 3.0 mm wide) on one of the block faces. Two 0.5 mm diameter holes are drilled in from the opposite side to connect at a 60º angle to the channel ends. A cover glass is attached to the holder with a thin layer of silicone rubber adhesive to form a ~10 μL volume flow cell. A slot is milled above the channel, leaving

only 0.5 mm of polymer above the illumination area to limit back scattering of the excitation source. After flushing in solutions, data are collected under static conditions, with entrance and exit holes covered by tape. All cover glass (22 x 22 mm Corning (Corning, NY) No. 1-1/2) is cleaned by soaking overnight in concentrated nitric acid followed by thorough rinsing with NANOPure water (Barnstead International, Dubuque, IA) and a 45 min treatment with ozone/UV light (UVO cleaner model 42, Jelight Company, Irvine, CA). The cover glass surface is passivated by a 120 μL flush and 10 minute incubation of bovine serum albumin (10 mg/mL bovin serum albumin), followed by a 1 mL buffer rinse, and then loaded by flowing in 200 μL of the desired RNA solution.

For freely diffusing studies, the laser focus is positioned 15 μm above the (BSA)-passivated cover glass in solution to observe freely diffusing molecule at powers of 50-100 μW. At such powers, photobleaching of freely diffusing molecules is minimal in our experimental apparatus, eliminating the need for an enzymatic oxygen scavenging system (118,134). Labeled RNA diffusing (100 pM) through the confocal excitation region is readily monitored via isolated bursts of fluorescence photons, resulting in a mean occupancy of about λ = 0.03 molecules in a ~0.5 fL confocal volume. The probability of m molecules in the detection volume is $P(m) = \dfrac{\lambda^m e^{-\lambda}}{m!}$, translating into < 0.044% probability for events with $m \geq 1$.

3.3.3 FRET Analysis of Time Traces of Freely Diffusing RNA

Fluorescence photons stored using TCSPC methods are recalled and sorted into 1 ms bins. This is comparable to the mean transit time through the confocal volume, as determined directly from fluorescence correlation spectroscopy ($t_{transit}$ = 0.38 ± 0.05 ms) (158,159). Based on rate constants obtained from our previous single-molecule studies, this time binning is considerably faster than the conformational dynamics (i.e., $1/k_{dock}$ or $1/k_{undock}$ > 5 ms) and therefore predominantly samples single docking/undocking events (106). We can analyze the photon bursts either at the single bin level or with a full-burst algorithm to group photons that most likely arrived from the same molecule (118,158,160). The full-burst-

level method, in which several time bins are combined to account for all emission collected from a single molecule's passage through the confocal volume, increases the signal-to-noise ratio but risks averaging docking/undocking events because of the decreased time resolution (161). For this reason, we choose to analyze bursts at the single-bin level, following the methods proposed by Schultz, Weiss and co-workers (118,160).

To distinguish RNA emission events from background, a threshold fluorescence signal is deter-mined by a minimum sum of donor and acceptor photons per bin. This criterion requires signal levels (typically 25-35 kHz) to be > 10 times the standard deviation of the background, which is sufficient to distinguish labeled RNA constructs with high statistical significance (118). The choice of threshold has been systematically varied and exhibits negligible effects on the results presented here. The FRET effi-ciency (E_{FRET}),

$$E_{FRET} = \frac{I_A{}^c / Q_A}{I_A{}^c / Q_A + I_D{}^c / Q_D},$$

3.1

can be calculated for each event from the *corrected* donor and acceptor intensities ($I_D{}^c$ and $I_A{}^c$), with Q_D and Q_A as the corresponding quantum yields. The corrected donor and acceptor intensities differ from background corrected intensities on the detection channels ΔI_1 and ΔI_2, respectively, because of (i) col-lection efficiencies of donor and acceptor emission on both channels (i.e., $\beta_A{}^1$, $\beta_A{}^2$, $\beta_D{}^1$, $\beta_D{}^2$) and (ii) non-negligible direct excitation of the acceptor (i.e., α_D vs. α_A, where $\alpha_D + \alpha_A = 1$). Corrections have been for-mulated for bulk FRET and single molecule studies, but not including each of the above contributing fac-tors (162,163). Weiss and coworkers have developed elegant methods for implementing such correc-tions using dual-laser excitation (164). Expressions for these corrections suitable under single-laser conditions are presented below.

Cross talk correction is implemented by expressing the experimentally observed intensity vec-tor (ΔI_1, ΔI_2) in terms of the actual donor/acceptor emission vector (I_D, I_A) via a 2 x 2 matrix equation,

$$\begin{pmatrix} \Delta I_1 \\ \Delta I_2 \end{pmatrix} = \begin{pmatrix} \beta_D{}^1 & \beta_A{}^1 \\ \beta_D{}^2 & \beta_A{}^2 \end{pmatrix} \begin{pmatrix} I_D \\ I_A \end{pmatrix},$$

3.2

where absolute collection efficiencies of the donor and acceptor on channels 1 and 2, $\beta_A{}^1$ (0.00000 ± 0.00003), $\beta_A{}^2$ (0.0242 ± 0.0018), $\beta_D{}^1$ (0.0269 ± 0.0024), $\beta_D{}^2$ (0.00211 ± 0.00018), are obtained from observed (ΔI_1 and ΔI_2) signals for known emission rates from donor and acceptor only constructs. The corresponding expressions for I_D, I_A, and I_D/I_A are obtained via matrix inversion of Eq. 3.2:

$$I_D = \frac{1}{\{\beta_D{}^1\beta_A{}^2 - \beta_A{}^1\beta_D{}^2\}}[\beta_A{}^2\Delta I_1 - \beta_A{}^1\Delta I_2], \qquad 3.3$$

$$I_A = \frac{1}{\{\beta_D{}^1\beta_A{}^2 - \beta_A{}^1\beta_D{}^2\}}[-\beta_D{}^2\Delta I_1 + \beta_D{}^1\Delta I_2], \qquad 3.4$$

$$I_D / I_A = [\beta_A{}^2\Delta I_1 - \beta_A{}^1\Delta I_2]/[-\beta_D{}^2\Delta I_1 + \beta_D{}^1\Delta I_2]. \qquad 3.5$$

Correction for direct excitation of the acceptor is most simply incorporated in a modified expression for E_{FRET}:

$$E_{\text{FRET}} = \frac{N_A - \alpha_A N_0}{N_A - \alpha_A N_0 + N_D} = \frac{I_A/Q_A - \alpha_A(I_A/Q_A + I_D/Q_D)}{I_A/Q_A - \alpha_A(I_A/Q_A + I_D/Q_D) + I_D/Q_D}$$

$$= \frac{1 - \dfrac{\alpha_A}{\alpha_D}\dfrac{Q_A}{Q_D}\dfrac{I_D}{I_A}}{1 + \dfrac{Q_A}{Q_D}\dfrac{I_D}{I_A}}, \qquad 3.6$$

where $N_0 = N_A + N_D = (I_A/Q_A + I_D/Q_D)$ represents the sum of acceptor and donor excitations and therefore $\alpha_A N_0$ is the fractional direct excitation of the acceptor. The quantum yield ratio (Q_A/Q_D) = 1.2 ± 0.3 ratios for Cy3 and Cy5 is determined independently via studies of singly labeled constructs. The fraction of acceptor direct excitation (α_A = 0.07 ± 0.01, α_D = 0.93 ± 0.01) can be obtained from extinction coefficients of singly labeled donor and acceptor at the direct excitation wavelength, $\alpha_A = \varepsilon_{A \text{ at 532 nm}}/(\varepsilon_{A \text{ at 532 nm}} + \varepsilon_{D \text{ at 532 nm}})$. Thus, a corrected E_{FRET} can be readily calculated from the experimental ΔI_1 and ΔI_2 signals via Eqs. 3.5 and 3.6 when collection efficiencies and quantum yields are measured.

Note that cross talk and direct excitation corrections in the above expressions treat all experi-mentally observed intensities equivalently, irrespective of whether photons came from donor-acceptor labeled RNA constructs or donor-only molecules. Consequently, $I_A{}^c$ is *overcorrected* by the removal of acceptor direct excitation when the acceptor is *missing* or *bleached*. Thereby, the correction yields *nega-tive* E_{FRET} values for donor-only molecules, which can be readily identified and isolated in E_{FRET} histo-grams. As a consistency check, the correction procedure can be tested by removing this direct excitation correction for a sample of donor-only molecules, which, as expected, yields $E_{FRET} = 0.01 \pm 0.01$, centered within uncertainty around zero.

3.4   Results and Discussion

3.4.1 Mg^{2+}-Induced Tetraloop and Receptor Docking in Freely Diffusing RNA

Burst fluorescence methods are used to observe [Mg^{2+}]-dependent folding due to a single GAAA te-traloop–receptor interaction in freely diffusing RNA constructs with Cy3 donor (Cy5 acceptor) fluores-cence tags positioned near the tetraloop(receptor), respectively. Docking of the tetraloop and receptor brings the dye pair in closer proximity; from $E_{FRET}(R) = R_0^6/(R_0^6 + R^6)$ this translates into readily meas-urable changes in FRET efficiency, where $R_0 \sim 53$ Å is the calculated Cy3/Cy5 Förster radius for 50% energy transfer probability. Fluorescence time trajectories are obtained as RNA molecules freely diffuse through the confocal detection volume. The left panel in Figure 3.2 *A* shows sample fluorescence traces as a function of time at low and high [Mg^{2+}]. A clear dominance of donor bursts is evident at low [Mg^{2+}] (Figure 3.2 *A* top), which shifts to predominantly acceptor emission under high [Mg^{2+}] conditions (Figure 3.2 *A* bottom).

Analysis of these fluorescence bursts provides statistical information on conformational proba-bilities, generated from all time bins with burst counts above the intensity threshold and representing $\approx 10^4$ RNA constructs passing through the detection region. The corresponding E_{FRET} histograms for the tetraloop–receptor constructs at low and high [Mg^{2+}] are shown in Figure 3.2 *B*, top and bottom, respec-tively. These E_{FRET} distributions reveal three distinct populations for the tetraloop–receptor construct
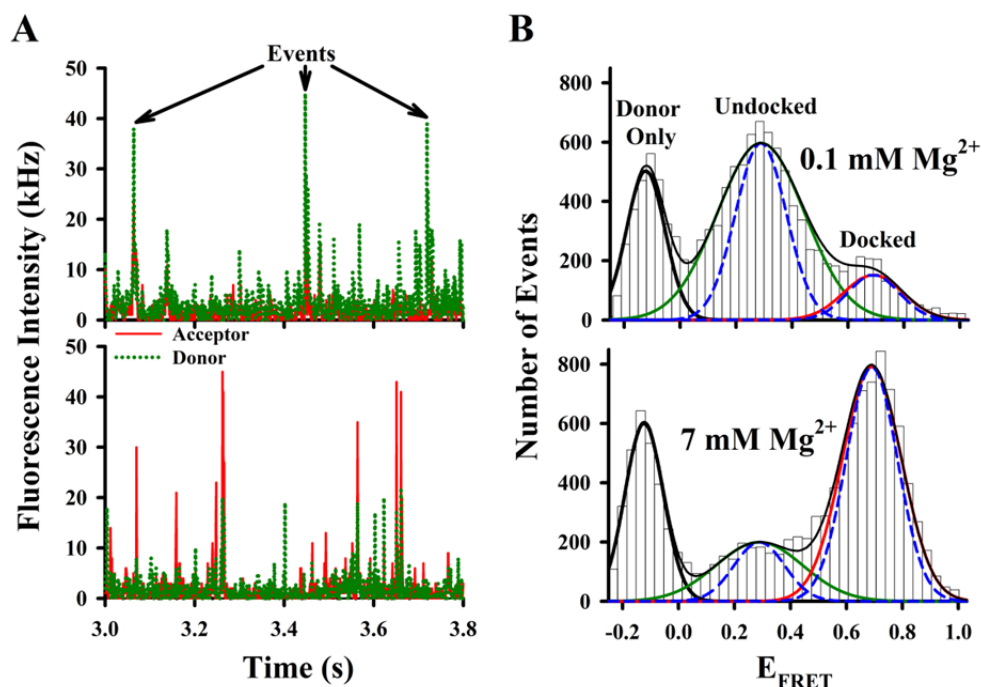
Figure 3.2   Sample smFRET data analysis of freely diffusing tetraloop–receptor RNA. (*A*) Sample time traces of donor (*dotted green lines*) and acceptor (*solid red lines*) fluorescence intensities at 0.1 mM Mg^{2+} (*top*) and 7 mM Mg^{2+} (*bottom*) indicate photon burst events as a molecule traverses the laser focal volume. (*B*) FRET efficiency (E_{FRET}) histograms generated from events that exceed a 25 kHz threshold at 0.1 mM Mg^{2+} (*top*) and 7 mM Mg^{2+} (*bottom*) fit to a sum of three Gaussian distributions (*black line*). The individual Gaussian components reveal distinct populations of donor-only ($E_{FRET} < 0$, *thick black),* undocked (*green*) and docked (*red*) constructs. Dashed blue lines represent shot-noise limited line-shape predictions.

well fit by a sum of Gaussians (160), with the individual components shown in Figure 3.2 *B* (*black*, *green* and *red lines*). As predicted, the left-most histogram peak ($\langle E_{FRET} \rangle = -0.124 \pm 0.003$) arises from donor-only molecules, resulting in $E_{FRET} < 0$ when corrected for direct excitation of the missing Cy5. Such donor-only bursts arise from incomplete constructs lacking an acceptor labeled strand, properly assembled RNA constructs lacking the fluorescent acceptor, as well as from free dye and acceptor photobleaching. The integrated number of such donor-only burst events is typically < 30 % of the total and clearly resolved from each of the physically relevant peaks with $E_{FRET} > 0$.

In a simple two-state model, the two well-resolved peaks at $E_{FRET} > 0$ correspond to the conformations of the tetraloop–receptor construct. The peaks centered at a low E_{FRET} (0.28 ± 0.01) and high E_{FRET} (0.687 ± 0.005) reflect the undocked and docked populations, consistent with predictions ($E_{FRET} \approx$

0.2, $E_{\text{FRET}} \approx 0.8$) based on modeling the donor/acceptor separation with helical RNA constraints. Dependence of tetraloop–receptor docking on [Mg^{2+}] is evident in the E_{FRET} histograms; the docked state is far more stable at high [Mg^{2+}], whereas the undocked state is favored at low [Mg^{2+}] (Figure 3.2 *B*).

Closer inspection of the peak widths in Figure 3.2 *B* provides additional insight into the nature of docked vs. undocked states for the tetraloop–receptor interaction. The curves (*dashed, blue lines*) in Figure 3.2 *B* represent shot-noise limited peaks, where the width, σ_{SN}, is the standard deviation of the Gaussian distribution about a mean E_{FRET} ($\langle E_{\text{FRET}} \rangle$) due to finite photon statistics for donor and acceptor intensities. From standard-error propagation, this shot-noise broadening is $\sigma_{\text{SN}} = (E_{\text{m}}(1-E_{\text{m}})/T)^{1/2}$, where E_m is the $\langle E_{\text{FRET}} \rangle$, and T is the photon threshold for event identification (161). Gaussian fits to the experimental histograms reveal the undocked and docked peak widths to be 1.8 ± 0.2 and 1.1 ± 0.1 times the shot-noise limit, respectively. Deviation from the shot-noise-limited behavior is therefore substantial for the undocked peak, whereas the docked peak indicates no significant broadening beyond shot noise (Figure 3.2 *B*).

There are many possible dynamical sources of such E_{FRET} broadening effects. For example, this broadening can arise from acceptor photobleaching or blinking during passage through the laser focus, as demonstrated in studies by Weiss, Seidel and co-workers (165,166). However, such broadening is small with respect to what we measure, and would contribute to asymmetric tailing of undocked and docked populations towards the donor-only peak (165,166), which is not evident in the data. Other broadening mechanisms include hindered rotational motion, spectral diffusion, and/or quantum-yield fluctuations of the dye labels. However, each of these have been ruled out as likely broadening mechanisms in Alexa 488-Cy5 FRET pairs by Antonik *et al.*, for which we anticipate behavior similar to Cy3-Cy5 (166). Most relevantly, the aforementioned broadening sources should affect both the undocked and docked states. However, linewidth broadening is evident for only the *undocked* state, with nearly shot-noise limited predictions for the *docked* state, and therefore is not likely originating from such photophysical effects. Finally, as will be discussed later, only the *undocked* peak widths and E_{FRET} values exhibit sensitivity to *monovalent* (Na^+) cation concentration.

A simple model for the observed broadening is that the FRET values reflect a distribution over conformational fluctuations in the nominally two-state picture of docked and undocked states. Docking of the tetraloop with the receptor confines the donor and acceptor to a smaller range of distances and directions, thereby greatly reducing variation in the observed FRET efficiencies. Thus, one would expect negligible dynamical contributions to the docked state widths, in agreement with the nearly shot-noise limited values observed experimentally. Conversely, since flexible motion of the tetraloop linker allows the undocked state to sample a variety of conformations and distances, one would anticipate the *undocked* E_{FRET} peak to be dynamically broadened beyond the shot-noise limit. Most importantly, this model provides a physical basis for understanding [Na$^+$]-dependent shifts in the mean E_{FRET} values for the undocked state due to ionic strength and Debye shielding effects, which will be addressed later.

For the moment, however, we focus on [Mg^{2+}]-dependent trends in the tetraloop–receptor docking. Systematic evolution of the docked and undocked populations under standard HEPES buffer conditions (125 mM Na$^+$, pH 7.5) is demonstrated in Figure 3.3 for sample concentrations along the [Mg^{2+}] titration by the growth of the high E_{FRET} peak and concomitant reduction of the low E_{FRET} peak, where the E_{FRET} histograms contain the same quantity of total positive E_{FRET} events. Since this occurs over a very small change in total ionic strength, we treat the GAAA tetraloop–receptor interaction as a two-state system, i.e., with distinguishable undocked and docked RNA subpopulations. There are no systematic shifts in the E_{FRET} peak positions or widths as function of [Mg^{2+}] (Figure 3.3). Therefore, to determine the fraction of burst events in these subpopulations, the complete set of E_{FRET} histograms has been fit simultaneously to a sum of three Gaussians, with peak widths and centers as adjustable but common parameters for all [Mg^{2+}] conditions. This combined fit procedure permits more accurate characterization of docked vs. undocked subpopulations, particularly for the small undocked and docked fractions obtained at the extreme values of [Mg^{2+}]. The fractional populations of undocked and docked species are then readily determined by integration over the respective peaks (118,161).
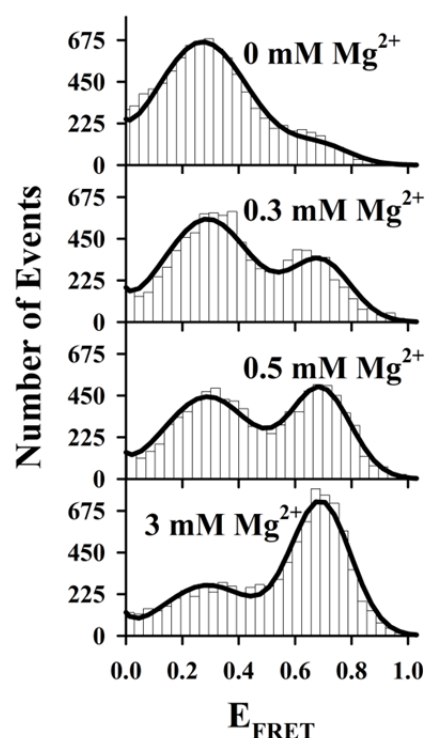
Figure 3.3 E_{FRET} population histograms as a function of [Mg^{2+}] with Gaussian fits superimposed. The tetraloop–receptor interaction is promoted by Mg^{2+}, as evidenced by the shift in the relative populations from undocked (low E_{FRET}) to docked (high E_{FRET}) states.

The fractional docked population under freely diffusing conditions, denoted by f_{free} = $N_{docked}/(N_{docked} + N_{undocked})$, where N_{docked} and $N_{undocked}$ are determined from the integrated Gaussian area of the docked and undocked peaks (149), is plotted versus [Mg^{2+}] as circles in Figure 3.4. The data rise smoothly between 0 mM and 1 mM Mg^{2+} i) from a small but nonzero intercept and ii) reach an asymptotic value less than unity under saturating Mg^{2+} concentrations. Based on a simple two-state kinetic model summarized in Figure 3.5 *A*, this Mg^{2+}-dependent docked fraction, f_{free}, can be reasonably well fit to a standard Hill binding equation (149), with the modification that only a *fixed fraction f_{max}* of the molecules are able to dock,

$$ f_{free} = \frac{f_{max}[M]^n}{[M]^n + K_D{}^n} , $$

3.7
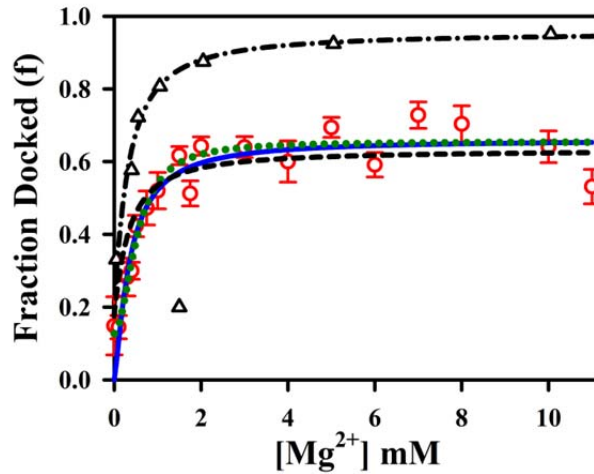
Figure 3.4   Comparison of Mg^{2+}-dependent fractional docked population for freely diffusing (*black circles*) and immobilized tetraloop–receptor constructs (*gray triangles* and *dash-dotted line*). $f_{immobilized}$ is calculated from the kinetic rate constants observed in tethered actively docking/undocking constructs, where a non-docking population (32 ± 1%) was previously observed. f_{free} is fit to Eq. 3.7 (*solid, gray line*), where n = 1.3 ± 0.3, K_D = 0.36 ± 0.6 mM, f_{max} = 0.66 ± 0.03. Linear scaling of $f_{immobilized}$ to f_{free} (Eq. 3.9) yields 66 ± 2% constructs are actively docking under freely diffusing conditions (*dotted*, *black line*). f_{free} is also fit to Eq. 3.10 (*solid, black line*) derived from the model in Figure 3.5 *C*, which allows for a nonzero docked fraction at 0 mM Mg^{2+} due to 125 mM Na$^+$.

where [M] is the metal ion concentration. This analysis yields a Mg^{2+} dissociation constant K_D = 0.36 ± 0.6 mM, a Hill coefficient of n = 1.3 ± 0.3 (see Figure 3.4, *solid blue line*) consistent with noncooperative behavior, and an asymptotic docking fraction of f_{max} = 0.66 ± 0.03. This model is inconsistent with the finite docked fraction (15 ± 8%) experimentally observed at [Mg^{2+}] = 0 mM in Figure 3.4, which requires a more complex mechanism (discussed later) involving both divalent Mg^{2+}- and monovalent Na$^+$-mediated docking pathways. Of more immediate relevance, however, the data indicate (1 − f_{max}) = 0.34 ± 0.03, which would be consistent with 34% of the RNA constructs unable to dock and undock. This non-unity asymptote could in principle suggest that docking proceeds via a Mg^{2+}-bound undocked intermediate, for which the asymptote would reflect the equilibrium of the Mg^{2+}-activated intermediate with the docked state. However, previous studies of surface-*immobilized* RNA constructs also identified such a subpopulation of non-docking species (106), which is shown below to be quantitatively consistent with the freely diffusing RNA data.

A $U + n\,M \underset{K_D}{\rightleftharpoons} D(M)_n$

B $U + n\,Mg^{2+} \underset{K_{Mg}}{\rightleftharpoons} U(Mg)_n$ *Low* E_{FRET}

$k_1 \downarrow\uparrow k_{-1}$ $k_2 \downarrow\uparrow k_{-2}$

   $D + n\,Mg^{2+} \underset{K'_{Mg}}{\rightleftharpoons} D(Mg)_n$ *High* E_{FRET}

C   *Low* E_{FRET} *High* E_{FRET}

   $U + n\,Mg^+ \underset{K_D^{Mg2+}}{\rightleftharpoons} DMg_n$

   $+$

   $m\,Na^+$

   $\downarrow\uparrow K_D^{Na+}$

   DNa_m *High* E_{FRET}
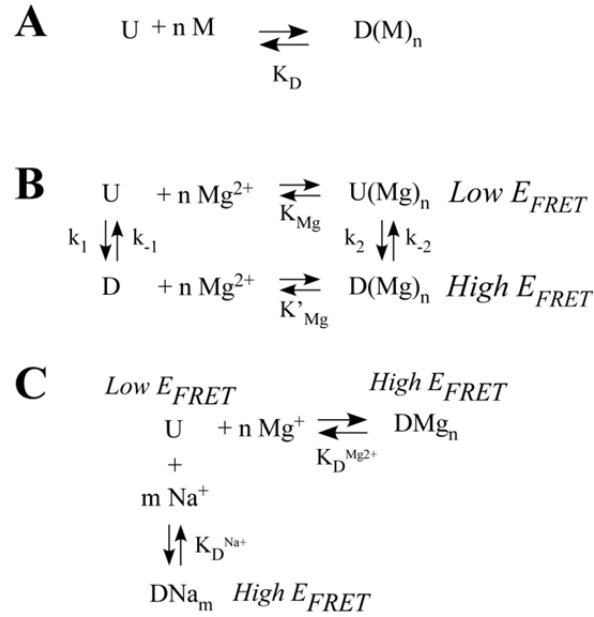
Figure 3.5   (*A*) Nominal two-state picture for cooperative binding of metal ions (M) to an undocked state (U), enabling progression to a docked state ($D(M)_n$) with metal ion dissociation constant, K_D. (*B*) Mechanism to describe docking of the GAAA tetraloop and receptor with and without Mg^{2+}, where K_{Mg} and K'_{Mg} are Mg^{2+}-dissociation constants and the rate constants reflect docking and undocking resolved by FRET. (*C*) Simplified parallel model to describe [Na^+] and [Mg^{2+}]-dependence for the observed fraction of docked molecules with Mg^{2+} and Na^+ dissociation constants.

3.4.2   Freely Diffusing versus Immobilized RNA: Kinetics and Heterogeneity

Single-molecule studies of *immobilized* tetraloop–receptor constructs indicate that the observed docking (k_{dock}) and undocking (k_{undock}) rate constants *increase* by 12-fold and *decrease* by 3-fold, respectively, over a 0–10 mM [Mg^{2+}] range (106), which motivated a more complex kinetic description than a traditional cooperative-binding model (Figure 3.5 *A*). Of particular relevance to the burst fluorescence results is that these previous tethered studies also provided evidence for kinetic heterogeneity; specifically, a significant fraction (32 ± 1%) of RNA molecules remained permanently undocked on the time scale of photobleaching even under saturating [Mg^{2+}] conditions. An important question to consider is whether the observed heterogeneity in kinetic activity is influenced by proximity of immobilized RNA molecules to the protein-passivated glass tethering surface. With burst fluorescence detection, we can now ad-

dress this question directly by comparing results for freely diffusing RNA with previous data for the immobilized constructs.

To make a quantitative comparison, we invoke the four-state kinetic model described by Kim *et al.* (148) that includes Mg^{2+}-dependent and -independent pathways for the docking of the tetraloop and receptor (Figure 3.5 *B*). According to this model, Mg^{2+} exchange occurs much faster than subsequent RNA folding/unfolding; thus Mg^{2+}-bound and -free states rapidly equilibrate with dissociation constants K_{Mg} and K'_{Mg} for undocked and docked states, respectively (Figure 3.5 *B*). Furthermore, the Mg^{2+}-bound and -free forms of the undocked and docked states are experimentally indistinguishable by FRET. As a result, the experimental rate constant reflect the combination of k_1 and k_2 for docking and k_{-1} and k_{-2} for undocking, which represent the Mg^{2+}-independent (dependent) pathways, respectively (106,148). Under these conditions, the observed docking and undocking rate constants (Figure 3.5 *B*) are well described by n, k_1, k_2, k_{-1}, k_{-2}, K_{Mg}, and K'_{Mg}, which have already been determined by least squares analysis of the [Mg^{2+}] dependence of k_{dock} and k_{undock} for the tethered RNA constructs (106).

Due to rapid diffusion, burst methods yield equilibrium population distributions rather than explicit docking and undocking events. An appropriate metric for predicting fractional docked populations from the immobilized studies is therefore

$$f_{\text{immobilize d}} = \frac{[D]}{[D]+[U]} = \frac{k_{\text{dock}}}{k_{\text{dock}} + k_{\text{undock}}},$$ 3.8

where k_{dock} and k_{undock} are experimentally determined and fit as a function of [Mg^{2+}] using the four-state kinetic model (Figure 3.5 *B*), and [U] and [D] are the undocked and docked populations. We can now use data and fits to *calculate* the equilibrium docking fraction as a continuous function of [Mg^{2+}] under burst conditions. The results are summarized in Figure 3.4, where open triangles represent the docked fraction calculated from rate constants for immobilized species, while the dash-dotted line represents the corresponding predictions from least-squares fits of k_{dock} and k_{undock}. At high [Mg^{2+}], the freely diffusing data saturate at a docking fraction less than unity, in contrast to the kinetic predictions based on tethered but actively folding/unfolding molecules. More subtly, both the burst data and tethered predictions indicate a *finite* docking fraction at low [Mg^{2+}], again supporting a Mg^{2+}-*independent* pathway for folding.

In tethered studies, the rate data are obtained from RNA constructs actively docking and un-docking on the millisecond-to-multiple second time scale. Though this correctly reflects the majority RNA population, it does not include the 32% of constructs that exhibit no folding prior to photobleaching. In the burst studies, *all* RNAs diffusing through the confocal volume are sampled, including both actively docking/undocking species ([*D*], [*U*]) as well as RNAs not able to dock on the time scale of the experiment ([*ND*]). Therefore, the freely diffusing and tethered curves in Figure 3.4 should be proportional to each other, with a constant scale factor.

$$\alpha = \frac{[U]+[D]}{[U]+[D]+[ND]},$$ 3.9

i.e., $f_{immobilized} = \alpha f_{free}$. We can make this quantitative by least squares fitting the burst data with a linearly scaled version of the kinetic predictions from the tethered samples, as shown in Figure 3.4 (*black, dashed line*) yielding α = 66 ± 2%. This implies a *non-docking* fraction of 34 ± 2% in the freely diffusing RNA, in agreement with the 32 ± 1% value reported previously for immobilized-RNA constructs. In particular, this confirms that the docking kinetic heterogeneity previously observed is an intrinsic property of the RNA construct and not an artifact of RNA surface immobilization.

3.4.3 Na⁺-Induced Docking of the Tetraloop and Receptor in Freely Diffusing RNA

The finite intercepts in Figure 3.4 indicate the presence of both [Mg^{2+}]-dependent and -independent pathways for tetraloop–receptor docking. As a likely source of this [Mg^{2+}]-independent channel, we note that all folding experiments addressed thus far occur in solutions with 125 mM Na⁺ (100 mM NaCl plus 25 mM Na⁺ from the 50 mM hemisodium HEPES buffer). Although much weaker than Mg^{2+}, Na⁺ has been known to play a role in RNA folding through electrostatic screening and specific binding (6,12,40,41,80,136,138). For example, Na⁺-induced folding of the *Tetrahymena* group I intron requires ~1,700-fold higher concentrations of Na⁺ than Mg^{2+} (K_D = 460 ± 6 vs. 0.270 ± 0.001 mM) (167). Furthermore, at 2,500-fold higher concentrations than required for Mg^{2+}, Na⁺ alone can successfully stabilize the folded conformation of the 16S ribosomal RNA junction (148). In this section, we demonstrate

that in the absence of *both* Na+ and Mg2+, the docking process is effectively blocked by exploiting burst

fluorescence time traces for freely diffusing RNA constructs as a function of [Na+].

The resulting E_{FRET} distributions from freely diffusing tetraloop–receptor constructs at 0 mM

Mg2+ are shown in Figure 3.6 for low ([Na+] = 25 mM) and high ([Na+] = 1.0 M) monovalent ion concen-

trations. At the lower limit of 25 mM Na+ arising from the HEPES buffer, the E_{FRET} distributions are dom-

inated by the undocked conformation (Figure 3.6 *A*), with the docked population visible only very weak-

ly. At [Na+] = 1.0 M, on the other hand, the docked E_{FRET} peak appears (Figure 3.6 *B*) quite prominently,

confirming that Na+ can induce docking of the tetraloop and receptor in the absence of Mg2+. However,

stabilization of the docked state with Na+ is much weaker than for divalent Mg2+, as indicated by the ∼

1000-fold higher range of concentrations required. At the very lowest Na+ concentrations, dissociation
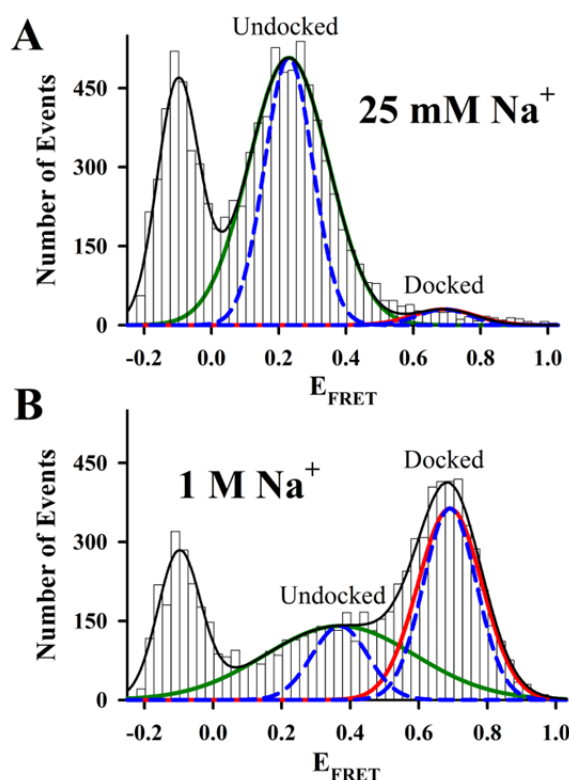
Figure 3.6 E_{FRET} distributions and Gaussian fits (*black*) showing, donor-only (*leftmost peak*), undocked (*green*) and docked (*red*) populations at (*A*) 25 mM Na+ and (*B*) 1.0 M Na+, with shot-noise predictions for each peak (*blue dashed lines*). Note that the undocked peak shifts to higher center E_{FRET} value and broadens with increasing [Na+] (see text for details).

of the RNA construct is a possible concern, which would be signaled by an increase in the donor-only population. However, we have tested for this and find no systematic difference in the fractional donor-only population over the range of Mg^{2+} or Na^+ concentrations sampled.

Relative populations of the docked and undocked states are determined by fitting the Na^+-dependent E_{FRET} histograms to Gaussian distributions and calculating fractional populations from integrated areas. Total Gaussian fits (*black lines*) are shown in Figure 3.6 *A* and *B*, indicating the *undocked* (*green*) and *docked* (*red*) components. Shot-noise limited width contributions (*blue*, *dashed lines*) are also shown, again indicating a greater sensitivity of peak widths in the undocked vs. docked species. Similar to the Mg^{2+} results, the *docked* peak center and width are unaffected by $[Na^+]$. However in stark contrast with the Mg^{2+} titration, the *undocked* peak both *shifts* and *broadens* (Figure 3.6). More quantitatively, the docked state center E_{FRET} obtained from a global fit of the Na^+-titration is 0.691 ± 0.001, in perfect agreement with the E_{FRET} value of 0.687 ± 0.005 obtained from the Mg^{2+}-dependent study. On the other hand, for similar changes in $[Na^+]$, the *undocked* peak shifts by nearly ≈ 0.200 FRET units and exhibits a $\approx$ 2-fold increase in peak width.

The fraction of docked tetraloop–receptor constructs (f_{free}) is plotted as a function of $[Na^+]$ in Figure 3.7 *A*. The data can be well fit by a cooperative-binding model (Eq. 3.7), yielding an asymptotic saturation $f_{max} = 0.55 \pm 0.05$, a Hill coefficient 1.3 ± 0.3, and an apparent dissociation constant $K_D = 180 \pm 30$ mM. The finite docking fraction at the lowest Na^+ concentration in Figure 3.6 and Figure 3.7 *A* is now seen to be completely consistent with residual 25 mM Na^+ in the HEPES buffer. However, the K_D with respect to Na^+ is ~500-fold larger than that observed for Mg^{2+}, indicating a much lower efficiency for monovalent Na^+-mediated docking and comparing quite well with ensemble studies, $K_D = 220 \pm 9$ mM (124). Interestingly, f_{max} is nearly equal for both Na^+- and Mg^{2+}-dependent studies, 0.55 ± 0.05 and 0.66 ± 0.03, respectively. This is again consistent with a constant fraction of RNA constructs unable to dock, but now established over an even broader range of both salt concentration and type of cation.

As a consistency check, the Na^+ titration data can predict the non-zero intercept value previously noted for $[Mg^{2+}] = 0$ mM with $[Na^+] = 125$ mM. For these purposes, the divalent and monovalent

Figure 3.7   (*A*) Least squares fits of fractional docked population ($N_{docked}/(N_{docked} + N_{undocked})$) versus [Na⁺] to Eq. 3.7, resulting in f_{max} = 0.55 ± 0.05, a Hill coefficient 1.3 ± 0.3, and K_D = 180 ± 30 mM. The asymptotic value (f_{max}) is consistent with Mg²⁺ studies in Figure 3.5, suggesting a ≈ 32–34% non-docking RNA subpopulation. (*B*) f_{free} as a function of [Mg²⁺] at low [Na⁺] (25 mM) with a fit to Eq. 3.10 that also allows for a [Na⁺] docking pathway (Figure 3.5 *C*), yielding f_{max} = 0.55 ± 0.04, n = 8 ± 2, $K_D^{Mg^{2+}}$ = 1.06 ± 0.03 mM, and demonstrating high cooperativity with respect to Mg²⁺ observed under minimal Na⁺ conditions.

docking processes can be assumed to be independent, as denoted by the kinetic scheme in Figure 3.5 *C*.

At equilibrium, the fraction of docked molecules for independent Na⁺ and Mg²⁺-induced folding can be

easily shown to be

$$f_{free} = \frac{f_{max}\left(\dfrac{[Mg^{2+}]^n}{(K_D^{Mg})^n} + \dfrac{[Na^+]^m}{(K_D^{Na})^m}\right)}{\left(\dfrac{[Mg^{2+}]^n}{(K_D^{Mg})^n} + \dfrac{[Mg^{2+}]^n}{(K_D^{Na})^n} + 1\right)} , \qquad 3.10$$

Based on least squares fitting of the Na⁺ titration, the predicted intercept for the fraction docked Mg²⁺

titration is f_{free} = 0.2 ± 0.1 at 0 mM Mg²⁺, 125 mM Na⁺, which is in excellent agreement with the experi-

mental values $f_{immobilized}$ = 0.19 ± 0.07 from Eqs. 8 and 9 and the f_{free} = 0.15 ± 0.08. Consequently, the fold-

ing in the absence of Mg²⁺ observed in immobilized and freely diffusing studies can be quantitatively

attributed to the presence of 125 mM Na⁺. Furthermore, this model allows us to describe the nonzero

intercept of the [Mg²⁺]-dependent fraction docked (Figure 3.4), whereas the simple cooperative binding

model (Eq. 3.7) could not. Using Eq. 3.10 to fit f_{free} yields a more precise and physical description of the

observed [Mg^{2+}]-dependent docking of the tetraloop–receptor with $K_{\mathrm{D}}^{\mathrm{Mg}^{2+}}$ = 0.46 ± 0.04 and n = 2.0 ± 0.4 (Figure 3.4, *green dotted line fit to the red open circles*).

Although Eq. 3.10 and the model in Figure 3.5 *C* assume independent docking pathways for Na$^+$ and Mg^{2+}, there is evidence that this is not true at low Na$^+$ concentrations. For example, Figure 3.7 *B* displays a titration curve as a function of Mg^{2+} for low Na$^+$ concentrations (25 mM). As expected from Eq. 3.10, the intercept nicely matches the value for 25 mM Na$^+$ and 0 mM Mg^{2+} in Figure 3.7 *A*. Furthermore, at high Mg^{2+} levels the docked fraction rises to the typical asymptotic value ($f_{\mathrm{max}} \approx 0.6$) seen in other titrations, which again originates from the heterogeneous presence of non-docking RNA constructs. Quite different, however, is the dramatically sigmoidal shape of the titration curve, which implies a much higher level of Mg^{2+} *cooperativity* in the absence of Na$^+$. More quantitatively, the data in Figure 3.7 *B* can be fit to the Hill curve of Eq. 3.10, yielding $K_{\mathrm{D}}^{\mathrm{Mg}^{2+}}$ = 1.06 ± 0.03 mM and a Hill coefficient of n = 8 ± 2. By way of comparison, when the data at higher Na$^+$ 125 mM are fit to the same model (Figure 3.4, *green dotted line fit to red open circles*), these values decrease (i.e., become *less* cooperative) to $K_{\mathrm{D}}^{\mathrm{Mg}^{2+}}$ = 0.46 ± 0.04 and n = 2.0 ± 0.4. Alternatively stated, the trends in affinities and Hill coefficients indicate that monovalent Na$^+$ greatly diminishes or eliminates docking cooperativity with respect to Mg^{2+}. Examples of high Mg^{2+} cooperativity at very low Na$^+$ have been noted in previous investigations of the *Tetrahymena* ribozyme and its P4-P6 domain (167,168). However, the remarkable feature in the present study is that such high cooperativities can be exhibited in much simpler RNA constructs docking via a *single* tertiary interaction. This would also suggest that electrostatic shielding of the phosphate groups by monovalent ions is necessary to facilitate efficient, *non-cooperative* Mg^{2+}-induced docking as discussed in the following two sections.

3.4.4   Electrostatic Relaxation of Tetraloop–Receptor RNA

The importance of electrostatic screening and relaxation effects in the tetraloop-receptor system is evidenced by a systematic *increase* in E_{FRET} (and thus *decrease* in the fluorophore separation) in

the undocked state with cation concentration (Figure 3.6). To investigate this further, we plot experi-

mental shifts in the undocked ⟨E_{FRET}⟩ as a function of sodium ion concentration (Figure 3.8 *A*). Upon ad-

dition of Na⁺, the undocked ⟨E_{FRET}⟩ increases significantly, consistent with partial relaxation of the RNA

structure and a statistically closer approach of donor and acceptor. Specifically, E_{FRET} center changes by

≈ 60%, corresponding to ≈ 13% (≈ 7 Å) decrease in Cy3-Cy5 separation.   A Hill-type expression com-

monly employed in ensemble FRET studies of two-state systems (124) proves convenient to character-

ize the [cation]-dependent relaxation of the tetraloop–receptor RNA,

Figure 3.8   Evidence for a cation-induced increase in electrostatic compaction and conformational sampling of the undocked state tetraloop–receptor construct. (*A*) and (*B*) Systematic shift in mean E_{FRET} of the undocked peak with increasing [Na⁺] and [Mg²⁺], respectively, fit by a Hill-type model (Eq. 3.11) with $E_{FRET}^0(Na^+, Mg^{2+})$ = 0.227 ± 0.004, 0.227 ± 0.003; $E_{FRET}^0(Na^+, Mg^{2+})$ = 0.15 ± 0.02, 0.07 ± 0.02, n(Na⁺, Mg²⁺) = 2.1 ± 0.4, 2.6 ± 0.8 and K_D(Na⁺, Mg²⁺) = 180 ± 20 mM, 0.9 ± 0.2 mM. (*C*) and (*D*) Systematic shifts in undocked E_{FRET} peak widths as function of [Na⁺] and [Mg²⁺], respectively, yielding K_D(Na⁺, Mg²⁺) = 0.23 M ± 0.02, 1.2 ± 0.4 mM; n(Na⁺, Mg²⁺) = 3.6 ± 0.8, 2.7 ± 1.2; and $\Delta\sigma$(Na⁺, Mg²⁺) = 0.10 ± 0.01, 0.07 ± 0.03, respectively.

$$\langle E_{\text{FRET}}\rangle = E_{\text{FRET}}^{0} + \frac{\Delta E_{\text{FRET}}[M]^{n}}{[M]^{n} + K_{\text{D}}^{\ n}}\ , \qquad\qquad 3.11$$

which for the [Na$^+$]-dependent data yields K_{D} = 180 ± 20 mM, n = 2.1 ± 0.4, E_{FRET}^{0} = 0.227 ± 0.004, and

ΔE_{FRET} = 0.15 ± 0.02. It is worth noting that the K_{D} from fits to Na$^+$-mediated FRET peak shifts is identical within uncertainty (K_{D} = 180 ± 30 mM, Figure 3.7 *A*) to that observed for the undocking/docking transition itself, in support of a common electrostatic origin for both phenomena. Although much smaller in magnitude, a similar Hill-type dependence is observed as a function of Mg^{2+} for low 25 mM Na$^+$ conditions, as plotted in Figure 3.8 *B*. Once again, the overall highly cooperative shape and affinities (K_{D} = 0.9 ± 0.2 mM) are nearly identical to the corresponding values (K_{D} = 1.06 ± 0.03 mM, Figure 3.7 *B*) obtained for Mg^{2+}-dependent docking at low [Na$^+$]. Interestingly, we do not see appreciable mean E_{FRET} shifts as a function of 0–11 mM Mg^{2+} under "normal" HEPES buffer, i.e., 125 mM Na$^+$ (see Figure 3.2 *B*). However, from Figure 3.8 *A* one can easily see that Na$^+$-induced FRET shifts for the undocked peak under these conditions dominate any Mg^{2+}-induced shifts.

It is worth considering whether these [cation]-dependent peak shifts and broadenings could be induced by changes in donor versus acceptor quantum yields. However, these effects are minor based on both the *magnitude* and *selectivity* with which the undocked vs. docked peak shift. In the case of Na$^+$, for example, to match $\Delta E_{\text{FRET}} \approx 0.14$ for the undocked peak would require $Q_{\text{A}}/Q_{\text{D}}$ to be reduced by > 50%, which is observed neither in the donor or acceptor signal intensities nor analysis of the fluorescence lifetimes. Furthermore, such a reduction would predict a corresponding shift of $\Delta E_{\text{FRET}} \approx 0.08$ in the docked peak, i.e., which is nearly an order of magnitude larger than our experimental uncertainties and not observed. Thus, shifts in the undocked $\langle E_{\text{FRET}}\rangle$ values must correspond to statistically significant structural changes in average donor-acceptor distances for the RNA construct.

Further confirmation of a cation-mediated structural flexibility in the undocked RNA constructs is also provided by the *broadening* of the undocked E_{FRET} peak with increased [Na$^+$]. Specifically, the *undocked* Gaussian peak width, σ, *increases* from (1.7 ± 0.1)- to (2.7 ± 0.3)-fold excess of shot noise from 25 mM to 1 M Na$^+$ (Figure 3.8 *C*), i.e., greatly exceeding other potential broadening contributions (e.g., tri-

plet formation and photobleaching, as described above) under these experimental conditions. By way of contrast, the *docked* peak maintains a *fixed* width near the shot-noise limit (1.2 ± 0.1 fold) over the same range of [Na$^+$]. More quantitatively, the undocked widths (σ) are fit to a Hill-type model (Eq. 3.11), yielding K_D = 230 ± 20 mM, n = 3.6 ± 0.8 and $\Delta\sigma$ = 0.10 ± 0.01 (in Figure 3.8 *C*). Once again, this behavior occurs over a comparable range of K_D values observed for both i) docking fraction (K_D = 180 ± 30 mM) and ii) undocked peak shifts (K_D = 180 ± 20 mM), consistent with a common origin of electrostatic screening.

A similar analysis of the undocked peak linewidths as a function of [Mg^{2+}] under low 25 mM Na$^+$ conditions is plotted in Figure 3.8 *D*. The data indicate a systematic increase in line widths, with an increased K_D (1.2 ± 0.4 mM) and a large Hill coefficient (n = 2.7 ± 1.2) that again illustrate high levels of cooperativity with respect to Mg^{2+}. This behavior agrees quantitatively with what was seen previously under low Na$^+$ conditions for Mg^{2+} dependent i) docking fraction (K_D = 1.06 ± 0.03 mM, n = 8 ± 2, Figure 3.7 *B*) and ii) undocked peak shifts (K_D = 0.9 ± 0.2 mM, n = 2.6 ± 0.8, Figure 3.8 *B*), underscoring the importance of electrostatic screening in conformational dynamics for the undocked constructs. It is worth reiterating that the undocked peak contains contributions from both actively docking/undocking RNA and non-docking species, as described above, but at high [Na$^+$] the *non-docking* species is the dominant component. Therefore, electrostatic relaxation of the *undocked* structure must be responsible for the peak broadening and shifts with increased cation concentration, rather than any fast dynamics due to rapid tetraloop–receptor docking/undocking.

We interpret the origin of the peak shifting and broadening for Na$^+$ as arising from structural relaxation in the undocked RNA constructs, induced by electrostatic screening in an aqueous electrolytic medium. Poisson Boltzmann or counterion condensation theories may be used to describe the role of monovalent vs. divalent ions in structural relaxation of RNA (169). However, due to the polyanionic nature of RNA, a more sophisticated nonlinear Poisson-Boltzmann analysis is sometimes necessary for quantitative determination of the electrostatic potential throughout the entire RNA structure (40,84,85,170-174). However, such numerically intensive schemes (84,85,175) often fail to quantita-

tively model the role of higher order valences (e.g. Mg^{2+}) in RNA folding (176). Therefore, to provide a qualitative physical interpretation for Mg^{2+} and Na^+-induced conformational changes in the tetraloop–receptor construct, we employ a simpler Debye shielding model of point charges screened as a function of ionic strength. Such a Debye treatment slightly underestimates the actual electrostatic screening, though matching all Poisson-Boltzmann trends as a function of ionic strength (177). The relevant parameter in such a model is the 1/e Debye length, $\lambda_D = \sqrt{\dfrac{\varepsilon_0 \varepsilon_r k_B T}{2 N_A q_e^2 I}}$, where ε_0 is the permittivity of free space (8.85 x 10^{-12} C^2/Nm^2), ε_r the dielectric constant of water (80.4), q_e the elementary charge (1.6 x 10^{-19} C), k_B Boltzmann's constant, T absolute temperature, N_A Avogadro's number, and $I = \dfrac{1}{2}\sum_i C_i Z_i^2$ represents ionic strength for a C_i molar concentration of each ion with charge Z_i. To illustrate ionic strength effects on charge screening, λ_D is plotted in Figure 3.9 for relevant Na^+ and Mg^{2+} concentrations in 50 mM hemisodium HEPES. The Debye length decreases dramatically from $\lambda_D \sim 20$ Å to $\lambda_D < 5$ Å (Figure 3.9 *A*) over a comparable range ([Na^+] ≈ K_D ≈ 180 mM) for which the undocked structure relaxes and broadens (Figure 3.8 *A* and *C*). By way of contrast, the effect of [Mg^{2+}] on Debye length is completely masked in HEPES buffer with the standard 100 mM NaCl (λ_D ≈ 8 Å over the entire [Mg^{2+}] range), but becomes much more relevant when NaCl is excluded (Figure 3.9 *B*), consistent with experimental observations in Figure 3.7 *B* and Figure 3.8 *B* and *D*.

These monovalent ion effects can be rationalized by efficient screening of phosphate groups in the RNA backbone, leading to reduction in the persistence length of single-stranded regions. This reduction permits greater flexibility and access to more compact RNA conformations, thereby shifting the average E_{FRET} peak for the undocked construct. Likewise, broadening of the E_{FRET} peak can originate from reduced repulsion of the tetraloop and linker from the receptor domain and tether regions of the construct, resulting in a wider distribution of accessible conformations. The observed E_{FRET} peak shifts and broadening effects both occur over a range of [Na^+] ∼ 0 to 0.5 M, over which the characteristic Debye length decreases by ≈ 5-fold (Figure 3.9 *A*).
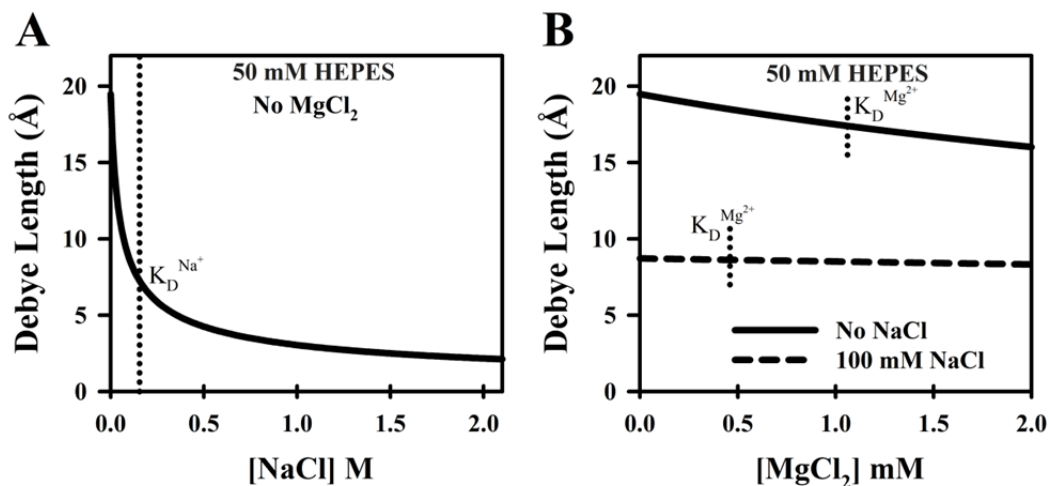
Figure 3.9   Calculated Debye shielding lengths in the presence of 50 mM hemisodium HEPES buffer with the addition of (*A*) [NaCl] in the absence MgCl$_2$ and (*B*) [MgCl$_2$] without and with 100 mM NaCl. Also shown (*dotted vertical lines*) are the observed K_D values for (*A*) Na$^+$- and (*B*) Mg^{2+}-facilitated docking.

Electrostatic shielding can also be used to interpret the different influences of Mg^{2+} and Na$^+$ on tetraloop–receptor docking. In the P4–P6 domain of the *Tetrahymena* ribozyme, both the tetraloop–receptor and adenosine-rich bulge tertiary interactions pack the opposing helical phosphate groups within 8-9 Å (62), requiring tetraloop and receptor helices to be effectively screened for stabilization of the folded structure. The experimentally observed K_D for Na$^+$ is consistent with the λ_D values necessary to allow tetraloop and receptor proximity (see Figure 3.9 *A*). On the other hand, much lower [Mg^{2+}] concentrations are required to achieve the same docked state, with only minor impact on ionic strength and λ_D (Figure 3.9 *B*). This is consistent with the fact that divalent Mg^{2+}, as a consequence of compact size and high charge density, can intimately localize and coordinate along the RNA at regions of high negative electrostatic potential, whereas monovalent Na$^+$ must rely on the less-efficient mechanism of bulk electrostatic screening (12,172,174,176,178). Similarly, the striking growth in cooperativity for Mg^{2+}-promoted docking can be attributed to the much larger Debye shielding length without NaCl (Figure 3.9 *B*). With insufficient monovalent ionic strength to screen phosphate repulsions, *multiple* Mg^{2+} ions must be recruited for relaxing the RNA structure prior to the more local task of promoting docking, therefore

resulting in higher cooperativity and increased K_D values. This analysis offers a simple physical model for monovalent-induced structural relaxation of the undocked RNA consistent with the experimental data, though more rigorous treatment of the electrostatic environment will be necessary to make such comparisons quantitative.

3.4.5   Na⁺ and Mg²⁺ Synergistically Promote Tetraloop–Receptor Docking

Closer comparison of the cation-dependent population distributions presented in this work also permits one to explore possible *competition* or *synergism* between Na⁺ and Mg²⁺-induced folding of the tetraloop–receptor constructs. Bokinsky *et al*. showed that in the presence of 500 mM Na⁺ the docking transition of the hairpin ribozyme saturates at much *higher* [Mg²⁺] than in the absence of Na⁺, suggesting that Mg²⁺ and Na⁺ *compete* for interaction with the hairpin RNA (8). For the tetraloop and receptor construct, however, we find precisely the opposite scenario; $K_D^{\mathrm{Mg}^{2+}}$ *decreases* from 1.06 ± 0.03 mM at 25 mM Na⁺ to 0.46 ± 0.04 mM at 125 mM Na⁺, while at the same time cooperativity with respect to Mg²⁺ is virtually eliminated (Figure 3.7 *B* and Figure 3.4). This suggests that Na⁺ *enhances* the ability of Mg²⁺ to promote docking, which would imply a more complex description for cation induced folding than the independent pathways depicted in Figure 3.5 *C*. Synergism between Na⁺ and Mg²⁺ in the promotion of the tetraloop–receptor tertiary interaction is best demonstrated by comparing the RNA docking fraction (f_{free}) for a series of Na⁺ and Mg²⁺ concentration pairs. The left-most two bars in Figure 3.10 present data for limits of i) Na⁺-dominated (125 mM Na⁺, 0.0 mM Mg²⁺) and ii) Mg²⁺ dominated (25 mM Na⁺, 0.5 mM Mg²⁺) docking, respectively, with the third bar based simply on the additive Mg²⁺ and Na⁺ promoted docking predictions (i + ii). In fact, experimental results (fourth bar) exceed this by nearly 2-fold, a direct indication that Na⁺ and Mg²⁺ interact with *positive* synergy in promoting the docking event.

Electrostatic shielding again provides a simple physical model for positive synergy in the folding of the RNA constructs. At low ionic strengths, additional Mg²⁺ ions are required to sufficiently relax the RNA prior to docking, which translates into a large Hill coefficient and increased affinity $K_D^{\mathrm{Mg}^{2+}}$. In the presence of only 100 mM NaCl, the Debye length decreases dramatically, cooperativity with respect to
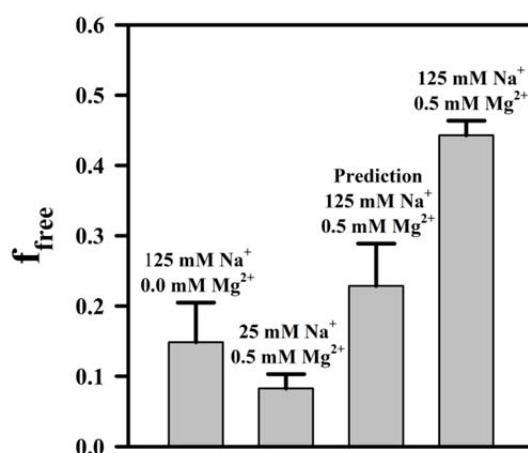
Figure 3.10   Evidence for positive Na+ and Mg2+-synergy in promoting tetraloop–receptor docking; f_{free} for combined Mg2+ and Na+ (*right most bar*) is significantly greater than the prediction (*third bar*) based on a simple additive model of individual Na+ and Mg2+ results (*left two bars*).

Mg2+ is eliminated (as in Figure 3.7 *B* vs. Figure 3.4), and $K_{\text{D}}^{\text{Mg}^{2+}}$ decreases by ≈ 2-fold because of the increased ionic strength. It would be interesting to test if this *synergy* changes into *competition* for docking promotion at even higher Na+ concentrations, as experimentally observed for folding of the hairpin ribozyme (8). Alternatively, the observed Na+-Mg2+ synergy could reflect a specific monovalent binding site that, when filled, facilitates tetraloop–receptor docking. Indeed, Basu *et al.* have identified a monovalent binding site in the tetraloop–receptor interaction of the P4–P6 domain of the Azoarcus Group I Intron using X-ray crystallography (80). Increased occupancy of this monovalent site could stabilize tetraloop–receptor binding in a way that complements Mg2+-induced folding and lowers $K_{\text{D}}^{\text{Mg}^{2+}}$. Either scenario would be interesting and underscores the importance of studying both the full ribozyme as well as individual tertiary binding interactions at the single-molecule level.

3.5   Summary and Conclusions

The [Mg^{2+}]- and [Na^+]-dependence of docking for a single GAAA tetraloop–receptor tertiary interaction has been studied under single-molecule fluorescence conditions for freely diffusing RNAs. General expressions are presented that include corrections for cross talk, collection efficiency, quantum yield, and direct excitation of the acceptor for single-laser excitation. Gaussian fits to the E_{FRET} histograms identify distinct populations corresponding to well-resolved docked versus undocked populations with E_{FRET} distributions providing a snapshot of equilibrium populations at the single-molecule level.

The fractional docked vs. undocked populations have been explored as a function of [Mg^{2+}], which exhibits a nonzero intercept, followed by a rapid increase in folding probability with concentration. The nonzero intercept at [Mg^{2+}] = 0.0 M is inconsistent with a two-state Hill analysis and is shown to arise from monovalent Na^+-mediated folding in the buffer solution. Independent Na^+ titrations indicate efficient folding of the tetraloop–receptor, but with a K_D ~500 times larger than for Mg^{2+}. This shows that both Na^+ and Mg^{2+} lead to equivalent formation of the tetraloop–receptor contact, as supported by identical E_{FRET} values in the docked state. Significant peak shifts ($\langle E_{FRET} \rangle$) and broadening beyond the shot-noise limit are observed in the undocked but not docked E_{FRET} distributions. This cation-induced peak broadening and shift is interpreted in terms of Debye shielding of the negatively charged RNA backbone, which provides significantly more flexibility to the undocked structures. Furthermore, Mg^{2+} and Na^+ behave synergistically in promoting tetraloop–receptor docking. The presence of even ~100 mM monovalent Na^+ results in a significant *decrease* in the Mg^{2+} dissociation constant, as well as eliminating cooperativity as a function of Mg^{2+}. This interdependence of monovalent and divalent ion concentrations raises issues not considered, to our knowledge, in previous kinetic models for tetraloop–receptor docking.

Finally, analysis of the burst studies reveal the existence of a 34 ± 2% subpopulation of *non-docking* RNA molecules under freely diffusing conditions, in quantitative agreement with smFRET studies of tethered RNA constructs. These results demonstrate that [Mg^{2+}]- and [Na^+]-dependent influences on folding kinetics can be studied quantitatively for isolated tetraloop–receptor RNA tertiary interac-

tions under both free and tethered conditions without surface interference. Additional kinetic investigations of tethered RNA should prove invaluable in further testing mechanisms for monovalent/divalent promotion of the tetraloop–receptor interaction.

3.6   Acknowledgments

Chapter 4   Enthalpy-Driven RNA Folding: Single-Molecule Thermodynamics of Tetraloop–Receptor Tertiary Interaction

This chapter is published and reprinted with permission from *Biochemistry* (2009) 48:2550-2558, Fiore *et al.*, © 2009 American Chemical Society. [4]

4.1   Abstract

RNA-folding thermodynamics are crucial for structure prediction, which requires characterization of both enthalpic and entropic contributions of tertiary motifs to conformational stability. We explore the temperature-dependence of RNA folding due to the ubiquitous GAAA tetraloop–receptor docking interaction, exploiting immobilized and freely diffusing single-molecule fluorescence resonance energy transfer (smFRET) methods. The equilibrium constant for intramolecular docking is obtained as function of temperature (T = 21 to 47 °C), from which a van't Hoff analysis yields the enthalpy (ΔH°) and entropy (ΔS°) of docking. Tetraloop–receptor docking is significantly exothermic and entropically unfavorable in 1 mM MgCl$_2$ and 100 mM NaCl, with excellent agreement between immobilized (ΔH° = −17.4 ±1.6 kcal/mol and ΔS° = −56.2 ± 5.4 cal/mol/K) and freely diffusing (ΔH° = −17.2 ± 1.6 kcal/mol and ΔS° = −55.9 ± 5.2 cal/mol/K) species. Kinetic heterogeneity in the tetraloop–receptor construct is unaffected over the temperature range investigated, indicating a large energy barrier for interconversion between the actively docking and nondocking subpopulations. Formation of the tetraloop–receptor interaction can account for ∼60% of the ΔH° and ΔS° of P4–P6 domain folding in the *Tetrahymena* ribozyme, suggesting that it may act as a thermodynamic clamp for the domain. Comparison of the isolated tetraloop–

---

receptor and other tertiary folding thermodynamics supports a theme that enthalpy vs. entropy-driven folding is determined by the number of hydrogen-bonding and base-stacking interactions.

4.2   Introduction

RNA folding is generally hierarchical, with tertiary structure occurring through interactions of pre-formed secondary elements (6,35,36). As a result, the kinetics and thermodynamics of tertiary interactions are crucial to understanding RNA folding and functionality as well for accurate structural predictions (6,123). Toward this end, individual folding motifs must be characterized both in isolation and in combination for a unifying thermodynamic description of RNA folding to emerge. Among the most widespread long-range RNA interactions are A-minor motifs (59,60,179), which are commonly manifested in GNRA tetraloop–receptor binding, where N is any nucleotide and R is a purine (127,179). The ubiquitous GAAA tetraloop–receptor interaction has been of particular interest because it properly forms outside of the ribozyme's framework (74,76). The GAAA tetraloop docks into an 11-nucleotide internal receptor (62), the bound and unbound structures of which have been extensively studied (50,62,73,74). There have been previous thermodynamic investigations of the tetraloop–receptor interaction in a variety of RNA folding contexts (77,92,106,124,126,127). However, none of these studies has been able to isolate the enthalpy and entropy changes associated with tetraloop–receptor tertiary docking.

Enthalpies and entropies of tertiary structure formation can be obtained from the temperature dependence of equilibrium constants and/or precision calorimetry measurements. Differential scanning calorimetry has revealed enthalpy and free energy changes for pseudoknot tertiary folding (180). However, thermodynamics for tertiary structure formation can be challenging to isolate by methods such as isothermal titration calorimetry (ITC)[5], since studies are performed as a titration of two species, for which resolution of secondary and tertiary contributions are achieved through clever experimental

[5]Abbreviations: HPLC, high-performance liquid chromatography; HEPES, *N*-(2hydroxyethyl)piperazine-*N'*-2-ethanesulfonic acid; EDTA, ethylenediaminetetraacetate; ITC, isothermal titration calorimetry; NMR, nuclear magnetic resonance.

design (125,181-184). Other ensemble methods to dissect thermodynamic contributions to folding, such as temperature gradient gel electrophoresis, have had restricted application because experimental conditions are constrained by the limited resolution of gel shifts (11). Time-resolved ensemble FRET methods have been elegantly employed to probe tertiary folding thermodynamics, though data interpretation often requires pre-conceived models of fluorophore distance distributions, which potentially mask rare subpopulations (185).

Single- molecule fluorescence resonance energy transfer (smFRET) methods have been widely applied to RNA folding (104,105). Such methods allow isolation and direct characterization of RNA tertiary-conformational dynamics with subpopulations that exhibit different rate constants and conformations readily distinguished under a wide range of experimental conditions, e.g., varying [Mg^{2+}] (3,106,107). Despite the potential for temperature-dependent smFRET measurements of tertiary RNA folding thermodynamics, there have been surprisingly few such studies, specifically only P1 helix docking into the prefolded core of the *Tetrahymena* ribozyme and four-way helical junction folding (103,108). Additionally, both of these smFRET studies have utilized surface-immobilized constructs.

In this work, we combine a microfluidic temperature-controlled stage with time-correlated single-photon counting methods in a confocal smFRET microscope to extract thermodynamic parameters for RNA folding under both freely diffusing and surface-immobilized conditions. We determine the standard state free energy ($\Delta G°$), enthalpy ($\Delta H°$), and entropy ($\Delta S°$) changes for intramolecular GAAA tetraloop–receptor docking under physiologically relevant cation conditions and identify the folding process as enthalpically driven but at a large entropic cost. We extract tertiary folding enthalpies and entropies by freely diffusing smFRET methods, which are advantageous because they require only trace amounts of sample while avoiding possible perturbative effects of surface tethering. Such freely diffusing methods have already yielded great insight into the secondary folding of RNA hairpins (186). We interpret the GAAA-tetraloop tertiary interaction in the context of previous thermodynamics studies to illuminate a possible enthalpic and entropic paradigm for RNA folding.

4.3   Experimental Procedures

4.3.1   RNA Preparation

Cy3-Cy5-labeled tetraloop–receptor constructs depicted in Figure 4.1 are prepared as previously de-
scribed (106,124). Briefly, synthetic 5′ amino-modified RNA oligomers (Dharmacon, Lafayette, CO) are
labeled with Cy3 and Cy5 N-succinimidyl esters (Amersham Biosciences, Piscataway, NJ) and HPLC puri-
fied. The sequences of the RNA oligonucleotides are 5′ Cy5-GCC GAU AUG GAC GAC ACG CCC UCA GAC
GAG UGC G 3′ and 5′ Cy3-GGC GAA AGC CAA AAA AAC GUG UCG UCC UAA GUC GGC 3′. The complete
construct (Figure 4.1) is formed by annealing the Cy3 (1 μM) and Cy5 (1.5 μM) RNA oligomers with 2
μM biotinylated DNA oligomer (5′ biotin-CGC ACT CGT CTG AG 3′, Integrated DNA Technologies, Coral-
ville, IA) by heating to 70 °C and cooling slowly to room temperature in an annealing buffer of 50 mM
HEPES, 100 mM NaCl, 100 μM EDTA, pH 7.5. The secondary structure of the Cy3 strand forms the te-
traloop with an A_7 linker connecting it to the receptor domain created by the hybridized Cy3 and Cy5
strands. Molecules can be tethered to streptavidin-coated glass surfaces with the biotinylated extension
formed by base pairing of the DNA and Cy5 strands. The micromolar stock of annealed RNA is diluted in
working buffer that differs from the annealing buffer by the addition of 1 mM $MgCl_2$.
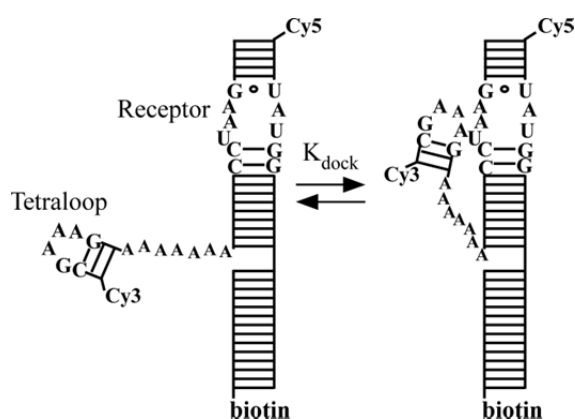
Figure 4.1   Secondary structure depiction of the tetraloop–receptor RNA docking/undocking equilibri-
um observable by Cy3-Cy5 FRET. The undocked (*left*) GAAA tetraloop docks (*right*) into the receptor via
a flexible A_7 linker, resulting in an increased FRET efficiency. Biotinylation of the RNA allows for immo-
bilization on streptavidin-coated coverglass.

4.3.2   Single-Molecule Fluorescence Measurements

Immobilized or freely diffusing molecules are observed using a time-resolved confocal microscope system (Microtime 200, PicoQuant GmbH, Berlin) based on an inverted microscope (Olympus IX-71) equipped with a time-correlating single photon-counting module and objective-scanning mode (P-721 PIFOC objective nano-positioner with x-y scanner P-733.2CL, Physik Instrumente, Karlsruhe, Germany). A 1.2 N.A. water-immersion objective (Olympus UPLSAPO 60XW) is used to focus a frequency-doubled pulsed-picosecond (40 MHz) semiconductor laser at 532 nm (PicoTA, PicoQuant GmbH) onto a glass surface for immobilized studies (1 μW at the microscope back plane, 1.7 kW/cm^2 at the focus) and 15 μm above the glass surface for freely diffusing studies (100 μW at back plane, 170 kW/cm^2 at focus). Fluorescence collected through the same objective is separated from the excitation source by a dichroic beamsplitter (Z532/635, Chroma Technology, Rockingham, VT) and focused through a 50 μm pinhole. Donor and acceptor emission are separated by a dichroic beamsplitter (645DCXR, Chroma Technology) and transmitted through bandpass filters (HQ585/70M and HQ700/75M, Chroma Technology) for detection by single-photon counting avalanche photodiodes (SPCM-AQR-14, Perkin Elmer). Data acquisition is achieved with a photon-counting module (PicoHarp 300, PicoQuant GmbH) in the time–tagged time-resolved mode, enabling the recording of every detected photon with its individual timing and detection channel information and processing with the SymPhoTime Software (PicoQuant GmbH).

Precise temperature control of the sample is achieved by implementation of a heated flow cell (FCS2, Bioptechs, Butler, PA) with internal volume of 31 μL defined by a gasket spacer (0.1 x 14 x 27 mm) between the coverglass and a resistively heated, indium-tin oxide-coated slide (Bioptechs). The objective is heated to the same temperature as the flow cell to prevent a temperature gradient with respect to the immersion optic, with the objective thermally isolated from the microscope turret by a thermal spacer (Bioptechs). The reported temperatures and standard deviations are from an average over four point measurements: the top of the heated slide, on the coverglass and on the objective in immersion droplets, and on the side of the objective. Temperatures are stable within ± 0.2 °C during a given measurement, which are performed from ambient to 47.4 °C.

Round No. 1.5 coverslips for the flow cell assembly are cleaned by rinsing with acetone, ethanol, HPLC grade water, then methanol followed by baking at 500 °C for 5 hours. All experiments are performed in buffer of 50 mM hemisodium HEPES (pH 7.5 at 25 °C), 100 mM NaCl, 100 μM ETDA, and 1 mM MgCl$_2$. HEPES is a temperature stable buffer with only minor changes in pH with temperature (ΔpKa = −0.014/°C), indicating a pH reduction to 7.22 at 45° (187). Immobilized samples are prepared in the FCS2 holder using a biotinylated BSA–streptavidin-biotinylated RNA-tethering scheme (147). An enzymatic oxygen scavenging solution of glucose (9 mg/mL), glucose oxidase (0.43 mg/mL), catalase (0.072 mg/mL), and 2 mM Trolox is added to the buffer to reduce fluorophore photobleaching and photophysics (129). For freely diffusing measurements, the glass surfaces are passivated with BSA and experiments are performed with 200 pM RNA solution in the holder (128). The oxygen scavenging system is not used in the freely diffusing experiments as it was found to offer no advantage at the powers used (128).

4.3.3   Single-Molecule FRET Efficiency Analysis

Immobilized single-molecule trajectories are analyzed with 10 ms data binning, which clearly resolves the undocked and docked events (106,128). Freely diffusing time traces are analyzed with 1 ms integration times, which is on the order of a molecule's dwell time in the laser focus (106,128). As previously derived, the corrected intensity-based FRET efficiency (E_{FRET}) is calculated from the background subtracted signals on the two channels, ΔI_1 and ΔI_2, designed primarily for donor and acceptor detection, respectfully. Corrections are implemented for (i) collection efficiencies and crosstalk of the donor and acceptor emission on channels 1 and 2 ($\beta_1{}^A$, $\beta_2{}^A$, $\beta_1{}^D$, $\beta_2{}^D$) (ii) differential quantum yields of the donor and acceptor (Q_D and Q_A), and (iii) fractional direct excitation of the acceptor vs. donor (α_A, where $1-\alpha_D = \alpha_A$),

$$E_{FRET} == \frac{\beta_1{}^D \Delta I_2 - \beta_2{}^D \Delta I_1 - \dfrac{\alpha_A}{\alpha_D}\dfrac{Q_A}{Q_D}(\beta_2{}^A \Delta I_1 - \beta_1{}^A \Delta I_2)}{\beta_1{}^D \Delta I_2 - \beta_2{}^D \Delta I_1 + \dfrac{Q_A}{Q_D}(\beta_2{}^A \Delta I_1 - \beta_1{}^A \Delta I_2)}. \qquad 4.1$$

Quantum yield ratios and collection efficiencies are determined in independent measurements of singly labeled constructs, $(Q_A/Q_D) = 1.2 \pm 0.3$, β_1^A (0.00000 ± 0.00003), β_2^A (0.0242 ± 0.0018), β_1^D (0.0269 ± 0.0024), $\beta_2^D = 0.00211 \pm 0.00018$ (128). Fractional direct laser excitation of the acceptor and donor is calculated from the extinction coefficients at 532 nm, $\alpha_A = 0.07 \pm 0.01$, $\alpha_D = 0.93 \pm 0.01$ (128). In contrast to the immobilized studies, donor-only species in the freely diffusing studies are necessarily included and appear at negative E_{FRET} values due to correction for acceptor direct excitation (128).

4.4 Results

4.4.1 Temperature Dependence of Tetraloop–Receptor-Docking Equilibrium Revealed in Single-Molecule Trajectories

The docking/undocking equilibrium of the GAAA tetraloop and receptor is readily monitored by FRET between the donor (Cy3) and acceptor (Cy5) in the RNA construct (Figure 4.1), as described elsewhere (106). The tetraloop is flexibly attached to the receptor domain with a single-stranded A_7 linker, which was previously shown to allow for an isolated thermodynamic study of a tertiary interaction (106,124). Temperature-dependent conformational dynamics in single-immobilized molecules at physiological conditions (100 mM NaCl, 1 mM MgCl$_2$, 50 mM HEPES, pH 7.5) are observed in E_{FRET} trajectories calculated ratiometrically from real-time donor- and acceptor-fluorescence signals (see *Experimental Procedures*) (106,128). Sample trajectories portraying the temperature dependence of the tetraloop–receptor-docking equilibrium are shown in Figure 4.2. Fluctuations between undocked (low E_{FRET}) and docked (high E_{FRET}) conformations are visible in each of the time traces (106). Probability distributions from the E_{FRET} trajectories allow two states to be distinguished (right panel, Figure 4.2).

The E_{FRET} probability histograms are well described by Gaussian distributions (128); fitting the histograms to a sum of Gaussians permits quantitation of relative docked vs. undocked populations. From cumulative probability distributions of many molecules, we determine $\langle E_{FRET} \rangle = 0.29 \pm 0.02$ and 0.70 ± 0.02 and width $\sigma = 0.14 \pm 0.01$ and 0.094 ± 0.003 for the undocked and docked states, respectively, in excellent agreement with previous studies of this construct (128). The E_{FRET} centers ($\langle E_{FRET} \rangle$) and
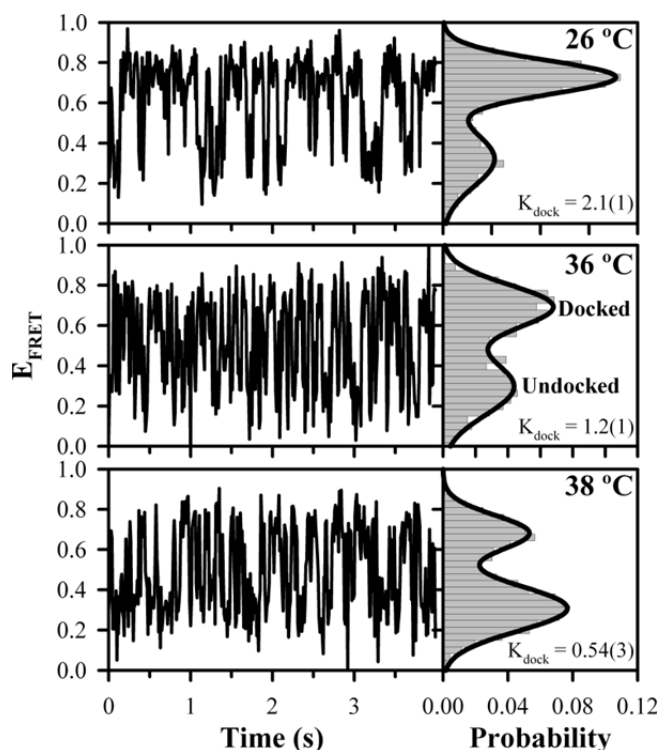
Figure 4.2   Immobilized single-molecule E_{FRET} trajectories and the corresponding probability distributions at 26 °C (*top*), 36 °C (*middle*) and 38 °C (*bottom*). The low and high E_{FRET} peaks correspond to the undocked and docked states, respectively. Integrated areas of the undocked and docked peaks are determined from the superimposed two-Gaussian fits with the equilibrium constant for docking, K_{dock}, calculated as the ratio of the docked to undocked area. Increasing temperature shifts the equilibrium toward the undocked state.

Gaussian widths (σ) for the molecules shown in Figure 4.2 at 26, 36, and 38 °C show no systematic shifts as the temperature is increased, indicating that (i) obscuring of docked and undocked peaks due to finite binning is negligible and (ii) the tetraloop and receptor secondary structure are maintained upon heating, since disruption of these regions would lead to shifts in the E_{FRET} peaks. Secondary structure is expected be unaffected in this study as both the tetraloop and receptor domains are extremely stable; the melting temperature (T_m) of the tetraloop domain was experimentally determined at ~64 °C (188) and T_m of the receptor domain is predicted to be ~62 °C from the DINAMelt server (189). Furthermore, the robustness of the E_{FRET} peak position suggests fluorophore quantum yields are sufficiently independent of temperature over the range measured.

The effect of temperature on fluorophore quantum yields is independently assessed on singly-labeled constructs at the experimental extremes. Comparison of relative quantum yields at 21 °C and 45 °C shows a decrease of ~20% for both Cy3 and Cy5 as the RNA is heated over this range; the quantum yield effect is reversible as observed by temperature cycling. A similar diminution of Cy3 quantum yield has also been observed by Levitus and co-workers for Cy3-DNA conjugates (116). Quantum yield changes induce a two-fold effect on the observed E_{FRET}, where $E_{\text{FRET}} = R_0^6/(R_0^6+R^6)$, both through the Förster radius (R_0) and in the E_{FRET} calculation from emission intensities (see Eq. 4.1) for a defined distance (R). The latter contribution is negligible in the ratiometric determination of E_{FRET} as the relative ratio Q_A/Q_D does not change (see Eq. 4.1). The R_0 effect is also subtle because R_0 is proportional to $Q_D^{1/6}$ (111). For the observed decrease of ~20% in donor quantum yield, we predict an E_{FRET} center decrease of only ~0.03 for undocked and docked peaks over the entire temperature range investigated (see *Supporting Information*, Figure 4.5). Such a peak shift may be present but is on the order of E_{FRET}- center reproducibility (± 0.02) and not relevant within experimental uncertainty. Furthermore, absolute positions of these peaks have no effect on the analysis presented here as we seek only the relative areas of the docked and undocked peaks, which are populated based on the fractional dwell time of the molecule in each configuration.

The calculated fluorophore distance from the corrected E_{FRET} in the docked conformation is consistent with proper formation of the tetraloop–receptor interaction as observed in the X-ray crystal structure of the P4–P6 domain of the Tetrahymena ribozyme (62). The inter-phosphate distance between the corresponding nucleotides to which the fluorophores would be attached in the P4–P6 domain is 31.6 Å. Our observed E_{FRET} of 0.29 corresponds to $R = 46$ Å in the docked state for an R_0 of 53.4 Å [$R^6 = R_0^6(1/E_{\text{FRET}} -1)$]. The functionalized fluorophores are attached to amino-modified RNA with a 3 carbon linker, which places 12 atoms between each fluorophore and its nucleotide phosphate. This added distance can easily account for the additional 14.4 Å between the Cy3 and Cy5 as compared to the inter-phosphate distance in P4–P6, though the exact position of the fluorophores in this construct is not structurally known. The observed E_{FRET} for the undocked conformation also indicates a distance ($R = 62$ Å)

that is consistent with a right angle extension of the tetraloop A_7 arm with respect to the receptor helix ($R \sim 70$ Å). However, inferring absolute distances from FRET efficiencies is challenging because fluoro-phore rotation is potentially hindered upon covalent attachment to biomolecules. Hindered rotation affects R_0, which is calculated assuming free rotation (i.e., κ^2 is assumed to be $\frac{2}{3}$), though this assumption is spurious as shown in recent studies by Lilley and coworkers (113). Nevertheless, observed E_{FRET} values are in agreement with the anticipated unfolded and folded conformations of the tetraloop–receptor construct.

The undocked and docked conformation of the tetraloop–receptor are clearly assigned, allowing for determination of equilibrium constants from the ratios of integrated peak areas (K_{dock} = area docked/area undocked), which is the fractional dwell time in the docked vs. undocked state (92). An enhanced tendency for tetraloop–receptor constructs to be in the undocked conformation with increasing temperature is apparent from the single-molecule traces, with K_{dock} decreasing from 2.1 ± 0.1 at 26 °C to 0.54 ± 0.03 at 38 °C (Figure 4.2). Cumulative K_{dock} values for many molecules are in good agreement with the individual molecules (see Figure 4.2) and summarized over the full temperature range (21.0–47.4 °C) in Table 4-1[K_{dock} (immobilized)], indicating a strongly exothermic-folding process.

Table 4-1   Temperature dependence of K_{dock} and thermodynamic parameters for intermolecular te-traloop–receptor docking via an A_7 linker from immobilized and freely diffusing (free) single-molecule methods

T (°C)	K_{dock} (immobilized)	K_{dock} (free)	$\Delta G°$ (kcal/mol) = $-RT \ln K_{dock}$
21 ± 0.2	4.18 ± 0.18	3.7 ± 1.8	−0.80 ± 0.05
26 ± 0.2	2.3 ± 0.1	–	−0.50 ± 0.03
29 ± 0.2	2.2 ± 0.2	1.7 ± 0.3	−0.4 ± 0.1
34 ± 2	1.32 ± 0.08	0.87 ± 0.16	−0.05 ± 0.16
36 ± 2	1.17 ± 0.09	–	−0.10 ± 0.05
38 ± 2	0.64 ± 0.03	0.60 ± 0.11	0.30 ± 0.03
43 ± 3	–	0.55 ± 0.10	0.43 ± 0.12
47.4 ± 3.6	–	0.42 ± 0.09	0.58 ± 0.12
	Immobilized	free	
$\Delta H°$ (kcal/mol)	−17.4 ± 1.6	−17.2 ± 1.6	
$\Delta S°$ (cal/mol/K)	−56.2 ± 5.4	−55.9 ± 5.2	

Significant kinetic heterogeneity exists in the tetraloop–receptor system, as discussed previous-ly (106). This heterogeneity can be adequately described as two non-interconverting populations: (i) 68% of species actively dock/undock and (ii) a 32% minority population of molecules that exhibit no folding events. The presence of "nondocking" molecules was also confirmed under freely diffusing con-ditions, yielding a 32% subpopulation in quantitative agreement with the tethered studies (128). Since these molecules show no docking events over the range of time scales and temperatures sampled, they are excluded from the above equilibrium analysis. The molecular origin of the species is not yet known, e.g., is the kinetic trap of secondary or tertiary origin. These molecules never achieve the bound form of the tetraloop receptor so cannot be included to assess tetraloop–receptor thermodynamics. Studies have suggested that the nondockers can be removed from the sample by native-gel electrophoresis (124), however we retain them in the study, to assess whether or not they can interconvert at increased temperatures. Temperatures hotter than investigated may allow the nondockers to fold, which would be consistent with the remarkable robustness of subpopulations in the hairpin ribozyme (99). Further confirmation of the temperature insensitivity of the nondocking RNA subpopulation is provided by stud-ies under freely diffusing conditions as described below.

4.4.2 Temperature Dependence of Docking Equilibrium in Freely Diffusing RNA

For temperature-dependent measurements of the tetraloop–receptor interaction without surface im-mobilization, the equilibrium conformational distribution of the tetraloop–receptor construct is ac-quired from individual molecules freely diffusing through the confocal detection volume. This method allows for rapid sampling of many molecules, but with the ability to still discern conformational states (128). Low concentrations (200 pM) of RNA ensure that we are in the single-molecule detection regime, which is confirmed by cross correlation analysis of the donor and acceptor channels for freely diffusing time traces (see *Supporting Information*, Figure 4.6) yielding an average occupancy of the focal volume of 0.71 ± 0.09 molecules. Cross correlations also indicate the concentration of donor-acceptor labeled species is maintained upon heating, indicating that RNA construct remains intact. Time traces are

binned at 1 ms integration for analysis, which is on the order of the dwell time of a molecule in the detection volume (128). This bin time is much shorter than the typical duration of a docking or undocking event; thus a molecule can be cleanly identified in either the undocked or docked conformation (106). Fluorescence emission from labeled RNA molecules is distinguished from background via an intensity threshold ($I_{threshold}$ = 10 $\sigma_{background}$ = 30 kHz) of a minimum sum of photons above background levels with typically >10^4 events acquired to ensure high-quality statistics (118). In freely diffusing studies, all fluorescent species in the solution are sampled. Therefore signals from donor-only molecules give rise to a
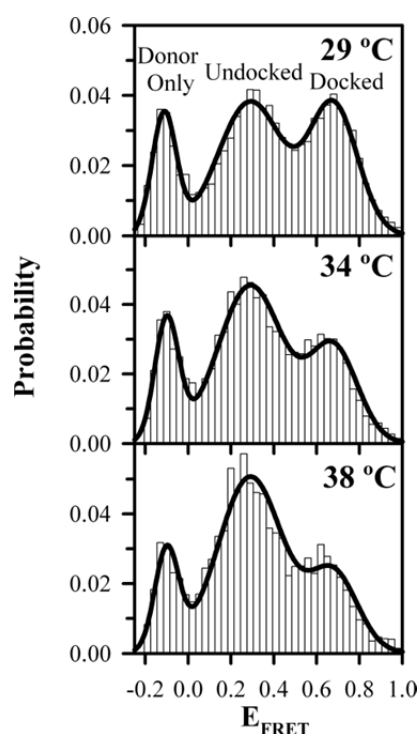
Figure 4.3   Temperature dependence of tetraloop–receptor docking shown in E_{FRET} histograms generated from photon bursts occurring when freely diffusing RNAs traverse the laser focal volume (see *Experimental Procedures*). Three distinct populations with the peaks corresponding to the donor-only (E_{FRET} < 0), undocked (center E_{FRET} = 0.291 ± 0.003) and docked RNA (center E_{FRET} = 0.679 ± 0.004) at 29 °C, 34 °C and 38 °C. There is a marked increase in the relative fraction of the undocked population with temperature indicating disruption of the tertiary interaction, which is quantified by fitting the histograms to a sum of Gaussian distributions (*solid black lines*).

peak at negative E_{FRET} due to direct excitation correction for the missing/photobleached acceptor (see *Experimental Procedures*) (128). As illustrated in Figure 4.3, the donor-only population shows no

change with temperature, consistent with the RNA constructs remaining hybridized throughout temperature cycling.

The freely diffusing E_{FRET} histogram data can be globally fit over all temperatures by a sum of Gaussians distributions with common widths and centers, confirming that the docked/undocked peak shapes and positions are not significantly affected by heating. Such an analysis yields $\langle E_{\text{FRET}}\rangle_{\text{undocked}}$ = 0.291 ± 0.003, σ_{undocked} = 0.153 ± 0.004, and $\langle E_{\text{FRET}}\rangle_{\text{docked}}$ = 0.679 ± 0.004, σ_{docked} = 0.112 ± 0.003, agreeing with the above immobilized results and previous freely diffusing studies at room temperature ($\langle E_{\text{FRET}}\rangle$ = 0.28 ± 0.01 and 0.687 ± 0.005 and σ = 0.165 ± 0.004 and 0.106 ± 0.002 for undocked and docked states, respectively) (128).   Temperature and immobilization, therefore, do not introduce any observable changes in FRET centers and widths for both the undocked and docked RNA conformations, indicating that the system is well described by a two-state reaction scheme under all conditions in these studies. As demonstrated in Figure 4.3, the docked and undocked E_{FRET} amplitude*s* do display temperature sensitivity. Specifically, increased temperature shifts the tetraloop–receptor construct to the undocked conformation, which is quantified by extracting K_{dock} from the equilibrium distributions. Extracting equilibrium constants from the threshold event distribution is potentially only an estimate of K_{dock} because of preferential detection of molecules dominantly emitting photons from the fluorophore with higher net collection efficiency. However, systematic variation of the threshold (15-75 kHz) reveals negligible effect on the results presented. Such a result is expected as the collection efficiencies/quantum yields of donor and acceptor are nearly equivalent in our experimental apparatus (see *Experimental Procedures*).

The freely diffusing detection method samples an equilibrium conformational distribution of all tetraloop–receptor constructs, including both actively docking (68%) and nondocking (32%) subpopulations. The nondocking subpopulation can be incorporated in the analysis by noting that the fractional population in the docked state (F_{free}) must scale linearly (i.e., $F_{\text{free}} = \alpha F_{\text{immobilized}}$) with the fractional docked state populations under immobilized conditions, where nondocking molecules are excluded (128). This fraction can be written explicitly as,

$$F_{\text{free}} = \frac{N_{\text{docked}}}{N_{\text{docked}} + N_{\text{undocked}} + N_{\text{nondock}}} = \alpha \frac{N_{\text{docked}}}{N_{\text{docked}} + N_{\text{undocked}}} = \alpha \frac{K_{\text{dock}}}{K_{\text{dock}}+1}, \qquad 4.2$$

where N_{docked} and $N_{\text{undocked}} + N_{\text{nondock}}$ are integrated areas of docked and undocked peaks in freely diffusing studies, respectively. The actively docking fraction, α, has been previously measured both for immobilized ($\alpha = 0.68 \pm 0.01$) (106) and freely diffusing ($\alpha = 0.66 \pm 0.02$) species (128). As a result, K_{dock} is readily extracted from F_{free} by

$$K_{\text{dock}} = \frac{F_{\text{free}}}{\alpha - F_{\text{free}}}, \qquad 4.3$$

where the nondocking fraction is identified by immobilized studies to be independent of temperature. As summarized in Table 4-1, K_{dock} under diffusing conditions decreases from 3.7 ± 1.8 at 21 °C to 0.42 ± 0.09 at 47 °C. The equilibrium constants are in agreement for freely diffusing and immobilized constructs confirming that the nondocking fraction is constant over this temperature range.

4.4.3   Thermodynamics for Tetraloop-Receptor Docking

To extract thermodynamic information from these two independent methods, we analyze both the immobilized and diffusing data sets according to the van't Hoff equation,

$$R \ln K_{\text{dock}} = -\frac{\Delta H^\circ}{T} + \Delta S^\circ, \qquad 4.4$$

from which a plot of $R \ln K_{\text{dock}}$ vs. $1/T$ yields a slope of $-\Delta H^\circ$ and an intercept of ΔS° (see Figure 4.4), where R is the gas constant (1.987 cal/mol/K). Both data sets yield straight line van't Hoff plots, of which least-squares fits weighted for uncertainty of the dependent and independent variables yield $\Delta H^\circ_{\text{immobilized}} = -17.4 \pm 1.6$ kcal/mol, $\Delta H^\circ_{\text{free}} = -17.2 \pm 1.6$ kcal/mol, $\Delta S^\circ_{\text{immobilized}} = -56.2 \pm 5.4$ cal/mol/K, and $\Delta S^\circ_{\text{free}} = -55.9 \pm 5.2$ cal/mol/K, which can be converted to standard international units with the conversion factor 4.184 J/cal. Freely diffusing and immobilized approaches for characterizing the thermodynamics of the tetraloop–receptor tertiary interaction are in excellent agreement (Figure 4.4). Docking of the tetraloop with the receptor results in a substantial decrease in enthalpy with unfavorable entropy change (i.e., "enthalpy-driven folding"), where exothermicity is balanced at room temperature
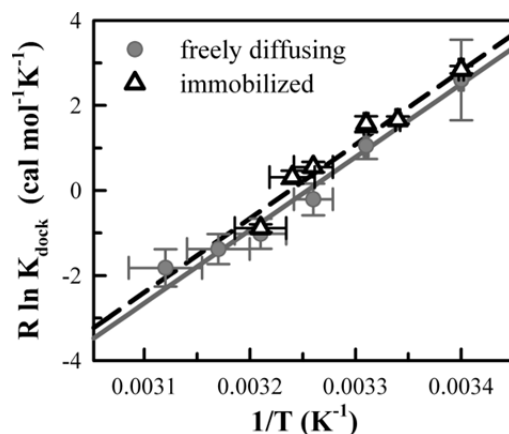
Figure 4.4   Thermodynamics of the tetraloop–receptor docking/undocking equilibrium from van't Hoff plots (see Eq. 4.4). The temperature (T) dependence of the docking equilibrium constant (K_{dock}) is shown for freely diffusing (*gray circles*) and immobilized (*black open triangles*) molecules. Linear fits of $R \ln K_{dock}$ vs. $1/T$ yield a slope of $-\Delta H°$ and intercept of $\Delta S°$ (see Table 4-1) for the freely diffusing (*gray solid line*) and immobilized (*black dashed line*) data.

by a high entropic cost. More quantitatively, the standard state free energy ($\Delta G°$) for forming the te-

traloop–receptor interaction can be directly calculated as a function of temperature from the equilibri-

um constants ($\Delta G° = -RT \ln K_{dock}$), as summarized in Table 4-1. At 21 °C the docked state is only margin-

ally favored ($\Delta G° = -0.08$ kcal/mol ± 0.05), with docking becoming thermodynamically unfavorable at

physiological temperatures ($\approx$ 37 °C).

4.5   Discussion

Accurate thermodynamic characterization of tertiary RNA folding is clearly important but has

been extremely limited (103,106,124,126). We demonstrate the applicability of temperature-dependent

smFRET methods under both immobilized and freely diffusing conditions to deconstruct the docking

thermodynamics for an isolated tetraloop–receptor tertiary motif. Freely diffusing techniques are par-

ticularly useful for monitoring RNA folding thermodynamics in the absence of surface tethering as was

demonstrated for the exothermic and entropically disfavored secondary folding of isolated RNA hairpins

(186). We determine the standard state enthalpic and entropic components of the tetraloop–receptor

binding, revealing that the tertiary interaction like the secondary hairpin formation is enthalpy driven,

but at a large entropic cost. Extension of these methods to other isolated tertiary interactions and comparison to RNAs with multiple interactions should permit valuable insights into, for example, the thermodynamic origin of tertiary cooperativity (39,92). Such dissections of RNA folding interactions are applicable to well-defined, modular components as in this scheme. Such a "bottom up" approach complements studies in large RNAs by avoiding potentially deleterious mutations that may disrupt the tertiary interaction through changes in global structure.

NMR and X-ray crystal structures of the docked and undocked tetraloop and receptor reveal structural contributions to the entropic cost and enthalpic favorability of tetraloop–receptor association. Specifically, tetraloop–receptor docking induces structural elements that have been generally recognized as enthalpically stabilizing and entropically unfavorable, namely hydrogen-bonding and base-stacking interactions (62,74,120,190). The large entropic penalty for folding may also originate from the loss of free orientational flexibility of the tetraloop in the undocked state to the specific orientation required for docking (62,73). The proposition that the flexibility in an unfolded structure is correlated with the entropic cost of folding is supported by thermodynamic studies of the hairpin ribozyme, in which a four-way junction (4WJ) greatly decreased the entropic penalty of folding as compared to the two-way junction (2WJ). The 4WJ contributes significant rigidity and orientation to the unfolded state, leading to a decrease in disorder prior to the docking event (185). Solvent effects, such as disruption of ordered water molecules, yield positive enthalpic and entropic folding contributions; this is opposite to what is observed for tetraloop–receptor folding and therefore not the dominant effect (191).

Although the measured enthalpy and entropy changes are dominated by the tetraloop-receptor interaction, we also consider thermodynamic contributions arising from the RNA construct design. Specifically, contributions could arise from the flexible A_7 linker because of small differences in its base stacking in the docked and undocked forms. Additional negative entropic contributions to folding due to the reduced conformational space accessible to the linker in the docked vs. undocked states are expected to be modest because this region is sufficiently long to maintain disorder. By way of confirma-

tion, previous studies of linker length (A_7 vs. A_{14}) and composition (A vs. U) dependence of tetraloop–receptor docking/undocking revealed minimal changes in the equilibrium constant (106,124).

Our results allow us to place the tetraloop–receptor interaction in thermodynamic context with global folding of the P4–P6 domain in the *Tetrahymena* ribozyme, which has been studied at the ensemble level by temperature-gradient gel electrophoresis under similar cation conditions (see Table 4-2) (11). Results for the global fold indicate large net enthalpy and entropy decreases of $\Delta H° = -28 \pm 3$ kcal/mol and $\Delta S° = -91 \pm 8$ cal/mol/K, respectively, i.e., also enthalpically driven with only slightly larger magnitude than the tetraloop–receptor values. Although there are additional tertiary interactions and structural complexity to consider in the full P4–P6 folding problem, it would appear that the GAAA tetralop–receptor alone may contribute $\approx 60\%$ of the $\Delta H°$ and $\Delta S°$. The tetraloop–receptor interaction has been proposed as a thermodynamic clamp for stabilizing the global fold of large RNAs (126). This work supports such a picture by identifying that thermodynamic stability may originate from the strongly favorable tetraloop–receptor docking enthalpy. Further inspection will be necessary to evaluate how the complexity of large RNAs affects the folding thermodynamics, with assessment of the role of global architecture, secondary elements, and solvent accessibility.

Table 4-2   Thermodynamic parameters for intermolecular tetraloop–receptor docking via A_7 linker and other tertiary folding (4.184 J/cal)

RNA Tertiary Interaction	Conditions	$\Delta H°$ (kcal/mol)	$\Delta S°$ (cal/mol/K)	$\Delta G°$ at 37 °C $=\Delta H° - T\Delta S°$ (kcal/mol)
Tetraloop–receptor docking	immobilized, 1 mM MgCl$_2$, 0.1 M NaCl	-17.4 ± 1.6	-56.2 ± 5.4	0.02 ± 2.3
	freely diffusing, 1 mM MgCl$_2$, 0.1 M NaCl	-17.2 ± 1.6	-55.9 ± 5.2	0.13 ± 2.3
P4–P6 domain folding (11)	0.9 mM MgCl$_2$, 10 mM NaCl	-28 ± 3	-91 ± 8	0.21 ± 3.9
Hairpin ribozyme folding (185)	2WJ, 1 mM MgCl$_2$	-9.7 ± 4.5	-34 ± 17	0.84 ± 7
	4WJ, 1 mM MgCl$_2$	-4.1 ± 2.3	-9.6 ± 5.7	-1.1 ± 2.9
P1 helix docking into *Tetrahymena* *g*roup I intron core	10 mM MgCl$_2$ (103)	8 ± 2	25 ± 8	0.25 ± 3.2
	5 mM MgCl$_2$, 135 mM NaCl (192)	8	40	-4
	10 mM MgCl$_2$ (193)	19 ± 9	62 ± 30	-0.22 ± 13
3WJ folding of hammerhead ribozyme core (182)	10 mM MgCl$_2$, 0.1 M NaCl	41 ± 1	120 ± 10	3.8 ± 3.3

However, the docked form of the tetraloop–receptor is thermodynamically disfavored (i.e., $\Delta G^{\circ}_{\text{immobilized, free}}$ = 0.022 ± 2.3, 0.13 ± 2.3 kcal/mol) under physiologically relevant conditions (37 °C, 1 mM $MgCl_2$, 100 mM NaCl), with similar results for the P4–P6 domain (ΔG° = 0.21 ± 3.9 kcal/mol). This borderline stability of even strongly enthalpy-driven folding underscores that RNA folding dynamics must be addressed for a complete understanding of functionality. Folding of both P4–P6 and the tetraloop–receptor become favorable with increased Mg^{2+} concentration (92,106).

The thermodynamic parameters for P4–P6 and tetraloop–receptor folding are consistent with that of other RNA folding, specifically global folding of tRNA (194), the 1051–1108 rRNA fragment (195), the hairpin ribozyme (185), pseudoknot tertiary structure formation (180), ligand-induced riboswitch folding (196), and secondary loop formation (48,186). Although differing in magnitude, each of these systems exhibits significant enthalpic gain (ΔH° < 0) and entropic penalty (ΔS° < 0) for folding. Furthermore, in a recent ITC study, Reymond *et al*. demonstrated through systematic mutations that various steps in the folding pathway of the hepatitis delta virus ribozyme are exothermic and entropically disfavored (184). However, these observations of enthalpy-driven RNA folding are not a universal trend, as evidenced in Table 4-2. The final stage of *Tetrahymena* ribozyme folding, i.e., the docking of the P1 duplex into the prefolded core, was investigated by smFRET of immobilized molecules and found to be enthalpically disfavored (ΔH° = 8 ± 2 kcal/mol) and entropically favored (ΔS° = 25 ± 8 cal/mol/K), in reasonable agreement with ensemble methods (192,193). ITC studies of the three-way junction (3WJ) in the minimal hammerhead ribozyme core reveal a similar degree of endothermicity and entropic advantage (Table 4-2) (182). As yet a third case, smFRET investigation of the hairpin ribozyme's 4WJ revealed no temperature dependence in the folding equilibrium constants, from which one can infer the processes to be nearly thermoneutral ($\Delta H^{\circ} \approx$ 0 kcal/mol) with only a modest decrease in disorder ($\Delta S^{\circ} \approx$ −2 ± 1cal/mol/K) between the various folded conformations (108).

As illustrated in Table 4-2 each tertiary association has ΔG° near zero; thus RNA folding can be regarded as either enthalpy or entropy driven. Folding events in the tetraloop–receptor, P4–P6 domain,

and hairpin ribozyme result in many hydrogen bonds and base stacking interactions, of which tetraloop–receptor docking forms the fewest, i.e., 10 hydrogen bonds and improved base stacking (62,74,197). By way of contrast, such enthalpy-driven interactions are much less prevalent in the entropy-driven P1 docking, which nets only 4–5 hydrogen bonds and no base stacking (198). Reduced prevalence of these secondary-like interactions may account for the shift from enthalpy-driven to entropy-driven (solvent-driven) folding. The net entropic drive for this folding has been proposed to arise from displacement of ordered water molecules and/or Mg^{2+} ions (192,193). Furthermore, the *Tetrahymena* ribozyme is already highly structured prior to P1 docking, which reduces the entropic cost associated with the final folding transition.

Folding of the minimal hammerhead ribozyme is more difficult to categorize in terms of base stacking and hydrogen bonding because the structure of the core prior to coaxial stacking of the helices is not clear. However, counting hydrogen bonds and base-stacking interactions that are contingent on helix alignment, i.e., not including the 2 noncanonical base pairs (A_9–G_{12} and G_8–A_{13}) that likely form in the core prior to helical stacking, we identify a net of ~7 interactions (58,199). This is intermediate albeit closer to the case of P1 than tetraloop–receptor docking; therefore solvent effects, e.g., displacement of ordered water molecules, are not overcome and entropy increases in minimal hammerhead ribozyme folding (182,192,193). Endothermic changes are associated with rearrangement of core residues of the hammerhead ribozyme (182), which may explain the additional enthalpic penalty for hammerhead core folding vs. P1 docking. In summary, the data suggest that categorization of enthalpy- vs. entropy-driven RNA folding arises from competing roles of solvent and hydrogen bonding/stacking interactions, with a net ~11 interactions providing sufficient negative $\Delta H°$ and $\Delta S°$ to overcome solvent effects, i.e., displacement of order H_2O. Furthermore, each of the entropy-driven RNA folds (see Table 4-2), as well as the thermoneutral 4WJ, were measured at significantly higher concentrations of Mg^{2+} than the enthalpy-driven cases (> 10 mM vs. ~ 1 mM Mg^{2+}, see Table 4-2. A Mg^{2+} environment that is more amenable to RNA folding, though less physiological, may also contribute to the observation of entropy-driven folding by increasing the magnitude of solvent effects. Further systematic study of isolated tertiary motifs with

temperature-dependent smFRET methods will be invaluable in further elucidating the competing roles of solvent and hydrogen bonding/stacking interaction to categorize enthalpy- vs. entropy-driven RNA folding.

As a final comment, in the course of preparing this manuscript, we became aware of a simultaneous and independent ITC investigation of tetraloop–receptor binding by the Butcher group. Specifically, the Butcher group had cleverly designed pairs of RNA constructs, based on back-to-back copies of tetraloops (TT) and receptors (RR). These constructs can form a duplex (TRTR) through dual tetraloop–receptor docking, which permits association to be initiated by mixing of the two constructs. This work by Vander Meulen *et al*. published elsewhere (125) yields $\Delta H° = -33.2 \pm 2.0$ kcal/mol at 45 °C for double tetraloop–receptor formation under similar solvent conditions (2 mM MgCl$_2$, 20 mM KCl). Assuming zero enthalpy cooperativity between the two tertiary interactions this translates into $\Delta H°_{ITC} = -16.9 \pm 1.0$ kcal/mol for single tetraloop–receptor binding, which agrees with the present single molecule values of $\Delta H°_{free} = -17.2 \pm 1.6$ kcal/mol and $\Delta H°_{immobilized} = -17.4 \pm 1.6$ kcal/mol. This agreement between isolated and dual tetraloop–receptor docking enthalpies supports a picture that tertiary structure formation in RNA may be largely enthalpically noncooperative. In turn, this would imply an entropic origin of tertiary cooperativity between the tetraloop–receptor and metal-core interactions observed in folding of the complete P4–P6 domain (92).

4.6   Conclusion

Thermodynamics of the isolated GAAA tetraloop-receptor tertiary interaction are investigated at the single-molecule level, exploiting the combination of temperature-dependent FRET and confocal microscopy methods. Results for the GAAA-tetraloop receptor motif are shown to be enthalpically driven ($\Delta H°_{free} = -17.2 \pm 1.6$ kcal/mol and $\Delta H°_{immobilized} = -17.4 \pm 1.6$ kcal/mol), yet balanced by entropically unfavorable ($\Delta S°_{free} = -55.9 \pm 5.2$ cal/mol/K and $\Delta S°_{immobilized} = -56.2 \pm 5.4$ cal/mol/K) conformational changes upon docking. These results for tetraloop–receptor folding already achieve 60% of the values obtained from ensemble studies for folding of the complete P4–P6 domain, supporting previous views

that the tetraloop–receptor provides a dominant source of thermodynamic stabilization. However, under physiological temperature conditions, this isolated tertiary motif is found to be marginally unstable ($\Delta G°$ = + 0.02 and 0.13 kcal/mol from immobilized and freely diffusing), indicating that conformational change may still play a key role in RNA functionality. Further work on temperature and cation dependences of the rates, k_{dock} and k_{undock}, will be necessary to illuminate the transition-state enthalpy and entropy changes associated with tertiary contact formation, as well as help develop a consensus picture for the thermodynamic origin of cation-mediated RNA folding.

4.7 Acknowledgments

4.8 Supporting Information

Prediction of donor quantum yield (Q_D) effect on the observed FRET efficiency (E_{FRET}) as a function of Cy3-Cy5 distances (R) (Figure 4.5) (111) and mean cross correlations of donor and acceptor channels for the same sample containing tetraloop–receptor constructs under freely diffusing single-molecule conditions at 21 and 45 °C (Figure 4.6) (131).
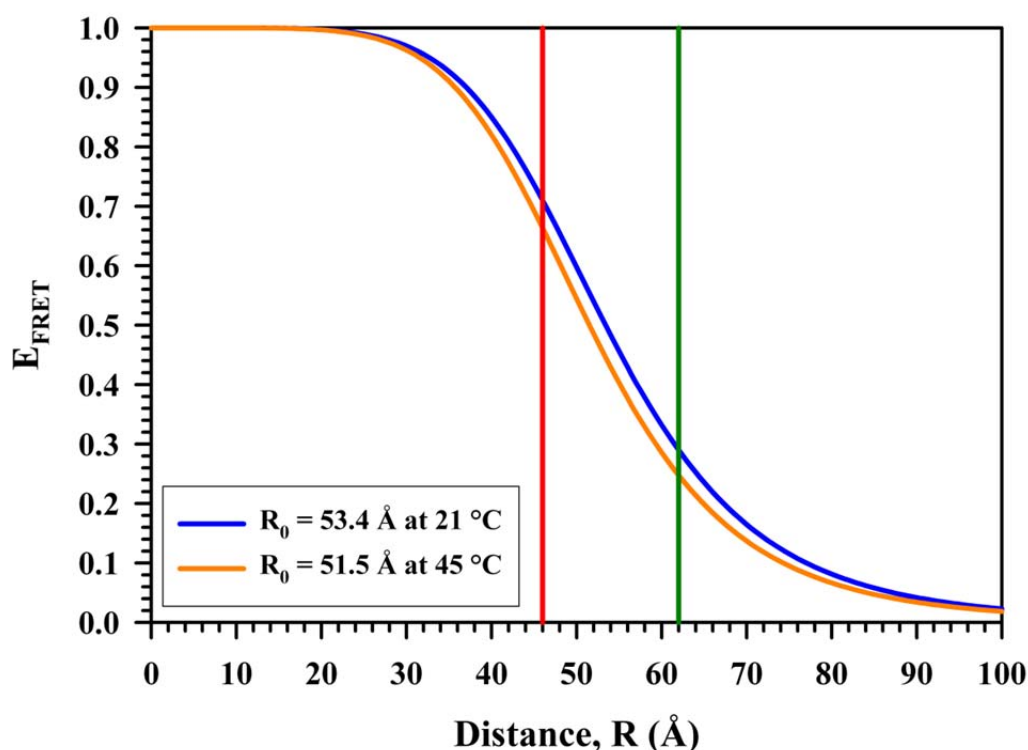
Figure 4.5   Prediction of donor quantum yield (Q_D) effect on the observed FRET efficiency (E_{FRET}) as a function of Cy3-Cy5 distances (R). E_{FRET} vs. R is calculated for 21 °C and 45 °C, where $E_{FRET} = R_0^6/(R_0^6+R^6)$ with a Förster radius, R_0, which is proportional to $Q_D^{1/6}$. The room temperature R_0 is calculated to be 53.4 Å from experimental spectra of Cy3- and Cy5-only labeled tetraloop–receptor constructs. The Cy3 quantum yield decreases by ~20% when heated to 45 °C, which theoretically decreases R_0 by 3% to 51.5 Å. Such a reduction in R_0 corresponds to a prediction of a ~0.03 shift in the mean E_{FRET} for the docked and undocked conformations from the 21 °C values. This decrease in E_{FRET} is indicated by the vertical lines at the observed E_{FRET} value for the docked and undocked states at 21 °C (46 Å and 62 Å). Such a small decrease in E_{FRET} is on the order of the reproducibility of peak centers in the experiment and therefore not observable over the temperature range investigated. Cy5 and Cy3 display similar quantum yield reductions over this temperature range and therefore Q_A/Q_D does not change; the corrected E_{FRET} (see. Eq. 4.1 in the text) calculated ratiometrically from the donor and acceptor emissions is negligibly affected by the changes in the quantum yield ratio.
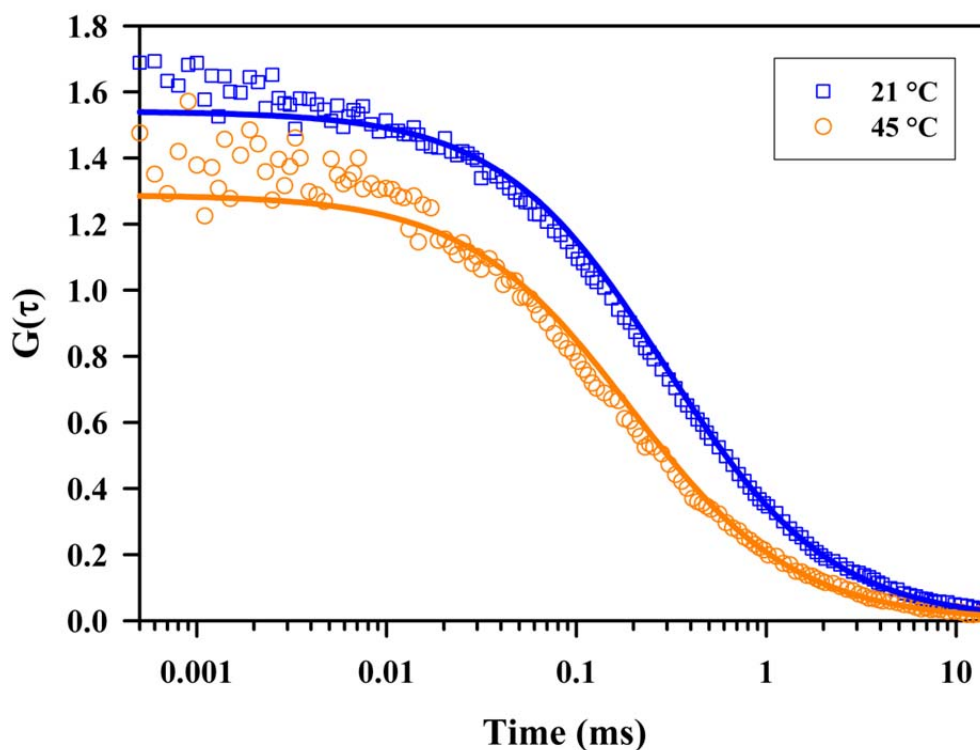
Figure 4.6   Mean cross correlations of donor and acceptor channels for the same sample containing tetraloop–receptor constructs under freely diffusing single-molecule conditions (see *Experimental Procedures*) at 21 and 45 °C. Cross correlations, $G(\tau) = \dfrac{\delta F_i(t)\delta F_j(t+\tau)}{\langle F_i(t)\rangle\langle F_j(t)\rangle}$, are calculated by a software analysis of 60 s time-correlated single-photon counting data traces; six such cross correlations are averaged from the same sample to calculate the mean cross correlation and standard deviations of each data point (σ). $G(\tau)$ is fit to the equation for 3D diffusion, $G(\tau) = \dfrac{1}{N}\left(1+\dfrac{4D\tau}{r_0^{\,2}}\right)^{-1}\left(1+\dfrac{4D\tau}{z_0^{\,2}}\right)^{-\frac{1}{2}}$, with $1/\sigma^2$ weighting. N is the mean occupancy of donor-acceptor labeled molecules in the focal volume, D is the diffusion coefficient, and r_0 and z_0 are the characteristic 3D-Gaussian dimensions in the lateral and axial directions (131). The dimensions are calibrated by measurements of fluorophore solutions, for which D and concentration are known, i.e., TMR solutions. The resulting diffusion coefficients for the tetraloop–receptor RNA at 45 °C and 21 °C are $D = 44 \pm 2\ \mu m^2/s$ and $29 \pm 1\ \mu m^2/s$, respectively. An increase in the diffusion coefficient with temperature is expected from Stoke-Einstein diffusion. The mean occupancies of effective focal volume are 0.78 ± 0.01 and 0.65 ± 0.01 at 45 °C and 21 °C, respectively, ensuring that freely diffusing measurements are performed in the single molecule detection regime and that the hybridization of donor and acceptor labeled strands is maintained with heating.

Chapter 5   An Entropic Origin of Mg²⁺-Facilitated RNA Folding

5.1   Abstract

Mg^{2+} is essential for the proper folding and catalytic activity of RNA, though the effect of Mg^{2+} on the underlying free energy, enthalpy and entropy landscapes of RNA folding is unknown. This work exploits temperature-controlled single-molecule FRET methods to address the thermodynamics of RNA folding pathways by probing the intramolecular docking/undocking kinetics of the ubiquitous GAAA tetraloop–receptor tertiary interaction as a function of $[Mg^{2+}]$. From these measurements, we obtain the barrier and standard state enthalpies, entropies, and free energies of an elementary RNA folding transition, revealing the thermodynamic origin of $[Mg^{2+}]$-facilitated tertiary folding. The docking transition state is "early" or reactant-like (i.e., $\Delta H^{\ddagger}_{dock} \approx 0$) though rate-limited in formation by a large entropic barrier (i.e., $-T\Delta S^{\ddagger}_{dock} >> 0$). Surprisingly, these studies support that increasing Mg^{2+} concentration promotes tetraloop–receptor tertiary interaction by reducing the entropic barrier ($-T\Delta S^{\ddagger}_{dock}$) and the overall entropic penalty ($-T\Delta S^{o}_{dock}$) for docking, with essentially negligible effects on both the activation enthalpy ($\Delta H^{\ddagger}_{dock}$) and overall exothermicity (ΔH^{o}_{dock}). These observations contrast sharply with the common expectation that increasing $[Mg^{2+}]$ facilitates folding by minimizing electrostatic repulsion of opposing RNA helices, which would incorrectly predict a decrease in $\Delta H^{\ddagger}_{dock}$ and ΔH^{o}_{dock} with $[Mg^{2+}]$. Instead we propose that higher $[Mg^{2+}]$ can aid RNA folding by decreasing the entropic penalty of counterion uptake requisite for the tertiary folding transition and by reducing disorder of the unfolded conformational ensemble.

5.2   Introduction

The folding of RNA into a compact, biochemically competent structure proceeds hierarchically, whereby secondary helical structure is formed rapidly and subsequent slow helical packing is mediated by tertiary interactions, such as loops and bulges (6,36). RNA secondary structure prediction from the known thermodynamics is quite reliable (5), though correspondingly accurate prediction of tertiary structure remains a major challenge (6). Static tertiary structure data alone are also not enough, as time-dependent structural dynamics occurring during biochemical processes determine function (2,3). As a result, one needs the full free energy, enthalpy and entropy landscape for folding. A major road block in achieving a predictive understanding of RNA folding landscapes, in particular the energy barriers for folding, is that they are often "rugged", i.e., with alternative conformations acting as kinetic traps resulting in slow and/or heterogeneous folding and/or unfolding rates (99,100). Moreover, the electrostatic challenge of folding a charged biopolymer (7-11) highlights the particularly critical role of Mg^{2+} and other counterions in thermodynamically influencing the overall folding free energy landscape.

Characterization of folding transition states—and the role of Mg^{2+} in stabilizing transition states—remains a crucial bottleneck for reconciling the kinetics and thermodynamics of RNA folding (7,8,10,101-103). Some insight into the free energy landscapes for RNA folding can be obtained from temperature-dependent stop-flow kinetic studies, which in principle offer the ability to deconstruct free energy barriers ($\Delta G^{\ddagger}$) into enthalpic ($\Delta H^{\ddagger}$) and entropic ($-T\Delta S^{\ddagger}$) components. However, with ensemble methods, generally only the net rate constant (i.e., $k_{total} = k_{fold} + k_{unfold}$) for approach to equilibrium can be observed, which requires additional strong assumptions (e.g., that $k_{fold} \gg k_{unfold}$ or k_{unfold} is temperature independent) to permit accurate extraction of transition-state barrier heights (9,10,124,200). Single-molecule FRET methods avoid such kinetic restrictions by providing both folding and unfolding rate constants under equilibrium conditions, though smFRET transition-states studies of RNA folding are scarce (8,103,108). Furthermore, despite the well-known role of Mg^{2+} in promoting the structural assembly of RNA, there is remarkably limited information even from ensemble studies on enthalpic vs.

entropic contributions to Mg^{2+} stabilized tertiary structures (11) and absolutely no such studies (ensemble or single-molecule) addressing the origin of Mg^{2+}-accelerated folding.

In this work, we exploit temperature-controlled single-molecule FRET (smFRET) microscopy to explore the $[Mg^{2+}]$-dependent thermodynamics of RNA folding/unfolding by characterizing enthalpy and entropy changes associated with the elementary formation of an isolated tertiary interaction. Specifically, we measure the temperature dependence of the equilibrium and rate constants for intramolecular docking and undocking of a GAAA tetraloop with its 11 nucleotide receptor via a flexible U_7 linker as function of $[Mg^{2+}]$ (Figure 5.1 *A*). The tetraloop–receptor interaction is a ubiquitous modular motif (62,63,77). The structures of the docked and undocked forms of the tetraloop and receptor are known
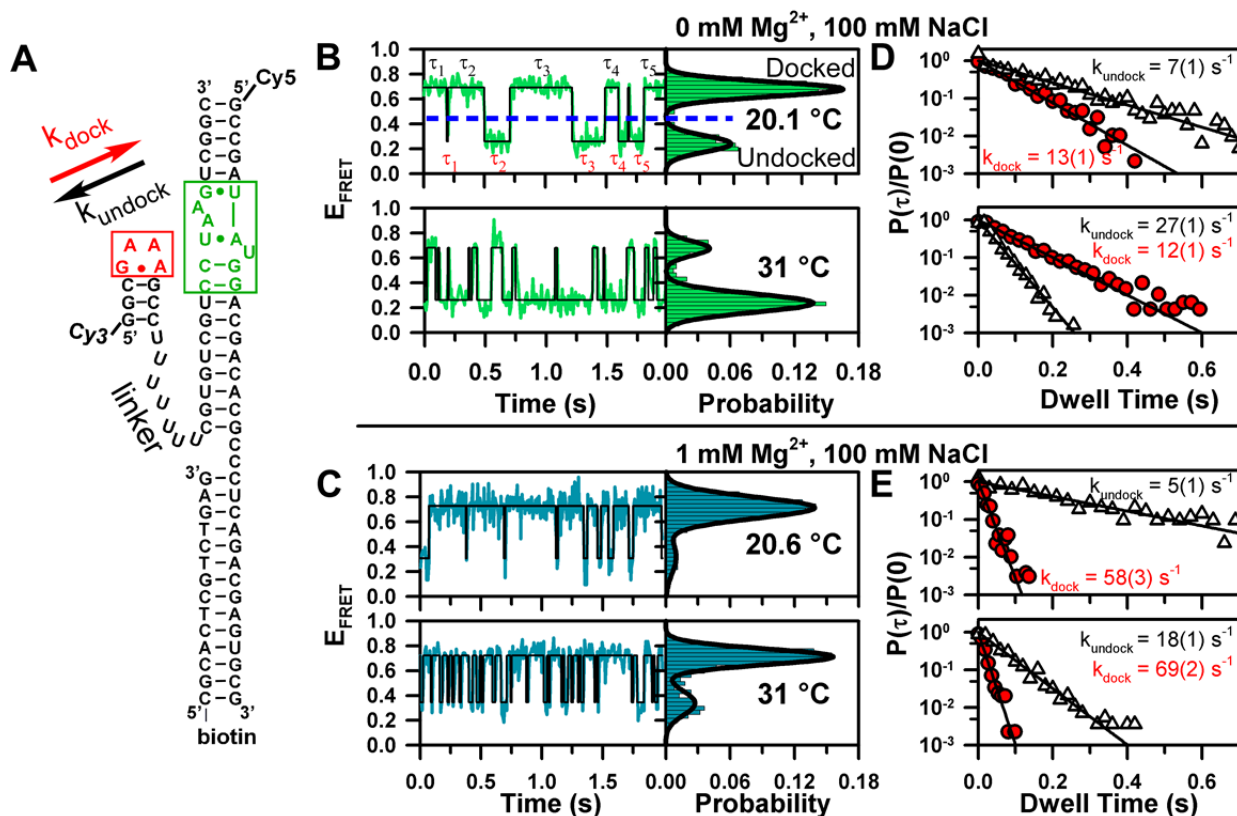
Figure 5.1   Single-molecule observation of intramolecular GAAA tetraloop and receptor docking and undocking. (*A*) Tetraloop–receptor (TL–R) construct in which docking/undocking are monitored by FRET between the donor (Cy3) and acceptor (Cy5), yielding rate constants, k_{dock} and k_{undock}. (*B-C*) Temperature-dependent single-molecule E_{FRET} trajectories and probability histograms 0 mM and 1 mM $MgCl_2$, respectively. *(D-E)* Dwell time probability densities from many molecules at varying $[Mg^{2+}]$ and temperatures yield k_{dock} and k_{undock} from single exponential fits of the undocked (*red filled circles*) and docked (*open triangles*) dwell times, respectively.

(51,62,73,74), allowing for direct correlation of structure with the observed thermodynamics (117,125).

By determining the activation enthalpies and entropies for TL–R docking and undocking, we show that

the free energy barrier ($\Delta G^{\ddagger}_{dock}$) for TL–R interaction is dominated by entropy ($-T\Delta S^{\ddagger}_{dock} \gg 0$) with an

"early" (i.e., $\Delta H^{\ddagger}_{dock} \approx 0$) transition state, while the overall reaction is extremely exothermic ($\Delta H^{o}_{dock} \ll$

0). These observations support a paradigm that RNA folding transition states lack significant hydrogen

bond formation (8,101-103). Most importantly, the results reveal that [Mg^{2+}]-based promotion of TL–R

docking is of an *entropic origin*, specifically by reduction of the entropic barrier ($-T\Delta S^{\ddagger}_{dock}$), and decrease

in the overall entropic cost of folding ($-T\Delta S^{o}_{dock}$). This is a fundamentally surprising result, in opposition

with the traditional notion of increased [Mg^{2+}] facilitating RNA folding via electrostatic screening of the

negatively charged phosphate backbone, which would incorrectly predict such an enhancement to be of

enthalpic origin—by decreasing $\Delta H^{\ddagger}_{dock}$ and ΔH^{o}_{dock} for the barrier and docked state, respectively.

5.3   Results

5.3.1 Mg^{2+} Increases the Melting Temperature of the Tetraloop–Receptor Interaction

Temperature and [Mg^{2+}]-dependent docking of an isolated GAAA tetraloop with its 11 nucleotide recep-

tor are explored using the RNA construct shown in Figure 5.1 *A*. Linked by a flexible single-stranded

poly(U) junction, the GAAA tetraloop (TL) facilely and specifically docks into its receptor (R), modulat-

ing the fluorescence resonance energy transfer (FRET) between the donor (Cy3) and acceptor (Cy5)

fluorophores (*Supporting Information* (*SI*) Figure 5.7) (106,124). The efficiency of energy transfer

(E_{FRET}) is monitored by single-molecule confocal microscopy—calculated ratiometrically from the donor

and acceptor emission intensities (see *Materials and Methods*) (106,128). Such real-time E_{FRET} traces

permit exploration of both kinetic and equilibrium properties for tertiary folding as the RNA vacillates

between two well-resolved states, high (docked) and low (undocked) E_{FRET} states (Figure 5.1 *B* and *C, SI*

Figure 5.2, and *SI* Figure 5.3). The U_7 linker behaves as a random coil at the moderate ionic strengths

and temperature ranges explored in this work (201), which therefore permits these studies to isolate

the thermodynamic effects of Mg^{2+} on the TL–R tertiary interaction alone, with the linker merely serving

to constrain the local concentration of the tetraloop and receptor (124). This represents a critical differ-ence over previous studies based on poly(A) linked constructs, which in principle could also be influ-enced by [Mg^{2+}] and temperature-dependent base-stacking in the A_7 junction (202,203).

Both [Mg^{2+}] and temperature affect the population distribution of the undocked and docked states of the TL–R construct as seen in the single-molecule E_{FRET} trajectories (Figure 5.1 *B* and *C*). Histo-grams of these time traces yield corresponding E_{FRET} probability distributions, which can be accurately described by two Gaussian distributions. Superimposed on the trajectories are Hidden Markov fits, clearly demonstrating molecular transitions between two distinguishable states (204). These data (and cumulative histograms of many molecules, *SI* Figure 5.8, and *SI* Figure 5.9) reveal the two central effects of [Mg^{2+}] and temperature on the TL–R docking equilibrium: (i) increasing [Mg^{2+}] favors the docked state, while (ii) increasing temperature favors the undocked state. In the absence of Mg^{2+}, increasing from 20.1 to 31 °C shifts the equilibrium from favoring the docked to the undocked state, while at 1 mM Mg^{2+} this same temperature excursion only modestly affects the population distribution. Thus, Mg^{2+} (i) enhances TL–R docking and (ii) increases the melting temperature of the interaction.

5.3.2 Dwell Time Analysis Yields Rate Constants and Equilibrium of Docking and Undocking (k_{dock}, k_{undock}, and K_{dock}) as a Function of [Mg^{2+}]

The TL–R tertiary interaction can be explored in more detail by determining k_{dock} and k_{undock} as a func-tion of [Mg^{2+}]. Dwell times for the TL–R construct in the docked and undocked states are defined in the real-time trajectory by crossings of a threshold set at the minimum of the bimodal E_{FRET} distribution (Figure 5.1 *B*) (106). To achieve a larger dynamic range and statistical accuracy, we calculate probabil-ity densities from cumulative histograms (~30 molecules and >300 transitions) of these docked and undocked dwell times under each experimental condition, i.e., $P(\tau_i) \approx H(\tau_i)/[0.5(\tau_{i+1} - \tau_{i-1})]$, where $H(\tau_i)$ is the standard histogram value and τ_i represents an ordered list of nonzero time bins (106). The result-ing normalized docked and undocked dwell time probability densities, $P(\tau)/P(0)$, are well described by a single-exponential decays over 3 orders of magnitude, confirming the validity of describing the kinetics as a two-state system and yielding high quality rate constants. Hidden Markov modeling is also pursued
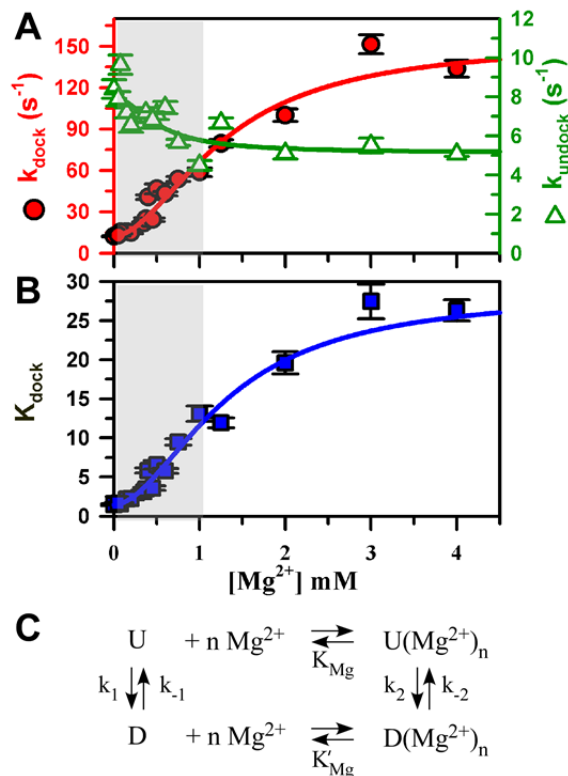
Figure 5.2   [Mg^{2+}]-dependence of the TL–R RNA docking via a U_7 linker: *(A)* k_{dock}, k_{undock} and *(B)* K_{dock} = k_{dock}/k_{undock} described by *(C)* a four-state kinetic model allowing for Mg^{2+} dependent and independent docking pathways (U = undocked, D = docked). The Mg^{2+} free and bound D and U states are indistinguishable by FRET. From this model the [Mg^{2+}]-dependence of k_{dock} and k_{undock} is k_{dock} = $\{k_1(K_{Mg})^n$ + $k_2[Mg^{2+})^n\}/\{(K_{Mg})^n$ + [Mg^{2+}). A simultaneous fit of the k_{dock} and k_{undock} titrations with the detailed balance constraint that K'_{Mg} = $(k_1 k_{-2}/(k_{-1}k_2))^{1/n}$ K_{Mg}, yields n = 1.8 ± 0.2, k_1 = 12.6 ± 0.9 s^{-1}, k_2 = 156 ± 23 s^{-1}, k_{-1} = 8.6 ± 0.7 s^{-1}, k_{-2} = 5.4 ± 0.2 s^{-1}, k_{Mg} =1.3 ± 0.3 mM, and K'_{Mg}, =0.23 ± 0.08 mM.

as method for determining rate constants and yields identical rate constants within uncertainties, as expected for well resolved states (204). Exponential fits of the normalized probability densities (Figure 5.1 *D* and *E*) reveal that k_{dock} is i) largely insensitive to temperature and *increases* with [Mg^{2+}], while k_{undock} ii) dramatically increases with temperature and *decreases* with [Mg^{2+}]. As developed more explicitly in the accompanying sections, the temperature insensitivity of the folding rate immediately implies that the pronounced lowering of $\Delta G^{\ddagger}_{dock}$ by Mg^{2+} is *entropic* rather than *enthalpic* in origin.

The [Mg^{2+}] dependence of k_{dock} and k_{undock} at 20 ± 1 °C is shown in Figure 5.2 *A*. Increasing [Mg^{2+}] dramatically accelerates k_{dock} (~12 fold) and slightly decelerates k_{undock} (~ 1.6 fold) , echoing similar trends observed at room temperature in an A_7 linked TL-R construct (106). For both k_{dock} to in-

crease and k_{undock} to decrease with [Mg^{2+}], the free energy for docking (ΔG°_{dock}) must drop more quickly with [Mg^{2+}] than the barrier for the forward docking reaction ($\Delta G^{\ddagger}_{dock}$). The combined effect on k_{dock} and k_{undock} results in a rapid increase in the docking equilibrium constant ($K_{dock} = k_{dock}/k_{undock}$) with [Mg^{2+}], as seen Figure 5.2 *B*. These trends can be fit to a four-state kinetic model with Mg^{2+}-dependent and -independent pathways for docking (Figure 5.2 *C*) (106,128,148). We now exploit the temperature dependence of k_{dock}, k_{undock}, and K_{dock} over the range where it has the most effect (0 to 1 mM Mg^2, highlighted in Figure 5.2) to explore the origin of Mg^{2+}-facilitated docking, i.e., how the enthalpy, entropy and free energy of the (i) overall folding and (ii) transition state are influenced by [Mg^{2+}].

5.3.3 Van't Hoff Analysis Yields Enthalpies (ΔH°_{dock}) and Entropies (ΔS°_{dock}) for Tetraloop–Receptor Docking as Function of [Mg^{2+}]

To explore how increasing [Mg^{2+}] stabilizes the overall TL−R interaction, we extract enthalpic and entropic information from the equilibrium constants ($\Delta G^{\circ}_{dock} = -RT \ln K_{dock}$) by employing the van't Hoff equation,

$$\ln K_{dock} = -\frac{\Delta H^{\circ}_{dock}}{RT} + \frac{\Delta S^{\circ}_{dock}}{R},$$ 5.1

where R is the ideal gas constant. From Eq. 5.1, a linear least squares fit of $\ln K_{dock}$ vs. 1/T yields a slope of $-\Delta H^{\circ}_{dock}/R$ and intercept of $\Delta S^{\circ}_{dock}/R$. Such an analysis assumes negligible temperature-dependent changes in ΔH°_{dock} and ΔS°_{dock}, which is supported by calorimetry measurements of a dual tetraloop−receptor construct over the temperature range investigated in this work (125).

The resulting van't Hoff plots (Figure 5.3) demonstrate that increasing [Mg^{2+}] has a negligible effect on the van't Hoff slope ($\Delta(\Delta H^{\circ}_{dock}) \approx 0$) while substantially increasing the offset ($\Delta(\Delta S^{\circ}_{dock} > 0$), as summarized in the top section of Table 5-1. Such behavior is reminiscent of salt-dependent formation of DNA duplexes (205). At all cation concentrations, TL−R docking is strongly exothermic and entropically disfavored. What is quite surprising, however, is that such an increase in folding propensity with Mg^{2+} is achieved exclusively by *entropic* rather than *enthalpic stabilization* of the docked RNA, i.e. $\Delta(-T\Delta S^{\circ}_{dock}) < 0$. This is in fundamental contrast with conventional expectations that increasing [Mg^{2+}] results in
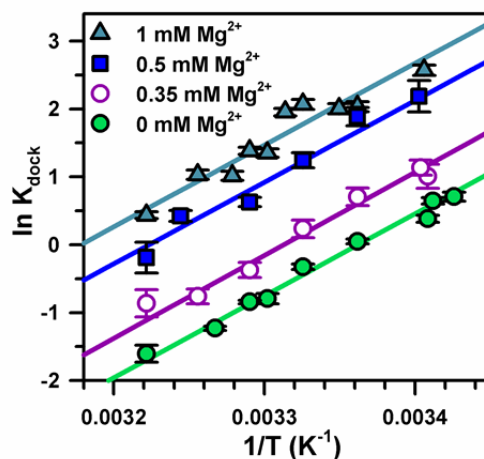
Figure 5.3 Temperature dependence (van't Hoff plot) of equilibrium constant (K_{dock}) for TL–R docking via a U_7 linker as a function of $[Mg^{2+}]$ at 100 mM NaCl yields standard state enthapies and entropies of docking (Table 5-1).

electrostatic stabilization of the folded RNA by better screening of the negatively charged helices, which

would incorrectly predict an increased exothermicity at higher $[Mg^{2+}]$.

Table 5-1 Thermodynamic parameters for tetraloop–receptor docking

[MgCl$_2$] mM	[NaCl] mM	$\Delta H°_{dock}$ kcal/mol	$\Delta S°_{dock}$ cal/mol	$\Delta G°_{dock}$ at 37 °C ($\Delta H°_{dock}$–T$\Delta S°_{dock}$) kcal/mol
TL–R Docking via U$_7$ linker				
0.00	100	−24.0 ± 0.5	−80.7 ± 1.7	1.0 ± 0.7
0.35	100	−24.3 ± 0.8	−80.3 ± 2.6	0.6 ± 0.7
0.50	100	−23.9 ± 0.9	−77.0 ± 3.2	−0.03 ± 1.3
1.00	100	−23.9 ± 0.8	−76.0 ± 2.6	−0.34 ± 1.1
1.00	25	−21.9 ± 1.2	−76.8 ± 4.8	1.9 ± 2.3
2.00	25	−16.0 ± 0.7	−50.7 ± 2.2	−0.3 ± 0.7
TL–R Docking via A$_7$ linker				
0.35	100	−25 ± 2	−84 ± 7	1.0 ± 3.0
0.50	100	−23 ± 1	−76 ± 5	0.2 ± 1.8
1.00[1]	100	−15 ± 1	−47 ± 4	−0.43 ± 1.6
2.00	100	−11 ± 1	−34 ± 5	−0.46 ± 1.8
TL–R Interaction in Bimolecular Construct[2]				
2.00	20	−15.1 ± 0.6	−	−

Unless indicated buffer also contains 50 mM hemisodium HEPES (pH 7.5) and 100 μM EDTA.
[1]This measurement was previously made, but is reported here with improved precision achieved by a large increase in the data set and improved temperature accuracy (117).
[2] 2 mM MgCl$_2$, 20 mM KCl, 20 mM HEPES, pH 7.0, 30 °C. The reported value is −30.1 ± 1.2 for bimolecular association of a dual tetraloop to a dual receptor construct (125).

5.3.4 Temperature Dependence of k_{dock} and k_{undock} Yields Activation Enthalpies and Entropies for Docking and Undocking

To further elucidate the thermodynamic origin of this increased folding stability with [Mg²⁺], we determine k_{dock} and k_{undock} as function of temperature. Arrhenius plots for k_{dock} and k_{undock} are shown in Figure 5.4, revealing a steep increase in k_{undock} (Figure 5.4 *B*) and yet only a very slight decrease in k_{dock} (Figure 5.4 *A*) with temperature. As in the van't Hoff plots (Figure 5.3), the *slopes* of these Arrhenius plots are independent of [Mg²⁺], while the *offsets* increase. To extract the activation enthalpies and entropies from the Arrhenius plots, we invoke a transition state analysis.
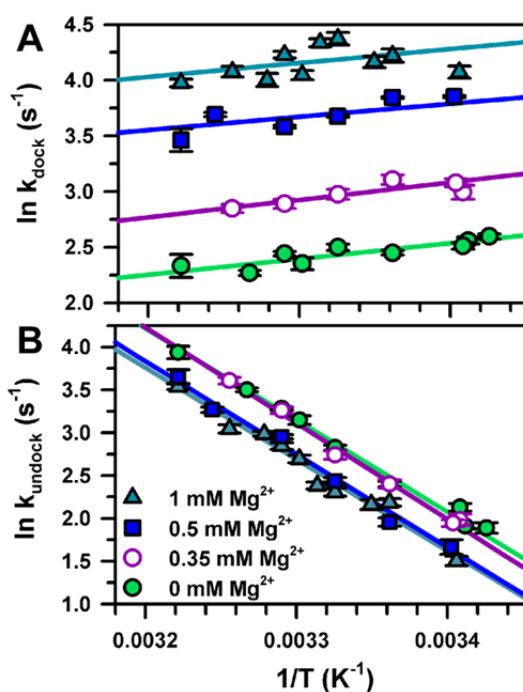
Figure 5.4   Temperature dependence k_{dock} and k_{undock} as function of [Mg²⁺]. Transition-state analysis yields activation enthalpies ($\Delta H^{\ddagger}$) and entropies ($\Delta S^{\ddagger}$) for docking/undocking dynamics from linear least squares fits of ln(k_{dock}) and ln(k_{undock}) vs $1/T$ (Eq. 5.3), summarized in Table 5-2.

From generalized transition-state theory, the reaction rate constant (e.g., k_{dock} or k_{undock}) can be written as,

$$k = ve^{(-\Delta G^{\ddagger}/RT)},$$

5.2

where $\Delta G^{\ddagger}$ is the activation free energy and ν is the attempt frequency for escape from the reactant energy well (see *SI 5.7.1*) (206,207). Thus, we can rewrite the reaction rate constant in terms of the enthalpic and entropic components as,

$$\ln k = \ln \nu + \frac{\Delta S^{\ddagger}}{R} - \frac{\Delta H^{\ddagger}}{RT},$$ 5.3

where $\Delta S^{\ddagger}$ and $\Delta H^{\ddagger}$ are the activation entropy and enthalpy obtainable from the slope and intercept of linear least-squares fits of $\ln k$ vs. $1/T$ plots. Extraction of the activation enthalpy ($\Delta H^{\ddagger}$) is unambiguous, while determination of the absolute value of $\Delta S^{\ddagger}$ from the experimental intercepts requires some knowledge of ν. However, since the dependence of the reaction rate on this parameter is only logarithmic, an estimate of $\upsilon \approx 10^{13}$ s^{-1} proves sufficient for our purposes, based on typical frequencies (~ 300 cm^{-1}) for low frequency skeletal motions (see *SI 5.7.1*) (208,209). The measured dependence of $\Delta S^{\ddagger}$ on [Mg^{2+}] (i.e., $\Delta(-T\Delta S^{\ddagger})$) are completely independent of ν, and therefore obtained rigorously from experiment.

Least squares fits of the Arrhenius plots in Figure 5.4 to Eq. 5.3 yield some striking observations, with the results summarized in the top section of Table 5-2. First of all, the slopes for $\ln k_{dock}$ vs. $1/T$ are small and *positive*, which indicates that achieving the transition state has no enthalpic barrier and in fact is a weakly *exothermic* process ($\Delta H^{\ddagger}_{dock} < 0$). However, a considerable activation entropy ($-T\Delta S^{\ddagger}_{dock} \gg 0$), results in a large free energy barrier ($\Delta G^{\ddagger}_{dock} \sim 16$ kcal/mol at 37 °C, Table 5-2) that slows folding more than 10 orders of magnitude below ν. Conversely, undocking is limited by a large enthalpic barrier ($\Delta H^{\ddagger}_{undock} \gg 0$) with a favorable entropy gain ($-T\Delta S^{\ddagger}_{undock} < 0$) in achieving the transition state. Most importantly, the effect of increasing [Mg^{2+}] is, once again, not by any change in *transition state enthalpy* ($\Delta\Delta H^{\ddagger}_{dock} \approx 0$) but rather by *increasing the entropy* of the transition state relative to the unfolded conformational ensemble (i.e., $\Delta\Delta S^{\ddagger}_{dock} > 0$) (Figure 5.4 *A* and Table 5-2). Similarly, increasing [Mg^{2+}] appears to decrease k$_{undock}$, not by any change in barrier enthalpy, but rather by reducing the entropic reward for achieving the transition state barrier. These trends are underscored by the constant slopes, yet

increasing intercepts for the Arrhenius plots (Figure 5.4 *B*), and quantitatively by the *decrease* in

$\Delta S^{\ddagger}_{undock}$ at 0 vs 1 mM Mg^{2+} as summarized in Table 5-2.

Table 5-2   Transition-state thermodynamics for tetraloop–receptor docking/undocking

[Mg^{2+}] mM	[NaCl] mM	$\Delta H^{\ddagger}_{dock}$ kcal/mol	$\Delta S^{\ddagger}_{dock}$ cal/mol/K	$\Delta G^{\ddagger}_{dock}$ (37 °C) cal/mol/K	$\Delta H^{\ddagger}_{undock}$ kcal/mol	$\Delta S^{\ddagger}_{undock}$ cal/mol/K	$\Delta G^{\ddagger}_{undock}$ (37 °C) cal/mol/K
U_7 linker							
0	100	−2.9 ± 0.3	−64.0 ± 1.0	17.0 ± 0.4	21.2 ± 0.4	17 ± 1	15.9 ± 0.5
0.35	100	−3.1 ± 0.6	−63.9 ± 2.0	16.7 ± 0.9	22.2 ± 0.6	20 ± 2	16.0 ± 0.8
0.5	100	−2.3 ± 0.7	−59.9 ± 2.1	16.3 ± 1.0	21.6 ± 0.7	17 ± 2	16.3 ± 0.9
1.0	100	−2.5 ± 0.6	−59.4 ± 1.8	15.9 ± 0.8	21.3 ± 0.5	16 ± 2	16.3 ± 0.8
1.0	25	0.03 ± 0.90	−56.5 ± 3.0	17.5 ± 1.3	21.2 ± 0.8	18 ± 3	15.6 ± 1.2
2.0	25	0.8 ± 0.5	−49.4 ± 1.5	16.1 ± 0.7	17.1 ± 0.5	3 ± 2	16.2 ± 0.8
A_7 linker							
0.35	100	−6.5 ± 4.2	−76 ± 14	17.1 ± 6.0	24.0 ± 3.7	25 ± 12	16.2 ± 5.2
0.5	100	−1.6 ± 1.6	−59 ± 5	16.7 ± 2.2	21.4 ± 0.7	17 ± 3	16.1 ± 1.1
1.0	100	2.0 ± 1.2	−46 ± 4	16.3 ± 1.7	16.7 ± 0.8	1 ± 3	16.4 ± 1.2
2.0	100	7.4 ± 0.5	−27 ± 2	15.8 ± 0.8	19.2 ± 1.3	9 ± 4	16.4 ± 1.8

Uncertainties are the standard deviation of the fits (Eq. 5.3). Uncertainty in $\Delta S^{\ddagger}$ is logarithmically dominated by imprecise knowledge of the attempt frequency (ν). Estimates of $\Delta S^{\ddagger}$ from Eq. 5.3 are based on typical value low frequency vibrational values of $\nu \approx 10^{13}$ s^{-1} (see text). Note, however, that any [Mg^{2+}]-dependent *changes* in $\Delta S^{\ddagger}$ ($\Delta\Delta S^{\ddagger}$) are obtained with full experimental accuracy.

5.4   Discussion

5.4.1   "Early" Transition States as a Paradigm for RNA Folding Pathways

Characterizing RNA folding transition states has been difficult due the ruggedness of folding landscapes,

i.e., stable misfolded intermediates (9,103,140). The structural simplicity of the isolated TL–R interac-

tion eliminates any such kinetic traps and therefore permits cleaner interpretation of free energy barrier

contributions. The absence of an enthalpic barrier (Figure 5.5 *A*, Table 5-1, and Table 5-2) is indicative

of an "early" or reactant-like transition state, where enthalpic rearrangements such as hydrogen bond-

ing between the tetraloop and receptor (62,74) are largely unformed. This interpretation can be further

corroborated by Φ-analysis, whereby mutational effects on equilibrium/rate constants for folding help

determine which interactions are formed in the transition state. Φ-analyses of the P4–P6 domain fold-

ing (101,126) and P1 substrate docking in the *Tetrahymena* ribozyme (103), tRNA unfolding (102), and

folding of the hairpin ribozyme (8) have concluded that tertiary interactions are largely unformed in the
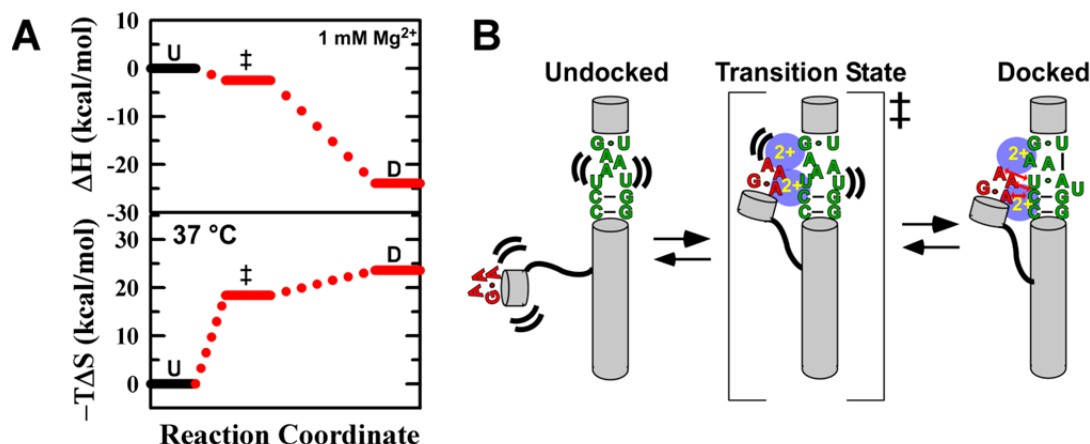
Figure 5.5   Proposed schematic mechanism for Mg^{2+}-facilitated TL–R folding. (*A*) The entropic and enthalpic reaction coordinate for TL–R docking, where U, ‡, and D indicate the undocked, transition, and docked states. (*B*) The transition state is early and "compact", i.e., requiring entropically disfavored proximity of the tetraloop and receptor and localization of counterions (e.g., Mg^{2++}, *blue circles*), yet hydrogen bonding and base-stacking interactions in the docked state (*red lines*) are largely unformed. Docking increases the charge density of the RNA, permitting further Mg^{2+} localization.

transition states. Mutational studies in the P4–P6 domain indicate that the TL-R contributes only to

overall thermodynamic stability of the folded domain rather than the transition state (101,126), in

agreement with this work. Indeed the TL-R interaction (Table 5-1) can account for the entire exother-

micity of the P4–P6 domain ($\Delta H^{\circ}_{P4–P6}$ = −25 ± 3 kcal/mol at 1.1. mM $MgCl_2$, 10 mM NaCl) (11). Ensem-

ble studies suggest that folding of the hairpin ribozyme also lacks an enthalpic barrier, again in support

of an early transition state (200). These studies together lend strong support to an emerging paradigm

that *early transition states* are a characteristic property of RNA folding (8).

        We consider now the origin of the large entropic barrier for folding and its remarkable depend-

ence on [Mg^{2+}]. One simple explanation would be that the TL–R transition state is "compact" (i.e., te-

traloop and receptor in close proximity/alignment), the achievement of which is limited by conforma-

tional search within the radius of the U_7 linker (Figure 5.1 *A* and Figure 5.5). Support for a compact

transition state can be obtained from studies of an alternative linker, e.g., A_7 vs. U_7 (Figure 5.6 *A*). In con-

trast to poly U, single-stranded poly A has a propensity for [Mg^{2+}]-dependent helix formation, which can

increase order in the unfolded RNA (201,202). As shown in the Arrhenius (Figure 5.6 *A*, sample data
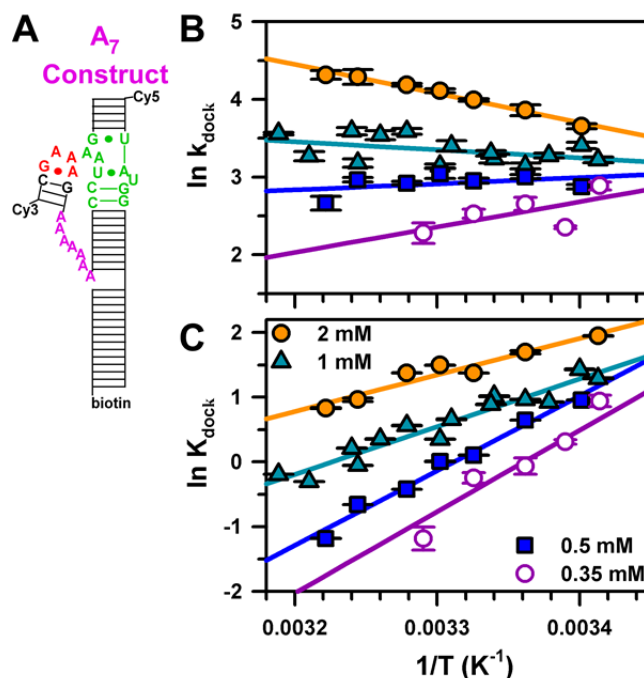
Figure 5.6   Thermodynamic analysis of TL–R docking in an A$_7$ linked construct. *(A)* A$_7$ TL–R construct. (*B*) Temperature dependence of k_{dock} and (*C*) K_{dock} as function [Mg^{2+}] yielding standard state and transition state enthalpies as summarized in Table 5-1 and Table 5-2. Sample data and analysis are shown in *SI* Figure 5.10.

shown in *SI* Figure 5.10) and van't Hoff (Figure 5.6*B*,) plots for docking, the A$_7$ construct displays a much greater range of intercepts than the U$_7$ construct (Figure 5.3 and Figure 5.4)—the entropic barrier decreases substantially with [Mg^{2+}] (Table 5-2). This trend supports the notion that increasing [Mg^{2+}] increases order in the undocked state, i.e., limits the conformational search of the tetraloop for the receptor by encouraging base-stacking interactions in the A$_7$ linker. This ordering is accompanied by an enthalpic penalty (Figure 5.6 and Table 5-2), which supports our assertion that the tetraloop must gain proximity to the receptor domain in the transition state, which can only be achieved by stretching the A$_7$ junction and breaking A–A base stacking (202).

In the simplest model, search for a compact oriented transition state is rate-limited by diffusion (209). Measurements and estimates for the end-to-end contact formation in similarly sized polymer systems would predict this rate to be ≈ 10^7 s^{-1} (see *SI 5.7.2*) (210-212), 5 orders of magnitude faster than the observed TL–R docking rate at 1 mm Mg^{2+} (Figure 5.2). Based on a diffusion-controlled rate of 10^7

s^{-1} and an attempt frequency of $\nu \approx 10^{13}$ s^{-1}, Eq. 5.3 can be rearranged to estimate a diffusional contribution to the transition state entropy of $\Delta S^{\ddagger}_{diffusion} \approx -27$ cal/mol/K, which is less than half of the experimental value of $\Delta S^{\ddagger}_{dock} = -59.4 \pm 1.8$ cal/mol/K at 1 mM Mg^{2+} (Table 5-2). Clearly, there must be additional entropy sources beyond simple diffusion of the tetraloop to the receptor, which we consider below.

The first source is a possible organization of the tetraloop and receptor units in the undocked vs transition state conformation. The tetraloop by itself is already rigidly structured (50,62,74) and therefore not anticipated to contribute to loss of entropy in the transition state. The free receptor, by way of contrast, is considerably less organized, and must undergo rearrangement upon docking (Figure 5.5 *B*) (74,125). The slight exothermicity of $\Delta H^{\ddagger}_{dock}$ (Table 5-2), suggests that some tertiary contact is formed, such as a hydrogen bond, which would be accompanied by a small entropic cost. This scenario is consistent with the proposal that the free receptor becomes more "rigid" in response to the tetraloop (74), which could contribute to a small increase in the barrier ($\Delta S^{\ddagger}_{receptor}$) (74,117,125). However, the lack of significant hydrogen bonding in the transition state ($\Delta H^{\ddagger}_{dock} \sim 0$ and $\Delta H^{o}_{dock} << 0$) suggests that the receptor is still quite dynamic, and therefore does not dominate entropy loss in the transition state. Procession to the docked state from the transition state along the reaction coordinate is accompanied by an additional entropic cost ($|\Delta S^{\ddagger}_{dock}| < |\Delta S^{o}_{dock}|$), consistent with the exothermic formation of the hydrogen bonded tertiary interaction and concomitant loss of entropy in the receptor structure (Figure 5.5 *A*) (117,125).

A second source of entropy decrease can arise from an uptake of Mg^{2+} and/or Na^{+} in the transition state. Folding increases the negative charge density of RNA, and is therefore frequently accompanied by an uptake of counterions (8,174). Formation of a compact transition state (Figure 5.5) requires an interface of cations to electrostatically shield the closely packed negatively charged helices (8), otherwise an enthalpic barrier would be observed. For an ideal solution, the entropic cost of localizing Mg^{2+} on the RNA can be simply estimated from

$$\Delta S^{\ddagger}_{ion\ uptake} = -n\ R\ \ln\left(\frac{[Mg^{2+}]_{ion\ atmosphere}}{[Mg^{2+}]_{bulk}}\right),\qquad\qquad 5.4$$

where R is the gas constant, n is the number ions taken up, with $[Mg^{2+}]_{ion\ atmosphere}$ and $[Mg^{2+}]_{bulk}$ representing $[Mg^{2+}]$ close in and far away from the RNA, respectively (213,214). $[Mg^{2+}]_{ion\ atmosphere}$ can be estimated by assuming that every phosphate charge is effectively neutralized by counterions at 1 mM Mg^{2+}, as supported by studies of the local ion atmosphere on DNA duplexes (176). To illustrate that the penalty of counterion uptake can be quite large in the transition, we make a simple calculation. Treating the volume as helical rod with dimensions of 100 by 26 Å around the 84 nucleotide RNA construct, this predicts $[Mg^{2+}]_{ion\ atmosphere} \approx 1.3$ M. This value is consistent with prediction of ~ 2 M ion atmosphere of monovalent cations on RNA helices in the absence of Mg^{2+} (213). From the analysis of k_{dock} (Figure 5.2), the Hill coefficient (n) of 1.8 ± 0.2 can be utilized to estimate that ~1.8 Mg^{2+} are taken up in the transition state (215). For this uptake and concentration gradient, Eq. 4 predicts an entropy loss of $\Delta S^{\ddagger}_{ion\ uptake}$ ≈ -26 cal/mol/K at 1 mM Mg^{2+}. Thus, the entropy from diffusion and cation uptake ($\Delta S^{\ddagger}_{ion\ uptake}$ + $\Delta S^{\ddagger}_{diffusion} \approx -55$ cal/mol/K) is on the order of the experimentally observed $\Delta S^{\ddagger}_{dock}$ of -59.4 ± 1.8 cal/mol/K at 1 mM Mg^{2+} (Table 5-2), even neglecting contributions from $\Delta S^{\ddagger}_{receptor}$. Thereby, as depicted schematically depicted in Figure 5.5 *B*, (i) cation localization and (ii) intramolecular diffusion of the tetraloop to the receptor are the likely candidates for the large *entropic* rather than *enthalpic* barrier to docking.

5.4.2 An Entropic Origin of $[Mg^{2+}]$-Facilitated RNA Folding

We return to the focus issue of how $[Mg^{2+}]$ facilitates RNA folding. Dissecting the free energy of the TL–R docking reaction coordinate (ΔG^{o} and $\Delta G^{\ddagger}$) into the enthalpic and entropic components (i.e., ΔH^{o}, $\Delta H^{\ddagger}$, ΔS^{o}, and $\Delta S^{\ddagger}$) as a function of $[Mg^{2+}]$ provides the relevant information (see Table 5-1 and Table 5-2). Simply summarized, the overall docking reaction is exothermic and yet entropically costly, which can be attributed to a dominant contribution of tertiary hydrogen bonding and base stacking (117,216). The free energy barrier ($\Delta G^{\ddagger}_{dock} = 15.9$ ± 0.8 kcal/mol at 1 mM $MgCl_2$, 100 mM NaCl, 37 °C) for achieving the

transition state is quite large and dominated by entropy. Increasing [Mg^{2+}] increases the docking rate constant and the overall equilibrium constant (Figure 5.2) by a corresponding *reduction* in the entropic barrier ($\Delta(-T\Delta S^{\ddagger}_{dock}) < 0$) and entropic cost of docking ($-\Delta(T\Delta S^{o}_{dock}) < 0$), while $\Delta H^{\ddagger}_{dock}$ and ΔH^{o}_{dock} remain relatively unchanged (top sections, Table 5-1 and Table 5-2).

One might have anticipated that *enhanced electrostatic screening* with increasing [Mg^{2+}] to be a relevant contributor to [Mg^{2+}]-assisted folding mechanism. However, such a mechanism would require a [Mg^{2+}]-dependent *decrease* in the *enthalpy* for docking, which is not observed. This result is surprising and motivated our further exploration in an even less charge screened environment by reducing the background [NaCl] from 100 mM to 25 mM, at which charge repulsion should be enhanced (62,128). In fact, we do see the correct sign of the effect, i.e., the enthalpy contribution to the barrier height does increase from weakly exothermic ($\Delta H^{\ddagger}_{dock}$ = −2.5 kcal/mol) to thermoneutral ($\Delta H^{\ddagger}_{dock} \approx 0$ kcal/mol) at fixed 1mM Mg^{2+} with reduction in [NaCl] (Table 5-1, Table 5-2, *SI* Figure 5.11). However, *change* in the enthalpic barrier with increasing [Mg^{2+}] is ~0. More importantly, the entropic effects in Mg^{2+} facilitated RNA folding become even more pronounced (see Table 5-1 and Table 5-2). The 25 mM NaCl, 2 mM Mg^{2+} conditions are most similar to calorimetry studies by Butcher and coworkers for bimolecular association of "dual" tetraloop−receptor constructs (ΔH^{o}_{dock} = −15.1 ± 0.6 kcal/mol), which is in excellent quantitative agreement with our results (ΔH^{o}_{dock} = −16 ± 1 kcal/mol).

Therefore, [Mg^{2+}]-facilitated TL−R docking is of an entropic origin even at very low ionic strength. Thus, we consider possible mechanisms for this entropic effect. As mentioned above, from Eq. 5.4, one can see that increasing [Mg^{2+}]$_{bulk}$ decreases the entropic penalty of counterion uptake. Moreover increasing [Mg^{2+}]$_{bulk}$ could increase the [Mg^{2+}]$_{ion}$ atmosphere, which would decrease the number of cations (*n*) that need to be taken up with folding. To show that this effect can easily account for the magnitude of the $\Delta\Delta S^{\ddagger}_{dock}$ observed with increasing [Mg^{2+}], we again estimate that ~1.8 Mg^{2+} ion are taken up with docking. The favorable entropy gain for this uptake in higher salt (e.g. 0.35 vs. 1 mM Mg^{2+}) yields $\Delta\Delta S^{\ddagger}_{uptake} \sim 4$ cal/mol/k, which is sufficient to account for the observed $\Delta\Delta S^{\ddagger}_{dock}$ (Table 5-2).

Other effects of increasing [Mg^{2+}] that would be entropically beneficial to the docking pathway would be compaction of the unfolded RNA or stabilization of the tetraloop receptor. Compaction of the RNA would increase the sampling rate of the native state (217). The undocked E_{FRET} of the TL–R construct (Figure 5.1and *SI* Figure 5.8) shifts with [Mg^{2+}], which indicates compaction. However, the Mg^{2+} induced E_{FRET} shift only corresponds to an ~1.2 fold increase in the diffusion-controlled collision frequency of the tetraloop and receptor (see *SI* 5.7.2), which is negligible in comparison to the 12 fold increase in k_{dock} over the same [Mg^{2+}]. The undocked E_{FRET} peak does not broaden with increased [Mg^{2+}] (*SI* Figure 5.8), thus there is also no evidence for Mg^{2+}-slowed interconversions amongst conformations in the unfolded ensemble. Stabilization of the free receptor with increasing [Mg^{2+}] is also not probable because even at high (125 mM Mg^{2+}), the native (bound) form of the receptor is undetectable (75). Furthermore, stabilization of a non-native receptor structure should lead to an increase in the enthalpy of docking. Thus, we propose that the dominant mechanism for increased [Mg^{2+}] facilitating folding in the U$_7$ TL–R RNA (Figure 5.1 *A*) is a decreased entropic penalty of the counterion localization.

In summary, the above analysis reveals the entropic cost of intramolecular diffusion ($\Delta S^{\ddagger}_{diffusion}$) in the TL–R RNA accounts for nearly half the overall entropic penalty for achieving the transition state. This contribution is not an intrinsic property of the TL–R interaction, as demonstrated by comparison of a U$_7$ vs A$_7$ connected construct. We argue that the remaining portion of this entropic barrier comes from a combination of (i) need for additional ion uptake into a "compact" TL–R transition state ($\Delta S^{\ddagger}_{ion\ uptake}$) and (ii) ordering of the receptor ($\Delta S^{\ddagger}_{receptor}$), which are intrinsic properties of the tertiary interaction. We also conclude that the major source of [Mg^{2+}]-facilitated docking for the isolated TL–R is reduced entropic penalty for counterion uptake. However, in natural RNAs, organization of the unfolded RNA, e.g., stabilization of linking regions (junctions), may be a more dominant and universal mechanism of [Mg^{2+}]-facilitated folding. For example, the [Mg^{2+}]-dependent decrease of entropic cost for TL–R docking is much more pronounced in the A$_7$ vs U$_7$ constructs (Table 5-1 and Table 5-2)—suggesting that Mg^{2+}-dependent rigidification of the linker is of large entropic benefit. Rigidification of the junction in the hairpin ribozyme also resulted in a considerable decrease in the entropic penalty of folding (185). It has

also been shown that junctions can limit the conformational accessibility (or entropy) of unfolded states (218,219). This work clearly reveals that [Mg^{2+}]-mediated formation of an isolated TL–R tertiary interaction is of an *entropic origin* and involves a significantly complex interplay between the ion atmosphere and the docked vs. undocked RNA structures along the folding pathway.

5.5   Materials and Methods

5.5.1   RNA Preparation

Synthetic Cy3-Cy5-labeled TL–R RNA constructs (Figure 5.1) are prepared and immobilized on the glass surface of a microfluidic flow cell by biotin-streptavidin chemistry as previously described (106,124). Unless otherwise specified, experiments are performed in 50 mM hemisodium HEPES buffer (pH 7.5 at 25 °C) with 100 mM NaCl, 0.1 mM EDTA, and an oxygen scavenger of 60 nM protocatechuic acid, 5 mM protocatechuate-3,4-dioxygenase, and 2 mM Trolox (130) with the specified [Mg^{2+}] from added $MgCl_2$. As reported previously, there is kinetic heterogeneity in the TL-R constructs under all ionic strength and temperature conditions; experiments are performed on the majority (∼70 %) population of actively docking species (106,124,128). Experiments are performed with a background of 100 mM NaCl (unless otherwise noted) to relax secondary structure (e.g., rigid, extended helices) (128,220).

5.5.2   Temperature Controlled Single-Molecule FRET Measurements

Emission from the donor (Cy3) and acceptor (Cy5) emission is spectrally separated for time-correlated single-photon counting detection using an inverted confocal microscope system (106,128) with temperature control (see *SI* 5.7.3) . Fluorescence trajectories of the donor and acceptor signal from single molecules are binned at 3–10 ms integration to clearly resolve undocking/docking. The FRET efficiency, E_{FRET}, is calculated ratiometrically from the acceptor and donor signals [$E_{FRET} = I_A/(I_A + \gamma I_D)$], where γ is the quantum yield ratio of Cy5 to Cy3 and I_A and I_D are the acceptor and donor fluorescence intensities corrected for background, direct laser excitation of the acceptor, and collection efficiencies/crosstalk (106,128).

5.7   Supporting Information

5.7.1   Transition-State Analysis

From generalized transition-state theory, the reaction rate constant (e.g., k_{dock} or k_{undock}) can be written as

$$k = \frac{k_B T}{h} \frac{q_\perp^{\ddagger}}{q_{ll}^R q_\perp^R} e^{(-\Delta U^{\ddagger}/RT)}, \qquad\qquad 5.5$$

where q_{ll}^R are $q_\perp^R$ are the partition functions of the reactant parallel and perpendicular to the reaction coordinate, $q_\perp^{\ddagger}$ is the partition function of the transition state perpendicular to the reaction coordinate, k_B is the Boltzmann constant, h is Planck's constant, T is the temperature in Kelvin, R is the gas constant and $\Delta U^{\ddagger}$ is the potential energy barrier, which due to the fact that pV work is negligible, is simply $\Delta H^{\ddagger}$, or the activation enthalpy. Assuming to reaction energy well is harmonic, in the high temperature (semi-classical limit), $q_{ll}^R = 2\pi\, k_B T/h\omega$, where ω is the angular attempt frequency for escape for the well and $\omega/2\pi = \upsilon$, which we define as the attempt frequency. Therefore, Eq. 5.5 reduces to

$$k = v \frac{q_\perp^{\ddagger}}{q_\perp^R} e^{(-\Delta H^{\ddagger}/RT)} = v e^{\Delta S^{\ddagger}/R} e^{(-\Delta H^{\ddagger}/RT)} = v e^{(-\Delta G^{\ddagger}/RT)}, \qquad\qquad 5.6$$

where ΔG‡ is the activation free energy, ΔS‡ the activation entropy , and ΔH‡ is the activation enthalpy (206,207). Rewriting this transition-state equation in logarithmic form generates

$$\ln k = \ln \nu + \frac{\Delta S^{\ddagger}}{R} - \frac{\Delta H^{\ddagger}}{RT}.$$

5.7

The celebrated Kramer's theory introduces an additional factor κ, the transmission coefficient (κ ≤ 1), to scale the prexponential factor υ to account for because not every achievement of the transition state results in product formation in a viscous medium (207,208,221). Determination of the free energy barrier height can still be reasonably made without inclusion of this parameter and it has no effect the experimentally determined transition state enthalpies (208). A reasonable estimate for the pre-exponential factor is 10^{13} s^{-1} based on low frequency vibrations of the polymer skeleton (96,222) and is conveniently on similar to the often assumed Eyring prefactor $k_B T/h$, where h is Planck's constant, yielding 6.6 X 10^{12} s^{-1}. The pre-exponential is most easily envisioned by the undocking of the tetraloop and receptor, where the attempt frequency of for escape from this free energy minimum would be based on the molecular vibrations that coordinate with breaking of the hydrogen bonds defining the energy minimum (208). Note that the diffusion-controlled (enthalpically barrierless) folding rates are much slower than the attempt frequency, highlighting the importance of entropic barriers in folding problems (208)—the pre-exponential factor should not be confused with the rate for intramolecular diffusion, which is coupled into the entropic barrier for a transition. Kramers' theory and experiment tend to agree with estimates for diffusional controlled polymer folding rates of 10^6–10^7 s^{-1} for collapse of large RNAs, end to end diffusion times in polypeptides, and base-pairing-interactions (209-212,222).

5.7.2 Diffusion-Controlled Rate of Intramolecular Tetraloop−Receptor Collision

The diffusion controlled time until collision of the tetraloop and receptor can be estimated from the volume constraints imposed by the flexible/passive U_7 linker (124) and estimates of the diffusion coefficient of the tetraloop as spherical from Stokes-Einstein diffusion theory. The translation diffusion coefficient, D_{trans}, of spherical object of radius r, in a viscous medium can be estimated as

$$D_{trans} = \frac{k_B T}{6 \pi \eta\, r}\,,$$

<div align="right">5.8</div>

where k_B is the Boltzmann constant, T is temperature in Kelvin, and η is the viscosity of the solution (0.89 X 10^{-3} kg/m/s for water at room temperature). We assume that the tetraloop is a sphere with the diameter of a standard A-form helix (26 Å) or 13 Å radius. If this tetraloop sphere makes a random walk in 3D to the receptor a distance x away, the standard statistical mechanics diffusion time (τ_{trans}) is

$$\tau_{trans} = x^2 / 6D\,.$$

<div align="right">5.9</div>

The average distance between tetraloop and receptor can be calculated from the local concentration of the two domains as constrained by the linker. To assess this concentration, we realize that the linker (Figure 5.1 *A*) constrains the volume of the tetraloop and receptor to a sphere of radius ~33 Å (11 bp x 3 Å/bp) minus an excluded volume from the receptor helix (diameter 26 Å and height of 2 x 33 Å) or ~14 mM. This local concentration of tetraloop and receptor corresponds to an average distance (x) of 49 Å between the tetraloop and receptor domains. Substituting these parameters into Eqs. 5.8 and 5.9 yields τ_{trans} ~14– 25 ns at 313 and 293 K. Viewing this diffusion time as the "rate" of collision between and undocked tetraloop to a receptor, or ($1/\tau_{trans}$) corresponds to $k_{diffusion}$ ~ 4-7 x 10^7 s-10^7 s^{-1} for 293–313 K.

In the transition state the tetraloop may also need to gain a specific orientation with respect to the receptor. Thus, we also consider the time scale of tetraloop rotational diffusion. Similar to translation, Stokes-Einstein theory (223) predicts the rotational diffusion constant (D_{rot}) to be:

$$D_{rot} = \frac{k_B T}{8 \pi \eta\, r^3}\,.$$

<div align="right">5.10</div>

with the variable definitions the same as in Eq. 4. The time to make one rotation (τ_{rot}) is:

$$\tau_{rot} = 1/6D\,.$$

<div align="right">5.11</div>

Using again a tetraloop radius (r) of 13 Å, one obtains τ_{rot} = 1.5–2.3 ns for 313–293 K. This value agrees well with the rotation diffusional time measured for a similarly sized hairpin loop-stem of 0.6 ns (224).

Rhodamine 6G (r = 5.6 Å) was measured to have τ_{rot} = 200 ps (225) and bovine serum albumin BSA (r = 34 Å) has τ_{rot} = 45 ns (223). Using the relation that τ_{rot} scales with r^3 (Eq. 4), we can also use these known values to estimate τ_{rot} for a 13 Å object. Both measurements would predict τ_{rot} = 2.5 ns for the tetraloop, in excellent agreement with the prediction from Eqs. 5.10 and 5.11.

From these estimates we can see that rotational diffusion is on a much faster (>10-fold) time scale the translational diffusion. Thus, folding rates of 10^7 s^{-1}, as observed for end to end contact rates in similarly sized polymers systems (210-212), is a reasonable estimate for the rate at which tetraloop localizes near the receptor. The barrier entropy depends only logarthimically on the rate (Eq. 5.7). Thus, we can make a reasonable estimate of the entropic barrier for localizing the tetraloop near the receptor using $k_{diffusion}$ = 10^7 s^{-1} to estimate. Assuming this diffusion process is enthalpically barrierless, substitution of $k_{diffusion}$ into Eq. 5.7 yields $\Delta S_{diffusion} \approx$ −27 cal/mol/K (212).

The slight increase from 0.26 to 0.36 E_{FRET} with [Mg^{2+}] units from 0 to 4 mM corresponds to a roughly 5 Å compaction (R = $R_0^6(1-E_{FRET})/E_{FRET}$) with R_0 ~55 Å for the Cy3−Cy5 pair. To estimate the diffusion controlled rate, we estimated that based on the helical constraints of the RNA the average the tetraloop−receptor distance is 49 Å and the time to collide is proportional to the square of the distance traveled (Eq. 5.9). Thus, the increased collisional rate of the tetraloop and receptor would be ~$49^2/45^2$, which corresponds to only a ~1.2 fold increase in docking rate due to RNA compaction.

5.7.3   Temperature Controlled Single-Molecule Measurements

Temperature control (± 0.2 °C) of the sample is achieved by placing the microfluidic flow cell (128) into a heated stage (HSC60, Instec, Boulder, CO) onto the previously described confocal microscope system (106,128). The objective is thermally isolated from the microscope turret and is resistively heated through a collar (Bioptechs, Butler, PA) to 1 degree cooler than the stage to compensate for differential thermal flow characteristics. A thermocouple inserted directly into the buffer-filled flow cell is used to calibrate the temperature of the fully assembled sample-heating microscope system.

5.7.4   Supporting Figures
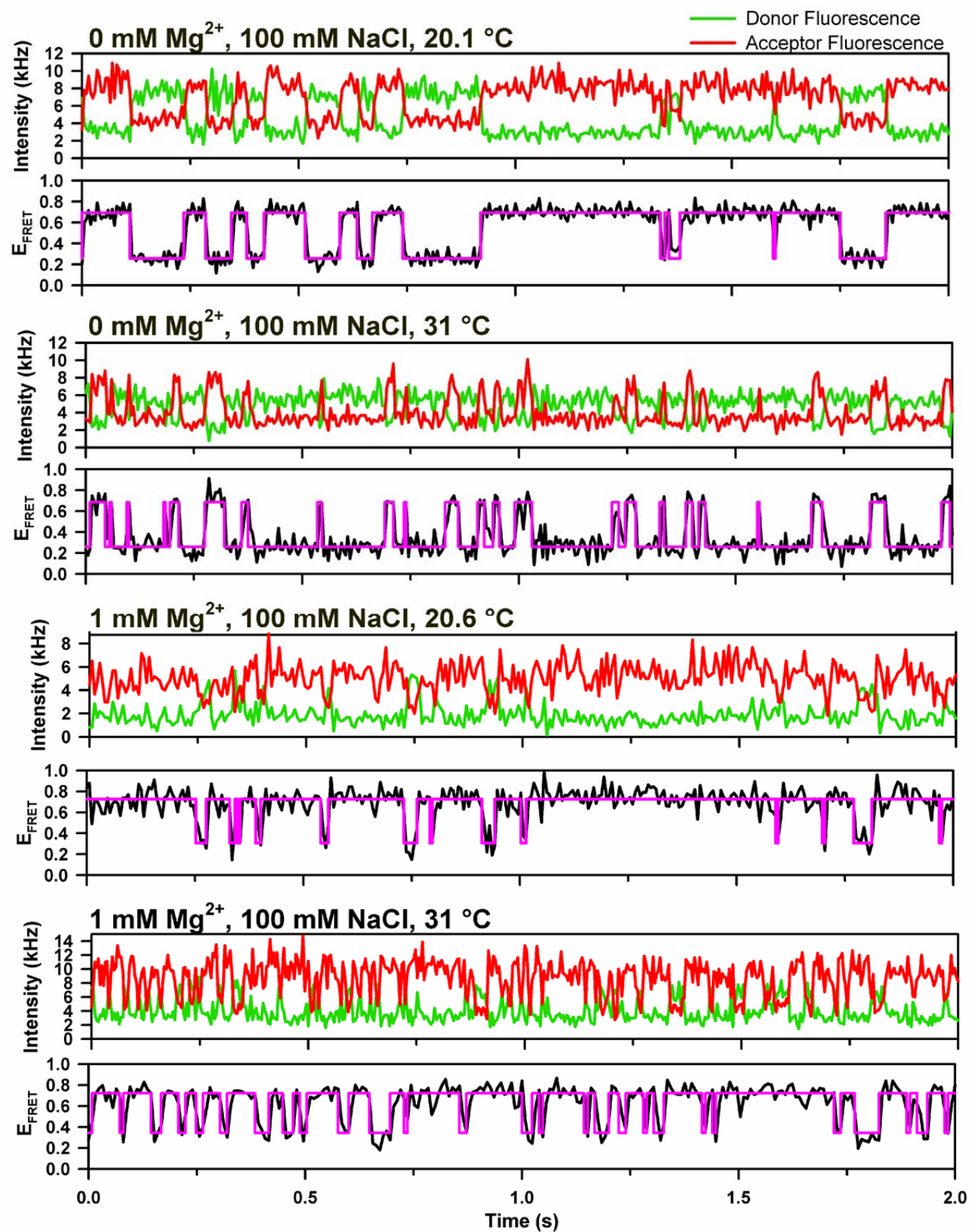
Figure 5.7   Sample donor/acceptor intensity and E_{FRET} traces for molecules in Figure 5.1.
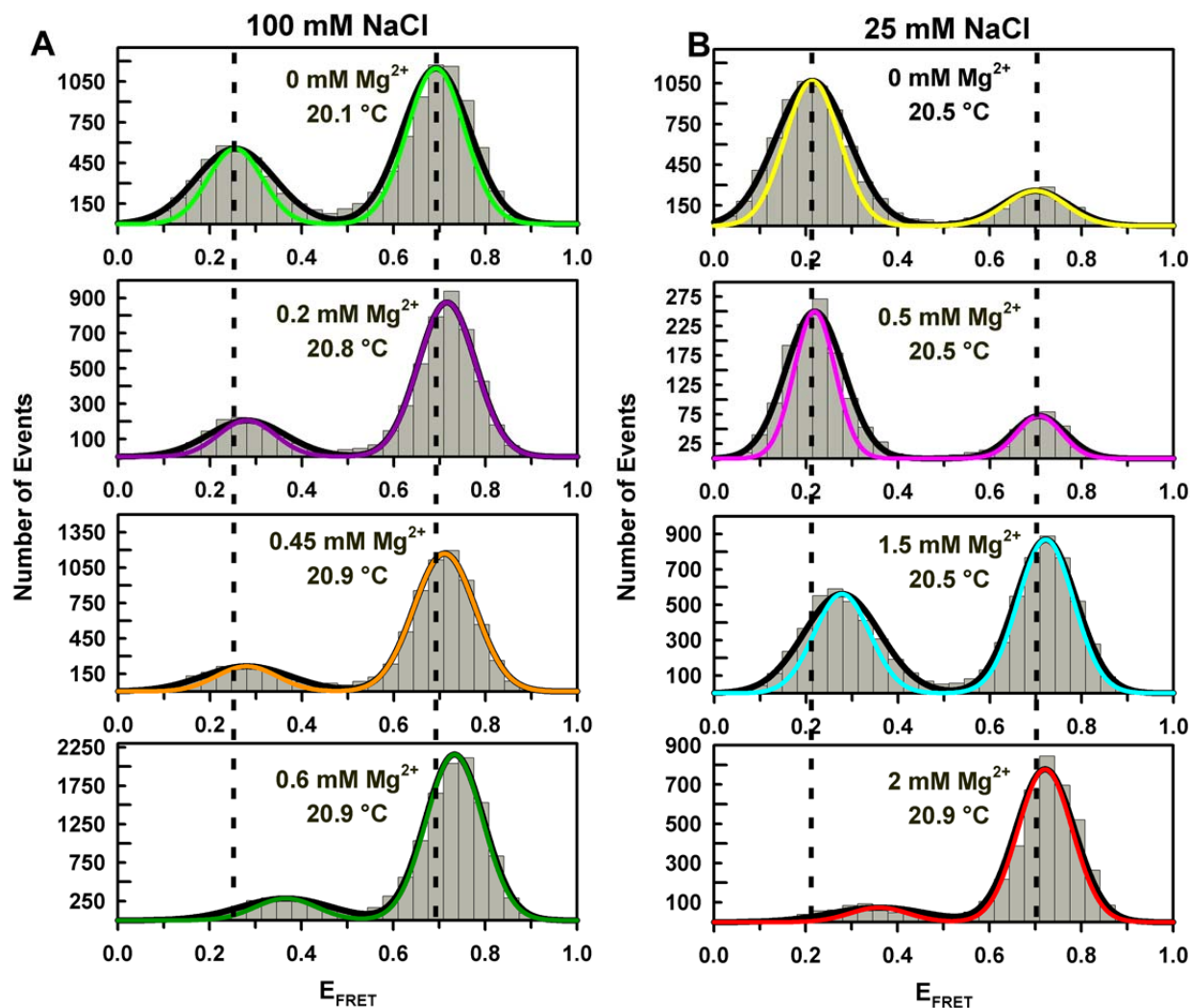
Figure 5.8   Increasing [Mg^{2+}] promotes tetraloop–receptor docking as seen in cumulative E_{FRET} population histograms (10–40 molecules each) in a (*A*) 100 mM NaCl or (*B*) 25 mM NaCl background. At 100 mM NaCl, over a range of 0 to 4 mM Mg^{2+}, the undocked and docked peaks (low and high E_{FRET}) shift from 0.26 ± 0.02 to 0.36 ± 0.02 and 0.69 ± 0.02 to 0.73 ± 0.02, respectively. At 25 mM NaCl, increasing [Mg^{2+}] from 0 to 3 mM shifts the undocked peak from 0.21 ± 0.02 to 0.36 ± 0.02, and the docked peak from 0.70 ± 0.02   to 0.72 ± 0.02. Comparison of the peak widths to shot-noise broadened width prediction reveals that the peaks do not broaden with increasing [Mg^{2+}] (shot-noise limited width predictions (*colored lines*) and Gaussian fits (*black lines*). The shot-noise predicted width, $\sigma_{SN,}$ is calculated from the standard-error propagation of finite photon counting statistic, as $\sigma_{SN} = (E_m(1-E_m)/T)^{1/2}$, where E_m is the center of the Gaussian peak and T is the minimum average count rate (photons/bin) for molecules included in the analysis (161). The undocked state is significantly broadened beyond shot noise ($\sigma_{undocked}/\sigma_{SN} = 1.4 ± 0.1$) while the docked peaks are nearly shot-noise limited ($\sigma_{docked}/\sigma_{SN} = 1.06 ± 0.08$). This observation is̍ consistent with a larger conformational accessibility of the tetraloop in the undocked vs docked states, as was discussed in detail elsewhere (128).

Figure 5.9   Cumulative histograms (~20 molecules) show that increasing temperature destabilizes the tetraloop-receptor interaction as seen by the decreased population in the high E_{FRET} state (docked) vs. low E_{FRET} state (undocked) and shown for individual molecules in Figure 5.1. The E_{FRET} peak positions and width are independent of temperature. There is no indication of broadening with temperature, the shot-noise limited peak expectations are shown in color vs. the Gaussian fits as described in Figure 5.8.
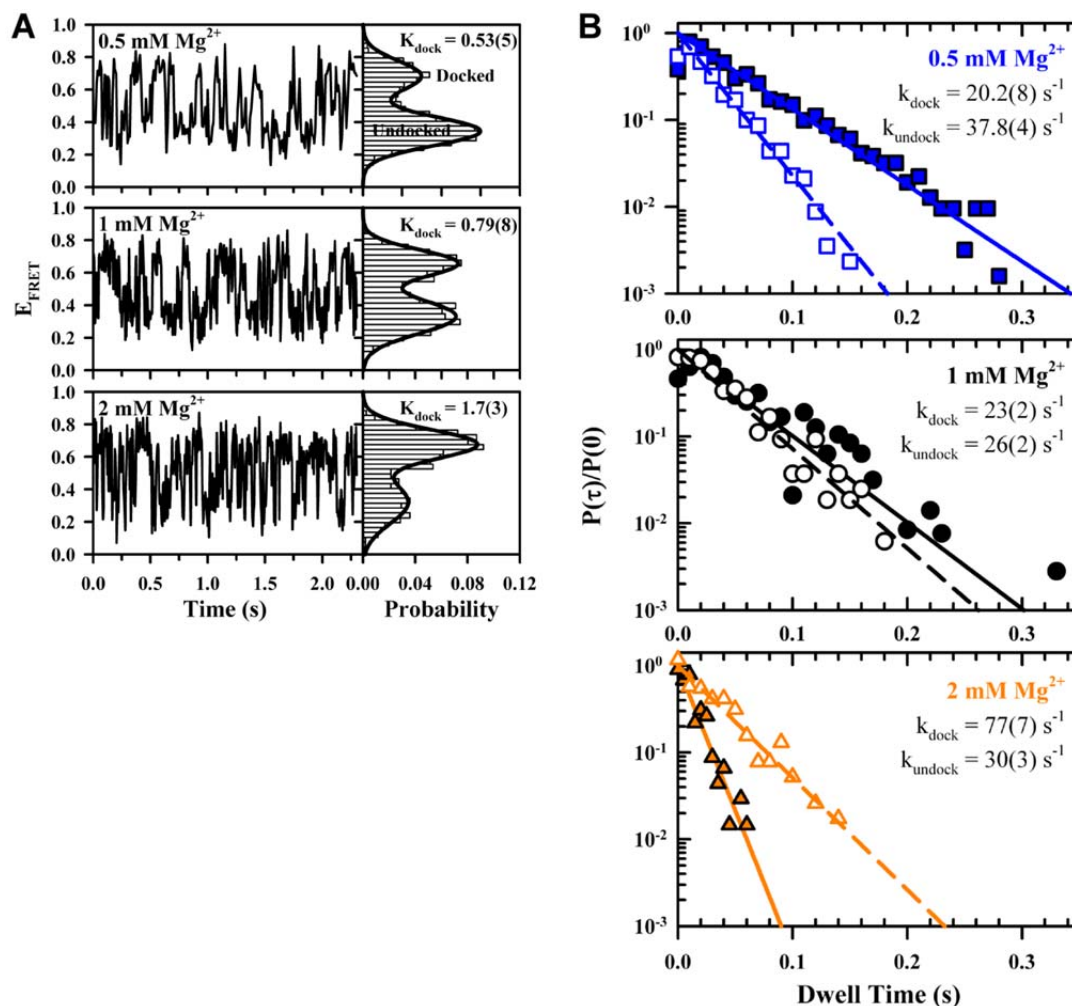
Figure 5.10 Sample data and analysis for the A$_7$ tetraloop–receptor construct at elevated temperature (35 °C). (*A*) Single-molecule E_{FRET} trajectories reveal that Mg^{2+} increases the melting temperature of the tertiary interaction, as indicated by the increased dwell time in the high E_{FRET} state (docked) vs. low E_{FRET} state (undocked) in the corresponding probability histograms. (*B*) Dwell time probability densities at varying [Mg^{2+}] yield rate constants for docking and undocking from the undocked and docked dwell times, respectively. The probability densities are well fit by single exponential decays, which are predicted from a two-state model for the observed process. [Mg^{2+}] increases k_{dock} and decreases k_{undock}.
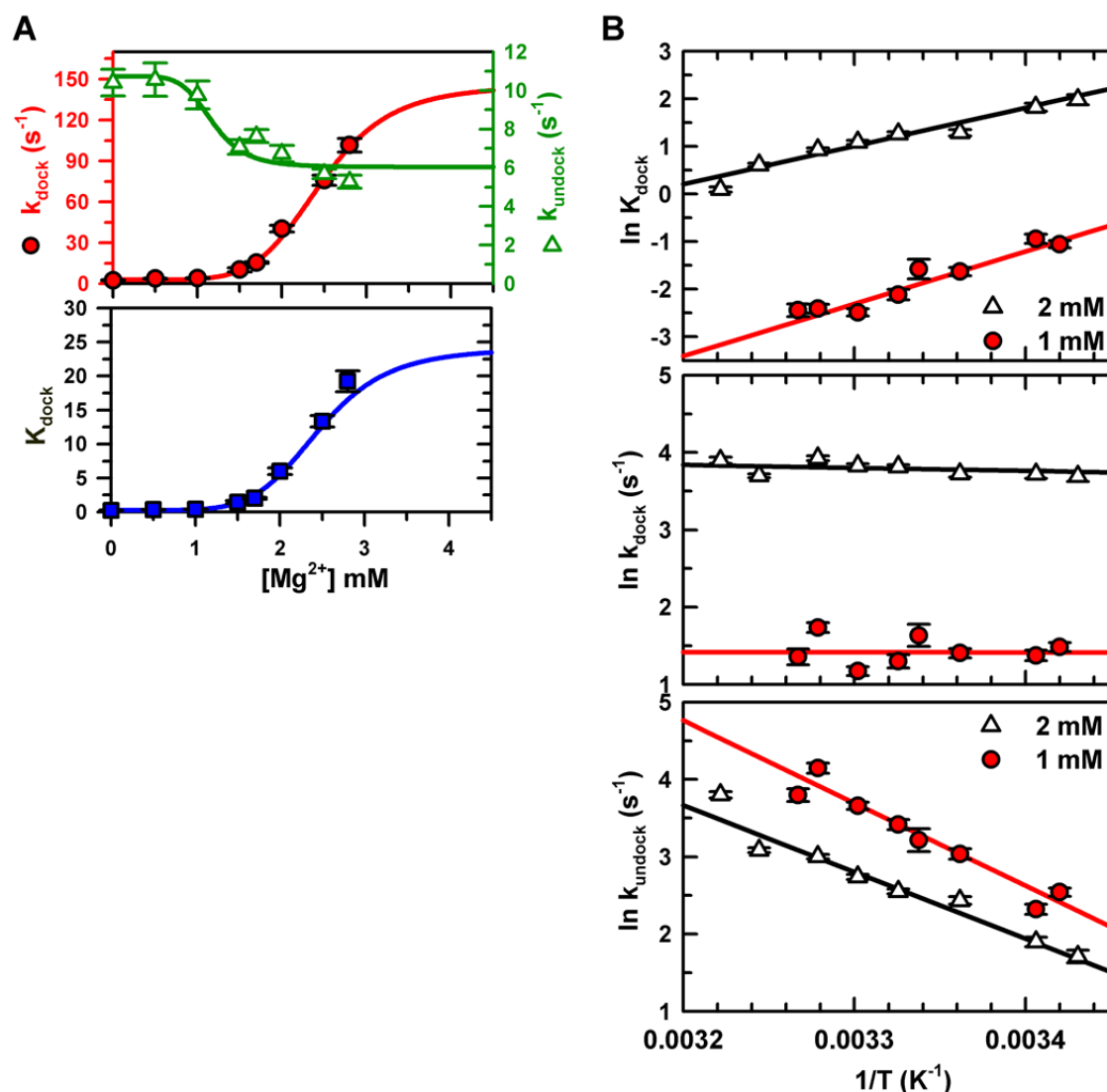
Figure 5.11 [Mg^{2+}] and temperature dependence of the tetraloop–receptor docking/undocking (U$_7$ construct, Figure 5.1 *A*) at low [NaCl] (25 mM). (*A*) [Mg^{2+}] for k_{dock}, k_{undock} and K_{dock} fit to the four-state kinetic scheme as in Figure 5.2, yielding n = 6 ± 1, k_1 = 2.9 ± 0.3 s^{-1}, k_2 = 145 ± 50 s^{-1}, k_{-1} = 11 ± 1 s^{-1}, k_{-2} = 6.0 ± 0.3 s^{-1}, K_{Mg} =2.5 ± 0.3 mM, and K'_{Mg}, =1.2 ± 0.2 mM. At low [Na$^+$] the Mg^{2+} cooperativity (n) is substantially greater than at 100 mM NaCl (Figure 5.2). (*B*) van't Hoff and Arrhenius plots yield the enthalpies and entropies of tetraloop–receptor docking/undocking at low [NaCl], as summarized in Table 5-1 and Table 5-2.

Chapter 6   The Role of Counterion Valence and Size in GAAA Tetraloop–Receptor Docking/Undocking Kinetics

6.1   Abstract

RNA overcomes electrostatic frustrations and folds into compact, ordered structures with the aid of counterions. A physical understanding of the counterion-assisted folding process requires addressing how cations kinetically and thermodynamically alter a tertiary-folding equilibrium. Using single-molecule FRET techniques, we explore the cation concentration dependence of a simple RNA folding system, i.e., the intramolecular docking/undocking of a GAAA tetraloop and its 11 nucleotide receptor. The tetraloop–receptor interaction is a ubiquitous tertiary motif, contributing to the proper folding and activity of a variety of structured RNAs. We extract the [cation]-dependent rate constants for docking (k_{dock}) and undocking (k_{undock}), which are obtained for cations of varying size and valence, specifically Na^+, K^+, Mg^{2+}, Ca^{2+}, $Co(NH_3)_6^{3+}$, and spermidine^{3+}. Increasing cation concentration dramatically acceler- ates k_{dock} while slightly decreasing k_{undock}. We describe these trends by a four-state kinetic model with cation dependent and independent pathways, allowing for separation of the kinetics of folding from the energetics of cation interaction with the undocked and docked states, respectively. This kinetic analysis reveals a markedly higher affinity for cations in the docked vs. undocked RNA, which leads to a change in net cation uptake as a function of cation concentration. We observe that the free energy of cation-RNA interaction is inversely proportional to cation valence, consistent with counterion condensation theory. The number of cations taken up with folding also decreases with the charge density of the cation. Inter- estingly, the same degree of free energy stabilization of the transition state and docked state is achieved by each cation investigated. The one exception is spermidine^{3+}, whose smaller effect on the docking

equilibrium can be ascribed to excluded volume effects and charge separation that prevent complete charge neutralization of the RNA.

6.2 Introduction

The ability of RNA molecules to assemble into compact, functional structures depends inherently on counterion neutralization of the negatively charged phosphate backbone (12,40). The cellular environment is composed of a multitude of cations that can promote folding, e.g., K^+ and Mg^{2+} are present in concentrations $\sim$150 mM and $\sim$0.5–1 mM, respectively (226-228). Polyamines are also abundant in sufficient concentrations to promote folding, and are implicated in a number of cellular processes (229-231). However, because metal ions can interact with RNA through both specific coordination and non-specific delocalized interactions, the relative efficacy of monovalents, divalents, and trivalents in aiding RNA folding is not well understood (6,12,40,41,58,62,81,136-139). The most common type of counterion interactions with RNA are non-specific and delocalized, i.e., interactions with "diffuse", fully hydrated, mobile yet localized cations (41). The roles of Mg^{2+} and Mn^{2+} in RNA folding are particularly intriguing, since one or the other is often required for catalytic activity, as in the case of the *Tetrahymena thermophila* ribozyme (232,233). A quantitative prediction of the dependence of folding on cation concentration is challenging because the irregularity of RNA structure produces a nonperiodically varying electrostatic potential along the molecule (231). Though much progress has been made in using nonlinear Poisson-Boltzmann theory to describe nucleic acid electrostatic potentials and the corresponding spatial distributions of cations (84,170,172,176), these models still fail to accurately predict the ion atmosphere of multivalent cations around even well-defined DNA helices (176,234), with clear deficiencies noted for coupling the ion atmosphere to RNA folding transitions (234). Furthermore, current theories are still limited in describing the folding dependence on cation size (176,235-237). For these reasons, simpler models, such as Manning counterion condensation theory, have proven useful in interpreting the cation dependence of *Tetrahymena* ribozyme folding rates and equilibrium (167,213). An even more complicated task is describing mixed cationic environments, which can lead to both cation compe-

tition and synergy (8,128). Though large unexplored, flexible oligomeric cations, such as polyamines, can also perturb folding (229). Studies of the *Tetrahymena* ribozyme providing particularly valuable insights into polyamine assisted folding (167,213,238).

However, RNA structure is a dynamic property, with the charge distributions changing as the molecule folds and unfolds (40). Folding increases the negative charge density of RNA, and is therefore frequently accompanied by an uptake of counterions (8,174,239). Thereby, cations interact more strongly with the folded vs unfolded conformation. Thus, knowledge of the differential affinity of counterions for the native and unfolded conformations is needed to fully address cation-mediated folding. A major theoretical challenge in describing the attraction of cations to the unfolded RNA is that the unfolded state exists as an ensemble of possible configurations, sampling of which can be altered by local ion atmosphere (217). Though much work has been done in the realm of the cation dependence of equilibrium folding distributions, it is less well understood how a cation-induced stabilization kinetically correlates with RNA folding and unfolding rates (40). In ensemble studies, the effects of counterions on both the folding and unfolding rate constants have not been identified, precluding a mechanistic identification of the role of the cations. In single-molecule FRET studies, on the other hand, folding and unfolding rates can be extracted at equilibrium, where folding of a fluorescently labeled RNA brings a dye pair in closer proximity. From FRET efficiency, $E_{\text{FRET}}(R) = R_0^6/(R_0^6 + R^6)$, this translates into readily measurable real-time changes in, where R_0 is the Förster radius for 50% energy transfer probability and R is the inter-dye distance. Single-molecule FRET studies of RNA folding have revealed particularly rich cation-dependent folding kinetics. For example, an investigation of [Mg^{2+}]-dependent folding in a three-helix junction has shown that metal-ion induced stability does not originate from a simple two-state binding scheme (148). The RNA folding picture has been significantly advanced by the folding/unfolding kinetic studies of the hairpin ribozyme (3,8,146), RNase P (240), the *Tetrahymena* ribozyme (103,119,150), a three and four-helix junctions (108,148), and a group II intron (241).

Since RNA folding proceeds hierarchically, with tertiary structure proceeding through association of well-defined secondary elements (6,35), frequently manifested tertiary motifs have emerged re-

curring themes in RNA (42). Thereby, the eventual goal of predicting RNA structure can also be approached from a reductionist perspective, e.g., the counterion-dependence of formation of an isolated tertiary interaction. Toward this end, we explore the kinetics of the ubiquitous and structurally characterized GAAA tetraloop–receptor tertiary interaction (Figure 6.1 *A* and *B*), which contributes to the proper folding and activity of a variety of structured RNAs, including group I and II introns and RNaseP (63,70,242,243). The structures of the free and bound forms of the tetraloop and receptor have been determined (50,62,73,74). The tetraloop is structurally not altered by binding, but the receptor under goes significant rearrangement, as depicted schematically in Figure 6.1 A (50,62,73,74). The tetraloop–receptor interaction can form outside the context of large RNAs (73,76) and in a vast range of ionic conditions (82), making it an ideal system for isolated study.

In this work, we continue a systematic characterization of simplified RNA folding systems, by studying the intramolecular docking of GAAA tetraloop with its 11 nt canonical receptor connected by an A_7 single-stranded linker (Figure 6.1 *A*). In previous ensemble measurements, we showed that the tetraloop–receptor interaction is formed with the aid of many different cations (124). Although such studies yield important qualitative trends for cation-facilitated docking, they are unable to elucidate the kinetic origin of the perturbed folding equilibrium. As a result, single-molecule FRET studies have revealed that both the undocking and docking rates are affected by $[Mg^{2+}]$(106). Here, we further explore these cationic effects and show that monovalents (Na^+ and K^+) , divalents (Mg^{2+} and Ca^{2+}) , and trivalents ($Co(NH_3)_6^{3+}$ and spermidine^{3+}) affect the folding landscape similarly—by increasing the docking rate constant, k_{dock}, while reducing the undocking rate constant, k_{undock}. Spermidine$^{3+,}$ however, promotes folding to a lesser extent than the other cations investigated. Cation charge is the major distinction between the affinities, with cation size, i.e., K^+ vs. Na^+, also playing a minor role in the stoichiometry of cation uptake with folding. These observations are consistent with RNA folding facilitated by counterion condensation. The origin of the RNA folding enhancement in the presence of each investigated cation is shown to be of the same mechanism, i.e., the thermodynamic stability of the docked form (ΔG°_{dock}) increases more rapidly than the corresponding barrier height for docking decreases. An analysis of the
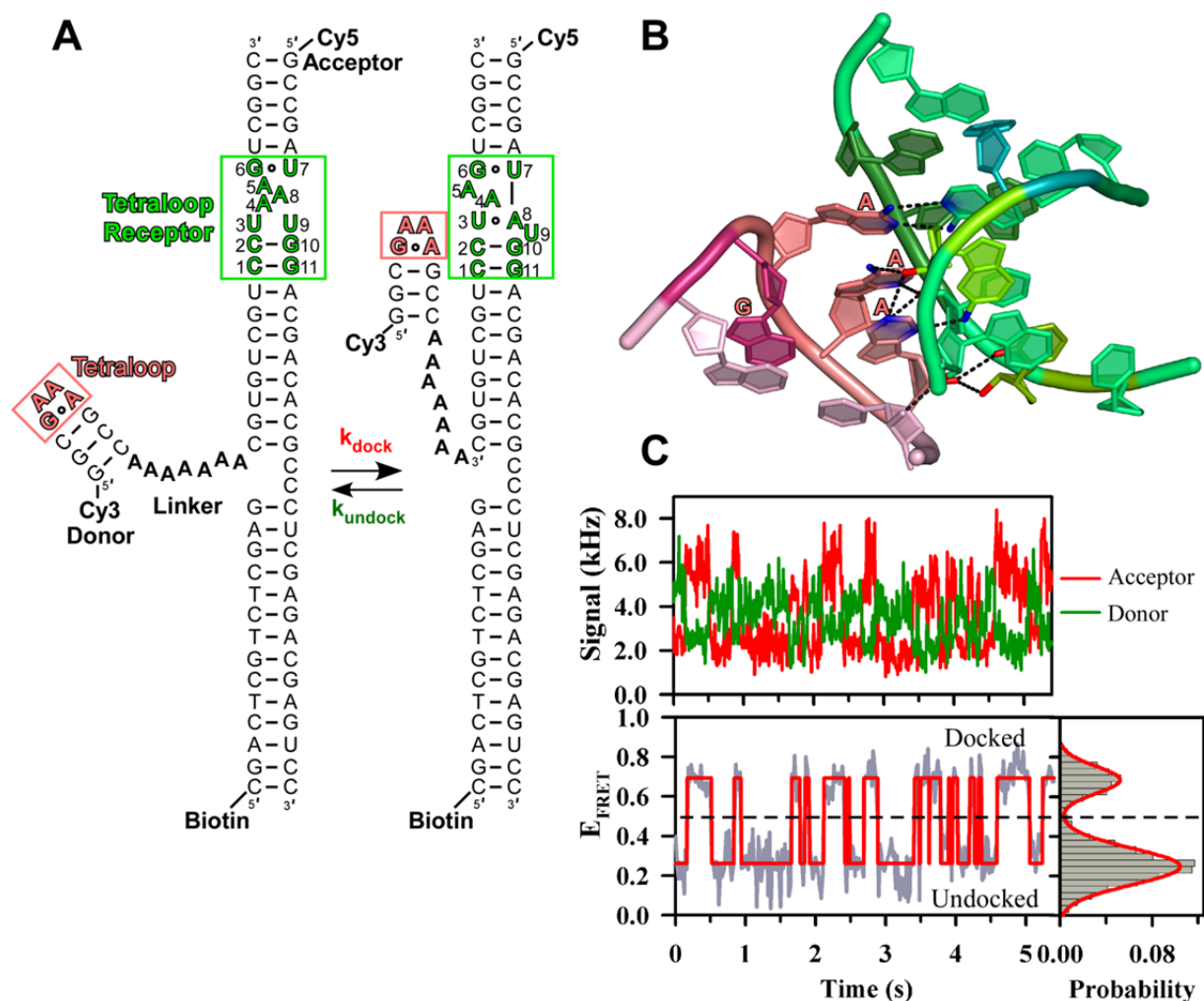
Figure 6.1   GAAA tetraloop–receptor docking/undocking interaction. (A) Schematic of the observable RNA folding transition in anRNA construct isolating the tetraloop–receptor interaction, characterized by rate constants, k_{dock} and k_{undock}. Changes in FRET efficiency between the Cy3 and Cy5 allow monitoring of GAAA tetraloop docking into its receptor. The RNA is immobilized on glass surfaces with biotin-streptavidin binding. (*B*) Structure of the GAAA tetraloop (AAA shown in *salmon*, G in *magenta*, closing base pair in *light pink*) and its canonical 11 nt receptor (*green*) in the *Tetrahymena* riobyzme's P4–P6 domain. 10 hydrogen bonds form between the tetraloop and receptor regions, shown as black dotted lines, *blue* = nitrogen, *red* = oxygen (hydrogens not shown) (PDB ID 1HR2). (*C*) Monitoring tetraloop–receptor docking/undocking by FRET as seen by the anti-correlated donor and acceptor fluorescence signal and corresponding E_{FRET} (*gray lines*) trajectory with Hidden Markov two-state fit shown in *red*. The probability distribution of the E_{FRET} traces reveals well-resolved docked and undocked states. The sample data conditions are 100 mM KCl, 50 mM HEPES, and 0.1 mM EDTA at pH 7.5 and 21 °C.

change in the preferential interaction coefficient ($\Delta\Gamma_M$) of the RNA reveals that the number of cations

taken up with folding dramatically changes as a function of cation concentration.

6.3   Experimental Procedures

6.3.1   RNA Preparation

Cy3-Cy5 labeled tetraloop–receptor constructs depicted in Figure 6.1 are prepared as previously described (106,124). Briefly, synthetic 5'-three carbon amino-modified RNA oligomers (Dharmacon, Lafayette, CO) are labeled with Cy3 and Cy5 N-succinimidyl esters (Amersham Biosciences, Piscataway, NJ) and HPLC purified. Annealing the Cy3 (1 μM) and Cy5 (1.5 μM) RNA oligomers with 2 μM biotinylated DNA oligomer (Integrated DNA Technologies, Coralville, IA) by heating to 70 °C and cooling slowly to room temperature in an annealing buffer of 50 mM HEPES, 100 mM NaCl, 100 μM EDTA, pH 7.5 forms the complete construct (Figure 6.1 A). The secondary structure of the Cy3 strand forms the tetraloop with an A_7 linker connecting it to the receptor domain created by the hybridized Cy3 and Cy5 strands. Molecules are tethered to streptavidin-coated glass surfaces with the biotinylated extension formed by the hybridized DNA and Cy5 strands.

6.3.2   Single-Molecule FRET Measurements

To enable smFRET studies of immobilized RNA, we have applied the previously described scanning confocal microscope system (106,128). Excitation is enabled by a 532 nm 82 MHz pulsed laser (Model 3800, Spectra Physics) with donor and acceptor emission discriminated by a dichroic beamsplitter and band-pass filters for respective detection on single-photon counting avalanche photodiodes (SPCM-AQR-14, Perkin-Elmer Optoelectronics, Fremont, CA) (106). Fluorescence trajectories are acquired for individual RNA constructs located on the coverglass by an intensity search algorithm with time traces acquired using a time-correlated single-photon counting (TCSPC) module (SPC-134 Becker & Hickl, Berlin).

RNA constructs are immobilized on a coverglass in a microfluidic flow-cell assembly (106,128) and observed in the diffraction-limited laser focus (objective numerical aperture 1.2) at 1–1.7 μW powers measured at the back plane of the microscope. A protocatechuic acid/protocatechuate-3,4-dioxygenase (PCA/PCD) enzymatic oxygen scavenging solution (60 nM PCA, 5 mM PCA) with 2 mM

Trolox is added for fluorophore photostability (129,130). Divalent (Mg^{2+} and Ca^{2+}) and trivalent ($Co(NH_3)^{3+}$ and spermidine^{3+} titrations are performed in a standard buffer containing 50 mM hemisodium HEPES (pH 7.5 at 25 °C), 100 mM NaCl and 0.1 mM EDTA with varying concentrations of hexamminecobalt trichloride and spermidine trihydrochloride. Freshly prepared spermidine^{3+} stock solutions are aliquoted and kept frozen to avoid deamination, as suggested by the manufacturer. Monovalent (K^+ and Na^+)-dependent studies of the tetraloop–receptor motif are performed in 50 mM hemisodium HEPES (pH 7.5) and 0.1 mM EDTA, with [NaCl] and [KCl] varied. This buffer contains 25 mM Na^+ independent of any added NaCl. Reported cation concentrations refer to the added concentration of the respective salt. All buffers are 0.2 μm sterile filtered and prepared using LC-MS Chromosolv H_2O. After flushing in solutions, data are collected under static conditions, with entrance and exit holes covered by tape. All experiments are performed at 21 °C.

6.3.3 Single-Molecule Trajectory Analysis

The FRET efficiency, E_{FRET}, is calculated ratiometrically from the donor and acceptor signals for time trajectories binned at 5–10 ms, which clearly resolves the undocked and docked states (106,128). The corrected intensity-based FRET efficiency (E_{FRET}) is calculated from the background subtracted signals on the two channels, ΔI_1 and ΔI_2, designed primarily for donor and acceptor detection, respectively. Corrections are implemented for (i) collection efficiencies and crosstalk of the donor and acceptor emission on channels 1 and 2 (β_1^A, β_2^A, β_1^D, β_2^D) (ii) differential quantum yields of the donor and acceptor (Q_D and Q_A), and (iii) direct laser excitation of the acceptor (α_A, where $1-\alpha_D = \alpha_A$),

$$E_{FRET} = \frac{\beta_1^D \Delta I_2 - \beta_2^D \Delta I_1 - \dfrac{\alpha_A}{\alpha_D}\dfrac{Q_A}{Q_D}(\beta_2^A \Delta I_1 - \beta_1^A \Delta I_2)}{\beta_1^D \Delta I_2 - \beta_2^D \Delta I_1 + \dfrac{Q_A}{Q_D}(\beta_2^A \Delta I_1 - \beta_1^A \Delta I_2)}. \qquad 6.1$$

Quantum yield ratios and collection efficiencies are determined with independent measurements (128). Fractional-direct laser excitation of the acceptor and donor is calculated from the extinction coefficients at 532 nm (128). Donor-only species, which are clearly identifiable by absence of ac-

ceptor emission, are disregarded. Day-to-day reproducibility of the center E$_{FRET}$ values is ±0.02. The quantum ratio (Q_A/Q_D) is independent of the cation concentration, as assessed by fluorescence lifetime measurements of singly labeled constructs (see *Supporting Information* section 6.8.1). Cy3 and Cy5 are subtly quenched by cobalt hexamine. However, the extent of quenching on both fluorophores is equal, such that the quantum yield ratio is maintained. Furthermore, the Förster radius (R_0) remains constant (see *Supporting Information* Figure 6.10) (111). Thus, the E$_{FRET}$ values between measurements at various [cation] can be compared to assess the effect of the cation on the undocked and undocked states.

6.3.4 Determination of Rate Constants for Docking and Undocking from Single-Molecule Trajectories

To determine docking and undocking rate constants at each cation concentration, dwell times of the tetraloop–receptor construct in the docked and undocked conformation are defined by the crossing of a threshold set at the minimum of the bimodal E$_{FRET}$ distribution in the real-time trajectory. Histograms of the dwell times in the docked and undocked states at a given cation concentration are converted to a probability density, P(τ_i) ≈ H(τ_i)/[0.5(τ_{i+1} - τ_{i-1})], where H(τ_i) is the standard histogram value and the τ_i represents an ordered list of nonzero time bins, enhancing the dynamics range of observable time scales (106,244). The resulting dwell time histograms and normalized probability densities (P(τ)/P(0)) can be described by a single-exponential decay, representing the observed processes of docking and undocking. Least squares single exponential fits yield rate constants, k$_{dock}$ and k$_{undock}$. Data sets contain >200 events (typically ~500). Hidden Markov modeling is also pursued as a method for determining rate constants and yields identical rate constants within uncertainties, as expected for two well resolved states (204).

Tetraloop–receptor folding is heterogeneous, consisting of three previously described populations: i) the aforementioned population that actively switches between the docked and undocked conformation (68%), ii) a minority population that does not dock on the time scale of observation (32%), and iii) third minority population of molecules always in a high E$_{FRET}$ state (~1%) (106). In that the minority populations show no docking/undocking events, they do not represent formation of tertiary structure and are therefore excluded from the analysis. As discussed elsewhere, the origin of this heter-

ogeneity is not known (117,128), but is observed in equivalent fractions over the vast ranges the cation-ic environments explored here.

6.4   Results and Analysis

6.4.1   Tetraloop–Receptor Docking is Promoted by Monovalent, Divalent, and Trivalent Cations

Intramolecular docking of the GAAA tetraloop into the tetraloop receptor (Figure 6.1) is monitored as a function of cationic environment to investigate the importance of counterion valence and size for tertiary RNA folding. Specifically, we examine the effect of $[Na^+]$, $[K^+]$, $[Mg^{2+}]$, $[Ca^{2+}]$, $[Co(NH_3)_6^{3+}]$, and $[spermidine^{3+}]$ on the tetraloop–receptor docking/undocking kinetics. Tetraloop–receptor docking and undocking are monitored by single-molecule FRET, which reveals two well-resolved states, docked and undocked (low and high E_{FRET}) in real time, Figure 6.1 C (106). At low monovalent concentration (100 mM KCl, Figure 6.1 C), the RNA spends the majority of its time undocked. Increasing the concentration of monovalent cation to 300 mM shifts the tetraloop–receptor equilibrium to favor docking (Figure 6.2). A similar shift in the equilibrium can be achieved with just 1 mM divalent, or 0.1 mM trivalent cations, as shown in the E_{FRET} trajectories displayed in Figure 6.2. Thus, docking is effectively promoted by a large array of cations.

As seen in these trajectories, the RNA fluctuates between the undocked and docked states (low and high E_{FRET}), well described by Gaussian distributions with mean centers of $\langle E_{FRET} \rangle = 0.26 \pm 0.02$ and 0.69 ± 0.01, respectively, consistent with previously observed values for these states at other conditions (106,128). It appears that the tetraloop–receptor has no specific cation requirement for docking. However, the concentration of cation necessary to promote docking is highly dependent on ion valence, [monovalent]>[divalent]>[trivalent], as noted in ensemble measurements, but now seen at the single-molecule level (124). Na^+ and K^+ appear to have identical distributions at 300 mM. However, Ca^{2+} is slightly less effective than Mg^{2+} at 1 mM, and cobalt hexamine induces more docking than spermidine^{3+} at 100 μM. Explicit comparison of the induced stability is obtainable by extraction of the rate constants
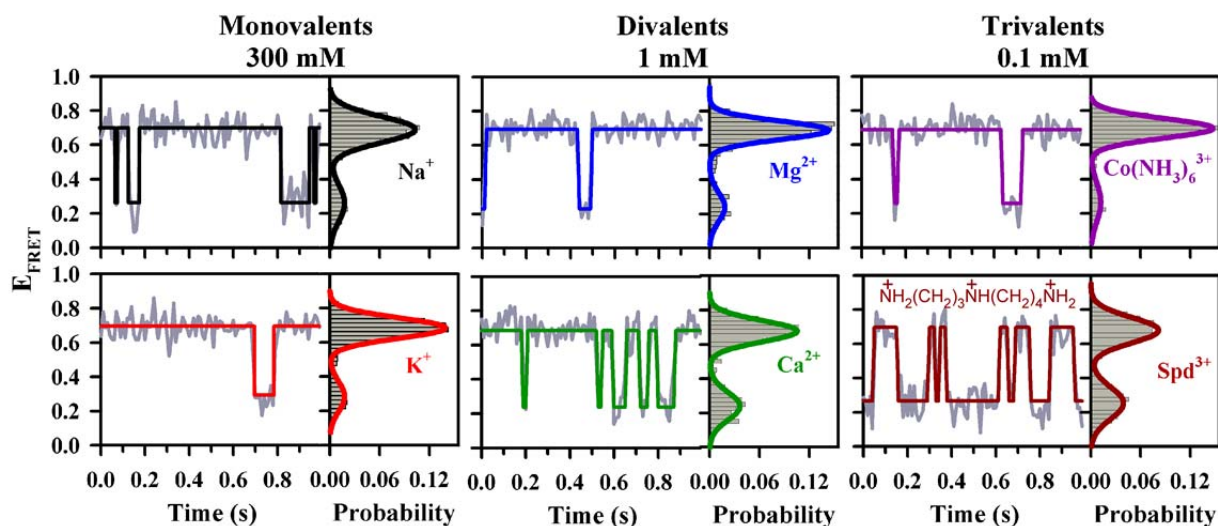
Figure 6.2   Sample FRET efficiency trajectories and probability distributions for tetraloop–receptor docking in concentrations of 300 mM monovalent (Na$^+$ or K$^+$), 1 mM divalent (Mg^{2+} or Ca^{2+}), or 100 μM trivalent (Co(NH$_3$)$_6^{3+}$ or spermidine^{3+} (Spd^{3+})). The RNA fluctuates between high and low E_{FRET} states. Data are shown in *gray* with Hidden Markov fits overlaid in color. Positively charged amino groups are covalently linked by hydrocarbon chains in spermidine^{3+} (*lower right panel*).

for docking and undocking (k_{dock} and k_{undock}) from the single-molecule time trajectories as a function of

cation concentration.

6.4.2   Cation dependence of k_{dock} and k_{undock}

Each cation investigated leads to successful docking of the tetraloop and receptor, with similar two-state

behavior as shown in Figure 6.2. Therefore, one can compare rate constants for docking and undocking

of the tetraloop into its receptor, k_{dock} and k_{undock}, as function of [cation]. Sample probability densities

for the dwell times in the undocked and docked states (see *Experimental Procedures*) are shown in Fig-

ure 6.3 for the same concentrations displayed in the sample trajectories in Figure 6.2. The semi-log plots

are linear, indicating an observed first-order (two-state) process, which corresponds to a monoexponen-

tial decay of the dwell time probability. Semi-logarithmic fits of the probability densities of the compila-

tion of the dwell times from many molecules (~10–30 molecules, 200–2000 events) yield k$_{dock}$ and k$_{un-}$

$_{dock}$ from the dwell time undocked and docked, respectively. There is some evidence of deviation from
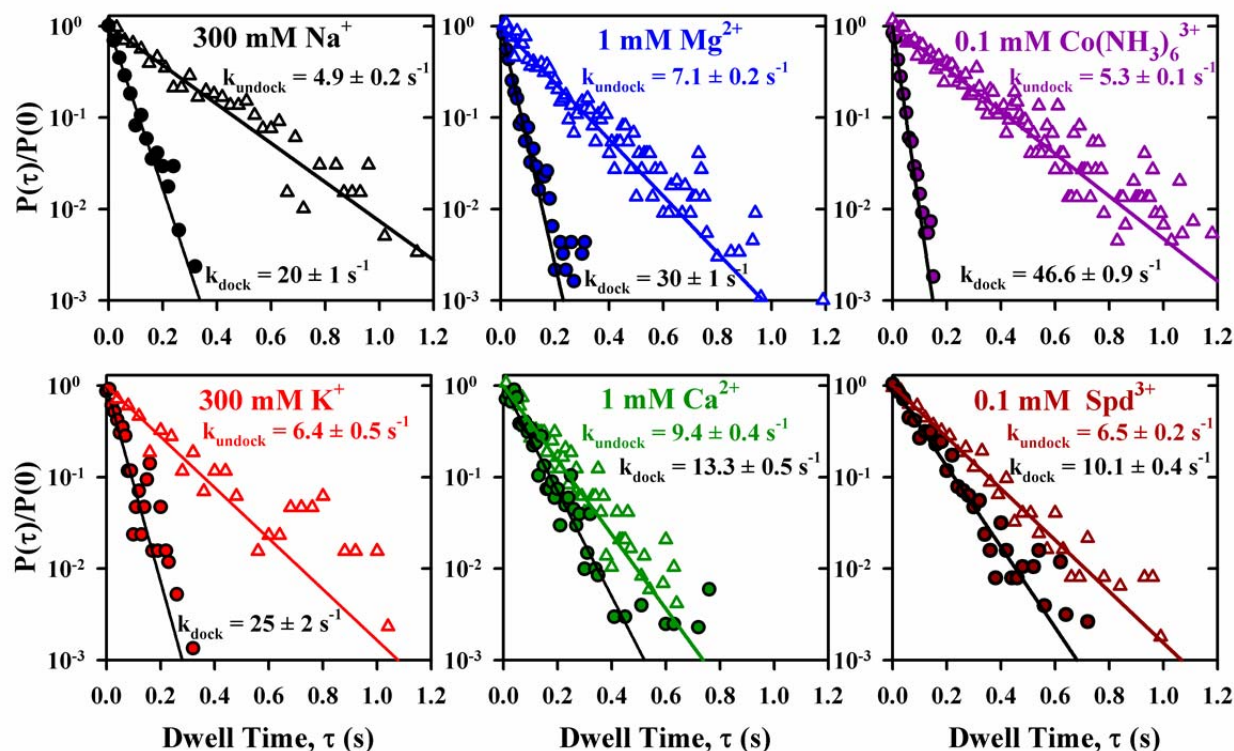
Figure 6.3   Cumulative normalized probability densities for the dwell time (τ) in the docked (*open triangles,* △) and undocked (*filled circles*, ●) at 300 mM monovalent (Na⁺ and K⁺), 1 mM divalent (Mg²⁺ and Ca²⁺), and 100 μM trivalent (Co(NH₃)₆³⁺ and spermidine³⁺ (Spd³⁺) concentration. Single exponential fits of the undocked and docked dwell times yield the rate constants for docking (k_{dock}, *black lines*) and undocking (k_{undock}, *colored lines).* Error bars are obtained from least squares fits of the probility densites compiled from >10 molecules and >220 events.

single exponential behavior at very long event durations (see undocked dwell times for Ca²⁺ in Figure 6.3). This deviation could be caused by potential kinetic heterogeneity or alternative folding pathways (106).

The dependence of k_{dock} and k_{undock} on [cation] is plotted in Figure 6.4. A stark increase in k_{dock} and concomitant decrease in k_{undock} is observed for increasing [cation]. The asymptotic values of the titration corresponds to an ~12 fold increase in k_{dock} and ~3 fold decrease in k_{undock} for the multivalent cations, which are performed from a background starting concentration of 100 mM NaCl. The background of 100 mM NaCl is included because it is physiologically relevant, enables clearer interpretations of the effect of multivalent cations on folding by maintaining constant activity because of the Cl- excess
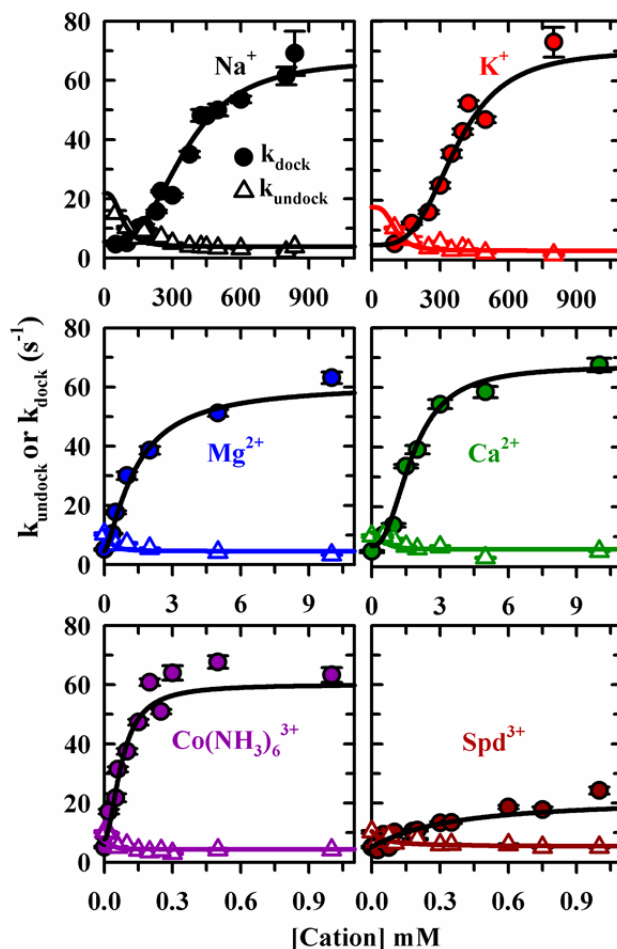
Figure 6.4 [Cation] dependence of tetraloop–receptor docking and undocking. (*A*) k_{dock} (*filled circles*) and k_{undock} (*open triangles*) are plotted *vs.* [cation]. Monovalents affect folding on the molar range, divalents on the 10 millimolar range, trivalents on the 100 micromolar range. The dependence of k_{dock} and k_{undock} are fit to a four-state kinetic model Figure 6.6 and Table 6-1. Note that the Spd^{3+} fit is to data that extends to 1.9 mM.

(215). Furthermore, the focus of this study is on tertiary structure formation, and a 100 mM monovalent background ensures relaxation of the secondary structure (e.g., rigid, extended helices), without contributing significantly to tertiary structure formation (128,220). The monovalent titrations, however, extend below 100 mM, thus the magnitude of changes for k_{dock} and k_{undock} are greatest (Figure 6.4). The exception of the counterions is spermidine^{3+}, for which there is only a 4-fold increase in k_{dock} and 1.7-fold decrease in k_{undock}. In terms of the equilibrium constant, $K_{dock} = k_{dock}/k_{undock}$, this would imply that spermidine^{3+} can only shift the equilibrium constant to ~4 vs. ~20 for the other cations. The effect

of this equilibrium constant difference is most clearly illustrated by the fraction of time the molecules spends docked,

$$f_d = \frac{k_{dock}}{k_{dock} + k_{undock}} \, .$$

6.2

At saturating [spermidine^{3+}], the RNA only spends ~80% of its time docked vs. 95% in the presence of Mg^{2+} (Figure 6.5). This analysis yield another very interesting effect; the combination of k_{dock} increasing and k_{undock} decreasing with [cation], results in a very steep saturation of the fraction docked as compared to the dependence of k$_{dock}$ or k$_{undock}$ alone. In a fraction docked analysis, the midpoint concentration is readily characterized. Thus, we can compare the single-molecule measurements to ensemble FRET studies, as is done in the following section.

There are orders of magnitude differences in the ranges over which the cations affect the rate constants, with trivalents in the hundreds of micromolar range, divalents in the mM range, and monovalents in the M range for saturation. In the case of the monovalents and divalents, there is no obvious difference in the effect of cation size at the same valence on k_{dock} and k_{undock}, i.e., Na$^+$ resembles K$^+$. However, there is a dramatic difference in the ability of spermidine^{3+} to induced docking and prevent undocking as compared to Co(NH$_3$)$_6$$^{3+}$ (Figure 6.4). These results suggest that the cation valence is the determining factor for the effective concentration range of the cation. A more quantitative analysis is required to assess the role of counterion charge density in promoting docking, as is discussed in the following section.

Figure 6.5   Fraction docked (Eq. 6.2), fractional dwell time of the molecule in the docked state with the four-state kinetic fit from Figure 6.4 overlaid. The midpoint concentration of the fraction docked is indicated in Table 6-2 and compared to ensemble values.

6.4.3 Kinetic Model for Cation-Facilitated Tetraloop–Receptor Docking

To explain the origin of the cation-dependence of k_{dock} and k_{undock} in the nominally two-state scheme (Figure 6.4 *A*) a minimally four-state model is required. In Figure 6.6 *B*, a model that involves a cation (M)-dependent and cation-independent pathway is shown (106). In this model, cation exchange occurs much faster than experimentally observable docking or undocking transition. The rapid equilibrium of the cation free and bound forms of undocked and docked states are described by the apparent Hill coef-

Figure 6.6   (*A*) Two states are observed for the tetraloop receptor docking undocking equilibrium. Both states undocked (U) and docked (D) are affected by the ion atmosphere, thus k_{dock} and k_{undock} are dependent on [cation]. (*B*) A four-state kinetic model for describing the origin of the cation-dependence of k_{dock} and k_{undock}, where K_M and K'_M are dissociation constants for the cation (M) and the rate constants reflect docking and undocking resolved by FRET for the cation dependent and independent pathways. The free energy changes associated with each transition are labeled. The observed docking/undocking rate constants are a combination of the M dependent and independent pathways, the relative contribution of which are perturbed by [M].

ficient, *n*, and dissociation constants K_M and K'_M, respectively (Figure 6.6B). Furthermore, the cation-bound and cation-free forms of the undocked and docked states are experimentally indistinguishable by FRET. As a result, the experimentally observed rate constant originating from a monoexponential decay of the dwell time probabilities (Figure 6.2) reflect the combination of k_1 and k_2 for docking and k_{-1} and k_{-2} for undocking (106,148), i.e.,

$$k_{dock} = \frac{k_1 (K_M)^n + k_2 [M]^n}{(K_M)^n + [M]^n}$$ 6.3

$$k_{undock} = \frac{k_{-1} (K_M^{'})^n + k_{-2} [M]^n}{(K_M^{'})^n + [M]^n}.$$ 6.4

This model has been previously applied to conformation changes in an RNA three-helix junction (148). The thermodynamic cycle described in Figure 6.6 *B* has also been employed to theoretically link Mg^{2+} binding with RNA folding in yeast tRNA^{Phe} and a ribosomal RNA fragment (174). This scheme can be energetically described by ΔG^{o}_M and $\Delta G^{o}_M{'}$, the free energies for cation binding, and $\Delta G^{o}_{dock,0}$ and $\Delta G^{o}_{dock,M}$, the folding in the presence and absence of cation (84,174). This model allows for different affinities for the cation in the folded vs. unfolded conformers. Since folding can proceed through the combination of the two pathways, the net uptake of cations with docking can vary with concentration of cation, as will be discussed later. A drawback of two-state binding mechanisms is the requirement that n cations be taken up with folding at every [cation]. The observation that k_{dock} increases with [M] implies that $k_1 < k_2$, while the decrease in k_{undock} with [M] mandates that $k_{-2} < k_{-1}$.

Using this four-state kinetic model to describe the [cation] dependence of k_{dock} and k_{undock}, the cation titrations can be explicitly compared. The Hill coefficient in this model, $n,$ is constrained to be common to both k_{undock} and k_{dock} as indicated in the model. Also, k_{undock} and k_{dock} are fit simultaneously to a detailed balance constraint, i.e., $\Delta G^{o}_M + \Delta G^{o}_{dock,M} = \Delta G^{o}_{dock,0} + \Delta G^{o'}_M$, or $K'_M = (k_1 k_{-2}/(k_{-1} k_2))^{1/n} K_M$. The resulting parameters, $k_1, k_{-1}, k_2, k_{-2}, K_M, K'_M$, and n, from weighted-least squares fits for each cation titration (shown in Figure 6.6), are summarized in Table 6-1. The most striking observation is that k_2 and k_{-2} are identical, with the exception of spd^{3+}. Thus, each cation has an equivalent ability to induce tetraloop–receptor docking. The dissociation constants (K_M, K'_M) for each cation increase with respect to valence, with a 200-fold higher affinity for divalents vs. monovalent, and 10-fold higher affinity for trivalents vs. divalents. The zero cation limits, k_1 and k_{-1}, are slower and faster, respectively, for monovalents than multivalents because the initial conditions of the multivalent titrations contain 100 mM NaCl.

Table 6-1 Cation-dependence of tetraloop–receptor docking/undocking from four-state kinetic model (Figure 6.6)

	k_1 (s^{-1})	k_{-1} (s^{-1})	k_2 (s^{-1})	k_{-2} (s^{-1})	K_M (mM)	K'_M (mM)	n
[b]Na$^+$	5.5 ± 1.2	21.9 ± 9.0	67 ± 11	3.8 ± 0.3	357 ± 53	82 ± 28	2.9 ± 0.5
[b]K$^+$	5.0 ± 1.4	17.7 ± 14.4	70 ± 6	3.1 ± 0.4	371 ± 17	102 ± 37	3.4 ± 0.7
[a,c]Mg^{2+}	6.7 ± 1.7	11.7 ± 2.7	60 ± 11	4.5 ± 0.5	1.5 ± 0.7	0.24 ± 0.18	1.7 ± 0.5
[c]Ca^{2+}	7.0 ± 1.9	11.5 ± 2.6	67 ± 12	5.8 ± 0.5	1.8 ± 0.4	0.53 ± 0.24	2.4 ± 0.7
[c]Co(NH$_3$)$_6$$^{3+}$	7.8 ± 1.0	12.3 ± 2.1	60 ± 6	4.4 ± 0.4	0.08 ± 0.01	0.02 ± 0.01	1.9 ± 0.4
[c]Spd^{3+}	5.3 ± 1.0	10.5 ± 1.8	22 ± 6	5.3 ± 0.5	0.34 ± 0.26	0.05 ± 0.05	1.1 ± 0.4
[c,e]Mg^{2+} (U$_7$)	12.6 ± 0.9	8.6 ± 0.7	156 ± 23	5.4 ± 0.2	1.3 ± 0.3	0.25 ± 0.08	1.8 ± 0.2

Parameters are determined from fits of the cation titration to a four-state model (Figure 6.6) with a detailed balance constraint of $K'_M = (k_1 k_{-2}/(k_{-1} k_2))^{1/n} K_M$.
[a] Data refit from Ref (106) with a detailed balance constraint.
[b] Titrations performed in 50 hemisodium HEPES (pH 7.5), 100 µM EDTA. Fits of the Na$^+$ and K$^+$ titrations give independent measurements of the docking/undocking rate constants in the absence of added cation (k_1 and k_{-1}) at these buffer conditions.
[c] Titrations performed in 50 hemisodium HEPES (pH 7.5), 100 µM EDTA, 100 mM NaCl. Each titration fit at these conditions is an independent measurement of k_1 and k_{-1}, the docking and undocking rate constants in the absence of added cation.
[e] The A$_7$ linker in the RNA construct (Figure 6.1) is replaced with a U$_7$ linker (see Chapter 5).

In terms of affinity (referred to here as K_M or K'_M) the cation valence is the most important factor. The cation affinity is also greater to the docked vs. undocked conformation, i.e., K_M is larger than K'_M (see Table 6-1). There also appears to be an increase in the cooperativity with size of the cation, with the exception of spermidine^{3+}, e.g., Ca^{2+} appears slightly more cooperative than Mg^{2+}. In spermidine^{3+}, the charges are separated by covalent bonds, and thus size is not the only factor distinguishing it from the other cations. Replacement of the A$_7$ linker with a U$_7$ alternative (Figure 6.1 *A*) has no effect on K_M, K'_M, or n, suggesting that the cation uptake is intrinsic to the tetraloop–receptor docking interaction, rather than the linker in the construct design. However, the kinetics of docking are slowed in the A$_7$ construct, which supports that the more rigid A$_7$ linker can hinder achievement of the transition state (Chapter 5).

The fraction docked analysis can be directly compared to bulk E$_{FRET}$ measurements, where a bulk E$_{FRET}$ value is simply the population-weighted average of molecules in the docked vs. undocked states, and therefore should directly mirror a fraction docked plot. There is good agreement from our kinetic data with the previous bulk studies (Table 6-2). However, there does appear to be a tendency of the bulk value to overestimate the midpoint. We believe this is likely due to a less precise method of de-

termining E_{FRET}. In particular, if one does not account for differences in collection efficiency of the donor and acceptor imposed to calculate E_{FRET}, the midpoint would shift. In this case, for example a higher collection of donor signal would make it appear as if it required a higher [cation] to shift the equilibrium. Nevertheless, the single-molecule and bulk midpoints are in good agreement. Our observations for midpoints are in also excellent agreement with our freely diffusing single-molecule studies (Table 6-2).

Table 6-2 Comparison of the cation midpoint concentration for the fraction of tetraloop–receptor docked from single-molecule (immobilized and freely diffusing) and bulk FRET methods

	Immobilized $M_{1/2}$ [a] (mM)	Bulk $M_{1/2}$ [b] (mM)	Diffusing $M_{1/2}$ [d] (mM)	Immobilized n	Bulk n	Diffusing n
[c]**Na+**	140 ± 53	220 ± 9	180 ± 30	2.9 ± 0.5	2.04 ± 0.1	1.3 ± 0.3
[c]**K+**	157 ± 66	200 ± 19	nd	3.4 ± 0.7	1.89 ± 0.1	nd
[d]**Mg^{2+}**	0.42 ± 0.32	0.87 ± 0.11	0.46 ± 0.04	1.7 ± 0.5	0.95 ± 0.05	2.0 ± 0.4
[d]**Ca^{2+}**	0.77 ± 0.36	0.83 ± 0.02	nd	2.4 ± 0.7	1.05 ± 0.05	nd
[d]**Co(NH$_3$)$_6$$^{3+}$**	0.032 ± 0.16	0.017 ± 0.03	nd	1.9 ± 0.4	0.99 ± 0.1	nd
[d]**Spd^{3+}**	0.11 ± 0.12	0.26 ± 0.05[e]	nd	1.1 ± 0.4	1.0 ± 0.12[e]	nd
[d]**Mg^{2+} (U$_7$)**	0.33 ± 0.11	0.63 ± 0.07	nd	1.8 ± 0.2	1	nd

nd = not determined.
[a] $M_{1/2}$ is the midpoint concentration between the initial and maximum value of the fraction docked in the single-molecule system (Figure 6.5), where $M_{1/2} = \{[k_2/(k_2+k_{-2})]/[k_1/(k_1+k_{-1})]\}^{1/n} K'_M$ as solved for steady state condition of the kinetic scheme in Figure 6.6 *B*.
[b] Previously reported bulk FRET measurements of the midpoint (124)
[c] Titrations performed in 50 hemisodium HEPES (pH 7.5), 100 µM EDTA.
[d] Titrations performed in 50 hemisodium HEPES (pH 7.5), 100 µM EDTA, 100 mM NaCl.
[e] This value was not previously reported and is determined in this work using identical methods to *b* (124). Data is shown in *SI* Figure 6.11.

6.5   Discussion

RNA folding implicitly requires overcoming inter-phosphate electrostatic repulsion in the RNA backbone. The non-periodic charge distribution of folded RNAs make this a fundamentally complicated problem to address from theoretical electrostatic perspective (12,231). Although much progress is being made in predicting spatial distributions of cations around structurally characterized RNA molecules (84,170,172,231,245), experimental data on the magnitude and origin of cation stabilization is lacking (231). Moreover, it is becoming increasingly clear that the structural ensemble needs to be considered

to predict the effect that cations will have on the folding equilibrium (217). In the tetraloop–receptor construct, the folding landscape is significantly simplified, in that the molecule can only undock and dock, which minimizes the potential for kinetic traps and folding intermediates that are so often observed in the rugged folding landscape of RNAs (39). With single-molecule FRET methods, extraction of k_{dock} and k_{undock} allows one to observe the kinetic origin of cationic effects on a folding equilibrium. In particular, the changes in k_{dock} and k_{undock} allow one to assess if the equilibrium is perturbed by way of stabilization of a folded structure, the transition state for folding, or both.

Cations can interact through specific coordination to RNA structures or diffusely through a hydration shell. The latter can explain the majority RNA-ion interactions (41). Previous work showed that the GAAA tetraloop–receptor structure is identical over a vast range of cationic environments (82). An ensemble FRET study also revealed that tetraloop–receptor docking could be promoted with various cations (124). Cobalt hexamine is unable to directly coordinate to RNA structures, whereas for example a hexahydrated Mg^{2+} can replace an H_2O ligand to make a direct contact with RNA (41,246). The ensemble FRET studies revealed that $Co(NH_3)_6{}^{3+}$ can effectively promote tetraloop–receptor, showing that diffuse ion interactions can satisfy the role of cations in docking (124). Therefore, studying the explicit docking and undocking of a GAAA tetraloop and its receptor is a novel way to investigate the role of diffuse cation interactions in the kinetics of RNA tertiary structure formation. The cations surveyed in this work accelerate tetraloop–receptor docking, while decelerating undocking, though valence and size clearly play a role, as is examined in the following sections.

6.5.1 Effect of Cations on Docking Reaction Coordinate

One can use simple arguments to characterize the thermodynamics of the tetraloop–receptor docking scheme (Figure 6.6 *B*). First, we consider the case of [Mg^{2+}]. In the absence of Mg^{2+}, the equilibrium constant (K_{dock}) or ratio of docking to undocking rates (k_1/k_{-1}) is 0.57(2) (Table 6-1). From equilibrium considerations, $K_{dock} = e^{-\Delta G^{o}{}_{dock}/RT}$, which implies that $\Delta G^{o}{}_{dock,0}$ = 0.3(2) kcal/mol, i.e. the docked state is higher in free energy than the undocked state. At high [Mg^{2+}], on the other hand, the ratio of the docking

to undocking rate (k_2/k_{-2}), yields $\Delta G^{\circ}_{dock,M}$ = −1.5(1) kcal/mol. Thus, Mg^{2+} bound docked state is substantially *lower* in free energy than the Mg^{2+} bound undocked state (Figure 6.6 *B*). The thermodynamic benefit of docking via the Mg^{2+} folding pathway ($\Delta G^{\circ}_{dock,M}$) over the initial pathway ($\Delta G^{\circ}_{dock,0}$) is therefore to decrease in the docking free energy ($\Delta\Delta G^{\circ}_{dock}$) by 1.8(2) kcal/mol. This analysis is applied to all of the cations, as summarized in Table 6-3. The cation stabilization is similar for all cation with the exception of spermidine[3+]. Similarly, we can calculate the free energy of binding one cation ($\Delta G^{\circ}_{M}/n$ and $\Delta G^{\circ}_{M}/n$) from the apparent dissociation constant, which reveals that the multivalent cations bind more favorably to the RNA than monovalents (Table 6-3), i.e., the affinity is dictated by valence, as will be discussed in more detail in the following section.

Table 6-3 Dependence of tetraloop–receptor docking cycle on cations of varying charge density (ξ)

	Radius (Å)	ξ (Å³)	$\Delta G^{\circ}_{M}/n$[d]	$\Delta G^{\circ}_{M}{}'/n$[d]	$\Delta G^{\circ}_{dock,0}$[d]	$\Delta G^{\circ}_{dock,m}$[d]
[a]Na[+]	2.5	0.0153	−0.60 ± 0.08	−1.5 ± 0.2	0.8 ± 0.3	-1.7 ± 0.1
[a]K[+]	2.7	0.0121	−0.58 ± 0.03	−1.3 ± 0.2	0.7 ± 0.5	-1.8 ± 0.1
[a]Mg[2+]	2.07	0.0538	−3.8 ± 0.3	−4.9 ± 0.4	0.3 ± 0.2	-1.5 ± 0.1
[a]Ca[2+]	2.33	0.0377	−3.7±0.1	−4.4 ± 0.3	0.3 ± 0.2	-1.4 ± 0.1
[b]Co(NH$_3$)$_6$[3+]	1.97	0.0937	−5.51 ± 0.08	−6.3 ± 0.3	0.3 ± 0.1	-1.5± 0.1
[c]Spd[3+]	–	0.0089	−4.7 ± 0.4	−4.8 ± 0.1	0.4 ± 0.1	-0.8 ± 0.2

[a] Radius is defined as the metal–oxygen distance for hydrated cations (7,247,248)
[b] Radius is defined as the Co-N bond length (249).
[c] Density calculated using van der Waals volume of 337 Å³ (238).
[d] The free energy for 1 metal binding $\Delta G^{\circ}_{M}/n = -RT \ln ([M]/K_M)$ and $\Delta G^{\circ}_{M}{}'/n = -RT \ln ([M]/K'_M)$, where $[M]$ = 1 molar at standard state conditions and T is room temperature (294 K).
[e] $\Delta G^{\circ}_{dock,0} = -RT \ln k_1/k_{-1}$ and $\Delta G^{\circ}_{dock,M} = -RT \ln k_2/k_{-2}$

With explicit determination of k_{dock} and k_{undock}, we can also extract the effect of cations on the free energy barrier heights for both docking and undocking. From generalized transition-state theory, the reaction rate constant (e.g., k_{dock} or k_{undock}) can be written as:

$$k = v e^{(-\Delta G^{\ddagger}/RT)},$$

6.5

where $\Delta G^{\ddagger}$ is the activation free energy and v is the attempt frequency for barrier crossing (206,207). Accurate determination of $\Delta G^{\ddagger}$ is limited by knowledge of v. However, since the dependence of the reac-

tion rate on this parameter is logarithmic,

$$\frac{\Delta G^{\ddagger}}{RT} = \ln \upsilon - \ln k \text{ ,}$$

6.6

an estimate of $\upsilon \approx 10^{13}$ s^{-1} proves sufficient for our purposes, based on typical frequencies ($\sim$300 cm^{-1}) for low frequency skeletal motions (208,209). Furthermore, changes in the barrier height with [cation], i.e., $\Delta\Delta G^{\ddagger}$, are independent of υ because this offset parameter falls out upon subtraction (Eq. 6.6). For Ca^{2+}, Mg^{2+}, and Co(NH$_3$)$_6$$^{3+}$, the same docking rates are observed for the cation dependent and independent pathways, Table 6-1. Thus, the $\sim$12-fold increase in docking rate with increasing concentration of these cations (Figure 6.4 and Table 6-1) translates into a 1.26(2) kcal/mol *decrease* in the activation barrier, from 16.32(2) kcal/mol to 15.06(1) kcal/mol, whereas the docked state drops 1.78(3) kcal/mol (Figure 6.7). K$^+$ and Na$^+$ also achieve the same docking rate (k_2), but the cation independent docking rate (k_1) is slightly lower because the 100 mM NaCl background aids k_1 in the multivalent scenario. Large barriers, as observed, here are common for RNA tertiary folding, for example, the barrier height for folding of the P4–P6 domain, is 16 kcal/mol at 35 °C at 10 mM Mg^{2+} (101).

Although, spermidine^{3+} affects tetraloop–receptor folding in the same concentration range as
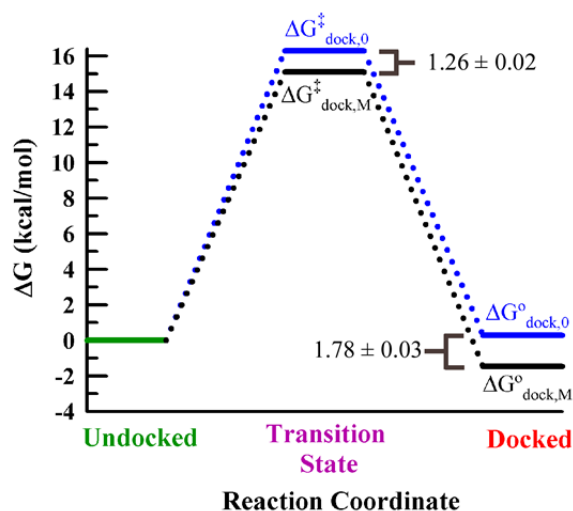
Figure 6.7   Change in free energy (kcal/mol) for tetraloop–receptor docking/undocking in the presence and absence of added cation (see Figure 6.6 *B*). The barrier and overall free energy changes are calculated from the average values of the rate constants k_1, k_2, k_{-1}, and k_{-2} for Mg^{2+}, Ca^{2+}, and Co(NH$_3$)$_6$$^{3+}$ (see Table 6-1) from Eq. 6.6.

Co(NH$_3$)$_6$$^{3+}$, it has a much smaller effect on k_{dock} and k_{undock}. With only a ~4 fold increase in docking rate, the activation barrier only drops by 0.8(2) kcal/mol with saturating [spermidine^{3+}] vs. 1.26(2) for Mg^{2+}. Similarly, the docked state with spd^{3+} is less stabilized than with Mg^{2+}—1.2(2) vs. 1.78(3) kcal/mol. The origin of this decreased ability of spermidine^{3+} to enable docking can be rationalized in terms of its structure. Spermidine^{3+} is a large flexible molecule with charges distributed along a chain (Figure 6.2). Therefore, when counterions condense on the RNA, significant volumes are excluded from charge screening. The large spd^{3+} may not be able properly localize in regions of high charge density to fully screen the negative repulsions. With its large size, repulsions between spermidine^{3+} cations can also become important and prevent further condensation of the additional spermidine cations that would be required to fully screen the RNA backbone. In such a picture the RNA molecule is saturated with spermidine^{3+}, yet might not be effectively screened, thus k_{dock} and k_{undock} saturate with less dynamic range than for the smaller cations. This effect is not unprecedented; for example spermidine^{3+} was shown to stabilize the folded *Tetrahymena* ribozyme less effectively and induced less compact structures than Co(NH$_3$)$_6$$^{3+}$ (167,213). Additionally, in studying the role of polyamines on the folding of the *Tetrahymena* ribozyme, it was observed that when excluded volumes become large, folding does not go to completion (238), similar to what we see here for tetraloop–receptor folding. By transition state investigations it was shown that the larger the polyamine, the broader and less compact the transition state (7), which could be a possible origin of the reduced effect spermidine^{3+} vs. the other cations for lowering the barrier height for tetraloop–receptor folding.

In our previous work, we showed that the drop in barrier height with increasing [Mg^{2+}] was of entropic origin. We ascribed this effect to reduced entropic cost of counterion uptake with folding and reduction of disorder in the unfolded RNA (Chapter 5). The question remains as to whether this effect will hold for the other cations explored in the work, but the quantitative similarities between the effects of the cations on the RNA folding reaction coordinate (Figure 6.7 and Table 6-1) suggest a similar origin for cation-facilitated folding.

6.5.2 Role of Valence and Size in Counterion Binding Affinities

The apparent binding affinities (K_M or $K_M{}'$) for the various cations explored in this work are correlated with the cation charge. For cations of the same charge, there is no discernible effect of ion size on the ion-RNA attraction strength (e.g., $K_M(Na^+) \sim K_M(K^+)$) (Table 6-1). According to Manning (counterion) condensation theory, increasing counterion charge increases a cation's ability to condense on polyanionic species such as RNA (250). Specifically, small ions, like Mg^{2+}, can condense on the RNA at much lower concentrations than monovalents, and thereby, neutralize phosphate charges with no effect on the solution's ionic strength (12,128,172,174,176,178). Furthermore, with a greater charge density and less excluded volume, multivalent cations more effectively screen negative charges. Likewise, fewer multivalent than monovalents cations must condense to achieve similar charge neutralization (167,213,214,250). From Manning counterion condensation theory, the effective residual charge per RNA phosphate after condensation (υ) is:

$$\upsilon = -\ln\phi \frac{R_G}{l_b N} \frac{1}{Z},$$

6.7

where N is the number of nucleotides in the RNA, R_G is the radius of gyration of the polymer, l_b is the Bjerrum length (~7.1 Å in water), Z is the counterion charge, and Φ is the volume fraction of counterions ($\Phi = N_A C V_C$, with N_A Avogadro's number, C the concentration of the cation, and V_C the volume per counterion) (167,250). Thus, for the same charge neutralization to be achieved for cations of different valence requires that $\ln \Phi$ scale with $1/Z$. Since Φ is proportional to counterion concentration, $\ln C$ must increase by $1/Z$ to maintain the same reduction in net charge on the RNA if Z is decreased. Comparison of the cation dissociation constants offers a means to compare the relative efficacy of the cations in condensation, since these values correspond to a similar effect on the folding transition (Figure 6.4). Thus, from Eq. 6.7, we might expect a correlation of $\ln (K_M)$ and $\ln (K_M{}')$ with $1/Z$. Indeed, a simple linear increase of $\ln (K_M)$ and $\ln (K_M{}')$ with $1/Z$ is observed (Figure 6.8 *A* and *B*). This trend implies also that the free energy of the ion-RNA interaction ($\Delta G°_M/n$ and $\Delta G°_M{}'/n$) becomes linearly less favorable (increases) with $1/Z$ (Table 6-3). A similar valence dependence of the midpoint of a folding titration the

Figure 6.8   Trends in apparent cation dissocation constants and hill coefficients for cations with charge (Z) and charge density determined from the four-state kinetic model for tetrlaoop receptor docking (Figure 6.6 *B*). Symbols are color coded as Na$^+$ (*black*), K$^+$ (*red*), Ca^{2+} (*green*), Co(NH$_3$)$_6$$^{3+}$ (*purple*), Spd^{3+} (*dark red*) and the larger and smaller cation for a given charge as triangles and squares, respectively. (A) and (B) The *ln K$_M$* and *ln K'$_M$* vary linearly with 1/*Z*. (C) The Hill coefficient (*n*) decreases a function of charge density. Spermidine^{3+} is not shown because the charge density is not a comparable quantity for the polymer with charge distributed across the change. An exponential fit is shown to guide the eye. For charge density determination see Table 6-3.
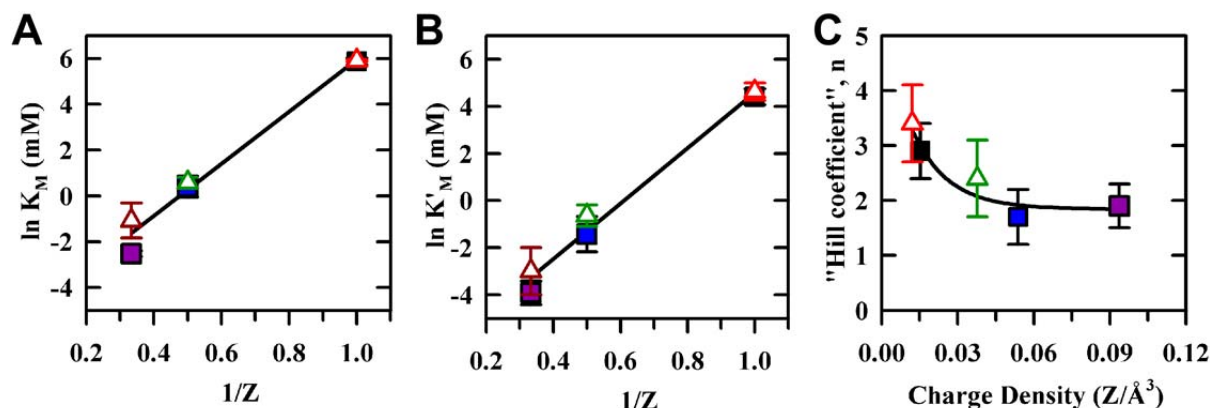
Tetrahymena ribozyme was also observed (251). Thus, with knowledge of a single titration, one can easily extrapolate the dissociation constants (binding energies) as function of valence. Melting studies of a dual tetraloop–receptor construct also supported that Na$^+$ and K$^+$ have similar affinities for the RNA (83). However, the observed dependence of k_{dock} and k_{undock} is not entirely independent of ion size. The stoichiometry (*n*) of the ion folding pathway shows sensitive to charge density (Figure 6.8 *C*)—*n* decreases with charge density.

This stoichiometry can be correlated with the binding pockets in the receptor. Two cation binding sites have been crystollographically identified in the receptor (i) a magnesium coordinated to the G10 phosphate oxygen (Figure 6.1 *A*) of the receptor (62,64,78,79) (ii) a monovalent ion (K$^+$) coordination site below the adenosine platform nucleotides (A4 and A5,) (64,71,80). An NMR studies revealed that both of these sites could be filled by Mn^{2+}, with the G10 binding also satisfied by CO(NH$_3$)$_6$$^{3+}$ (82). It was shown by Draper and coworkers in tRNA that the stoichiometry of a four-state binding model is correlated with the crystallographic binding sites, supporting that the stoichiometry reflects the localization of a small number of fully hydrated, localized cations (173). The crystallographic sites correlate

with pockets of negative electrostatic potential according to NLPB calculations in both tRNA and a tetraloop–receptor complex (82,173). The cooperativity of ~2 for the divalent cations and $Co(NH_3)_6^{3+}$ (Figure 6.8 *C*) can support that the stoichiometry corresponds to the two structurally identified binding sites. Spermidine^{3+} may simply be too large to localize efficiently in the binding pockets, which may be the origin of the decrease in *n* for this cation. In the absence of divalents, it would be expected that more monovolents would be required to satisfy the same charge neutralization, consistent with the ~1.7 fold increase in *n* from Mg^{2+} to Na^+ and in agreement with Poisson Boltzmann calculations that Mg^{2+} can replace 1.9 Na^+ ions (174). Alternatively, the increase in cooperatively may simply reflect that in the multivalent studies the presence of the 100 mM NaCl background may aid in filling of the putative monovalent binding sites or reduce electrostatic frustration of the RNA (128). Although the stoichiometry can offer insights into cation binding sites, this inference is limited because the ion atmospheres of both the docked and undocked states are significantly altered as cations are added. For example, in the absence of Mg^{2+}, the RNA is predominantly decorated with Na^+ ions. Increasing $[Mg^{2+}]$ displace Na^+, with Mg^{2+} condensation differing for the docked and undocked states. Thereby, the observed cation uptake with folding changes as a function of Mg^{2+} concentration. Therefore, not only does the stoichiometric amount of Mg^{2+} taken up change with $[Mg^{2+}]$, the initial (Mg^{2+} unbound state) is surrounded by more Mg^{2+}. Furthermore, changes in ion atmosphere can alter the unfolded structural ensemble. However, it has been shown that in the presence of high concentrations of monovalents, the hill coefficient can be correlated with specific binding sites (252).

According to theories for counterion distributions around RNA, increasing the bulk concentration of cations causes an excess of cations to condense on RNA (178,214,250,253,254). The benefit of this condensation to RNA folding is proposed to be a decrease in the electrostatically repulsive barrier to helix packing (178,214,253). However, we have recently shown that the origin of $[Mg^{2+}]$ facilitated tetraloop–receptor docking is entropic (Chapter 5). Recent studies have shown that nucleic acid duplexes have such a propensity to localize counterions that charge neutralization in the local ion atmosphere is achieved even under very modest [salt] (176). Thus, we proposed that as RNA folds, it recruits cations

to combat electrostatic repulsion. Increasing [cation] decreases the entropic penalty for localizing additional cations with folding or eliminating the need for cation uptake (Chapter 5). The role of concentration in cation uptake with RNA folding is explored in more detail in the following section.

6.5.3   Accumulation of Cations on the Docked vs Undocked RNA ($\Delta\Gamma_M$)

The net cation uptake with folding is not the same quantity as the Hill coefficient in the four-state folding scheme, as it is for a two-state binding scheme (215). In the kinetic model used to describe folding of the tetraloop–receptor interaction, the RNA can fold via two pathways, with both the undocked and docked states accumulating cations with increasing cation concentration (Figure 6.6 *B*).   The fraction of the undocked population with cation (*M*) bound, fUM_n, can be written as,

$$fUM_n = \frac{[UM_n]}{[UM_n]+[U]} = \frac{[M]^n}{[M]^n + K_M^{\,n}},$$

$$6.8$$

where [UM_n] and [U] are the undocked populations with M bound and unbound states, respectively. A similar expression can be written for the docked state, fractional population of the docked state with M bound, fDM_n is

$$fDM_n = \frac{[DM_n]}{[DM_n]+[D]} = \frac{[M]^n}{[M]^n + K_M'^{\,n}}.$$

$$6.9$$

For an excess of Cl- in the solution, the preferential interaction coefficient (Γ_M) can be used to quantify chemical potential of the RNA as a function of salt concentration (215,255,256). Γ_M is a parameter that characterizes the excess counterions around the RNA relative to the bulk solution. The net cation uptake with folding, $\Delta\Gamma_M$, is the difference between this interaction coefficient for the folded vs unfolded RNA ($\Delta\Gamma_M = \Gamma_{M,docked} - \Gamma_{M,undocked}$). $\Delta\Gamma_M$ has been linked to the change in free energy for the observed folding process (or $\Delta G°_{dock}$ in this case) as a function of cation concentration ([*M*])and can be written as:

$$\Delta\Gamma_M \approx -\frac{1}{RT}\frac{\partial \Delta G°_{dock}}{\partial \ln[M]}.$$

$$6.10$$

For the four-state kinetic model (Figure 6.6 *B*), $\Delta\Gamma_M$ can also be written as

$$\Delta\Gamma_M \approx (fDM_n)\, n - (fUM_n)\, n.$$

<div align="right">6.11</div>

where fDM_n and fUM_n are determined at a given $[M]$ from the parameters, K_M, $K_M{}'$, and n per Eqs. 6.8 and 6.9. Thus, using either Eq. 6.10 or 6.11 one can characterize the net cation uptake.

A plot of $\Delta G^o{}_{dock}$ vs $ln\,[M]$ (Figure 6.9 A) as calculated from the fits of k_{dock} and k_{undock} as a function of [cation], where $K_{dock} = k_{dock}/k_{undock}$ and $\Delta G^o{}_{dock} = -RT \ln K_{dock}$. The initial slope of the plot is ~0 for each cation, thus one can anticipate the net cation uptake ($\Delta\Gamma_M$) is ~0 according to Eq. 6.10. The slope of the plots steepens ($\Delta\Gamma_M > 0$), then flattens ($\Delta\Gamma_M \sim 0$). The calculated cation uptake ($\Delta\Gamma_M$) vs. $[M]$ is shown for each of the cations in Figure 6.9 B, as determined from Eq. 6.11. Utilization of Eq. 6.11 yields indistinguishable results; Eq. 6.10 is selected as it enabled propagation of error from the uncertainties in K_M, $K_M{}'$, and n. At low [cation] concentrations the uptake is ~0, followed by a steep rise that saturates at a value near the Hill coefficient, before again dropping. At high [cation] the uptake again decreases, since the equilibrium for cation accumulation of the undocked and docked states is saturated. As expected from the Hill coefficients, the uptake correlates with charge density (Figure 6.9). Measurements of $\Delta\Gamma_M$ for a series of group I monovalent cations (e.g., Na+ and K+) for the bimolecular dual tetraloop−receptor complex also support that charge density (or ion size) alters ion interactions with the RNA (83). Our observed trends in cation uptake are consistent with explicit measurement of the excess Mg^{2+} by the unfolded and folded RNA, where the initial Mg^{2+} is zero and approaches a value near the Hill coefficient (215). Analysis of the tetraloop−receptor docking process in terms of the four-state kinetic model, thus captures that cation uptake can vary as a function of salt and that charge density—even for cations of the same valence—affects RNA-ion interactions (83,215). This observation supports that RNA-ion interactions theories must incorporate size effects to explain the curvature of a cation titration (234,237). Furthermore, these results emphasize that cation uptake with folding can be significant and therefore, the unfolded and folded states must both be addressed to predict the thermodynamics of folding as a function of [salt].

From this analysis of the net cation uptake, the thermodynamic effect of increasing cation concentration becomes clear. Increasing cation concentration can decrease the net entropic penalty of coun-

Figure 6.9   Free energy for tetraloop–receptor and net uptake of cation with increasing [cation]. (*A*) A plot of ΔG^{o}_{dock} vs ln [cation] as calculated from the fits of k_{dock} and k_{undock} (Figure 6.4) as a function of [cation], where $K_{dock} = k_{dock}/k_{undock}$ and $\Delta G^{o}_{dock} = -RT \ln K_{dock}$. (*B*) The net cation uptake ($\Delta \Gamma_M$) as a function of [cation] calculated from Eq. 6.11. Error bars (shown in *gray*) are propagated from the uncertainties in K_M, $K_M{'}$, and n.

terion uptake. Consider, for example, the multivalent titration, which begin at 100 mM NaCl. At these

initial conditions, the RNA takes up ~1.7 Na⁺ cations to fold (Figure 6.9). Increasing the [Mg²⁺] lessens

the number of Na+ needed as the RNA can preferentially take up Mg²⁺ to which it is more attracted

(Figure 6.9). Mg^{2+} can satisfy the cation uptake with fewer ions, and is therefore entropically beneficial, as seen by the maximum uptake; at saturation, the net accumulation of Mg^{2+} is ~1.1 Mg^{2+} ions vs. ~2.4 Na^+ ions. At high [Mg^{2+}] both the docked and undocked states are saturated with Mg^{2+} such that no cation uptake is needed for folding. Thus, increasing [cation] can aid folding by decreasing the entropic penalty of counterion uptake and by decreasing the number of cations needed. The overall interactions of the cations with the RNA is, of course, more complicated than this, as Na^+ and Mg^{2+} can compete, and anions can be depleted from the region around the RNA (174). Furthermore, increasing [cation] can change the unfolded structures of the RNA, e.g., by organizing the receptor or increasing base stacking in the poly(A)-linker, which can alter RNA folding thermodynamics by reducing the conformational entropy of the undocked state (202,203). Thus, dissection of the thermodynamics parameters determined for the cation-mediated tetraloop–receptor docking cycle still remains a challenge because the explicit ion atmosphere and structures must be considered. These studies, however give quantitative insight into energetics and kinetics of cation-mediated folding.

6.6   Conclusions

A two-state metal binding scheme is insufficient to explain the origin of a [cation]-dependent k_{dock} and k_{undock} for the tetraloop–receptor interaction. A four-state kinetic model can describe the increase in k_{dock} and decrease in k_{undock} with [cation], whereby the cations are also more attracted to the docked than undocked RNA. With this model we can separate the free energies of cation-RNA interactions from the kinetic of docking/undocking. Furthermore, this model allows for quantitation of the differential affinities for cations in the folded and unfolded conformations of the RNA, an essential measurement for understanding RNA-ion interactions. We also determine the effective cation uptake with folding by way of the preferential interaction coefficient. Ion size and valence affect the preferential interaction of the cation with the docked vs. undocked RNA, which should prove to be a useful measurement for testing theoretical models for ion-RNA interactions (40). Cation valence is the major determinant in predicting the efficiency of a cation to induce folding. Each of the cations studied promotes the same folding rate

and equivalently stabilizes the docked state, with the exception that spermidine^{3+} is less effective at promoting docking. The size and structure of spermidine^{3+} likely prevent the localization of cations needed to properly aid folding. Utilization of a four-state kinetic model reveals a possible physical origin of the cation dependence in terms of the decreased entropic penalty of cation uptake with folding. It would be interesting, if the [cation] dependence is affected by receptor sequence, to test the role of the metal ion binding sites in cation localization/uptake. Furthermore, investigation of the underlying enthalpic and entropic landscapes in the presence of various cations will give insights into the structural and electrostatic role of the cations in the folding process.

6.7 Acknowledgments

6.8 Supporting Information

6.8.1 Fluorescence Lifetime Measurements: Cy3/Cy5 Quantum Yield Ratio is Unaffected by Cationic Environment

To compare the undocked and docked conformations inferred from FRET for the tetraloop–receptor construct, as depicted in Figure 6.1, changes in donor and acceptor emission properties induced by increasing [cation] be assessed. In particular, calculating E_{FRET} ratiometrically from donor and acceptor fluorescence intensities requires knowledge of the quantum yields (Eq. 6.1). To ascertain whether or not the quantum yields are affected by the varying [cations] at experimental buffer conditions, we moni-

tor the fluorescence decays of Cy3 in Cy3-only labeled constructs and Cy5 in Cy3-Cy5 labeled tetraloop–

receptor across the range of cationic concentrations explored in this work, which reveal that the ratio of

Figure 6.10   Fluorescence lifetimes of donor Cy3 (*A*) and acceptor Cy5 (*B*) as function of [cation]. 10 mM Mg^{2+} has no effect on the fluorescence lifetimes (left). Cobalt hexamine quenches the donor and acceptor to same extent. Lifetimes are fit to bi-exponential decays convoluted with the instrument response function (IRF shown in the *left* panel). The population-weighted average (τ_{avg}) of the two lifetime components is shown. Uncertainties are indicated in parentheses. The effect of the added cation on the donor (Q_D) and acceptor (Q_A) is shown at the top of each graph (see supporting text).

Cy3/Cy5 quantum yields are unaffected over the cation ranges investigated (Figure 6.10).

Fluorescence lifetime measurements are performed in the aforementioned confocal fluorescence microscope with 500 pM solutions of Cy3 only and Cy3-Cy5 labeled constructs. Cy3 fluorescence

lifetime is measured for donor (Cy3) only constructs, whereas acceptor (Cy5) is monitored for full constructs based on donor excitation in standard buffer conditions with varying concentrations of the cations of interest with laser powers of 60 μW (102 μW/cm^2) with 30 s integration times using TCSPC with 1024 ADC channel resolution. Fluorescence decay curves are fit with a convolution of the instrument response function measured from instantaneous Raman scattering from water with exponential decays. The fluorescence lifetimes are monitored in the standard buffer conditions (with oxygen scavenger) and the specified amount of added salt.

To determine if the quantum yields are affected by the varying [cations], we monitor the fluorescence decays of Cy3-only and Cy3-Cy5 labeled tetraloop–receptor at the extreme cationic concentrations. Cy3-only constructs are necessary to monitor the donor decay without influence from FRET, which is affected by [cation]. Quantum yield is proportional to the fluorescence lifetime, i.e., $Q = k_{rad}\tau_{fluor}$, where Q is the quantum yield, k_{rad} is the radiative rate, and τ_{fluor} is the fluorescence which is $1/(k_{rad} + k_{nonrad})$. In general, environmental effects can perturb the quantum yield by increasing nonradiative decay rates. Thereby, relative changes in quantum yield are observable by the relative fluorescence lifetimes.

The fluorescence lifetimes of Cy3 and Cy5 are unaffected by even high concentration of Mg^{2+}, Ca^{2+}, Na$^+$, K$^+$, and spermidine^{3+} (Figure 6.10 *A* and *B*, *left* panel, 10 mM Mg^{2+} is shown as the example). Only cobalt hexamine affects the fluorescence lifetimes of Cy3 and Cy5 at our experimental conditions (Figure 6.10 *A* and *B*, *right* panel). There is an apparent shortening of the Cy3 and Cy5 lifetimes with the addition of [Co(NH$_3$)$_6$$^{3+}$]. Deconvolution of the instrument response function with the exponential decays reveals the magnitude of fluorescence quenching induced by the presence [Co(NH$_3$)$_6$$^{3+}$], allowing for quantification of the quenching magnitude.

The fluorescence lifetime of Cy3 has been established as a multiexponential process (113,116) . In the case of multiexponential behavior, the observed quantum yield (*Q*) on the time scale >ns is the population weighted lifetime, or average lifetime, i.e., ($Q = k_{rad}(P_1\tau_1 + P_2\tau_2 + ...) = k_{rad}\tau_{avg}$, where P_i, the probability for a given lifetime, is:

$$P_i = \frac{\int_0^{\infty} A_i e^{-\tau_i t}}{\int_0^{\infty} (A_1 e^{-\tau_1 t} + A_2 e^{-\tau_2 t} + ...)} = \frac{A_i \tau_i}{A_1 \tau_1 + A_2 \tau_2 + ...}, \qquad 6.12$$

where A_i is the amplitude for a lifetime of τ_i, and t is time. The direct proportionality of τ_{avg} and quantum yield assumes that k_{rad} is constant. Changes in the nonradiative rate can be explained by a varying environment of the fluorophore on timescales longer the fluorescence lifetime, such as a base-stacked vs freely rotation dye, as suggested by Lilley and coworkers (113). The observed Cy3 fluorescence lifetime can be well-described by double exponential decay with $\tau_1 = 0.26 \pm 0.04$ and $\tau_2 = 1.6 \pm 0.1$, as seen in Figure 6.10 and consistent with previous work (113,116). These values result in a population-weighted lifetime of $\tau_{average} = 0.94 \pm 0.12$ ns in standard buffer condition (with oxygen scavenger). Interestingly, we also see a bi-exponential character for Cy5, yielding $\tau_{average} = 1.18 \pm 0.06$ ns, though in the absence of oxygen scavenger the Cy5 lifetimes is well described by a single exponential decay (data not shown).

For Cy3, $\tau_{average}$ is reduced from 0.94 ± 0.12 ns to 0.74 ± 0.08 with the addition of 1 mM $Co(NH_3)_6{}^{3+}$, or only a 1.3 ± 0.2 -fold effect on Q_D. In the case of Cy5, $\tau_{average}$ is decreased from 1.18 ± 0.06 ns to 0.91 ± 0.05— a 1.3 ± 0.1-fold reduction in Q_A. Bulk lifetime measurements agree well with the lifetimes measured on immobilized single molecules (data not shown). These quenching effects at even the extreme condition (1 mM) are small, thus have only a negligible effect on the Förster radius (111). Furthermore, the quantum yield ratio of donor to acceptor is maintained, thus E_{FRET} can be determined with the same quantum yield correction factor in Eq. 6.1 for all experimental conditions explored in this work.

6.8.2 Ensemble Fluorometry of [Spermidine[3+]]-Dependent Tetraloop−Receptor Docking

Fluorescence spectra for a 20 nM RNA solution in standard buffer conditions are collected as a function of [spermidine[3+]] using a FluoroLog - Modular Spectrofluorometer (Horiba Scientific/Jobin Yvon) with excitation at 500 nm (4 nm bandwidth) and emission measured in 1 nm steps from 525 to 700 nm (8 nm bandwidth, 300 ms integration time), as previously described (124). The bulk E_{FRET} for each fluores-

Figure 6.11   Ensemble FRET measurements of [spermidine[3+]]-dependent tetraloop–receptor docking. (*A*) Sample fluorescence spectra of tetralooop–receptor construct at specified [spd[3+]]. The donor (Cy3) emission decreases and the acceptor (Cy5) emission increases as function of [spd[3+]], indicating an increase in E_{FRET}. (*B*) Bulk E_{FRET} value calculated from the fluorescence spectra and fit to Eq. 6.13 yielding e_i = 0.15 ± 0.02, Δe = 0.230 ±0.005, n = 1.0 ± 0.2, and $M_{1/2}$ = 0.26 ±0.05.

cence spectrum is estimated as $I_A/(I_D + I_A)$, where I_A and I_D are the integrated acceptor and donor emission from 655 to 700 nm and 549 to 620 nm, respectively.

$$E_{FRET} = e_i + \Delta e \frac{[M]^n}{M_{1/2}^n + [M]^n} ,$$
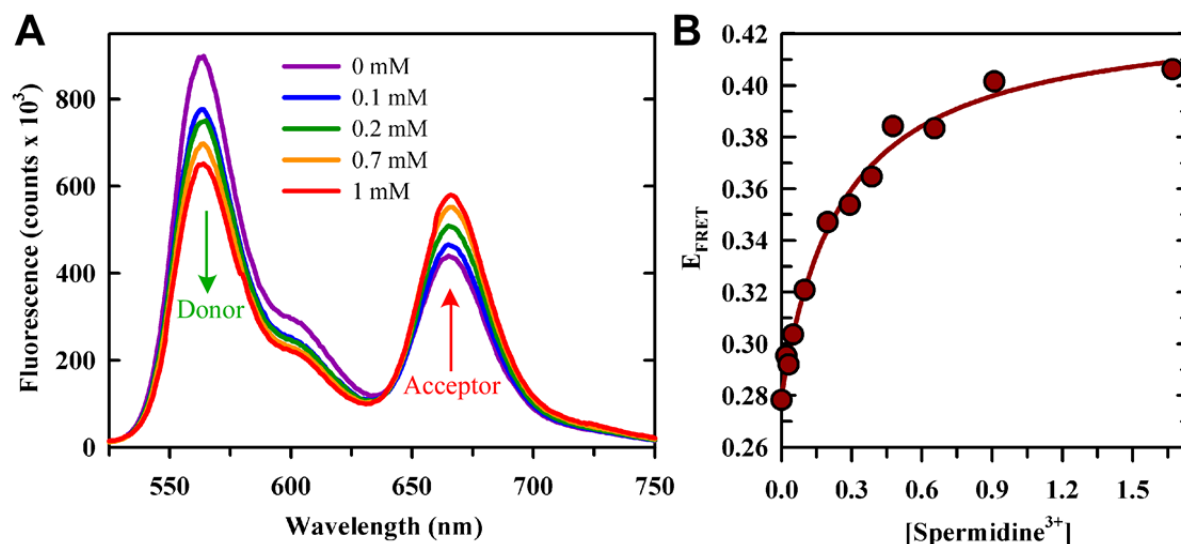
6.13

where $[M]$ is the metal ion concentration, e_i is the FRET efficiency with no metal ion added, Δe is the

maximum change in FRET efficiency induced by addition of metal ion, $M_{1/2}$ the midpoint concentration,

and *n* is the apparent Hill coefficient (124). The titration midpoint and fitting results are shown in Fig-

ure 6.11 and compare well with the single molecule observations as shown in Table 6-2.

Chapter 7   Conclusions and Future Directions

7.1   Conclusions

Structural assembly is critical to RNA biological functionality. However, the current energetic and molecular descriptions of how RNA folds pales in comparison to the understanding of protein folding, in large part because the RNA folding landscape is rugged and mediated by counterions. To gain molecular insight into the energetic barriers along RNA folding pathways, we explored the kinetics and thermodynamics for the formation of a single tertiary interaction as a function of cationic environment. Such studies have yielded information about RNA folding transition states and the role of cations in facilitating folding.

We focused on the kinetics and thermodynamics of the GAAA tetraloop −11 nucleotide receptor tertiary interaction. We determined the underlying free energy, enthalpy, and entropy of tetraloop−receptor docking as a function of [Mg^{2+}]. Quite surprisingly, we revealed an entropic origin of Mg^{2+}-facilitated RNA folding. We also showed that the tetraloop−receptor folding transition state is "early" or unbound-like and is dominated by an entropic barrier, which may be a general feature of RNA folding. The overall docking reaction is exothermic and entropically costly, consistent with the large number of hydrogen bonding and base stacking interactions that occur in this tertiary interaction. The tetraloop−receptor interaction alone can account for the exothermicity and entropic cost folding in the *Tetrahymena* P4−P6 domain (11). It was also shown that the tetraloop−receptor interaction does not require specific cations to fold, supporting the idea that the interaction of counterions with this RNA are "diffuse", i.e., not requiring specific coordination. A four-state kinetic model proved useful for quantifying the effects of cation concentration on the rate constants for docking and undocking of tertiary inter-

actions. With this model, we extracted the free energy of cation binding to the tetraloop–receptor RNA, which was shown to increase (become less favorable) linearly with the inverse charge of the cation. These kinetic studies also allowed us to identify a higher affinity for counterions to the docked vs undocked state. Ion size and valence both affect the preferential interaction of the cation with the docked and undocked RNA.

These studies also reveal a possible paradigm for the cooperativity of tertiary folding observed in large RNAs. An example of cooperativity is that the disruption of a tetraloop–receptor interaction destabilize tertiary interactions throughout a bacterial group I intron (39). Similarly, cooperativity of tertiary interactions in the *Tetrahymena* P4–P6 domain has also been observed—the overall ΔG° of folding was much more favorable than the sum of the ΔG°'s of individual tertiary interactions (92). A comparison between our isolated tetraloop–receptor docking thermodynamics and a complex containing two tetraloop–receptor interaction supports a picture that the enthalpies of multiple tertiary interactions within an RNA are additive. However, the free energy for forming two (inter-molecular) tetraloop–receptor interactions is much more favorable than twice the ΔG° of binding for a bimolecular association due to a single interaction (125). Since our data support that the enthalpies of the tertiary interactions are additive, the entropic cost of forming the second tertiary interaction must be greatly reduced by the first (125). In turn, this would imply an entropic origin of tertiary cooperativity. All of these observations may aid prediction capabilities of RNA tertiary structures and the effect of cations therein. Toward that end, other tertiary interactions will need to be studied.

7.2 Future Directions

Although the GAAA tetraloop–11 nt receptor interaction has been subjected to extensive biophysical characterization, many questions still remain. A number of 11 nt receptor mutations have been investigated for their effect on the thermodynamics of bimolecular tetraloop–receptor binding (68), but never have they been explored for an intra-molecular interaction. Mutant studies may yield insight into the evolution of the RNA enzymes that utilize these tetraloop–receptor interactions. The metal ion depend-

ence of any mutants would also further elucidate the role of the cation binding sites in the folding pathway. No mutations of the receptor or tetraloop have been explored in terms of enthalpic/entropic or kinetic effects for an isolated tetraloop–receptor interaction. Such a study could give more insight into the transition state for docking.

To assess the origin of tertiary cooperativity in large RNAs, other individual and combined tertiary interactions must be studied. In particular, the A-rich bulge interaction (Figure 1.12 *C* and *D*) can readily be characterized using the methods presented in this thesis. Considerable effort has already been made in this direction. It is crucial to characterize more tertiary interactions at the level achieved for the GAAA tetraloop–receptor interaction, both together and in isolation, to develop predictive capabilities of RNA tertiary structure. Even little is known about thermodynamics and kinetics of other types of tetraloop–receptor interactions. Furthermore, the junction (linker) can also play a crucial role in the folding outcome and counterion dependence. Thus, the role of the junctions in folding thermodynamics should be the subject of further investigations. In terms of the tetraloop–receptor interaction, it would be interesting to replace the single-stranded linker with an uncharged polymer chain, e.g., polyethylene glycol, to eliminate potential structural and counterion effects of the linker region.

It is critical that folding studies be coupled with theoretical efforts. In particular molecular dynamics simulations, when coupled with structural studies, can be incredibly powerful in elucidating RNA folding pathways (257) and will hopefully give a glimpse into the role of metal ions in RNA folding landscapes. The simplified tetraloop–receptor system could be a target for molecular dynamics simulations, such as coarse grained and nudged elastic-band methods (257,258). Such studies could help to identify the contribution of conformational entropy and ion atmosphere to RNA folding. On the same token, theories for describing the counterion mediation of electrostatic potentials, (e.g., Poisson Boltzmann) of the simplified RNA systems would aid in defining the role of counterions. In that regard, theory would need to account for differences in charge distributions in folded vs. unfolded RNA.

Lastly, other environments for RNA folding must be considered. For instance the crowded, viscous cellular medium could dramatically alter folding thermodynamics. Understanding such effects will

be essential to correlate RNA structural dynamics with function. Many RNA folding systems rely on protein folding chaperones and cofactors to fold. GNRA tetraloop–receptor interactions have been implicated in this process of RNP assembly, so the kinetic and thermodynamic origins of these effects would help elucidate the mechanisms of RNA folding in the cell.

Bibliography

1.   Sullenger, B. A., and E. Gilboa. 2002. Emerging clinical applications of RNA. *Nature*. 418:252-258.

2.   Serganov, A., and D. J. Patel. 2007. Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nat. Rev. Genet.* 8:776-790.

3.   Zhuang, X. W., H. Kim, M. J. B. Pereira, H. P. Babcock, N. G. Walter, and S. Chu. 2002. Correlating structural dynamics and function in single ribozyme molecules. *Science*. 296:1473–1476.

4.   Famulok, M., and S. Verma. 2002. In vivo-applied functional RNAs as tools in proteomics and genomics research. *Trends Biotechnol.* 20:462-466.

5.   Mathews, D. H., and D. H. Turner. 2006. Prediction of RNA secondary structure by free energy minimization. *Curr. Opin. Struct. Biol.* 16:270-278.

6.   Tinoco, I., and C. Bustamante. 1999. How RNA folds. *J. Mol. Biol.* 293:271–281.

7.   Koculi, E., D. Thirumalai, and S. A. Woodson. 2006. Counterion charge density determines the position and plasticity of RNA folding transition states. *J. Mol. Biol.* 359:446-454.

8.   Bokinsky, G., D. Rueda, V. K. Misra, M. M. Rhodes, A. Gordus, H. P. Babcock*, et al.* 2003. Single-molecule transition-state analysis of RNA folding. *Proc. Natl. Acad. Sci. U. S. A.* 100:9302–9307.

9.   Rook, M. S., D. K. Treiber, and J. R. Williamson. 1999. An optimal Mg^{2+} concentration for kinetic folding of the Tetrahymena ribozyme. *Proc. Natl. Acad. Sci. U. S. A.* 96:12471–12476.

10.   Fang, X. W., P. Thiyagarajan, T. R. Sosnick, and T. Pan. 2002. The rate-limiting step in the folding of a large ribozyme without kinetic traps. *Proc. Natl. Acad. Sci. U. S. A.* 99:8518-8523.

11.   Szewczak, A. A., E. R. Podell, P. C. Bevilacqua, and T. R. Cech. 1998. Thermodynamic stability of the P4-P6 domain RNA tertiary structure measured by temperature gradient gel electrophoresis. *Biochemistry*. 37:11162–11170.

12.   Woodson, S. A. 2005. Metal ions and RNA folding: A highly charged topic with a dynamic future. *Curr. Opin. Chem. Biol.* 9:104–109.

13.   Crick, F. 1970. Central dogma of molecular biology. *Nature*. 227:561-563.

14.   Doudna, J. A., and T. R. Cech. 2002. The chemical repertoire of natural ribozymes. *Nature*. 418:222-228.

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

16.   Cech, T. R. 2000. Structural biology: The ribosome is a ribozyme. *Science*. 289:878-879.

17.   Gesteland, R. F., T. R. Cech, and J. F. Atkins. 2006. The RNA world: The nature of modern RNA suggests a prebiotic RNA, Cold Spring Harbor Laboratory Press.

18.   Baltimore, D. 1970. Viral RNA-dependent DNA polymerase: RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature*. 226:1209-1211.

19.   Ahlquist, P. 2002. RNA-dependent RNA polymerases, viruses, and RNA silencing. *Science*. 296:1270-1273.

20.   Carninci, P., J. Yasuda, and Y. Hayashizaki. 2008. Multifaceted mammalian transcriptome. *Curr. Opin. Cell Biol.* 20:274-280.

21.   Birney, E., J. A. Stamatoyannopoulos, A. Dutta, R. Guigo, T. R. Gingeras, E. H. Margulies*, et al.* 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*. 447:799-816.

22.   Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin*, et al.* 2001. Initial sequencing and analysis of the human genome. *Nature*. 409:860-921.

23.   Mattick, J. S., and M. J. Gagen. 2001. The evolution of controlled multitasked gene networks: The role of introns and other noncoding RNAs in the development of complex organisms. *Mol. Biol. Evol.* 18:1611-1630.

24.   Gilbert, W., M. Marchionni, and G. McKnight. 1986. On the antiquity of introns. *Cell*. 46:151-154.

25.   Doolittle, W. F., and C. Sapienza. 1980. Selfish genes, the phenotype paradigm and genome evolution. *Nature*. 284:601-603.

26.   Orgel, L. E., and F. H. C. Crick. 1980. Selfish DNA: The ultimate parasite. *Nature*. 284:604-607.

27.   Grabowski, P. J., A. J. Zaug, and T. R. Cech. 1981. The intervening sequence of the ribosomal-RNA precursor is converted to a circular RNA in isolated-nuclei of *Tetrahymena*. *Cell*. 23:467-476.

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

29.   Steitz, T. A., and P. B. Moore. 2003. RNA, the first macromolecular catalyst: The ribosome is a ribozyme. *Trends Biochem.Sci.* 28:411-418.

30.   Gilbert, W. 1986. Origin of life: The RNA world. *Nature*. 319:618-618.

31.   Sharp, P. A. 1987. Splicing of messenger: RNA precursors. *Science*. 235:766-771.

32.   Cech, T. R., A. J. Zaug, and P. J. Grabowski. 1981. In vitro splicing of the ribosomal RNA precursor of tetrahymena: Involvement of a guanosine nucleotide in the excision of the intervening sequence. 27:487.

33.   Guo, F., A. R. Gooding, and T. R. Cech. 2004. Structure of the *Tetrahymena* ribozyme: Base triple sandwich and metal ion at the active site. *Mol. Cell*. 16:351-362.

34.   Quigley, G. J., and A. Rich. 1976. Structural domains of transfer-RNA molecules. *Science*. 194:796-806.

35.   Greenleaf, W. J., K. L. Frieda, D. A. N. Foster, M. T. Woodside, and S. M. Block. 2008. Direct observation of hierarchical folding in single riboswitch aptamers. *Science.* 319:630–633.

36.   Brion, P., and E. Westhof. 1997. Hierarchy and dynamics of RNA folding. *Annu. Rev. Biophys. Biomolec. Struct.* 26:113–137.

37.   Hendrix, D. K., S. E. Brenner, and S. R. Holbrook. 2005. RNA structural motifs: building blocks of a modular biomolecule. *Q. Rev. Biophys.* 38:221-243.

38.   Chauhan, S., G. Caliskan, R. M. Briber, U. Perez-Salas, P. Rangan, D. Thirumalai, *et al.* 2005. RNA tertiary interactions mediate native collapse of a bacterial group I ribozyme. *J. Mol. Biol.* 353:1199–1209.

39.   Chauhan, S., and S. A. Woodson. 2008. Tertiary interactions determine the accuracy of RNA folding. *J. Am. Chem. Soc.* 130:1296–1303.

40.   Draper, D. E. 2004. A guide to ions and RNA structure. *RNA-Publ. RNA Soc.* 10:335–343.

41.   Draper, D. E., D. Grilley, and A. M. Soto. 2005. Ions and RNA folding. *Annu. Rev. Biophys. Biomolec. Struct.* 34:221–243.

42.   Batey, R. T., R. P. Rambo, and J. A. Doudna. 1999. Tertiary motifs in RNA structure and folding. *Angew. Chem.-Int. Edit.* 38:2327–2343.

43.   Tamura, M., and S. R. Holbrook. 2002. Sequence and structural conservation in RNA ribose zippers. *J. Mol. Biol.* 320:455-474.

44.   Woese, C. R., S. Winker, and R. R. Gutell. 1990. Architecture of ribosomal-RNA: Constraints on the sequence of Tetra-loops. *Proc. Natl. Acad. Sci. U. S. A.* 87:8467-8471.

45.   Michel, F., and E. Westhof. 1990. Modeling of the 3-Dimensional architecture of group-I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.* 216:585-610.

46.   Michel, F., M. Hanna, R. Green, D. P. Bartel, and J. W. Szostak. 1989. The guanosine binding-Site of the *Tetrahymena* ribozyme. *Nature.* 342:391-395.

47.   Klosterman, P. S., D. K. Hendrix, M. Tamura, S. R. Holbrook, and S. E. Brenner. 2004. Three-dimensional motifs from the SCOR, structural classification of RNA database: extruded strands, base triples, tetraloops and U-turns. *Nucleic Acids Res.* 32:2342-2352.

48.   Antao, V. P., and I. Tinoco. 1992. Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops. *Nucleic Acids Res.* 20:819–824.

49.   Antao, V. P., S. Y. Lai, and I. Tinoco. 1991. A thermodynamic study of unusually stable RNA and DNA hairpins. *Nucleic Acids Res.* 19:5901-5905.

50.   Jucker, F. M., and A. Pardi. 1995. GNRA tetraloops make a U-Turn. *RNA.* 1:219–222.

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

52.   Santalucia, J., R. Kierzek, and D. H. Turner. 1992. Context dependence of hydrogen-bond free-energy revealed by substitutions in an RNA hairpin. *Science.* 256:217-219.

53. Correll, C. C., and K. Swinger. 2003. Common and distinctive features of GNRA tetraloops based on a GUAA tetraloop structure at 1.4 angstrom resolution. *RNA-Publ. RNA Soc.* 9:355-363.

54. Correll, C. C., A. Munishkin, Y. L. Chan, Z. Ren, I. G. Wool, and T. A. Steitz. 1998. Crystal structure of the ribosomal RNA domain essential for binding elongation factors. *Proc. Natl. Acad. Sci. U. S. A.* 95:13436-13441.

55. Jucker, F. M., H. A. Heus, P. F. Yip, E. H. M. Moors, and A. Pardi. 1996. A network of heterogeneous hydrogen bonds in GNRA tetraloops. *J. Mol. Biol.* 264:968-980.

56. Jaeger, L., F. Michel, and E. Westhof. 1994. Involvement of a GNRA tetraloop in long-range tertiary interactions. *J. Mol. Biol.* 236:1271–1276.

57. Pley, H. W., K. M. Flaherty, and D. B. McKay. 1994. Model for an RNA tertiary interaction from the structure of an intermolecular complex between a GAAA tetraloop and an RNA helix. *Nature*. 372:111-113.

58. Pley, H. W., K. M. Flaherty, and D. B. McKay. 1994. 3-Dimensional structure of a hammerhead ribozyme. *Nature*. 372:68–74.

59. Nissen, P., J. A. Ippolito, N. Ban, P. B. Moore, and T. A. Steitz. 2001. RNA tertiary interactions in the large ribosomal subunit: The A-minor motif. *Proc. Natl. Acad. Sci. USA.* 98:4899–4903.

60. Doherty, E. A., R. T. Batey, B. Masquida, and J. A. Doudna. 2001. A universal mode of helix packing in RNA. *Nat. Struct. Biol.* 8:339–343.

61. Xin, Y. R., C. Laing, N. B. Leontis, and T. Schlick. 2008. Annotation of tertiary interactions in RNA structures reveals variations and correlations. 14:2465-2477.

62. Cate, J. H., A. R. Gooding, E. Podell, K. H. Zhou, B. L. Golden, C. E. Kundrot, *et al.* 1996. Crystal structure of a group I ribozyme domain: Principles of RNA packing. *Science*. 273:1678–1685.

63. Costa, M., and F. Michel. 1995. Frequent use of the same tertiary motif by self-folding RNAs. *Embo J.* 14:1276–1285.

64. Adams, P. L., M. R. Stahley, A. B. Kosek, J. M. Wang, and S. A. Strobel. 2004. Crystal structure of a self-splicing group I intron with both exons. *Nature*. 430:45-50.

65. Toor, N., K. S. Keating, S. D. Taylor, and A. M. Pyle. 2008. Crystal structure of a self-spliced group II intron. *Science*. 320:77-82.

66. Krasilnikov, A. S., X. J. Yang, T. Pan, and A. Mondragon. 2003. Crystal structure of the specificity domain of ribonuclease P. *Nature*. 421:760-764.

67. Costa, M., and F. Michel. 1997. Rules for RNA recognition of GNRA tetraloops deduced by in vitro selection: Comparison with in vivo evolution. 16:3289-3302.

68. Geary, C., S. Baudrey, and L. Jaeger. 2008. Comprehensive features of natural and in vitro selected GNRA tetraloop-binding receptors. *Nucleic Acids Res.* 36:1138–1152.

69. Collins, R. A. 1988. Evidence of natural-selection to maintain a functional domain outside of the core in a large subclass of group-I introns. *Nucleic Acids Res.* 16:2705-2715.

70. Murphy, F. L., and T. R. Cech. 1994. GAAA tetraloop and conserved bulge stabilize tertiary structure of a group-I intron domain. 236:49–63.

71.   Cate, J. H., A. R. Gooding, E. Podell, K. H. Zhou, B. L. Golden, A. A. Szewczak*, et al.* 1996. RNA tertiary structure mediation by adenosine platforms. *Science.* 273:1696-1699.

72.   Mondragon, P. A., A. S. Krasilnikov, and T. Pan. 2003. Structural studies of the specificity domain of ribonuclease P. *Biochemistry*. 42:11.

73.   Davis, J. H., M. Tonelli, L. G. Scott, L. Jaeger, J. R. Williamson, and S. E. Butcher. 2005. RNA helical packing in solution: NMR structure of a 30 kDa GAAA tetraloop-receptor complex. *J. Mol. Biol.* 351:371–382.

74.   Butcher, S. E., T. Dieckmann, and J. Feigon. 1997. Solution structure of a GAAA tetraloop receptor RNA. *Embo J.* 16:7490–7499.

75.   Qin, P. Z., J. Feigon, and W. L. Hubbell. 2005. Site-directed spin labeling studies reveal solution conformational changes in a GAAA tetraloop receptor upon Mg^{2+}-dependent docking of a GAAA tetraloop. 351:1-8.

76.   Qin, P. Z., S. E. Butcher, J. Feigon, and W. L. Hubbell. 2001. Quantitative analysis of the isolated GAAA tetraloop/receptor interaction in solution: A site-directed spin labeling study. *Biochemistry*. 40:6929–6936.

77.   Jaeger, L., and N. B. Leontis. 2000. Tecto-RNA: One-dimensional self-assembly through tertiary interactions. *Angew. Chem.-Int. Edit.* 39:2521–2524.

78.   Juneau, K., E. Podell, D. J. Harrington, and T. R. Cech. 2001. Structural basis of the enhanced stability of a mutant ribozyme domain and a detailed view of RNA-solvent interactions. *Structure*. 9:221-231.

79.   Adams, P. L., M. R. Stahley, M. L. Gill, A. B. Kosek, J. M. Wang, and S. A. Strobel. 2004. Crystal structure of a group I intron splicing intermediate. *RNA-Publ. RNA Soc.* 10:1867-1887.

80.   Basu, S., R. P. Rambo, J. Strauss-Soukup, J. H. Cate, A. R. Ferre-D'Amare, S. A. Strobel*, et al.* 1998. A specific monovalent metal ion integral to the AA platform of the RNA tetraloop receptor. *Nat. Struct. Biol.* 5:986–992.

81.   Cate, J. H., and J. A. Doudna. 1996. Metal-binding sites in the major groove of a large ribozyme domain. *Structure*. 4:1221–1229.

82.   Davis, J. H., T. R. Foster, M. Tonelli, and S. E. Butcher. 2007. Role of metal ions in the tetraloop-receptor complex as analyzed by NMR. *RNA-Publ. RNA Soc.* 13:76-86.

83.   Lambert, D., D. Leipply, R. Shiman, and D. E. Draper. 2009. The influence of monovalent cation size on the stability of RNA tertiary structures. 390:791-804.

84.   Misra, V. K., R. Shiman, and D. E. Draper. 2003. A thermodynamic framework for the magnesium-dependent folding of RNA. *Biopolymers*. 69:118–136.

85.   Misra, V. K., and D. E. Draper. 2001. A thermodynamic framework for Mg^{2+} binding to RNA. *Proc. Natl. Acad. Sci. U. S. A.* 98:12456–12461.

86.   Downey, C. D., R. L. Crisman, T. W. Randolph, and A. Pardi. 2007. Influence of hydrostatic pressure and cosolutes on RNA tertiary structure. *J. Am. Chem. Soc.* 129:9290-9291.

87.   Pace, U., and J. W. Szostak. 1991. Mutations in a semiconserved region of the *Tetrahymena* Intron. *FEBS Lett.* 280:171-174.

88.   Van der horst, G., A. Christian, and T. Inoue. 1991. Reconstitution of a group-I intron self-splicing reaction with an activator RNA. *Proc. Natl. Acad. Sci. U. S. A.* 88:184-188.

89.   Joyce, G. F., G. Vanderhorst, and T. Inoue. 1989. Catalytic activity is retained in the *Tetrahymena* group-I intron despite removal of the large extension of element P5. *Nucleic Acids Res.* 17:7879-7889.

90.   Flor, P. J., J. B. Flanegan, and T. R. Cech. 1989. A conserved base pair within helix-P4 of the *Tetrahymena* ribozyme helps to form the tertiary structure required for self-splicing. *Embo J.* 8:3391-3399.

91.   Murphy, F. L., and T. R. Cech. 1993. An independently folding domain of RNA tertiary structure within the *Tetrahymena* ribozyme. *Biochemistry*. 32:5291-5300.

92.   Sattin, B. D., W. Zhao, K. Travers, S. Chut, and D. Herschlag. 2008. Direct measurement of tertiary contact cooperativity in RNA folding. *J. Am. Chem. Soc.* 130:6085–6087.

93.   Jack, A., J. E. Ladner, D. Rhodes, R. S. Brown, and A. Klug. 1977. Crystallographic study of metal-binding to yeast phenylalanine transfer-RNA. *J. Mol. Biol.* 111:315-328.

94.   Holbrook, S. R., J. L. Sussman, R. W. Warrant, G. M. Church, and S. H. Kim. 1977. RNA-ligand interactions - (I)Magnesium binding-sites in yeast transfer-RNAphe. *Nucleic Acids Res.* 4:2811-2820.

95.   Ikawa, Y., T. Yoshimura, H. Hara, H. Shiraishi, and T. Inoue. 2002. Two conserved structural components, A-rich bulge and P4 XJ6/7 base-triples, in activating the group I ribozymes. *Genes to Cells.* 7:1205-1215.

96.   Zwanzig, R., A. Szabo, and B. Bagchi. 1992. Levinthals paradox. *Proc. Natl. Acad. Sci. U. S. A.* 89:20-22.

97.   Levinthal, C. 1968. Are there pathways for protein folding. *J. Chim. Phys.-Chim. Biol.* 65:44-&.

98.   Shcherbakova, I., S. Mitra, A. Laederach, and M. Brenowitz. 2008. Energy barriers, pathways, and dynamics during folding of large, multidomain RNAs. *Curr. Opin. Chem. Biol.* 12:655-666.

99.   Ditzler, M. A., D. Rueda, J. J. Mo, K. Hakansson, and N. G. Walter. 2008. A rugged free energy landscape separates multiple functional RNA folds throughout denaturation. *Nucleic Acids Res.* 36:7088-7099.

100.   Solomatin, S. V., M. Greenfeld, S. Chu, and D. Herschlag. 2010. Multiple native states reveal persistent ruggedness of an RNA folding landscape. *Nature*. 463:681-686.

101.   Silverman, S. K., and T. R. Cech. 2001. An early transition state for folding of the P4-P6 RNA domain. *RNA-Publ. RNA Soc.* 7:161–166.

102.   Maglott, E. J., J. T. Goodwin, and G. D. Glick. 1999. Probing the structure of an RNA tertiary unfolding transition state. *J. Am. Chem. Soc.* 121:7461-7462.

103.   Bartley, L. E., X. W. Zhuang, R. Das, S. Chu, and D. Herschlag. 2003. Exploration of the transition state for tertiary structure formation between an RNA helix and a large structured RNA. *J. Mol. Biol.* 328:1011–1026.

104.   Li, P. T. X., J. Vieregg, and I. Tinoco. 2008. How RNA unfolds and refolds. *Annu. Rev. Biochem.* 77:77–100.

105.   Bokinsky, G., and X. W. Zhuang. 2005. Single-molecule RNA folding. *Accounts Chem. Res.* 38:566–573.

106. Hodak, J. H., C. D. Downey, J. L. Fiore, A. Pardi, and D. J. Nesbitt. 2005. Docking kinetics and equilibrium of a GAAA tetraloop-receptor motif probed by single-molecule FRET. *Proc. Natl. Acad. Sci. USA.* 102:10505–10510.

107. Xie, Z., N. Srividya, T. R. Sosnick, T. Pan, and N. F. Scherer. 2004. Single-molecule studies highlight conformational heterogeneity in the early folding steps of a large ribozyme. *Proc. Natl. Acad. Sci. USA.* 101:534–539.

108. Hohng, S., T. J. Wilson, E. Tan, R. M. Clegg, D. M. J. Lilley, and T. J. Ha. 2004. Conformational flexibility of four-way junctions in RNA. *J. Mol. Biol.* 336:69–79.

109. Stryer, L., and R. P. Haugland. 1967. Energy Transfer - A Spectroscopic Ruler. *Proc. Natl. Acad. Sci. U. S. A.* 58:719-&.

110. Karunatilaka, K. S., and D. Rueda. 2009. Single-molecule fluorescence studies of RNA: A decade's progress. *Chem. Phys. Lett.* 476:1-10.

111. Förster, T. 1948. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys.-Berlin.* 2:55-75.

112. Dale, R. E., J. Eisinger, and W. E. Blumberg. 1979. Orientational freedom of molecular probes - Orientation factor in Intra-molecular energy-transfer. *Biophys. J.* 26:161-193.

113. Iqbal, A., S. Arslan, B. Okumus, T. J. Wilson, G. Giraud, D. G. Norman*, et al.* 2008. Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nucleic acids. *Proc. Natl. Acad. Sci. U. S. A.* 105:11176-11181.

114. Nie, S. M., and R. N. Zare. 1997. Optical detection of single molecules. *Annu. Rev. Biophys. Biomolec. Struct.* 26:567-596.

115. Ha, T. 2001. Single-molecule fluorescence resonance energy transfer. *Methods.* 25:78–86.

116. Sanborn, M. E., B. K. Connolly, K. Gurunathan, and M. Levitus. 2007. Fluorescence properties and photophysics of the sulfoindocyanine Cy3 linked covalently to DNA. *J. Phys. Chem. B.* 111:11064-11074.

117. Fiore, J. L., B. Kraemer, F. Koberling, R. Erdmann, and D. J. Nesbitt. 2009. Enthalpy-driven RNA folding: Single-molecule thermodynamics of tetraloop–receptor tertiary interaction. *Biochemistry.* 48:2550-2558.

118. Deniz, A. A., M. Dahan, J. R. Grunwell, T. J. Ha, A. E. Faulhaber, D. S. Chemla*, et al.* 1999. Single-pair fluorescence resonance energy transfer on freely diffusing molecules: Observation of Forster distance dependence and subpopulations. *Proc. Natl. Acad. Sci. USA.* 96:3670–3675.

119. Zhuang, X. W., L. E. Bartley, H. P. Babcock, R. Russell, T. J. Ha, D. Herschlag*, et al.* 2000. A single-molecule study of RNA catalysis and folding. *Science.* 288:2048–2051.

120. Serra, M. J., D. H. Turner, L. J. Michael, and G. K. Ackers. 1995. Predicting thermodynamic properties of RNA. *Methods Enzymol.* 259:242–261.

121. Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31:3406-3415.

122. Qin, H., T. R. Sosnick, and T. Pan. 2001. Modular construction of a tertiary RNA structure: The specificity domain of the Bacillus subtilis RNase P RNA. 40:11202-11210.

123.   Leontis, N. B., A. Lescoute, and E. Westhof. 2006. The building blocks and motifs of RNA architecture. *Curr. Opin. Struct. Biol.* 16:279–287.

124.   Downey, C. D., J. L. Fiore, C. D. Stoddard, J. H. Hodak, D. J. Nesbitt, and A. Pardi. 2006. Metal ion dependence, thermodynamics, and kinetics for intramolecular docking of a GAAA tetraloop and receptor connected by a flexible linker. *Biochemistry*. 45:3664–3673.

125.   Vander Meulen, K. A., J. H. Davis, T. R. Foster, T. Record, and S. E. Butcher. 2008. Thermodynamics and folding pathway of tetraloop receptor-mediated RNA helical packing. *J. Mol. Biol.* 384:702-717.

126.   Young, B. T., and S. K. Silverman. 2002. The GAAA tetraloop-receptor interaction contributes differentially to folding thermodynamics and kinetics for the P4-P6 RNA domain. *Biochemistry*. 41:12271–12276.

127.   Abramovitz, D. L., and A. M. Pyle. 1997. Remarkable morphological variability of a common RNA folding motif: The GNRA tetraloop-receptor interaction. *J. Mol. Biol.* 266:493–506.

128.   Fiore, J. L., J. H. Hodak, O. Piestert, C. D. Downey, and D. J. Nesbitt. 2008. Monovalent and divalent promoted GAAA tetraloop-receptor tertiary interactions from freely diffusing single-molecule studies. *Biophys. J.* 95:3892-3905.

129.   Rasnik, I., S. A. McKinney, and T. Ha. 2006. Nonblinking and longlasting single-molecule fluorescence imaging. *Nat. Methods*. 3:891–893.

130.   Aitken, C. E., R. A. Marshall, and J. D. Puglisi. 2008. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys. J.* 94:1826-1835.

131.   Schwille, P., F. J. MeyerAlmes, and R. Rigler. 1997. Dual-color fluorescence cross-correlation spectroscopy for multicomponent diffusional analysis in solution. *Biophys. J.* 72:1878-1886.

132.   Schwille, P., and E. Haustein. 2002. Fluorescence Correlation Spectroscopy. *Biophysics Textbook Online*.

133.   Rothwell, P. J., S. Berger, O. Kensch, S. Felekyan, M. Antonik, B. M. Wohrl*, et al.* 2003. Multiparameter single-molecule fluorescence spectroscopy reveals heterogeneity of HIV-1 reverse transcriptase: primer/template complexes. *Proc. Natl. Acad. Sci. U. S. A.* 100:1655–1660.

134.   Grunwell, J. R., J. L. Glass, T. D. Lacoste, A. A. Deniz, D. S. Chemla, and P. G. Schultz. 2001. Monitoring the conformational fluctuations of DNA hairpins using single-pair fluorescence resonance energy transfer. *J. Am. Chem. Soc.* 123:4295–4303.

135.   Iqbal, A., L. Wang, K. C. Thompson, D. M. J. Lilley, and D. G. Norman. 2008. The structure of cyanine 5 terminally attached to double-stranded DNA: Implications for FRET studies. *Biochemistry*. 47:7857-7862.

136.   Anderson, C. F., and M. T. Record. 1995. Salt nucleic-acid interactions. *Annu. Rev. Phys. Chem.* 46:657–700.

137.   Correll, C. C., B. Freeborn, P. B. Moore, and T. A. Steitz. 1997. Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. *Cell*. 91:705–712.

138.   Draper, D. E., and V. K. Misra. 1998. RNA shows its metal. *Nat. Struct. Biol.* 5:927–930.

139.   Misra, V. K., and D. E. Draper. 1998. On the role of magnesium ions in RNA stability. *Biopolymers*. 48:113–135.

140. Treiber, D. K., M. S. Rook, P. P. Zarrinkar, and J. R. Williamson. 1998. Kinetic intermediates trapped by native interactions in RNA folding. *Science.* 279:1943–1946.

141. Treiber, D. K., and J. R. Williamson. 1999. Exposing the kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.* 9:339–345.

142. Treiber, D. K., and J. R. Williamson. 2001. Beyond kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.* 11:309–314.

143. Thirumalai, D., N. Lee, S. A. Woodson, and D. K. Klimov. 2001. Early events in RNA folding. *Annu. Rev. Phys. Chem.* 52:751–762.

144. Thirumalai, D., and S. A. Woodson. 1996. Kinetics of folding of proteins and RNA. *Accounts Chem. Res.* 29:433–439.

145. Sosnick, T. R., and T. Pan. 2003. RNA folding: Models and perspectives. *Curr. Opin. Struct. Biol.* 13:309–316.

146. Tan, E., T. J. Wilson, M. K. Nahas, R. M. Clegg, D. M. J. Lilley, and T. Ha. 2003. A four-way junction accelerates hairpin ribozyme folding via a discrete intermediate. *Proc. Natl. Acad. Sci. USA.* 100:9308–9313.

147. Ha, T., X. W. Zhuang, H. D. Kim, J. W. Orr, J. R. Williamson, and S. Chu. 1999. Ligand-induced conformational changes observed in single RNA molecules. *Proc. Natl. Acad. Sci. USA.* 96:9077–9082.

148. Kim, H. D., G. U. Nienhaus, T. Ha, J. W. Orr, J. R. Williamson, and S. Chu. 2002. Mg^{2+}-dependent conformational change of RNA studied by fluorescence correlation and FRET on immobilized single molecules. *Proc. Natl. Acad. Sci. U. S. A.* 99:4284–4289.

149. Pljevaljcic, G., D. P. Millar, and A. A. Deniz. 2004. Freely diffusing single hairpin ribozymes provide insights into the role of secondary structure and partially folded states in RNA folding. *Biophys. J.* 87:457–467.

150. Russell, R., X. W. Zhuang, H. P. Babcock, I. S. Millett, S. Doniach, S. Chu, *et al.* 2002. Exploring the folding landscape of a structured RNA. *Proc. Natl. Acad. Sci. U. S. A.* 99:155–160.

151. Xie, S. N. 2001. Single-molecule approach to enzymology. *Single Mol.* 2:229–236.

152. Xie, X. S., and H. P. Lu. 1999. Single-molecule enzymology. *J. Biol. Chem.* 274:15967–15970.

153. Yasuda, R., T. Masaike, K. Adachi, H. Noji, H. Itoh, and K. Kinosita. 2003. The ATP-waiting conformation of rotating F-1-ATPase revealed by single-pair fluorescence resonance energy transfer. *Proc. Natl. Acad. Sci. U. S. A.* 100:9314–9318.

154. Talaga, D. S., W. L. Lau, H. Roder, J. Y. Tang, Y. W. Jia, W. F. DeGrado, *et al.* 2000. Dynamics and folding of single two-stranded coiled-coil peptides studied by fluorescent energy transfer confocal microscopy. *Proc. Natl. Acad. Sci. U. S. A.* 97:13021–13026.

155. Okumus, B., T. J. Wilson, D. M. J. Lilley, and T. Ha. 2004. Vesicle encapsulation studies reveal that single molecule ribozyme heterogeneities are intrinsic. *Biophys. J.* 87:2798–2806.

156. Boukobza, E., A. Sonnenfeld, and G. Haran. 2001. Immobilization in surface-tethered lipid vesicles as a new tool for single biomolecule spectroscopy. *J. Phys. Chem. B.* 105:12165–12170.

157. Rhoades, E., E. Gussakovsky, and G. Haran. 2003. Watching proteins fold one molecule at a time. *Proc. Natl. Acad. Sci. U. S. A.* 100:3197–3202.

158. Fries, J. R., L. Brand, C. Eggeling, M. Kollner, and C. A. M. Seidel. 1998. Quantitative identification of different single molecules by selective time-resolved confocal fluorescence spectroscopy. *J. Phys. Chem. A*. 102:6601–6613.

159. Eigen, M., and R. Rigler. 1994. Sorting single molecules: Application to diagnostics and evolutionary biotechnology. *Proc. Natl. Acad. Sci. U. S. A.* 91:5740–5747.

160. Dahan, M., A. A. Deniz, T. J. Ha, D. S. Chemla, P. G. Schultz, and S. Weiss. 1999. Ratiometric measurement and identification of single diffusing molecules. *Chem. Phys.* 247:85–106.

161. Deniz, A. A., T. A. Laurence, M. Dahan, D. S. Chemla, P. G. Schultz, and S. Weiss. 2001. Ratiometric single-molecule studies of freely diffusing biomolecules. *Annu. Rev. Phys. Chem.* 52:233–253.

162. Jia, Y. W., D. S. Talaga, W. L. Lau, H. S. M. Lu, W. F. DeGrado, and R. M. Hochstrasser. 1999. Folding dynamics of single GCN4 peptides by fluorescence resonant energy transfer confocal microscopy. *Chem. Phys.* 247:69–83.

163. Gordon, G. W., G. Berry, X. H. Liang, B. Levine, and B. Herman. 1998. Quantitative fluorescence resonance energy transfer measurements using fluorescence microscopy. *Biophys. J.* 74:2702–2713.

164. Lee, N. K., A. N. Kapanidis, Y. Wang, X. Michalet, J. Mukhopadhyay, R. H. Ebright, *et al.* 2005. Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophys. J.* 88:2939–2953.

165. Nir, E., X. Michalet, K. M. Hamadani, T. A. Laurence, D. Neuhauser, Y. Kovchegov, *et al.* 2006. Shot-noise limited single-molecule FRET histograms: Comparison between theory and experiments. *J. Phys. Chem. B*. 110:22103–22124.

166. Antonik, M., S. Felekyan, A. Gaiduk, and C. A. M. Seidel. 2006. Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis. *J. Phys. Chem. B*. 110:6970–6978.

167. Heilman-Miller, S. L., D. Thirumalai, and S. A. Woodson. 2001. Role of counterion condensation in folding of the Tetrahymena ribozyme. I. Equilibrium stabilization by cations. *J. Mol. Biol.* 306:1157–1166.

168. Deras, M. L., M. Brenowitz, C. Y. Ralston, M. R. Chance, and S. A. Woodson. 2000. Folding mechanism of the Tetrahymena ribozyme P4-P6 domain. *Biochemistry*. 39:10975–10985.

169. Stigter, D. 1995. Evaluation of the counterion condensation theory of polyelectrolytes. *Biophys. J.* 69:380–388.

170. Sharp, K. A., R. A. Friedman, V. Misra, J. Hecht, and B. Honig. 1995. Salt effects on polyelectrolyte-ligand binding: Comparison of Poisson-Boltzmann, and limiting/law counterion binding models. *Biopolymers*. 36:245–262.

171. Honig, B., and A. Nicholls. 1995. Classical electrostatics in biology and chemistry. *Science*. 268:1144–1149.

172. Chin, K., K. A. Sharp, B. Honig, and A. M. Pyle. 1999. Calculating the electrostatic properties of RNA provides new insights into molecular interactions and function. *Nat. Struct. Biol.* 6:1055–1061.

173. Misra, V. K., and D. E. Draper. 2000. Mg^{2+} binding to tRNA revisited: The nonlinear Poisson-Boltzmann model. *J. Mol. Biol.* 299:813–825.

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

175. Rocchia, W., E. Alexov, and B. Honig. 2001. Extending the applicability of the nonlinear Poisson-Boltzmann equation: Multiple dielectric constants and multivalent ions. *J. Phys. Chem. B*. 105:6507–6514.

176. Bai, Y., M. Greenfeld, K. J. Travers, V. B. Chu, J. Lipfert, S. Doniach*, et al.* 2007. Quantitative and comprehensive decomposition of the ion atmosphere around nucleic acids. *J. Am. Chem. Soc.* 129:14981–14988.

177. Stigter, D. 1975. Charged colloidal cylinder with a Gouy double-layer. *J. Colloid Interface Sci.* 53:296–306.

178. Misra, V. K., and D. E. Draper. 1999. The interpretation of Mg^{2+} binding isotherms for nucleic acids using Poisson-Boltzmann theory. *J. Mol. Biol.* 294:1135–1147.

179. Lee, J. C., R. R. Gutell, and R. Russell. 2006. The UAA/GAN internal loop motif: A new RNA structural element that forms a cross-strand AAA stack and long-range tertiary interactions. *J. Mol. Biol.* 360:978–988.

180. Nixon, P. L., and D. P. Giedroc. 2000. Energetics of a strongly pH dependent RNA tertiary structure in a frameshifting pseudoknot. *J. Mol. Biol.* 296:659–671.

181. Hammann, C., A. Cooper, and D. M. J. Lilley. 2001. Thermodynamics of ion-induced RNA folding in the hammerhead ribozyme: An isothermal titration calorimetric study. *Biochemistry*. 40:1423–1429.

182. Mikulecky, P. J., J. C. Takach, and A. L. Feig. 2004. Entropy-driven folding of an RNA helical junction: An isothermal titration calorimetric analysis of the hammerhead ribozyme. *Biochemistry*. 43:5870–5881.

183. Feig, A. L. 2007. Applications of isothermal titration calorimetry in RNA biochemistry and biophysics. *Biopolymers*. 87:293–301.

184. Reymond, C., M. Bisaillon, and J.-P. Perreault. 2009. Monitoring of an RNA multistep folding pathway by isothermal titration calorimetry. *Biophys. J.* 96:132.

185. Klostermeier, D., and D. P. Millar. 2000. Helical junctions as determinants for RNA folding: Origin of tertiary structure stability of the hairpin ribozyme. *Biochemistry*. 39:12970–12978.

186. Gell, C., T. Sabir, J. Westwood, A. Rashid, D. A. M. Smith, S.-A. Harris*, et al.* 2008. Single-molecule fluorescence resonance energy transfer assays reveal heterogeneous folding ensembles in a simple RNA stem-loop. *J. Mol. Biol.* 384:264–278.

187. Good, N. E., G. D. Winget, W. Winter, T. N. Connolly, S. Izawa, and R. M. M. Singh. 1966. Hydrogen ion buffers for biological research. *Biochemistry*. 5:467-477.

188. Hernández, B., V. Baumruk, N. Leulliot, C. Gouyette, T. Huynh-Dinh, and M. Ghomi. 2003. Thermodynamic and structural features of ultrastable DNA and RNA hairpins. 651–653:67–74.

189. Markham, N. R., and M. Zuker. 2005. DINAMelt web server for nucleic acid melting prediction. 33:W577–581.

190.   Silverman, S. K., and T. R. Cech. 1999. Energetics and cooperativity of tertiary hydrogen bonds in RNA structure. *Biochemistry*. 38:8691–8702.

191.   Mikulecky, P. J., and A. L. Feig. 2006. Heat capacity changes associated with nucleic acid folding. *Biopolymers*. 82:38–58.

192.   Li, Y., P. C. Bevilacqua, D. Mathews, and D. H. Turner. 1995. Thermodynamic and activation parameters for binding of a pyrene-labeled substrate by the *Tetrahymena* ribozyme: Docking is not diffusion-controlled and is driven by a favorable entropy. *Biochemistry*. 34:14394–14399.

193.   Narlikar, G. J., and D. Herschlag. 1996. Isolation of a local tertiary folding transition in the context of a globally folded RNA. *Nat. Struct. Biol.* 3:701–710.

194.   Stein, A., and D. M. Crothers. 1976. Conformational changes of transfer-RNA: Role of magnesium(II). *Biochemistry*. 15:160–168.

195.   Lu, M., and D. E. Draper. 1994. Bases defining an ammonium and magnesium ion-dependent tertiary structure within the large subunit ribosomal-RNA. *J. Mol. Biol.* 244:572–585.

196.   Stoddard, C. D., S. D. Gilbert, and R. T. Batey. 2008. Ligand-dependent folding of the three-way junction in the purine riboswitch. *RNA-Publ. RNA Soc.* 14:675–684.

197.   Rupert, P. B., and A. R. Ferre-D'Amare. 2001. Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature*. 410:780–786.

198.   Szewczak, A. A., L. Ortoleva-Donnelly, S. P. Ryder, E. Moncoeur, and S. A. Strobel. 1998. A minor groove RNA triple helix within the catalytic core of a group I intron. *Nat. Struct. Biol.* 5:1037–1042.

199.   Scott, W. G., J. T. Finch, and A. Klug. 1995. The crystal structure of an All-RNA hammerhead ribozyme: A proposed mechanism for RNA catalytic cleavage. *Cell*. 81:991–1002.

200.   Pljevaljcic, G., D. Klostermeier, and D. P. Millar. 2005. The tertiary structure of the hairpin ribozyme is formed through a slow conformational search. *Biochemistry*. 44:4870-4876.

201.   Seol, Y., G. M. Skinner, and K. Visscher. 2004. Elastic properties of a single-stranded charged homopolymeric ribonucleotide. *Phys. Rev. Lett.* 93.

202.   Seol, Y., G. M. Skinner, K. Visscher, A. Buhot, and A. Halperin. 2007. Stretching of homopolymeric RNA reveals single-stranded helices and base-stacking. *Phys. Rev. Lett.* 98.

203.   Dewey, T. G., and D. H. Turner. 1979. Laser Temperature-Jump Study of Stacking In Adenylic Acid Polymers. *Biochemistry*. 18:5757-5762.

204.   McKinney, S. A., C. Joo, and T. Ha. 2006. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys. J.* 91:1941-1951.

205.   SantaLucia, J. 1998. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. U. S. A.* 95:1460-1465.

206.   Zhou, H. X. 2010. Rate theories for biologists. *Q. Rev. Biophys.* 43:219-293.

207.   Hanggi, P., P. Talkner, and M. Borkovec. 1990. Reaction-rate theory: 50 years after Kramers. *Rev. Mod. Phys.* 62:251-341.

208.   Frauenfelder, H., and P. G. Wolynes. 1985. Rate theories and puzzles of Heme protein kinetics. *Science*. 229:337-345.

209.   Zhou, H. X. 2003. Theory for the rate of contact formation in a polymer chain with local conformational transitions. *J. Chem. Phys.* 118:2010-2015.

210.   Porschke, D., Uhlenbec.Oc, and F. H. Martin. 1973. Thermodynamics and kinetics of helix-coil transition of pligomers containing GC base pairs. *Biopolymers*. 12:1313-1335.

211.   Thirumalai, D., and C. Hyeon. 2005. RNA and protein folding: Common themes and variations. *Biochemistry*. 44:4957–4970.

212.   Hagen, S. J., J. Hofrichter, A. Szabo, and W. A. Eaton. 1996. Diffusion-limited contact formation in unfolded cytochrome c: Estimating the maximum rate of protein folding. *Proc. Natl. Acad. Sci. U. S. A.* 93:11615-11617.

213.   Heilman-Miller, S. L., J. Pan, D. Thirumalai, and S. A. Woodson. 2001. Role of counterion condensation in folding of the Tetrahymena ribozyme II. Counterion-dependence of folding kinetics. *J. Mol. Biol.* 309:57–68.

214.   Record, M. T., C. F. Anderson, and T. M. Lohman. 1978. Thermodynamic analysis of ion effects on binding and conformational equilibria of proteins and nucleic acids: Roles of ion association or release, screening, and ion effects on water activity. *Q. Rev. Biophys.* 11:103-178.

215.   Leipply, D., and D. E. Draper. 2010. Dependence of RNA tertiary structural stability on Mg^{2+} concentration: Interpretation of the Hill equation and coefficient. *Biochemistry*. 49:1843-1853.

216.   Meulen, K. A. V., J. H. Davis, T. R. Foster, M. T. Record, and S. E. Butcher. 2008. Thermodynamics and Folding Pathway of Tetraloop Receptor-Mediated RNA Helical Packing. 384:702-717.

217.   Bai, Y., V. B. Chu, J. Lipfert, V. S. Pande, D. Herschlag, and S. Doniach. 2008. Critical assessment of nucleic acid electrostatics via experimental and computational investigation of an unfolded state ensemble. *J. Am. Chem. Soc.* 130:12334-12341.

218.   Bailor, M. H., X. Y. Sun, and H. M. Al-Hashimi. 2010. Topology links RNA secondary structure with global conformation, dynamics, and adaptation. *Science*. 327:202-206.

219.   Chu, V. B., J. Lipfert, Y. Bai, V. S. Pande, S. Doniach, and D. Herschlag. 2009. Do conformational biases of simple helical junctions influence RNA folding stability and specificity? *RNA-Publ. RNA Soc.* 15:2195-2205.

220.   Takamoto, K., R. Das, Q. He, S. Doniach, M. Brenowitz, D. Herschlag, *et al.* 2004. Principles of RNA compaction: Insights from the equilibrium folding pathway of the P4-P6 RNA domain in monovalent cations. *J. Mol. Biol.* 343:1195–1206.

221.   Kramers, H. A. 1940. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*. 7:284-304.

222.   Szabo, A., K. Schulten, and Z. Schulten. 1980. 1st passage time approach to diffusion controlled reactions. *J. Chem. Phys.* 72:4350-4357.

223.   Loman, A., I. Gregor, C. Stutz, M. Mund, and J. Enderlein. Measuring rotational diffusion of macromolecules by fluorescence correlation spectroscopy. *Photochem. Photobiol. Sci.* 9:627-636.

224.   Xi, X. M., Y. Sun, C. B. Karim, V. M. Grigoryants, and C. P. Scholes. 2008. HIV-1 nucleocapsid protein NCp7 and its RNA stem loop 3 partner: Rotational dynamics of spin-labeled RNA stem loop 3. *Biochemistry*. 47:10099-10110.

225.   Klein, U. K. A., and H. P. Haar. 1978. Picosecond time-dependent rotational diffusion of rhodamine 6G in micellar solution. *Chem. Phys. Lett.* 58:531-536.

226.   Froschauer, E. M., M. Kolisek, F. Dieterich, M. Schweigel, and R. J. Schweyen. 2004. Fluorescence measurements of free [Mg^{2+}] by use of mag-fura 2 in Salmonella enterica. *FEMS Microbiol. Lett.* 237:49-55.

227.   Cayley, S., B. A. Lewis, H. J. Guttman, and M. T. Record. 1991. Characterization of the cytoplasm of Escherichia-Coli-K-12 as a function of external osmolarity: Implications for protein DNA interactions in vivo. *J. Mol. Biol.* 222:281-300.

228.   London, R. E. 1991. Methods for measurement of intracellular magnesium: NMR and fluorescence. *Annu. Rev. Physiol.* 53:241-258.

229.   Igarashi, K., and K. Kashiwagi. 2000. Polyamines: Mysterious modulators of cellular functions. *Biochem. Biophys. Res. Commun.* 271:559-564.

230.   Tabor, C. W., and H. Tabor. 1984. Polyamines. *Annu. Rev. Biochem.* 53:749-790.

231.   Draper, D. E. 2008. RNA folding: Thermodynamic and molecular descriptions of the roles of ions. *Biophys. J.* 95:5489-5495.

232.   Grosshans, C. A., and T. R. Cech. 1989. Metal-ion requirements for sequence-specific endoribonuclease activity of the *Tetrahymena* ribozyme. *Biochemistry*. 28:6888-6894.

233.   McConnell, T. S., D. Herschlag, and T. R. Cech. 1997. Effects of divalent metal ions on individual steps of the Tetrahymena ribozyme reaction. *Biochemistry*. 36:8293-8303.

234.   Chu, V. B., Y. Bai, J. Lipfert, D. Herschlag, and S. Doniach. 2008. A repulsive field: advances in the electrostatics of the ion atmosphere. *Curr. Opin. Chem. Biol.* 12:619-625.

235.   Koculi, E., C. Hyeon, D. Thirumalai, and S. A. Woodson. 2007. Charge density of divalent metal cations determines RNA stability. *J. Am. Chem. Soc.* 129:2676–2682.

236.   Chu, V. B., and D. Herschlag. 2008. Unwinding RNA's secrets: advances in the biology, physics, and modeling of complex RNAs. *Curr. Opin. Struct. Biol.* 18:305-314.

237.   Chu, V. B., Y. Bai, J. Lipfert, D. Herschlag, and S. Doniach. 2007. Evaluation of ion binding to DNA duplexes using a size-modified Poisson-Boltzmann theory. *Biophys. J.* 93:3202-3209.

238.   Koculi, E., N. K. Lee, D. Thirumalai, and S. A. Woodson. 2004. Folding of the *Tetrahymena* ribozyme by polyamines: Importance of counterion valence and size. *J. Mol. Biol.* 341:27–36.

239.   Soto, A. M., V. Misra, and D. E. Draper. 2007. Tertiary structure of an RNA pseudoknot is stabilized by "diffuse" Mg^{2+} ions. *Biochemistry*. 46:2973-2983.

240.   Smith, G. J., K. T. Lee, X. H. Qu, Z. Xie, J. Pesic, T. R. Sosnick, *et al.* 2008. A large collapsed-state RNA can exhibit simple exponential single-molecule dynamics. *J. Mol. Biol.* 378:943–953.

241.   Steiner, M., K. S. Karunatilaka, R. K. O. Sigel, and D. Rueda. 2008. Single-molecule studies of group II intron ribozymes. *Proc. Natl. Acad. Sci. U. S. A.* 105:13853-13858.

242. Massire, C., L. Jaeger, and E. Westhof. 1998. Derivation of the three-dimensional architecture of bacterial ribonuclease P RNAs from comparative sequence analysis. *J. Mol. Biol.* 279:773-793.

243. Tanner, M. A., and T. R. Cech. 1995. An important RNA tertiary interaction of group-I and group-II introns is implicated in Gram-Positive RNase-P RNAs. *RNA-Publ. RNA Soc.* 1:349-350.

244. Kuno, M., D. P. Fromm, H. F. Hamann, A. Gallagher, and D. J. Nesbitt. 2001. "On"/"off" fluorescence intermittency of single semiconductor quantum dots. *J. Chem. Phys.* 115:1028-1040.

245. Herschlag, D. 2001. Channels in the RNA folding landscape. *Biochemistry*. 40:8635–8635.

246. Young, K. J., F. Gill, and J. A. Grasby. 1997. Metal ions play a passive role in the hairpin ribozyme catalysed reaction. *Nucleic Acids Res.* 25:3760-3766.

247. Ohtaki, H., and T. Radnai. 1993. Structure and dynamics of hydrated ions. *Chem. Rev.* 93:1157-1204.

248. Sakuma, H., and K. Kawamura. Structure and dynamics of water on Li^+-, Na^+-, K^+-, Cs^+-, H_3O^+-exchanged muscovite surfaces: A molecular dynamics study. *Geochim. Cosmochim. Acta*. 75:63-81.

249. Meek, D. W., and J. A. Ibers. 1970. Crystal structure of hexaamminecobalt(III) tetrachlorozincate(II) chloride, [Co(NH3)6][ZnCl4] Cl. 9:465.

250. Manning, G. S. 1978. Molecular theory of polyelectrolyte solutions with applications to electrostatic properties of polynucleotides. *Q. Rev. Biophys.* 11:179-246.

251. Silverman, S. K., M. L. Deras, S. A. Woodson, S. A. Scaringe, and T. R. Cech. 2000. Multiple folding pathways for the P4-P6 RNA domain. *Biochemistry*. 39:12465–12475.

252. Das, R., K. J. Travers, Y. Bai, and D. Herschlag. 2005. Determining the Mg^{2+} stoichiometry for folding an RNA metal ion core. *J. Am. Chem. Soc.* 127:8272–8273.

253. Tan, Z.-J., and S.-J. Chen. 2011. Salt contribution to RNA tertiary structure folding stability. 101:176.

254. Stigter, D., and K. A. Dill. 1996. Binding of ionic ligands to polyelectrolytes. 71:2064.

255. Record, M. T., W. T. Zhang, and C. F. Anderson. 1998. Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: A practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects of salts. *In* Advances In Protein Chemistry, Vol 51. 281-353.

256. Grilley, D., A. M. Soto, and D. E. Draper. 2006. Mg^{2+}-RNA interaction free energies and their relationship to the folding of RNA tertiary structures. *Proc. Natl. Acad. Sci. U. S. A.* 103:14003-14008.

257. Venditti, V., L. Clos, N. Niccolai, and S. E. Butcher. 2009. Minimum-energy path for a U6 RNA conformational change involving protonation, base-pair rearrangement and base flipping. *J. Mol. Biol.* 391:894-905.

258. Whitford, P. C., A. Schug, J. Saunders, S. P. Hennelly, J. N. Onuchic, and K. Y. Sanbonmatsu. 2009. Nonlocal helix formation is key to understanding S-Adenosylmethionine-1 riboswitch function. *Biophys. J.* 96:L7-L9.

Appendix

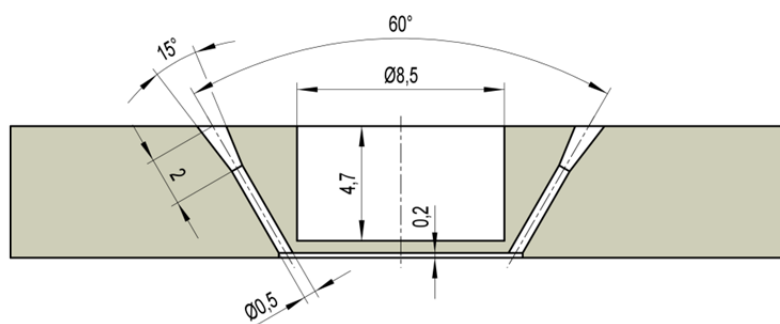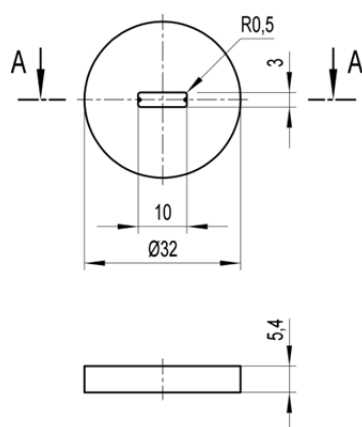## A.   Microfluidic Flow-Cell Sample Holders

Figure A.1   Technical drawing of flow cell holder used for observing single RNA molecules in a small chamber volume (Section 2.2.2). Round versions of this sample holder fit into the temperature stage.

Mountable Flow Cell

Material: Polychlortrifluorethylen (PCTFE)
Brandname: Kel-F, Neoflon

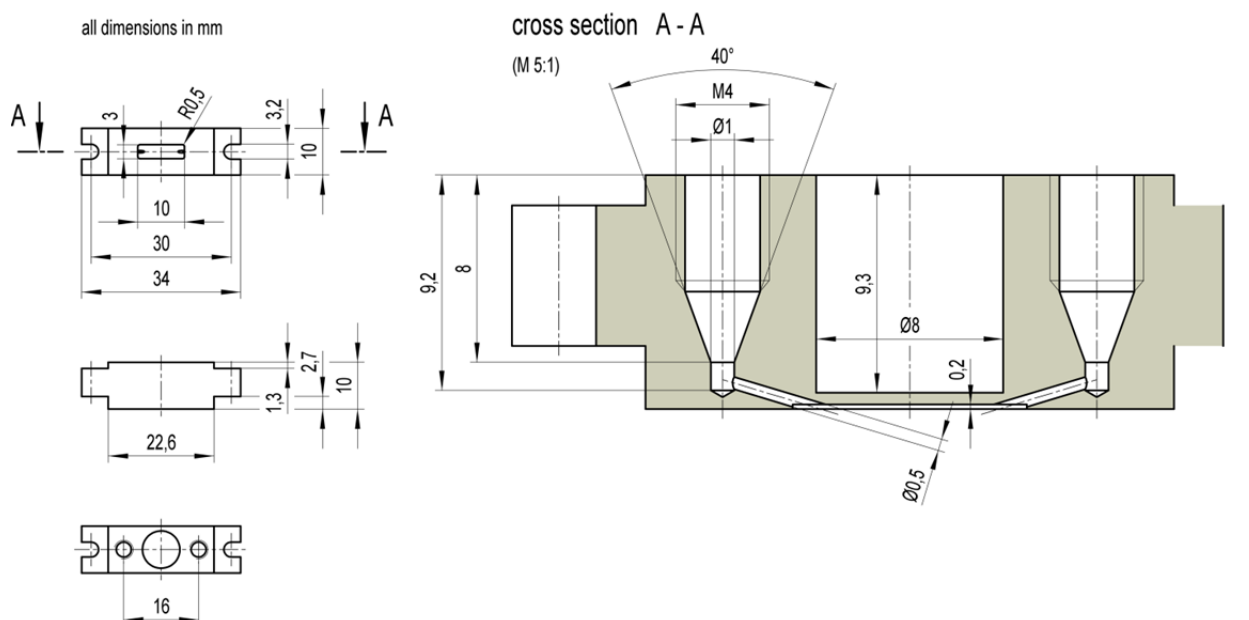all dimensions in mm

cross section   A - A

(M 5:1)

Figure A.2   Technical drawing of the mountable sample holder used for observation of the same molecules under different solution conditions (Section 2.2.2). The inlets are designed such that tubing can be inserted. The experimenter can insert the pipette tip into the tubing and flush in a new solution without disrupting the position of the cell on the stage.

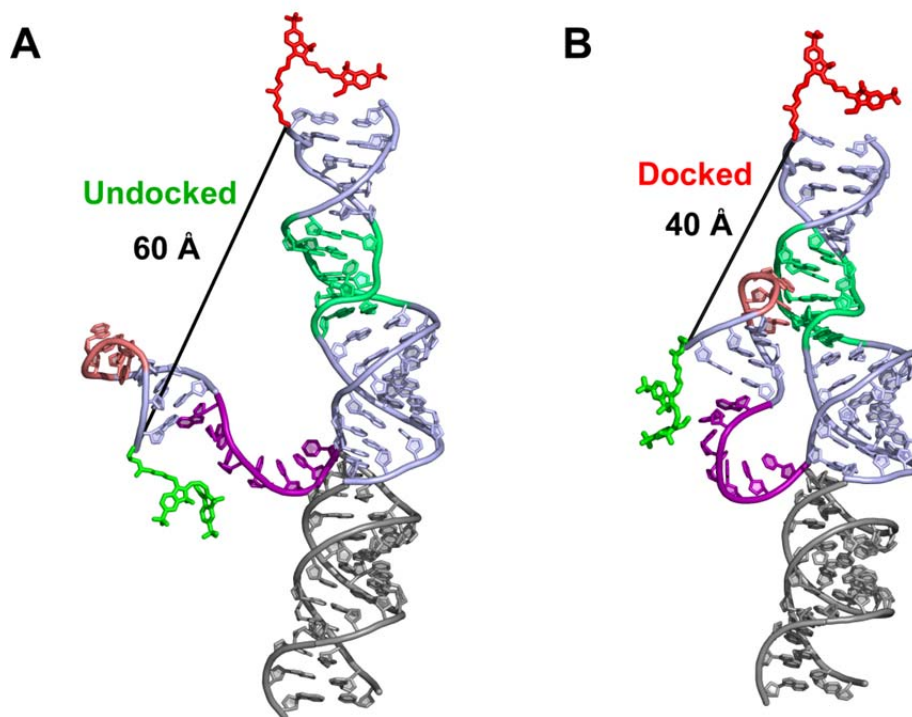## B.   Structural Model of the GAAA Tetraloop–Receptor Construct

Figure B.1   Model of the undocked (*A*) and docked (*B*) tetraloop—receptor construct. The model is built in *Pymol* by aligning the phosphates of various helical fragments that are the same number of nucleotides as the helices in the RNA construct (Figure 1.12 *A*). The coordinates for the docked tetraloop (*pink*), receptor (*green*), and helices (*light blue*) are from the NMR structure of a tetraloop–receptor complex (PDB 2ADT). The other helical regions (*purple* and *gray*) are from typical RNA helices (1QC2). The undocked structure (*A*) is generated by dragging the tetraloop/linker away from the receptor. Cy3 and Cy5 NHS esters structures are shown superimposed in *light green* and *red*, respectively. The RNA structures are *not* energy minimized. *Black lines* indicate the interphosphate distance between the terminal nucleotides to which the Cy3 and Cy5 are attached. From this, we estimate the Cy3–Cy5 distance is 60 Å in the undocked state and 40 Å in the docked state, corresponding to an E_{FRET} of ~0.3 and 0.8, respectively.