# CHAPTER IV

# Dissecting the Effects of the Intracellular Environment on the Hairpin Ribozyme[1]

## 4.1 Introduction

Substantial work has been done on the structural dynamics of the hairpin ribozyme under *in vitro* conditions. The hairpin ribozyme is a small, self-cleaving RNA molecule originally isolated from the negative-sense strand of the satellite RNA of the tobacco ringspot virus, where it plays an important role in the replication of the satellite RNA genome.[56] The active conformation of the hairpin ribozyme is formed when domains A and B, which make up the minimal motif, interact with each other in what is known as the docked state (Fig. 4.1).[66,93] In the undocked state, the internal loops of domains A and B are not interacting with each other, domains A and B are farther apart from each other and the active site is not formed.[66,93] Hence, the ribozyme is inactive in the undocked state.[66,93] From *in vitro* smFRET and single molecule kinetic fingerprinting experiments, the hairpin ribozyme's conformational dynamics have been well characterized such that rate constants have been determined for the

[1]Experiments pertaining to PEG in Tris-standard conditions (Fig. 4.5a) and the yeast extract without PEG condition (Fig. 4.6c) were performed and analyzed solely by Dr. May Daher. Experiments and analyses for Fig. 4.5b were done collaboratively by both Dr. May Daher and Wendy Tay. All other experiments were performed and analyzed by Wendy Tay. All figures were made by Wendy Tay.

known steps of the kinetic mechanism.[66,69,71] Additionally, heterogeneity has been observed in docking and undocking kinetics.[66,69,71] These observations are consistent with several other experiments documenting the heterogeneous nature of RNA folding as well.[66,189,210–216]

However, the intracellular environment is very different from a simple *in vitro* environment. Inside a cell, there is a high degree of macromolecular crowding from proteins, nucleic acids and other metabolites.[119–124] Cells typically contain 300-400 g/l of RNA and protein that take up 20 to 40% of a cell's volume.[119–124,134–136] RNA can be greatly affected by the excluded volume resulting from the highly crowded environment.[217–220] Additionally, there can be specific and nonspecific interactions between the components inside a cell.[129,221] For example, the highly crowded environment is thought to induce nonspecific attractive and repulsive interactions between macromolecules.[119,123,129,221] Also, there are proteins known to bind nonspecifically with RNA albeit at a lower affinity than for specific interactions.[128,130–133] These nonspecific interactions can affect a macromolecule's structure, dynamics and function.[119,129–133] Other key differences between intracellular and *in vitro* conditions are the composition of metal ions and the redox environment. The free magnesium concentration inside the cell is between 0.1 mM and 3.5 mM,[222–225] while *in vitro* experiments for the hairpin ribozyme typically use 10 to 100 mM $MgCl_2$.[66,69,71,77,84,89,93,198,210,226] There is also a high concentration of monovalent metal ions such as $K^+$ (130-170 mM) and $Na^+$ (10-30 mM) that are not typically present in *in vitro* conditions.[127,227,228] Inside the cell, there is a reducing environment due to the flow of electrons from oxidizable organic compounds to oxygen, thus providing energy for the cell.[126,229] This reducing environment protects cellular components against oxidative damage[229] and is speculated to trace back to when metabolic processes occurred in the absence of an oxygen-abundant atmosphere.[230] The behavior of the hairpin ribozyme in an intracellular environment has not yet been thoroughly characterized.

Several studies have already begun to elucidate the effects of macromolecular crowding on RNA. Macromolecular crowders have been shown to stabilize active conformations of RNA structures.[137–142] For instance, polyethylene glycol (PEG) stabilizes the association of subunits in the ribosome.[138] Also, studies on the hammerhead ribozyme indicate that protein, RNA, PEG and other kinds of crowders stabilize the active structure, resulting in higher activities in the presence of crowder.[139,140] These studies also suggest that crowders, while stabilizing the hammerhead ribozyme tertiary structure, simultaneously destabilize base pairing interactions, which thus allow for more rapid rearrangements of inactive conformations to reform into active structures.[139,140] More studies with the hammerhead ribozyme have shown that crowders increase ribozyme activity such that even with physiological magnesium concentrations, the ribozyme displays efficient catalysis.[141,142] Macromolecular crowders have also been shown to promote small, compact RNA structures due to the excluded volume effect of the crowder.[137,143,144] Studies by Kilburn et al. have shown that crowding due to PEG favors more compact conformations for both the folded and unfolded *Azoarcus* ribozyme.[143,144] Computational studies have shown that more compact structures of RNA are favored in crowded conditions.[231,232] Crowding effects, though, can depend on the size and electrostatic properties of the crowder.[140,145] For example, small organic osmolytes actually destabilize RNA structures.[145] In contrast to larger, inert crowders, like PEG, that promote entropic stabilization, osmolyte crowders, such as sugars, amino acids and methylamines, interact with the RNA surface and thereby affect the RNA structure.[145,146] Furthermore, a study by Nakano et al. showed that hammerhead ribozyme activity is higher when the crowder is PEG, as opposed to Ficoll, glycerol or dextran, and that a higher molecular weight PEG tends to result in a higher ribozyme activity.[140] However, macromolecular crowding is but one feature of the intracellular environment.

In this study, we performed smFRET experiments to gain insight into the con-

tribution of three major features of the intracellular environment on the structure and dynamics of the hairpin ribozyme. Macromolecular crowding alone should not be the only consideration when mimicking inside the cell.[233] Therefore, the first feature we investigated was the buffer environment inside the cell excluding crowding effects and any interactions with cellular components. The high $K^+$ concentration in the intracellular buffer outcompetes $Mg^{2+}$ for binding to the RNA, therefore causing destabilization of the docked state. The second feature was macromolecular crowding due to a commonly used inert crowder, PEG-8000. The crowder stabilizes the docked state, likely due to the excluded volume effect. Finally, the third feature we investigated was the interactions with cellular components present in cell extracts. These interactions are seen to be more potent than crowding by PEG-8000 in promoting the docked state. By considering these three features of the intracellular environment, we have gained understanding of the importance of buffer environment, macromolecular crowding and interactions with cellular components on RNA behavior.

## 4.2 Materials and Methods

### RNA preparation

RNA sequences were based on those from Rueda et al.[71] and are shown in Fig. 4.1. RNA strands were purchased with 2′-TBDMS protection groups from Keck Oligonucleotide Synthesis Facility (Yale University, New Haven, CT). RNA was deprotected according to the protocol from the manufacturer. Following deprotection, RNA was purified by C8-reverse-phase HPLC chromatography as described previously.[179] Purified RNA strands were reconstituted in double-distilled $H_2O$ and stored at -20°C. The RzA strand was ordered with a 5′ Cy5 and a 3′ amino C7 modifier, which was used to attach a Cy3 in-house (using a Cy3 succinimidyl ester, GE Amersham) as described previously.[66,93,181] The RzB strand was synthesized with a 5′ biotin. Also, a noncleav-

able substrate analog was used throughout, which contains a 2′OMe modification at the A-1 residue to prevent dissociation of the cleavage products from interfering with studies of domain docking and undocking.

### Tris-standard and intracellular-mimic buffers

The standard *in vitro* buffer used for hairpin ribozyme experiments, referred to as "Tris-standard", was 50 mM Tris-HCl (pH 7.5) with 12 mM $MgCl_2$ (unless otherwise specified). The intracellular-mimic buffer was designed to mimic near physiological conditions. The intracellular-mimic buffer, based on a buffer used by Edmonds et al.,[234] consists of 30 mM PIPES (pH 7.5), 10 mM NaCl, 130 mM $K^+$gluconate (pH ~7), 10 mM dithiothreitol (DTT), 1 mM ATP and 2.5 mM MgOAc (unless otherwise specified). Intracellular-mimic buffer was made fresh and filtered with a 0.2 $\mu$m Ultra-free-MC filter (Millipore) before each experiment.

### HeLa and yeast cell extracts

HeLa S100 (cytosolic) extract used for smFRET experiments was purchased from SPEED BioSystems. According to the manufacturer's website, the S100 extract contains a protein concentration of 6 mg/ml and is stored in dialysis buffer (20 mM HEPES (pH 7.9), 20% glycerol, 42 mM $(NH_4)_2SO_4$, 0.5 mM DTT, 0.2 mM EDTA). Upon receipt, extract was thawed on ice, divided into ~12 $\mu$l aliquots, flash frozen with liquid nitrogen and stored at -80°C until use.

Yeast whole extract was prepared in-house using the following protocol (Matthew Kahlscheuer, unpublished). One litre of yeast was grown in YPD (yeast extract peptone dextrose) media until an $OD_{600}$ of 1.6 to 2.0 was reached (approximately 2 days). Yeast cells were then harvested by centrifuging at 4,500 rpm for 15 min. Pellets from the 1 L culture were resuspended in 35 ml of AGK buffer (10 mM HEPES-KOH (pH 7.9), 1.5 mM $MgCl_2$, 200 mM KCl, 10% v/v glycerol, 0.5 mM DTT), then

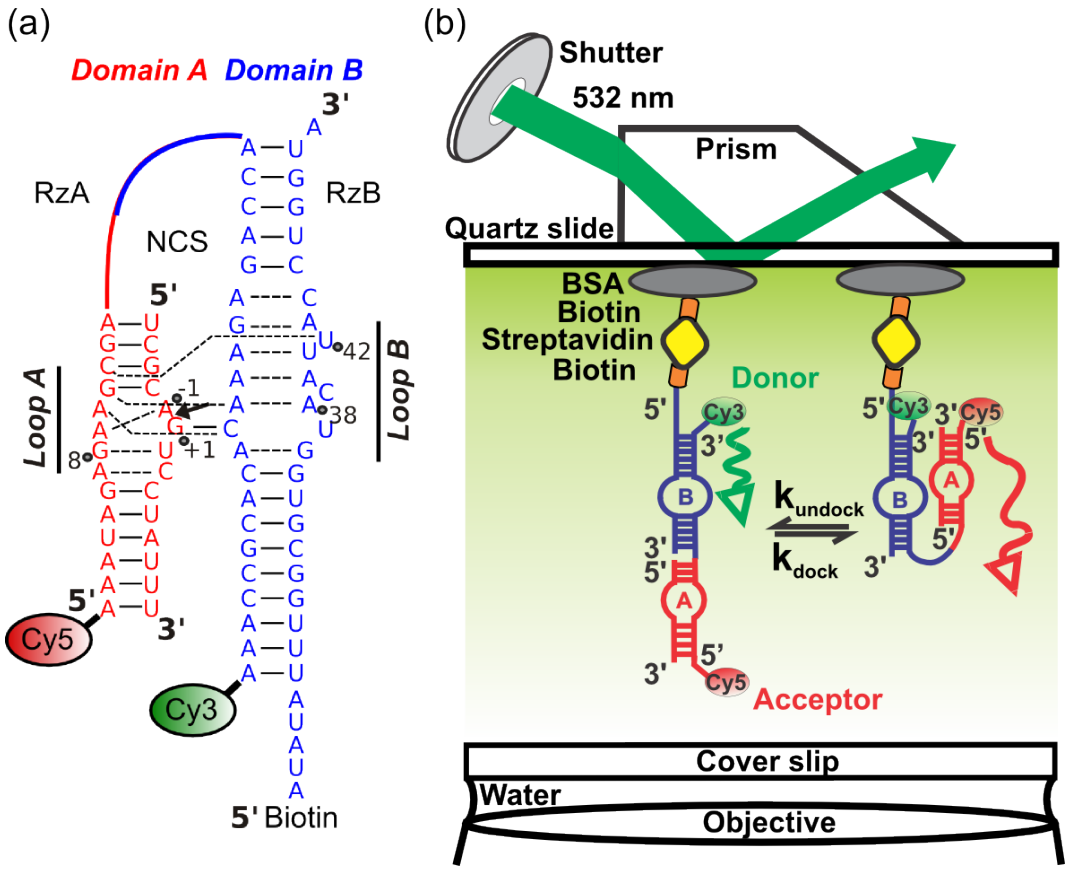


**Figure 4.1: The three-stranded hairpin ribozyme and smFRET setup. (a)** The three-stranded hairpin ribozyme used in the smFRET experiments. The ribozyme was made noncleavable by use of a noncleavable substrate analog (NCS) containing a 2′OMe modification at the A-1 residue. The positions of Cy3, Cy5 and biotin are also indicated. Short, solid lines indicate canonical hydrogen bonding; dashed lines indicate non-canonical hydrogen bonding. **(b)** The smFRET experimental setup. Ribozymes are attached to a quartz slide via biotin-streptavidin interactions. Illumination of the molecules is achieved by the evanescent wave produced by total internal reflection of fluorescence (TIRF) of the excitation laser light. The hairpin ribozyme is shown in both its docked and undocked conformations. Modified from an original image by Professor Nils Walter.

centrifuged again at 4,500 rpm for 15 min at 4°C. All pellets were combined and resuspended in 30 ml of AGK buffer, 600 $\mu$M PMSF and 1.5 mM benzamidine, then centrifuged in a swinging bucket rotor at 2,500 rpm for 10 min. The supernatant was completely decanted, and 0.4 volumes (based on the volume of the pellet) of AGK buffer with 600$\mu$M PMSF and 1.5 mM benzamidine was added and vortexed to make a thick cell suspension. The suspension was frozen down as pellets by dropwise addition of the cell suspension to a 50 ml tube filled with liquid nitrogen. The frozen pellets were mixed with liquid nitrogen and then ground up using a mortar and pestle pre-cooled in liquid nitrogen. The yeast cells were ground in liquid nitrogen until a fine powder was formed. The powder was then thawed at 4°C in a water-ice bath for 30 to 40 min. The cell lysate was centrifuged at 17,000 rpm for 30 min at 4°C using a SS-34 fixed angle rotor (Thermo Scientific). The supernatant was transfered to another pre-chilled ultracentrifuge tube and spun at 37,000 rpm for 1 h and 20 min at 4°C in a type 70 Ti rotor (Beckman Coulter). No more than two-thirds of the clear pale phase in the middle of the centrifuge tube was transferred to a pre-chilled 15 ml conical tube. The second layer from the top was extracted from the conical tube and dialyzed twice against 2 L of dialysis buffer (20 mM HEPES-KOH (pH 7.9), 0.2 mM EDTA, 50 mM KCl, 20% v/v glycerol, 0.5 mM DTT) at 4°C for 1.5 h at a time. After dialysis, cell extract was microcentrifuged at 13,000 rpm for 5 min at 4°C then frozen in liquid nitrogen in 50 $\mu$l aliquots and stored at -80°C until use. Protein concentration in the extracts was determined by Bradford assay. Nucleic acid concentration was determined from absorbance at 260 nm.

**PEG-8000 as a macromolecular crowder**

For conditions with PEG-8000 (PEG with an average molecular weight of 8,000 Da), a 50% w/v PEG-8000 solution was prepared fresh on the day of the experiment. PEG was dissolved in double-distilled water and stirred with gentle heating to dis-

solve. Once dissolved, the PEG solution was filtered with a 0.2 $\mu$m nitrocellulose filter (Bio-Rad).

**Single molecule FRET**

The RzA and RzB strands were annealed using 2 $\mu$M of RzA and 1 $\mu$M of RzB in either Tris-standard buffer or intracellular-mimic buffer with 3% PEG. The annealing solution was heated to 90°C for 45 s in the absence of magnesium or PEG. Magnesium and PEG (where applicable) was then added and the solution was allowed to cool to room temperature for 15 min. The annealed ribozyme was diluted to ~50 pM and flowed onto a streptavidin-coated quartz slide. Streptavidin was attached via biotin-streptavidin interactions. For Tris-standard conditions, biotinylated-bovine serum albumin (BSA) was used to adhere to the slide. For intracellular-mimic conditions, slides used had biotin attached to the slide via PEG-succinimidyl valerate (biotin-PEG-SVA, MW 5,000, Laysan Bio) and were passivated with PEG (mPEG-SVA, MW 5,000, Laysan Bio). The solution in which the ribozyme was imaged, referred to as "imaging solution" contained 400 nM of the noncleavable substrate analog, 4.3 mM Trolox (Fisher) to reduce photoblinking, and an oxygen scavenging system consisting of 4.3 mM protocatechuic acid (Sigma-Aldrich) and 43 nM protcatechuate 3,4-dioxygenase (Sigma-Aldrich). Imaging solution for Tris-standard conditions also contained Tris-standard buffer. Imaging solution for intracellular-mimic conditions contained intracellular-mimic buffer and 80 units of recombinant RNasin ribonuclease inhibitor (Promega) as well as, depending on the condition, cell extract and/or PEG. Imaging solution was incubated at room temperature for 5 min before flowing onto the slide, followed by another 5 minute incubation at room temperature before imaging.

The slide was mounted onto a prism-based TIRF (total internal reflection of fluorescence) microscope (Olympus IX71) with a 60X water-immersion objective (Olympus). Figure 4.1 shows a schematic of the experimental setup. The Cy3 fluorophore

was directly excited using a 532 nm solid-state continuous wave laser and the Cy5 fluorophore was also directly excited for part of the time using 635 nm solid-state continuous wave laser. Emission from both Cy3 and Cy5 was collected simulataneously using a CCD camera (I-PentaMAX HQ) at 100 ms time resolution. Time traces of the Cy3 and Cy5 fluorescence intesities were extracted from the raw movies using IDL (Research Systems) and then manually analyzed using Matlab (The Math Works).[174] Traces kept for further analysis were greater than 50 frames and showed single-step photobleaching to ensure that a single molecule was being tracked. A FRET ratio was calculated using the equation $F = I_A/(I_A + I_D)$, where $I_A$ and $I_D$ represent the background corrected fluorescence intensities of the acceptor (Cy5) and donor (Cy3) fluorophores, respectively. Histograms of FRET distributions were plotted using Origin (OriginLab 8.1). Ratios of static to dynamic molecules were determined manually. Only traces greater than 500 frames were included in this analysis. Also, in order to compare traces of drastically different lengths, traces were assessed as dynamic or static based on between 500 and 1000 frames. Dynamic molecules had to meet the criteria that there was at least one FRET transition, excluding FRET changes due to photobleaching or blinking.

**Native polyacrylamide gel electrophoresis**

The RzA and RzB strands were annealed as described in the smFRET section above. The annealed ribozyme was diluted to ~56 nM (as determined by the concentration of RzB) and mixed with 2.8 $\mu$M of non-cleavable substrate analog, intracellular-mimic buffer, 24 units of RNasin and either 68% v/v or 15% v/v of HeLa S100 extract (Jena Bioscience). The solution was incubated at 25°C using a water bath for 5 min and for 1 h and 45 min. After incubation, aliquots of the solution were quenched with a stop solution containing 10% Contrad-70 at pH 9.3 and immediately placed on ice. Samples were run on a nondenaturing 8% polyacrylamide (19:1

acrylamide:bisacrylamide) gel containing 2.5 mM MgOAc and 50 mM Tris-acetate. Prior to running the samples, the gel was pre-run for at least 2 h. Then, the samples were run for 7 h at 4°C using 13 W maximum power. The gel was imaged using a Typhoon 9410 Variable Mode Imager to detect Cy3 (excitation at 532 nm, 600 PMT, 580 ± 30 nm band-pass emission filter) and Cy5 (excitation at 633 nm, 600 PMT and 670 ± 30 nm band-pass emission filter).

## 4.3 Results

### The intracellular-mimic buffer in the absence of crowders or cell extract favors the low FRET state

We first determined the effect of the intracellular environment, excluding crowding and cellular interactions, on the hairpin ribozyme FRET distribution. To do so, we made use of an intracellular-mimic buffer, a buffer intended to simulate the physiological conditions inside a cell. The intracellular-mimic condition contains a much higher amount of monovalent ions (130 mM of $K^+$ and 10 mM $Na^+$) compared to the amount of $Mg^{2+}$ present and, therefore, was predicted to destabilize the docked state due to competition between $K^+$ and $Na^+$ with $Mg^{2+}$ for electrostatic interactions with RNA.[25,77,235–237] Intracellular-mimic conditions used for these experiments contained 5 mM MgOAc instead of 2.5 mM because there were not enough docking events in the lower $Mg^{2+}$ condition and hence did not result in usable smFRET data. The rest of the intracellular-mimic buffer components were as described in the Materials and Methods. Compared to Tris-standard conditions, intracellular-mimic conditions favor the lower FRET state (Figs. 4.2a,d), consistent with our hypothesis that the high concentration of monovalent ions outcompetes $Mg^{2+}$, which is needed for docking to occur, for binding to the ribozyme. The high and low FRET states we see are consistent with previously observed values for the docked and undocked states

respectively.[66,69,71] However, low concentrations of $Mg^{2+}$, even in Tris-standard buffer (that contains 12 mM $MgCl_2$), favor the undocked state (Fig. 4.2b). Therefore, sm-FRET experiments were performed for the intracellular-mimic condition with 12 mM MgOAc in addition to 5 mM MgOAc to confirm that favoring of the undocked state is not due solely to the lower $Mg^{2+}$ concentration (Fig. 4.2c). With 12 mM MgOAc, the intracellular-mimic condition also favors the undocked state (Fig. 4.2c). A series of experiments was done where each component of the intracellular-mimic buffer was added in a stepwise fashion (Fig. 4.3). The shift in FRET distribution from a greater proportion in the docked state to a greater proportion in the undocked state is seen after addition of the 130 mM potassium gluconate, confirming that the observed effect from the intracellular-mimic buffer is primarily due to metal competition for RNA binding between the high amount of $K^+$ compared to 12 mM $Mg^{2+}$.

Regarding the effect on dynamics by the intracellular-mimic buffer, the number of static and dynamic single molecule traces were counted for each condition. Static traces do not show any FRET transition while dynamic traces show at least one FRET transition (not due to photobleaching of either the Cy3 or Cy5 fluorophores). The intracellular-mimic conditions cause the single molecule traces to be shorter-lived compared to Tris-standard conditions (Fig. 4.4). Therefore, in order to have a fair comparison of static vs. dynamic traces across the different conditions, traces were only considered in the static/dynamic analysis if they were longer than 500 frames (50 seconds) and traces were not analyzed past 1000 frames (100 seconds). Generally, there were less dynamic molecules (labeled "D" in Fig. 4.2 histograms) compared to static molecules (labeled "S" in Fig. 4.2 histograms) in the intracellular-mimic buffer, though the effect is more clearly seen in the presence of 5 mM $Mg^{2+}$ ("S" and "D" labels in Fig. 4.2). Comparing just the Tris-standard conditions, a lower $Mg^{2+}$ concentration results in more dynamic molecules compared to static molecules (indicated by the "S" and "D" labels in Figs. 4.2a,b). These results can be explained

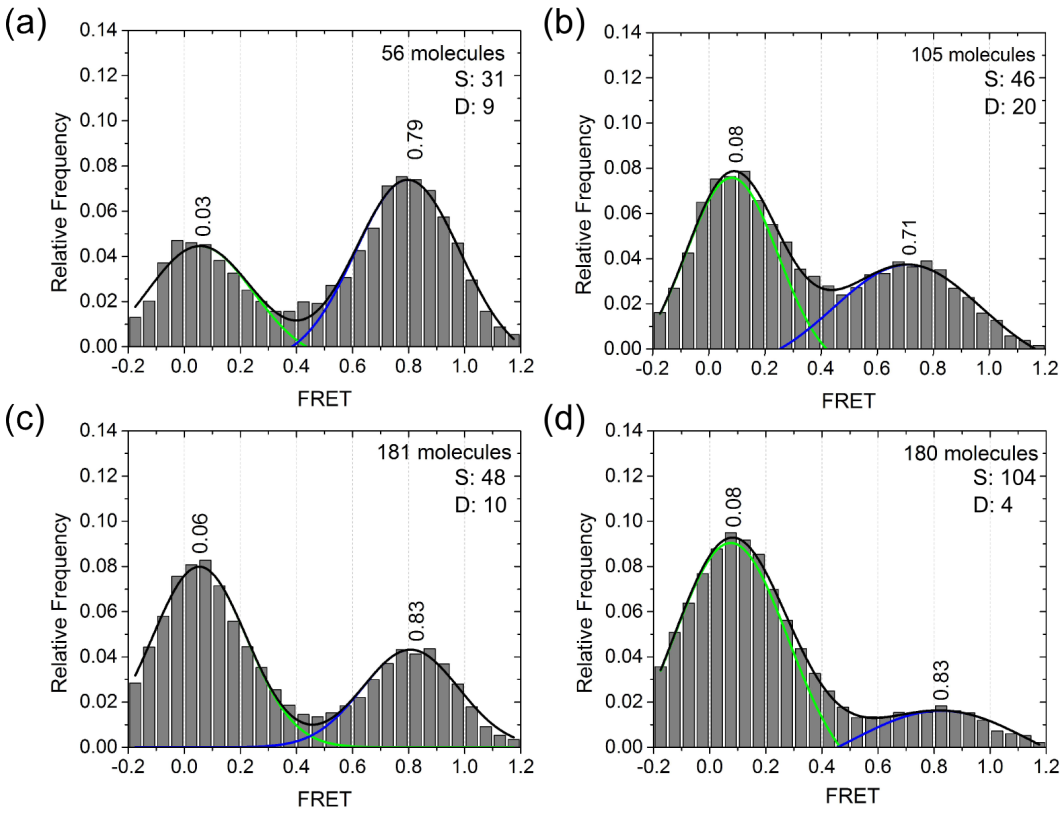


**Figure 4.2: Histograms of the FRET distribution in Tris-standard and intracellular-mimic conditions.** **(a)** Tris-standard conditions (50 mM Tris-HCl (pH 7.5), 12 mM $MgCl_2$). **(b)** Tris-standard conditions with 5 mM $MgCl_2$. **(c)** Intracellular-mimic conditions (30 mM PIPES (pH 7.5), 10 mM NaCl, 130 mM $K^+$gluconate (pH ~7), 10 mM DTT, 1 mM ATP) with 12 mM MgOAc. **(d)** Intracellular-mimic conditions with 5 mM MgOAc. Histograms are fitted to a two-Gaussian distribution and the means of each Gaussian are indicated. "S" and "D" indicate the number of static and dynamic molecules respectively.

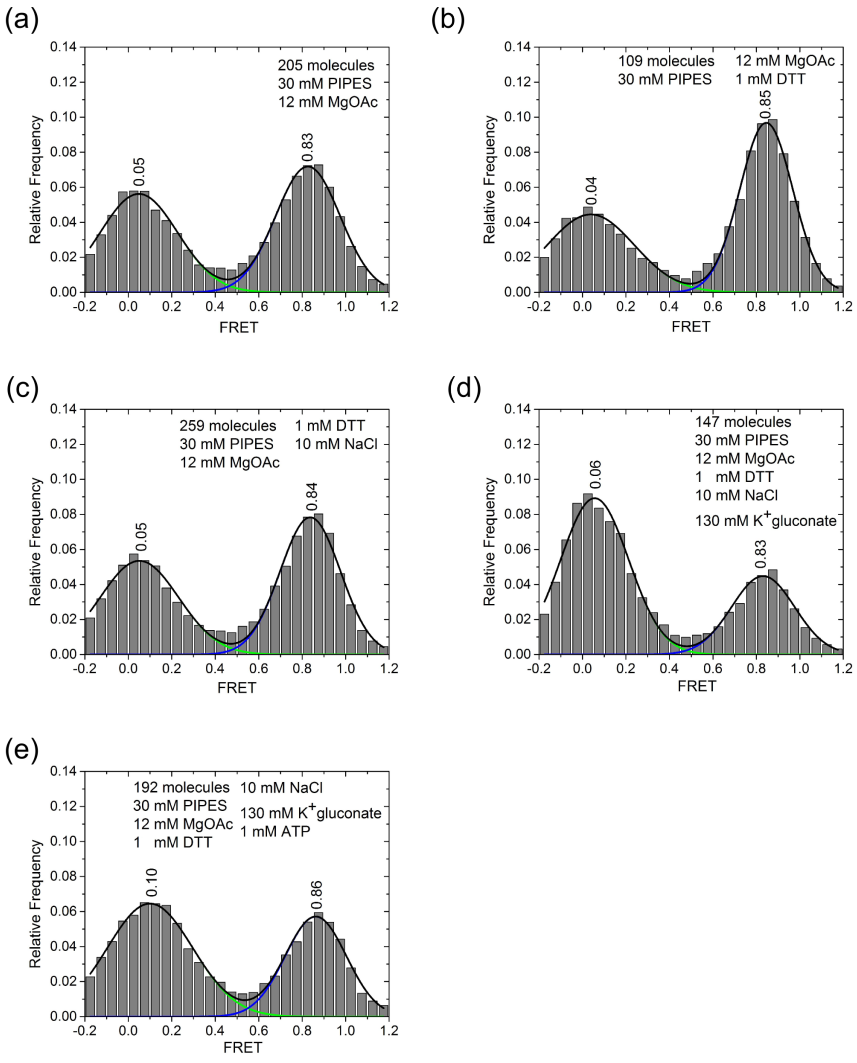


**Figure 4.3: Stepwise addition of the components in the intracellular-mimic buffer.** **(a)**-**(e)** Addition of each component of the intracellular-mimic buffer. Histograms are fitted to a two-Gaussian distribution and the means of each Gaussian are indicated.

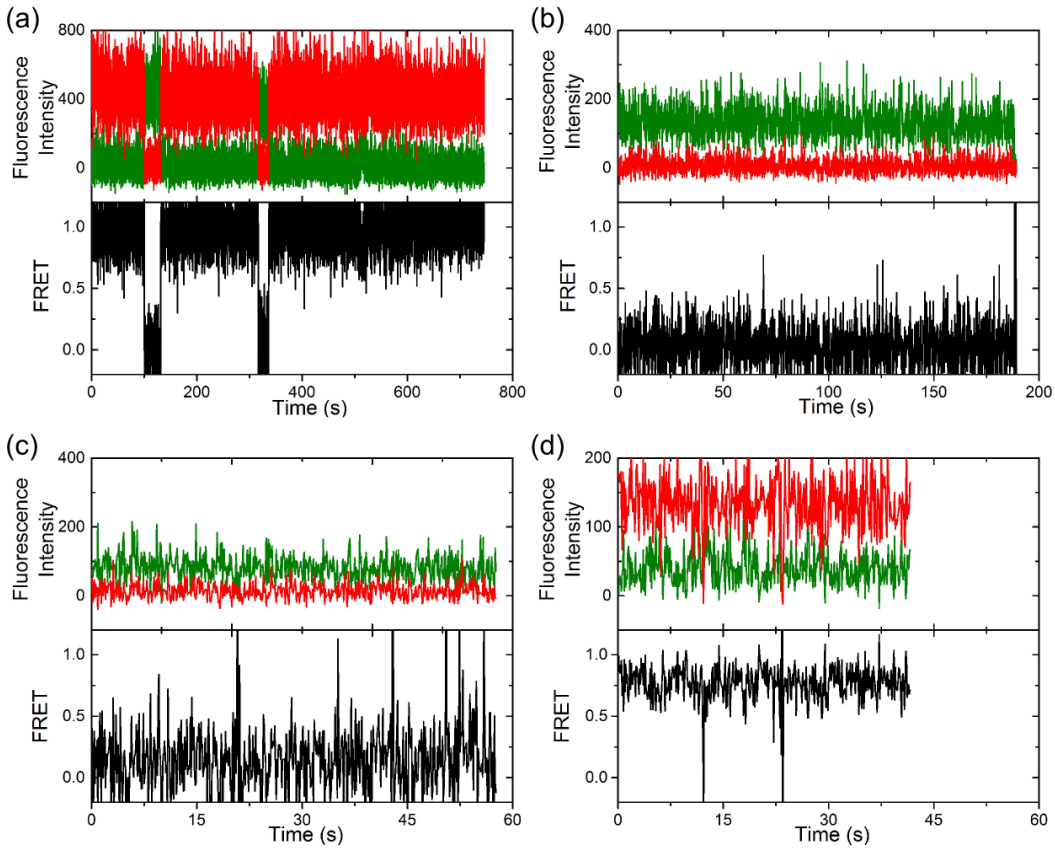


**Figure 4.4: Representative smFRET traces for various experimental conditions.** **(a)** Tris-standard conditions. **(b)** Intracellular-mimic conditions with 12 mM MgOAc. **(c)** Intracellular-mimic conditions with 3% PEG. **(d)** Intracellular-mimic conditions with 3% PEG and ~13 mg/ml yeast whole cell extract. Green traces indicate Cy3 emission and red traces indicate Cy5 emission with arbitrary units for fluorescence intensity. Black traces indicate the calculated FRET values.

by a lower propensity of 5 mM $Mg^{2+}$ to stabilize the docked state compared with 12 mM $Mg^{2+}$, thereby causing more fluctuations between the two states. However, in the intracellular-mimic condition, a lower $Mg^{2+}$ concentration results in less dynamic molecules compared to static molecules (indicated by the "S" and "D" labels in Figs. 4.2c,d). For the intracellular-mimic conditions, 5 mM $Mg^{2+}$ may not have been enough to compete with $Na^+$ and $K^+$ ions to even cause enough docking events resulting in the majority of molecules staying statically in the undocked state. With 12 mM $Mg^{2+}$, the $Mg^{2+}$ concentration may be high enough to cause more docking events though it is unable to fully outcompete the inhibitory effects of $Na^+$ and $K^+$ in order to stabilize the docked state to the same extent as in Tris-standard buffer.

**Crowding by PEG favors the more compact docked state**

We next looked at how macromolecular crowding by PEG affects the hairpin ribozyme. PEG-8000 is commonly used to simulate macromolecular crowding.[140–144,238–242] We chose to use PEG because of its relatively weak, minimal interactions with RNA.[144,145] Based on previous studies discussing how crowders promote more compact, smaller RNA structures,,[137,143,144,231,232] we predicted to see in the presence of PEG a favoring of the docked state, which is more compact than the undocked state. Crowding inside the cell is most similar to 20% PEG conditions, but lower concentrations of PEG were also assessed. In Tris-standard conditions, increasing the percentage of PEG from 3% to 20% (w/v) increases the proportion of the high FRET or docked state (Fig. 4.5a), consistent with our prediction. Intracellular-mimic conditions with 1.5%, 3% and 20% PEG show the same trend of higher PEG concentrations favoring the docked state (Fig. 4.5b). A high enough amount of PEG of 20% even overcomes the low 2.5 mM $Mg^{2+}$ concentration in the intracellular-mimic buffer to favor the docked state (Fig. 4.5b). Additionally, in the Tris-standard conditions, the proportion of dynamic to static molecules slightly decreases with increasing PEG

concentration (indicated by the "S" and "D" labels in Fig. 4.5a). The increased crowding due to PEG can conceivably be inhibiting conformational fluctuations. In contrast, more PEG in the intracellular-mimic conditions increases the proportion of dynamic molecules (indicated by the "S" and "D" labels in Fig. 4.5b). The 3% PEG condition in intracellular-mimic buffer ("S" and "D" labels in Fig. 4.5b, middle) shows a particularly high proportion of dynamic molecules compared to the expected trend based on the proportion of dynamic molecules seen in the 1.5% and 20% PEG conditions ("S" and "D" labels in Figs. 4.5b, top and bottom), but the reason is unclear. The competing effects of the PEG crowder favoring the docked state and the intracellular-mimic buffer favoring the undocked state may be the cause of the more dynamic traces seen when compared to conditions with either PEG or the intracellular-mimic buffer alone.

For the Tris-standard conditions with increasing PEG concentrations, a couple of minor, but interesting, trends are observed. First, the mean value of the low FRET state shifts 0.1 units to a higher FRET value from 3% to 20% PEG (Fig. 4.5a). This trend suggests the favoring of a slightly more compact undocked structure with increasing PEG, consistent with predictions based on the excluded volume effect of PEG.[137,143,144] Secondly, and perhaps more intriguingly, the mean value of the high FRET state subtly decreases with increased PEG, though to a lesser degree than the shift in the low FRET state (Fig. 4.5a). We know that the hairpin ribozyme displays heterogeneity in the docked and undocked states.[25,66,69,71] Therefore, one possible explanation for this trend is that the excluded volume effect is not the only consideration from the presence of PEG. PEG may actually be stabilizing a specific docked sub-state, as proposed by previous studies for other RNAs,[137,139,140,142] which has a FRET value closer to 0.67 than 0.72.

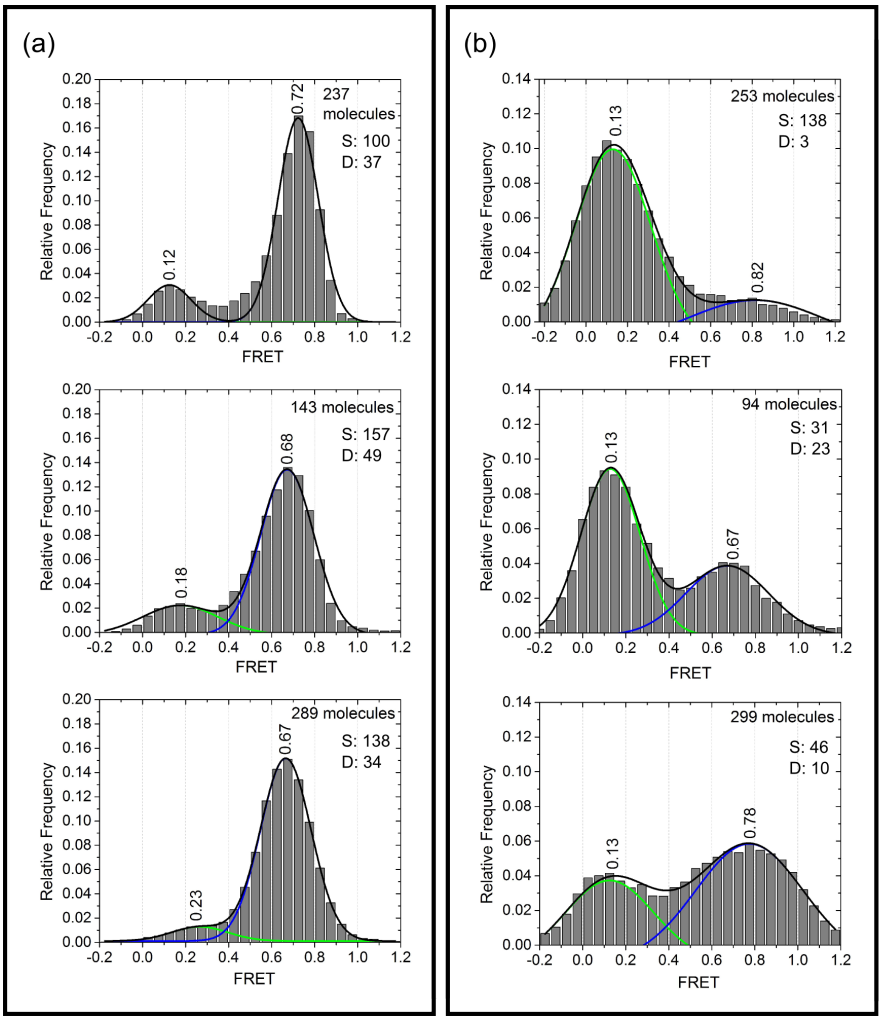


**Figure 4.5: The effect of PEG-8000 on the FRET distribution and dynamics of the hairpin ribozyme. (a)** Tris-standard conditions with 3% (top), 8% (middle) and 20% (bottom) w/v PEG. Experiments performed by Dr. May Daher. **(b)** Intracellular-mimic conditions (containing 2.5 mM MgOAc) with 1.5% (top), 3% (middle) and 20% (bottom) PEG. Histograms are fitted to a two-Gaussian distribution and the means of each Gaussian are indicated. The middle histogram for (b) fit best to a one-Gaussian distribution. "S" and "D" indicate the number of static and dynamic molecules respectively. Experiments performed together with Dr. May Daher.

**The presence of cell extract favors the docked state and dynamic molecules**

The previous section involving PEG probed the effects of macromolecular crowding without (or with minimal) interactions with the RNA surface. In this section, we wanted to investigate the effect on the hairpin ribozyme FRET distribution and dynamics of interactions between the ribozyme and cellular components. Hence, we first tested the effect of 0.9 mg/ml (protein) HeLa S100 (cytosolic) extract in intracellular-mimic buffer with 5 mM MgOAc. Compared to intracellular-mimic conditions with 5 mM MgOAc in the absence of extract (Fig. 4.2d), the presence of HeLa extract results in a greater proportion of the high FRET state (Fig. 4.6a). Also, the proportion of dynamic to static molecules drastically increases in the presence of extract (indicated by the "S" and "D" labels in Fig. 4.6).

We ran a nondenaturing polyacrylamide gel to show that the hairpin ribozyme did interact with components of the HeLa cell extract (Fig. 4.7). The hairpin ribozyme was incubated for 5 min or 1 h and 45 min in either Tris-standard or intracellular-mimic buffer containing either ~0.9 mg/ml or ~4 mg/ml (protein) HeLa extract. After incubation, the solutions were separated on an 8% nondenaturing polyacrylamide gel and scanned for Cy3 and Cy5 fluorescence. The supershifted bands present only in conditions with extract indicate interactions of the hairpin ribozyme with cellular components even at a low concentration of ~0.9 mg/ml extract (Fig. 4.7). Although an incubation of 5 min is enough to see supershifted bands, the longer incubation at 1 h and 45 min shows a less intense free ribozyme band, which runs the same distance as the band in the Tris-standard lanes, presumably because of increased association with cellular components (Fig. 4.7, blue arrow) and some minor degradation (Fig. 4.7, red arrow).

However, the protein concentration of the HeLa extract is very low compared to the 200-300 mg/ml typically found inside cells.[119–124,134–136] We hypothesized that use of a more concentrated extract would have a more pronounced effect on the FRET

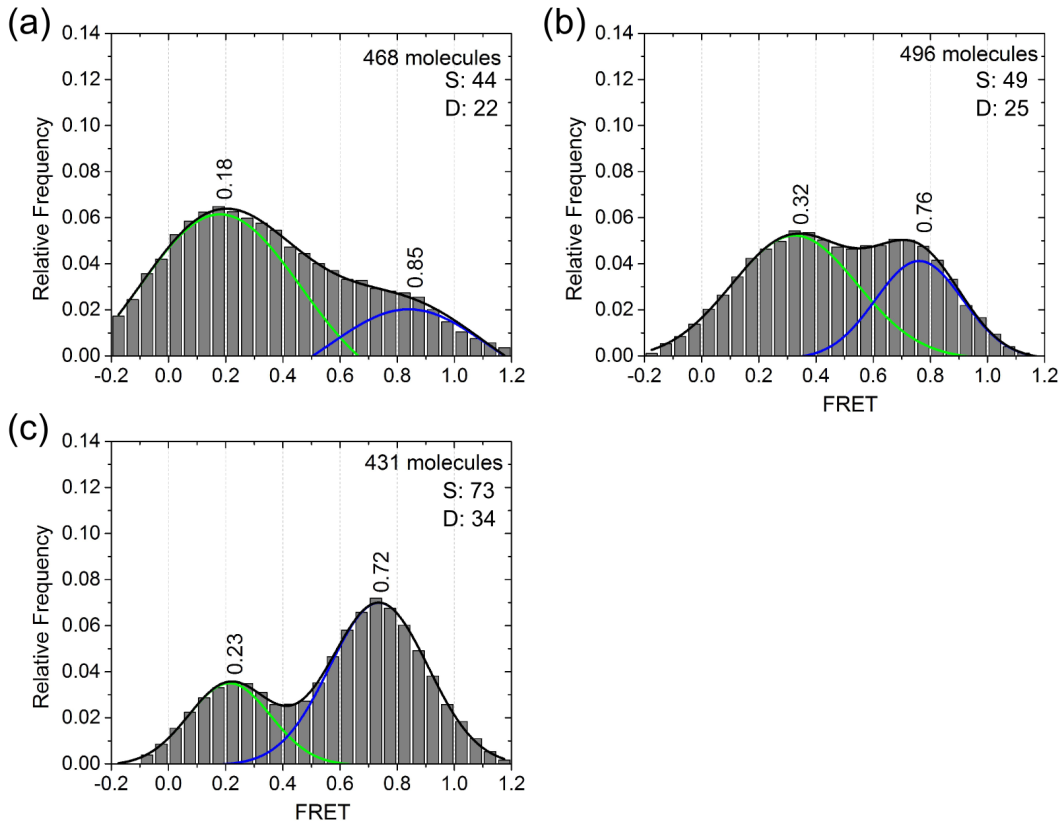


**Figure 4.6: FRET distributions and the number of static and dynamic molecules for conditions with cell extract.** **(a)** HeLa S100 (cytosolic) extract at a concentration of ~0.9 mg/ml (protein). **(b)** Yeast whole cell extract at a concentration of ~13 mg/ml (protein) with 3% PEG. **(c)** Yeast whole cell extract at a concentration of ~11 mg/ml (protein, ~14 mg/ml including protein and nucleic acids). All three conditions are in intracellular-mimic buffer with 5 mM MgOAc for (a) and 2.5 mM MgOAc for (b) and (c). Histograms were fitted to a two-Gaussian distribution and the means of each Gaussian are indicated. "S" and "D" indicate the number of static and dynamic molecules respectively. Experiments and analysis for (c) were done by Dr. May Daher.

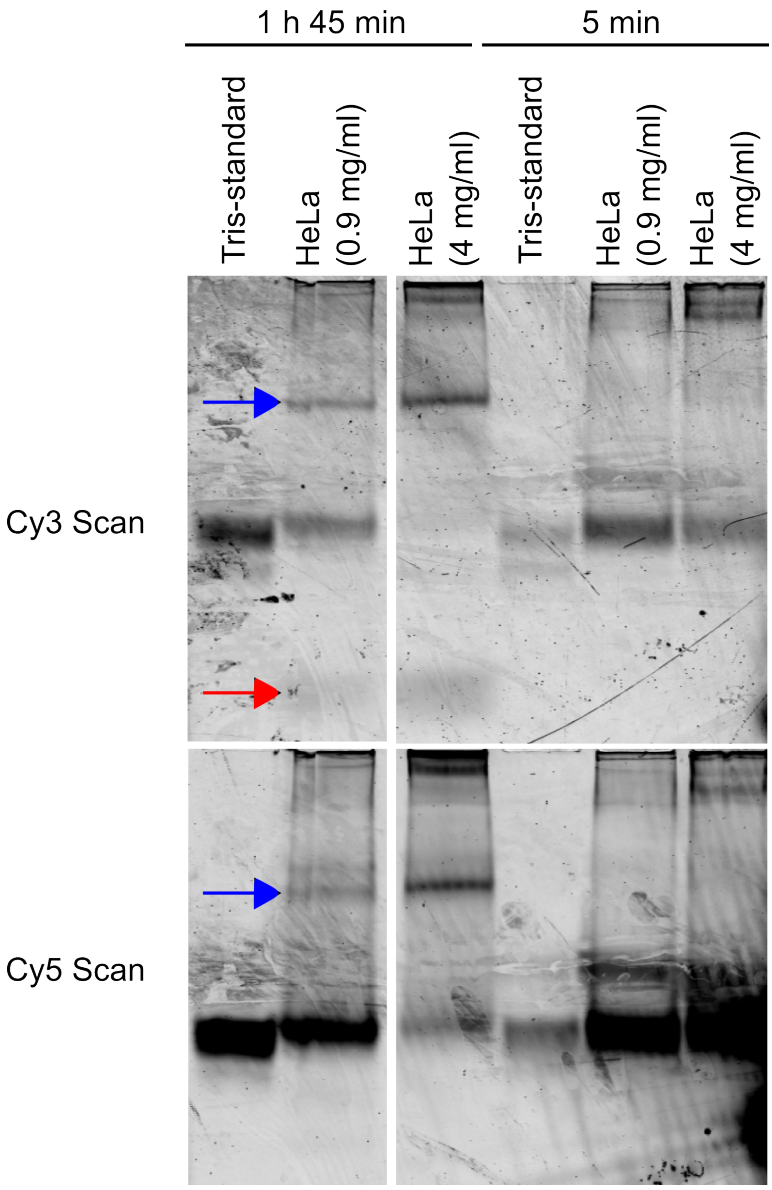


**Figure 4.7: Nondenaturing gel of the hairpin ribozyme incubated with Tris-standard buffer or with intracellular-mimic buffer and HeLa cell extract.** The hairpin ribozyme was incubated with either Tris-standard buffer or intracellular-mimic buffer with ~0.9 mg/ml or ~4 mg/ml (protein concentrations) HeLa cell extract. The solutions were then run on an 8% polyacrylamide (19:1 acrylamide:bisacrylamide) gel containing 2.5 mM MgOAc and 50 mM Tris-acetate. The blue arrow indicates a supershifted band that only appears in the 1 h 45 minute incubation but not in the 5 minute incubation condition. The red arrow indicates minor degradation seen in the Cy3 scan in the 1 h 45 minute incubation condition.

distribution. We therefore used yeast whole cell extract made in-house with a higher protein concentration of ~16-20 mg/ml (stock). Incorporating the maximum volume of extract into our experiments, we were able to conduct experiments with ~13 mg/ml protein. These experiments with yeast extract also used a more physiologically relevant $Mg^{2+}$ concentration of 2.5 mM. In order to collect enough traces that sampled the docked state, we found we needed some crowder present and thus used 3% PEG when using 2.5 mM $Mg^{2+}$. The FRET distribution from using ~13 mg/ml extract shows, as predicted, a higher proportion of docked molecules (Fig. 4.6b) than in conditions containing 0.9 mg/ml HeLa extract (Fig. 4.6a) or in intracellular-mimic buffer with 3% PEG (Fig. 4.5b, middle). Moreover, the mean FRET value of the low FRET state in the yeast extract conditions is particularly higher than in all other conditions studied (Fig. 4.6b). This higher-value low FRET state may be a due to the formation of a more compact conformation of the undocked state or even an alternative fold. The higher concentration of yeast extract or a component specifically present in yeast extract, but not in HeLa extract, in combination with the presence of PEG may be the reasons for the higher mean value low FRET state. Additionally, similar to when HeLa extract is used, yeast extract also favors more dynamic molecules. Compared to intracellular-mimic buffer with 3% PEG, the condition with ~13 mg/ml yeast extract shows a much higher proportion of dynamic molecules ("S" and "D" labels in Fig. 4.6b). However, the proportion of dynamic molecules is more or less the same when comparing 0.9 mg/ml HeLa extract with 5 mM MgOAc to ~13 mg/ml yeast extract with 2.5 mM MgOAc and 3% PEG (Figs. 4.6a,b).

We have so far determined that both the presence of cell extract and PEG, in intracellular-mimic buffer, promote the docked state as well as an increased proportion of dynamic relative to static molecules. In order to compare the degree of impact by cell extract versus PEG, we tested intracellular-mimic conditions containing 2.5 mM MgOAc and ~14-15 mg/ml of either PEG-8000 or yeast cell extract (this time taking

into account the concentration of both protein and nucleic acids). The yeast extract condition much more prominently favors the high FRET state as well as a greater proportion of dynamic molecules compared to the 1.5%, or 15 mg/ml, PEG condition (Fig. 4.6c, Fig. 4.5b, middle). Therefore, we conclude that the interactions of the hairpin ribozyme with cell extract can have a much more pronounced effect than the crowding effects from a similar concentration of PEG. This pronounced effect can only begin to be comparable by macromolecular crowding, at least in terms of FRET distribution, when the environment is very crowded, such as in 20%, or 200 mg/ml, PEG conditions. Even then, the ribozyme's nonspecific interactions with ~14 mg/ml of yeast whole cell extract still have a greater impact on the proportion of dynamic molecules than 20% PEG (Fig. 4.6c, Fig. 4.5b, middle).

## 4.4 Discussion

We have investigated the effects of three major features of the intracellular environment. The first feature was the environment inside the cell but excluding crowding effects or nonspecific interactions with cellular components. A comparison between the Tris-standard and intracellular-mimic buffers show that a cell-like composition destabilizes the high FRET state, or undocked state, primarily due to competition with $Mg^{2+}$ by $K^{+}$. There are also less dynamic molecules in the intracellular-mimic buffer compared to the Tris-standard buffer. The second feature was the highly crowded environment of the cell, simulated by 20% w/v PEG-8000. In both Tris-standard and intracellular-mimic conditions, increasing concentrations of PEG results in increased ratios of docked, or high FRET, state to undocked state. Regarding dynamics, however, PEG promotes more dynamic molecules in the intracellular-mimic conditions but more static molecules in the Tris-standard conditions. Finally, the third major feature we investigated was the effect of interactions of the hairpin ribozyme with components in cell extract. We find that both HeLa cytosolic and yeast whole cell

extracts have a significant impact on promoting the docked state. The presence of extract also promotes more dynamic molecules than in the absence of extract. A comparison of an equivalent concentration in mg/ml of PEG and cellular extract also shows that the interactions of the hairpin ribozyme with cellular components have a more pronounced effect on both the FRET distribution and dynamics of the hairpin ribozyme than any crowding effects at that concentration. A very crowded environment simulated by 20% PEG can produce a similar FRET distribution to conditions with ~14 mg/ml yeast extract, but does not promote dynamic molecules as much as the extract does.

The hairpin ribozyme is a good model system for exploring the effects of the intracellular environment on RNA structure and dynamics. Firstly, there are already many *in vitro* studies on the dynamics of the hairpin ribozyme.[25,66,69,71,93] Therefore, knowledge gained about intracellular contributions on RNA structure and dynamics can be directly compared to the *in vitro* results from these previous experiments. It would be useful to ascertain major differences between experiments under *in vitro* conditions and intracellular-like conditions in order to determine how biologically relevant the results are. The second reason why the hairpin ribozyme is a useful system is because it displays clearly distinguishable docked and undocked conformations that interconvert on the second time scale easily detectable by smFRET.[66,69,71,93,243] Previous experiments established the undocked, docked and substrate-free states as having mean FRET values of approximately 0.15, 0.8 and 0.4, respectively.[66] In this study, the undocked and docked states, taking into account all conditions, similarly display mean FRET values of ~0.13 ± 0.1 and ~0.76 ± 0.09, respectively. The substrate-free state is not observed.

In our hands, macromolecular crowding, the most commonly investigated feature of the intracellular environment, produces results that are consistent with previous experiments.[137,143,144] Although the effect of 20% PEG on dynamics depends on whether

Tris-standard or intracellular-mimic buffer is used, the effect on FRET distribution is the same: the higher FRET state is favored. Previous experiments using PEG-8000 and other crowders also show favoring of a more compact, smaller RNA structure because of the excluded volume effect.[137,143,144] Previous crowding studies demonstrated that PEG destabilizes the secondary RNA structure and stabilizes the active tertiary structure, which minimizes inactive misfolding and thus increases activity.[137,139,140,142] The effects are proposed to happen via a "macromolecular collision effect" or weak, transient, nonspecific interactions between the RNA and PEG.[140] These effects may be happening with the hairpin ribozyme as well and may be the explanation for trends such as the subtle decrease observed in the mean FRET value of the high FRET state as the PEG concentration is increased in Tris-standard conditions (Fig. 4.5a). However, our studies using smFRET and a noncleavable version of the hairpin ribozyme cannot speak directly about misfolded structures or destabilized secondary structures, especially if such structures produce similar FRET values to the correctly folded undocked and docked states. Therefore, while we do not exclude the possibility of macromolecular collision effects, we can clearly attest to the excluded volume effect by PEG on the hairpin ribozyme. Given the complicated make-up of the macromolecular crowders inside the cell,[119–124] PEG-8000 is a reasonable mimic of crowding in the cell. Certainly other sizes and shapes of crowders can be used, such as PEG-200, Ficoll or even bovine serum albumin, which may give different results.[129,137,144,145] Our use of PEG-8000, a commonly-used molecular crowder,[140–144,238–242] has perhaps but spearheaded a series of experiments about the effects of various crowders on the hairpin ribozyme. Such experiments may be explored in future studies.

Compared to crowding, the impact of the intracellular-mimic buffer, without crowding or cellular components, is very different. Our data show that the intracellular-mimic buffer favors the low FRET state compared to Tris-standard buffer conditions primarily due to the high $K^+$ concentration compared to $Mg^{2+}$. These results are

consistent with many other studies on how nucleic acids in the presence of high ionic strength, like that of the cellular environment, are more compact due to the shielding by ions of the negatively charged backbone.[25,77,235–237] Additionally, our data show that the intracellular-mimic buffer favors generally less dynamic molecules compared to Tris-standard buffer conditions. The intracellular-mimic buffer as a whole is not overwhelmed by, nor does it overpower, the effects by PEG and cell extract. Instead, the intracellular-mimic buffer has a significant influence along with crowders and cellular component interactions on the behavior of the hairpin ribozyme.

Finally, in the investigation of the effect of the hairpin ribozyme's interactions with cellular components, we found that both HeLa and yeast cell extract favors the docked state as well as a greater proportion of dynamic molecules. Cellular components may be destabilizing secondary structures but stabilizing the catalytically-relevant tertiary structure, similar to previously proposed explanations for the stabilization of the hammerhead ribozyme by PEG.[139,140] The interactions with cellular components may therefore be akin to a chaperone by promoting unfolding of misfolded conformations, while stabilizing the correctly folded conformer.[139,140] Use of the more concentrated yeast cell extracts, compared to the very dilute HeLa extract, has a greater effect on promoting the docked state (Fig. 4.6). The concentrations of cell extract used are kept much more dilute than what is found inside cells, however, in order to directly probe the effect of interactions with cellular components without contribution from macromolecular crowding. We did attempt to probe crowded conditions in the presence of extract by mixing 20% PEG with yeast cell extract. However, the solution turned cloudy upon mixing of PEG with extract and could not be used. Even cell extract with 10% and 5% PEG turned slightly cloudy and could not produce usable smFRET traces. These observations are likely due to the precipitation of proteins,[244,245] DNA[246] and long RNAs[247] by PEG. In fact, a propensity of PEG to cause precipitation of cellular extract components may be linked to two observations

in the condition containing yeast extract with 3% PEG. In this condition, the proportion of the docked population is less than in the condition containing yeast extract without PEG (Figs. 4.6b,c). Precipitation of cellular components may explain why there is not the expected greater proportion of the docked state in the presence of both PEG and extract compared to in the presence of 14 mg/ml of extract alone since PEG and extract each, at least separately, promote the docked state. Secondly, the mean value of the low FRET state is much higher in the condition containing yeast extract with 3% PEG than in all other conditions (Fig. 4.6b). The higher FRET value may reflect a more compact undocked state, a less compact substrate-free state or a different conformation altogether that is not present in other conditions. There thus may be some precipitation-related effects, even at 3% PEG, that may be influencing the ribozyme's FRET distribution. Finally, because yeast and HeLa cells are quite different from each other, there likely are differences in the kinds of interactions that occur in each extract but these differences are not distinguishable in the FRET distribution histograms or the ratios of static to dynamic molecules. Future experiments may look at determining the identity of the components from yeast and HeLa extracts that interact with the ribozyme by doing mass spectrometry analyses on supershifted bands cut out from nondenaturing gels of the ribozyme incubated with cell extract. Furthermore, experiments where known RNA binding components, such as the yeast membrane-associating RNA binding protein She2p,[248] the yeast stress granule RNA binding protein Whi3,[249] or even the HIV-1 nucleocapsid protein NCp7,[250] can be added instead of cell extract to determine whether such components alone can reproduce results seen from using extract.

In summary, we have parsed apart three defining features of the intracellular environment and interrogated their effects on the FRET distribution and dynamics of the hairpin ribozyme. These three features are the intracellular buffer environment without crowding or cell components, molecular crowding with little surface inter-

actions and cellular components that do nonspecifically interact with the RNA. The work in this chapter has gone beyond looking at just macromolecular crowding, as many studies have done before,[137–146,217,231,238–240] and has provided insight into the contribution of each of these features as parts of a whole. The intracellular environment is very complex. We have been able to study the hairpin ribozyme in an *in vitro* buffer for many years already.[25,66,69,71,93] Now, we can aim to understand how RNA behaves in a cell-like environment, which would not only be useful from a molecular biology point of view, but may also provide relevant considerations for ribozyme-based therapeutics.

## 4.5 Acknowledgements

Both Dr. May Daher and Matthew Kahlscheuer are acknowledged for their generous contributions of yeast extract. Additionally, Dr. Sethuramasundaram Pitchiaya and Dr. Vishalakshi Krishnan are also acknowledged for their insights regarding the use of HeLa cell extracts.

# CHAPTER V

# Summary, Conclusions and Future Directions

## 5.1 Summary and Conclusions

The variety of crucial roles RNA has in biology makes it a molecule worth studying in depth. RNA was once considered a molecule that supported protein synthesis, being either an intermediary in the flow of genetic information (mRNA) or structural components (tRNA and rRNA).[3–8] But with the discovery of ribozymes, it was realized that RNA could do so much more than originally thought.[1,2] In fact, ribozymes were found to be responsible for several life-essential processes. For instance, gene splicing,[13–15] peptide bond formation,[12] tRNA maturation[2] and sub-viral genome replication[11,57–59] are some examples of processes catalyzed by ribozymes. Bioinformatic searches have also suggested that there are many more ribozymes to be discovered in the genomes of several organisms including humans, fish, plants and insects.[23,24] Artificial ribozymes have even been engineered to perform chemical reactions, like the Diels-Alder reaction,[95] and therapeutic functions, such as for treatment of HIV.[96] Clearly, the importance of RNA in biology and its potential to be used in new contexts have been demonstrated. The appreciation we now have for RNA makes RNA research a top scientific priority.

In order to learn more about catalytic RNA, the work in this thesis focused on studies of the HDV and hairpin ribozymes. Originally isolated from the human

satellite virus HDV (hepatitis delta virus), the HDV ribozyme has a crucial role in HDV genome replication.[11] The HDV ribozyme has a minimal motif of ~85 nucleotides,[21,32,33] displays a double-pseudoknot secondary structure,[11,34,37] requires a divalent cation for cleavage[50,51] and has a deeply electronegative, solvent-exposed active site.[50,51] In contrast, the hairpin ribozyme was originally isolated from the satellite RNA of the tobacco ringspot virus, where it is a key component in satellite RNA replication.[10,56–59] The hairpin ribozyme has a minimal motif of ~50 nucleotides,[31,66,67] consists of two loop-containing helices joined together by a helical junction,[31,66,67] does not require metal ions for catalysis[78–80] and has a deeply electronegative active site secluded from bulk solvent.[10,74,75] The relatively small sizes and simple, yet commonly-found, structures[20,21,25–28] make the HDV and hairpin ribozymes excellent model systems for studying catalytic RNA. The knowledge learned from studying the HDV and hairpin ribozymes is directly applicable to other ribozymes as well.

The overall goal of this thesis was to gain insight into the structural dynamics of the HDV and hairpin ribozymes. Global dynamics and local dynamics, as well as their interdependency, are important for understanding the functions of molecules.[98,99] Scientists have learned key aspects about various biological processes from looking at dynamics.[102–109] I used ensemble FRET and smFRET techniques to probe the global dynamics, or the millisecond and slower sub-domain motions,[98,99] of the HDV and hairpin ribozymes. MD simulations were used to study ribozyme local dynamics, which are the atomic-level vibrations and fluctuations localized to a few groups of atoms occurring on the femto- to nanosecond timescale.[98] In the following sections, I summarize the key take-away points from each study and how they fit into the overall focus of this thesis.

**Heterogeneity and a rugged landscape along the catalytic pathway for the HDV ribozyme**

Both the global and local dynamics of the most recently crystallized *trans*-acting HDV ribozyme[53] were interrogated in **Chapter II**. This HDV ribozyme crystal structure has a product-like active site, prompting further studies on its dynamic behavior. From our solution FRET studies, done at physiological pH, we show that upon conversion from the precursor to the product state, the HDV ribozyme undergoes a lengthening of ~8 Å in the end-to-end distance along the P1-P4 axis. This is comparable to previous observations for *trans*-acting HDV ribozymes.[37,41,54,55] Interestingly, kinetic assays and electrophoretic mobility shift assays (EMSAs) both indicate heterogeneity in the precursor population that was removed upon crystallization at low pH. Nevertheless, the precursor population can largely convert to the product ribozyme, indicating the importance of structural dynamics in converting inactive to active conformations given the heterogeneity of the precursor population. Analysis of 1.8 $\mu$s total of MD simulations resulted in a number of insights into the active site structure and dynamics of the HDV ribozyme. We find that the hammerhead ribozyme active site, which was modeled into the HDV ribozyme active site to compensate for the lack of electron density for the U-1 and scissile phosphate, results in a catalytically fit active site. A protonated C75 also results in a more catalytically fit active site for this version of the HDV ribozyme. Removal of several 2′-oxygens in the active site compromises the catalytic in-line fitness and so does the presence of crystal contacts, which directly interacts with active site residues. Furthermore, a *cis*-acting version of the ribozyme exhibits a more dynamic active site, while a U-1G mutation causes poor fitness, suggesting molecular rationalizations to biochemically observed changes in catalytic activity. Finally, integrating both solution and computational data, we propose a unifying model suggesting a rugged landscape for the folding of the HDV ribozyme along the catalytic pathway. The use of the catalytic in-line fitness pa-

rameter originally discussed by Soukup et al.[175] was also demonstrated as a way to link computational simulations on the nanosecond timescale to kinetically relevant timescales that are much slower.

### Internal waters in the hairpin ribozyme are not easily disrupted

MD simulations were used in **Chapter III** to identify modifications in the hairpin ribozyme active site that can disrupt the active site water chain. Crystal structures and MD simulations show a string of 4 to 6 long-residing internal waters within the otherwise solvent-excluded core of the hairpin ribozyme, but the function of this water chain remains uncertain.[31,84,86,92,198] Identification of a ribozyme variant with a disrupted active site water chain would be very useful for comparing to the wild type ribozyme in the endeavor to determine the role of the water chain. We find that the water chain is quite robust and difficult to disrupt. Single-atom modifications aimed at disrupting hydrogen bond interactions between active site residues and the water chain are not able to cause significant disruption of the water chain. Only one ribozyme variant, with both N1dA9 and 4SU42 modifications, most consistently shows disruption of the water chain. Therefore, we establish that successful water chain disruption requires the simultaneous interference of RNA-water interactions at two separate locations in the active site. These results also illustrate the application of MD simulations for predicting interesting variants for experimental testing. Another two-modification variant bearing N1dA9 and N1dA10 shows disruption of the water chain in addition to perturbation of the active site but, curiously, only when A38 is protonated. However, significant structural disruption of the active site is not seen for the N1dA9_4SU42 variants, suggesting a model where the water chain is not acting as a structural glue that holds the active site together. Additionally, the N1dA9_N1dA10 variant with unprotonated A38 barely shows an effect on the water chain compared to wild type simulations. Simulations of protonated and unprotonated A38 were

performed because it remains uncertain which state is present in the active hairpin ribozyme.[67,74,75,83,86–88,90,195–197] Interestingly, in our MD simulations, unprotonated A38 promotes better in-line fitness but protonated A38 promotes more stable docking hydrogen bonding. These results lend support to the hypothesis of a bifunctional role for A38,[86] where both the protonated and unprotonated states are important for ribozyme function.

**Cellular components potently promote the docked hairpin ribozyme**

In **Chapter IV**, smFRET was used to study the docking and undocking conformational dynamics of the hairpin ribozyme in an intracellular-like environment. The intracellular environment is very different from the *in vitro* environment used in traditional laboratory experiments,[119,125–129] thus leading to the hypothesis that global structures and dynamics of ribozymes are also different in these two environments. With an established body of *in vitro* data on FRET distributions as well as docking and undocking dynamics,[66,69,71,93] the hairpin ribozyme is an excellent candidate for exploring the contribution of various intracellular features on global RNA structure and dynamics. To determine the contribution of the intracellular environment excluding any macromolecular crowding or interactions with cellular components, we compared results using standard Tris buffer to those using a buffer that mimics the intracellular environment. We find that the intracellular-mimic buffer favors the undocked state primarily due to competition between $Mg^{2+}$ and $K^{+}$ for electrostatic binding to the ribozyme.[25,77,235–237] The intracellular-mimic buffer also favors less dynamic molecules compared to the standard Tris buffer. Next, the contribution by macromolecular crowding was assessed by addition of 20% PEG-8000 to both intracellular-mimic and Tris buffer conditions. PEG favors the docked state, which is consistent with previous studies discussing the excluded volume effect of crowders in promoting a more compact RNA structure.[137,143,144] PEG also promotes more dy-

namic molecules in intracellular-like conditions but more static molecules in *in vitro* buffer conditions. Finally, the contribution of interactions between the ribozyme and cellular components was probed by incorporating HeLa cytosolic extract and yeast whole cell extract in the smFRET experiments. The interactions between the ribozyme and cellular components ultimately cause favoring of the docked state with a more pronounced effect than PEG alone. In addition, both HeLa and yeast extracts promote a greater proportion of dynamic traces seen. We therefore propose a model where interactions between cellular components and the ribozyme are destabilizing misfolded conformations and promote the active conformer, such as have been proposed for interactions between the hammerhead ribozyme and PEG.[139,140]

## 5.2 Future Directions

Research on ribozymes in 2014 has come a long way from the initial discovery of catalytic RNA ~30 years ago.[1,2] This thesis exemplifies the two current trends in our continued exploration of ribozymes. The first trend is the integration of computational and experimental work. MD simulations can provide atomic insights that suggest rationalizations for experimental results, which was especially depicted by the work in Chapter II. The femto- to nanosecond dynamics probed by MD simulations also complement the data obtained from experimental techniques, like smFRET, that probe slower (millisecond) dynamics. Use of parameters like in-line fitness[175] provide chemical relevance, such as the propensity for catalytic activity, to observations in MD simulations. MD simulations can also be used as a tool to first screen for interesting variants of a molecule that can then be experimentally tested, as was done in Chapter III. A combination of both computational and experimental techniques thereby enables us to do a more comprehensive studies than we would be able to otherwise. The second trend is the interest in understanding ribozymes in their native, more complicated cellular environments. The cellular environment is complex and

in our endeavor to understand the behavior of each cellular component, we typically isolate components and study them in simpler environments. But, eventually, we do need to put these components back into the cell to acquire a comprehensive, big-picture understanding of how the biological system works. An understanding of the interactions between the individual components, after all, are as equally important as an understanding of the isolated component itself. Taking what can be called a "post-reductionist" approach,[251] the work done in Chapter IV paves the way toward understanding the behavior of ribozymes in the more complex, but biologically relevant, cellular environment. The next few sections will outline some of the future directions for the studies in this thesis as they relate to the two major trends in the future of ribozyme research.

**Exploring the rugged terrain of the heterogeneous HDV ribozyme precursor population**

Following the lead of Chapter II, a new set of experimental and computational tools can be employed to further interrogate the heterogeneity in the precursor population of the most recently crystallized *trans*-acting HDV ribozyme.[53] We can employ smFRET to get an idea of the predominant FRET states present in this heterogeneous population as well as to determine the kinetics of transitions between these states. Optimization may be needed to determine the best placement of the donor and acceptor fluorophores that would best capture the structural differences present in the heterogeneity of the precursor population. Based on data from the smFRET studies, we can then use enhanced sampling computational techniques, such as replica exchange,[252,253] to model the structures for the FRET states and the transitions between these structures.[254] Additionally, classical all-atom MD simulations and principal component analysis can be used to construct a free-energy landscape for the HDV ribozyme precursor population as described by a 2009 study by Riccardi et

al.[255] Based on the ability for the heterogeneous precursor population to largely convert to the product state, as seen in our EMSAs, the rate constants for transitions between many of the heterogeneous conformations should be relatively fast compared to the rate constant for the chemical catalytic step. Several other studies have shown that heterogeneity is a common feature of RNA folding.[66,189,210–216] An in-depth characterization of the heterogeneity in the precursor HDV ribozyme can be valuable for better understanding the RNA folding landscape.

**Experimental studies with the N1dA9_4SU42 hairpin ribozyme variant**

With the computational identification of the N1dA9_4SU42 variant as a likely candidate in having a disrupted active site water chain, future experimental work can be done to further characterize this variant as it compares to the wild type hairpin ribozyme. Chapter III focused solely on MD simulations, which served as a springboard for future experimental work to tease out the functional role of the water chain. Crystallization of the N1dA9_4SU42 variant may be used to verify the disruption of the water chain observed in MD simulations. The individual rate constants for the known kinetic steps of the hairpin ribozyme (including $k_{ligated,dock}$, $k_{ligated,undock}$, $k_{cleave}$, $k_{ligate}$, $k_{cleaved,dock}$, $k_{cleaved,undock}$) can be determined using a single molecule kinetic fingerprinting protocol previously published by Liu et al.[69] If the water chain does have a functional role, we predict that a disruption in the water chain would be reflected in at least one of these rate constants when compared to those of the wild type ribozyme. The results can also be compared to those from N1dA9 or 4SU42 single-atom modified variants to confirm that the effects are due specifically to the disrupted water chain. Additionally, based on our hypothesis that the water chain is not acting predominantly as a structural glue that keeps the active site from falling apart, we predict that the N1dA9_4SU42 variant can still dock and form the active site. Therefore, future smFRET experiments on the N1dA9_4SU42

variant are predicted to show the occurrence of a high FRET state as seen in previous smFRET experiments.[25,66,69,71] However, because U42(O4) is involved in docking, the docked state of the N1dA9_4SU42 variant may not be as stable as in the wild type, which may be reflected by a greater proportion of the low FRET, or undocked, state from the FRET distributions generated by smFRET experiments. Together with the MD studies, these experiments can provide further insight into the role of the hairpin ribozyme active site water chain.

**Delving deeper into exploring the behavior of the hairpin ribozyme in a cellular environment**

There are three prominent directions we can take to follow up the work in Chapter IV, which spearheaded a parsing apart of the impact of various features of the cellular environment on the structure and dynamics of the hairpin ribozyme. The first direction that can be explored is characterization of the components that interact with the ribozyme. We can incubate the ribozyme with cell extract and run the mixtures out on a nondenaturing gel, as has been done in Chapter IV using HeLa cytosolic extract. Mass spectrometry analyses can then be performed on bands supershifted relative to where the ribozyme runs in conditions without the presence of extract to determine the identity of the components that interact with the ribozyme. Additionally, smFRET experiments with various fractions of yeast extract, instead of using whole cell extract, can be performed to narrow down the components responsible for the potent favoring of the docked state and the dynamic molecule traces in the presence of extract. It is currently uncertain whether these components are interacting with the ribozyme specifically or nonspecifically. Therefore, an experimental approach that can be performed to distinguish between these types of interactions may include a series of smFRET experiments utilizing cell extract at various concentrations. We predict that, if the interactions are primarily nonspecific, the degree to

which the docked state is favored would depend on the extract concentration whereas if the interactions are primarily specific, there would be an extract concentration after which there would not be an effect on the FRET distribution. In conjunction with varying the cell extract concentration, the hairpin ribozyme can be incubated separately with a variety of known RNA binding components, including the yeast membrane-associating RNA binding protein She2p,[248] the yeast stress granule RNA binding protein Whi3,[249] or even the HIV-1 nucleocapsid protein NCp7,[250] instead of with cell extract to assess whether interactions with each of these RNA binding components produces the results seen from using cell extracts. If the interactions with the extract cellular components are primarily nonspecific, we predict that all the results from using a variety of such RNA binding components would be similar. However, it would be unlikely that all RNA binding components would produce similar results if a particular component in cell extract is responsible for the observed results from using cell extract.

The second direction we can explore is extending the types of crowders and extracts used. Inside the cell, there are many different kinds of molecules that contribute toward molecular crowding.[119–124] PEG-8000 is but one molecule that can be used to emulate crowding. Other classes of crowders, such as Ficoll, dextran, other sugars, proteins and amino acids, can be tried. It is known that crowding effects can depend on the size and electrostatic properties of the crowder.[129,140,140,145,146] Therefore, we predict that not all crowders may favor the docked state of the hairpin ribozyme. With the use of small organic osmolytes, for instance, based on previous crowding studies,[145] we predict that there would be destabilization of the secondary structure of the hairpin ribozyme although the effect on the tertiary structure is less predictable. Use of a UV thermal melting analysis in conjunction with solvent accessible surface area calculations, such as used by Lambert and Draper,[145] would give an indication of the effects of small osmolytes, like urea and amino acids, on the hairpin ribozyme

secondary and tertiary structures. Furthermore, we can extend our smFRET experiments towards using other kinds of cell extracts. Both HeLa and yeast extracts were used in Chapter IV and both extracts favored the docked state and promoted a greater proportion of dynamic molecules. Also, the hairpin ribozyme is a plant virus pathogen,[56] so one could hypothesize that the docked (active) state would also be favored in plant extract, but smFRET experiments to verify this hypothesis have yet to be performed. However, it may not be the case that all cell extracts also favor the docked state of the hairpin ribozyme or promote more dynamic molecules. The identity of components in other extracts that interact with the ribozyme can also be determined and, from these data, we may be able to piece together common trends about the types of components that interact with the hairpin ribozyme and their resulting effects on ribozyme structural dynamics. With the array of protocols for preparing various extracts,[256–260] future studies can look at the effect of different extracts on the structural distribution and dynamics of the hairpin ribozyme.

The third direction is to study the hairpin ribozyme dynamics inside intact, living cells. As useful as extracts are, they do not represent exactly the same environment as that inside a cell.[261] We ultimately want to know what the behavior of the hairpin ribozyme is inside a living cell because that, and not an emulated cellular environment, is its native environment. In fact, I have outlined, in Appendix A, a protocol and some preliminary data for a technique that may allow for smFRET experiments inside mammalian cells. There is still much work to be done to get this protocol working, but if successful, this protocol can then be expanded to incorporate use of different cell lines and/or different ribozymes. The ability to study the structural dynamics of ribozymes inside living cells and compare the results with those from *in vitro* conditions would lead to three major gains. First, we would better understand how ribozymes behave in their native environment. Second, we would gain insight into the effect of being inside a living cell on RNA dynamics. And, finally, such as technical

protocol would pave the way to new applications for screening and testing ribozyme therapeutics, which are ultimately designed to function inside living cells.

### Beyond ribozymes and the modern RNA world

The work in this thesis is applicable to more than just ribozymes; it is applicable to other non-protein coding RNA (ncRNA) as well. Beyond ribozymes, ncRNA has important roles in processes ranging from stress response,[262–264] to regulation of gene expression by RNA interference,[265] to organization of the nuclear architecture.[266] To better understand how these other ncRNAs function at the molecular level, we can apply knowledge about structural heterogeneity in RNA populations, impact of local dynamics within the active site on overall structure and function, and the contribution of various features of the intracellular environment on RNA structural dynamics. But perhaps more importantly, we can apply the knowledge gained from studies of model systems like the HDV and hairpin ribozymes to RNA therapeutics. As interest grows in the area of RNA therapeutics,[96,267–272] it becomes increasingly relevant to leverage the insights from simpler ribozyme model systems in order to more efficiently design, screen and perform preliminary tests of new therapeutics.

In a 2001 review, Sean Eddy wrote about the "modern RNA world".[18] The notion behind the modern RNA world is that the roles RNA has are optimal for RNA, and that proteins would not necessarily be more efficient.[18] Unlike in the original RNA world hypothesis,[16] RNA is no longer seen as merely a prebiotic molecule that "preran" biochemical reactions before proteins took over. Unlike in the central dogma laid out by Francis Crick,[3,4] we now know that RNA does a lot more than be the intermediary between DNA genes and proteins. Both proteins and RNA carry out essential processes in the cell and each type of molecule has evolved to excel at its function. With the modern understanding of molecular biology, the limelight falls on both proteins as well as RNA.

# APPENDICES

# APPENDIX A

# Paving the Way to Intracellular Single Molecule FRET with the Hairpin Ribozyme[1]

## A.1 Introduction

The intracellular environment is a unique environment not easily replicated by *in vitro* conditions. The defining characteristics of the intracellular environment are that there is macromolecular crowding, high concentrations of monovalents (specifically $K^+$), a reducing environment and the potential for specific and nonspecific interactions with many other molecules.[119,125–129] Even cell extract, which was used in Chapter IV, cannot accurately represent the environment inside a living cell because, at the very least, it is challenging to reach macromolecule concentrations as high as those inside the cell.[261] Therefore, in order to really study ribozyme behavior inside living cells, we must do just that - study ribozymes inside living cells.

Single molecule fluorescence resonance energy transfer (smFRET) is an extremely informative technique for studying global structural dynamics.[102,243,273] One can learn

[1]Data for the SDS-PAGE of HeLa cell extracts (Fig. A.5) were collected solely by Hanna Wagner. Data for Fig. A.4 were collected by both Hanna Wagner and Wendy Tay. Figs. A.4 and A.5 were adapted from figures made by Hanna Wagner. Experiments with U2OS cells (Fig. A.6) were done together with Dr. Laurie Heinicke. All other work and analyses done by Wendy Tay.

about the relative magnitude of the conformational changes from the changes in FRET states, the distributions of the conformations in an ensemble population and the kinetics of the conformational changes.[274] Ideally, in order to do smFRET experiments, the molecule of interest must be adhered to a surface so that it is within range to be illuminated by the evanescent wave generated upon total internal reflection of the incident laser light.[243,273] To apply smFRET to living cells, the simplest solution would be to localize the molecule of interest to the basal plasma membrane. This is because cultured cells can grow in a single layer adhered to a glass bottom plate. Therefore, at least some of the molecules of interest would be localized for some time to the plasma membrane adhered to the glass bottom and it would be these molecules that would be illuminated by the evanescent wave while inside a living cell. The question really then boils down to this: How does one localize a ribozyme to the inner leaflet of the plasma membrane such that two-dimensional diffusion is limited?

There have been few reported studies of smFRET inside living cells.[275,276] One study, by Murakoshi et al. in 2004, looked at the activation of the small G protein Ras inside KB cells (human epidermoid mouth carcinoma).[275] Murakoshi et al. found that upon Ras activation, Ras diffusion was greatly slowed, likely due to the formation of large signalling complexes.[275] The H- and K-ras proteins used in the study localize to the inner leaflet of the plasma membrane, establishing that it was possible to do intracellular smFRET on molecules if they were associated with the plasma membrane.[275] In 2010, Sakon and Weninger used intracellular smFRET to study the various assembled complexes of recombinant soluble N-ethylmaleimidesensitive factor attachment protein receptor (SNARE) proteins in BS-C-1 cells (kidney epithelial cells of the African green monkey).[276] SNARE proteins also natively associate with the membrane, as they are involved in membrane fusion, reinforcing that plasma membrane localization is a viable method to adhere molecules for smFRET.[276] Like in the Murakoshi et al. study, the smFRET traces collected by Sakon and Weninger

were much shorter than what has been collected for *in vitro* conditions, making it difficult to determine the kinetic rate constants of conformational changes.[66,69,71,274–276] Due to the challenging nature of intracellular smFRET, which includes considerations such as autofluorescence and fast photobleaching,[276] there are not many studies employing this technique and no known studies focused on RNA. An intracellular FRET protocol to study RNA would be a great advance for this field as another single molecule fluorescence tool for studying intracellular RNAs.[277]

To this end, we have developed a novel protocol for localizing the hairpin ribozyme to the inner leaflet of the plasma membrane in mammalian cells (Fig. A.1). We used the hairpin ribozyme as our model ribozyme because it exhibits well-defined conformational changes[66,69,71,93] but the protocol could eventually be applicable to a wide variety of RNA molecules. Our protocol involves adaptation of the SNAP-CaaX system formerly available from New England Biolabs (the product was discontinued 2011). The SNAP-tag protein (20 kDa) is a mutant version of the human $O^6$-alkylguanine-DNA alkyltransferase.[278] The native $O^6$-alkylguanine-DNA alkyltransferase enzyme plays a role in DNA repair.[278] The SNAP-tag specifically reacts to form a covalent bond with the suicide inhibitor $O^6$-benzylguanine (BG).[278] $O^6$-benzylguanine can be incorporated into a number of molecules, such as fluorescent probes, giving rise to a whole range of benzylguanine derivatives that are specific substrates for the SNAP-tag.[279] With the SNAP-CaaX protein, the SNAP-tag has been modified to contain the C-terminal sequence from the human H-ras protein, a protein involved in signal transduction.[280] This C-terminal sequence contains what is known as a CaaX box, where CaaX refers to cysteine ("C") to which the post-translational prenyl modification is attached, followed by two aliphatic residues ("a"), followed by any amino acid ("X"), the identity of which determines the type of prenyl modification added.[281–283] In the case of SNAP-CaaX, a farnesyl group is added to the cysteine in the CaaX box and palmitoyl groups are also added to two cysteines upstream of the H-ras CaaX

box, which in turn cause SNAP-CaaX to localize specifically to the inner leaflet of the mammalian plasma membrane.[281] Although the SNAP-CaaX protein was initially intended to be a control system for other SNAP-tag experiments, with the availability of BG maleimide that can be used to label thiol-modified RNA, the potential for applying this system to RNA became apparent.

## A.2 Materials and Methods

### The hairpin ribozyme

The hairpin ribozyme used for these experiments is shown in Fig. A.1. The ribozyme was ordered from Keck Oligo Synthesis Facility (Yale University, New Haven, CT) with a number of modifications: 5′ Cy5, an internal deoxyribothymidine with an amino C6 modifier, five 2′O-methyl ribonucleotides at the 3′ end and a 3′ thiol C6 modifier. This hairpin ribozyme was also designed with an A8G mutation to prevent cleavage during the intracellular smFRET experiments. The oligo was first deprotected using the protocol for TOM deprotection on the manufacturer's website (http://medicine.yale.edu/keck/oligo/services/protocols/RNA.aspx). Following deprotection, the oligo was ethanol precipitated and re-suspended in double-distilled water. Intact oligo was then purified from an 8% 19:1 acrylamide:bisacrylamide denaturing gel containing 8 M urea. The purified hairpin oligo was stored at -20°C in double-distilled water.

### Replacing the H-ras C-terminal sequence with that of K-ras

The SNAP-CaaX plasmid, pSNAP-CaaX, was generously given to us for free from New England Biolabs. As discussed in the introduction, SNAP-CaaX contains the C-terminal sequence from the human H-ras protein. However, we were also interested in creating a SNAP-tag with the K-ras C-terminal sequence in place of the H-ras C-

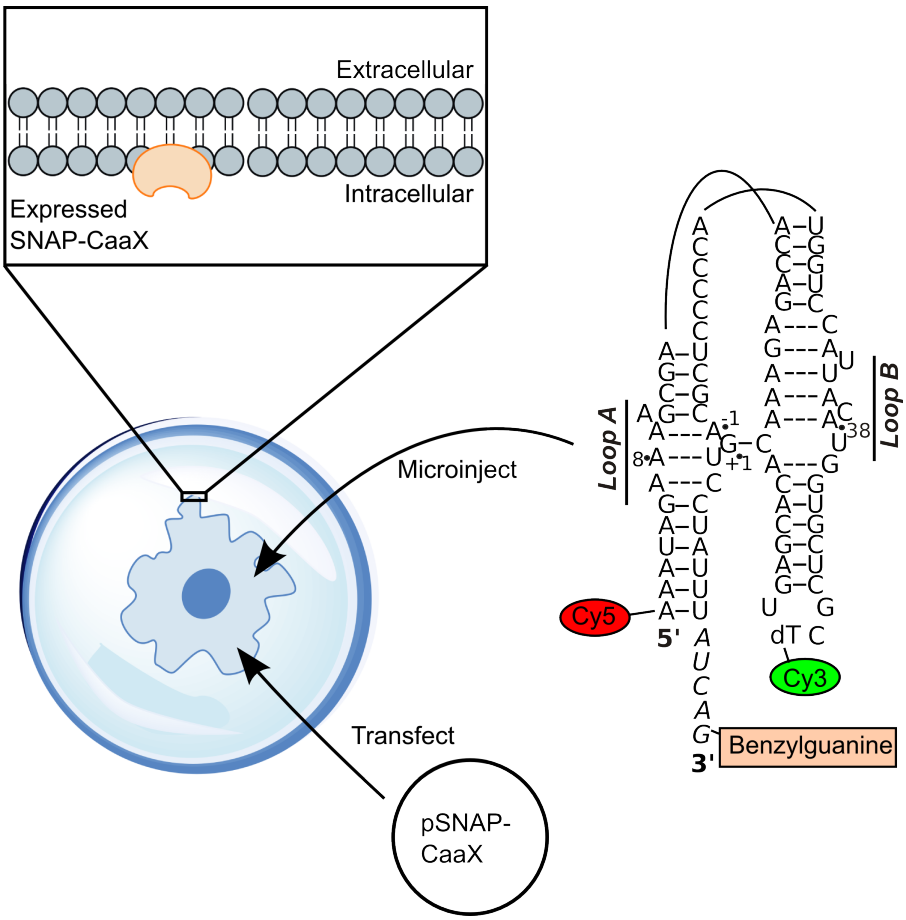


**Figure A.1: Strategy for localizing the hairpin ribozyme to the inner leaflet of mammalian plasma membranes.** An overview of the strategy using SNAP-CaaX to localize the hairpin ribozyme to the plasma membrane is depicted. The pSNAP-CaaX plasmid is first transfected into mammalian cells, where the SNAP-CaaX protein is expressed, post-translationally modified and localized to the membrane. The hairpin ribozyme containing a benzylguanine moiety is then microinjected into the cell where it should bind to the SNAP-CaaX protein. The hairpin ribozyme is also labeled with Cy3 (FRET donor) and Cy5 (FRET acceptor) fluorophores.

terminal sequence because K-ras more consistently localized to the plasma membrane compared to H-ras (Fig. A.2a).[281,284] The K-ras C-terminal sequence was encoded in a double-stranded DNA oligomer ordered from IDT. This oligomer, referred to as the K-ras C-terminal insert, has the following sequence (given for just one of the complementary DNA strands): 5′-CGA CCT GCA GTA AAA ATG AGC AAG GAC GGC AAG AAG AAG AAG AAG AAG AAG AGC AAG ACC AAG TGC GTG ATC ATG TGA GGA TCC GGC C-3′ (underlined residues indicate PstI and BamHI restriction sites respectively). PstI and a BamHI sites were encoded in the K-ras C-terminal insert because the H-ras C-terminal sequence in pSNAP-CaaX is flanked by these restriction enzyme sites. Therefore, we used PstI and BamHI to cut out the H-ras C-terminal sequence and replace it with the K-ras C-terminal sequence, which also contained sticky ends generated by PstI and BamHI. Both the pSNAP-CaaX and K-ras C-terminal insert were digested with PstI and BamHI restriction enzymes for 1 h at 37°C following the manufacturer's protocol (New England Biolabs). To inactivate the restriction enzymes, the mixture containing pSNAP-CaaX was purified using QIAquick PCR purification kit (Qiagen) and the K-ras C-terminal insert was subjected to phenol-chloroform extraction followed by ethanol precipitation. Digested pSNAP-CaaX was treated with antarctic phosphatase as described by the manufacturer's protocol (New England Biolabs). Ligation of digested pSNAP-CaaX and K-ras C-terminal insert was accomplished using T4 DNA ligase following the manufacturer's protocol (New England Biolabs). The ligation reaction was transformed into XL10-Gold Ultracompetent cells (Agilent Technologies) with ampicillin selective pressure. Ten colonies from the transformation were mini-prepped and sequenced to verify the replacement of the H-ras C-terminal sequence with that of K-ras. For the rest of this appendix, SNAP-CaaX-Hras and SNAP-CaaX-Kras will refer to the SNAP-tag with the H-ras and K-ras CaaX sequences respectively.

(a)

| C-terminal Tail | |
|---|---|
| H-ras Tail | GCMSCK***CVLS*** |
| K-ras Tail | VKMSKDGKKKKKKSKTK***CVIM*** |

(b)

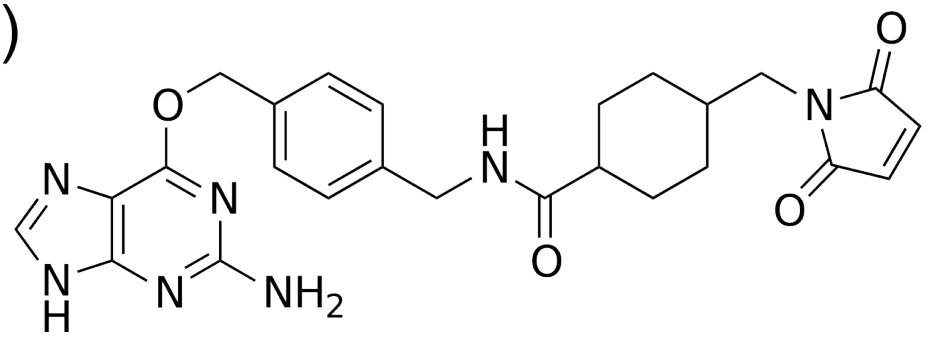


**Figure A.2: Ras C-terminal sequences and the chemical structure of benzylguanine-maleimide.** **(a)** C-terminal sequences for both H-ras and K-ras. The CaaX sequences are highlighted in red with the italicized C indicating the amino acid that modified with the prenyl group. The polybasic region of K-ras is highlighted in blue. **(b)** The structure of benzylguanine-maleimide that can react with thiol groups.

### Cell culture

We worked with two cell lines derived from human carcinomas: HeLa (cervical cancer) and U2OS (osteosarcoma).[285,286] Both of these cells are optimally shaped for single molecule fluorescence studies and for microinjection. Cells are maintained at 37°C with 5% CO and 95% relative humidity. The media used for HeLa cells is Dulbecco's Modified Eagle Medium (Life Technologies) supplemented with 10% fetal bovine serum (FBS) (Life Technologies) and 1X non-essential amino acids (Life Technologies) and 100 U/ml of penicillin and 0.1 mg/ml of streptomycin. The media used for U2OS cells is McCoy's 5A basal media (Life Technologies) supplemented with 10% FBS and 100 U/ml of penicillin and 0.1 mg/ml of streptomycin.

### Transient expression of SNAP-CaaX-Hras and SNAP-CaaX-Kras in HeLa and U2OS cells

Cells to be transiently transfected with pSNAP-CaaX-Hras or pSNAP-CaaX-Kras were plated onto a DeltaT dish (~100,000 cells/dish) one day before transfection. Each dish of cells was then transfected with 1 $\mu$g of pSNAP-CaaX-Hras or pSNAP-CaaX-Kras using FuGENE HD (Promega) following the manufacturer's protocol. Cells were then incubated with the transfection mixture for 24 h, after which, the media was replaced. Forty-eight hours after transfection, cells were imaged with SNAP-Cell Oregon Green dye (New England Biolabs, see below).

### Stable expression of SNAP-CaaX-HRas and SNAP-CaaX-KRas in U2OS Cells

To create monoclonal stably-transfected U2OS cells expressing either SNAP-CaaX-Hras or SNAP-CaaX-Kras, we used a protocol based on that from Kedersha et al.[287] We first plated approximately 150,000 U2OS cells per well in a 6-well plate. About 24 h after plating, cells were ready for transfection. In a separate tube (note: amounts

given are for two wells and can be scaled as needed), 1.5 $\mu$g of either pSNAP-CaaX-Hras or pSNAP-CaaX-Kras was mixed with 5.25 $\mu$l of Fugene HD (Promega) and 68 $\mu$l Opti-MEM reduced serum media (Life Technologies). The DNA mixture was incubated for 15 min at room temperature. Meanwhile, media from the cells were removed and replaced with 2.075 ml of Opti-MEM containing 10% fetal bovine serum. After the 15 min incubation, 25 $\mu$l of the DNA mixture was added to each well and mixed. The final volume of each well was 2.1 ml. The cells were then incubated at 37°C overnight. The antibiotic selection marker, G418 (Life Technologies), was added to a final concentration of 0.5 mg/ml.(The appropriate concentration of G418 was determined from a kill curve.) Twenty-four hours after adding selection marker, the cells in a well were replated to a T75 flask. Media containing G418 was replaced every 1 to 2 days until cells reached 75% confluency (about 1 week). Cells were also trypsinized and replated frequently to avoid formation of "colonies".

To get a monoclonal population of stably-transfected cells, a serial dilution technique based on a protocol from Corning was used.[288] Serial dilutions were set up for three 96-well plates. Single colonies from the 96-well plates were transferred to a 24-well plate, then a 12-well plate then two 6-well plates. Each well of cells was grown to ~95% confluency and then frozen in vials and stored in liquid nitrogen. A portion of the cells from each well was also replated and maintained to test for SNAP-CaaX-Hras or SNAP-CaaX-Kras expression.

**Live cell imaging and microinjection**

For cells that were imaged using the SNAP-Cell Oregon Green dye, the dye first had to be resuspended in DMSO. One tube of SNAP-Cell Oregon Green was dissolved in 25 $\mu$l of DMSO, of which 2.5 $\mu$l was added to 500 $\mu$l McCoy's 5A basal media (with 10% FBS and 100 U/ml penicillin and 0.1 mg/ml streptomycin). This media was incubated with cells ready for imaging for 30 min at 37°C. After incubation, cells were

washed three times with media, then incubated for another 30 min at 37°C in media without SNAP-Cell Oregon Green (New England Biolabs). Cells were then washed three times and imaged with phenol-red free balanced salt solution (BSS) containing reduced FBS.[289] For cells that were to be microinjected with hairpin ribozyme, they were simply washed three times and imaged in BSS.

Cells were imaged using an Olympus IX81 microscope with either a 20X objective or 60X 1.45 NA oil-immersion objective (Olympus). An additional 1.6X magnification (internal to the microscope) was used for all images. Cells in the DeltaT dishes were kept at 37°C during imaging usng a DeltaT open dish system and heated lid (Bioptechs). Images were acquired using an EVOLVE EM-CCD camera (Photometrics) at 100 ms exposure time with a pixel area of 512 x 512. Sample excitation was done using solid state lasers with wavelengths at 405 nm, 488 nm, 532 nm and 640 nm. The lasers were directed through an acousto-optical tunable filter (AOTF), which enabled the user to select the desired excitation wavelength. All laser lines were filtered with a 10 nm bandwidth filter. The power at the objective was 0.8 mW, 1.2 mW, 7 mW and 8 mW for the 405 nm, 488 nm, 532 nm and 640 nm lasers respectively. SNAP-Cell Oregon Green and Cy5 were excited by the 488 nm and 640 nm lasers respectively and their emission was detected by a dual-band filter cube containing a z491/639rpc dichoric filter (Chroma). Imaging was controlled using the software program MetaMorph (Molecular Devices).

A Femtojet pump (Eppendorf) and Injectman NI2 micromanipulator (Eppendorf) mounted next to the stage were used to do the microinjections. The tips used for microinjection were the Femtotips II (Eppendorf), which have an inner diameter of 500 nm and were positioned 45° to the plane of the sample. Just prior to microinjection, the solution to be microinjected was transferred to the capillary tip. Solutions of hairpin ribozyme to be microinjected was made up at a concentration of ~0.4 M and centrifuged at 16,200 x g at 4°C for 10 min to spin down any debris that could

clog the capillary.

### SDS-PAGE of HeLa cell lysates

Lysates from HeLa cells transiently expressing SNAP-CaaX-Hras and incubated with SNAP-Cell Oregon Green were harvested as follows. First, cells were washed twice with ice-cold Dulbecco's Phosphate-Buffered Saline (DPBS). Cells were then treated with 200 $\mu$l of cold RIPA buffer containing protease inhibitors (Roche) and 1 mM phenylmethylsulfonyl fluoride). The plates were gently shaken for ~5 min to equally diperse the RIPA buffer. Cells were then collected using a cell scraper, transferred into a microfuge tube and stored at -80°C until analysis by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). When lysate was ready to be analyzed by SDS-PAGE, lysate was thawed and then mixed with Laemmli buffer. Samples were then loaded into a 15% SDS-PAGE gel (with 5% stacking gel) and run at 90 V for ~15 min, then at 115 V for 1.5 h. Visualization of the bands of interest was by fluorescence scanning at 488 nm excitation wavelength and 520 nm emission filter sets using a Typhoon 9410 Variable Mode Imager.

## A.3 Results

### Challenges with the one-strand hairpin ribozyme

It was difficult to work with a concentrated enough sample of the hairpin ribozyme. Fig. A.3 shows that the stock solution of the ribozyme after deprotection contained a significant amount of RNA smearing and the presence of many low molecular weight bands when scanned for Cy5 (lane 2). The topmost band as seen from the Cy5 scan ran at the expected molecular weight and is likely to be the intact hairpin ribozyme. Following deprotection, intact hairpin ribozyme was purified from a 8% denaturing polyacrylamide gel. However, even after gel purification, a single clean band was not

seen. Instead, the bands were broad and smeared (Fig. A.3, lanes 3 and 4). The degree of smearing seen for this particular hairpin ribozyme is likely due to its length (see Discussion). As a result, the yields of intact, full-length hairpin ribozyme were low. In turn, the low yields proved difficult to use in a double labeling protocol to incorporate both BG-maleimide and Cy3-succinimidyl ester. Expression of SNAP-CaaX was therefore assessed using incubation with SNAP-Cell Oregon Green dye instead of microinjected hairpin ribozyme.

**Transiently expressed SNAP-CaaX-Hras in HeLa cells**

We first wanted to make sure that, in our hands, HeLa cells could be successfully transfected with pSNAP-CaaX-Hras. Therefore, HeLa cells were transiently transfected with pSNAP-CaaX-Hras and assessed for protein expression using SNAP-Cell Oregon Green dye. Non-transfected and non-labeled HeLa cells did not show any fluorescent cells, while transfected and labeled cells clearly showed fluorescence, indicating that SNAP-CaaX-Hras was being expressed (Fig. A.4a-h). Additionally, transfected cells showed especially dense fluorescence near the plasma membrane, suggesting that there was membrane localization of SNAP-CaaX-Hras.

**SNAP-CaaX-Hras and native $O^6$-alkylguanine-DNA alkyltransferase both bind to SNAP-Cell Oregon Green**

As further validation that SNAP-CaaX-Hras was expressed in HeLa cells and bound BG derivative substrates, we performed an SDS-PAGE gel analysis on lysate from HeLa cells. The cell lysate was harvested from transfected and non-transfected HeLa cells that had already incubated with SNAP-Cell Oregon Green. When the gel was scanned using 488 nm excitation and 520 nm emission filter sets, the lane containing lysate from transfected cells showed two bands running at ~20 kDa and a fainter band at ~25 kDa (Fig. A.5). These two bands can be explained by the

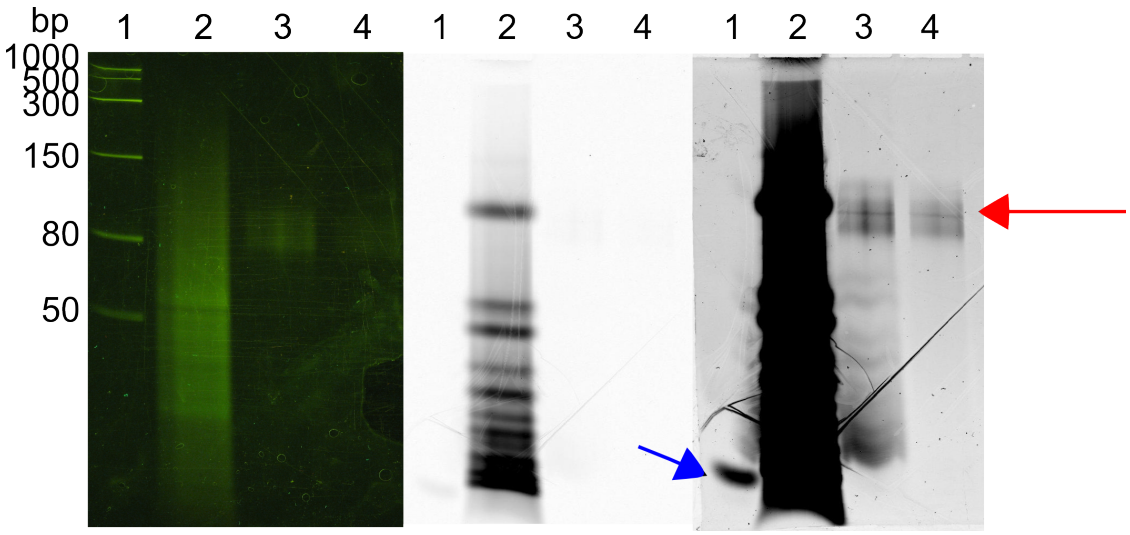


**Figure A.3: Challenges with the one-strand hairpin ribozyme.** Three visualizations are shown for an 8% denaturing polyacrylamide gel: SYBR Gold staining (leftmost), 532 nm scan with 350 mV PMT (middle), 532 nm scan with 500 mV PMT (rightmost). Lane 1 shows the low range single-stranded RNA ladder (New England Biolabs). Lane 2 shows the deprotected but unpurified hairpin ribozyme stock. Lanes 3 and 4 show ~100 pmol and ~50 pmol gel-purified hairpin ribozyme respectively. The red arrow indicates where the intact hairpin ribozyme ran and the blue arrow indicates xylene cyanol dye.

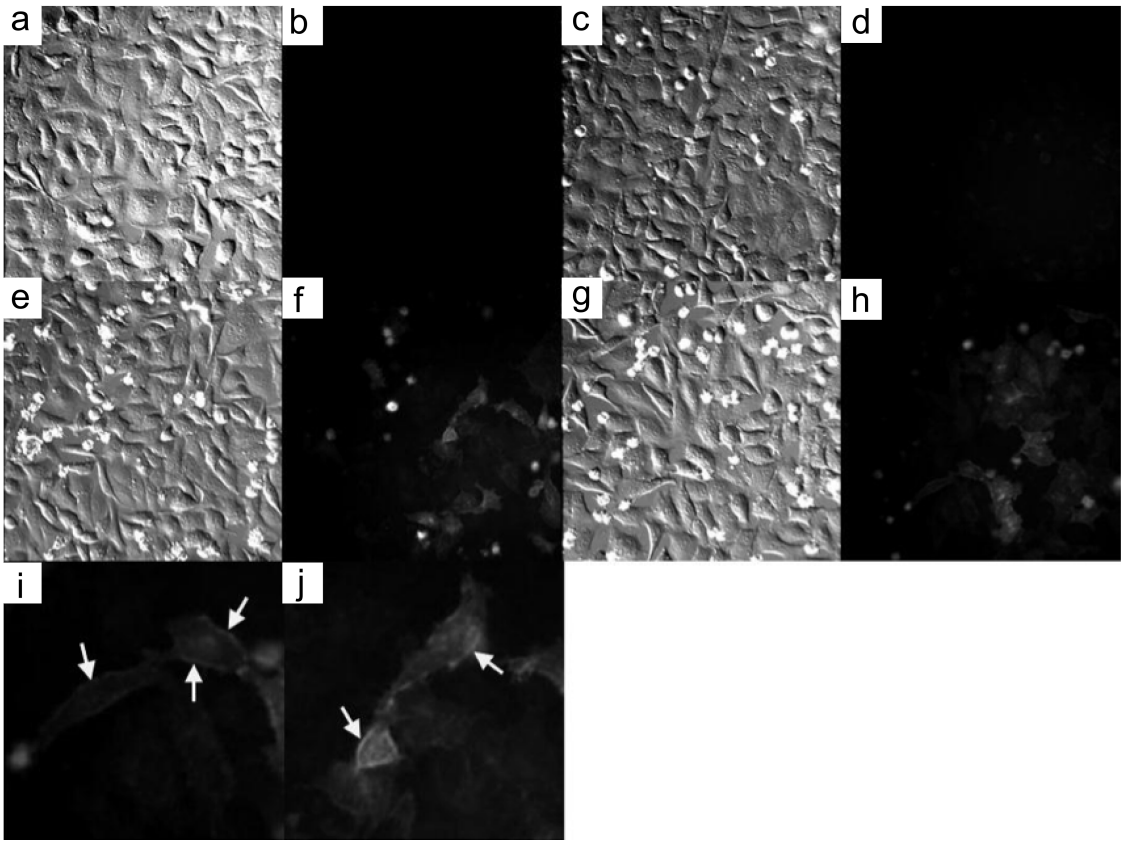


**Figure A.4: Transient expression of SNAP-CaaX-Hras in HeLa cells.** Images taken with bright field or 488 nm illumination. **(a),(b)** Non-transfected and non-labeled cells. **(c),(d)** Non-transfected cells labeled with SNAP-Cell Oregon Green. **(e)-(h)** Transfected cells labeled with SNAP-Cell Oregon Green. **(i),(j)** Close up images of (f) and (h) respectively. The white arrows indicate the higher density of fluorescence observed at the plasma membrane of cells. Images courtesy of Hanna Wagner.

binding of SNAP-Cell Oregon Green dye by both SNAP-CaaX-Hras (~20 kDa) and the endogenous $O^6$-alkylguanine-DNA alkyltransferase, which is ~24 kDa.[290] Furthermore, the 24-25 kDa band was also faintly present in lysate from non-tranfected cells, consistent with the idea that it was endogenous $O^6$-alkylguanine-DNA alkyltransferase (Fig. A.5). Although the BG derivatives are supposed to bind specifically to SNAP-CaaX-Hras, its similarity to the native $O^6$-alkylguanine-DNA alkyltransferase makes it conceivable that there is some cross-reactivity, hence the presence of a faint band from the alkyltransferase. These results confirm that SNAP-CaaX-Hras was being expressed in transfected cells and that SNAP-Cell Oregon Green was binding preferentially to the SNAP-CaaX-Hras protein.

**Transient transfection of SNAP-CaaX-Hras and SNAP-CaaX-Kras in U2OS cells**

Before attempting to make U2OS cells stably transfected with SNAP-CaaX-Hras, a transient transfection was done to ensure the SNAP-CaaX system also works well in U2OS cells. U2OS cells were transiently transfected with pSNAP-CaaX-Hras and imaged after incubation with SNAP-Cell Oregon Green (Fig. A.6c). Unlike in HeLa cells, transfected U2OS cells did not show as clearly a dense fluorescent cell perimeter, although this perimeter was visualized for some cells suggesting membrane localization of SNAP-CaaX-Hras (Fig. A.6c). There was, however, more autofluorescence present in the non-transfected and no-plasmid transfection controls in U2OS cells compared to HeLa cells (Fig. A.6a,b).

We were also interested in swapping the H-ras C-terminal sequence with the K-ras C-terminal sequence to increase the likelihood of SNAP-CaaX localization to the plasma membrane. Native H-ras can be found in both the plasma membrane and the endoplasmic reticulum (ER).[281,284] H-ras localization is driven by palmitoylation of another cysteine upstream of the CaaX box (Fig. A.2a).[281,284] Palmitoylation of the

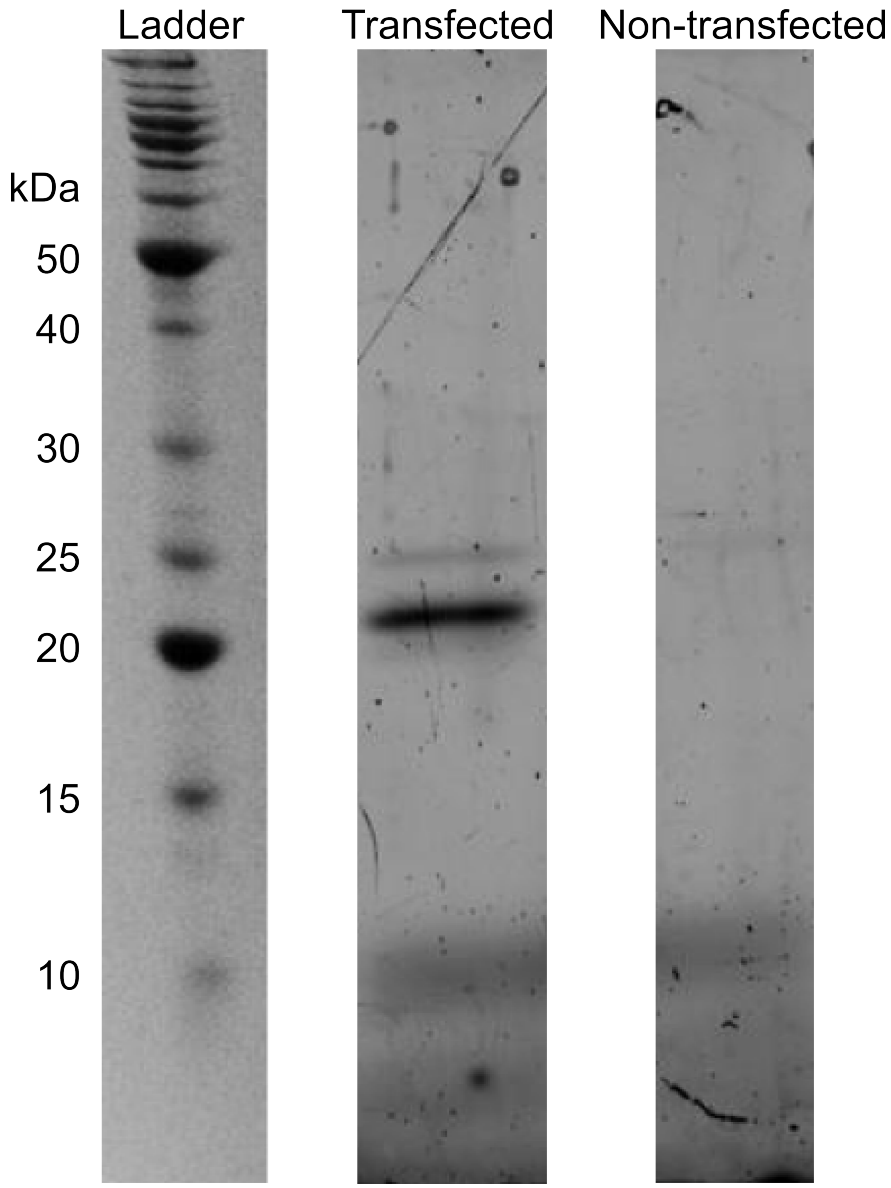


**Figure A.5: SDS-PAGE gel of HeLa cell lysates.** Lysates HeLa cells that were either non-transfected or transfected with pSNAP-CaaX-Hras were harvested and run on a 15% SDS-PAGE gel. The gel was scanned at 488 nm with 520 nm emission filter sets to detect the SNAP-Cell Oregon Green dye. The BenchMark protein ladder (Life Technologies) was also run for comparison and visualized by staining with Coomassie Blue.

upstream cysteine is required for H-ras to localize to the plasma membrane; without the palmitoyl signal, H-ras localizes to the ER.[284] K-ras, in contrast, does not have palmitoylation signals upstream of the CaaX sequence.[281] Instead, there is a polybasic sequence composed of several lysine residues upstream of the CaaX sequence (Fig. A.2), which results in K-ras being localized only to the plasma membrane.[281] Because maximal localization at the plasma membrane would be optimal, we cloned in the K-ras C-terminal sequence to replace the H-ras C-terminal sequence. U2OS cells were transiently transfected with pSNAP-CaaX-Kras (Fig. A.6d). Transfected cells showed clearly defined shapes of cells with many cells, more than for U2OS cells expressing SNAP-CaaX-Hras, exhibiting a densely fluorescent outline (Fig. A.6d).

**Stably transfected U2OS cell lines**

We next pursued creation of U2OS cells stably transfected with either pSNAP-CaaX-Hras or pSNAP-CaaX-Kras. Stably transfected cell lines are generally more convenient to use than transiently transfected cells. Monoclonal populations can be isolated for stably transfected cells where all the cells in the population express the desired protein to the same degree. The use of monoclonal populations eliminates a lot of experimental variation from the varying protein expression found in a pool of transiently transfected cells. Furthermore, transiently transfected cells only express the desired protein for up to ~48 h after transfection because the plasmid is not replicated and passed on to daughter cells as cells divide.[291] We have frozen down a polyclonal population of U2OS cells likely stably transfected with pSNAP-CaaX-Hras. Also, several colonies of U2OS cells isolated from the serial dilution protocol, and therefore predicted to be monoclonal, were also frozen down. Experiments involving incubation with the SNAP-Cell Oregon Green dye to confirm stable transfection and to analyze of the degree of SNAP-CaaX-Hras and SNAP-CaaX-Kras expression are still to be performed. These frozen colonies of cells are currently stored in liquid

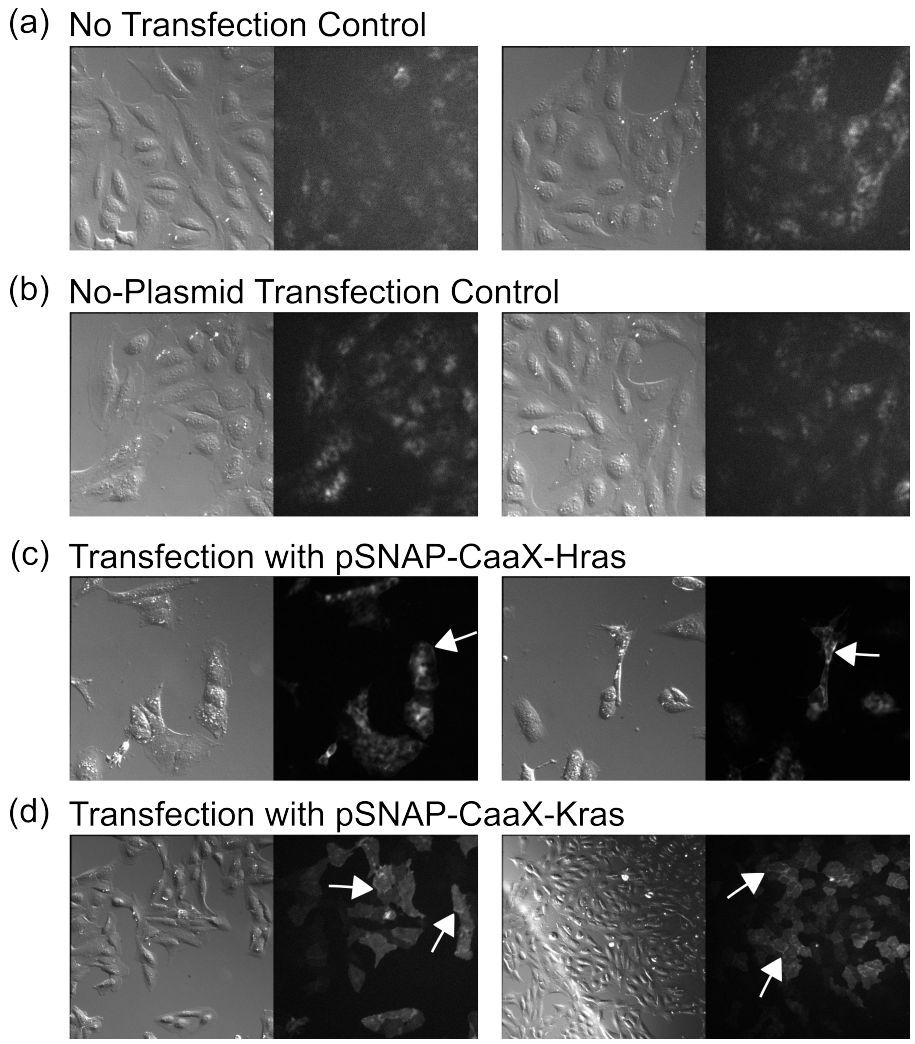


**Figure A.6: Transiently expressed SNAP-CaaX-Hras and SNAP-CaaX-Kras in U2OS cells.** Images of U2OS taken with bright-field and 488 nm illumination. **(a)** No transfection control. U2OS cells were not transfected with any plasmid but were still incubated with SNAP-Cell Oregon Green. **(b)** No-plasmid transfection control. U2OS cells were subjected to transfection conditions but no DNA was transfected. **(c)** U2OS cells transiently transfected with pSNAP-Caax-Hras. **(d)** U2OS cells transiently transfected with pSNAP-CaaX-Kras. White arrows indicate a few examples of the densely fluorescent perimeter present in several cells.

nitrogen.

## A.4 Discussion

The unique environment inside a living cell makes it especially challenging to do experiments that would be much more straightforward in an *in vitro* setting.[275,276] The intracellular environment can have a profound effect on the dynamics and function of molecules, so it is becoming increasingly important to further develop tools that enable us to do experiments inside living cells.[119,125,128,129,137,144,193] Additionally, intracellular tools will enable us to more accurately study biological systems that contain many components and are very complicated to replicate *in vitro*.[275,276,287,289,292,293] Finally, potential therapeutics, including ribozymes,[96] can greatly benefit from intracellular experiments early in the development pipeline, which can then significantly impact where in the molecule optimization efforts should be focused. Therefore, though such experiments are not trivial, the advances in developing intracellular experimental tools have important consequences. With our intracellular smFRET experiments, we faced technical challenges that we still have to overcome.

One of the major setbacks in our experimental design was the hairpin ribozyme we used. While it was initially thought that a one-strand ribozyme would be a simpler design than a two-strand ribozyme, the one-strand ribozyme turned out to be perhaps too long for practical use, especially with the number of modifications that had to be incorporated onto the RNA. It was possible there was some degradation of the hairpin ribozyme. However, the degree of smearing seen for this particular hairpin ribozyme (and not other versions of the ribozyme used in the lab) suggested that the majority of the problem lay with the synthesis and deprotection of such a long RNA. Due to the presence of many side-products that were not easily purified away and the low yield of the full-length hairpin ribozyme, we were not able to doubly-label a high enough concentration of ribozyme for microinjection into cells.

However, there are a number of future directions that can be attempted for this project. Instead of a one-strand hairpin ribozyme, two-strand or three-strand ribozymes, such as have been used in previous experiments, can be experimented with.[25,69,71] The modifications can then be divided up between the multiple strands such that the thiol modification for labeling with BG maleimide is not on the same strand as the primary amine group for labeling with Cy3-succinimidyl ester. The separation of these modifications would also address the generally inefficient maleimide-succinimidyl ester labeling scheme that has to be used for a one-strand construct. Alternatively, the ribozyme could be synthesized in two or three strands, separately modified and then ligated together, using a DNA splint and T4-DNA ligase, for instance, to produce a one-strand ribozyme.[294,295] Alternative membrane localization techniques could also be explored. For example, the ribozyme could be coupled to cholesterol for membrane association.[296]

The work in this appendix has laid out the potential for an intracellular smFRET protocol for RNA. In addition to components of the protocol already discussed, future directions will also need to consider, among other things, how to inhibit RNA degradation while in the cell and how to delay fluorophore photobleaching, all the while without using reagents that would drastically alter the health of the cell. Development of a protocol for intracellular smFRET of ribozymes is undoubtedly a very complicated challenge. We have established that HeLa and U2OS cells can be used with the SNAP-CaaX system. We also have results suggesting that SNAP-CaaX-Hras localizes to the plasma membrane in HeLa cells and that both SNAP-CaaX-Hras and SNAP-CaaX-Kras both localize to the plasma membrane in U2OS cells. Finally, though experiments to confirm are still needed, we have likely isolated monoclonal populations of U2OS cells stably transfected with pSNAP-CaaX-Kras and we have likely made polyclonal populations of U2OS cells stably transfected with pSNAP-CaaX-Hras. Once further characterized with regards to SNAP-CaaX-Hras

and SNAP-CaaX-Kras expression, they can be used for future experiments.

## A.5 Acknowledgements

For the generous donation of pSNAP-CaaX-Hras, New England Biolabs is acknowledged. Eka Melani is also acknowledged for her assistance on the cloning of the K-ras C-terminal CaaX sequence.

# APPENDIX B

# Bash Scripts for Processing Simulaid Hydrogen Bonding Occupancy Outputs

The following bash scripts were used to determine the presence of hydrogen bonding at each time point over the course of a molecular dynamics simulation. Scripts are based on data extracted by the Simulaid program.[173]

**1. getframe** - Used to determine the time points that have a hydrogen bond. Hydrogen bonds are defined as being ≤3.5 Å and ≥120°. Input for this script is the .hbn output after running the Simulaid program.

```
#!/bin/csh
# USAGE: ./getframe SimulaidOutput.hbn Atom1 Res1 Atom2 Res2

set file = $1
set atom1 = $2
set res1 = $3
set atom2 = $4
set res2 = $5

awk '\
      {     if ($1=="Configuration"){\
                  frame=$3\
            }if ($2==atom1 && $3==res1 && $7==atom2 && $8==res2){\
                  printf "%5d %s\n", frame, "\t" 0\
```

```
                }\
      }\
' atom1=$atom1 res1=$res1 atom2=$atom2 res2=$res2 $file

#end
```

**2. getframe_extra** - Used to determine the time points that, after analysis by the getframe script, do not have a hydrogen bond. Use with the all50000.dat file for simulations up to 50 ns.

```
#!/bin/bash

INPUT="[Input file name]"

awk '{printf "%5d\n", $1}' < $INPUT > temp.1
awk '{printf "%5d\n", $1}' < all50000.dat > temp.2
comm -1 -3 temp.1 temp.2 > $INPUT.comm
cat $INPUT $INPUT.comm > combined

sort -n combined -o $INPUT.sorted

rm temp.1 temp.2

awk '{\
       if ($2 == 0)\
         {\
            print $0\
         }\
       if ($2 != 0)\
         {\
            printf "%5d %s\n", $1, "\t" 1\
         }\
    }\
' < $INPUT.sorted > $INPUT.out

rm $INPUT.comm
rm combined
rm $INPUT.sorted
#end
```

**3. all50000.dat** - Used in conjunction with getframe_extra.

```
1       1
2       1
3       1
4       1
5       1
6       1
.       .
.       .
.       .
49997   1
49998   1
49999   1
50000   1
```

**4. Hydrogen bonding occupancy Matlab script** - Used to visually plot out the hydrogen bonding occupancy over time. A value of "0" corresponds to a hydrogen bond present and is indicated by the color red. A value of "1" corresponds to no hydrogen bond and is indicated by the color white. Script adapted from original written by Dr. Jana Sefcikova.

```
clear all;

sim_root_list = ({'[name of simulation]'});
sim_num_list = ({'[name of simulation]'});
DIST = 1;

num_cols = 1;
num_rows = ceil(length(sim_root_list)/num_cols);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%


color_map = [1.0000 0.0000 0.0000;
    1.0000 0.6666 0.0000;
    0.7843 0.7843 0.7843;
    1.0000 1.0000 1.0000];


figure(1); clf;
set(1,'papersize',[6 5],'paperposition',[0 0 6 5]);

for i = 1:length(sim_root_list)
```

```
    sim_root = sim_root_list{i};

    x = []; y = [];
    data1 = [];
    for j = 1:length(sim_num_list)
        sim_num = sim_num_list{j};

        file1 = ...
            sprintf('%sPlot_%s.out',sim_root,sim_num);

        data1 = load(file1);
        x = data1(:,1)*1e-3;
        y(1:size(data1,1),j) = data1(:,2);
    end

  img = zeros(size(y,1),size(y,2),3);

  for m = 1:size(img,1)
      for n = 1:size(img,2)
          if y(m,n) < 1
              index = 1;
%          elseif y(m,n) < 3.0
%              index = 2;
%          elseif y(m,n) < 3.5
%              index = 3;
          else
              index = 4;
          end

          img(m,n,:) = color_map(index,:);
      end
  end


  %SubM = [2 4];
   %     [6 8];
    %    [10 12];
    %    [14 16];
    %    [18 20];
    %    [22 24];
    %    [26 28];
    %    [30 32];
  subplot(num_rows,num_cols,i);
  imagesc(x,1:size(y,1),permute(img,[2 1 3]));
  %imagesc(x,1:size(y,1),img);
  xlim([0 50]);
  set(gca,'xtick',(0:10:50));
```

```
  set(gca,'ytick',(1:length(sim_num_list)));
  %Sets each Hbond distance as a "number" along the x-axis
  set(gca,'yticklabel',sim_num_list);
  set(gca, 'yticklabel',[]);
  set(gca, 'clim',[2.25 5.75]);
  for m = 1:length(sim_num_list)-1
  end
  set(gca,'linewidth',3);
  grid off;
  xlabel('Time (ns)','fontsize',18,'fontweight','bold');
end

return;
figure(2); clf;
imagesc((1:4),1,reshape(color_map,[],1,3));
set(gca,'xtick',[],'ytick',[]);
set(gca,'linewidth',3');
  for m = 1:length(sim_num_list)-1
    line(xlim,m*[1 1]+.5,'color','k','linewidth',3);
  end
xlim([-.5 .5]);
axis xy;

return;
```

# APPENDIX C

# Protocol and Scripts for Analysis of trFRET Data Collected Using the ISS Alba Confocal Fluorescence Microscope and VistaVision

The following protocol and scripts were used to convert the .csv files produced by VistaVision (ISS) to the "new" file format used by the etrans program, which is used to obtain the trFRET distance distributions.

1. Import the [name].csv file into Excel.
2. Transpose only the third row of the [name].csv file.
3. Save the file as [name]_tpsd.csv.
4. Run split.exe and get an output file with [name]_split.csv.

```
#!/bin/bash

FILE=[file name]
awk -F, ‘ { \
            for (i=1;i<=NF;i++) \
                { \
                  printf("%s%s",$i,i%8?",":"\n") \
                } \
}’ $FILE\_tpsd.csv > $FILE\_split.csv
```

5. Import the [name]_split.csv file into Excel again.

6. Insert a new column at the beginning and copy the column below and paste it in the [name]_split.csv file.

```
0
8
96
104
.
.  (Increase by increments of 8)
.
4064
4072
4080
4088
```

7. Save as [name].txt (tab delimited format).

8. Open [name].txt in Vim and get rid of the ^ M carriage returns.

9. Run convert_4_etrans.exe to get an output with the extension .mc2.

```
#!/bin/bash

FILE=[file name]

awk ‘{ \
          printf "%6d%7d%7d%7d%7d%7d%7d%7d%7d\n", $1, $2, $3, \
          $4, $5, $6, $7, $8, $9 \
}’ $FILE\.txt > $FILE\.mc2
```

10. Paste the header below into the .mc2 file. The file is then ready to be read as input by the etrans program.

```
DATE
GARBAGE
   12.207031
Collection Time
BLANK
```

# BIBLIOGRAPHY

# BIBLIOGRAPHY


[1] Cech, T. R.; Zaug, A. J.; Grabowski, P. J. *Cell* **1981**, *27*, 487–496.

[2] Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. *Cell* **1983**, *35*, 849–857.

[3] Crick, F. *Symp Soc Exp Biol* **1958**, *12*, 138–163.

[4] Crick, F. *Nature* **1970**, *227*, 561–563.

[5] Brenner, S.; Jacob, F.; Meselson, M. *Nature* **1961**, *190*, 576–581.

[6] Zimmermann, R. A.; Dahlberg, A. E. *Ribosomal RNA: Structure, Evolution, Processing, and Function in Protein Biosynthesis*; CRC Press: Boca Raton, FL, 1996.

[7] Hoagland, M. B.; Stephenson, M. L.; Scott, J. F.; Hecht, L. I.; Zamecnik, P. C. *J Biol Chem* **1958**, *231*, 241–257.

[8] Soll, D.; RajBhandary, U. L. *tRNA: Structure, Biosynthesis, and Function*; ASM: Washington, DC, 1995.

[9] Lilley, D. M. *RNA* **2004**, *10*, 151–158.

[10] Walter, N. G.; Burker, J. M. *Curr Opin Chem Biol* **1998**, *2*, 24–30.

[11] Lai, M. M. *Annu Rev Biochem* **1995**, *64*, 259–286.

[12] Cech, T. R. *Science* **2000**, *289*, 878–879.

[13] Lambowitz, A. M.; Caprara, M. G.; Zimmerly, S.; Perlman, P. S. *The RNA World*, second edition ed.; Cold Spring Harbor Laboratory Press, 1999; pp 451–485.

[14] Fica, S. M.; Tuttle, N.; Novak, T.; Li, N.-S.; Lu, J.; Koodathingal, P.; Dai, Q.; Staley, J. P.; Piccirilli, J. A. *Nature* **2013**, *503*, 229–234.

[15] Valadkhan, S. *RNA Biol* **2010**, *7*, 345–353.

[16] Gilbert, W. *Nature* **1986**, *319*, 618.

[17] Waldrop, M. M. *Science* **1989**, *246*, 1248–1249.

[18] Eddy, S. R. *Nat Rev Genet* **2001**, *2*, 919–929.

[19] Mattick, J. S. *BioEssays* **2003**, *25*, 930–939.

[20] Ferré-D'Amaré, A. R.; Scott, W. G. *Cold Spring Harb Perspect Biol* **2010**, *2*, a003574.

[21] Doherty, E. A.; Doudna, J. A. *Annu Rev Biochem* **2000**, *69*, 597–615.

[22] Strobel, S. A.; Cochrane, J. C. *Curr Opin Chem Biol* **2007**, *11*, 636–643.

[23] Salehi-Ashtiani, K.; Lupták, A.; Litovchick, A.; Szostak, J. W. *Science* **2006**, *313*, 1788–1792.

[24] Webb, C. H.; Riccitelli, N. J.; Ruminski, D. J.; Lupták, A. *Science* **2009**, *326*, 953.

[25] Bokinsky, G.; Rueda, D.; Misra, V. K.; Rhodes, M. M.; Gordus, A.; Babcock, H. P.; Walter, N. G.; Zhuang, X. *Proc Natl Acad Sci U S A* **2003**, *100*, 9302–9307.

[26] Bratty, J.; Chartrand, P.; Ferbeyre, G.; Cedergren, R. *Biochim Biophys Acta Gene Struct Expr* **1993**, *1216*, 345–359.

[27] Lilley, D. M. J.; Eckstein, F. *Ribozymes and RNA Catalysis*; Royal Society of Chemistry, 2008; pp 1–10.

[28] Ban, N.; Nissen, P.; Hansen, J.; Moore, P. B.; Steitz, T. A. *Science* **2000**, *289*, 905–920.

[29] Servant-Delmas, A.; Le Gal, F.; Gallian, P.; Gordien, E.; Laperche, S. *J Clin Virol* **2014**, *59*, 126–128.

[30] Ke, A.; Zhou, K.; Ding, F.; Cate, J. H.; Doudna, J. A. *Nature* **2004**, *429*, 201–205.

[31] Salter, J.; Krucinska, J.; Alam, S.; Grum-Tokars, V.; Wedekind, J. E. *Biochemistry* **2006**, *45*, 686–700.

[32] Been, M. D.; Wickham, G. S. *Eur J Biochem* **1997**, *247*, 741–753.

[33] Shih, I. H.; Been, M. D. *Annu Rev Biochem* **2002**, *71*, 887–917.

[34] Been, M. D. *Curr Top Microbiol Immunol* **2006**, *307*, 47–65.

[35] Flores, R.; Ruiz-Ruiz, S.; Serra, P. *Semin Liver Dis* **2012**, *32*, 201–210.

[36] Taylor, J. M. *Semin Liver Dis* **2012**, *32*, 195–200.

[37] Pereira, M. J.; Harris, D. A.; Rueda, D.; Walter, N. G. *Biochemistry* **2002**, *41*, 730–740.

[38] Been, M. D. *Trends Biochem Sci* **1994**, *19*, 251–256.

[39] Luptak, A.; Ferre-D'Amare, A. R.; Zhou, K.; Zilm, K. W.; Doudna, J. A. *J Am Chem Soc* **2001**, *123*, 8447–8452.

[40] Harris, D. A.; Rueda, D.; Walter, N. G. *Biochemistry* **2002**, *41*, 12051–12061.

[41] Jeong, S.; Sefcikova, J.; Tinsley, R. A.; Rueda, D.; Walter, N. G. *Biochemistry* **2003**, *42*, 7727–7740.

[42] Tinsley, R. A.; Harris, D. A.; Walter, N. G. *J Am Chem Soc* **2003**, *125*, 13972–13973.

[43] Tinsley, R. A.; Harris, D. A.; Walter, N. G. *Biochemistry* **2004**, *43*, 8935–8945.

[44] Das, S. R.; Piccirilli, J. A. *Nat Chem Biol* **2005**, *1*, 45–52.

[45] Tinsley, R. A.; Walter, N. G. *Biol Chem* **2007**, *388*, 705–715.

[46] Walter, N. G.; Perumal, S. *Springer Ser Biophys* **2009**, *13*, 103–127.

[47] Nakano, S.; Chadalavada, D. M.; Bevilacqua, P. C. *Science* **2000**, *287*, 1493–1497.

[48] Gong, B.; Chen, J. H.; Chase, E.; Chadalavada, D. M.; Yajima, R.; Golden, B. L.; Bevilacqua, P. C.; Carey, P. R. *J Am Chem Soc* **2007**, *129*, 13335–13342.

[49] Ferré-D'Amaré, A. R.; Zhou, K.; Doudna, J. A. *Nature* **1998**, *395*, 567–574.

[50] Murray, J. B.; Seyhan, A. A.; Walter, N. G.; Burke, J. M.; Scott, W. G. *Chem Biol* **1998**, *5*, 587–595.

[51] Rosenstein, S. P.; Been, M. D. *Biochemistry* **1990**, *29*, 8011–8016.

[52] Perrotta, A. T.; Shih, I.; Been, M. D. *Science* **1999**, *286*, 123–126.

[53] Chen, J. H.; Yajima, R.; Chadalavada, D. M.; Chase, E.; Bevilacqua, P. C.; Golden, B. L. *Biochemistry* **2010**, *49*, 6508–6518.

[54] Harris, D. A.; Tinsley, R. A.; Walter, N. G. *J Mol Biol* **2004**, *341*, 389–403.

[55] Tanaka, Y.; Tagaya, M.; Hori, T.; Sakamoto, T.; Kurihara, Y.; Katahira, M.; Uesugi, S. *Genes Cells* **2002**, *7*, 567–579.

[56] Roossinck, M. J.; Sleat, D.; Palukaitis, P. *Microbiol Rev* **1992**, *56*, 265–279.

[57] Buzayan, J. M.; Gerlach, W. L.; Bruening, G. *Nature* **1986**, *323*, 349–353.

[58] Kaper, J. M.; Tousignant, M. E.; Steger, G. *Biochem Biophys Res Commun* **1988**, *154*, 318–325.

[59] Branch, A. D.; Robertson, H. D. *Science* **1984**, *223*, 450–455.

[60] Hampel, A.; Tritz, R. *Biochemistry* **1989**, *28*, 4929–4933.

[61] Feldstein, P. A.; Buzayan, J. M.; Bruening, G. *Gene* **1989**, *82*, 53–61.

[62] Haseloff, J.; Gerlach, W. L. *Gene* **1989**, *82*, 43–52.

[63] Chowrira, B. M.; Burke, J. M. *Biochemistry* **1991**, *30*, 8518–8522.

[64] Duckett, D. R.; Murchie, A. I. H.; Lilley, D. M. J. *Cell* **1995**, *83*, 1027–1036.

[65] Rupert, P. B.; Ferré-D'Amaré, A. R. *Nature* **2001**, *410*, 780–786.

[66] Zhuang, X.; Kim, H.; Pereira, M. J. B.; Babcock, H. P.; Walter, N. G.; Chu, S. *Science* **2002**, *296*, 1473–1476.

[67] Rupert, P. B.; Massey, A. P.; Sigurdsson, S. T.; Ferré-D'Amaré, A. R. *Science* **2002**, *298*, 1421–1424.

[68] Ryder, S. P.; Strobel, S. A. *J Mol Biol* **1999**, *291*, 295–311.

[69] Liu, S.; Bokinsky, G.; Walter, N. G.; Zhuang, X. *Proc Natl Acad Sci U S A* **2007**, *104*, 12634–12639.

[70] Kuzmin, Y. I.; Da Costa, C. P.; Fedor, M. J. *J Mol Biol* **2004**, *340*, 233–251.

[71] Rueda, D.; Bokinsky, G.; Rhodes, M. M.; Rust, M. J.; Zhuang, X.; Walter, N. G. *Proc Natl Acad Sci U S A* **2004**, *101*, 10066–10071.

[72] Donahue, C. P.; Yadava, R. S.; Nesbitt, S. M.; Fedor, M. J. *J Mol Biol* **2000**, *295*, 693–707.

[73] Ferré-D'Amaré, A. R. *Biopolymers* **2004**, *73*, 71–78.

[74] Nam, K.; Gao, J.; York, D. M. *RNA* **2008**, *14*, 1501–1507.

[75] Nam, K.; Gao, J.; York, D. M. *J Am Chem Soc* **2008**, *130*, 4680–4691.

[76] Hampel, K. J.; Walter, N. G.; Burke, J. M. *Biochemistry* **1998**, *37*, 14672–14682.

[77] Chowrira, B. M.; Berzal-Herranz, A.; Burke, J. M. *Biochemistry* **1993**, *32*, 1088–1095.

[78] Hampel, A.; Cowan, J. A. *Chem Biol* **1997**, *4*, 513–517.

[79] Nesbitt, S.; Hegg, L. A.; Fedor, M. J. *Chem Biol* **1997**, *4*, 619–630.

[80] Young, K. J.; Gill, F.; Grasby, J. A. *Nucleic Acids Res* **1997**, *25*, 3760.

[81] Fedor, M. J. *J Mol Biol* **2000**, *297*, 269–291.

[82] Pinard, R.; Hampel, K. J.; Heckman, J. E.; Lambert, D.; Chan, P. A.; Major, F.; Burke, J. M. *EMBO J* **2001**, *20*, 6434–6442.

[83] Kuzmin, Y. I.; Da Costa, C. P.; Cottrell, J. W.; Fedor, M. J. *J Mol Biol* **2005**, *349*, 989–1010.

[84] Spitale, R. C.; Volpini, R.; Heller, M. G.; Krucinska, J.; Cristalli, G.; Wedekind, J. E. *J Am Chem Soc* **2009**, *131*, 6093–6095.

[85] Bevilacqua, P. C. *Biochemistry* **2003**, *42*, 2259–2265.

[86] Ditzler, M. A.; Šponer, J.; Walter, N. G. *RNA* **2009**, *15*, 560–575.

[87] Torelli, A. T.; Krucinska, J.; Wedekind, J. E. *RNA* **2007**, *13*, 1052–1070.

[88] MacElrevey, C.; Salter, J. D.; Krucinska, J.; Wedekind, J. E. *RNA* **2008**, *14*, 1600–1616.

[89] Cottrell, J. W.; Kuzmin, Y. I.; Fedor, M. J. *J Biol Chem* **2007**, *282*, 13498–13507.

[90] Torelli, A. T.; Spitale, R. C.; Krucinska, J.; Wedekind, J. E. *Biochem Biophys Res Commun* **2008**, *371*, 154–158.

[91] Mlýnský, V.; Banáš, P.; Hollas, D.; Réblová, K.; Walter, N. G.; Šponer, J.; Otyepka, M. *J Phys Chem B* **2010**, *114*, 6642–6652.

[92] Rhodes, M. M.; Réblová, K.; Šponer, J.; Walter, N. G. *Proc Natl Acad Sci U S A* **2006**, *103*, 13380–13385.

[93] Walter, N. G.; Hampel, K. J.; Brown, K. M.; Burke, J. M. *EMBO J* **1998**, *17*, 2378–2391.

[94] Tang, J.; Breaker, R. R. *Proc Natl Acad Sci U S A* **2000**, *97*, 5784–5789.

[95] Jäschke, A. *Curr Opin Struct Biol* **2001**, *11*, 321–326.

[96] Burnett, J.; Rossi, J. *Chem Biol* **2012**, *19*, 60–71.

[97] Scherer, L. J.; Rossi, J. J. *Nat Biotech* **2003**, *21*, 1457–1465.

[98] Ramanathan, A.; Savol, A. J.; Langmead, C. J.; Agarwal, P. K.; Chennubhotla, C. S. *PLoS ONE* **2011**, *6*, e15827.

[99] Al-Hashimi, H. M.; Walter, N. G. *Curr Opin Struct Biol* **2008**, *18*, 321–329.

[100] Bahar, I.; Lezon, T. R.; Yang, L.-W.; Eyal, E. *Annu Rev Biophys* **2010**, *39*, 23–42.

[101] Heyduk, E.; Heyduk, T.; Lee, J. C. *J Biol Chem* **1992**, *267*, 3200–3204.

[102] Zhuang, X.; Bartley, L. E.; Babcock, H. P.; Russell, R.; Ha, T.; Herschlag, D.; Chu, S. *Science* **2000**, *288*, 2048–2051.

[103] Xu, Z.; Sigler, P. B. *J Struct Biol* **1998**, *124*, 129–141.

[104] Merlino, A.; Vitagliano, L.; Antoine Ceruso, M.; Di Nola, A.; Mazzarella, L. *Biopolymers* **2002**, *65*, 274–283.

[105] Rasmussen, B. F.; Stock, A. M.; Ringe, D.; Petsko, G. A. *Nature* **1992**, *357*, 423–424.

[106] McDowell, S. E.; Jun, J. M.; Walter, N. G. *RNA* **2010**, *16*, 2414–2426.

[107] Itoh, K.; Sasai, M. *Proc Natl Acad Sci U S A* **2010**, *107*, 7775–7780.

[108] Kim, D.-N.; Nguyen, C.-T.; Bathe, M. *J Struct Biol* **2010**, *173*, 261–270.

[109] Conway, J. F.; Wikoff, W. R.; Cheng, N.; Duda, R. L.; Hendrix, R. W.; Johnson, J. E.; Steven, A. C. *Science* **2001**, *292*, 744–748.

[110] Leach, A. R. *Molecular Modelling Principles and Applications*, 2nd ed.; Pearson Educated Limited, 2001.

[111] Perez, A.; Marchan, I.; Svozil, D.; Šponer, J.; Cheatham, T. E. r.; Laughton, C. A.; Orozco, M. *Biophys J* **2007**, *92*, 3817–3829.

[112] Šponer, J.; Špačková, N. *Methods* **2007**, *43*, 278–290.

[113] Banáš, P.; Sklenovský, P.; Wedekind, J. E.; Šponer, J.; Otyepka, M. *J Phys Chem B* **2012**, *116*, 12721–12734.

[114] Krasovska, M. V.; Sefcikova, J.; Špačková, N.; Šponer, J.; Walter, N. G. *J Mol Biol* **2005**, *351*, 731–748.

[115] Zgarbová, M.; Otyepka, M.; Šponer, J.; Mládek, A.; Banáš, P.; Cheatham, T. E. r.; Jurečka, P. *J Chem Theory Comput* **2011**, *7*, 2886–2902.

[116] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M. J.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J Am Chem Soc* **1995**, *117*, 5179–5197.

[117] Perez, A.; Luque, F. J.; Orozco, M. *Acc Chem Res* **2012**, *45*, 196–205.

[118] Saxena, A.; Sept, D. *J Chem Theory Comput* **2013**, *9*, 3538–3542.

[119] Ellis, R. J. *Trends Biochem Sci* **2001**, *26*, 597–604.

[120] Ellis, R. J. *Curr Opin Struct Biol* **2001**, *11*, 114–119.

[121] Ellis, R. J.; Minton, A. P. *Nature* **2003**, *425*, 27–28.

[122] Zimmerman, S. B.; Trach, S. O. *J Mol Biol* **1991**, *222*, 599–620.

[123] Zimmerman, S. B.; Minton, A. P. *Annu Rev Biophys Biomol Struct* **1993**, *22*, 27–65.

[124] Luby-Phelps, K. *Int Rev Cytol* **2000**, *192*, 189–221.

[125] Fessl, T.; Adamec, F.; Polívka, T.; Foldynová-Trantírková, S.; Vácha, F.; Trantírek, L. *Nucleic Acids Res* **2012**, *40*, e121.

[126] Schafer, F. Q.; Buettner, G. R. *Free Radic Biol Med* **2001**, *30*, 1191–1212.

[127] Knight, A. B.; Welt, L. G. *J Gen Physiol* **1974**, *63*, 351–373.

[128] Tijerina, P.; Bhaskaran, H.; Russell, R. *Proc Natl Acad Sci U S A* **2006**, *103*, 16698–16703.

[129] Ando, T.; Skolnick, J. *Proc Natl Acad Sci U S A* **2010**, *107*, 18457–18462.

[130] Herschlag, D.; Khosla, M.; Tsuchihashi, Z.; Karpel, R. L. *EMBO J* **1994**, *13*, 2913–2924.

[131] Singh, R.; Valcarcel, J. *Nat Struct Mol Biol* **2005**, *12*, 645–653.

[132] Kaminski, A.; Ostareck, D. H.; Standart, N. M.; Jackson, R. J. In *RNA-Protein Interactions : A Practical Approach*; Smith, C. W. J., Ed.; Oxford University Press Inc.: New York, 2002; pp 137–160.

[133] Magee, J.; Warwicker, J. *Nucleic Acids Res* **2005**, *33*, 6694–6699.

[134] Guigas, G.; Kalla, C.; Weiss, M. *FEBS Lett* **2007**, *581*, 5094–5098.

[135] Nakano, S.-i.; Karimata, H.; Ohmichi, T.; Kawakami, J.; Sugimoto, N. *J Am Chem Soc* **2004**, *126*, 14330–14331.

[136] Wang, Y.; Benton, L. A.; Singh, V.; Pielak, G. J. *J Phys Chem Lett* **2012**, *3*, 2703–2706.

[137] Nakano, S.-i.; Miyoshi, D.; Sugimoto, N. *Chemical Reviews* **2014**, *114*, 2733–2758.

[138] Zimmerman, S. B.; Trach, S. O. *Nucleic Acids Res* **1988**, *16*, 6309–6326.

[139] Nashimoto, M. *Eur J Biochem* **2000**, *267*, 2738–2745.

[140] Nakano, S.-i.; Karimata, H. T.; Kitagawa, Y.; Sugimoto, N. *J Am Chem Soc* **2009**, *131*, 16881–16888.

[141] Nakano, S.-i.; Kitagawa, Y.; Karimata, H. T.; Sugimoto, N. *Nucleic Acids Symp Ser* **2008**, *52*, 519–520.

[142] Karimata, H.; Nakano, S.-i.; Sugimoto, N. *Nucleic Acids Symp Ser* **2006**, *50*, 81–82.

[143] Kilburn, D.; Roh, J. H.; Guo, L.; Briber, R. M.; Woodson, S. A. *J Am Chem Soc* **2010**, *132*, 8690–8696.

[144] Kilburn, D.; Roh, J. H.; Behrouzi, R.; Briber, R. M.; Woodson, S. A. *J Am Chem Soc* **2013**, *135*, 10055–10063.

[145] Lambert, D.; Draper, D. E. *J Mol Biol* **2007**, *370*, 993–1005.

[146] Pincus, D. L.; Hyeon, C.; Thirumalai, D. *J Am Chem Soc* **2008**, *130*, 7364–7372.

[147] Webb, C. H.; Lupták, A. *RNA Biol* **2011**, *8*, 719–727.

[148] Been, M. D.; Wickham, G. S. *Eur J Biochem* **1997**, *247*, 741–753.

[149] Perrotta, A. T.; Nikiforova, O.; Been, M. D. *Nucleic Acids Res* **1999**, *27*, 795–802.

[150] Wadkins, T. S.; Perrotta, A. T.; Ferré-D'Amaré, A. R.; Doudna, J. A.; Been, M. D. *RNA* **1999**, *5*, 720–727.

[151] Perrotta, A. T.; Been, M. D. *Biochemistry* **2006**, *45*, 11357–11365.

[152] Perrotta, A. T.; Wadkins, T. S.; Been, M. D. *RNA* **2006**, *12*, 1282–1291.

[153] Nakano, S.; Cerrone, A. L.; Bevilacqua, P. C. *Biochemistry* **2003**, *42*, 2982–2994.

[154] Chadalavada, D. M.; Cerrone-Szakal, A. L.; Bevilacqua, P. C. *RNA* **2007**, *13*, 2189–2201.

[155] Cerrone-Szakal, A. L.; Chadalavada, D. M.; Golden, B. L.; Bevilacqua, P. C. *RNA* **2008**, *14*, 1746–1760.

[156] Gong, B.; Chen, Y.; Christian, E. L.; Chen, J. H.; Chase, E.; Chadalavada, D. M.; Yajima, R.; Golden, B. L.; Bevilacqua, P. C.; Carey, P. R. *J Am Chem Soc* **2008**, *130*, 9670–9672.

[157] Chen, J. H.; Gong, B.; Bevilacqua, P. C.; Carey, P. R.; Golden, B. L. *Biochemistry* **2009**, *48*, 1498–1507.

[158] Gong, B.; Chen, J. H.; Bevilacqua, P. C.; Golden, B. L.; Carey, P. R. *Biochemistry* **2009**, *48*, 11961–11970.

[159] Sefcikova, J.; Krasovska, M. V.; Špačková, N.; Šponer, J.; Walter, N. G. *Biopolymers* **2007**, *85*, 392–406.

[160] Sefcikova, J.; Krasovska, M. V.; Šponer, J.; Walter, N. G. *Nucleic Acids Res* **2007**, *35*, 1933–1946.

[161] Ferré-D'Amaré, A. R.; Doudna, J. A. *J Mol Biol* **2000**, *295*, 541–556.

[162] Walter, N. G. *Curr Protoc Nucleic Acid Chem* **2002**, *11*, 11.10.1–11.10.23.

[163] Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. *Nat Methods* **2012**, *9*, 671–675.

[164] Case, D. A. et al. AMBER 10. 2008.

[165] Case, D. A. et al. AMBER 11. 2010.

[166] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J Chem Phys* **1983**, *79*, 926–935.

[167] Beveridge, D. L.; Cheatham, T. E. r.; Mezei, M. *J Biosci* **2012**, *37*, 379–397.

[168] Lavery, R. et al. *Nucleic Acids Res* **2010**, *38*, 299–313.

[169] Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J Chem Phys* **1984**, *81*, 3684–3690.

[170] Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J Comput Phys* **1977**, *23*, 327–341.

[171] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. *J Am Chem Soc* **1993**, *115*.

[172] Frisch, M. J. et al. Gaussian 03. 2004.

[173] Mezei, M. *J Comput Chem* **2010**, *31*, 2658–2668.

[174] MATLAB, *version 7.10.0 (R2010a)*; The MathWorks Inc.: Natick, MA, 2010.

[175] Soukup, G. A.; Breaker, R. R. *RNA* **1999**, *5*, 1308–1325.

[176] Zuker, M. *Nucleic Acids Res* **2003**, *31*, 3406–3415.

[177] Wadkins, T. S.; Been, M. D. *Cell Mol Life Sci* **2002**, *59*, 112–125.

[178] Walter, N. G.; Burke, J. M. *Methods Enzymol* **2000**, *317*, 409–440.

[179] Walter, N. G. *Methods* **2001**, *25*, 19–30.

[180] Walter, N. G.; Harris, D. A.; Pereira, M. J.; Rueda, D. *Biopolymers* **2001**, *61*, 224–242.

[181] Walter, N. G.; Burke, J. M.; Millar, D. P. *Nat Struct Biol* **1999**, *6*, 544–549.

[182] Walter, N. G.; Chan, P. A.; Hampel, K. J.; Millar, D. P.; Burke, J. M. *Biochemistry* **2001**, *40*, 2580–2587.

[183] Rueda, D.; Wick, K.; McDowell, S. E.; Walter, N. G. *Biochemistry* **2003**, *42*, 9924–9936.

[184] Ditzler, M. A.; Otyepka, M.; Šponer, J.; Walter, N. G. *Acc Chem Res* **2010**, *43*, 40–47.

[185] Krasovska, M. V.; Sefcikova, J.; Réblová, K.; Schneider, B.; Walter, N. G.; Šponer, J. *Biophys J* **2006**, *91*, 626–638.

[186] McDowell, S. E.; Spackova, N.; Sponer, J.; Walter, N. G. *Biopolymers* **2007**, *85*, 169–184.

[187] Banáš, P.; Hollas, D.; Zgarbová, M.; Jurečka, P.; Orozco, M.; Cheatham, T. E. r.; Sponer, J.; Otyepka, M. *J Chem Theory Comput* **2010**, *6*, 3836–3849.

[188] Lee, T. S.; Giambasu, G.; Harris, M. E.; York, D. M. *J Phys Chem Lett* **2011**, *2*, 2538–2543.

[189] Marek, M. S.; Johnson-Buck, A.; Walter, N. G. *Phys Chem Chem Phys* **2011**, *13*, 11524–11537.

[190] Nakano, S.; Proctor, D. J.; Bevilacqua, P. C. *Biochemistry* **2001**, *40*, 12022–12038.

[191] Bevilacqua, P. C.; Brown, T. S.; Nakano, S.; Yajima, R. *Biopolymers* **2004**, *73*, 90–109.

[192] Walter, N. G. *Mol Cell* **2007**, *28*, 923–929.

[193] Woodson, S. A. *RNA Biol* **2010**, *7*, 677–686.

[194] Mlýnský, V.; Banáš, P.; Walter, N. G.; Šponer, J.; Otyepka, M. *J Phys Chem B* **2011**, *115*, 13911–13924.

[195] Lide, D. R. *CRC Handbook of Chemistry and Physics*, 83rd ed.; CRC Press: Boca Raton, FL, 2003.

[196] Guo, M.; Spitale, R. C.; Volpini, R.; Krucinska, J.; Cristalli, G.; Carey, P. R.; Wedekind, J. E. *J Am Chem Soc* **2009**, *131*, 12908–12909.

[197] Cottrell, J. W.; Scott, L. G.; Fedor, M. J. *J Biol Chem* **2011**, *286*, 17658–17664.

[198] Spitale, R. C.; Volpini, R.; Mungillo, M. V.; Krucinska, J.; Cristalli, G.; Wedekind, J. E. *Biochemistry* **2009**, *48*, 7777–7779.

[199] Park, H.; Lee, S. *J Chem Theory Comput* **2006**, *2*, 858–862.

[200] Geissler, P. L.; Dellago, C.; Chandler, D.; Hutter, J.; Parrinello, M. *Science* **2001**, *291*, 2121–2124.

[201] Mohammed, O. F.; Pines, D.; Dreyer, J.; Pines, E.; Nibbering, E. T. J. *Science* **2005**, *310*, 83–86.

[202] Martick, M.; Scott, W. G. *Cell* **2006**, *126*, 309–320.

[203] Klein, D. J.; Ferré-D'Amaré, A. R. *Science* **2006**, *313*, 1752–1756.

[204] Schmeing, T. M.; Huang, K. S.; Kitchen, D. E.; Strobel, S. A.; Steitz, T. A. *Mol Cell* **2005**, *20*, 437–448.

[205] Prashar, V.; Bihani, S.; Das, A.; Ferrer, J.-L.; Hosur, M. *PLoS ONE* **2009**, *4*, e7860.

[206] Tashiro, M.; Stuchebrukhov, A. A. *J Phys Chem B* **2004**, *109*, 1015–1022.

[207] McGregor, A.; Rao, M. V.; Duckworth, G.; Stockley, P. G.; Connolly, B. A. *Nucleic Acids Res* **1996**, *24*, 3173–3180.

[208] Humphrey, W.; Dalke, A.; Schulten, K. *J Molec Graphics* **1996**, *14*, 33–38.

[209] Meyer, E. *Protein Science* **1992**, *1*, 1543–1562.

[210] Ditzler, M. A.; Rueda, D.; Mo, J.; Håkansson, K.; Walter, N. G. *Nucleic Acids Res* **2008**, *36*, 7088–7099.

[211] Desai, R.; Kilburn, D.; Lee, H.-T.; Woodson, S. A. *J Biol Chem* **2013**, *289*, 2972–2977.

[212] Alam, S.; Grum-Tokars, V.; Krucinska, J.; Kundracik, M. L.; Wedekind, J. E. *Biochemistry* **2005**, *44*, 14396–14408.

[213] Greenfeld, M.; Solomatin, S. V.; Herschlag, D. *J Biol Chem* **2011**, *286*, 19872–19879.

[214] Solomatin, S. V.; Greenfeld, M.; Chu, S.; Herschlag, D. *Nature* **2010**, *463*, 681–684.

[215] Pereira, M. J.; Behera, V.; Walter, N. G. *PLoS One* **2010**, *5*, e12953.

[216] Russell, R. *Front Biosci* **2008**, *13*, 1–20.

[217] Thirumalai, D.; Klimov, D. K.; Lorimer, G. H. *Proc Natl Acad Sci U S A* **2003**, *100*, 11195–11197.

[218] Minton, A. P. *Biopolymers* **1981**, *20*, 2093–2120.

[219] Minton, A. P. *Molecular crowding: analysis of effects of high concentrations of inert cosolutes on biochemical equilibria and rates in terms of volume exclusion*; Academic Press, 1998; Vol. 295; pp 127–149.

[220] Zhou, H.-X.; Qin, S. *Phys Rev E Stat Nonlin Soft Matter Phys* **2010**, *81*, 031919.

[221] Zhou, H.-X.; Rivas, G.; Minton, A. P. *Annu Rev Biophys* **2008**, *37*, 375–397.

[222] Grubbs, R. D. *Biometals* **2002**, *15*, 251–259.

[223] Beeler, T.; Bruce, K.; Dunn, T. *Biochim Biophys Acta* **1997**, *1323*, 310–318.

[224] Kozak, J. A.; Cahalan, M. D. *Biophys J* **2003**, *84*, 922–927.

[225] Chen, C.; Nakatani, K.; Koutalos, Y. *J Physiol* **2003**, *553*, 125–135.

[226] Spitale, R. C.; Wedekind, J. E. *Methods* **2009**, *49*, 87–100.

[227] Benos, D. J. *Developmental Biology of Membrane Transport Systems*; Current Topics in Membranes; Academic Press Inc.: San Diego, 1991; Vol. 39.

[228] Sperelakis, N. *Cell Physiology Sourcebook: A Molecular Approach*, 3rd ed.; Academic Press Inc.: San Diego, 2001.

[229] Packer, L.; Yodoi, J. *Redox Regulation of Cell Signaling and Its Clinical Application*; Marcel Dekker Inc.: New York, 1999.

[230] Krauth-Siegel, R. L.; Jockers-Scherübl, M. C.; Becker, K.; Schirmer, R. H. *Biochem Soc Trans* **1989**, *17*, 315–317.

[231] Denesyuk, N. A.; Thirumalai, D. *J Am Chem Soc* **2011**, *133*, 11858–11861.

[232] Tan, Z.-J.; Chen, S.-J. *Biophys J* **2012**, *103*, 827–836.

[233] Elcock, A. H. *Curr Opin Struct Biol* **2010**, *20*, 196–206.

[234] Edmonds, B. T.; Wyckoff, J.; Yeung, Y. G.; Wang, Y.; Stanley, E. R.; Jones, J.; Segall, J.; Condeelis, J. *J Cell Sci* **1996**, *109*, 2705–2714.

[235] Record, M. T. *Biopolymers* **1975**, *14*, 2137–2158.

[236] Record Jr., M.; Zhang, W.; Anderson, C. *Adv Protein Chem* **1998**, *51*, 281–353.

[237] Manning, G. S. *Quarterly reviews of biophysics* **1978**, *11*, 179.

[238] Karimata, H.; Nakano, S.-i.; Ohmichi, T.; Kawakami, J.; Sugimoto, N. *Nucleic Acids Symp Ser* **2004**, *48*, 107–108.

[239] Tan, C.; Saurabh, S.; Bruchez, M. P.; Schwartz, R.; LeDuc, P. *Nat Nanotechnol* **2013**, *8*, 602–608.

[240] Rajendran, A.; Nakano, S.-i.; Sugimoto, N. *Chem Commun* **2010**, *46*, 1299–1301.

[241] Ge, X.; Luo, D.; Xu, J. *PLoS One* **2011**, *6*, e28707.

[242] Sanders, G. M.; Kassavetis, G. A.; Geiduschek, E. P. *Proc Natl Acad Sci U S A* **1994**, *91*, 7703–7707.

[243] Ha, T. *Methods* **2001**, *25*, 78–86.

[244] Atha, D. H.; Ingham, K. C. *J Biol Chem* **1981**, *256*, 12108–12117.

[245] Fahie-Wilson, M.; Halsall, D. *Ann Clin Biochem* **2008**, *45*, 233–235.

[246] Humphreys, G. O.; Willshaw, G. A.; Anderson, E. S. *Biochim Biophys Acta* **1975**, *383*, 457–463.

[247] Nilsen, T. W. *Cold Spring Harb Protoc* **2011**, *2012*, pdb.prot072322.

[248] Genz, C.; Fundakowski, J.; Hermesh, O.; Schmid, M.; Jansen, R.-P. *J Biol Chem* **2013**,

[249] Holmes, K. J.; Klass, D. M.; Guiney, E. L.; Cyert, M. S. *PLoS ONE* **2013**, *8*, e84060.

[250] Gabus, C.; Derrington, E.; Leblanc, P.; Chnaiderman, J.; Dormont, D.; Swietnicki, W.; Morillas, M.; Surewicz, W. K.; Marc, D.; Nandi, P.; Darlix, J.-L. *J Biol Chem* **2001**, *276*, 19301–19309.

[251] Gierasch, L. M.; Gershenson, A. *Nat Chem Biol* **2009**, *5*, 774–777.

[252] Hukushima, K.; Nemoto, K. *J Phys Soc Jpn* **1996**, *65*, 1604–1608.

[253] Swendsen, R. H.; Wang, J.-S. *Phys Rev Lett* **1986**, *57*, 2607–2609.

[254] Koplin, J.; Mu, Y.; Richter, C.; Schwalbe, H.; Stock, G. *Structure* **2005**, *13*, 1255–1267.

[255] Riccardi, L.; Nguyen, P. H.; Stock, G. *J Phys Chem B* **2009**, *113*, 16660–16668.

[256] Gilmour, S. J.; Zeevaart, J. A. D.; Schwenen, L.; Graebe, J. E. *Plant Physiol* **1986**, *82*, 190–195.

[257] Nilsen, T. W. *Cold Spring Harb Protoc* **2013**, *2013*, pdb.prot075176.

[258] Lasda, E. L.; Kuersten, S.; Blumenthal, T. *Cold Spring Harb Protoc* **2011**, *2011*, pdb.prot5574.

[259] Ben-Hail, D.; Shoshan-Barmatz, V. *Cold Spring Harbor Protocols* **2014**, *2014*, pdb.prot073130.

[260] Kigawa, T.; Yabuki, T.; Matsuda, N.; Matsuda, T.; Nakajima, R.; Tanaka, A.; Yokoyama, S. *J Struct Funct Genomics* **2004**, *5*, 63–68.

[261] Visweswaraiah, J.; Dautel, M.; Sattlegger, E. *Protocol Exchange* **2011**, doi:10.1038/protex.2011.212.

[262] Mattick, J. S.; Makunin, I. V. *Hum Mol Genet* **2006**, *15*, R17–R29.

[263] Espinoza, C. A.; Allen, T. A.; Hieb, A. R.; Kugel, J. F.; Goodrich, J. A. *Nat Struct Mol Biol* **2004**, *11*, 822–829.

[264] Allen, T. A.; Von Kaenel, S.; Goodrich, J. A.; Kugel, J. F. *Nat Struct Mol Biol* **2004**, *11*, 816–821.

[265] Fire, A. *Trends Genet* **1999**, *15*, 358–363.

[266] Hacisuleyman, E. et al. *Nat Struct Mol Biol* **2014**, *21*, 198–206.

[267] Yang, L.; Froberg, J. E.; Lee, J. T. *Trends in Biochem Sci* **2014**, *39*, 35–43.

[268] Barry, G. *Mol Psychiatr* **2014**, *19*, 410–416.

[269] Pereira, D. M.; Rodrigues, P. M.; Borralho, P. M.; Rodrigues, C. M. P. *Drug Discovery Today* **2013**, *18*, 282–289.

[270] Xu, L.; Anchordoquy, T. *J Pharm Sci* **2010**, *100*, 38–52.

[271] Haussecker, D. *Mol Ther Nucleic Acids* **2012**, *1*, e8.

[272] Rossbach, M. *Curr Mol Med* **2010**, *10*, 361–368.

[273] Roy, R.; Hohng, S.; Ha, T. *Nat Methods* **2008**, *5*, 507–516.

[274] Blanco, M.; Walter, N. G. **2010**, *472*, 153–178.

[275] Murakoshi, H.; Iino, R.; Kobayashi, T.; Fujiwara, T.; Ohshima, C.; Yoshimura, A.; Kusumi, A. *Proc Natl Acad Sci U S A* **2004**, *101*, 7317–7322.

[276] Sakon, J. J.; Weninger, K. R. *Nat Methods* **2010**, *7*, 203–205.

[277] Pitchiaya, S.; Heinicke, L. A.; Custer, T. C.; Walter, N. G. *Chem Rev* **2014**, *114*, 3224–3265.

[278] Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. *Nat Biotech* **2003**, *21*, 86–89.

[279] Keppler, A.; Pick, H.; Arrivoli, C.; Vogel, H.; Johnsson, K. *Proc Natl Acad Sci U S A* **2004**, *101*, 9955–9959.

[280] Parikh, C.; Subrahmanyam, R.; Ren, R. *Cancer Res* **2007**, *67*, 7139–7146.

[281] Apolloni, A.; Prior, I. A.; Lindsay, M.; Parton, R. G.; Hancock, J. F. *Mol Cell Biol* **2000**, *20*, 2475–2487.

[282] Gao, J.; Liao, J.; Yang, G.-Y. *Am J Transl Res* **2009**, *1*, 312–325.

[283] Zverina, E. A.; Lamphear, C. L.; Wright, E. N.; Fierke, C. A. *Curr Opin Chem Biol* **2012**, *16*, 544–552.

[284] Arozarena, I.; Matallanas, D.; Berciano, M. T.; Sanz-Moreno, V.; Calvo, F.; Muñoz, M. T.; Egea, G.; Lafarga, M.; Crespo, P. *Mol Cell Biol* **2004**, *24*, 1516–1530.

[285] Landry, J. J. M.; Pyl, P. T.; Rausch, T.; Zichner, T.; Tekkedil, M. M.; Sttz, A. M.; Jauch, A.; Aiyar, R. S.; Pau, G.; Delhomme, N.; Gagneur, J.; Korbel, J. O.; Huber, W.; Steinmetz, L. M. *G3* **2013**, *3*, 1213–1224.

[286] Niforou, K. N.; Anagnostopoulos, A. K.; Vougas, K.; Kittas, C.; Gorgoulis, V. G.; Tsangaris, G. T. *Cancer Genomics Proteomics* **2008**, *5*, 63–77.

[287] Kedersha, N.; Tisdale, S.; Hickman, T.; Anderson, P.; Lynne, E. M.; Megerditch, K. **2008**, *448*, 521–552.

[288] Ryan, J. A. Cell cloning by serial dilution in 96 well plates.

[289] Pitchiaya, S.; Krishnan, V.; Custer, T. C.; Walter, N. G. *Methods* **2013**, *63*, 188–199.

[290] Margison, G. P.; Povey, A. C.; Kaina, B.; Santibáñez Koref, M. F. *Carcinogenesis* **2003**, *24*, 625–635.

[291] Watson, J. D. *Recombinant DNA*, 2nd ed.; Scientific American Books: New York, NY, 1998; pp 216–232.

[292] Warkocki, Z.; Odenwalder, P.; Schmitzova, J.; Platzmann, F.; Stark, H.; Urlaub, H.; Ficner, R.; Fabrizio, P.; Luhrmann, R. *Nat Struct Mol Biol* **2009**, *16*, 1237–1243.

[293] Fei, J.; Wang, J.; Sternberg, S. H.; MacDougall, D. D.; Elvekrog, M. M.; Pulukkunat, D. K.; Englander, M. T.; Gonzalez Jr, R. L.; Nils, G. W. **2010**, *472*, 221–259.

[294] Moore, M. J.; Sharp, P. A. *Science* **1992**, *256*, 992–997.

[295] Kershaw, C. J.; O'Keefe, R. T. In *Recombinant and In Vitro RNA Synthesis*; Conn, G. L., Ed.; Methods in Molecular Biology; Humana Press, 2012; Vol. 941; pp 257–269.

[296] Sarveswaran, K.; Hu, W.; Huber, P. W.; Bernstein, G. H.; Lieberman, M. *Langmuir* **2006**, *22*, 11279–11283.