EXPLORING THE FOLDING FREE ENERGY LANDSCAPE
OF SMALL RNA MOLECULES BY SINGLE-PAIR
FÖRSTER RESONANCE ENERGY TRANSFER

ANDREI YU. KOBITSKI
Institute of Biophysics, University of Ulm, Albert-Einstein-Allee 11
89081 Ulm, Germany

ALEXANDER NIERTH, MARTIN HENGESBACH,
ANDRES JÄSCHKE and MARK HELM
Institute of Pharmacy and Molecular Biotechnology, University of Heidelberg
Im Neuenheimer Feld 364, Heidelberg, 69120, Germany

G. ULRICH NIENHAUS
Institute of Biophysics, University of Ulm, Albert-Einstein-Allee 11,
89081 Ulm, Germany
Department of Physics, University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA
uli@uiuc.edu

Received 10 June 2008
Revised 26 September 2008

Proteins and RNA are biological macromolecules built from linear polymers. The process
by which they fold into compact, well-defined, three-dimensional architectures to perform
their functional tasks is still not well understood. It can be visualized by Brownian
motion of an ensemble of molecules through a rugged energy landscape in search of
an energy minimum corresponding to the native state. To explore the conformational
energy landscape of small RNAs, single pair Förster resonance energy transfer (spFRET)
experiments on solutions as well as on surface-immobilized samples have provided new
insights. In this review, we focus on our recent work on two FRET-labeled small RNAs,
the Diels-Alderase (DAse) ribozyme and the human mitochondrial tRNA Lys. For both
RNAs, three different conformational states can be distinguished, and the associated
mean FRET efficiencies provide clues about their structural properties. The systematic
variation of their free energies with the concentration of Mg^{2+} counterions was analyzed
quantitatively by using a thermodynamic model that separates conformational changes
from Mg^{2+} binding. Furthermore, time-resolved spFRET studies on immobilized DAse
reveal slow interconversions between intermediate and folded states on the time scale of
∼100 ms. The quantitative data obtained from spFRET experiments may likely assist
in the further development of theories and models addressing the folding dynamics and
(counterion-dependent) energetics of RNA molecules.

Keywords: Fluorescence microscopy; single molecule detection; Förster resonance energy
transfer (FRET); ribonucleic acids; RNA folding; RNA thermodynamics.
1. Introduction

Proteins are linear polymers of amino acids the sequence of which is encoded in the genome. To understand how they fold into well-defined, three-dimensional structures has been a challenging task for decades. More recently, concepts borrowed from the statistical physics of polymers and glasses have led to the description of protein folding as Brownian motion of protein molecules in a complex, rugged energy landscape with many local minima separated by energy barriers.\textsuperscript{1,2} The overall shape of this landscape resembles a funnel that guides the molecules toward a free energy basin corresponding to a multitude of properly folded, native states.\textsuperscript{3,4} Similar concepts have also been applied to the problem of ribonucleic acid (RNA) folding in recent years.\textsuperscript{5} Like proteins, RNA molecules also fold into a compact three-dimensional structure, and self-assembly, dynamics and function of RNA can be visualized in terms of an energy landscape. However, there are also crucial differences that arise from the different physico-chemical properties of both biopolymers. RNA is a polyanion, and an understanding of its interaction with mono- or divalent counterions is key to understanding its thermodynamic stability.\textsuperscript{6} At low counterion concentration or high temperature, it fluctuates among all possible conformations. Under conditions favorable for folding, a counterion-mediated collapse is observed, and base pairing of complementary nucleotides governs the formation of locally stable structural motifs such as hairpins, bulges, and loops. These elements, which by themselves form autonomous units, subsequently assemble into larger tertiary structures.\textsuperscript{7} The folding process is thus hierarchical in nature, and relatively stable secondary structures may even form misfolded configurations prior to attaining the native structure. Apparently, structural, dynamic and energetic issues need to be addressed for a thorough understanding of RNA folding.

Experiments at the single-molecule level, especially by fluorescence or force microscopy, have provided novel insights into the structure and dynamics of biomolecules (Ref. 8 and references therein). Its main advantage lies in the fact that different subpopulations within a molecular ensemble can be resolved and investigated selectively. Single-pair Förster (or fluorescence) resonance energy transfer (spFRET) is a structure-sensitive spectroscopic technique that allows conformational changes to be probed in real time. Based on the coupling of optical transition dipoles of two dye molecules, referred to as the donor and the acceptor, attached to the biomolecule of interest with a mutual separation of \(\sim 20–100\) Å, spFRET is sensitive to inter-dye distance variations in the ångström range. Thus, the technique is well suited to monitor conformational motions, for example the transition of small and medium-sized RNA molecules from a completely unfolded random coil to the compact folded state. Over the past decade, single-molecule FRET studies have been performed on a variety of RNA and DNA molecules, such as an RNA three-helix junction,\textsuperscript{9,10} Tetrahymena group I intron,\textsuperscript{11,12} hairpin\textsuperscript{13–20} and hammerhead\textsuperscript{21} ribozymes, ribonuclease P,\textsuperscript{22,23} and the 8–17 deoxyribozyme.\textsuperscript{24} Extensive investigations of ribozymes revealed highly complex folding kinetics of substrate docking...
and undocking and cleavage/ligation reactions even for fairly small RNA molecules such as the hairpin ribozyme.\textsuperscript{25–28}

However, our understanding of the mechanistic relationships between RNA structure, folding and function is far from complete. Here, we address thermodynamic issues of truly intramolecular RNA folding without interference from docking or undocking events. Small RNA molecules are in the focus of this review; they serve as model systems to study the details of an energy landscape without having to deal with the complication of major kinetic traps.

\section{Single-Pair FRET}

\subsection{Design of FRET constructs}

To observe reversible folding–unfolding transitions of RNA molecules by single-pair (sp)FRET, covalent attachment of donor and acceptor dyes to the same RNA strand is required. An additional anchor is needed in the case of experiments on surface immobilized molecules. Even for small RNAs, the design of FRET-labeled constructs is a challenging task, since several prerequisites have to be met. The attached fluorophores should not interfere with the formation of any intermediate or folded states, and the natural function of the biomolecule should remain unperturbed. For example, large hydrophobic dye moieties can hinder or even block access to catalytic pockets. Therefore, available structural information has to be taken into careful consideration. One possibility to minimize interferences is the placement of dyes in helical regions, where they protrude out of the grooves. Incorporation of flexible linkers promotes a maximum degree of rotational freedom of the dyes. However, the choice of the attachment sites is limited by conjugation chemistry and commercial availability. For example, postsynthetic coupling by NHS chemistry of fluorophores to primary amines has proven to be a reliable method. Available nucleotide building blocks for RNA solid phase synthesis include deoxy- or ribothymidine and deoxyguanosine, which carry amino linkers at the C5 and C8 atoms of the respective bases. As a general strategy, attachment sites are chosen such that maximum changes in the FRET efficiency between folded and intermediate states can be anticipated on the basis of the available structural information. Further fine-tuning of the FRET efficiency can be accomplished either by slightly varying the dye attachment sites or by exchanging the dyes to obtain a more suitable Förster radius.

The two RNAs which are discussed in this work, namely the Diels-Alderase (DAse) ribozyme and the unmodified human mitochondrial lysine transfer RNA (human mt tRNA\textsuperscript{Lys}), are shown in Fig. 1. The design and construction by splint ligation of the dye-labeled constructs have been described in detail in Refs. 29 and 30. Briefly, RNA synthesis started from short oligoribonucleotides, synthesized by customary solid phase chemistry. After annealing on a complementary splint DNA template, RNA fragments were joined by T4 DNA ligase. After removal of the DNA by DNase I, RNA constructs were purified by denaturing PAGE, excised,
eluted and precipitated using standard procedures. Complementary DNA covering the full length of the target RNA is the key ingredient in the synthesis because the sequence of the DNA splint is used to align the fragments in their correct order, akin to pearls on a string, so as to assemble multiple fragments in a single enzymatic step. This approach is flexible, as single fragments may be exchanged at will, allowing the combinatorial synthesis of various constructs carrying virtually any combination of dyes and/or other modifications, e.g. biotin for surface immobilization. For both RNAs, Cy3 and Cy5 dyes were placed at various positions within the RNA sequences. The resulting panel of constructs was screened for suitability in spFRET measurements and for possible interference of the dyes with the regular structure and function of the RNAs. For the DAse ribozyme, the catalytic activity of the constructs was investigated in the Diels-Alder reaction of anthracene with N-pentyl-maleimide. For tRNA$^{Lys}$, the substrate properties of dye-labeled derivates toward maturation enzymes were compared to unlabeled controls.

2.2. Fluorescence microscopy apparatus

A variety of single-molecule techniques based on fluorescence microscopy or manipulation (force) have been developed in recent years.$^8,31^{–}33$ Single-molecule sensitivity in fluorescence microscopy is achieved by rigorous reduction of the detection volume either by introducing a confocal pinhole in the detection path (confocal...
microscopy) or by exciting molecules in a ∼100-nm layer close to a surface by using an evanescent field (total internal reflection fluorescence microscopy). Both methods are complementary. Whereas confocal microscopy offers a very high sensitivity and time resolution but is time-consuming, as the measurements are carried out one molecule at a time, total internal reflection microscopy allows data on large ensembles to be measured in parallel. Our laser scanning confocal setup based on a Zeiss Axiovert 135 TV inverted microscope (Zeiss, Germany) is shown schematically in Fig. 2(a). A detailed description can be found elsewhere.34–36 Briefly, various lasers are available as excitation sources; the excitation of the donor in the green spectral range is provided by an Ar$^+$ ion laser (514.5 nm) or a diode-pumped solid state (Nd:YAG) laser (532 nm); the red line of a He–Ne laser (633 nm) serves to directly excite the acceptor dye. An acousto-optical tunable filter (AOTF) is a convenient means to select the laser lines and also to alternate between them on a fast time scale if an alternating laser excitation (ALEX) scheme is used.37 Insertion of a single-mode fiber into the excitation path provides a gaussian-shaped laser beam profile. The fluorescence emitted by the sample is collected by a water immersion objective and passed through a confocal pinhole. A set of high-quality dichroic mirrors and filters rejects unwanted light and splits the emitted light into two detection channels. Careful design of the filters, taking into account the optical properties of the donor and acceptor dyes, ensures that the light is efficiently collected while maintaining minimal crosstalk between the two color channels. Finally, photons are detected by two avalanche photodiodes (APDs) and the events are registered using a data acquisition card in the computer.

2.3. Solution versus surface measurements

Confocal laser scanning fluorescence microscopy offers two types of single-molecule experiments, either on molecules diffusing freely in solution (Fig. 2(b)) or being immobilized on a surface (Fig. 2(c)). In the case of freely diffusing molecules, the signal from each individual molecule is collected during its brief sojourn in the sensitive volume. This volume is on the order of ∼1 fl, and fluorescence bursts persist for typically 0.1–1 ms for small RNA molecules. An advanced excitation scheme, consisting of rapid alternation (10–100 µs) between donor and acceptor excitation on a time scale faster than the diffusion time, allows the rejection of molecules with a non-functional acceptor in the subsequent analysis.38 The sample concentration has to be adjusted to ∼100–300 pM to minimize the presence of more than one molecule in the detection volume at the same time. Otherwise, the calculated FRET value would be an intensity-weighted average of two or more molecules, which would obviously be detrimental for the analysis. Thousands of bursts from individual molecules can be registered conveniently within 10–30 min, so that good statistics can be obtained. There is a drawback, however, because only 30 to 100 photons are typically detected from a single burst. Therefore, the shot-noise contribution to the FRET efficiency value is substantial, which gives rise to a broad
distribution rather than a sharp value. Moreover, the method is only sensitive to FRET changes due to dynamics on time scales shorter than the burst length; processes on longer time scales remain obscured. Surface immobilization allows the time of observation to be extended up to several minutes under continuous excitation, until the fluorophores fall victim to photobleaching. This interval, during
which \( \sim 10^4 - 10^5 \) photons are registered, can be further extended by using time-lapse illumination.\(^{39}\) However, surface immobilization can be challenging because the biomolecules should be specifically and stably tethered while exhibiting minimal non-specific interactions with the surface. This requirement has been difficult to meet in single-molecule protein folding studies,\(^ {40}\) which led us to introduce surface coatings using cross-linked star-shaped polyethylene glycol molecules for this application.\(^ {34, 41, 42}\) For RNA, these problems appear to be less severe and, consequently, we have as yet immobilized our RNA molecules using surface coatings that are easy to prepare on the microscope. In a first step, biotinylated bovine serum albumin (BSA) is physisorbed to a (hydrophilic) glass cover slip. Subsequently, the surface is exposed to streptavidin, a protein which strongly binds to biotin via one of its four specific binding sites. Finally, biotin-functionalized, FRET pair-labeled RNA molecules are attached to the surface via vacant biotin binding sites on the streptavidin molecules. A sparse coverage with RNA molecules, necessary to ensure that only a single FRET pair-labeled molecule is contained within the spot size (point spread function) of the microscope, is achieved by incubating with highly dilute RNA solutions (\( \sim 10 \) pM) for a few minutes. Images are acquired by continuous scanning of the sample across the confocal spot by using a piezoelectric stage. To select molecules with a functional dye pair, the same areas are scanned with donor excitation and a second time with acceptor excitation. To measure time traces of the fluorescence emission, molecules are selected automatically from the images using acceptor excitation and are subsequently moved into the focus. Then, the donor is excited and donor and acceptor fluorescence emission is recorded until both donor and acceptor are photobleached.

### 2.4. FRET histogram analysis and structural assignments of states

FRET histograms are compiled from FRET efficiency values calculated for each individual burst (in solution measurements) or spot (in the analysis of images of immobilized molecules). The FRET efficiency, \( E \), is calculated ratiometrically from the fluorescence intensities registered from the acceptor, \( I_A \), and the donor, \( I_D \), respectively, corrected for background and cross-talk between the channels,

\[
E = \frac{I_A}{I_A + \gamma I_D}.
\]

and \( \gamma \) is an experimental parameter that accounts for differences in the detection efficiencies of the two channels and quantum yields of the dyes. In addition, the fluorescence emission of acceptor upon direct excitation, \( I_A^{\text{Dir}} \), is measured to calculate the stoichiometry ratio, \( S \), defined as

\[
S = \frac{I_A + \gamma I_D}{I_A + \gamma I_D + I_A^{\text{Dir}}}.
\]
The excitation intensities of the green and red lasers are adjusted such that $S$ takes a value of 0.5 for a properly labeled molecule, and $S \approx 1$ (0) if the acceptor (donor) is missing.

Typical FRET histograms of the DAse ribozyme and tRNA$_{\text{Lys}}$ are shown for three different Mg$^{2+}$ concentrations in Fig. 3(a). For both RNAs, three subpopulations can be distinguished, located at low, intermediate and high FRET efficiencies. For the DAse ribozyme, these subpopulations have been assigned to the folded state ($F$), with $\langle E_F \rangle \approx 0.9$, an intermediate state ($I$), with $\langle E_I \rangle \approx 0.7$ and an unfolded state ($U$), with $\langle E_U \rangle \approx 0.35$. An unfolded state $U$ was also observed for tRNA$_{\text{Lys}}$.

The other two states at about 0.4 and 0.8 were assigned to a more extended hairpin ($E$) structure, as was suggested by earlier bulk studies, and to the native, highly compact, cloverleaf-based L-shape structure ($C$), respectively.

Structural information can be extracted from FRET measurements due to the connection between a spectroscopic observable, $E$, and a structural parameter, the inter-dye distance, $r$,

$$ E = \frac{R_0^6}{r^6 + R_0^6}. $$

The Förster radius, $R_0$, is a characteristic distance at $E = 0.5$. This equation, however, applies only to cases where both $r$ and $R_0$ are constant during measurements. It amounts to $\sim 53 \, \text{Å}$ for the Cy3/Cy5 dye pair (in aqueous solution) used in our studies, assuming that the dye orientation is averaged over the donor fluorescence lifetime. In biopolymers including proteins or RNA, the polymer chain is flexible, especially in the unfolded states, and dye reorientation is often geometrically restricted or hindered by interactions between the dye and the biopolymer. In this case, the FRET efficiency becomes an average over all possible realizations of the polymer chain and dipole–dipole couplings of the dyes,

$$ \langle E \rangle = \int_0^{l_c} \int_0^4 \frac{1}{1 + \frac{2\kappa^2}{\pi^2} \left( \frac{r^2}{R_0^2} \right)^6} P(r) P(\kappa^2) d\kappa^2 dr. $$

Here, the dipolar coupling is quantified by the orientation factor $\kappa^2$, $P(r)$ and $\rho(\kappa^2)$ are probability densities, and $l_c$ is the contour length of the polymer chain. As the probability densities are usually not a priori known, they have to be based on simple, statistical assumptions, and the calculated inter-dye distances should, therefore, be taken with a grain of salt.

For well-ordered states of biomolecules, however, a constant $r$ appears to be a viable approximation. To compare measured FRET distances with distances extracted from structural models, one also has to take into account that the dye molecules are usually connected to the attachment points with fairly long tethers, which may or may not contribute to the inter-dye distance, depending on the individual geometries. With dye-labeled DNA constructs, we had earlier found that the Förster radius is close to the fixed spatial separation of the attachment points.
Fig. 3. FRET histogram analysis. (a) Histograms of FRET efficiency values measured on immobilized DAse ribozyme (left) and freely diffusing tRNA\textsuperscript{Lys} (right) molecules in buffer solutions at three different Mg\textsuperscript{2+} concentrations. Dotted, dashed and solid thin lines represent best-fit model distributions for the three conformational states for the DAse ribozyme and tRNA\textsuperscript{Lys} molecules, respectively; the solid thick line gives the sum of the three distributions. (b) Mg\textsuperscript{2+} dependence of the fractional populations for each conformational state of the DAse ribozyme (left) and tRNA\textsuperscript{Lys} (right). The lines represent fractional populations obtained by fits based on the thermodynamic models depicted in Fig. 4 to the data.
This result was explained by the fact that the tethered dyes can fluctuate freely around these points.\textsuperscript{45}

For the folded (\(F\)) state of the DAs e ribozyme, the measured \(E\) corresponds to \(r = 34 \text{ Å}\) according to Eq. (3). In the DAs e ribozyme X-ray structure,\textsuperscript{47} the distance between the C5 atoms on the two modified uridine bases, to which the dyes are affixed, is 22 Å. The dye molecules were connected via 16-atom flexible spacers, which each endow them with freedom to fluctuate by \(\sim 7 \text{ Å}\) (treating the linkers as Gaussian chains) around the attachment points. The linkers appear to contribute significantly to the overall inter-dye distance in the DAs e ribozyme. For the \(C\) state of tRNA\textsuperscript{Lys}, the measured \(E\) corresponds to \(r = 46 \text{ Å}\), whereas the crystal structure of yeast tRNA\textsuperscript{Phe} yields a very similar distance of 42.8 Å between the C5-atoms of nucleotides 4 and 41.\textsuperscript{48} Here, the 11-atom-chain linkers do not appear to contribute much to the overall inter-dye separation. Regardless of the uncertainties involved, these estimations provide strong support for the assignment of the \(F\) and \(C\) states to the properly folded DAs e ribozyme and cloverleaf-based L-shape tRNA\textsuperscript{Lys} molecules, respectively.

Even though structural information is not available \textit{a priori} for the \(U\) state, it presumably has minimal residual structure and can be modeled as a semi-flexible, self-avoiding polymer chain characterized by an ionic strength-dependent persistence length, \(l_p\).\textsuperscript{49} The corresponding model is the worm-like chain (WLC) introduced by Kratky and Porod.\textsuperscript{50} Under the conditions chosen in our work, \(l_p\) is in the range 10–21 Å, which is much less than the contour length of the molecules. Therefore, the simpler Gaussian chain model is also valid and, by using \(P(r)\) from this model and assuming slow rotation of the dyes through all mutual orientations, \(\langle E \rangle\) can be calculated numerically with Eq. (4) and approximated by an analytical function.\textsuperscript{46}

With this calculation, the measured FRET efficiency \(\langle E \rangle = 0.06\) for the \(U\) state corresponds to an inter-dye distances of 151 Å for the DAs e ribozyme, whereas the simple Förster expression, Eq. (3), yields 84 Å, with rather large errors due to the weak distance dependence at low FRET efficiencies.\textsuperscript{45} For comparison, a theoretical calculation based on the Gaussian chain with a contour length between the dye attachment sites, \(l_{DA} = 246 \text{ Å}\), would yield an average inter-dye distance, \(R_a = 102\text{ Å}\). For tRNA\textsuperscript{Lys}, the measured FRET efficiency \(\langle E \rangle = 0.25\) corresponds to an inter-dye distance of 81 Å with Eq. (4). The Gaussian chain model yields (accidentally) the same value as for the DAs e ribozyme, 102 Å. In both cases, the assignment of the low-FRET distributions to unfolded RNA chains appears appropriate. Extraction of reliable inter-dye distances is problematic at low \(E\), however, and the distributions are broad and extend to high FRET values (Fig. 3(a)), indicating that the \(U\) ensemble contains RNA molecules with some secondary structure.

Most challenging is the assignment of intermediate states, as they contain some secondary and tertiary structure which little may be known about. Here, results from other measurements may be utilized to guess possible conformations. For example, comparative solution mapping studies of tRNA\textsuperscript{Lys} revealed that the RNA
can form an extended hairpin consisting of three helices connected by two loops.\textsuperscript{43} Moreover, it was found by transient electric birefringence measurements that the acceptor and anticodon stems form an angle of $\sim 140^\circ$.\textsuperscript{51} Based on these data and by taking a distance of 3.4 Å and 6.3 Å for each of the nucleotides in helices and loops, respectively, we estimated a total distance between the dye attachment sites of $\sim 75$ Å, what is in reasonable agreement with the distance of 65 Å calculated from the FRET efficiency value of $\langle E \rangle = 0.37$ based on the Gaussian chain relation.

3. Thermodynamics of RNA

3.1. Fractional populations of states

To quantify the change of fractional populations with Mg\textsuperscript{2+} concentration, we fitted the FRET efficiency distributions corresponding to the different conformations by continuous functions, log-normal functions for the low- and high-FRET states, and a standard Gaussian for the state at intermediate FRET efficiency. The use of the asymmetric log-normal function is motivated by the high curvature of the $E(r)$ dependence at extreme distances, which changes a symmetric distribution of inter-dye distances, $P(r)$, in real space into a highly asymmetric one in $E$ space. We further note that an asymmetric probability density of inter-dye distances also appears when modeling the $U$ state of RNA as a flexible polymer chain. For inter-dye separations close to the Förster distance, the relationship between $E$ and $r$ is essentially linear; therefore, the $E(r)$ dependence will not introduce a pronounced distortion of the $P(r)$ distribution. It is important to realize that experimental limitations such as shot noise of the detected fluorescence contribute significantly to the broadening of the observed distributions, thus hiding possible asymmetries in $P(r)$. Consequently, an approximation of the intermediate state with a Gaussian function appears sensible. The resulting individual and overall distributions are shown in Fig. 3(a). All distributions overlap significantly, which precludes a strict separation based on the FRET efficiency histograms alone. However, we found that the mean and width of the FRET efficiency distributions of the structurally well-defined states $F$, $E$ and $C$ are all — within the experimental error — invariant with respect to Mg\textsuperscript{2+} ion concentration, which appears reasonable. Consequently, to achieve a reliable separation, we fixed the positions and widths of these distributions and varied only their populations in a simultaneous, global fit of the histograms at all Mg\textsuperscript{2+} ion concentrations. State $U$ is mostly unstructured and the shape of the FRET efficiency distribution could in principle depend on the Mg\textsuperscript{2+} ion concentration, but changes were not resolvable under conditions used in our experiments. An interesting behavior is seen for the distribution of state $I$ of the DAse ribozyme (Fig. 3(a), left). Its mean, $\langle E_I \rangle$, clearly shifts to higher values with increasing Mg\textsuperscript{2+} concentration, indicating a compaction of the overall ensemble of conformational substates in this state.

In Fig. 3(b), the relative populations of the $U$, $I$, and $F$ states of the DAse ribozyme and $U$, $E$, and $C$ states of tRNA\textsuperscript{Lys} are plotted as functions of the Mg\textsuperscript{2+}
ion concentration. For the DAse ribozyme, a pronounced increase in the folded fraction with Mg$^{2+}$ was revealed at the expense of the two other states, $U$ and $I$. For tRNA$^{\text{Lys}}$, a pronounced drop of the $U$ state population was observed with increasing Mg$^{2+}$ concentration, with a midpoint at $\sim0.5$ mM, accompanied by an increase of the $E$ and $C$ state populations. At higher concentrations ($\sim100$ mM), a second transition is evident in which the $C$ state increases at the expense of the $E$ state.

3.2. Thermodynamic models

To analyze the Mg$^{2+}$ dependence of the RNA conformations, we have employed the thermodynamic framework developed by Draper and coworkers, in which the Mg$^{2+}$-induced RNA folding reaction was decomposed into an RNA folding reaction and a Mg$^{2+}$ binding reaction in each state. The corresponding six-state thermodynamic schemes are depicted in Fig. 4, where subscripts “0” and “Mg” correspond to Mg-free and Mg-bound conformations within the states. Both schemes are similar; the key difference is the unfavorable transition from $U$ to $F$ in the DAse ribozyme.

We also found that the ratio between the populations $U$ and $I$ and, thus, their free energy difference, $\Delta G_{UI}$, remained constant within the error upon changing Mg$^{2+}$. Therefore, the $U$ and $I$ states can be combined to a single state when modeling the Mg$^{2+}$ ion dependence. This treatment reduces the scheme in Fig. 4(a) to a four-state model. In the case of tRNA$^{\text{Lys}}$, the presence of three highly populated states, $U$, $E$ and $C$ and their complicated behavior upon varying the Mg$^{2+}$ concentration makes it necessary to consider all equilibria without additional simplifications.

![Thermodynamic schemes](image)

Fig. 4. Thermodynamic schemes describing the equilibria between the Mg-free and Mg-bound forms of the conformational states for the (a) DAse ribozyme and (b) human mt tRNA$^{\text{Lys}}$. 
The relative populations of the structurally different states are governed by equilibrium coefficients, \( K^0_{ij} \):

\[
K^0_{ij} = \frac{[j_0]}{[i_0]} = \exp \left( -\frac{\Delta G^0_{ij}}{RT} \right).
\]

Here, the brackets denote fractional populations, \( \Delta G^0_{ij} \) is the free energy difference between the Mg-free states \( i \) and \( j \), and \( R \) and \( T \) represent the gas constant and the absolute temperature, respectively. The Mg-free and Mg-bound fractional populations are described by the Mg\(^{2+}\)-dependent relation:

\[
K_i = \frac{[i_0]}{[i_{Mg}]} = \exp \left( -\frac{\Delta G_i(Mg)}{RT} \right).
\]

A general approach to describe these interactions without making any assumptions on their nature was introduced by Grilley et al.\(^{54}\) The free energy of the Mg\(^{2+}\)-RNA interaction at constant concentration of monovalent ions is given by

\[
\Delta G(Mg) \approx RT \int_0^{[Mg^{2+}]} \Gamma_{2+} d \ln [Mg^{2+}],
\]

where \( \Gamma_{2+} \) is the difference in the number of ions in proximity and far away from the RNA for a given bulk Mg\(^{2+}\) concentration. The simplest approximation is to consider only strong interactions and, thus, specific ligand binding. Then, \( \Gamma_{2+} \) is equivalent to the cooperativity parameter (Hill coefficient) of the transition, \( n \), which controls the sharpness of the transition and relates to the number of binding sites occupied in the cooperative transition.\(^{55}\) Hence, the free energy difference between Mg\(^{2+}\)-bound and unbound states can be written as

\[
\Delta G_i(Mg) \approx \Delta G^0_i + n_i RT \cdot \ln [Mg^{2+}],
\]

where \( \Delta G^0_i \) represents the standard free energy of each state \( i \). We note that our FRET experiment can only distinguish conformational states, but not Mg-bound and Mg-free conformations within each state; therefore, only the sums \([i_0 + i_{Mg}]\) can be accessed. The equilibrium between the observed states can be written as

\[
[i_0 + i_{Mg}] = \frac{1}{m - 1} \sum_{j \neq i}^{m-1} \left( [j_0 K^0_{ji} + j_{Mg} K^M_{ji}] \right),
\]

where \( m \) is the number of conformational states (2 and 3 for the DAse ribozyme and tRNA\(^{Lys}\), respectively). The global fit was performed by solving the set of linear equations, Eqs. (9); the results are plotted in Fig. 3(b) as lines.

The Hill coefficient of \( \sim 2 \) for the \( I \leftrightarrow F \) transition of the DAse ribozyme correlates well with the finding that two Mg\(^{2+}\) ions are crucial for proper folding.\(^{47}\) Thus, from six Mg\(^{2+}\) ions identified in the X-ray structure, Mg1 and Mg2 were found to be involved in the packing of helices I and II onto each other by forming multiple direct and water-mediated bonds. Interestingly, the compaction of the \( I \) state, which occurs at slightly lower Mg\(^{2+}\) concentration with respect to \( I \leftrightarrow F \) transition, is...
also strongly cooperative and characterized by a Hill coefficient of ∼4. For tRNA\textsuperscript{Lys}, Hill coefficients were found to be 0.7, 1.1 and 1.2 for the \textit{U}, \textit{E} and \textit{C} states. As one may have expected, the weakest interaction with Mg\textsuperscript{2+} ions is observed for the unfolded state, and the strongest for the most compact \textit{C} conformation. For both RNA molecules, the compaction of the Mg-free states increases the free energy due to electrostatic repulsion, which is compensated by binding of Mg\textsuperscript{2+} ions.

### 3.3. Structural dynamics of conformations

FRET histogram analysis reveals conformational heterogeneity, which may be either static or dynamic. While static heterogeneity is due to an inhomogeneous environment on the time scale of the experiment, dynamics is of particular relevance because essentially all biomolecular functionality is closely connected to conformational motions.\textsuperscript{56} The study of single-molecule FRET time trajectories allows the structural dynamics to be examined under equilibrium conditions. Heterogeneous structural dynamics has been shown to be an intrinsic property of medium-sized and large RNA molecules,\textsuperscript{26} and also small RNAs display this property. For a surface-immobilized DAse ribozyme molecule, FRET efficiency fluctuations between the intermediate and folded states at 5-mM Mg\textsuperscript{2+} concentration are plotted as a function of time (50-ms resolution) in Fig. 5(a), together with a FRET histogram that depicts the relative probabilities with which the different \textit{E} values occur in the trace. For the quantitative analysis of time trajectories, two different approaches may be employed, dwell-time or correlation analysis. Dwell-time analysis is more general and allows the folding/unfolding transition rates, \(k_f\) and \(k_u\), to be determined independently. The challenge of the method is to unambiguously resolve transitions between discrete FRET states by setting certain threshold levels. Recently, some sophisticated mathematical methods for identifying transitions within FRET trajectories have been developed and successfully applied to proteins and RNAs.\textsuperscript{39,57,58} For FRET transitions within two-level systems, straightforward correlation analysis can be applied, which results in an exponentially decaying cross-correlation function with a negative amplitude.\textsuperscript{10} Even for shot noise-dominated fluorescence fluctuations, the apparent rate coefficient, \(k_{app}\), can be extracted with high accuracy by counting a large number of photons. The cross-correlation function corresponding to the trace in Fig. 5(a) is plotted in Fig. 5(b). Its negative amplitude reflects the anticorrelated intensity variations of donor and acceptor due to conformational motions occurring on the 100-ms time scale. The apparent rate coefficients, \(k_{app}\), of many individual molecules were calculated from the corresponding cross-correlation curves and compiled into the histogram presented in Fig. 5(c). A rather broad distribution results, which may in part also arise from the heterogeneous surroundings that endow each molecule with a somewhat different rate coefficient. Nevertheless, the Mg\textsuperscript{2+} concentration governs the average time that the molecule spends in each conformational state. So, the mean apparent rate coefficients calculated from the distributions, taken at different Mg\textsuperscript{2+} concentrations, produce the chevron-shaped
curve seen in Fig. 5(d). Moreover, the width of distributions, represented by the error-bars in Fig. 5(d), becomes much narrower at very low and very high concentrations, which may indicate that the observed heterogeneous dynamics is intrinsic to the RNA molecule.

The dependence of the transition rate coefficients on the Mg\(^{2+}\) ion concentration can be analyzed with a two-state thermodynamic scheme. The microscopic rate coefficients, \(k_{ij}^0\) and \(k_{ji}^0\) (\(k_{ij}^{\text{Mg}}\) and \(k_{ji}^{\text{Mg}}\)), describe the transitions between the Mg-free (Mg-bound) forms \(i \rightarrow j\) and \(j \rightarrow i\), respectively. These rate coefficients are connected to the equilibrium coefficients \(K_{ij}^0\) and \(K_{ij}^{\text{Mg}}\) by the relations

\[
K_{ij}^0 = \frac{k_{ji}^0}{k_{ij}^0}, \quad K_{ij}^{\text{Mg}} = \frac{k_{ji}^{\text{Mg}}}{k_{ij}^{\text{Mg}}}.
\] (10)

![Fig. 5. (a) FRET efficiency time trace of an individual ribozyme molecule measured at 5-mM Mg\(^{2+}\) concentration and the corresponding histogram of FRET efficiencies, obtained with 50-ms dwell time binning. (b) Cross-correlation function calculated from the FRET trace in panel (a). (c) Histogram of apparent rate coefficients for an ensemble of DAse molecules, all at 5-mM Mg\(^{2+}\) concentration, binned on a logarithmic scale. (d) Means (open circles) and widths (error bars) of the apparent rate coefficient distributions are shown as a function of the Mg\(^{2+}\) concentration. The solid black line represents a curve calculated with the model discussed in the text. The gray lines represent the fractional contributions of \(k_u\) and \(k_f\) to \(k_{\text{app}}\).](image-url)
Then, the overall folding and unfolding rate coefficients, $k_f$ and $k_u$, are calculated according to Kim et al.\textsuperscript{10} as

$$k_f = \frac{K_i k_0^{ij} + k_{Mg}^{ij}}{K_i + 1}, \quad k_u = \frac{K_j k_{Mg}^{ji} + k_0^{ji}}{K_j + 1} \quad \text{and} \quad k_{app} = k_f + k_u. \quad (11)$$

Note that $k_f$ and $k_u$ depend on the Mg$^{2+}$ concentration through the equilibrium coefficients $K_i$ and $K_j$ (Eqs. (6) and (7)). We have applied this approach to the transitions between the $I$ and $F$ states of the DAse ribozyme, taking the equilibrium coefficients from the population data in Fig. 3(b) (left panel). The resulting Mg$^{2+}$ dependences of the rate coefficients $k_u$, $k_f$ and $k_{app}$ are depicted in Fig. 5(d). The simulated curve of $k_{app}$, which was scaled to pass through the point at 40 mM by adjusting the value of $k_{Mg}^{FI}$ to $\sim 40 \text{ s}^{-1}$, nicely describes the Mg$^{2+}$ concentration dependence of the experimental points. From the individual curves for $k_u$ and $k_f$, it is obvious that $k_u$ ($k_f$) mainly contributes at low (high) Mg$^{2+}$ concentrations to $k_{app}$; close to the midpoint concentration of the $I$–$F$ transition, they intersect and contribute equally.

4. Conclusion

Even rather small RNA molecules display a complicated folding behavior, including the formation of intermediate states with their individual secondary and tertiary structure features. The equilibrium and dynamics between unfolded, intermediate and folded states are governed by a complex free energy landscape which can be shaped by structural modifications and also by variation of the environmental conditions such as the concentration of counterions.

An important aspect in the understanding of RNA folding is the contribution of tertiary interactions to the stabilization of the three-dimensional fold. In this context, point mutations or even small modifications of certain nucleotides may cause large changes in the energy landscape. As an example, a single base change in the P3 domain of the Tetrahymena ribozyme disfavored the formation of the misfolded state and increased the folding rate to the native state as much as 50-fold.\textsuperscript{59} Another example is a single methyl group modification in human mitochondrial tRNA\textsuperscript{Lys}, which suffices to shift the thermodynamic equilibrium from a non-functional, extended hairpin conformation to a functional form.\textsuperscript{36,60} Quantitative information extracted from such experiments will be most interesting in comparison with computational approaches aimed at simulating the dynamics in the complex RNA energy landscape.\textsuperscript{61–63}

In studies of RNA energy landscapes, the power of single-molecule techniques has become very obvious. Especially the sensitivity of FRET to minute inter-dye distance changes has allowed discrete conformational states to be identified, and their structures and dynamics have been characterized individually and quantitatively. Variation of the counterion concentration has been a very useful approach in these studies, as specific binding of cations to each particular conformation governs
their energetics and allows the relative populations to be adjusted over wide ranges. Periodic counterion concentration jump experiments have recently allowed conformational states to be probed that are not easily detected under equilibrium conditions. A drawback of the FRET technique is that cation-bound and cation-free states cannot be directly distinguished. This problem can at least be partially addressed by simultaneous measurement of single-molecule FRET and monitoring the concentration of free cations by an indicator dye, for example, a Mg$^{2+}$ chelator such as 8-hydroxyquinoline-5-sulfonic acid (HQS). Clearly, single-molecule studies of RNA structure and dynamics are still at the beginning, and there is ample opportunity to further advance the experimental procedures in various ways so as to discover more and more details of their complex conformational energy landscapes.

**Acknowledgments**

Financial support from the Volkswagen Foundation, Deutsche Forschungsgemeinschaft (DFG, grants NI291/9, JA 794/3 and HE 3397/3), Fonds der Chemischen Industrie and the Landesgraduiertenförderung Baden-Württemberg is gratefully acknowledged.

**References**