Review
Focus on Function: Single Molecule RNA Enzymology

Mark A. Ditzler,[1] Elvin A. Alemán,[2] David Rueda,[2] Nils G. Walter[3]

[1] Biophysics Research Division, Single Molecule Analysis Group, University of Michigan, Ann Arbor, MI 48109

[2] Department of Chemistry, Wayne State University, Detroit, MI 48202

[3] Department of Chemistry, Single Molecule Analysis Group, University of Michigan, Ann Arbor, MI 48109

ABSTRACT:

The ability of RNA to catalyze chemical reactions was first demonstrated 25 years ago with the discovery that group I introns and RNase P function as RNA enzymes (ribozymes). Several additional ribozymes were subsequently identified, most notably the ribosome, followed by intense mechanistic studies. More recently, the introduction of single molecule tools has dissected the kinetic steps of several ribozymes in unprecedented detail and has revealed surprising heterogeneity not evident from ensemble approaches. Still, many fundamental questions of how RNA enzymes work at the molecular level remain unanswered. This review surveys the current status of our understanding of RNA catalysis at the single molecule level and discusses the existing challenges and opportunities in developing suitable assays. © 2007 Wiley Periodicals, Inc. Biopolymers 87: 302–316, 2007.

Keywords: single molecule microscopy; fluorescence resonance energy transfer; ribozyme; ribosome; catalytic RNA

WILEY
InterScience®
DISCOVER SOMETHING GREAT

This article was originally published online as an accepted preprint. The "Published Online" date corresponds to the preprint version. You can request a copy of the preprint by emailing the Biopolymers editorial office at biopolymers@wiley.com

INTRODUCTION

A detailed understanding of the biopolymer ribonucleic acid (RNA) is of great importance throughout the life sciences. RNA-coding genes are now recognized to be far more abundant in eukaryotes than their protein-coding counterparts and are essential to the central biochemical processes within all living cells.[1–3] RNA is responsible for the synthesis of all proteins within the cell, plays a central role in replication of many viruses, regulates gene expression in both bacteria and eukaryotes, is involved in the maintenance, processing, modification, and editing of genetic information, and probably carries out a host of still unknown cellular processes. The discovery of the catalytic capabilities of group I introns[4] and RNase P,[5] coupled with the knowledge that certain viral genomes are composed entirely of RNA, established RNA as unique in nature for its ability to both store genetic information and catalyze chemical reactions. The dual genetic and catalytic role of RNA lends tremendous support to the hypothesis that purely RNA-based life predated the emergence of both protein and DNA.[6–8] In addition to their important functions in nature, catalytic RNAs have been used to derive RNA-based therapeutics.[9,10] Our understanding of the molecular underpinnings of organisms, and possibly the origin of life, as well as the development of new medicines, therefore, significantly depend on our ability to dissect the fundamental properties of RNA enzymes.

Naturally occurring ribozymes can be divided into several groups based on their size: small self-cleaving RNAs (<200 nucleotides), medium-sized self-splicing introns, and larger
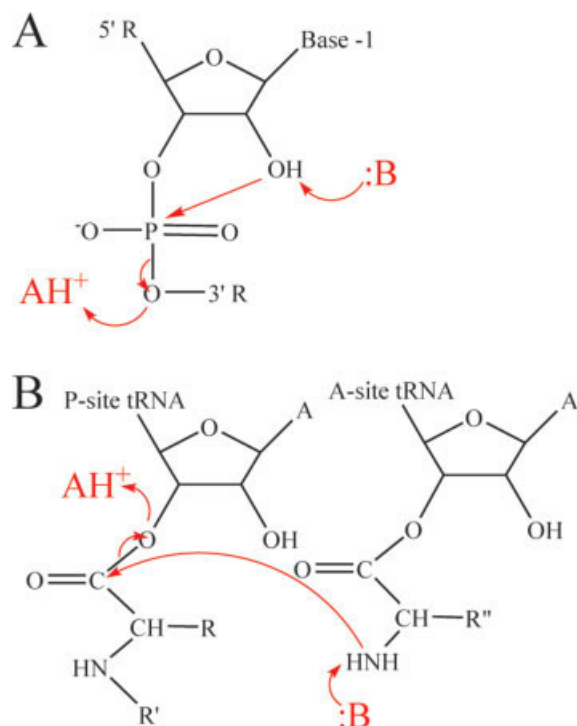
FIGURE 1 Reaction mechanism of the two ribozymes highlighted here. (A) Site-specific phosphodiester transfer as catalyzed by the self-cleaving small ribozymes, including the hairpin ribozyme. A suitably positioned base B deprotonates the 2′-OH of the upstream ribose, thereby activating the 2′-oxygen for nucleophilic in line attack on the scissile phosphodiester. The 5′-oxygen leaving group is protonated by a properly positioned acid AH+. (B) Peptide bond formation as catalyzed by the ribosome. A suitably positioned base deprotonates the amino acid esterified with the A-site tRNA, thereby activating the amino group for nucleophilic attack on the peptidyl-tRNA ester bond on the P-site tRNA. The 3′-oxygen leaving group is protonated by a properly positioned acid AH+.

catalytic ribonuclear-protein (RNP) complexes. The class of small ribozymes comprises the hairpin, hammerhead, hepatitis delta virus (HDV), Varkud satellite (VS), and *glmS* ribozymes. All of these ribozymes catalyze a site-specific RNA backbone cleavage reaction as well as the reverse ligation reaction. Cleavage is achieved through an S_N2-like reaction mechanism in which the 2′-hydroxyl (2′-OH) of the cleaved strand acts as the nucleophile, resulting in 2′,3′-cyclic-phosphate and 5′-OH termini on the 5′- and 3′-products, respectively (Figure 1A).[11–13] On the other end of the spectrum, large RNPs such as RNase P, the spliceosome, and the ribosome represent catalytic RNAs that recruit protein cofactors for optimal function in vivo (self-splicing introns are of intermediate complexity as some of them require protein cofactors and others do not). RNase P and the spliceosome carry out site-specific hydrolysis and transesterification reactions on RNA backbones, respectively, through mechanisms distinct from

that of the small ribozymes. The ribosome is unique among the naturally occurring ribozymes in that it generates a product that is not itself an RNA. The ribosome catalyzes peptide bond formation between amino acids coupled to tRNA adapters and so is responsible for the production of all cellular protein (Figure 1B). Evidence that the RNA rather than protein components of RNPs are catalytic stems from the observation of catalytic competence in the absence of protein and/or an active site composed of RNA only.[4,5,14,15]

Since their discovery a quarter-century ago, extensive investigations into the catalytic mechanisms of ribozymes have been conducted in the quest to understand and potentially exploit this essential and ubiquitous class of enzymes. Until recently, catalytic RNAs were studied in bulk solution, where the number of molecules present is many orders of magnitude larger than the low copy number typical of many RNAs and RNPs in a single cell ($1–10^3$, up to 10^6 in case of the ribosome). Recently, it has become increasingly common to study protein and RNA enzymes using single molecule methods, offering the ability to observe short-lived mechanistic intermediates and minor subpopulations often masked in the ensemble average. Single molecule approaches to understanding RNA include atomic force microscopy, optical tweezers, and single molecule fluorescence microscopy (for review please see Refs. 16 and 17). Of these, single molecule fluorescence resonance energy transfer (smFRET) has proven particularly effective in studying reaction pathways of ribozymes, as smFRET assays provide information on the global dynamics of molecules under native conditions. smFRET has therefore provided researchers with the unique opportunity to quantify the (equilibrium) kinetics of both directions in reversible reactions, which are commonly found in RNA.

In this review we first survey the insights gained from single molecule probing of catalysis by two representative ribozymes and focus on structural dynamics as a signature for catalysis. We then discuss the bottleneck presented by the need to develop suitable assays that probe specific steps on a reaction pathway, as well as proven or plausible routes to overcoming this obstacle to the broader use of single molecule techniques. Single molecule studies of RNA folding pathways have been thoroughly reviewed elsewhere.[17–19]

EXAMPLES OF SINGLE MOLECULE ENZYMOLOGY

Currently, the primary approach used in single molecule RNA enzymology is to monitor global conformational changes associated with individual steps on or off a reaction pathway such as substrate binding, tertiary structure (un)folding, chemical catalysis, and product release. In the following we will explore in detail single molecule investiga-

tions that highlight the scope and limitations of single molecule RNA enzymology. We will focus on two significant RNA catalysts at opposite ends of the spectrum, the hairpin ribozyme and the ribosome. The hairpin ribozyme is probably the most investigated RNA in single molecule enzymology. The ribosome is far more complex and has been subjected to fewer single molecule studies than the comparably simple hairpin ribozyme. However, the tremendous biological importance of protein biosynthesis has motivated substantial progress also on single molecule enzymology of the ribosome.

The Hairpin Ribozyme: Synergy Between Single Molecule and Ensemble Assays

The hairpin ribozyme (Figure 2A) is a small noncoding RNA that facilitates site-specific cleavage and ligation chemistry of its own backbone as part of the double-rolling circle replication of Nepovirus satellite RNAs. It serves as a convenient model system to study RNA catalysis, and a vast body of ensemble biochemical,[20–26] structural,[27–33] and computational data[34,35] is available, as are extensive single molecule analyses.[36–42] The drive toward a complete understanding of catalysis in this system has demonstrated and exploited the power of single molecule spectroscopy to uncover short-lived intermediates, minor subpopulations, and molecular heterogeneity, which otherwise are all hidden in the ensemble average. An effective application of single molecule techniques, conversely, requires correlation of statistically significant averages from stochastic single molecule events with observables from ensemble measurements. In fact, most successful approaches have relied on the availability of a thorough characterization
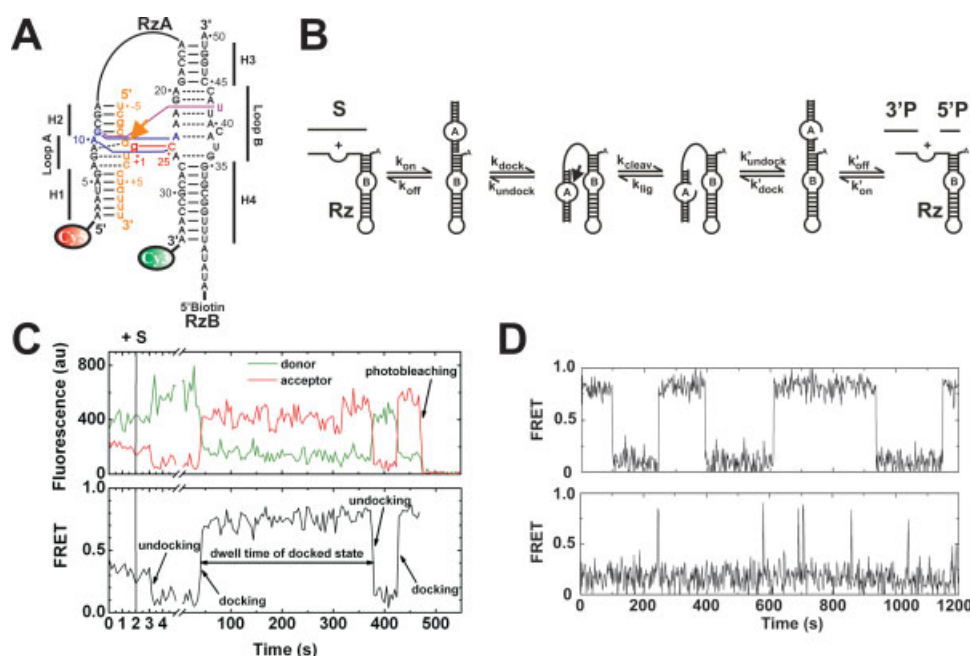

FIGURE 2 Single molecule FRET applied to hairpin ribozyme docking. (A) A two-stranded (RzA, RzB) hairpin ribozyme binds substrate (orange and small letters; arrow, cleavage site) to form internal loops A and B, each flanked by two helices (H1–H4) and connected between H2 and H3. Noncanonical base pairs are indicated by dashed lines. Tertiary structure docking occurs via a g+1:C25 Watson–Crick base pair (red), a ribose zipper (blue), and the U42 binding pocket (purple). Terminal Cy3 and Cy5 fluorophores serve as donor/acceptor FRET pair and biotin is used for surface immobilization through binding to streptavidin. (B) Multistep reaction pathway of the hairpin ribozyme with distinct kinetic steps identified by their rate constants. (C) Typical smFRET time trajectory monitoring donor and acceptor emission intensity, together with the resulting FRET = $I_A/(I_D + I_A)$ trace. Characteristic of a single molecule observation are the anticorrelated donor and acceptor signals and the single-step photobleaching; specific events are indicated. Rate constants are calculated from statistically significant numbers of state dwell times and corrected as described.[36,42] (D) Two FRET time trajectories from different molecules show dramatically different dwell times in the high-FRET docked state that reveal persistent heterogeneity between molecular subpopulations. Reproduced from Ref. 36, with permission from American Association for the Advancement of Science.

of ensemble behavior in order to interpret single molecule observations with confidence.

In the case of the hairpin ribozyme, extensive insights from ensemble techniques into the ribozyme's structural and kinetic properties have formed a solid platform for probing at the single molecule level. For example, ensemble FRET experiments in solution revealed the existence of two structural states at equilibrium—the catalytically inactive undocked and the active docked conformations.[20] Upon docking, the internal loops of domains A and B are brought into close contact, compacting the RNA (Figure 2A).[43,44] Several crystallographic studies showed that this docked state is stabilized by a number of well-characterized tertiary hydrogen bond and base-stacking interactions (Figure 2A).[30–33] In addition, ensemble enzymology approaches were applied extensively, yet the (presumably microreversible) mechanism of cleavage and ligation remains debated.[45] Nucleobase derived general acid–base catalysis,[21,30] water assisted acid–base catalysis,[33,34] and transition state charge stabilization[22–24,31,46] have all been invoked as possible mechanisms. The important contributions that a single nucleobase or even a functional group can make to proper RNA folding as well as catalysis[42] and the inherent ambiguity in the interpretation of enzymologic results[45] contribute to the difficulty of pinpointing the reaction mechanism and necessitate additional mechanistic probing tools.

smFRET based on biotin-streptavidin-mediated surface immobilization and total internal reflection fluorescence microscopy has been employed to dissect the reaction pathway of the hairpin ribozyme, which comprises substrate binding, interdomain docking, substrate cleavage, interdomain undocking, and finally product release (Figure 2B). By labeling the termini of the two interacting domains with a suitable smFRET donor/acceptor pair such as cyanine dyes Cy3/Cy5,[47] the docked, undocked, and product/substrate-free (unbound) states of the ribozyme display distinguishable FRET levels (defined as $I_A/(I_D + I_A)$, where I_A and I_D are the fluorescence signals from acceptor and donor, respectively) (Figures 2 and 3A).[36] Single-step photobleaching to background signal at the end of each smFRET time trajectory confirms that indeed a single RNA molecule is observed (Figure 2C). Since the cleavage products rapidly dissociate from the undocked state, cleavage and subsequent undocking result in a decrease in smFRET from the docked to the unbound state (Figure 3A). The good agreement between the rate of unbound state appearance in smFRET and that of product appearance as monitored by traditional (ensemble) electrophoretic separation further supports the assignment of states and the functional validity of single molecule trajectories (Figure 3A).[36]
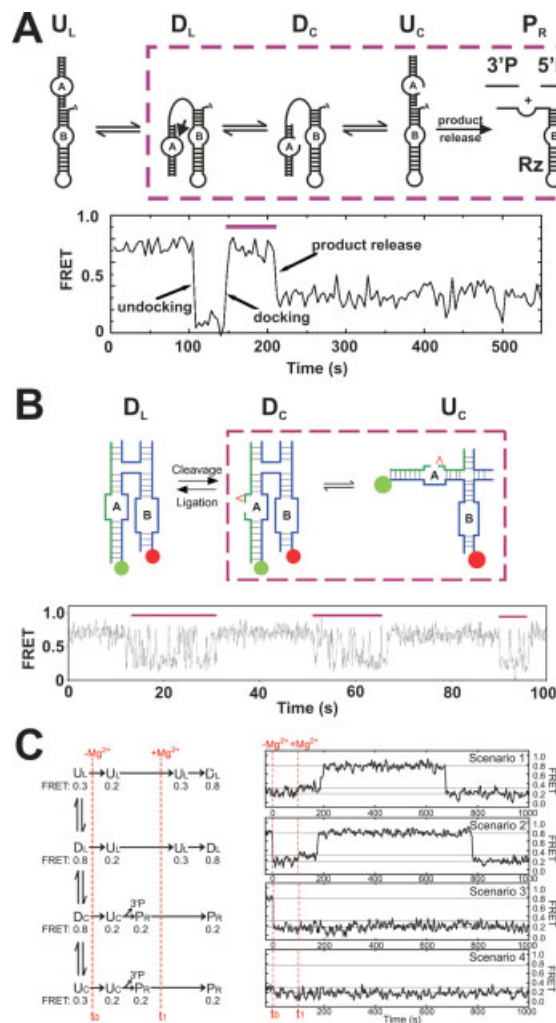
FIGURE 3 Accessing reaction chemistry of the hairpin ribozyme through single molecule FRET. (A) smFRET time trace of the two-way junction form of the hairpin ribozyme, showing the docked, undocked, and substrate/product free states at distinct FRET values.[36] The purple box and bar indicate equivalent processes on the reaction scheme and the experimental data, respectively. Individual states are indicated as U (undocked), D (docked), L (ligated), C (cleaved), and P (product). (B) smFRET time trace of the four-way junction form of the hairpin ribozyme, illustrating the difference in (un)docking dynamics before and after cleavage (i.e., in the ligated and cleaved forms). The purple box and bar highlight equivalent processes on the reaction scheme and in the experimental data, respectively.[39] (C) Schematic of the possible outcome scenarios from a double buffer-exchange experiment (first removal, then replenishment of 12 mM Mg^{2+}) on the two-way junction form of the hairpin ribozyme with the associated experimental smFRET readouts.[49] Reproduced from Refs. 36,39, and 49, with permission from American Association for the Advancement of Science, Nature Publishing Group, and National Academy of Sciences.

To determine (un)docking rate constants in the intact ribozyme–substrate complex unaffected by cleavage, a blocking 2′-O-methyl substitution was introduced into the active

site adenosine (A-1), based on its preservation of the sugar pucker preference and hydrogen bond acceptor capacity of the native 2′-OH.[36] Substrate dissociation is slow under standard conditions (pH 7.5, 12 mM Mg^{2+}), effectively isolating the docking/undocking steps from the remaining reaction pathway. Similarly, (un)docking rate constants in the isolated ribozyme-product complex can be determined by installing a ligation blocking 3′-phosphate on the 5′-product instead of the natural 2′,3′-cyclic phosphate. Rate constants can then be extracted by plotting the cumulative number, $N(t)$, of state dwell (residence) times that are shorter than time t and fitting with a multiexponential of the form:

$$N(t) = \sum_i A_i(1 - e^{k_{i,obs} \cdot t}) \tag{1}$$

The observed rate constants $k_{i,\text{obs}}$ need to be corrected for photobleaching, which shortens the observed dwell times; while the amplitudes A_i need to be corrected if indeed multiple rate constants are observed, to avoid bias toward shorter dwell times.[36,42] Corrected dwell times in the docked state then determine the undocking rate constant(s) k_{undock}, whereas corrected dwell times in the undocked state separately determine the docking rate constant(s) k_{dock} (Figure 2B). This ability to determine forward and reverse rate constants of a reversible reaction independently of each other is an important advantage of single molecule enzymology. In the ensemble average, only collective and synchronized relaxation of many molecules from one state to another can be observed, yielding an aggregate rate constant.

The fact that single molecule data are often best represented through multiple rate constants $k_{i,\text{obs}}$ is another distinction from ensemble averaging techniques. In the hairpin ribozyme (and commonly in RNA), an underlying molecular heterogeneity is apparent upon closer inspection (Figure 2D).[36,37,40] At least four distinct subpopulations of molecules are found in individual smFRET time trajectories,[36,37,42] each of which undocks with one of the four rate constants extracted from the docking dwell times. Representatives of each subpopulation are remarkably resistant to interconversion as they continue to exhibit the same undocking behavior even when probed at 3-h intervals at 25°C,[36] in various Mg^{2+} concentrations,[37] or in the presence of various RNA modifications.[42] Such static heterogeneity (or molecular memory) is also observed in the context of a four-way junction form of the ribozyme, with potentially even more subpopulations exhibiting heterogeneity in both docking and undocking kinetics,[38] and is independent of the strategy for RNA surface immobilization.[40] The four distinctly undocking molecular subpopulations map onto fast and slow phases

of biphasic ensemble cleavage[42] and folding assays,[25] and so have important consequences for the interpretation of data from ensemble measurements.

The Hairpin Ribozyme: Accessing Chemistry

Although docking and undocking rate constants of the hairpin ribozyme can be derived directly from dwell time analyses of smFRET trajectories, catalysis itself does not result in any discernable change in smFRET signal and thus calls for less direct inference. Three different approaches have been pursued so far. In the first approach, single molecule probing of inactivated ribozyme–substrate and –product complexes is combined with ensemble cleavage assays and classic mechanistic modeling.[36,42] Essentially, the intrinsic cleavage and ligation rate constants are derived by finding either a numerical[36] or analytical solution[42] to the set of differential equations that defines the reaction pathway in Figure 2B after substrate binding (where substrate binding is assumed to be irreversible):

$$
\begin{cases}
\dfrac{dN_{undock}^S}{dt} = -k_{dock}N_{undock}^S + k_{undock}N_{dock}^S \\[2mm]
\dfrac{dN_{dock}^S}{dt} = k_{dock}N_{undock}^S - (k_{undock} + k_{cleav})N_{dock}^S + k_{lig}N_{dock}^p \\[2mm]
\dfrac{dN_{dock}^P}{dt} = k_{cleav}N_{dock}^S - (k_{lig} + k_{undock}^P)N_{dock}^P + k_{dock}^p N_{undock}^P \\[2mm]
\dfrac{dN_{undock}^P}{dt} = k_{undock}^P N_{dock}^P - (k_{dock}^P + k_{off})N_{undock}^P + k_{on}N_{diss} \\[2mm]
\dfrac{dN_{diss}}{dt} = k_{off}N_{undock}^P - k_{on}N_{diss}
\end{cases}
$$

$$\tag{2}$$

where N denotes the population of molecules in a given state with the subscript indicating the conformation (docked or undocked) and the superscript indicating whether the ribozyme is bound to substrate (S) or product (P). Using the known overall cleavage kinetics and assuming the complete absence of interconversion between the molecular subpopulations throughout the catalytic cycle (which makes them independent of one another), a numerical fit yields upper and lower bounds for the intrinsic cleavage and ligation rate constants and an estimate of their ratio.[36] Alternatively, Eqs. 2 can be formulated in matrix notation so that an analytical simulation of the overall cleavage time course is derived by diagonalizing and solving the corresponding master equation.[42] If the chemical equilibrium constant is independently determined by, for example, running a ligation reaction to completion in the presence of excess reaction product, the problem of solving the master equation for the five unimolecular reactions that describe the kinetic pathway is reduced to a single-variable fit.[42] The use of substrate and product analogs

and matrix-algebra assisted kinetic simulations thus enables rapid relative comparison of single functional group variants. It was discovered that functional groups far from the docking interactions and active site directly impact both docking and chemistry. These findings led to the proposal that, similar to protein enzymes,[48] long-range coupled molecular motions exist in ribozymes that link the overall fold to the active site and contribute to RNA function.[42] Recent molecular dynamics (MD) simulations support this hypothesis.[34]

It is important to note that two additional assumptions are implicitly employed in this analysis. First, common intrinsic cleavage and ligation rate constants are assumed for all molecular subpopulations so that the derived rate constants represent averages over all molecules (a standard feature also of ensemble enzymology). A range of rate constants may exist, but kinetic modeling suggests that the chemical rate constants vary by less than threefold between the different subpopulations.[42] Second, in the analysis it is assumed that the docking and undocking rate constants obtained for the inactivated ribozyme complexes (with $2'$-O-methyl modified substrate and $3'$-phosphate modified $5'$-product analogs) closely resemble those of the active complexes. Recent evidence from single molecule studies in the presence of cleavable substrate suggests that undocking of the substrate complex is decelerated and that of the product complex accelerated in the presence of the native $2'$-OH and $2',3'$-cyclic phosphate, respectively, leading to a systematic overestimation of the intrinsic cleavage rate constant in the earlier analysis by approximately sevenfold (while docking and ligation rate constants are unaffected; see also the following discussion).[39,49]

The second single molecule approach used to access chemical rate constants of the hairpin ribozyme exploits the fact that undocking is slow in the native substrate complex, but fast in the cleaved product complex; thus, cleavage in situ can be fortuitously detected by an acceleration of the smFRET fluctuations between docked and undocked conformations (rather than a change in FRET level).[39] This change in dynamics is particularly pronounced at 1 mM Mg^{2+} in a four-way junction form of the ribozyme, in which both the $5'$- and $3'$-products are extended to prevent dissociation. Significant enhancement and suppression of the docking/undocking kinetics can thus be used as signatures for cleavage and ligation, respectively (Figure 3B). Still, extraction of a cleavage rate constant is complicated since, first, any given transition from docked to undocked state may originate from either the substrate or product complex; second, cleavage events followed by ligation before undocking will go undetected; and, third, the observed docked state dwell times are shortened by photobleaching. Using a succession of undocking events as indication of cleavage and correcting for

missed events as well as photobleaching then yields an estimate of the intrinsic cleavage rate constant. In addition, the ligation rate constant has to be corrected for the fact that a rapidly docking/undocking ribozyme spends only part of its time in the docked state where ligation can occur. The final corrected intrinsic chemistry rate constants indicate a stronger equilibrium bias toward the ligated state than did the use of chemistry blocking modifications.[39]

The third approach exploited to tease out the intrinsic chemistry rate constants that also uses cleavable substrate and sets up a succession of buffer exchanges to produce distinct time sequences of smFRET signal that serve as kinetic "fingerprints" of specific catalytic intermediates.[49] In concept, such an approach is analogous to pulse-chase experiments widely used in ensemble enzymology, but it gains from the ability to assign a specific state to each individual molecule and count the number of representatives. Figure 3C illustrates how the number of molecules in the undocked (U) and docked (D) states in the presence of either ligated (L) or cleaved (C) substrate is assessed. First, chemical equilibrium is reached upon incubation of the ribozyme in standard buffer (pH 7.5, 12 mM Mg^{2+}) in the presence of a saturating excess of $3'$-product. Upon addition of EDTA to remove Mg^{2+} at time t_0 and subsequent replenishment of Mg^{2+} at time t_1, distinct scenarios are observed depending on which of the four reaction intermediates U$_L$, D$_L$, D$_C$, or U$_C$ are observed. In particular, the two docked states undock (transit from high [0.8] to low [0.2] FRET) upon Mg^{2+} removal, while the two undocked states only slightly decrease in FRET (from 0.3 to 0.2). In addition, the two complexes involving ligated substrate will eventually dock again after the replenishment of Mg^{2+}, whereas the two cleaved complexes lose their $3'$-product under these conditions (the $5'$-product is covalently linked to the ribozyme) and thus can never dock after Mg^{2+} addition. Given a sufficient observation window after time t_1 (i.e., slow photobleaching), the four reaction intermediates can be unequivocally identified through their unique FRET versus time patterns (Figure 3C) and counted. Based on a sufficient number of molecule assignments P, yielding P(U$_L$) = 21, P(D$_L$) = 591, P(D$_C$) = 47, and P(U$_C$) = 165, the equilibrium constants of docking before and after cleavage and of internal chemistry are derived as ratios of the appropriate molecule counts.[49] In conjunction with the rate constant of the very last transition from high to low FRET under standard conditions (Figure 3A), which is a convolution of the undocking, cleavage, and ligation rate constants, the intrinsic cleavage and ligation rate constants can be calculated.

The latter two single molecule studies yield similar intrinsic cleavage and ligation rate constants and a consistent picture of how a ribozyme is optimized for its self-cleavage

function in Nepovirus replication—stable docking of the ligated ribozyme–substrate complex allows for ample time to cleave, while instable docking of the cleaved ribozyme–product complex results in rapid product release.[39,49] One may note that a structural explanation for this kinetic phenomenon is still outstanding. In addition, since a properly ligated RNA is important as a replication template, one may wonder whether alternating structures switch the RNA between preferably cleaved (active) and ligated (inactive) forms. Strikingly, standard ensemble measurements of the chemistry equilibrium position do not distinguish between the docked and the undocked ligated or cleaved states and thus lead to a significant underestimation of the ligation equilibrium constant (from $[P(U_L) + P(D_L)]/[P(D_C) + P(U_C)] \approx 2.9$[42] compared to its true value (which is defined as $P(D_L)/P(D_C) \approx 13$,[49] and a resulting overestimation of the intrinsic cleavage rate constant. Ensemble studies also average out important information on parallel (heterogeneous) reaction pathways, which are studied in isolation when observing single molecules, and on short-lived intermediates, which are identified by short smFRET bursts as long as they live longer than the experimental time resolution (Figure 2D).

In summary, single molecule enzymology studies of the hairpin ribozyme have demonstrated feasible routes toward determining rate and equilibrium constants of the chemical step in a fully reversible RNA reaction pathway that exhibits molecular heterogeneity. It thus has become possible to dissect the often surprisingly profound role(s) of individual residues and functional groups in structural dynamics and chemistry, may they be close to or far from the active site and/or tertiary structure interactions. Careful consideration needs to be given to the various types of modifications (fluorophore labeling, surface immobilization, functional group, and sequence changes) that have to be introduced into the RNA to address specific scientific questions. Powerful synergies arise from the use of multiple alternate approaches. Future advances in our understanding of the mechanism of site-specific backbone cleavage will require a careful integration of single molecule fluorescence approaches with those of, in particular, ensemble enzymology coupled with mutagenesis, X-ray crystallography, NMR spectroscopy, MD simulation, and quantum mechanical/molecular mechanical calculations.

The Ribosome: A Complex RNA-Protein Machinery

The largest ribozyme studied so far at the single molecule level is the protein biosynthetic machinery, the ribosome, arguably the most abundant enzyme on earth. Ribosomes are very large (in bacteria ~2.5 MDa) RNA-protein complexes that universally translate the sequence of a messenger RNA (mRNA) with high fidelity into a polypeptide chain using transfer RNA (tRNA) adaptors. Ribosomes are composed of a large and a small subunit (termed 50S and 30S in bacteria, respectively). Translation is initiated by the assembly of the two subunits into the 70S ribosome on an mRNA template. Protein synthesis is catalyzed by the ribosomal RNA (rRNA) component of the large subunit by transfer of the growing peptide chain from one tRNA onto the next aminoacyl-tRNA (aa-tRNA),[50] whose selection is determined by the small subunit. A sequence of three nucleotides in the tRNA, called the anticodon, base pairs with each mRNA codon; this short hybrid is proofread with high fidelity by the rRNA component of the small subunit.[51] Both subunits of the ribosome contribute to its three tRNA binding sites: the aminoacyl site (A site), the peptidyl site (P site), and the exit site (E site).[52]

The highly dynamic peptide elongation cycle is composed of at least the following steps (Figure 4A)[51,53–56]: (1) Initial binding of the aa-tRNA to the ribosome in a ternary complex with the elongation factor Tu (EF-Tu) and guanosine 5′-triphosphate (GTP). (2) Codon recognition where base pairs form between the tRNA anticodon and mRNA codon in the A site of the 30S subunit. (3) Stimulation of the GTPase activity of EF-Tu in response to correct codon–anticodon pairing with the cognate aa-tRNA. At this stage, mismatched (noncognate) aa-tRNAs are readily ejected from the ribosome, whereas near-cognate aa-tRNAs (only one base mismatch) remain bound.[51,57] (4) GTP hydrolysis by EF-Tu. (5) Conformational change of EF-Tu coupled with dissociation of inorganic phosphate (P_i). (6) Accommodation of the A-site aa-tRNA within the peptidyl transferase center (PTC) of the large subunit, during which EF-Tu:GDP dissociates. At this point, most near-cognate aa-tRNAs are rejected in a process known as proofreading.[51,58] (7) Peptidyl transfer, that is, formation of the peptide bond in the PTC, during which the polypeptide chain is rapidly transferred from the P-site to the A-site tRNA (see also Figure 1B). (8) The tRNAs proceed to two discernable hybrid states, where they remain bound to the mRNA in the A and P sites of the 30S subunit, while their 3′-ends bind to the P and E sites of the 50S subunit, respectively.[56] (9) Full translocation of the tRNA:mRNA complex from the P and A sites to the E and P sites of the ribosome, respectively, which is accelerated by GTP hydrolysis on the elongation factor G (EF-G). The ribosome is thus ready for the next round of elongation, where the A site is prepared to accept the next aa-tRNA ternary complex (Figure 4A).

The elongation cycle involves extensive conformation changes and the dynamic association of multiple RNA and protein complexes, which lend themselves to FRET-based
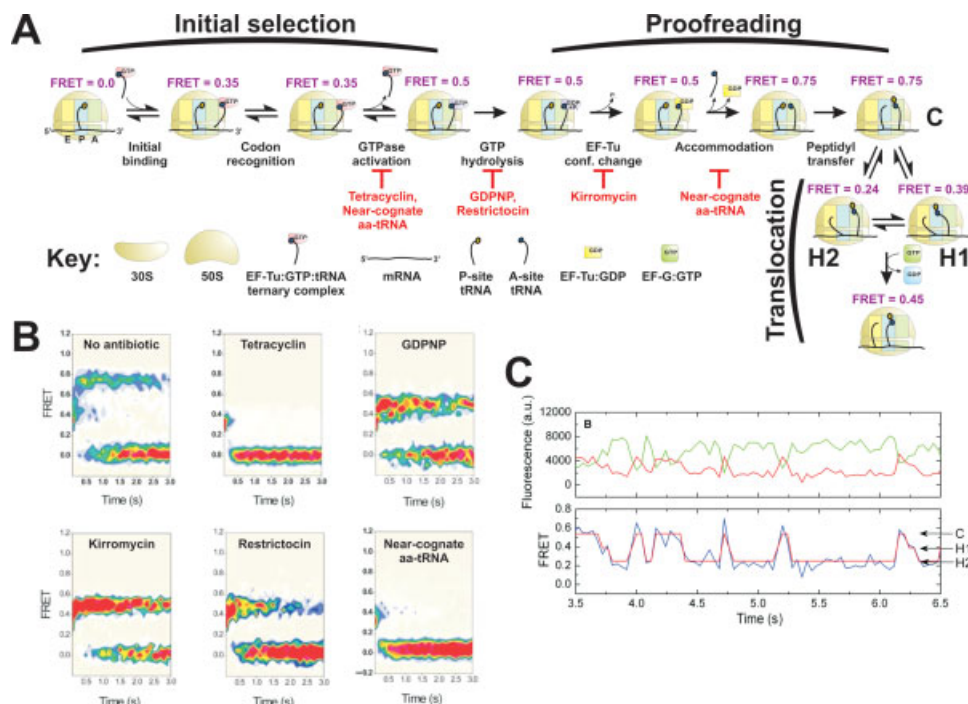
FIGURE 4 Dissecting the complex ribosomal translation cycle by single molecule FRET. (A) Schematic of the reaction cycle of the ribosome as illuminated through smFRET (FRET values are indicated in purple). The E (yellow), P (blue), and A (green) sites on the ribosome are shown as small rectangles. Specific steps can be inhibited by the antibiotics or additives indicated in red. (B) Contour plots of the postsynchronized time evolution of many individual smFRET trajectories in the presence of the indicated antibiotics and additives. Contours are plotted from tan (lowest population) to red (highest population).[54] (C) Donor and acceptor fluorescence signals and corresponding smFRET time trajectory displaying the classic (C), hybrid I (H1), and hybrid II (H2) states. The FRET data are hidden Markov modeled (red line) to determine dwell times in the three states and derive interconversion rate constants.[56] Reproduced from Refs. 54 and 56, with permission from Nature Publishing Group and Elsevier.

distance measurements. Single molecule analysis is particularly well suited to address such a multistep biological process where transitions between states do not remain synchronized within the ensemble and the resulting averaged signal cannot easily be deconvoluted. Finally, the recent explosion in structural and mechanistic insights into ribosome function (reviewed in, for example, Refs. 50,51,59, and 60) provide a vast knowledge base for indepth studies of the single molecule enzymology of translation.

So far, several single molecule fluorescence studies have been conducted on reconstituted ribosomes from *Escherichia coli*.[54–56,61] In particular, intermediates involved in aa-tRNA selection are resolved by combining stopped-flow mixing with the use of antibiotics that inhibit tRNA selection, a nonhydrolyzable GTP analog (GDPNP) and mutant ribosomes. Single *E. coli* ribosomes with a Cy3-labeled aa-tRNA (fMet-tRNA^{fMet}) in the P site are immobilized via a 5′-biotinylated mRNA on a streptavidin coated and otherwise passivated

quartz slide.[55] An EF-Tu:GTP:Phe-tRNA^{Phe} ternary complex, where the tRNA is labeled with Cy5, is delivered to the immobilized ribosomes via stopped-flow. The FRET level of single complexes proceeds stepwise from FRET = 0.35 in the initial bound state to FRET = 0.75 in the state with a fully A-site accommodated aa-tRNA (Figure 4A). Because of the asynchronous nature of initial aa-tRNA binding, individual time traces were "postsynchronized" to the first FRET $\geq$ 0.25 signal. The average time from this initial FRET state to complete accommodation is ~93 ms, in agreement with previous ensemble studies on aa-tRNA selection, which validates the single molecule approach.[62]

Classic approaches to stop translation at defined points along the elongation cycle provide an elegant means of assigning the FRET states and so further dissect ribosome kinetics.[54] Tetracycline is an antibiotic that inhibits selection of the ternary complex by the ribosomal A site (Figure 4A). smFRET traces in the presence of tetracycline show the tran-

sient sampling of a FRET = 0.35 state, with rare excursions into higher FRET ($\geq$ 0.5) states (Figure 4B).[54] Each FRET = 0.35 event is identified as an independent attempt of a single ternary complex to enter the ribosomal A site, upon which tetracycline interrupts aa-tRNA selection. Replacing GTP with its nonhydrolyzable analog GDPNP in the ternary complex is known to efficiently stall aa-tRNA selection prior to GTP hydrolysis (Figure 4A). Single molecule traces observed in the presence of GDPNP rapidly transit through the FRET = 0.35 state and show a stabilized FRET = 0.5 state (Figure 4B). In this prehydrolysis state, the ternary complex is in closer contact with the ribosome than in the 0.35 FRET state. Kirromycin inhibits the conformational change of EF-Tu after GTP hydrolysis (Figure 4A). In single molecule traces, kirromycin also stops the ribosome at a FRET = 0.5 state after readily passing through the 0.35 FRET state (Figure 4B). Restrictocin is an α-sarcin toxin homolog that specifically cleaves the sarcin–ricin loop of the ribosome, which is involved in activating the GTPase activity of EF-Tu. In the presence of this toxin, smFRET time traces of the ribosome again stall in a FRET = 0.5 state before aa-tRNA accommodation (Figure 4B).

The fidelity of initial aa-tRNA selection and proofreading is investigated by comparing encoded mRNAs with cognate (UUU), near-cognate (CUU), and noncognate (AAA) codons at the A site.[54] Analysis of smFRET traces shows that a noncognate aa-tRNA is effectively rejected at the codon recognition state (FRET = 0.35, Figure 4A). By contrast, 65% of bound cognate aa-tRNAs advance to a FRET state higher than 0.35, while only 11% of near-cognate aa-tRNAs move forward. Therefore, the FRET = 0.35 state represents an important initial selection step for correct anticodon–codon pairing that favors cognate over near-cognate aa-tRNAs by ~6:1 (Figure 4A). A (second) proofreading step is based on the stability of the A-site aa-tRNA bound to the rearranged ribosome after the release of the EF-Tu-GDP complex, which occurs during accommodation and is identified by smFRET as a transition from FRET = 0.5 to FRET = 0.75 (Figure 4A). Cognate aa-tRNAs are favorably accommodated in the PTC over near-cognate aa-tRNAs by a ratio of 24:1.[54]

After peptidyl transfer, fluctuations between FRET = 0.75 and 0.45 states are observed, which were initially attributed to a dynamic exchange between the classic (C) and translocation hybrid states (H), respectively (Figure 4A).[55] More recently, detailed smFRET analysis yielded evidence for the formation of two hybrid states (H1 and H2) in equilibrium with each other and the classic state (Figures 4A and 4C).[56] Rate constants for intermediate steps in aa-tRNA selection, proofreading, and translocation were obtained by evaluating the FRET states and determining their corresponding single

event dwell times based on the statistical evaluation afforded by hidden-Markov modeling (HMM; Figure 4C and discussion given later).

In summary, work on the ribosome has demonstrated how even very complex reaction cycles can be followed in real-time at the single molecule level to determine rate constants of individual reaction steps. Oftentimes only a single molecule approach can identify and characterize specific conformational states in the cycle. Future advances in our understanding of the ribosome mechanism will take advantage of new labeling sites to probe additional states by smFRET and specify the dynamic location of reaction participants hidden to the aa-tRNA distances monitored so far.

CHALLENGES

The examples given earlier demonstrate how previous knowledge from ensemble experiments can be used to inform single molecule experimental design and interpretation and thus dissect multistep enzymatic reaction pathways. smFRET observations do not require synchronicity between molecules to extract rate constants from statistical analyses of the underlying stochastic (even rare and/or brief) dwell times, making them uniquely suited for answering questions of RNA enzymology. (State dwell times often define the macroscopic rate constants since conformational transitions in RNA are typically fast, but rare. Only recently have attempts been made to measure an actual folding transition time.)[63] As with any enzymatic assay, however, challenges lie in developing appropriate probes, assay and signal detection conditions, strategies for extracting and interpreting experimental observables, and in deriving kinetic models. In the following we discuss some of the important challenges and highlight possible solutions.

Evaluating Dwell Times Correctly

In kinetic analyses of single molecule traces, dwell times must be measured with confidence. Large ribozymes may introduce an added level of complexity into the analysis, as they may not display simple two-state FRET behavior with well-defined dwell times. To deal with increasingly complex single molecule data sets a number of statistical methods have been implemented. Hidden-Markev Modeling (HMM), in particular, has recently been utilized to determine the number of distinct states and the transitions probabilities between states.[56,64,65] HMM in principle can identify a large number of states within a single molecule trajectory in a semiautomated, reproducible fashion, thus avoiding potential bias. However, the assumptions implicit in any statistical

model need to be kept in mind; for example, HMM requires that the transitions be Markovian in nature, i.e., the current state is independent of past states. This assumption requires transitions within a single trajectory to be dictated by a single rate constant, which is not necessarily the case for RNA. It is therefore wise to apply HMM independently to each single molecule trace to avoid masking any molecular heterogeneity that may be present. Furthermore, even individual trajectories may violate the assumption of Markovian behavior. Nevertheless, HMM has been successfully used to evaluate smFRET trajectories in the ribosome work discussed earlier (Figure 4C),[56] as well as in studies of Holliday junctions,[64] and RecA filament assembly on single-stranded DNA (Figures 5A and 5B).[65] The extent to which HMM can be used to distinguish FRET states is demonstrated by the latter study. smFRET trajectories monitoring sequential association of up to four RecA monomers into a filament were analyzed, lead-

ing to five discernable FRET states with eight transition densities and fundamental rate constants for the stepwise binding and dissociation rates of individual monomers (Figure 5B).[65]

In some cases, such as in a folding study of the catalytic domain of *Bacillus subtilis* RNase P RNA, gradual structural FRET changes that lack abrupt transitions may be observed, limiting the ability to define FRET states (Figure 5C).[66] While the fully unfolded (0 mM Mg^{2+}) and folded (5 mM Mg^{2+}) states of the catalytic domain display narrow distributions with FRET = 0.13 and 0.85, respectively, at intermediate and physiologically relevant Mg^{2+} concentrations (at least four FRET distributions are discerned; Figure 5D). Any particular RNA molecule is restricted to a limited range of FRET values (although this may in part be related to the unusually short observation window in this particular study, see also the discussion given earlier).[66] The authors therefore evaluate the
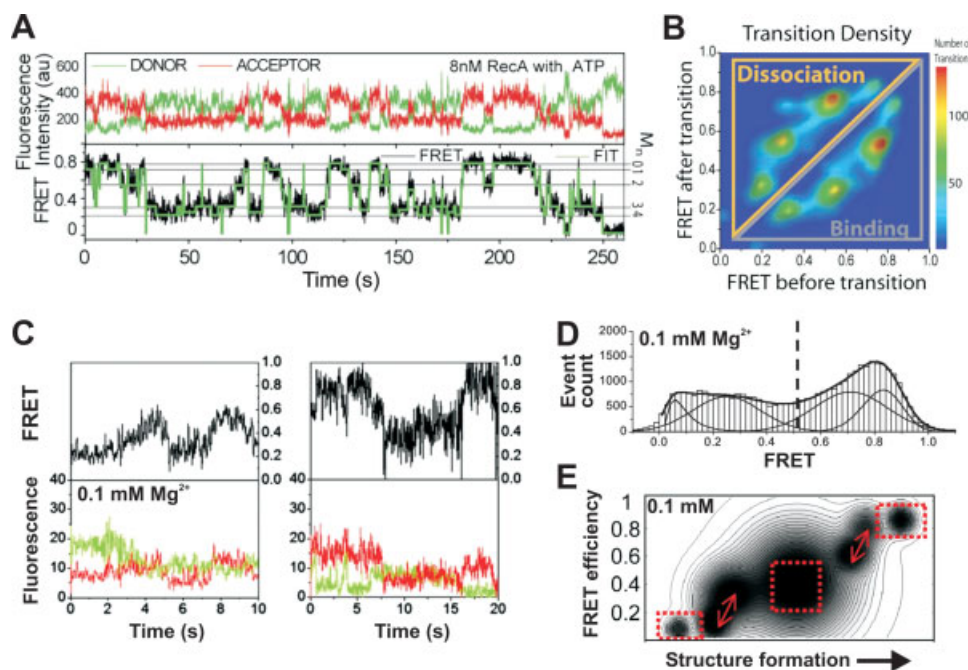
FIGURE 5 Challenge: Dealing with complex single molecule FRET kinetics. (A) Donor and acceptor signals and corresponding smFRET time trajectory upon assembling RecA in the presence of ATP onto the single-stranded 3′-extension of a double-stranded DNA. The FRET data are hidden Markov modeled (green line) to determine dwell times in five different states (M$_0$–M$_4$) and distinguish from acceptor dark states (FRET = 0). (B) A transition density plot of RecA binding and dissociation transitions observed on 82 DNA molecules shows five FRET states (FRET ≈ 0.2, 0.3, 0.55, 0.75, 0.85) that interconvert pairwise.[65] (C) smFRET time trajectories and donor and acceptor fluorescence signals from the catalytic domain of RNase P incubated at 0.1 mM Mg^{2+} reveal gradual transitions between poorly defined FRET states. (D) FRET distribution histogram from ∼50 smFRET time trajectories of the RNase P catalytic domain observed at 0.1 mM Mg^{2+}. (E) Resulting free-energy contour plot for the folding pathway of the RNase P catalytic domain as monitored by 3′- to 5′-end proximity. Two fluctuating classes, reflected by two pairs of closely connected basins (double arrows), as well as three nonfluctuating classes of smFRET states (dashed boxes) can be defined.[66] Reproduced from Refs. 65 and 66, with permission from Elsevier and National Academy of Sciences.

folding pathway using free-energy contour maps (Figure 5E). This approach assumes thermodynamic equilibrium, since free energies are derived from probabilities based on the relative population sizes of folding states in the ensemble. The authors conclude that early folding steps in the catalytic domain of RNase P RNA involve a series of intermediates that fold under the kinetic control of local conformational rearrangements. Similar free-energy contour map approaches may prove useful in evaluating progress along the reaction coordinate of RNA enzymes.

A significant limitation for any fluorescence-based single molecule study is the nonideal photophysical behavior of fluorophores; in particular, their photobleaching limits the total observation window (so that rate constants extracted from dwell times have to be corrected[42]) and long-lived dark states may persist for seconds. The donor–acceptor pair Cy3/Cy5 is often favored in single molecule studies because of its large wavelength difference and strong FRET signal that can be prolonged by enzymatic oxygen scavenger systems.[67] However, low FRET states have been observed in smFRET studies that arise from the acceptor temporarily visiting a dark state (Figure 5A).[42,65,68] This so-called "blinking" is of concern because it can potentially be misinterpreted as a conformational change in the labeled molecule. The problem is exacerbated by the observation that the blinking kinetics vary depending on the identity of the donor, interfluorophore distance, and buffer conditions.[69,70] Fortuitously, the resulting FRET = 0.0 is often sufficiently distinct from a "real" low FRET signal (Figure 5A). If this is not the case, rapidly alternating-laser excitation of the donor and acceptor fluorophores provides a solution, whereby acceptor activity is continuously probed as a control.[69,71] Other promising advances toward longer smFRET observation windows may be expected from additives such as Trolox (6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid), which suppress blinking and photobleaching of Cy5,[72] as well as the development of improved fluorophores.[73]

Multiple Turnover Kinetics: Michaelis–Menten Applied to Single Molecules

The assays described earlier are all single-turnover in nature, with one substrate turned over per single enzyme (although potentially multiple times). Traditional ensemble enzyme assays are often performed under multiple-turnover conditions, where the observed rate constants are not only affected by conformational change and reaction chemistry, but also by substrate binding and dissociation. However, multiple turnover assays at the single molecule level bear the potential to resolve slow conformational changes of enzymes as a potential basis for molecular heterogeneity. Since Michaelis and Menten's pioneering work in 1913 on invertase (nowadays called β-fructofuranosidase),[74] the multiple turnover properties of enzymes have been described using the Michaelis–Menten formalism, where substrate (S) binds reversibly to the enzyme (E) to form an ES complex, which reacts unimolecularly to yield the final product (P) and restore the original enzyme (E):

$$\text{E} + \text{S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \overset{k_2}{\rightarrow} \text{E} + \text{P} \tag{3}$$
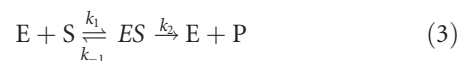
Michaelis and Menten found that the velocity v of an enzymatic reaction has a hyperbolic dependence on the substrate concentration [S]:

$$\frac{v}{[E]_0} = \frac{k_2[S]}{[S] + K_M} \tag{4}$$

The Michaelis constant K_M and the maximum rate constant $v_{\max}$ are defined as $K_M = \frac{k_{-1}+k_2}{k_1}$ and $v_{\max} = k_2[\text{E}]_0$, respectively, where $[\text{E}]_0$ is to the total enzyme concentration.

Xie and coworkers have developed approaches to describe multiple turnovers by single enzymes that focus on the stochastic dwell times for the enzyme to complete one turnover cycle.[75–78] Potentially all rate constants in reaction Eq. 3 may then be dependent on the multidimensional, fluctuating conformational coordinate r of the enzyme. The derived kinetic equation uses the mean waiting (dwell) time between consecutive catalytic events, $\langle t \rangle$, to describe the kinetics of the enzyme reaction:

$$\frac{1}{\langle t \rangle} = \frac{\chi_2[S]}{[S] + C_M} \tag{5}$$

The analogy between Eqs. 4 and 5 is obvious, although the apparent catalytic rate constant χ_2 and apparent Michaelis constant C_M relate to the classic k_2 and K_M values in ways that depend on the relative magnitudes of the rate constants in equation 3.[77] They also take on a new ensemble-averaged meaning.[79]

To test the validity of the single molecule Michaelis–Menten Eq. 5, single (tetrameric) β-galactosidase enzyme molecules were immobilized on beads for easy manipulation and monitored the continuous turnover of fluorogenic substrate molecules of resorufin-β-D-galacto-pyranoside.[78] At low substrate concentration, substrate binding and dissociation predominate and the waiting time distribution appears as relatively single-exponential, while at high concentration catalysis dominates the observed waiting times, and a clear multiexponential distribution is observed. This behavior is attributed to dynamic conformational heterogeneity, leading to fluctuations in catalytic rate constant over broad timescales (from milliseconds to tens of seconds).[78] (It should be noted that a recent reanalysis showed that quasiequilibrium conditions of substrate binding and dissociation can account

for all data so that conformational and thus catalytic heterogeneity consistently contributes, making enzyme turnover multiexponential at both low and high substrate concentrations.[77]) Based on Eq. (5), a linear Lineweaver–Burke plot of $\langle t \rangle$ as a function of $1/[S]$ yields $\chi_2 = 730 \pm 80 \, s^{-1}$ and $C_M = 390 \pm 60 \, \mu M$, values that are in excellent agreement with the ensemble-averaged classic Michaelis–Menten parameters $\frac{v_{max}}{[E]_{Tot}} = 740 \pm 60 \, s^{-1}$ and $K_M = 380 \pm 40 \, \mu M$, despite their different microscopic interpretation.

Observation of multiple substrate turnovers thus has proven valuable in detecting conformational fluctuations between various catalytic forms of a protein enzyme, but an application to RNA enzymes is still outstanding. The applicability of this approach of course depends on the necessity that observations on single enzymes be longer than the timescale of the conformational fluctuations to be probed (which is not easily accomplished, for example, in case of the very slowly interconverting molecular subpopulations of the hairpin ribozyme described earlier and in Figure 2D).

New Observables

Ensemble-based FRET probing of RNA global structures and reaction pathways has often been complimented by techniques that utilize fluorescent nucleoside analogs to detect local conformational changes.[80–95] Some of the unconventional nucleosides integrated into RNA for ensemble studies are shown with their spectroscopic properties in Figure 6. The most commonly used example is 2-aminopurine nucleoside, an adenosine isomer whose fluorescence intensity decreases dramatically when it stacks on nearby nucleotides in single- or double-stranded RNA.[80,96] Pyrrolo-C nucleoside, an analogue of cytidine, also decreases significantly in fluorescence when integrated into a single- or double-strand.[92,97] The

recently synthesized furan-conjugated uridine analog shows strong fluorescence free in aqueous solution and is threefold quenched within an RNA.[95] Tor and coworkers suggest four general requirements for the selection of a suitable nucleoside analog[95]: (1) It should preserve structural features of the natural nucleoside for isosteric replacement. (2) The emission maximum should be at long wavelengths (ideally in the visible range), where detection systems are most sensitive. (3) The extinction coefficient and fluorescence quantum yield should be high. (4) The photophysical properties must be sensitive to changes in the local microenvironment. If ways can be found to follow these guidelines and particularly improve on the typically low extinction coefficients in the ultraviolet wavelength range, nucleoside analogs incorporated into RNA could potentially be used to probe local conformations by single molecule fluorescence spectroscopy.

Expanding Single Molecule RNA Enzymology to New Systems

Future applications of single molecule RNA enzymology will be guided by questions of biological function. Studies of the single hammerhead, HDV, VS, and *glmS* ribozymes, as well as of the larger self-splicing introns and RNP complexes such as RNase P and the spliceosome are all within reach. All of these systems possess the necessary broad foundation of prior ensemble measurements that make them ideal candidates for single molecule approaches. Our current knowledge and targets of opportunity for smFRET are briefly summarized here for RNase P and the spliceosome.

RNase P is an evolutionarily ancient enzyme that universally catalyzes the removal of the $5'$ end of pre-tRNAs as part of tRNA maturation in all organisms. Activity of the ribo-
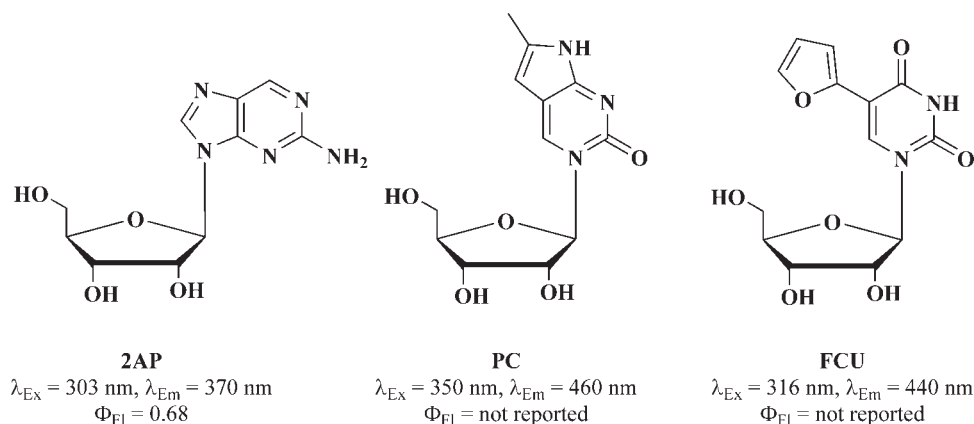
FIGURE 6 Fluorescent nucleoside analogs currently used to probe local structural dynamics in RNA ensemble assays. λ_{Ex} = excitation wavelength; λ_{Em} = emission wavelength; Φ_{Fl} = fluorescence quantum yield.

2AP
$\lambda_{Ex} = 303$ nm, $\lambda_{Em} = 370$ nm
$\Phi_{Fl} = 0.68$

PC
$\lambda_{Ex} = 350$ nm, $\lambda_{Em} = 460$ nm
Φ_{Fl} = not reported

FCU
$\lambda_{Ex} = 316$ nm, $\lambda_{Em} = 440$ nm
Φ_{Fl} = not reported

zyme involves the site-specific coordination of substrate, divalent metal ions, and at least one protein component by the RNase P RNA.[98–100] Even for the simple one-RNA-one-protein bacterial system, it has been difficult to clearly identify the role of metal ions and the cationic protein component in the reaction mechanism, since folding, substrate binding, and catalysis are all ion-dependent. Despite extensive study and several recent crystal structures of the RNA component, important details of the reaction pathway remain unclear.[100] smFRET, which so far has only been applied to folding of the isolated catalytic domain of RNase P RNA as described earlier, is ideally suited for dissecting this reaction pathway. Multiple labeling sites have already been established for ensemble time-resolved FRET studies and may also be suitable for single molecule approaches.[99]

The spliceosome is a massive RNP complex that assembles to excise introns with single-nucleotide precision from pre-mRNAs as part of mRNA maturation in all eukaryotes.[101–103] pre-mRNA splicing is accompanied by a complex series of conformational rearrangements of the central, presumably catalytic small nuclear (sn)RNAs, guided by a large number of protein cofactors including several RNA helicases. The presence of multiple global conformational changes along the splicing reaction pathway provides numerous opportunities to follow activity by smFRET. Additionally, the RNA helicases in the spliceosome are ATP-dependent, which provides a convenient means of controlling progress along the reaction trajectory.

CONCLUSIONS

Current single molecule FRET microscopy approaches that detect large-scale RNA conformational changes in real-time have been successfully employed to determine rate and equilibrium constants of folding as well as catalysis by ribozymes, including the small hairpin ribozyme and the large ribosome. Single molecule analysis exposes molecular heterogeneities and short-lived and/or rare intermediates otherwise masked in the ensemble average. Such single molecule RNA enzymology provides a fresh focus on the biological function of an increasing number of nonprotein coding RNAs discovered over the past 25+ years. Care has to be taken to develop suitable assays based on prior knowledge from ensemble studies. The preparation of proper controls, the conscientious analysis of statistically significant numbers of single molecule trajectories, and the comparison of single molecule results with ensemble data are all essential considerations. In the future, single molecule fluorescence microscopy will undoubtedly see expanded utility through the development of new approaches and the application to new RNA systems.

REFERENCES

1. Carninci, P.; Kasukawa, T.; Katayama, S.; Gough, J.; Frith, M. C.; Maeda, N.; Oyama, R.; Ravasi, T.; Lenhard, B.; Wells, C.; Kodzius, R.; Shimokawa, K.; Bajic, V. B.; Brenner, S. E.; Batalov, S.; Forrest, A. R.; Zavolan, M.; Davis, M. J.; Wilming, L. G.; Aidinis, V.; Allen, J. E.; Ambesi-Impiombato, A.; Apweiler, R.; Aturaliya, R. N.; Bailey, T. L.; Bansal, M.; Baxter, L.; Beisel, K. W.; Bersano, T.; Bono, H.; Chalk, A. M.; Chiu, K. P.; Choudhary, V.; Christoffels, A.; Clutterbuck, D. R.; Crowe, M. L.; Dalla, E.; Dalrymple, B. P.; de Bono, B.; Della Gatta, G.; di Bernardo, D.; Down, T.; Engstrom, P.; Fagiolini, M.; Faulkner, G.; Fletcher, C. F.; Fukushima, T.; Furuno, M.; Futaki, S.; Gariboldi, M.; Georgii-Hemming, P.; Gingeras, T. R.; Gojobori, T.; Green, R. E.; Gustincich, S.; Harbers, M.; Hayashi, Y.; Hensch, T. K.; Hirokawa, N.; Hill, D.; Huminiecki, L.; Iacono, M.; Ikeo, K.; Iwama, A.; Ishikawa, T.; Jakt, M.; Kanapin, A.; Katoh, M.; Kawasawa, Y.; Kelso, J.; Kitamura, H.; Kitano, H.; Kollias, G.; Krishnan, S. P.; Kruger, A.; Kummerfeld, S. K.; Kurochkin, I. V.; Lareau, L. F.; Lazarevic, D.; Lipovich, L.; Liu, J.; Liuni, S.; McWilliam, S.; Madan Babu, M.; Madera, M.; Marchionni, L.; Matsuda, H.; Matsuzawa, S.; Miki, H.; Mignone, F.; Miyake, S.; Morris, K.; Mottagui-Tabar, S.; Mulder, N.; Nakano, N.; Nakauchi, H.; Ng, P.; Nilsson, R.; Nishiguchi, S.; Nishikawa, S.; Nori, F.; Ohara, O.; Okazaki, Y.; Orlando, V.; Pang, K. C.; Pavan, W. J.; Pavesi, G.; Pesole, G.; Petrovsky, N.; Piazza, S.; Reed, J.; Reid, J. F.; Ring, B. Z.; Ringwald, M.; Rost, B.; Ruan, Y.; Salzberg, S. L.; Sandelin, A.; Schneider, C.; Schonbach, C.; Sekiguchi, K.; Semple, C. A.; Seno, S.; Sessa, L.; Sheng, Y.; Shibata, Y.; Shimada, H.; Shimada, K.; Silva, D.; Sinclair, B.; Sperling, S.; Stupka, E.; Sugiura, K.; Sultana, R.; Takenaka, Y.; Taki, K.; Tammoja, K.; Tan, S. L.; Tang, S.; Taylor, M. S.; Tegner, J.; Teichmann, S. A.; Ueda, H. R.; van Nimwegen, E.; Verardo, R.; Wei, C. L.; Yagi, K.; Yamanishi, H.; Zabarovsky, E.; Zhu, S.; Zimmer, A.; Hide, W.; Bult, C.; Grimmond, S. M.; Teasdale, R. D.; Liu, E. T.; Brusic, V.; Quackenbush, J.; Wahlestedt, C.; Mattick, J. S.; Hume, D. A.; Kai, C.; Sasaki, D.; Tomaru, Y.; Fukuda, S.; Kanamori-Katayama, M.; Suzuki, M.; Aoki, J.; Arakawa, T.; Iida, J.; Imamura, K.; Itoh, M.; Kato, T.; Kawaji, H.; Kawagashira, N.; Kawashima, T.; Kojima, M.; Kondo, S.; Konno, H.; Nakano, K.; Ninomiya, N.; Nishio, T.; Okada, M.; Plessy, C.; Shibata, K.; Shiraki, T.; Suzuki, S.; Tagami, M.; Waki, K.; Watahiki, A.; Okamura-Oho, Y.; Suzuki, H.; Kawai, J.; Hayashizaki, Y. Science 2005, 309, 1559–1563.

2. Katayama, S.; Tomaru, Y.; Kasukawa, T.; Waki, K.; Nakanishi, M.; Nakamura, M.; Nishida, H.; Yap, C. C.; Suzuki, M.; Kawai, J.; Suzuki, H.; Carninci, P.; Hayashizaki, Y.; Wells, C.; Frith, M.; Ravasi, T.; Pang, K. C.; Hallinan, J.; Mattick, J.; Hume, D. A.; Lipovich, L.; Batalov, S.; Engstrom, P. G.; Mizuno, Y.; Faghihi, M. A.; Sandelin, A.; Chalk, A. M.; Mottagui-Tabar, S.; Liang, Z.; Lenhard, B.; Wahlestedt, C. Science 2005, 309, 1564–1566.

3. Taft, R. J.; Pheasant, M.; Mattick, J. S. Bioessays 2007, 29, 288–299.

4. Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. Cell 1982, 31, 147–157.

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

6. Orgel, L. E. J Mol Biol 1968, 38, 381–393.

7. Crick, F. H. C. J Mol Biol 1968, 38, 367–379.

8. Gesteland, R. F.; Cech, T. R.; Atkins, J. F. E. The RNA World, 2nd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1999.

9. Sullenger, B. A.; Gilboa, E. Nature 2002, 418, 252–258.

10. Breaker, R. R. Nature 2004, 432, 838–845.

11. Doudna, J. A.; Lorsch, J. R. Nat Struct Mol Biol 2005, 12, 395–402.

12. Fedor, M. J.; Williamson, J. R. Nat Rev Mol Cell Biol 2005, 6, 399–412.

13. Winkler, W. C.; Nahvi, A.; Roth, A.; Collins, J. A.; Breaker, R. R. Nature 2004, 428, 281–286.

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

15. Valadkhan, S.; Manley, J. L. Nature 2001, 413, 701–707.

16. Zhuang, X.; Rief, M. Curr Opin Struct Biol 2003, 13, 88–97.

17. Zhuang, X. W. Annu Rev Biophys Biomol Struct 2005, 34, 399–414.

18. Bokinsky, G.; Zhuang, X. W. Acc Chem Res 2005, 38, 566–573.

19. Cornish, P. V.; Ha, T. ACS Chem Biol 2007, 2, 53–61.

20. Walter, N. G.; Hampel, K. J.; Brown, K. M.; Burke, J. M. EMBO J 1998, 17, 2378–2391.

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

22. Lebruska, L. L.; Kuzmine, I. I.; Fedor, M. J. Chem Biol 2002, 9, 465–473.

23. Kuzmin, Y.; Da Costa, C.; Cottrell, J.; Fedor, M. J Mol Biol 2005, 349, 989–1010.

24. Kuzmin, Y.; Da Costa, C.; Fedor, M. J Mol Biol 2004, 340, 233–251.

25. Pljevaljcic, G.; Klostermeier, D.; Millar, D. P. Biochemistry 2005, 44, 4870–4876.

26. Cottrell, J. W.; Kuzmin, Y. I.; Fedor, M. J. J Biol Chem 2007, 282, 13498–13507.

27. Cai, Z.; Tinoco, I., Jr. Biochemistry 1996, 35, 6026–6036.

28. Butcher, S. E.; Allain, F. H.; Feigon, J. Nat Struct Biol 1999, 6, 212–216.

29. Butcher, S. E.; Allain, F. H.; Feigon, J. Biochemistry 2000, 39, 2174–2182.

30. Rupert, P. B.; Ferre-D'Amare, A. R. Nature 2001, 410, 780–786.

31. Rupert, P. B.; Massey, A. P.; Sigurdsson, S. T.; Ferre-D'Amare, A. R. Science 2002, 298, 1421–1424.

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

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

34. Rhodes, M.; Reblova, K.; Sponer, J.; Walter, N. Proc Natl Acad Sci USA 2006, 103, 13380–13385.

35. Tang, C. L.; Alexov, E.; Pyle, A. M.; Honig, B. J Mol Biol 2007, 366, 1475–1496.

36. Zhuang, X.; Kim, H.; Pereira, M. J.; Babcock, H. P.; Walter, N. G.; Chu, S. Science 2002, 296, 1473–1476.

37. Bokinsky, G.; Rueda, D.; Misra, V. K.; Rhodes, M. M.; Gordus, A.; Babcock, H. P.; Walter, N. G.; Zhuang, X. Proc Natl Acad Sci USA 2003, 100, 9302–9307.

38. Tan, E.; Wilson, T. J.; Nahas, M. K.; Clegg, R. M.; Lilley, D. M.; Ha, T. Proc Natl Acad Sci USA 2003, 100, 9308–9313.

39. Nahas, M. K.; Wilson, T. J.; Hohng, S.; Jarvie, K.; Lilley, D. M.; Ha, T. Nat Struct Mol Biol 2004, 11, 1107–1113.

40. Okumus, B.; Wilson, T. J.; Lilley, D. M.; Ha, T. Biophys J 2004, 87, 2798–2806.

41. Pljevaljcic, G.; Millar, D. P.; Deniz, A. A. Biophys J 2004, 87, 457–467.

42. Rueda, D.; Bokinsky, G.; Rhodes, M. M.; Rust, M. J.; Zhuang, X.; Walter, N. G. Proc Natl Acad Sci USA 2004, 101, 10066–10071.

43. Hampel, K. J.; Walter, N. G.; Burke, J. M. Biochemistry 1998, 37, 14672–14682.

44. Pinard, R.; Lambert, D.; Walter, N. G.; Heckman, J. E.; Major, F.; Burke, J. M. Biochemistry 1999, 38, 16035–16039.

45. Bevilacqua, P. C. Biochemistry 2003, 42, 2259–2265.

46. Torelli, A. T.; Krucinska, J.; Wedekind, J. E. RNA 2007, 13, 1052–1070.

47. Walter, N. G. Curr Prot Nucleic Acid Chem 2002, 11.10, 11.10.11–11.10.23.

48. Benkovic, S. J.; Hammes-Schiffer, S. Science 2003, 301, 1196–1202.

49. Liu, S.; Bokinsky, G.; Walter, N. G.; Zhuang, X. Proc Natl Acad Sci USA 2007, 104, 12634–12639.

50. Rodnina, M. V.; Beringer, M.; Wintermeyer, W. Q Rev Biophys 2006, 39, 203–225.

51. Ogle, J. M.; Ramakrishnan, V. Annu Rev Biochem 2005, 74, 129–177.

52. Schuwirth, B. S.; Borovinskaya, M. A.; Hau, C. W.; Zhang, W.; Vila-Sanjurjo, A.; Holton, J. M.; Cate, J. H. Science 2005, 310, 827–834.

53. Ramakrishnan, V. Cell 2002, 108, 557–572.

54. Blanchard, S. C.; Gonzalez, R. L.; Kim, H. D.; Chu, S.; Puglisi, J. D. Nat Struct Mol Biol 2004, 11, 1008–1014.

55. Blanchard, S. C.; Kim, H. D.; Gonzalez, R. L., Jr.; Puglisi, J. D.; Chu, S. Proc Natl Acad Sci USA 2004, 101, 12893–12898.

56. Munro, J. B.; Altman, R. B.; O'Connor, N.; Blanchard, S. C. Mol Cell 2007, 25, 505–517.

57. Daviter, T.; Murphy, F. V. T.; Ramakrishnan, V. Science 2005, 308, 1123–1124.

58. Valle, M.; Zavialov, A.; Li, W.; Stagg, S. M.; Sengupta, J.; Nielsen, R. C.; Nissen, P.; Harvey, S. C.; Ehrenberg, M.; Frank, J. Nat Struct Biol 2003, 10, 899–906.

59. Noller, H. F. Science 2005, 309, 1508–1514.

60. Noller, H. F. Biochimie 2006, 88, 935–941.

61. Sytnik, A.; Vladimirov, S.; Jia, Y.; Li, L.; Cooperman, B. S.; Hochstrasser, R. M. J Mol Biol 1999, 285, 49–54.

62. Rodnina, M. V.; Wintermeyer, W. Annu Rev Biochem 2001, 70, 415–435.

63. Lee, T. H.; Lapidus, L. J.; Zhao, W.; Travers, K. J.; Herschlag, D.; Chu, S. Biophys J 2007, 92, 3275–3283.

64. McKinney, S. A.; Joo, C.; Ha, T. Biophys J 2006, 91, 1941–1951.

65. Joo, C.; McKinney, S. A.; Nakamura, M.; Rasnik, I.; Myong, S.; Ha, T. Cell 2006, 126, 515–527.

66. Xie, Z.; Srividya, N.; Sosnick, T. R.; Pan, T.; Scherer, N. F. Proc Natl Acad Sci USA 2004, 101, 534–539.

67. Ha, T. Methods 2001, 25, 78–86.

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

69. Sabanayagam, C. R.; Eid, J. S.; Meller, A. J Chem Phys 2005, 123, 224708.

70. Heilemann, M.; Margeat, E.; Kasper, R.; Sauer, M.; Tinnefeld, P. J Am Chem Soc 2005, 127, 3801–3806.

71. Kapanidis, A. N.; Laurence, T. A.; Lee, N. K.; Margeat, E.; Kong, X.; Weiss, S. Acc Chem Res 2005, 38, 523–533.

72. Rasnik, I.; McKinney, S. A.; Ha, T. Nat Methods 2006, 3, 891–893.

73. Conley, N. R.; Pomerantz, A. K.; Wang, H.; Twieg, R. J.; Moerner, W. E. J Phys Chem B 2007, 111, 7929–7931.

74. Michaelis, L.; Menten, M. L. Biochem Z 1913, 49, 333–369.

75. Min, W.; English, B. P.; Luo, G.; Cherayil, B. J.; Kou, S. C.; Xie, X. S. Acc Chem Res 2005, 38, 923–931.

76. Kou, S. C.; Cherayil, B. J.; Min, W.; English, B. P.; Xie, X. S. J Phys Chem B 2005, 109, 19068–19081.

77. Min, W.; Gopich, I. V.; English, B. P.; Kou, S. C.; Xie, X. S.; Szabo, A. J Phys Chem B 2006, 110, 20093–20097.

78. English, B. P.; Min, W.; van Oijen, A. M.; Lee, K. T.; Luo, G.; Sun, H.; Cherayil, B. J.; Kou, S. C.; Xie, X. S. Nat Chem Biol 2006, 2, 87–94.

79. Walter, N. G. Nat Chem Biol 2006, 2, 66–67.

80. Ward, D. C.; Reich, E.; Stryer, L. J Biol Chem 1969, 244, 1228–1237.

81. Walter, N. G.; Burke, J. M. Methods Enzymol 2000, 317, 409–440.

82. Menger, M.; Eckstein, F.; Porschke, D. Nucleic Acids Res 2000, 28, 4428–4434.

83. Kirk, S. R.; Luedtke, N. W.; Tor, Y. Bioorg Med Chem 2001, 9, 2295–2301.

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

85. O'Neill, M. A.; Barton, J. K. J Am Chem Soc 2002, 124, 13053–13066.

86. Lafontaine, D. A.; Wilson, T. J.; Zhao, Z. Y.; Lilley, D. M. J Mol Biol 2002, 323, 23–34.

87. Harris, D. A.; Rueda, D.; Walter, N. G. Biochemistry 2002, 41, 12051–12061.

88. Bradrick, T. D.; Marino, J. P. RNA 2004, 10, 1459–1468.

89. Shandrick, S.; Zhao, Q.; Han, Q.; Ayida, B. K.; Takahashi, M.; Winters, G. C.; Simonsen, K. B.; Vourloumis, D.; Hermann, T. Angew Chem Int Ed Engl 2004, 43, 3177–3182.

90. Clerte, C.; Hall, K. B. Biochemistry 2004, 43, 13404–13415.

91. Gondert, M. E.; Tinsley, R. A.; Rueda, D.; Walter, N. G. Biochemistry 2006, 45, 7563–7573.

92. Tinsley, R. A.; Walter, N. G. RNA 2006, 12, 522–529.

93. Lemay, J. F.; Lafontaine, D. A. RNA 2007, 13, 339–350.

94. Tam, V. K.; Kwong, D.; Tor, Y. J Am Chem Soc 2007, 129, 3257–3266.

95. Srivatsan, S. G.; Tor, Y. J Am Chem Soc 2007, 129, 2044–2053.

96. Jean, J. M.; Hall, K. B. Proc Natl Acad Sci USA 2001, 98, 37–41.

97. Berry, D. A.; Jung, K.-Y.; Wise, D. S.; Sercel, A. D.; Pearson, W. H.; Mackie, H.; Randolph, J. B.; Somers, R. L. Tetrahedron Lett 2004, 45, 2457–2461.

98. Hsieh, J.; Andrews, A. J.; Fierke, C. A. Biopolymers 2004, 73, 79–89.

99. Rueda, D.; Hsieh, J.; Day-Storms, J. J.; Fierke, C. A.; Walter, N. G. Biochemistry 2005, 44, 16130–16139.

100. Kazantsev, A. V.; Pace, N. R. Nat Rev Microbiol 2006, 4, 729–740.

101. Staley, J. P.; Guthrie, C. Cell 1998, 92, 315–326.

102. Villa, T.; Pleiss, J. A.; Guthrie, C. Cell 2002, 109, 149–152.

103. Stevens, S. W.; Abelson, J. Methods Enzymol 2002, 351, 200–220.

Reviewing Editor: Sarah Woodson