doi:10.1016/S0022-2836(03)00272-9 *J. Mol. Biol.* (2003) **328**, 1011–1026

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Exploration of the Transition State for Tertiary Structure Formation between an RNA Helix and a Large Structured RNA

Laura E. Bartley[1], Xiaowei Zhuang[2], Rhiju Das[1,2], Steven Chu[2]* and Daniel Herschlag[1]*

[1]*Department of Biochemistry B400 Beckman Center Stanford University, Stanford CA 94305-5307, USA*

[2]*Department of Physics Stanford University, Stanford CA 94305-4060, USA*

**Corresponding authors*

Docking of the P1 duplex into the pre-folded core of the Tetrahymena group I ribozyme exemplifies the formation of tertiary interactions in the context of a complex, structured RNA. We have applied Φ-analysis to P1 docking, which compares the effects of modifications on the rate constant for docking (k_{dock}) with the effects on the docking equilibrium (K_{dock}). To accomplish this we used a single molecule fluorescence resonance energy transfer assay that allows direct determination of the rate constants for formation of thermodynamically favorable, as well as unfavorable, states. Modification of the eight groups of the P1 duplex that make tertiary interactions with the core and changes in solution conditions decrease K_{dock} up to 500-fold, whereas k_{dock} changes by ≤2-fold. The absence of effects on k_{dock}, both from atomic modifications and global perturbations, strongly suggests that the transition state for docking is early and does not closely resemble the docked state. These results, the slow rate of docking of 3 s^{-1}, and the observation that a modification that is expected to increase the degrees of freedom between the P1 duplex and the ribozyme core accelerates docking, suggest a model in which a kinetic trap(s) slows docking substantially. Nonetheless, urea does not increase k_{dock}, suggesting that there is little change in the exposed surface area between the trapped, undocked state and the transition state. The findings highlight that urea and temperature dependencies can be inadequate to diagnose the presence of kinetic traps in a folding process. The results described here, combined with previous work, provide an in-depth view of an RNA tertiary structure formation event and suggest that large, highly structured RNAs may have local regions that are misordered.

Keywords: RNA folding; Tetrahymena ribozyme; single molecule assay; kinetic trap; phi analysis

Introduction

To function, many RNA molecules form complex tertiary structures characterized by the packing together of helical secondary structure elements interlaced by single-stranded regions. The processes by which RNAs form functional tertiary structures has been studied intensely for a number of large and small RNA molecules (for recent reviews, see Treiber & Williamson,[1,2] Tinoco & Bustamante,[3] and Woodson[4]), including RNase P,[5–7] the Tetrahymena ribozyme,[8–14] and its P4–P6 domain,[15–17] several tRNAs,[18–22] and the small viral hairpin ribozyme.[23,24] For one of the best understood of these, the ribozyme derived from the self-splicing group I intron of *Tetrahymena thermophila* (Figure 1), tertiary folding from

Present address: X. Zhuang, Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA.

Abbreviations used: L- followed by a number, as in L-21 and L-16, indicates the position of the 5′ most residue of the ribozyme relative to the 5′ end of the parent intron; P and J followed by numbers and letters, as in P1 and J1/2, refer to the base-paired regions and joining regions in the ribozyme secondary structure; FRET, fluorescence resonance energy transfer; IGS, internal guide sequence; S, oligonucleotide substrate; G, guanosine; modifications to the P1 duplex are indicated by d for 2′-deoxyribose, and m for 2′-methoxy-2′-deoxyribose followed by the number of the residue modified; wt indicates the unmodified, all-ribose L-21 ribozyme with the 3′ extension for single molecule studies; wt, wild-type.

E-mail addresses of the corresponding authors: schu@stanford.edu; herschla@cmgm.stanford.edu

0022-2836/03/$ - see front matter

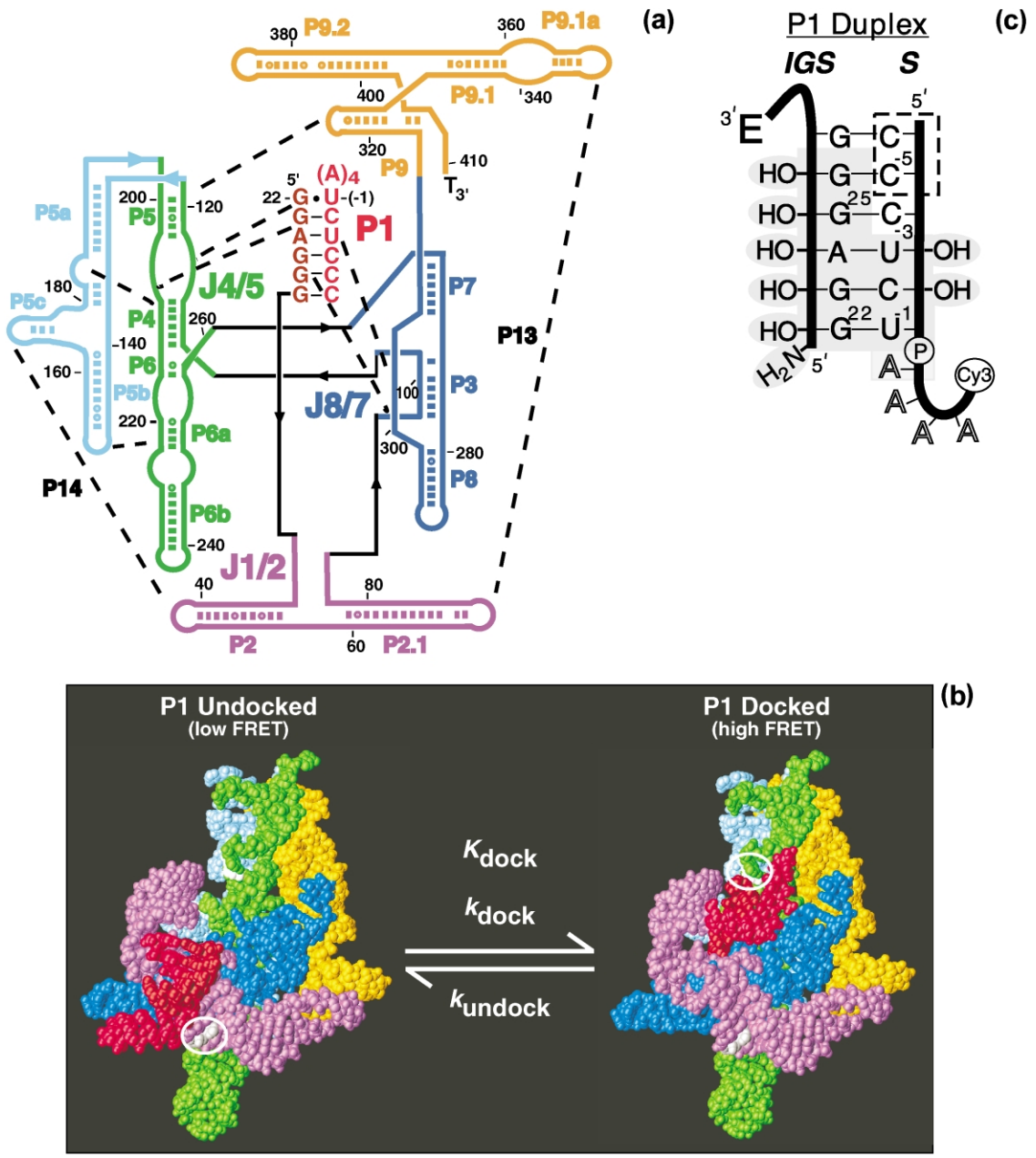

Figure 1. The Tetrahymena group I ribozyme and docking of its P1 duplex. (a) The ribozyme secondary structure colored according to structural sub-domains. Small numbers indicate the residue position relative to the cleavage site in the parent intron. Secondary structure elements are labeled sequentially with P, for paired region, and J, for joining region. Only the J regions that are discussed in the text are labeled. T, for tail, indicates the extension that is complementary to the Cy5-labeled tether used in surface attachment for single molecule measurements.[68] Prominent long-range tertiary contacts are indicated by broken lines.[51,52,119] (b) A model for P1 docking shown in the context of the tertiary structure model of the ribozyme.[52] Colors are as for (a). The undocked P1 duplex is shown near the cross-link to A88 (grey, circled) in P2 that is formed in the undocked state.[74] The docked P1 duplex is proximate to A114 (white, circled).[74,110] (c) The sequence of the P1 duplex, which is formed from the internal guide sequence (IGS) and the oligonuceotide substrate (S). S is the 5′-exon analog. The cleavage site phosphoryl group is represented by a circled P, and the remainder of the ribozyme is indicated by E. The Cy3 on the 3′ end of the substrate is the donor dye in FRET studies. The numbers indicate the residue positions relative to the cleavage site in the parent intron, where a minus sign (−) signifies that a residue is 5′ of the cleavage site. In previous studies each residue of the P1 duplex has been substituted with a deoxyribose residue.[41,47] Groups that were previously found to change the equilibrium for docking are shown explicitly and are indicated by the underlying grey shading. The L-16 ribozyme has an extended IGS (5′-GGUUUGGAGGG-3′), which is used with a complementary substrate (5′-CCCUCUAAACC-3′).[27] In the short substrate, CUCU, the residues within the broken box, C(−5) and C(−6), are absent.

secondary structure requires magnesium or another divalent metal ion,[25,26] occurs *via* multiple intermediates,[9,12,27] and can follow multiple pathways.[10,13,14,27] The first steps of folding involve rapid compaction[28] followed by formation of local and long-range tertiary interactions.[9,12] Depending on solution conditions, tertiary structure formation can be accompanied by secondary structure rearrangements.[27,29–32] The late steps of folding can include resolution of kinetically trapped species within the context of a molecule that is highly structured.[11,14,29]

The study of RNA folding is motivated by the desire to understand the processes by which RNAs attain functional structures. Here we examine what appears to be the last step of folding[33,34] and an initial step of the chemical reaction of the Tetrahymena ribozyme: docking of the P1 duplex formed between the ribozyme internal guide sequence (IGS) and the 5′ splice site analog (S) into the ribozyme active site (Figure 1(b) and (c)).[35,36] P1 duplex docking may allow examination of tertiary structure formation with isolated secondary structure in the absence of other complicating events. As docking occurs in the context of a highly structured RNA, it may be an especially good system for understanding late tertiary folding events.

A further motivation for understanding the attainment of structure by complex RNAs is the prevalence of multi-step reaction cycles involving native state conformational changes within biologically important RNA systems. Splicing, as carried out by group I and group II self-splicing introns as well as the ribonucleoprotein spliceosome, involves conformational changes between the first and second steps of splicing.[37] Throughout translation, numerous conformational changes of the ribosome have been observed (e.g. see Frank & Agarwaal[38]); the most conspicuous of these is the translocation of tRNAs through the core of ribosome particle.[39,40] As a reaction step of the Tetrahymena ribozyme, P1 duplex docking represents a conformational change within the native state of the ribozyme and may give insight into conformational changes in complex RNAs.

The features of the P1 duplex in the undocked state have led to a model that P1 is an isolated secondary structure element. The undocked P1 duplex has the same stability as a duplex free in solution.[41] Further, the P1 duplex forms only tenfold less rapidly than an isolated duplex[42,43] with a rate that is largely insensitive to temperature and length as expected for a free duplex (e.g. see Williams *et al.*[44]). Upon docking, biochemical studies suggest a model in which the minor groove of the P1 duplex[32,41,42,45–49] forms a triplex with the J4/5 and J8/7 segments of the ribozyme core.[45,50,51] The J4/5 and J8/7 regions in turn contact all other regions of the ribozyme core, either covalently or through tertiary interactions (Figure 1(b)).[50–52] Unlike many other binding reactions involving RNA, the P1 duplex and the native state of the Tetrahymena ribozyme core have been suggested to be largely preformed in the absence of P1 docking.[25,34,50,53–55]

To test this model and better understand the process of P1 duplex docking, we measured the rate and equilibrium constants for docking under a variety of conditions. We employed Φ-analysis, a powerful formalism developed by Fersht and colleagues[56] that compares the properties of the transition state to those of the initial and final states for the reaction of interest. Φ-analysis has been used extensively to examine protein folding (reviewed by Plaxco *et al.*[57] and Oliveberg[58])[59–65] and binding reactions[66,67] and has begun to be applied to RNA folding.[17,20,68] Here, the reaction is docking of the P1 duplex. The initial state is undocked and the final state is docked. $\Phi_{\text{Undocked}\rightarrow\text{Transition State }(\ddagger)}$, hereinafter Φ, compares the energetic effect of a modification on the rate constant for docking with the energetic effect of a modification on the equilibrium for docking (equation (1)).

$$\Phi = \frac{\Delta\Delta G(k_{\text{dock, modified}\rightarrow\text{unmodified}})}{\Delta\Delta G(K_{\text{dock, modified}\rightarrow\text{unmodified}})} \qquad (1)$$

A Φ-value of ~1 suggests that the interactions with the modified group are made at or before the rate limiting transition state, implying that the modified group is in a similar environment in the transition state as in the docked state. A Φ-value of ~0 is most simply interpreted to indicate that the interactions with the modified group have not yet been made in the transition state.

Our initial study, with substitution of a single group involved in tertiary contacts between P1 and the core, found no effect on k_{dock} relative to that for wild-type (wt), yielding a Φ-value of ~0.[68] Here, we have greatly expanded our examination of the docking process by measuring the effects of substitutions at all positions of P1 known to make tertiary interactions, as well as the effects of varying salt, temperature, and urea. The data suggest that the transition state for docking is highly dissimilar from the docked state. Further, we propose a model in which the ribozyme is not locally prearranged to allow docking, a situation that may generally hold even for large, cooperatively folded RNAs.

Results

Docking of the P1 duplex with the core of the Tetrahymena group I ribozyme was examined by measuring the effects of various perturbations on docking kinetics and thermodynamics. Rate and equilibrium constants were obtained with a single molecule FRET assay for docking that we established previously.[68] The single molecule assay allows states that do not accumulate to be directly observed (reviewed by Weiss[69]). Thus, the rate constants for both docking and undocking can be measured directly even when the equilibrium is

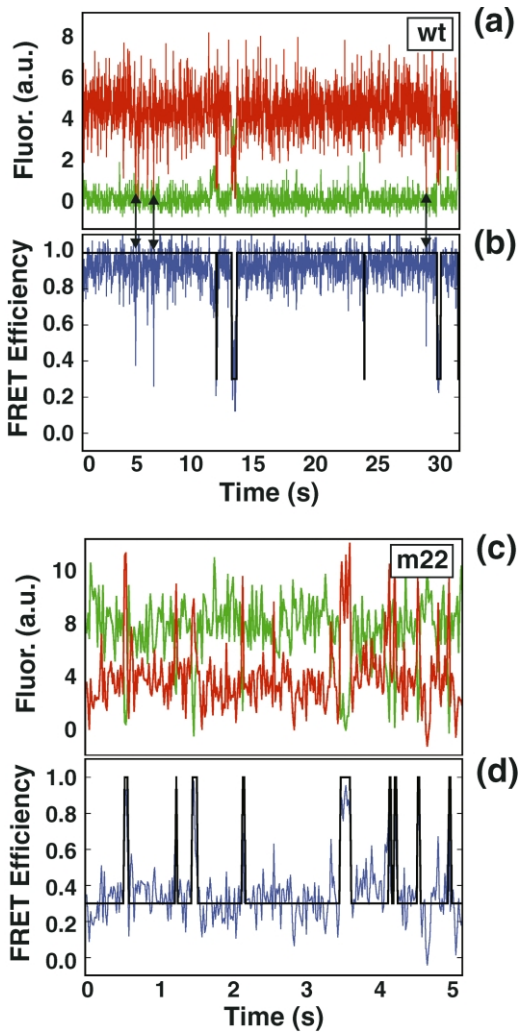

Figure 2. Representative fluorescence data for individual ribozymes with different P1 duplexes: (a) and (b) wild-type; (c) and (d) m22. In each pair of panels, the top panel (a, c) shows the fluorescence signals, after background correction, with the donor signal in green (D, Cy3) and the acceptor signal in red (A, Cy5). The bottom panels (b, d) show the FRET efficiency, defined as $[I_A/(I_A + I_D)]$, in blue, and the assignment of the docked and undocked states according to primary and secondary screening criteria, in black. Assignment of an event is based on the fluorescence characteristics at each time bin. Docked events are identified by a FRET value of ≥ 0.75 and undocked events are identified by a FRET value of ≤ 0.5. In addition to undocking, reductions in FRET are also caused by a photophysical "blinking" events in which either dye enters into a relatively long-lived ($\tau \approx 20$ ms) non-emitting state, as seen in the trace of wild-type (wt) (double-headed arrows). (Blinking events are also observed with other ribozyme complexes besides wt.) We distinguished low FRET signals due to undocking from those due to blinking primarily based on a loss in total fluorescence signal, which generally accompanies blinking but not undocking. However, signal-to-noise was not always sufficient to allow all blinking events to be excluded, resulting in a fast phase in some undocked time plots. Data analysis with respect to blinking is described in Supplementary Material.

far from unity. Modifications in data collection and analysis are outlined in Materials and Methods and elaborated further in Supplementary Material. Representative docking time traces for single ribozymes are shown in Figure 2, and the data described below are summarized in Table 1 and Figure 4.

Modifications of the P1 duplex

The effects on docking of modification of each of the eight groups on the P1 duplex known to make tertiary interactions are summarized in Table 1(A). Specific 2′-hydroxyl (−OH) groups of the ribose rings of P1 were replaced with 2′-deoxyribose (d; −H) and/or 2′-methoxyribose (m; $-OCH_3$) residues (Figure 1(c)). The role of the exposed exocyclic amine of the G·U wobble pair at the cleavage site was probed by replacement with a hydrogen atom in the I22 construct (I for inosine, in which the exocyclic amine is replaced with a hydrogen atom) and replacement of the U residue with C to form a G·C Watson–Crick base-pair. All of the modifications gave significant effects on the docking equilibrium under our experimental conditions. Quantitative differences between these and previously published results[32,41,47,49] are attributed to the different conditions and constructs employed and are discussed in Supplementary Material.

Figure 3(b) and (c) shows representative kinetic data for the wild-type (wt) and two modified constructs. For these and all other site-specific modifications that disrupt docking, the effect is exerted nearly entirely on the rate constant for undocking (k_{undock}) (Figure 3(c)), with $\leq$2-fold effects on the rate constant for docking (k_{dock}) (Figure 3(b)). The small apparent difference in k_{dock} value between the wild-type and modified constructs in Figure 3(b) results from photophysical effects, which have been assessed quantitatively (Figure 3(a)) and accounted for as described in Materials and Methods and Supplementary Material.

Other perturbations of docking

To gain further insight into the docking process, the effect of several other factors on the docking rate and equilibrium constants were also examined. Binding of the guanosine substrate (G) to the ribozyme increases the binding affinity of the oligonucleotide substrate.[70] To measure this effect we replaced the 2′-OH at U(−1), the cleavage site (Figure 1(c)), with hydrogen [d(−1)] to prevent cleavage of the oligonucleotide substrate (S) on the time scale of these measurements.[71] We also measured the docking kinetics for a substrate with a 2′-H at position −3 [d(−3)], as this substitution decreases the docking equilibrium constant to near 1 and thereby increases the accuracy in determination of the rate and equilibrium constants as explained in Supplementary Material. Guanosine

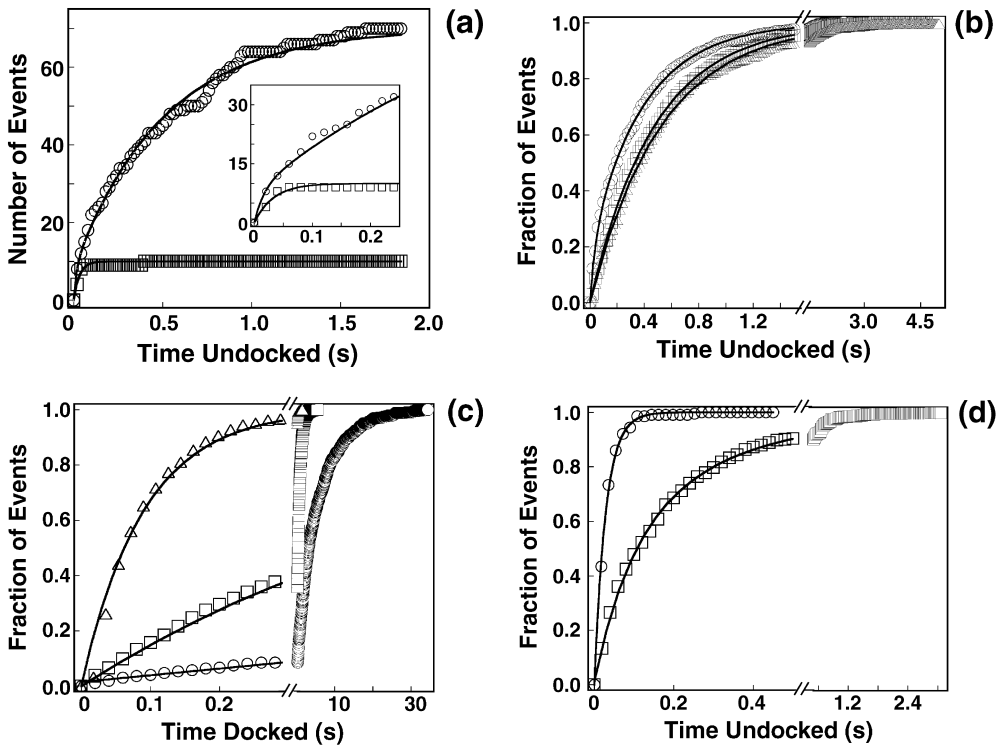

Figure 3. Analysis of docking and undocking kinetics. Plots show the cumulative number of events observed *versus* length of time spent in the docked or undocked state, or dwell time, for each event, i.e. the sum of the number of events observed with a dwell time less than or equal to the value on the *x*-axis. (a) Comparison of the low FRET events that pass the filtering criteria for undocking for wt ribozyme (○) and for a DNA duplex with a Cy3–Cy5 pair at one end (□) from 840 seconds of data taken at similar laser intensity. The inset is an expansion of the short times. The DNA data are fit by a single exponential, $N(t) = N[1 - \exp(-kt)]$, where N is the number of events observed. The ribozyme data are fit by a double exponential, $N(t) = N[1 - A^{(1)} \exp(-k^{(1)}t) - A^{(2)} \exp(-k^{(2)}t)]$. Consistent with the fast phase in the wt ribozyme data being caused by photophysics, a similar number of short events ($N = 10$, DNA, □; $A^{(1)} = 8$, ribozyme, ○) and magnitude of rate constant associated with them ($k = 30(\pm 10)\ s^{-1}$, DNA; $k^{(1)} = 60(\pm 20)\ s^{-1}$, ribozyme) was identified for the DNA and wt ribozyme. Thus, the rate constant of the fast phase in the undocked time histograms is attributed to blinking of the dyes and approximates the rate constant for leaving the blinked state, $k_{unblink}$. Irregularities in the data are the result of noise due to the small data set. (b and c) Plots of the times spent in the undocked (b) and docked states (c) for the wt (○), the m(−1) (□), and m22 (△) ribozymes. The number of events observed (N) is normalized to 1 so that the data may be easily compared. The data sets are composed of data from observations made on multiple days. In (b), the wt undocking data (○; $N = 680$) are fit by a double exponential ($A^{(1)} = 0.2$, $k^{(1)}_{dock} = 23(\pm 4)\ s^{-1}$, $A^{(2)} = 0.8$, $k^{(2)}_{dock} = 2.5(\pm 1)\ s^{-1}$), and the m(−1) (□; $N = 1856$; $k_{dock} = 2.1(\pm 0.1)\ s^{-1}$) and m22 (△; $N = 645$; $k_{dock} = 1.9(\pm 0.2)\ s^{-1}$) data are fit by a single exponential equation. The rate constant of the fast phase for wt, $k^{(1)}_{dock}$ likely reflects blinking, and thus is designated as $k_{unblink}$. Such a fast phase was also observed for most mutants examined but was able to be eliminated by secondary screening criteria from data sets for which $K_{dock} <\sim 5$ (see Supplementary Material). The data for m22 exclude three very long, statistically unlikely undocked events. In (c), the wt data (○; $N = 516$) are best fit by a double exponential, as shown ($A^1 = 0.21$, $k^{(1)}_{undock} = 0.7(\pm 0.2)\ s^{-1}$, $A^2 = 0.79$, $k^{(2)}_{undock} = 0.15(\pm 0.01)\ s^{-1}$); fitting by a single exponential gives $k_{undock} = 0.19(\pm 0.01)\ s^{-1}$. The m(−1) (□; $N = 1497$; $k_{undock} = 1.6(\pm 0.1)\ s^{-1}$) and m22 (△; $N = 666$; $k_{undock} = 11(\pm 1)\ s^{-1}$) data are fit by a single exponential equation. (d) Comparison of dwell times in the undocked state for the short S, d(−1)CUCU (○), *versus* full length (−1d)S (□) with the L-16 ribozyme. The short S data are fit by a single exponential (○; $N = 300$; $k_{dock} = 35(\pm 5)\ s^{-1}$). The full length S data are fit by a double exponential (□; $N = 563$) ($A^{(1)} = 0.5$, $k^{(1)}_{dock} = 11(\pm 2)\ s^{-1}$, $A^{(2)} = 0.5$, $k^{(2)}_{dock} = 3.3(\pm 0.4)\ s^{-1}$). Despite the similarity of k_{dock} for the short substrate to $k^{(1)}_{dock}$ (i.e. $k_{unblink}$) for full-length substrates (L-21 and L-16), the greater frequency of low FRET events for the short S relative to full-length S and the absence of a second, slower phase in the undocked time plot for the short S suggest that the reported rate constant is due to docking and not simply to blinking (see Supplementary Material for elaboration).

has an ~3-fold effect on K_{dock}, consistent with previous results at this temperature.[72] Within error, the effect of bound guanosine is manifested in a smaller k_{undock} value with no effect on k_{dock} (Table 1(B)).

Increasing the $MgCl_2$ concentration from 10 mM to 100 mM had only a ~2-fold effect on K_{dock} and had no effect within error on the rate constant for docking (Table 1(B)). Similarly, previous measurements detected no change in k_{dock} between 5 mM and 15 mM $MgCl_2$.[73] In contrast to the absence of effect with increased Mg^{2+}, addition of 500 mM Na^+ decreased the equilibrium for docking by ~50-fold, consistent with previous

Table 1. Perturbations of the thermodynamics and kinetics of docking

A. Effects of site-specific modifications to the P1 duplex on docking[a]

P1 duplex[b]	K_{dock}[c]	$\Delta\Delta G$ (K_{dock})[d] (kcal mol^{-1})	k_{dock}[e] (s^{-1})	$\Delta\Delta G^{\ddagger}$ (k_{dock})[f] (kcal mol^{-1})	k_{undock}[e] (s^{-1})	$-\Delta\Delta G^{\ddagger}$ (k_{undock})[g] (kcal mol^{-1})	Φ[h]
wt	25 ± 3[i]	–	2.5[j] ± 0.2	–	0.07[k] ± 0.02	–	–
d(−1)	22 ± 5	–	2.7[j] ± 0.4	–	0.13[l] ± 0.04	–	–
d(−1)[m]	32 ± 4	–	3.3[j] ± 0.4	–	0.15[k] ± 0.02	–	–
d(−2/−1)	0.8 ± 0.1	2.0 ± 0.1	1.5[n] ± 0.1	0.3 ± 0.1	1.9[o] ± 0.2	1.6 ± 0.2	0.2 ± 0.1
d(−3/−1)	0.36 ± 0.04	2.4 ± 0.1	1.6[n] ± 0.1	0.3 ± 0.1	3.7[o] ± 0.3	2.0 ± 0.2	0.13 ± 0.04
I22[m]	1.7 ± 0.2	1.7 ± 0.1	3.8[n] ± 0.3	−0.1 ± 0.1	1.7[o] ± 0.1	1.4 ± 0.4	−0.05 ± 0.05
C(−1)	0.40 ± 0.03	2.4 ± 0.1	2.6[n] ± 0.1	0.0 ± 0.1	5.6[o] ± 0.3	2.5 ± 0.2	−0.01 ± 0.02
d22[m]	0.37 ± 0.05	2.6 ± 0.1	2.6[n] ± 0.2	0.1 ± 0.1	6.2[o] ± 0.5	2.2 ± 0.4	0.05 ± 0.03
d23	12 ± 3	0.4 ± 0.1	3.8[n] ± 0.6	−0.3 ± 0.1	0.4[l] ± 0.1	1.0 ± 0.3	–[p]
d24	30 ± 7	−0.1 ± 0.2	4.1[n] ± 0.8	−0.3 ± 0.1	0.14[k] ± 0.06	0.4 ± 0.2	–[p]
d25	4 ± 1	1.0 ± 0.1	4.0 ± 0.7	−0.3 ± 0.1	1.0 ± 0.2	1.5 ± 0.2	−0.3 ± 0.2
d25[m]	1.1 ± 0.1	2.0 ± 0.1	3.6 ± 0.2	−0.1 ± 0.1	2.6[o] ± 0.2	1.7 ± 0.4	0.00 ± 0.01
d26	4.1 ± 0.6	1.1 ± 0.1	4.3 ± 0.5	−0.3 ± 0.1	1.2 ± 0.1	1.6 ± 0.2	−0.3 ± 0.1
m(−1)	1.3 ± 0.1	1.7 ± 0.1	2.2[n] ± 0.1	0.1 ± 0.1	1.6[o] ± 0.1	1.8 ± 0.2	0.04 ± 0.03
m(−3)	0.14 ± 0.01	3.0 ± 0.1	1.6[n] ± 0.1	0.3 ± 0.1	9.7[o] ± 0.5	2.9 ± 0.2	0.09 ± 0.02
m(−3)[m]	0.77 ± 0.05	2.2 ± 0.1	3.7 ± 0.2	−0.1 ± 0.1	5.2 ± 0.2	2.1 ± 0.1	−0.03 ± 0.02
m22	0.17 ± 0.02	2.9 ± 0.1	2.6[n] ± 0.2	0.0 ± 0.1	15[o] ± 1	3.1 ± 0.2	−0.01 ± 0.02
m25	0.04 ± 0.01	3.7 ± 0.1	2.7[n] ± 0.3	0.0 ± 0.1	62[o] ± 6	3.9 ± 0.2	−0.01 ± 0.02
C(−1)/d25	0.22 ± 0.03	2.8 ± 0.1	3.9[n] ± 0.4	−0.3 ± 0.1	18[o] ± 2	3.2 ± 0.2	−0.09 ± 0.03
m(−3/22)	0.07 ± 0.03	3.4 ± 0.3	2.6[n] ± 0.7	0.0 ± 0.2	31[o] ± 8	3.5 ± 0.2	−0.01 ± 0.05
m(−1/22)	0.12 ± 0.02	3.1 ± 0.1	3.5[n] ± 0.5	−0.2 ± 0.1	16[o] ± 2	3.2 ± 0.2	−0.06 ± 0.03
d(−1)CUCU[m]	90 ± 20	−0.6 ± 0.2	34 ± 5	−1.4 ± 0.1	0.3[l] ± 0.1	0.1 ± 0.1	2.4 ± 0.6[q]

B. Effects of guanosine binding and increased monovalent and divalent metal ion concentration on docking[a]

Conditions[r]	K_{dock}[c]	$\Delta\Delta G$ (K_{dock})[d] (kcal mol^{-1})	k_{dock}[e] (s^{-1})	$\Delta\Delta G^{\ddagger}$ (k_{dock})[f] (kcal mol^{-1})	k_{undock}[e] (s^{-1})	$-\Delta\Delta G^{\ddagger}$ (k_{undock})[f] (kcal mol^{-1})	Φ^{app} [s]
–	22 ± 5	–	2.7[j] ± 0.4	–	0.13[k] ± 0.04	–	–
2 mM G	47 ± 14	−0.4 ± 0.2	2.1[j] ± 0.5	0.1 ± 0.2	0.05 ± 0.04	−0.6 ± 0.5	−0.3 ± 0.4
–[t]	0.36 ± 0.04	–	1.6[n] ± 0.1	–	3.7[l] ± 0.3	–	–
2 mM G[t]	1.6 ± 0.3	−0.9 ± 0.1	1.7 ± 0.2	0.0 ± 0.1	1.3[o] ± 0.2	−0.6 ± 0.1	0.0 ± 0.1
–	22 ± 5	–	2.7[j] ± 0.4	–	0.13 ± 0.04	–	–
500 mM NaCl	0.4 ± 0.1	2.4 ± 0.2	3.6 ± 0.3	−0.2 ± 0.1	7.4 ± 0.4	2.4 ± 0.2	−0.07 ± 0.04
100 mM $MgCl_2$	40 ± 10	−0.4 ± 0.2	2.5[j] ± 1.0	0.0 ± 0.3	0.10 ± 0.05	−0.2 ± 0.4	−0.1 ± 0.8

[a] Measured at 22 °C, 10 mM $MgCl_2$ (pH 7.0), unless specified otherwise.
[b] See Figure 1(c) for a diagram of the P1 duplex and the abbreviations for a description of nomenclature.
[c] K_{dock} from total time spent in the docked state divided by total time spent in the undocked state.
[d] $\Delta\Delta G\ (K_{dock}) = -RT\ln(K_{dock}^{mut}/K_{dock}^{wt})$, where R is the gas constant, and T is the temperature in Kelvin.
[e] Rate constants are adjusted to account for systematic biases in data collection as described in Materials and Methods and Supplementary Material and noted here. The general form of equations used is given in the legend to Figure 3.
[f] $\Delta\Delta G\ (k_{dock}) = -RT\ln(k_{dock}^{mut}/k_{dock}^{wt})$.
[g] $-\Delta\Delta G\ (k_{undock}) = -RT\ln(k_{undock}^{wt}/k_{undock}^{mut})$.
[h] $\Phi = \Delta\Delta G^{\ddagger}\ (k_{dock})/\Delta\Delta G\ (K_{dock})$.
[i] Errors are two standard deviations based on Poisson statistics and are calculated as $1/\sqrt{N}$, where N is the number of events observed for a given perturbation.
[j] $k_{dock}^{(2)}$, the rate constant for the second phase of the undocked time histogram, which excludes photophysical events.
[k] $k_{undock} = k_{undock,obs}^{(2)} - k_{photobleach} - 1/(\text{trace length})$.
[l] $k_{undock} = k_{undock} - k_{photobleach} - 1/(\text{trace length})$.
[m] Measured in the context of the L-16 ribozyme which has a five residue extension of the IGS compared to the L-21 ribozyme.[27]
[n] k_{dock}, corrected for undercounting of short undocked events.
[o] k_{undock}, corrected for undercounting of short docked events.
[p] Φ-values are not given for modifications that have <1 kcal mol^{-1} effect on K_{dock}, as the error associated with these measurements is large relative to the size of the effect.
[q] This is a Φ^{app}-value because shortening the substrate is a complex modification (see footnote **s**).
[r] All measurements in the context of d(−1) substrate with the L-21 ribozyme, unless otherwise specified.
[s] $\Phi^{app} = \Delta\Delta G^{\ddagger}(k_{dock})/(\Delta\Delta G(K_{dock})$. The superscript app, for apparent, indicates that while the ratio of $\Delta\Delta G^{\ddagger}/\Delta\Delta G$ may still be used to compare the properties of the transition state with those of the ground state, these perturbations are not described by typically defined Φ-values.
[t] For the P1 duplex with the d(−1/−3) modifications.

cross-linking results.[74] The effect arises from an increase in the rate of undocking, whereas k_{dock} is affected <2-fold (Table 1(B)).

The global properties of the docking process were probed by measuring the urea and temperature dependencies of docking. Figure 4(a) shows that while the docking equilibrium is greatly reduced by urea, there is little effect on the docking rate constant; the entire effect is expressed in the undocking rate constant. Figure 4(b) shows that increasing temperature renders docking more favorable, and both the docking and undocking rate constants increase, in qualitative agreement with previous results.[74]

Lastly, the effect of increasing the length of the single-stranded region between the P1 duplex and the ribozyme was probed by shortening the oligonucleotide substrate by two residues, i.e. removing residues (−5) and (−6) (Figure 1(c)). Previous studies found effects of <3-fold on K_{dock} due to removal of the 2′-hydroxyl groups or the entire residue at these positions.[41,43,46,47] The shortened P1 duplex docks somewhat more strongly (~3-fold) than the full length duplex (Table 1(A)). Interestingly, the rate of docking of the short P1 duplex is ~10-fold faster than that of full length (Figure 3(d)). The undocking rate constant is also increased modestly.

Discussion

Docking of the P1 duplex into the pre-folded Tetrahymena ribozyme core provides an opportunity to examine a tertiary structure formation event in the context of a large, highly structured RNA (Figure 1(b)). As noted in the Introduction, this may also provide a model for conformational transitions of functional RNAs such as the ribosome and spliceosome.

To dissect the docking process we used Φ-analysis, which allows the development of a picture of the transition state for docking relative to the undocked and docked states, using equation (1).[56] It has been emphasized that modifications used in Φ-analysis should be conservative to reduce effects on the structure and solvation of the ground state structure.[56,59] Especially in the context of a rigid duplex, the single atom modifications made here are unlikely to have large effects on the undocked state. More generally, the rigidity of RNA secondary structure elements and the ability to modify individual functional groups may maximize the power of Φ-analysis.

The transition state for P1 docking

Figure 5 depicts simple models for docking in which different possible transition states contain progressively larger numbers of contacts between the P1 duplex and the ribozyme core. The three models shown represent a continuum of possi-

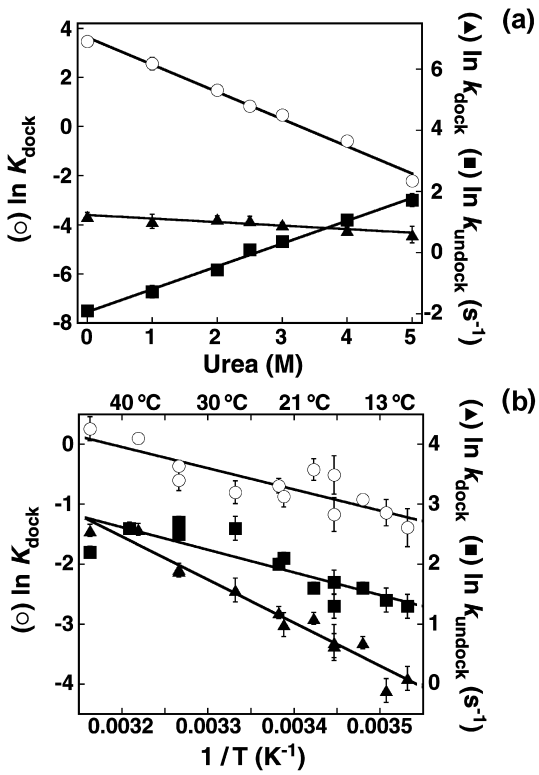

Figure 4. Effect of urea (a) and temperature (b) on the thermodynamics and kinetics of docking. K_{dock} (○) is from the total time spent in the docked state divided by the total time spent in the undocked state. k_{dock} (▲) and k_{undock} (■) are from fits of dwell time data as described for Figure 3. Measurements were conducted at 22 °C in 10 mM $MgCl_2$ (pH 7.0). Errors are two standard deviations from Poisson statistics and are smaller than the symbols if not shown. Tables with the individual values are provided in Supplementary Material. For (a) the lines are linear best fits to the log plot shown according to the equation: $\Delta G(\text{urea}) = \Delta G + m[\text{urea}]$. The resulting slopes (in kcal mol^{-1}) are $m(K_{dock}) = 0.65(\pm 0.05)$, $m^{\ddagger}(k_{dock}) = 0.07(\pm 0.04)$, and $m^{\ddagger}(k_{undock}) = -0.44(\pm 0.05)$. The relatively large errors in m-values and imperfect agreement between the m-values determined from the rate *versus* the equilibrium data arise largely from limitations in the data at 0 M and 5 M urea. At 0 M urea there is considerable error in k_{undock} due to the small number of events observed and other biases (see Supplementary Material), and at 5 M urea the data are consistent with the presence of two ribozyme populations with different rate and equilibrium behavior (data not shown). Omission of these points gives improved agreement but does not significantly change the m-values, nor does it effect the conclusions drawn. The ratio $m^{\ddagger}(k_{dock})/m(K_{dock})$, expressed as $\beta^{\ddagger}$, has a value of $0.1(\pm 0.1)$. Data are from four independent experiments. For (b) the equilibrium data are fit by the van't Hoff equation: $K_{dock} = \exp(\Delta H/RT - \Delta S/R)$, yielding $\Delta H = 8(\pm 2)$ kcal mol^{-1} and $\Delta S = 25(\pm 8)$ cal mol^{-1} K^{-1}. Rate constants are fit by the Eyring equation: $k = \alpha \exp(\Delta H^{\ddagger}/RT - \Delta S^{\ddagger}/R)$. $\Delta H^{\ddagger}(k_{dock}) = 12(\pm 2)$ kcal mol^{-1} and $\Delta H^{\ddagger}(k_{undock}) = 5(\pm 3)$ kcal mol^{-1}. $\Delta S^{\ddagger}(k_{dock}) = -14$ to 13 cal mol^{-1} K^{-1} and $\Delta S^{\ddagger}(k_{undock}) = -37$ to -10 cal mol^{-1} K^{-1}, where the lower activation entropies use a value for α that is estimated based on polymer folding in solution, and the higher are based on an estimate for covalent bond vibration (see footnote on page 1019†). Data are from three independent experiments.

bilities. In the first transition state model, the P1 duplex is localized near the active site but has not yet formed contacts with the core ($\ddagger_1$). The transition state is "early" in the docking process, and modifications that disrupt interactions in the docked state have no effect on the rate constant for docking ($\Phi = 0$). Alternatively, the transition state for docking may be later in the docking process, containing a few ($\ddagger_2$) or many ($\ddagger_3$) contacts between the P1 duplex and the core. In the simplest case, formation of a particular set of interactions in the transition state is required for docking. Modifications that interfere with those interactions decrease the rate constant for docking and destabilize the docked state to the same degree ($\Phi \approx 1$).

As summarized in Table 2, modifications that affect specific interactions in the docked state and changes in solution conditions that may act less site-specifically have large effects on the docking equilibrium, up to 3.7 kcal mol^{-1}, but negligible effects on the rate constant for docking (± 0.3 kcal mol^{-1}). Thus, a transition state for docking that is highly dissimilar from the docked state is suggested. Below we describe the significance of the individual measurements as well as limitations of this analysis. The subsequent sections use these and additional data to further develop our model for this folding process.

Removal of each of the functional groups on the P1 helix involved in tertiary interactions and a replacement with a hydrogen atom (e.g. 2′-OH → 2′-H) give small or negligible effects on the docking rate constant and Φ-values that are within error of 0.1 (Table 1(A)). The exception, d26, has a small negative Φ-value because docking appears to be slightly faster; nevertheless, this rate constant is within twofold of that for wild-type and provides no evidence for a transition state interaction. The possible significance of small variations in the docking rate constant upon modification is addressed below.

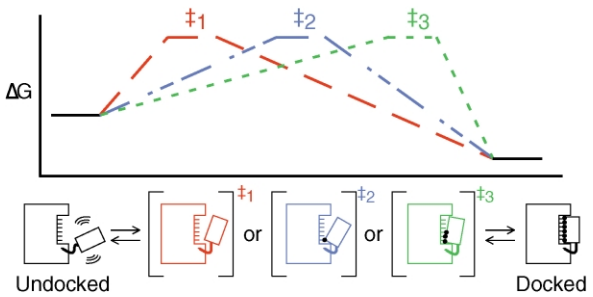

Figure 5. Simple models for the transition state for docking that represent the continuum of possibilities: ($\ddagger_1$) (broken line) is an "early" transition state in which no contacts are made with the core; ($\ddagger_2$) (dot-dash line) represents an "intermediate" transition state in which a few contacts are formed between P1 and the core; ($\ddagger_3$) (dotted line) represents a "late" transition state in which many contacts are formed between P1 and the core. Tertiary contacts between P1 and the core are depicted as bold dots.

As all positions gave Φ-values close to zero, the simplest interpretation is an absence of native contacts in the folding transition state. Nevertheless there are important alternative models for Φ-values that are not unity.[56] First, tertiary contacts that are formed in the transition state can give small Φ-values if there is less cost to accommodating the modification, because the surrounding environment is less structured in the transition state than in the final state.[61,75,76] Second, multiple folding pathways with distinct transition states will result in small Φ-values if a contact is formed in only a fraction of the transition states.[63] While the alternative models cannot be eliminated, the experiments described below designed to probe these possibilities gave no indication of their applicability to docking.

The analysis described above involved replacement of functional groups by hydrogen atoms. To test whether the energetic consequence of disrupting contacts is greatly reduced in the transition state and therefore undetected upon removal of the 2′-OH group, we also introduced 2′-OCH_3 substitutions at several positions (Table 1(A)). Despite the substantially larger equilibrium effects of the bulky substituents, the Φ-values for the 2′-OCH_3 substitutions remain near zero and do not increase relative to those for the 2′-H substitutions. These results provide no indication of neighboring contacts with reduced energetic consequences in the docking transition state.

To test whether the near-zero Φ-values arose from multiple folding transition states, we examined constructs with two modifications made simultaneously. Within error the double modifications have no effect on k_{dock}, similar to the single modifications. This provides no evidence for multiple transition states with different P1 contacts. Nevertheless, the experimental uncertainty in these measurements renders this test unable to discern situations in which there are more than a few competing transition states†.

To further probe the docking transition state, we followed the response of docking kinetics and thermodynamics to several types of more global perturbations. In each case, k_{dock} is insensitive to the perturbation, bolstering the view that the docking transition state is early and highly distinct from the docked complex. As the

† For example, if there are three distinct transition states, each containing distinct subsets of interactions, a modification that fully destabilizes one of the three transition states, preventing it from being accessed, would decrease the rate constant for docking by one-third, i.e. from 3 s^{-1} to 2 s^{-1}. Disruption of two of the three possible transition states, for example, by making two key modifications, would decreases the rate constant for docking by two-thirds, i.e. from 3 s^{-1} to 1 s^{-1}.

Table 2. Summary of perturbations and their effects on the transition state for docking

Perturbation	Class[a]	Effect on K_{dock}[b]	Effect on k_{dock}[b]	Fractional effect on U → ‡ *versus* U → D[c]
Interference with a tertiary interaction	Site-specific	$\leq +3.7$ kcal mol^{-1}	± 0.3 kcal mol^{-1}	~ 0 ($\Phi \leq 0.3$)
Bound guanosine	Site-specific	~ -0.6 kcal mol^{-1}	$+0.1$ kcal mol^{-1}	~ 0 ($\Phi^{app} \leq -0.3$)
Magnesium (100 mM)	Site-specific and/or global	< -0.3 kcal mol^{-1}	None	~ 0
Sodium (500 mM)	Site-specific and/or global	$+2.4$ kcal mol^{-1}	-0.2 kcal mol^{-1}	~ 0 ($\Phi^{app} < -0.1$)
Urea (0 to 5 M)	Global	$\leq +3.3$ kcal mol^{-1} $m = 0.65 \pm 0.05$ kcal mol^{-1} M^{-1}	$\leq +0.3$ kcal mol^{-1} $m^{\ddagger} = 0.07 \pm 0.04$ kcal mol^{-1} M^{-1}	– ~ 0 ($\beta^{\ddagger} = 0.1$)
Temperature (10 °C to 43 °C)	Global	$\Delta H = 8 \pm 2$ kcal mol^{-1} $\Delta S = 25 \pm 8$ cal mol^{-1} K^{-1}	$\Delta H^{\ddagger} = 12 \pm 2$ kcal mol^{-1} $\Delta S^{\ddagger}$[d] $= -14$ to 13 cal mol^{-1} K^{-1}	– –

[a] Class refers to whether the perturbation effects a specific interaction or set of interactions within the ribozyme (site-specific) or if the effects are less specific, interfering with all or many interactions (global).
[b] Positive energetic effects indicate a decrease in K_{dock} or k_{dock}, negative effects an increase.
[c] Effect on k_{dock} (undocked (U) to the transition state(‡)) divided by the effect on K_{dock} (U to docked (D).
[d] See the legend of Figure 4 for an explanation of the range of values.

effects of global perturbations on processes are not described by typically defined Φ-values, we use Φ^{app}, with the superscript app for apparent, to indicate that the ratio of $\Delta\Delta G^{\ddagger}/\Delta\Delta G$ may nonetheless be used to compare the properties of the transition state with those of the ground state.

Binding of guanosine, the other ribozyme substrate, is thermodynamically coupled to P1 docking *via* a metal ion that bridges the 2′-moiety of G and the reactive phosphoryl group of the oligonucleotide substrate strand of the P1 duplex.[70,77] The absence of an effect of G on the k_{dock} value suggests that the regions of the ribozyme active site responsible for coupling are not involved in the transition state for docking. Similarly, the destabilizing effect of Na^+ on the K_{dock} value is likely due to displacement of Mg^{2+} that support docking. The absence of an effect of Na^+ on the k_{dock} value suggests that these sites are not involved in the docking transition state ($\Phi^{app} \approx 0$) (Table 1(B)). Further, the absence of effects on k_{dock} of increased concentrations of Na^+ and Mg^{2+} provides evidence against large electrostatic differences between the undocked state and transition state.

Urea denaturation is a powerful probe for thermodynamic and kinetic studies of protein folding[78,79] and has seen increasing use in RNA folding.[7,11,80–84] Urea substantially destabilizes the docked state, with the slope of the urea *versus* ln *K* plot, or *m*-value, equal to 0.65 kcal mol^{-1}M^{-1} (Figure 4(a)). Shelton *et al.*[81] found a linear dependence of the *m*-values for the urea dependencies of duplex dissociation and tRNA unfolding on the change in solvent accessible surface area (ΔASA) between folded (duplexes) and unfolded (single-strands). From this empirical relationship, the *m*-value for docking corresponds to a change in solvent accessible surface area of ~4300 Å^2, equivalent to the ΔASA of forming ~5 base-pairs.[81] In contrast, there is little effect of urea on the rate of docking ($m^{\ddagger} \approx 0.07$ kcal mol^{-1} M^{-1}) (Figure 4(a)). The ratio between $m^{\ddagger}(k_{dock})$ and $m(K_{dock})$, referred to as $\beta^{\ddagger}$, has a value of ~0.1, and suggests that only ~10% of the ΔASA for the overall docking process has occurred in the transition state.[58,85]

Regarding the temperature dependence of docking, a difference in sign of ΔS and $\Delta S^{\ddagger}$ has been taken as evidence for a physical difference between the docked and transition states.[43] However, uncertainty in the exponential pre-factor in the Eyring equation that defines $\Delta S^{\ddagger}$ makes this parameter ambiguous and prevents interpretation of $\Delta S^{\ddagger}$ relative to ΔS†.

The multiple small Φ and Φ^{app}-values, and the small $\beta^{\ddagger}$ provide evidence that both locally and

† $\Delta S^{\ddagger}$ values are sensitive to the value of the pre-exponential factor, α, in the Eyring equation: $k = \alpha \cdot \exp(\Delta H^{\ddagger}/RT - \Delta S^{\ddagger}/R)$. For chemical reactions in a vacuum, α has been determined to be approximately equal to $k_B T/h = 6.2 \times 10^{12}$ s^{-1}, the frequency of vibration of the breaking bond (where k_B is Boltzman's constant and *h* is Planck's constant).[86] For a polymer in solution, this is certainly an overestimate of the maximum observable rate constant for an interaction.[17,87] Based on the measured rate of diffusion together of two residues of a short polypeptide, the α value for proteins has been roughly estimated at ~10^6 s^{-1}.[88] Interestingly, a dependence of the rate constant of DNA hairpin formation on loop size extrapolated to zero loop residues is also ~10^6 s^{-1}.[89] Nevertheless, it would be naive to expect a single value of α to be appropriate for all pairs of potentially interacting groups within a polymer or polymer type.

globally the properties of the transition state for docking are highly dissimilar from the docked state. This set of results strongly suggests that the transition state is not late in the docking process (Figure 5, $\ddagger_3$), and an "intermediate" transition state ($\ddagger_2$) is also unlikely. The remaining model, an early transition state for docking ($\ddagger_1$), is fully consistent with all of the data. This model is developed further and refined below.

An early transition state for docking: conformational search or kinetic trap?

The analysis described above strongly suggests that the transition state for docking is similar to the undocked state. Figure 6 presents alternative models to account for the early transition state for docking: a conformational search and a kinetic trap. Based on crystallographic, chemical protection, and small angle X-ray scattering data the undocked state is thought to be compact and highly structured, with most characteristics of the native, docked state.[25,34,45,54,90–92] If the undocked state of the ribozyme core were pre-organized for binding the P1 duplex, the simplest model for an early docking transition state would be a rate-limiting conformational search. However, comparisons with model systems and additional data described below suggest that docking is unlikely to be limited by a conformational search. Instead we propose that the existence of a kinetic trap, escape from which includes local reordering, slows docking. In principle, the barrier for escaping the kinetic trap could be large, so that escape from the kinetic trap is rate limiting. Alternatively, the kinetically trapped state could be thermodynamically favorable relative to other intermediates on the reaction coordinate, so that a step subsequent to breaking the interactions that stabilize the kinetic trap is rate limiting.

Macromolecular folding entails the establishment of native interactions, accompanied by a reduction in chain entropy and an increase in solvent entropy. Folding can be limited simply by the diffusion together of contact partners.[87,93,94] Such a rate-limiting conformational search may be slow because it is "frustrated" by intramolecular collisions between different regions of the polymer[95] or because infrequently sampled conformations are required for formation of native interactions.[96] Further along the continuum of possibilities, folding can be limited by escape from kinetic traps (e.g. see Pandya *et al.*[97]). This occurs when the requirement to break interactions formed in the course of folding slows the folding rate significantly. Indeed, many RNA folding processes have been suggested to be dominated by escape from kinetic traps.[1,11,14,29,98,99]

Experimentally, the hallmarks for kinetic traps in RNA folding have been considered to be the acceleration of folding upon addition of urea, increased temperature, reduced Mg^{2+}, or introduction of destabilizing mutations.[6,11,29,80] Recently, models

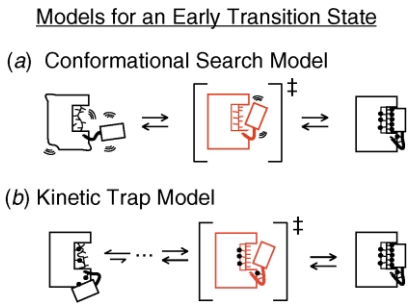

Figure 6. Alternative models to account for an early transition state for P1 duplex docking. Interactions are emphasized by dots. (a) A conformational search involving the P1 duplex and limited regions of the ribozyme core slows docking. (b) A kinetic trap slows docking of the P1 duplex. As depicted schematically, we conjecture that this kinetic trap may consist of interactions in J1/2 and/or interaction between the P1 duplex and the ribozyme, as well as local misfolding of regions in the core, such as J8/7.

with rate-limiting conformational searches have been proposed to account for RNA folding events that are not accelerated by urea or increased temperature.[2,83]

The absence of urea dependence for P1 docking superficially supports a model in which a conformational search limits docking. Alternative interpretations for the urea and temperature data are possible, however. Indeed, in at least two protein systems, intermediates with non-native interactions that slow overall folding have been described, but folding is still decelerated by denaturant.[100–102] More generally, if there is no net change in hydrogen bonding or solvent accessible surface area (in the case of urea[81]) or net bonding enthalpy (in the case of temperature) between the kinetically trapped state and the rate-limiting transition state for escape from the trapped state, then the observed rate constant would not be sensitive to these experimental parameters. For example, escape from the trapped state might involve a sliding motion with no net change in surface exposure or a large motion but with an early transition state, so that all exposure occurs after the rate-limiting barrier. We therefore investigated further whether a conformational search is rate limiting for docking.

Docking occurs on the hundreds of milliseconds time scale, with a rate constant of $\sim 3\ s^{-1}$. If the three adenosine residues between P1 and P2, termed J1/2 (Figure 1(a)), were acting as a simple tether, docking would be expected to occur in tens of microseconds or faster, similar to the folding rate of a DNA hairpin with a four residue loop ($k_{zip} \approx 10^5\ s^{-1}$).[103] Further, overall folding of the $\sim$400 residue Tetrahymena ribozyme can occur at $\sim 1\ s^{-1}$,[104] even though many more degrees of freedom are expected to be lost in overall folding than in docking from an already structured undocked state. Folding of other large RNAs occurs at similar or faster rates.[7,17] These data suggest that presumably extensive conformational searches can occur

much more rapidly than docking and argue against docking being limited by a conformational search between the P1 duplex and the preordered ribozyme core†. In addition, while temperature dependencies, like urea dependencies, are difficult to interpret, as expressed above, the observed positive $\Delta H^{\ddagger}$ value of 12(±2) kcal mol^{-1} is greater than that expected for a diffusion limited process[67] and consistent with an increase of temperature accelerating the breaking of interactions or rearrangement of solvent in escape from a kinetic trap, as suggested previously for docking.[105]

To test the conformational search model more directly we measured the kinetics of docking of a P1 duplex formed with a short oligonucleotide substrate that frees two guanine residues of the IGS from base-pairing interactions, thereby lengthening the tether between P1 and the ribozyme (Figure 1(c)). Analogous experiments have been carried out to assess the contribution of conformational entropy to protein folding.[62,106,107] If a conformational search were rate-limiting, lengthening the tether between the P1 duplex and the ribozyme in this manner is most simply expected to increase the degrees of freedom involved in the docking process and thus decrease the rate of docking.[108] However, k_{dock} does not decrease due to this modification; instead, the rate of docking is increased by ~tenfold (Table 1(A) and Figure 3(d)). This result provides experimental evidence against a rate-limiting conformational search for docking and supports instead the model of a kinetic trap slowing docking. Analogously, increasing the loop length of α_1-antitrypsin, a serpin protein that folds to a kinetically trapped state, was shown to prevent accumulation of the trapped state.[109]

The basis for faster docking of the shortened P1 is under exploration. The simplest hypothesis is that the shortened substrate relieves constraining interactions between P1 and other regions of the ribozyme resulting in accelerated docking. Consistent with conformational restriction of P1 in the undocked state, cross-linking experiments with a highly reactive azidophenacyl group at the end of the IGS show a predominant cross-link with a residue in P2 in the undocked state.[74,110] Additionally, a residue within J1/2 (Figure 1(a)) is modified strongly by dimethyl sulfate in the docked state but not in the undocked state,[34,111] suggesting a significant change in the conformation of J1/2 upon docking. Lengthening and shortening J1/2 have similar effects on the cleavage position within the P1 duplex providing further evidence that this region does not act as a simple tether.[36,112] The shortened P1 duplex may stack less well on J1/2 than the full length duplex, freeing P1 to achieve the docking transition state more frequently. Similarly, the observation of small (~twofold) increases (with modifications to the IGS) and small decreases (with modifications to S) in k_{dock} may be due to destabilization or stabilization, respectively, of non-native interactions between P1 and the ribozyme in the undocked state.

Importantly, the rate constant of ~30 s^{-1} for docking with the shortened P1 duplex remains two to three orders of magnitude slower than expected for formation of a simple intramolecular RNA structure.[89,103] Thus, there may be significant additional barriers to docking beyond the possible mispositioning of the P1 duplex and J1/2 in the undocked state. Other regions that might require rearrangement in the process of docking are those in the core with which P1 has been proposed to interact, namely the J4/5 loop and J8/7.[45,50,51] J8/7 is a particularly good candidate for a structural rearrangement as it is the longest unduplexed region in the ribozyme.[50] In principle J8/7 could be unstructured and require a conformational search to attain the docked state, perhaps to a state that includes infrequently sampled conformers. However, the observation of protection against hydroxyl radicals of the J8/7 backbone[25] suggests instead that J8/7 is structured. Thus, we conjecture that J8/7 has an incorrect structure(s) in the undocked state that must be disrupted to support docking. This model is consistent with the high propensity of each RNA residue to interact with each other one through stacking, hydrogen bonding between bases,[113] and hydrogen bonding with the 2′-hydroxyl groups of the ribose rings. Indeed, hairpin formation was proposed to be slowed by myriad local misfolded structures.[103] Further, in the context of the preformed ribozyme core, the propensity for formation of moderately stable structures may redouble as a result of prealignment of multiple residues.

In summary, a kinetic trap in the undocked state involving J1/2 and possibly J8/7 provides the simplest model to account for the dissimilarity between the transition state and the docked state, the slow rate of docking, the positive $\Delta H^{\ddagger}$, and the absence of rate decrease upon unpairing residues

† For a perspective on the magnitude of the rate constant for docking we can utilize a simple probabilistic model, in which the rate constant for an interaction varies approximately by $k_{fold} = k_o \cdot C^N$, where k_o is the maximal rate constant for interacting, C is the fraction of conformations of a group that can interact, and N is the number of independent ligands involved. If multiple independent regions must be correctly aligned to support docking (i.e. large N) and/or the probability of correct alignment of any or all of ligands is small (i.e. small C) then k_{fold} becomes very slow. Nonetheless, using a rough estimate of $k_o = 1\ \mu s^{-1}$ based on hairpin[89] and polypeptide data,[88] N and C must take on surprising values in order to approach the slow rate of docking. If $C = 1/3$, the estimated value of N that would give k_{dock} of 3 s^{-1} is $N \approx 10$. Ten independent regions that must come together for docking to occur is a large number given the evidence that the ribozyme is largely prefolded in the absence of the P1 duplex (Figure 1(a)).[34,50,53–55] Alternatively, if $N = 3$, for the regions involved in docking, namely P1, J1/2, and J8/7, then C becomes ~0.014, such that only 1/70th of conformations for each region would support docking.

within the P1 duplex. Nevertheless, conformational search models in which a large number of rare conformations must be attained to support docking and/or in which intra-chain collisions slow the search cannot be wholly excluded. Both classes of models (Figure 6) imply that the catalytic site of the ribozyme is not fully pre-organized for P1 binding. Indeed, recent experiments provide a strong precedent for rearrangements within the ribozyme active site: guanosine binds to the ribozyme several orders of magnitude slower than diffusional encounter, and rearrangements of functional group interactions have been identified within ribozyme complexes with substrates and products.[70,77,114] Another example of an alternative local structure in the context of a cooperatively folded RNA is the alternative pairing of P3 in a misfolded Tetrahymena ribozyme structure that is otherwise similar to the native fold by chemical protection.[12,13,29]

The emerging picture of misorder in the Tetrahymena ribozyme core suggests that even large, cooperatively folded RNAs can have local regions that lack native structure. These regions become appropriately ordered in the course of folding or reaction, much like the structural accommodation between binding partners described for numerous small RNAs.[115,116] Modifications to selected tertiary interactions involved in tRNA and P4–P6 folding were also observed to have no effect on the folding rates of these RNAs.[17,20] The apparent observation of early transition states in these processes as well as in docking may arise because rearrangement from alternative structures is a common feature of these and other RNA folding and binding events. The challenge remains to discern to what degree RNA's propensity for alternative conformations has interfered with biological function and, conversely, to what degree this propensity has enhanced function. Has the propensity for multiple states facilitated the use of RNA in complex reactions that utilize successive structural states, and has nature taken advantage of these multiple states to regulate processes?

Materials and Methods

Reagent preparation

Oligonucleotide substrates were purchased from Dharmacon, Inc. (Lafayette, CO) and labeled with Cy3 as described.[68] For single molecule measurements, the well-characterized L-21 *Sca*I ribozyme or a ribozyme with an extended IGS (L-16, having an additional 5′-GGUUU sequence)[27] was extended at the 3′ end by 27 residues (T; Figure 1(a)) for annealing to a biotinylated DNA tether.[68] Ribozymes with IGS modifications were created by ligation of a synthetic oligonucleotide to a transcribed L-38, T ribozyme with minor modifications to previous protocols[47,117] using low Mg^{2+} and a 44mer DNA oligonucleotide complementary to the 3′ end of the ribozyme to prevent 3′ end processing. After gel purification, ligated molecules were >90% of the population as judged by kinase exchange labeling.[118]

Single-molecule measurements

For single molecule fluorescence resonance energy transfer (FRET) assays, ribozymes were annealed with donor-labeled (Cy3) oligonucleotide substrate and acceptor-labeled (Cy5), biotinylated tether and bound to a streptavidin-coated glass surface for visualization with a confocal microscope (Figure 2). Minor modifications to previous protocols[68] are described in Supplementary Material. Unless stated otherwise, experiments were carried out at 22(±1) °C, 10 mM $MgCl_2$, 50 mM sodium Mops (3-[*N*-morpholino]propanesulfonic acid) (pH 7.0), with an oxygen scavenging system consisting of 10% (w/v) glucose, 1% (v/v) β-mercaptoethanol, ~50 μg/ml glucose oxidase (Roche Molecular Biochemicals), 10 μg/ml catalase (Roche Molecular Biochemicals). Control experiments with wild-type ribozyme showed that lower and higher glucose (3% and 13%) did not affect the rate and equilibrium constants for docking within error (data not shown). For experiments with urea, oxygen-scavenging enzymes were incubated with the buffer for five to ten minutes before addition of urea.

Docked and undocked states, indicated by high and low FRET events, respectively, were identified based on quantitative criteria for FRET and relative fluorescence signals as described in Supplementary Material. The reported equilibrium constants for docking (K_{dock}) are the ratio of the total time spent in the docked and undocked states. The rate constants for docking (k_{dock}) and undocking (k_{undock}) were determined by fitting plots of the cumulative number of undocked and docked events observed with each dwell time, respectively (Figure 3). The observed dwell times for events attributed to the undocked state for wild-type ribozyme and other complexes with K_{dock} greater than ~5 are not well described by a single exponential (Figure 3(a) and (b)). A similar number of short-duration, low-FRET events that fulfill the criteria for undocking were identified for a control DNA duplex with a Cy3–Cy5 pair at one end (Figure 3(a)). This and additional observations suggest that the fast phase in the undocked time plot is due to photophysical events, termed "blinking" events, and is thus unrelated to the ribozyme (see Supplementary Material). Other corrections to the reported rate constants were made for systematic biases due to time trace truncation and undercounting of short docked events (Supplementary Material). These modifications improve the accuracy of the reported rate constants by ≤twofold and the agreement between equilibrium constants calculated from total times and the ratio of rate constants, but are not necessary for the conclusions made herein.

Acknowledgements

We are especially grateful to H. D. Kim and H. P. Babcock for sharing their experimental expertise. We thank S. Strobel for the generous gift of a vector for over-expression of T4 DNA ligase, S. Strobel and A. Szewczak for sharing the coordinates of their model for the docked complex, and K. Travers, M. Peck, and the reviewers for comments on the manuscript. This work was supported by a Packard Interdisciplinary grant (to S.C. and D.H.), a National Institutes of Health

(NIH) grant (to D.H.), a National Science Foundation (NSF) and an Air Force Office of Scientific Research grant (to S.C.). X.Z. and L.E.B. received support from an NIH postdoctoral fellowship and an NIH biotechnology training grant, respectively. R.D. was supported by an Abbott Laboratories Stanford Graduate Fellowship and an NSF fellowship.

References

1. Treiber, D. K. & Williamson, J. R. (1999). Exposing the kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.* **9**, 339–345.
2. Treiber, D. K. & Williamson, J. R. (2001). Beyond kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.* **309-314**, 309–345. 21299342..
3. Tinoco, I., Jr & Bustamante, C. (1999). How RNA folds. *J. Mol. Biol.* **293**, 271–281.
4. Woodson, S. A. (2000). Recent insights on RNA folding mechanisms from catalytic RNA. *Cell. Mol. Life Sci.* **57**, 796–808.
5. Zarrinkar, P. P., Wang, J. & Williamson, J. R. (1996). Slow folding kinetics of RNase P RNA. *RNA*, **2**, 564–573.
6. Pan, T. & Sosnick, T. R. (1997). Intermediates and kinetic traps in the folding of a large ribozyme revealed by circular dichroism and UV absorbance spectroscopies and catalytic activity. *Nature Struct. Biol.* **4**, 931–938.
7. Fang, X. W., Pan, T. & Sosnick, T. R. (1999). Mg^{2+}-dependent folding of a large ribozyme without kinetic traps. *Nature Struct. Biol.* **6**, 1091–1095.
8. Emerick, V. L. & Woodson, S. A. (1994). Fingerprinting the folding of a group I precursor RNA. *Proc. Natl Acad. Sci. USA*, **91**, 9675–9679.
9. Zarrinkar, P. P. & Williamson, J. R. (1994). Kinetic intermediates in RNA folding. *Science*, **265**, 918–924.
10. Pan, J., Thirumalai, D. & Woodson, S. A. (1997). Folding of RNA involves parallel pathways. *J. Mol. Biol.* **273**, 7–13.
11. Treiber, D. K., Rook, M. S., Zarrinkar, P. P. & Williamson, J. R. (1998). Kinetic intermediates trapped by native interactions in RNA folding. *Science*, **279**, 1943–1946.
12. Sclavi, B., Sullivan, M., Chance, M. R., Brenowitz, M. & Woodson, S. A. (1998). RNA folding at millisecond intervals by synchrotron hydroxyl radical footprinting. *Science*, **279**, 1940–1943.
13. Russell, R. & Herschlag, D. (1999). New pathways in folding of the *Tetrahymena* group I RNA enzyme. *J. Mol. Biol.* **291**, 1155–1167.
14. Russell, R. & Herschlag, D. (2001). Probing the folding landscape of the *Tetrahymena* ribozyme: commitment to form the native conformation is late in the folding pathway. *J. Mol. Biol.* **308**, 839–851.
15. Silverman, S. K. & Cech, T. R. (1999). Energetics and cooperativity of tertiary hydrogen bonds in RNA structure. *Biochemistry*, **38**, 8691–8702.
16. Silverman, S. K., Deras, M. L., Woodson, S. A., Scaringe, S. A. & Cech, T. R. (2000). Multiple folding pathways for the P4–P6 RNA domain. *Biochemistry*, **39**, 12465–12475.
17. Silverman, S. K. & Cech, T. R. (2001). An early transition state for folding of the P4–P6 RNA domain. *RNA*, **7**, 161–166.
18. Cole, P. E. & Crothers, D. M. (1972). Conformational changes of transfer ribonucleic acid. Relaxation kinetics of the early melting transition of methionine transfer ribonucleic acid (*Escherichia coli*). *Biochemistry*, **11**, 4368–4374.
19. Crothers, D. M., Cole, P. E., Hilbers, C. W. & Shulman, R. G. (1974). The molecular mechanism of thermal unfolding of *Escherichia coli* formylmethionine transfer RNA. *J. Mol. Biol.* **87**, 63–88.
20. Maglott, E. J., Goodwin, J. T. & Glick, G. D. (1999). Probing the structure of an RNA tertiary unfolding transition state. *J. Am. Chem. Soc.* **121**, 7461–7462.
21. Fang, X., Littrell, K., Yang, X. J., Henderson, S. J., Siefert, S., Thiyagarajan, P. *et al.* (2000). Mg^{2+}-dependent compaction and folding of yeast $tRNA^{Phe}$ and the catalytic domain of the *B. subtilis* RNase P RNA determined by small-angle X-ray scattering. *Biochemistry*, **39**, 11107–11113.
22. Serebrov, V., Clarke, R. J., Gross, H. J. & Kisselev, L. (2001). Mg^{2+}-induced tRNA folding. *Biochemistry*, **40**, 6688–6698.
23. Walter, N. G., Hampel, K. J., Brown, K. M. & Burke, J. M. (1998). Tertiary structure formation in the hairpin ribozyme monitored by fluorescence resonance energy transfer. *EMBO J.* **17**, 2378–2391.
24. Zhuang, X., Kim, H., Pereira, M. J., Babcock, H. P., Walter, N. G. & Chu, S. (2002). Correlating structural dynamics and function in single ribozyme molecules. *Science*, **296**, 1473–1476.
25. Latham, J. A. & Cech, T. R. (1989). Defining the inside and outside of a catalytic RNA molecule. *Science*, **245**, 276–282.
26. Celander, D. W. & Cech, T. R. (1991). Visualizing the higher order folding of a catalytic RNA molecule. *Science*, **251**, 401–407.
27. Russell, R., Zhuang, X. W., Babcock, H. P., Millett, I. S., Doniach, S., Chu, S. & Herschlag, D. (2002). Exploring the folding landscape of a structured RNA. *Proc. Natl Acad. Sci. USA*, **99**, 155–160.
28. Russell, R., Millett, I. S., Tate, M. W., Kwok, L. W., Nakatani, B., Gruner, S. M. *et al.* (2002). Rapid compaction during RNA folding. *Proc. Natl Acad. Sci. USA*, **99**, 4266–4271.
29. Pan, J. & Woodson, S. A. (1998). Folding intermediates of a self-splicing RNA: mispairing of the catalytic core. *J. Mol. Biol.* **280**, 597–609.
30. Silverman, S. K., Zheng, M., Wu, M., Tinoco, I., Jr & Cech, T. R. (1999). Quantifying the energetic interplay of RNA tertiary and secondary structure interactions. *RNA*, **5**, 1665–1674.
31. Green, R. & Szostak, J. W. (1994). *In vitro* genetic analysis of the hinge region between helical elements P5–P4–P6 and P7–P3–P8 in the sunY group I self-splicing intron. *J. Mol. Biol.* **235**, 140–155.
32. Knitt, D. S., Narlikar, G. J. & Herschlag, D. (1994). Dissection of the role of the conserved G·U pair in group I RNA self-splicing. *Biochemistry*, **33**, 13864–13879.
33. Pyle, A. M., Murphy, F. L. & Cech, T. R. (1992). RNA substrate binding site in the catalytic core of the *Tetrahymena* ribozyme. *Nature*, **358**, 123–128.
34. Engelhardt, M. A., Doherty, E. A., Knitt, D. S., Doudna, J. A. & Herschlag, D. (2000). The P5abc peripheral element facilitates preorganization of the *Tetrahymena* group I ribozyme for catalysis. *Biochemistry*, **39**, 2639–2651.
35. Bevilacqua, P. C., Kierzek, R., Johnson, K. A. & Turner, D. H. (1992). Dynamics of ribozyme binding

of substrate revealed by fluorescence-detected stopped-flow methods. *Science*, **258**, 1355–1358.
36. Herschlag, D. (1992). Evidence for processivity and two-step binding of the RNA substrate from studies of J1/2 mutants of the *Tetrahymena* ribozyme. *Biochemistry*, **31**, 1386–1399.
37. Staley, J. P. & Guthrie, C. (1998). Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell*, **92**, 315–326.
38. Frank, J. & Agrawal, R. K. (2000). A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature*, **406**, 318–322.
39. Wilson, K. S. & Noller, H. F. (1998). Molecular movement inside the translational engine. *Cell*, **92**, 337–349.
40. Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N. & Noller, H. F. (1999). X-ray crystal structures of 70 S ribosome functional complexes. *Science*, **285**, 2095–2104.
41. Narlikar, G. J., Khosla, M., Usman, N. & Herschlag, D. (1997). Quantitating tertiary binding energies of 2′ OH groups on the P1 duplex of the *Tetrahymena* ribozyme: intrinsic binding energy in an RNA enzyme. *Biochemistry*, **36**, 2465–2477.
42. Bevilacqua, P. C. & Turner, D. H. (1991). Comparison of binding of mixed ribose–deoxyribose analogues of CUCU to a ribozyme and to GGAGAA by equilibrium dialysis: evidence for ribozyme specific interactions with 2′ OH groups. *Biochemistry*, **30**, 10632–10640.
43. Narlikar, G. J., Bartley, L. E., Khosla, M. & Herschlag, D. (1999). Characterization of a local folding event of the *Tetrahymena* group I ribozyme: effects of oligonucleotide substrate length, pH, and temperature on the two substrate binding steps. *Biochemistry*, **38**, 14192–14204.
44. Williams, A. P., Longfellow, C. E., Freier, S. M., Kierzek, R. & Turner, D. H. (1989). Laser temperature-jump, spectrscopic, and thermodynamic study of salt effects on duplex formation by dGCATGC. *Biochemistry*, **28**, 4283–4291.
45. Pyle, A. M. & Cech, T. R. (1991). Ribozyme recognition of RNA by tertiary interactions with specific ribose 2′-OH groups. *Nature*, **350**, 628–631.
46. Herschlag, D., Eckstein, F. & Cech, T. R. (1993). Contributions of 2′-hydroxyl groups of the RNA substrate to binding and catalysis by the *Tetrahymena* ribozyme. An energetic picture of an active site composed of RNA. *Biochemistry*, **32**, 8299–8311.
47. Strobel, S. A. & Cech, T. R. (1993). Tertiary interactions with the internal guide sequence mediate docking of the P1 helix into the catalytic core of the *Tetrahymena* ribozyme. *Biochemistry*, **32**, 13593–13604.
48. Strobel, S. A. & Cech, T. R. (1995). Minor groove recognition of the conserved G·U pair at the *Tetrahymena* ribozyme reaction site. *Science*, **267**, 675–679.
49. Strobel, S. A. & Cech, T. R. (1996). Exocyclic amine of the conserved G·U pair at the cleavage sit of the *Tetrahymena* ribozyme contributes to 5′-splice site selection and transition state stabilization. *Biochemistry*, **35**, 1201–1211.
50. Michel, F. & Westhof, E. (1990). Modeling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.* **216**, 585–610.
51. Szewczak, A. A., Ortoleva-Donnelly, L., Ryder, S. P., Moncoeur, E. & Strobel, S. A. (1998). A minor groove RNA triple helix within the catalytic core of a group I intron. *Nature Struct. Biol.* **5**, 1037–1042.
52. Lehnert, V., Jaeger, L., Michel, F. & Westhof, E. (1996). New loop–loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the *Tetrahymena thermophila* ribozyme. *Chem. Biol.* **3**, 993–1009.
53. Tanner, M. A. & Cech, T. R. (1997). Joining the two domains of a group I ribozyme to form the catalytic core. *Science*, **275**, 847–849.
54. Golden, B. L., Gooding, A. R., Podell, E. R. & Cech, T. R. (1998). A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme. *Science*, **282**, 259–264.
55. Chen, X., Gutell, R. R. & Lambowitz, A. M. (2000). Function of tyrosyl-tRNA synthetase in splicing group I introns: an induced-fit model for binding to the P4–P6 domain based on analysis of mutations at the junction of the P4–P6 stacked helices. *J. Mol. Biol.* **301**, 265–283.
56. Fersht, A. R., Matouschek, A. & Serrano, L. (1992). The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. *J. Mol. Biol.* **224**, 771–782.
57. Plaxco, K. W., Simons, K. T., Ruczinski, I. & Baker, D. (2000). Topology, stability, sequence, and length: defining the determinants of two-state protein folding kinetics. *Biochemistry*, **37**, 11177–11183.
58. Oliveberg, M. (2001). Characterisation of the transition states for protein folding: towards a new level of mechanistic detail in protein engineering analysis. *Curr. Opin. Sruct. Biol.* **11**, 94–100.
59. Matouschek, A., Kellis, J. T., Serrano, L. & Fersht, A. R. (1989). Mapping the transition-state and pathway of protein folding by protein engineering. *Nature*, **340**, 122–126.
60. Serrano, L., Matouschek, A. & Fersht, A. R. (1992). The folding of an enzyme. III. Structure of the transition state for unfolding of barnase analysed by a protein engineering procedure. *J. Mol. Biol.* **224**, 805–818.
61. Milla, M. E., Brown, B. M., Waldburger, C. D. & Sauer, R. T. (1995). P22 Arc repressor: transition-state properties inferred from mutational effects on the rates of protein unfolding and refolding. *Biochemistry*, **34**, 13914–13919.
62. Burton, R. E., Huang, G. S., Daugherty, M. A., Calderone, T. L. & Oas, T. G. (1997). The energy landscape of a fast-folding protein mapped by Ala → Gly substitutions. *Nature Struct. Biol.* **4**, 305–310.
63. Myers, J. K. & Oas, T. G. (1999). Contribution of a buried hydrogen bond to lambda repressor folding kinetics. *Biochemistry*, **38**, 6761–6768.
64. Goldberg, J. M. & Baldwin, R. L. (1999). A specific transition state for S-peptide combining with folded S-protein and then refolding. *Proc. Natl Acad. Sci. USA*, **96**, 2019–2024.
65. Gloss, L. M., Simler, B. R. & Matthews, C. R. (2001). Rough energy landscapes in protein folding: dimeric *E. coli Trp* repressor folds through three parallel channels. *J. Mol. Biol.* **312**, 1121–1134.
66. Taylor, M. G., Rajpal, A. & Kirsch, J. F. (1998). Kinetic epitope mapping of the chicken lysozyme-HyHEL-10 Fab complex: delineation of docking trajectories. *Protein Sci.* **7**, 1857–1867.
67. Frisch, C., Fersht, A. R. & Schreiber, G. (2001). Experimental assignment of the structure of the

transition state for the association of barnase and barstar. *J. Mol. Biol.* **308**, 69–77.
68. Zhuang, X., Bartley, L. E., Babcock, H. P., Russell, R., Ha, T., Herschlag, D. & Chu, S. (2000). A single molecule study of RNA catalysis and folding. *Science*, **288**, 2048–2051.
69. Weiss, S. (2000). Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. *Nature Struct. Biol.* **7**, 724–729.
70. McConnell, T. S., Cech, T. R. & Herschlag, D. (1993). Guanosine binding to the *Tetrahymena* ribozyme: thermodynamic coupling with oligonucleotide binding. *Proc. Natl Acad. Sci. USA*, **90**, 8362–8366.
71. Herschlag, D., Eckstein, F. & Cech, T. R. (1993). The importance of being ribose at the cleavage site in the *Tetrahymena* ribozyme reaction. *Biochemistry*, **32**, 8312–8321.
72. McConnell, T. S. & Cech, T. R. (1995). A positive entropy change for guanosine binding and for the chemical step in the *Tetrahymena* ribozyme reaction. *Biochemistry*, **34**, 4056–4067.
73. Li, Y. & Turner, D. H. (1997). Effects of Mg^{2+} and the 2′-OH of guanosine on steps required for substrate binding and reactivity with the *Tetrahymena* ribozyme reveal several local folding transitions. *Biochemistry*, **36**, 11131–11139.
74. Narlikar, G. J. & Herschlag, D. (1996). Isolation of a local tertiary folding transition in the context of a globally folded RNA. *Nature Struct. Biol.* **3**, 701–710.
75. Narlikar, G. J., Bartley, L. E. & Herschlag, D. (2000). Use of duplex rigidity for stability and specificity in RNA tertiary structure. *Biochemistry*, **39**, 6183–6189.
76. Fersht, A. R., Itzhaki, L. S., Elmasry, N., Matthews, J. M. & Otzen, D. E. (1994). Single versus parallel pathways of protein-folding and fractional formation of structure in the transition-state. *Proc. Natl Acad. Sci. USA*, **91**, 10426–10429.
77. Shan, S. O. & Herschlag, D. (2002). Dissection of a metal-ion-mediated conformational change in *Tetrahymena* ribozyme catalysis. *RNA*, **8**, 861–872.
78. Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280.
79. Ternstrom, T., Mayor, U., Akke, M. & Oliveberg, M. (1999). From snapshot to movie: phi analysis of protein folding transition states taken one step further. *Proc. Natl Acad. Sci. USA*, **96**, 14854–14859.
80. Rook, M. S., Treiber, D. K. & Williamson, J. R. (1998). Fast folding mutants of the *Tetrahymena* group I ribozyme reveal a rugged folding energy landscape. *J. Mol. Biol.* **281**, 609–620.
81. Shelton, V. M., Sosnick, T. R. & Pan, T. (1999). Applicability of urea in the thermodynamic analysis of secondary and tertiary RNA folding. *Biochemistry*, **38**, 16831–16839.
82. Fang, X., Pan, T. & Sosnick, T. R. (1999). A thermodynamic framework and cooperativity in the tertiary folding of a Mg^{2+}-dependent ribozyme. *Biochemistry*, **38**, 16840–16846.
83. Buchmueller, K. L., Webb, A. E., Richardson, D. A. & Weeks, K. M. (2000). A collapsed non-native RNA folding state. *Nature Struct. Biol.* **7**, 362–366.
84. Ralston, C. Y., He, Q., Brenowitz, M. & Chance, M. R. (2000). Stability and cooperativity of individual tertiary contacts in RNA revealed through chemical denaturation. *Nature Struct. Biol.* **7**, 371–374.
85. Tanford, C. (1968). Protein Denaturation. Parts a and b. *Advan. Protein Chem.* **23**, 121–282.
86. Sime, R. J. (1990). *Physical Chemistry: Methods, Techniques, and Experiments*, Saunders College Publisher, Philadelphia, PA.
87. Dill, K. A. & Chan, H. S. (1997). From Levinthal to pathways to funnels. *Nature Struct. Biol.* **4**, 10–19.
88. Hagen, S. J., Hofrichter, J., Szabo, A. & Eaton, W. A. (1996). Diffusion-limited contact formation in unfolded cytochrome *c*: estimating the maximum rate of protein folding. *Proc. Natl Acad. Sci. USA*, **93**, 11615–11617.
89. Bonnet, G., Krichevsky, O. & Libchaber, A. (1998). Kinetics of conformational fluctuations in DNA hairpin-loops. *Proc. Natl Acad. Sci. USA*, **95**, 8602–8606.
90. Russell, R., Millett, I. S., Doniach, S. & Herschlag, D. (2000). Small angle X-ray scattering reveals a compact intermediate in RNA folding. *Nature Struct. Biol.* **7**, 367–370.
91. Inoue, T. & Cech, T. R. (1985). Secondary structure of the circular form of the *Tetrahymena* ribosomal-RNA intervening sequence: a technique for RNA structure-analysis using chemical probes and reverse-transcriptase. *Proc. Natl Acad. Sci. USA*, **82**, 648–652.
92. Banerjee, A. R., Jaeger, J. A. & Turner, D. H. (1993). Thermal unfolding of a group I ribozyme: the low-temperature transition is primarily disruption of tertiary structure. *Biochemistry*, **32**, 153–163.
93. Plaxco, K. W. & Baker, D. (1998). Limited internal friction in the rate-limiting step of a two-state protein folding reaction. *Proc. Natl Acad. Sci. USA*, **95**, 13591–13596.
94. Makarov, D. E. & Plaxco, K. W. (2003). The topomer search model: a simple, quantitative theory of two-state protein folding kinetics. *Protein Sci.* **12**, 17–26.
95. Ansari, A., Jones, C. M., Henry, E. R., Hofrichter, J. & Eaton, W. A. (1992). The role of solvent viscosity in the dynamics of protein conformational changes. *Science*, **256**, 1796–1798.
96. Murthy, V. L., Srinivasan, R., Draper, D. E. & Rose, G. D. (1999). A complete conformational map for RNA. *J. Mol. Biol.* **291**, 313–327.
97. Pandya, M. J., Williams, P. B., Dempsey, C. E., Shewry, P. R. & Clarke, A. R. (1999). Direct kinetic evidence for folding via a highly compact, misfolded state. *J. Biol. Chem.* **274**, 26828–26837.
98. Uhlenbeck, O. C. (1995). Keeping RNA happy. *RNA*, 4–6.
99. Herschlag, D. (1995). RNA chaperones and the RNA folding problem. *J. Biol. Chem.* **270**, 20871–20874.
100. Kiefhaber, T. (1995). Kinetic traps in lysozyme folding. *Proc. Natl Acad. Sci. USA*, **92**, 9029–9033.
101. Lu, H., Buck, M., Radford, S. E. & Dobson, C. M. (1997). Acceleration of the folding of hen lysozyme by trifluoroethanol. *J. Mol. Biol.* **265**, 112–117.
102. Capaldi, A. P., Kleanthous, C. & Radford, S. E. (2002). Im7 folding mechanism: misfolding on a path to the native state. *Nature Struct. Biol.* **9**, 209–216.
103. Ansari, A., Kuznetsov, S. V. & Shen, Y. Q. (2001). Configurational diffusion down a folding funnel describes the dynamics of DNA hairpins. *Proc. Natl Acad. Sci. USA*, **98**, 7771–7776.
104. Zhuang, X. W., Ha, T., Kim, H. D., Centner, T., Labeit, S. & Chu, S. (2000). Fluorescence quenching:

a tool for single-molecule protein-folding study. *Proc. Natl Acad. Sci. USA*, **97**, 14241–14244.
105. Li, Y., Bevilacqua, P. C., Mathews, D. & Turner, D. H. (1995). Thermodynamic and activation parameters for binding of a pyrene-labeled substrate by the *Tetrahymena* ribozyme: docking is not diffusion-controlled and is driven by a favorable entropy change. *Biochemistry*, **34**, 14394–14399.
106. Ladurner, A. G. & Fersht, A. R. (1997). Glutamine, alanine or glycine repeats inserted into the loop of a protein have minimal effects on stability and folding rates. *J. Mol. Biol.* **273**, 330–337.
107. Viguera, A. R. & Serrano, L. (1997). Loop length, intramolecular diffusion and protein folding. *Nature Struct. Biol.* **4**, 939–946.
108. Cohen, S. B. & Cech, T. R. (1998). A quantitative study of the flexibility contributed to RNA structures by nicks and single-stranded gaps. *RNA*, **4**, 1179–1185.
109. Im, H., Ahn, I. Y. & Yu, M. H. (2000). Bypassing the kinetic trap of serpin protein folding by loop extension. *Protein Sci.* **9**, 1497–1502.
110. Wang, J. F., Downs, W. D. & Cech, T. R. (1993). Movement of the guide sequence during RNA catalysis by a group I ribozyme. *Science*, **260**, 504–508.
111. Hanna, R. L., Gryaznov, S. M. & Doudna, J. A. (2000). A phosphoramidate substrate analog is a competitive inhibitor of the *Tetrahymena* group I ribozyme. *Chem. Biol.* **7**, 845–854.
112. Young, B., Herschlag, D. & Cech, T. R. (1991). Mutations in a non-conserved sequence of the *Tetrahymena* ribozyme increase activity and specificity. *Cell*, **67**, 1007–1019.
113. Burkard, M. E., Turner, D. H. & Tinoco, I. Jr (1999). Appendix 1: strucutres of base pairs involving at least two hydrogen bonds. In *The RNA World* (Gesteland, R. F., Cech, T. R. & Atkins, J. F., eds), 2nd edit., pp. 675–680, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
114. Karbstein, K. & Herschlag, D. (2003). Extraordinarily slow binding of guanosine to the *Tetrahymena* group I ribozyme: implications for RNA preorganization and function. *Proc. Natl. Acad. Sci. USA*, **100**, 2300–2305.
115. Frankel, A. D. & Smith, C. A. (1998). Induced folding in RNA–protein recognition: more than a simple molecular handshake. *Cell*, **92**, 149–151.
116. Patel, D. J., Suri, A. K., Jiang, F., Jiang, L., Fan, P., Kumar, R. A. & Nonin, S. (1997). Structure, recognition and adaptive binding in RNA aptamer complexes. *J. Mol. Biol.* **272**, 645–664.
117. Moore, M. J. & Query, C. C. (2000). Joining of RNAs by splinted ligation. *Methods Enzymol.* **317**, 109–123.
118. Narlikar, G. J. & Herschlag, D. (1998). Direct demonstration of the catalytic role of binding interactions in an enzymatic reaction. *Biochemistry*, **37**, 9902–9911.
119. Szewczak, A. A., Ortoleva-Donnelly, L., Zivarts, M. V., Oyelere, A. K., Kazantsev, A. V. & Strobel, S. A. (1999). An important base triple anchors the substrate helix recognition surface within the *Tetrahymena* ribozyme active site. *Proc. Natl Acad. Sci. USA*, **96**, 11183–11188.

Edited by D. E. Draper

(Received 21 November 2002; received in revised form 18 February 2003; accepted 24 February 2003)

Supplementary Material for this paper comprising three Tables, two Schemes and one Figure, is available on Science Direct