Article No. jmbi.1999.3026 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (1999) 291, 1155–1167

JMB

New Pathways in Folding of the *Tetrahymena* Group I RNA Enzyme

Rick Russell and Daniel Herschlag*

Department of Biochemistry Stanford University, Stanford CA 94305-5307, USA

*Corresponding author

Previous studies have shown that the earliest detectable step in folding of the *Tetrahymena* ribozyme is tertiary structure formation of the peripheral element P5abc. This, along with other results, has suggested that P5abc may serve as a scaffold upon which additional tertiary structure is built. Herein we use the onset of oligonucleotide cleavage activity as a readout for native state formation and investigate the effect of P5abc on the rate of folding to the native structure. Despite the early folding of P5abc, its removal to give the $E^{\Delta P5abc}$ variant decreases the rate of attainment of an active structure less than fivefold (20-100 mM Mg^{2+}, 15-50 °C). Furthermore, P5abc added *in trans* is able to bind the folded $E^{\Delta P5abc}$ ribozyme and promote oligonucleotide cleavage at least tenfold more rapidly than folding of the wild-type ribozyme, indicating that $E^{\Delta P5abc}$ does not have to first unfold before productively binding P5abc to form the true native state. This suggests that a state with the overall tertiary structure formed but with P5abc unfolded represents a viable on-pathway intermediate for the wild-type ribozyme. These results provide strong evidence for the existence of two pathways to the native state: in one pathway P5abc forms tertiary structure first, and in another it forms late. The pathway in which P5abc forms first is favored because P5abc can fold quickly and because its tertiary structure is stable in the absence of additional structured elements, not because P5abc formation is required for subsequent folding steps. In the course of these experiments, we also found that most of the ribozyme population does not reach the native state directly under standard conditions *in vitro*, but instead forms an inactive structure that is stable for hours. Finally, the fraction that does fold to the native state folds with a single rate constant of 1 min^{-1}, suggesting that there are no significantly populated "fast-track" pathways that reach the native state directly by avoiding slow folding steps.

Keywords: group I intron; ribozyme; RNA folding; *Tetrahymena thermophila*

Introduction

The *Tetrahymena* ribozyme, derived from a self-splicing group I intron, has served as a valuable model system for understanding how an RNA molecule can achieve and maintain a discrete three-dimensional structure (for recent reviews, see Doudna & Cate, 1997; Strobel & Doudna, 1997; Brion & Westhof, 1997). The ribozyme is composed of a conserved catalytic core, containing the paired structural elements P4, P5, P6 and P3-P7-P8, and additional peripheral elements that are conserved among members of intron subclasses (Figure 1(a)). P4-P6, which includes the peripheral element P5abc, has been shown to constitute an independent folding domain, acquiring its native magnesium-dependent tertiary structure in the absence of other parts of the ribozyme (Murphy & Cech, 1993). This property has allowed determination of the crystal structure of the isolated P4-P6 domain, providing a high-resolution picture of this RNA structural element (Cate *et al.*, 1996a,b, 1997). In the

Abbreviations used: E, wild-type *Tetrahymena* ribozyme; $E^{\Delta P5abc}$, ribozyme variant lacking P5abc (Figure 1); G, guanosine; Mops, 3-(*N*-morpholino)propanesulfonic acid; P, oligonucleotide product; rP, the oligonucleotide product CCCUCU; S, oligonucleotide substrate; rSA_5, the oligonucleotide substrate $CCCUCUA_5$; $*rSA_5$, (5′-^{32}P)-labeled rSA_5.

E-mail address of the corresponding author: herschla@cmgm.stanford.edu

0022-2836/99/351155–13 $30.00/0

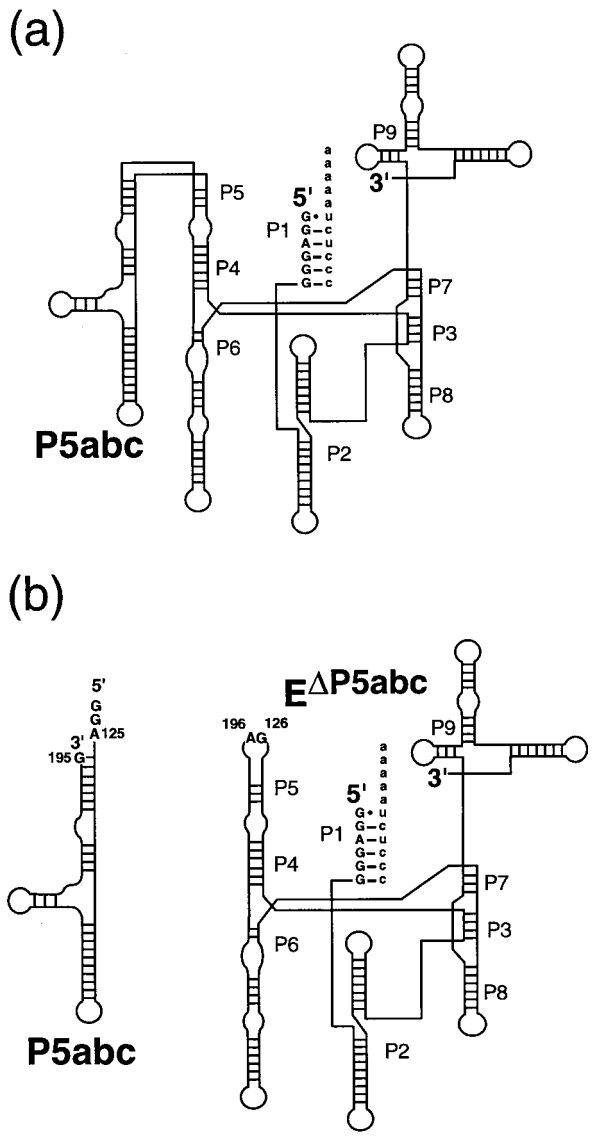


Figure 1. Secondary structure of the *Tetrahymena* ribozyme and $E^{\Delta P5abc}$ variant. (a) The ribozyme in an orientation that reflects the tertiary subdomains (Cech *et al.*, 1994). Paired structural elements are labeled with the abbreviations P1-P9. The catalytic core consists of P4, P5, P6 and P3-P7-P8, and the ribozyme additionally contains peripheral elements P2, P9, and P5abc (labeled in bold). P4, P5, P5abc, and P6 form a stable domain referred to as P4-P6. The oligonucleotide substrate rSA_5 ($CCCUCUA_5$, shown as lowercase letters) is base-paired to the ribozyme. (b) Secondary structure of the two-part ribozyme system, $E^{\Delta P5abc}$ and the separated P5abc.

context of the intact ribozyme, the P4-P6 domain acquires tertiary structure at lower concentrations of magnesium than the rest of the molecule (Celander & Cech, 1991) and maintains it at a higher temperature (Banerjee *et al.*, 1993). Furthermore, portions of P3-P7-P8 acquire stable tertiary structure only in the presence of P4-P6, suggesting that P4-P6 plays an essential role in the organization of the ribozyme's architecture (Doherty & Doudna, 1997).

The P4-P6 domain also appears to play a central role in the folding process of the ribozyme. Oligonucleotide hybridization and hydroxyl radical cleavage studies showed that P4-P6 forms tertiary contacts within seconds of initiating folding with the addition of Mg^{2+}, while portions of the catalytic core, including P3 and P7, require minutes to form stably (Zarrinkar & Williamson, 1994; Sclavi *et al.*, 1997). Rapid time-resolved hydroxyl radical cleavage studies have further refined the order of tertiary structure formation by showing that portions of P5abc become protected about twofold faster than other parts of P4-P6 (Sclavi *et al.*, 1998).

Both the thermodynamic and kinetic results have suggested a model for folding of the ribozyme in which P5abc acquires tertiary structure first and then allows the rest of the P4-P6 domain to become organized around it (Cate *et al.*, 1997; Sclavi *et al.*, 1998). This domain could then serve as a scaffold upon which the catalytic core is built in subsequent slower folding steps. Thus, the folding pathway may be hierarchical, with each step dependent on completion of the previous one.

A variant that lacks the P5abc subdomain, referred to here as $E^{\Delta P5abc}$ (Figure 1(b)), retains some catalytic activity at high concentrations of Mg^{2+} (Joyce & Inoue, 1987; Joyce *et al.*, 1989; M. Engelhardt, E. A. Doherty, D. S. Knitt, J. A. Doudna & D. Herschlag, unpublished results). Thus, folding to an active structure is not absolutely dependent on early formation of P5abc. However, folding without the early formation of P5abc could be severely compromised. The simplest expectation from the above model is that $E^{\Delta P5abc}$ would fold much more slowly than the wild-type.

To understand further the extent to which folding of the ribozyme is ordered, we have compared the rates of folding to the native state for the wild-type and $E^{\Delta P5abc}$ ribozymes. Techniques such as chemical protection and oligonucleotide hybridization are extremely powerful for monitoring the kinetics of structure formation (Zarrinkar & Williamson, 1994, 1996; Sclavi *et al.*, 1997, 1998). Herein, we use a complementary approach, exploiting the ability of the ribozyme to cleave a specific oligonucleotide as a probe for the native structure (Zarrinkar & Williamson, 1994; Pan & Sosnick, 1997). An advantage of this approach is that it monitors attainment of the native state specifically, distinguishing it from inactive forms that also possess tertiary structure. Remarkably, deletion of P5abc has only a small effect on the overall folding rate to the native state, demonstrating that this element is not required for the slower folding steps to proceed efficiently. We also found that under conditions used in this and several previous folding studies, most of the ribozyme population becomes trapped in an inactive

conformation rather than folding to the native state.

Results

Catalytic activity as a readout of native structure formation

The oligonucleotide cleavage activity of the *Tetrahymena* ribozyme was used to monitor attainment of the native structure for the wild-type ribozyme (E) and a variant in which P5abc is deleted ($E^{\Delta P5abc}$), as depicted in Figure 2. The ribozyme uses an exogenous guanosine (G) as a nucleophile, transferring the 3′ portion of an oligonucleotide substrate that mimics the 5′ splice site (S) to give a shorter oligonucleotide product (P) and GA_5 (equation (1) and Figure 2(a); reviewed by Cech & Herschlag, 1997; Narlikar & Herschlag, 1997):

$$\underset{S}{CCCUCUA_5} + G \rightharpoonup \underset{P}{CCCUCU} + GA_5 \quad (1)$$

In the presence of excess S and G, the reaction proceeds with multiple turnover (Zaug *et al.*, 1986, 1988).

Because the chemical cleavage step is much faster than folding of the wild-type ribozyme upon addition of Mg^{2+} (Herschlag & Cech, 1990; Zarrinkar & Williamson, 1994), the rate of formation of P during the first turnover is limited by native structure formation rather than by the chemical step itself (Figure 2(b), k_{fold}). That is, when Mg^{2+} is added to the ribozyme in the presence of saturating G and a small excess of the oligonucleotide substrate $CCCUCUA_5$ (rSA_5), a burst of the product CCCUCU (rP) is expected, and the observed rate constant for this burst is equal to the rate constant for folding to the native state, k_{fold}. The burst is expected to be followed by a slower steady-state increase in rP that is rate-limited by release of rP from the ribozyme (k^P_{off}), which is much slower than folding.

In contrast, release of rP from the $E^{\Delta P5abc}$ ribozyme is expected to be faster than folding because $E^{\Delta P5bc}$ releases rP much more quickly than E (M. Engelhardt *et al.*, unpublished results). In the presence of high concentrations of Mg^{2+}, chemical cleavage is also expected to be faster than folding ($k_c > k_{fold}$), so pre-steady-state formation of rP is limited by the rate of folding to the active state ($k_{obs} = k_{fold}$; Figure 2(c)). However, this pre-steady-state transient appears as a lag that precedes the faster steady-state formation of rP. The lag can be understood as an increase in the rate of product formation that occurs as the active ribozyme accumulates during folding. Lag kinetics are also expected using the wild-type ribozyme and a shorter substrate, $CUCUA_5$, as its product, CUCU, which lacks two residues that form base-pairs with the ribozyme, is released much faster than rP

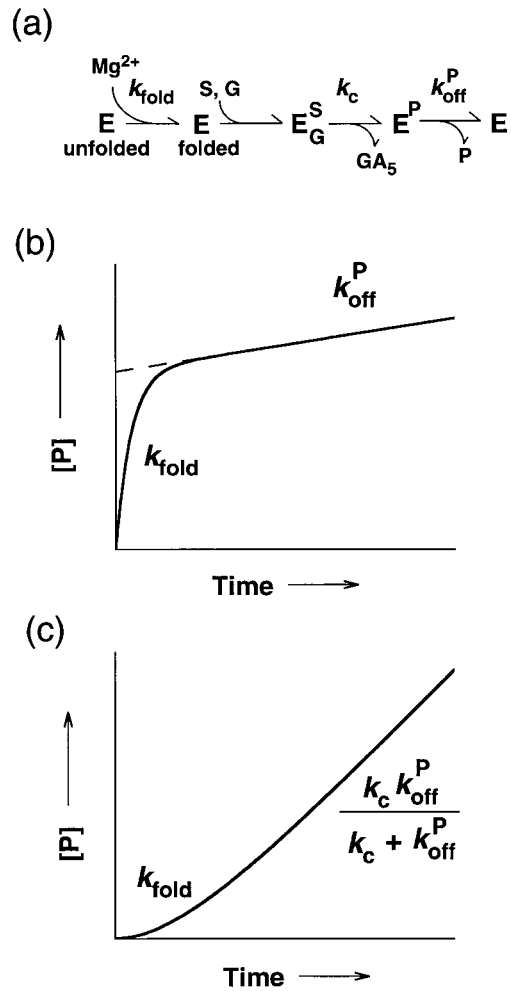


Figure 2. Ribozyme catalysis as a monitor for formation of the native state. (a) Upon addition of Mg^{2+}, the ribozyme folds to the native structure with a rate constant k_{fold}. In the presence of sufficient S and G both binding steps are fast relative to folding. The rate constant for the chemical cleavage step (k_c) is also large relative to k_{fold}. (The product GA_5 is released quickly, so that chemical cleavage is essentially irreversible with excess G (Herschlag & Cech, 1990).) The rapid cleavage of S relative to folding produces a burst of P that remains bound to E. Subsequently, P is released with a rate constant k^P_{off}, and E is able to catalyze another round of oligonucleotide cleavage. (b) Using the substrate rSA_5, product release is slower than folding ($k^P_{off} < k_{fold}$), so that a burst of rP is observed that is limited by k_{fold}, followed by a slower steady-state increase in rP. (c) The product of reaction with the substrate $CUCUA_5$ (CUCU) is released quickly, such that the rate constant for the steady-state reaction is greater than k_{fold}. This results in the appearance of a lag in the progress curve that gives k_{fold}. This scenario holds for the $E^{\Delta P5abc}$ variant with both the rSA_5 and $CUCUA_5$ oligonucleotide substrates, because product release is much faster for $E^{\Delta P5abc}$ than for the wild-type.

(Bevilacqua *et al.*, 1992; G. J. Narlikar, M. Khosla, L. E. Bartley & D. Herschlag, unpublished results).

Folding of the wild-type ribozyme

The rate constant for formation of the native state was determined by initiating folding with the addition of 10 mM Mg^{2+} in the presence of a small excess of ^{32}P-labeled rSA_5 (*rSA_5) and 1 mM G at 37 °C. A burst of product formation was observed with $k_{obs} = 1.5$ min^{-1} (Figure 3(a)). This rate constant is similar to the rate constant for global tertiary structure formation of 0.5-1 min^{-1} previously determined using oligonucleotide hybridization and hydroxyl radical cleavage (Zarrinkar & Williamson, 1994; Sclavi *et al.*, 1997, 1998). The observed rate constant was unaffected by a two-fold increase in rSA_5 concentration, by changes in G concentration (0.2-2 mM), and by changes in pH from 6.8 to 7.9 (data not shown), suggesting that k_{obs} is limited by the rate of native structure formation, rather than by binding of rSA_5 or G, or by the chemical step†. Consistent with this interpretation, an analogous reaction in which the ribozyme was pre-folded in 10 mM Mg^{2+} at 50°C for 30 minutes before adding rSA_5, conditions previously shown to result in maximal activity (Herschlag & Cech, 1990), gave a burst with a substantially larger rate constant (>8 min^{-1}; Figure 3(b)).

An equivalent folding experiment in which rSA_5 was replaced with the shorter substrate $CUCUA_5$ showed, as expected, an initial lag rather than a burst of product formation (Figure 3(c)). Nevertheless, data with the shorter substrate gave a value for k_{fold} of 1.1 min^{-1}, the same within error as the value of 1.5 min^{-1} determined with the longer substrate (Figure 3(a)).

To further characterize the kinetics of native state formation and to facilitate comparison with the $E^{\Delta P5abc}$ variant, the effects of changing temperature and Mg^{2+} concentration were examined. The observed rate constant increased ~40-fold from 15 °C to 50 °C (Table 1). The folding rate was independent of Mg^{2+} concentration from 5-100 mM at 37 °C (Table 2), confirming and extending previous results (Zarrinkar & Williamson, 1994).

Both for reactions containing rSA_5 and $CUCUA_5$, the onset of activity was well-described by a single rate constant (most clearly seen in Figure 3(a)). If a fraction of the ribozyme folded before the first time point in Figure 3(a) of two seconds, an unresolved burst of product formation from rSA_5 would be

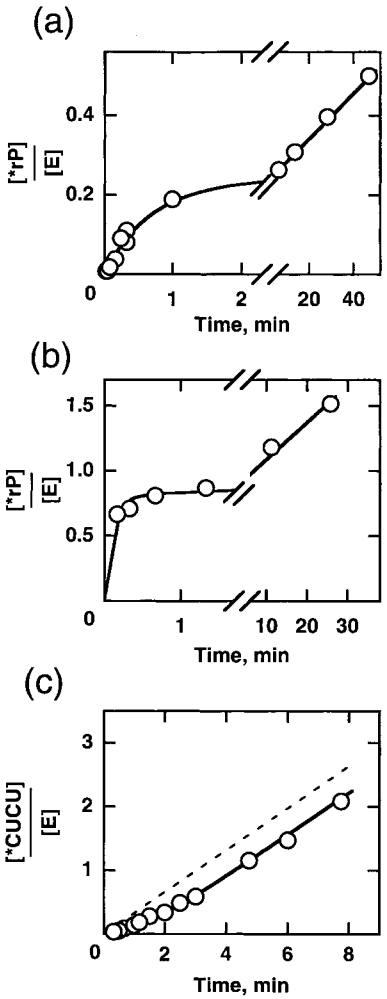


Figure 3. Tertiary folding of the wild-type ribozyme. (a) Folding of the ribozyme (200 nM) was initiated at 37 °C, pH 6.8 by the addition of 10 mM Mg^{2+} in the presence of 500 nM *rSA_5 and 1 mM G. Three independent determinations gave a value for k_{fold} of 1.5(±0.3) min^{-1}. The burst magnitude from three determinations was 0.27(±0.02) product/ribozyme, normalized to the equivalent reactions in which E was pre-folded (see below). The slower steady-state rate constant was 0.0086(±0.002) min^{-1}, uncorrected for ribozyme that folds to an inactive conformation (see Results). (b) E (200 nM) was pre-folded in the presence of 10 mM Mg^{2+} for 30 minutes at 50 °C. The cleavage reaction was subsequently initiated by adding 500 nM *rSA_5 and 1 mM G at 37 °C. The burst magnitude was 0.80(±0.03), and the steady-state rate constant was 0.031(±0.004) min^{-1}. (c) Folding of E (1 μM) measured using lag kinetics by initiating the folding reaction with 10 mM Mg^{2+} in the presence of 20 μM *$CUCUA_5$ and 1 mM G. Three independent determinations gave a value for k_{fold} of 1.1(±0.3) min^{-1}. The broken line shows the reaction progress that would occur in the absence of a lag for folding.

† The observed rate constant would be expected to be dependent on rSA_5 or G concentration if binding of either were rate-limiting for cleavage of rSA_5. Likewise, k_{obs} would be expected to vary log-linearly with pH if the chemical step were rate-limiting (Herschlag & Khosla, 1994). Further, independent determinations of the rate constants for rSA_5 binding and the chemical step (Herschlag & Cech, 1990) and determination of a minimum rate constant for G binding (Herschlag & Khosla, 1994) predict >50-fold larger observed rate constants for each of these steps than k_{obs} in these experiments, providing strong evidence that ribozyme folding is rate-limiting for the first round of chemical cleavage.

Table 1. Rate constants for native state formation as a function of temperature

	k_{fold}, (min^{-1})		
	Wild-type (E)		$E^{\Delta P5abc}$
Temperature (°C)	10 mM Mg^{2+}	50 mM Mg^{2+}	50 mM Mg^{2+}
15	0.14 ± 0.02^a	0.15^a	0.033^b
25	0.37 ± 0.10^a	0.30^a	0.10^b
37	1.5 ± 0.3^a	0.91 ± 0.1^a	0.40^a
	1.1 ± 0.3^b		0.54 ± 0.05^b
50	6.0 ± 1.5^a	5.1^a	3.4 ± 1.2^a

50 mM Na·Mops (pH 7.0) (determined at 25 °C). Values of k_{fold} reported with errors are the mean and standard deviations from three independent determinations. Values without errors are the results of a single determination.

[a] Reactions contained 200 nM E or $E^{\Delta P5abc}$ and 500-1200 nM *rSA_5. Values of k_{fold} for E are observed rate constants for the initial burst of product formation. Only a fraction of the ribozyme reached the native state on the timescale of these experiments. This fraction was determined by comparing the magnitude of the product burst with reactions in which the ribozyme was pre-folded in 10 mM Mg^{2+} at 50 °C for 30 minutes, with additional Mg^{2+} added subsequently for reactions performed with 50 mM Mg^{2+}. The fractions of E that folded directly to the native state were: 15 °C, 0.05; 25 °C, 0.12; 37 °C, 0.27; 50 °C, 0.45 (see Results). $E^{\Delta P5abc}$ was also observed to fold predominantly to an inactive conformation at low temperatures. The fractions of $E^{\Delta P5abc}$ that folded to the native state were: 15 °C, 0.09; 25 °C, 0.20. These values were determined by comparing the steady-state rates of product formation with and without the 30 minutes, 50 °C pre-incubation that allows folding to the active conformation. At temperatures greater than 25 °C, $E^{\Delta P5abc}$ converted from the inactive to active form efficiently, preventing accurate determination of the fraction that initially misfolded.

[b] Reactions contained 1 μM E or $E^{\Delta P5abc}$, and 20-50 μM *$CUCUA_5$. Values of k_{fold} are observed rate constants for the initial lag phase of product formation.

observed, giving a curve with a positive *y*-intercept. However, in Figure 3(a) and in all other cases the data were adequately fit by a curve that passes through the origin, indicating that if there is a population of fast-folding molecules, this population is a small fraction of the total (<1 %).

A large fraction of the ribozyme does not reach the native state

With rSA_5 in excess of E, the amount of product formed in the first turnover reflects the amount of active ribozyme. Therefore, to determine the fraction of the ribozyme that reached the native structure, the amplitude of the burst representing the first turnover was compared to that from a reaction in which the ribozyme was pre-folded in Mg^{2+} at 50 °C for 30 minutes before adding rSA_5. The pre-folded ribozyme gave a burst of product that was nearly stoichiometric with the ribozyme concentration determined spectrophotometrically (Figure 3(b); [*rP]/[E] =0.8). In contrast, the amplitude of the burst was about fourfold smaller without the pre-incubation (Figure 3(a)), suggesting that only 25 % of the ribozyme population attained the native structure. Similarly, the observation that the linear increase in product concentration was about fourfold faster with pre-folded E (Figure 3(a) and (b)) is most simply explained by a fourfold higher concentration of active ribozyme in reactions that had been pre-incubated with Mg^{2+}. Similar results were obtained under conditions identical to previous folding studies (see Materials and Methods). The fraction that folded to the native state was independent of ribozyme concentration (40 nM-2 μM, data not shown), suggesting that the inactive ribozyme is monomeric.

The steady-state phases, in reactions both with and without the pre-incubation, were linear for at least an hour (Figure 3(a) and (b)), and data not shown), suggesting that no significant conversion between the native and inactive forms occurs under these conditions on this time scale. If additional ribozyme were reaching the native state during this time, the slow phase would curve upward because the concentration of active ribozyme would be increasing with time. Further, independent experiments have confirmed that the conversion from inactive to active ribozyme does not occur to a significant extent on the time scale of these experiments (R.R. & D.H., unpublished results).

Table 2. Rate constants for native state formation as a function of Mg^{2+} concentration

	k_{fold} (min^{-1})	
[Mg^{2+}] (mM)	Wild-type (E)	$E^{\Delta P5abc}$
5	1.5	N.D.
10	1.5 ± 0.3	N.D.
20	0.92	0.59
30	0.90 ± 0.07	0.51
50	0.91 ± 0.10	0.54 ± 0.05
100	0.93 ± 0.03	0.46

37°C, 50 mM Na·Mops (pH 6.8). N.D., not determined. Values of k_{fold} reported with errors are the mean and standard deviations from three independent determinations. Values without errors are the results of a single determination.

The $E^{\Delta P5abc}$ variant folds with a rate similar to the wild-type

To examine the influence of P5abc on the kinetics of folding to the native state, the folding rate of the $E^{\Delta P5abc}$ variant was measured by following

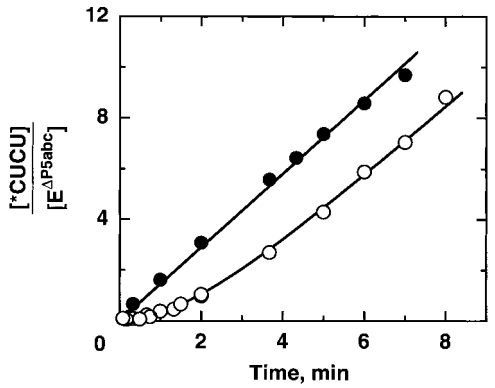


Figure 4. Folding of the $E^{\Delta P5abc}$ variant. *$CUCUA_5$ (50 μM) was added to 1 μM $E^{\Delta P5abc}$ at 37 °C (pH 6.8) with 50 mM Mg^{2+} (○) or after pre-folding $E^{\Delta P5abc}$ in the presence of 10 mM Mg^{2+} for 30 minutes at 50 °C, followed by the addition of Mg^{2+} to 50 mM (●). Three independent determinations gave a value for k_{fold} of 0.54(±0.02) min^{-1}. The oligonucleotide 5′-CUCG-3′ (0.5 mM), which binds to the ribozyme more tightly than G (Russell & Herschlag, 1999; M. Engelhardt *et al.*, unpublished results), was used as the nucleophile. This increases the steady-state rate, improving the resolution of the lag phase, without significantly affecting k_{fold} (data not shown). Although a substantial fraction of $E^{\Delta P5abc}$ folds to an inactive conformation, analogous to the wild-type ribozyme, the steady-state cleavage rate for $E^{\Delta P5abc}$ at 37 °C is the same whether or not the ribozyme is pre-folded to the active conformation. This is because $E^{\Delta P5abc}$ refolds to the active conformation on the same timescale as the initial folding lag. This renders the observed rate constant from the lag a lower limit for the initial folding rate constant. Because refolding is substantially slowed at lower temperatures, accurate determination of both the initial folding rate and the partitioning between active and inactive conformations is possible at temperatures below 37 °C (see Table 1).

the lag in the onset of maximal steady-state activity, as described above (Figure 2(c)). Folding reactions initiated by the addition of 50 mM Mg^{2+} gave a lag, followed by a faster steady-state rate of product formation (Figure 4, open symbols). Pre-incubation of $E^{\Delta P5abc}$ with Mg^{2+} eliminated the lag (Figure 4, filled symbols), suggesting that the lag reflects formation of an active structure.

The observed folding rate constant was 0.54 min^{-1} at 37 °C, within twofold of that for the wild-type under the same conditions (0.91 min^{-1}), and similar results were obtained from 15-50 °C (Table 1). The shorter substrate, $CUCUA_5$, was used instead of rSA_5 at temperatures <37 °C to increase the steady-state cleavage rate of $E^{\Delta P5abc}$, which is limited for reactions with rSA_5 by release of the product, rP. The steady-state rate was increased about twofold at 37 °C with $CUCUA_5$, and substantially more at lower temperatures (data not shown). The faster steady-state reaction led to a larger difference between the rate constant for folding and that for the steady-state reaction, enhancing the sensitivity in the determination of k_{fold}. The substrates rSA_5 and $CUCUA_5$ gave similar values for k_{fold} at 37 °C (0.4 min^{-1} and 0.5 min^{-1}; Table 1). At lower temperatures the steady-state reactions with rSA_5 were sufficiently slow so that lags were not detectable (data not shown); thus, only $CUCUA_5$ was used to determine k_{fold}.

For further comparison between $E^{\Delta P5abc}$ and the wild-type, k_{fold} for $E^{\Delta P5abc}$ was determined at several concentrations of Mg^{2+}. Similarly to the wild-type, no dependence on Mg^{2+} concentration was observed from 20-100 mM (Table 2). At concentrations of Mg^{2+} lower than 20 mM the chemical step is compromised such that it is expected to be slower than the rate of folding (M. Engelhardt *et al.*, unpublished results) and lags were not observed under these conditions (data not shown). The value of k_{fold} at 10 mM Mg^{2+} was estimated by initiating folding by the addition of Mg^{2+}, and then adding additional Mg^{2+} and P5abc *in trans* at various times to rescue activity of the fraction of $E^{\Delta P5abc}$ that had folded (P5abc can bind to $E^{\Delta P5abc}$ and rescue activity much faster than the folding of $E^{\Delta P5abc}$; see below). The fraction folded was determined from the amount of rP formed from rSA_5 in an initial burst, and gave an estimate for k_{fold} of 1.6(±0.7) min^{-1} (data not shown). Although this method was less precise than those described above, k_{fold} was similar to values for $E^{\Delta P5abc}$ obtained at higher Mg^{2+} concentrations, and to values of k_{fold} for E at all Mg^{2+} concentrations tested.

P5abc binds quickly to folded $E^{\Delta P5abc}$

The result that the $E^{\Delta P5abc}$ variant folds to an active conformation nearly as quickly as the wild-type demonstrates that formation of P5abc structure early in the folding pathway does not have a large effect on the rate of folding to a conformation that has enzymatic activity. Further, these results raise the possibility that there is an additional folding pathway for the wild-type ribozyme in which P5abc acquires tertiary structure last rather than first (Figure 5(a)). However, although the global folds are similar (Doherty *et al.*, 1999), the active conformation attained by $E^{\Delta P5abc}$ is not the native conformation reached by the wild-type, as the folded $E^{\Delta P5abc}$ continues to lack P5abc and is compromised in activity (Joyce *et al.*, 1989; M. Engelhardt *et al.*, unpublished results). Thus, it was necessary to determine whether a structure with the global fold but without a structured P5abc can readily convert to the native structure containing P5abc without first unfolding, forming P5abc tertiary structure early, and progressing along the established folding pathway (Figure 5(a)).

This possibility was tested by measuring the rate at which P5abc is able to bind to the pre-folded $E^{\Delta P5abc}$ and promote oligonucleotide cleavage when P5abc is added as a separate molecule (Figure 5(b)). If $E^{\Delta P5abc}$ were required to unfold

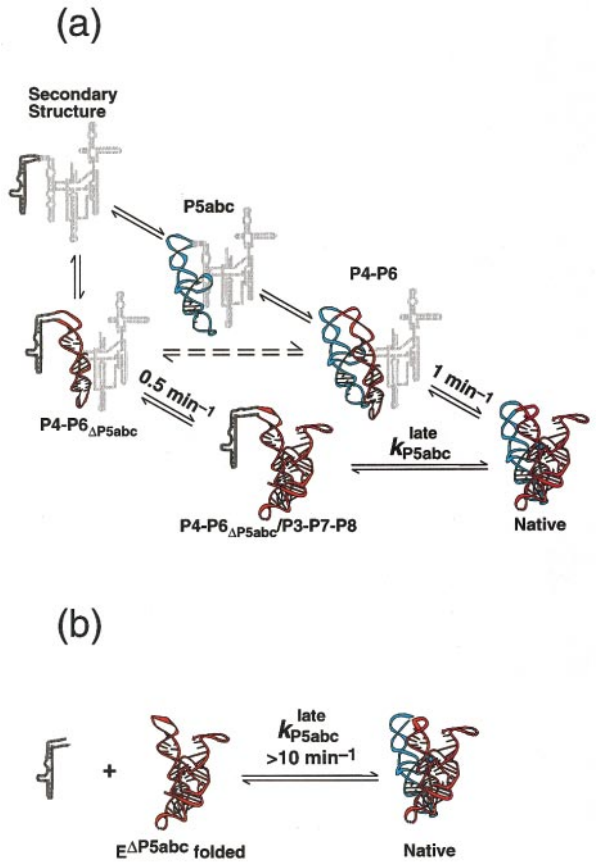


Figure 5. A branched folding pathway to the native state. (a) P5abc can form tertiary structure early in folding (top pathway), or late in folding, after the rate limiting folding step (bottom pathway). The broken arrow is included to emphasize the likelihood that P5abc can also fold at intermediate stages of the overall folding process. This is expected from the view that P5abc folding is largely independent from folding of other domains of the ribozyme, although it should be recognized that our experimental evidence addresses only the lower pathway in which P5abc folds last. Each structural element is depicted as changing color from gray to red as it acquires tertiary structure. P5abc is highlighted in dark gray for the unfolded state, and blue for the folded state. The rate constant k^{late}_{P5abc} represents the first-order rate constant for formation of P5abc late in folding. Only the ribozyme core and P5abc are shown for simplicity. (b) Schematic of P5abc binding to the folded $E^{\Delta P5abc}$, completing the lower folding pathway in (a) by forming a native structure with wild-type oligonucleotide cleavage activity. Results of an experiment in which P5abc was added in *trans* to the folded $E^{\Delta P5abc}$ provide an estimated lower limit for k^{late}_{P5abc} of 10 min^{-1}, because the observed rate constant includes both k^{late}_{P5abc} and a second-order rate constant for P5abc binding to $E^{\Delta P5abc}$ (see Results). (Adapted from an original drawing by L. Jaeger.)

prior to productively binding P5abc, the onset of activity could be no faster than the rate constant for wild-type folding (Figure 5(a); 1 min^{-1}). Conversely if $E^{\Delta P5abc}$ does not have to first unfold, the

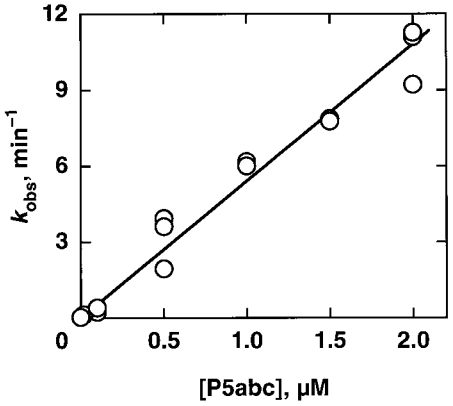


Figure 6. Unfolding of $E^{\Delta P5abc}$ is not required for binding and activation by P5abc. $E^{\Delta P5abc}$ (20-200 nM) was pre-folded in 10 mM Mg^{2+}, allowed to bind a trace amount of $*rSA_5$, and P5abc was added *in trans* with G. Observed rate constants were independent of $E^{\Delta P5abc}$ concentration (10-50 nM) as long as P5abc was in excess, and independent of pH from 6.8-7.9. The pH independence suggests that P5abc binding is rate-limiting for the cleavage reaction rather than the chemical step, which is log-linearly dependent on pH (Herschlag & Khosla, 1994). The slope of the fitted line gives an apparent second-order rate constant for P5abc binding of 5×10^6 M^{-1} min^{-1}.

simplest expectation is that cleavage activity would be limited by the rate at which P5abc binds to $E^{\Delta P5abc}$.

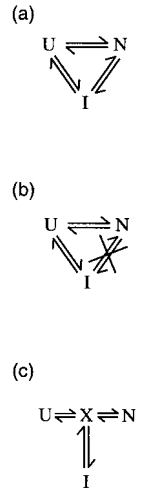
Thus, the $E^{\Delta P5abc}$ ribozyme was pre-folded in 10 mM Mg^{2+}, a concentration in which it is globally folded but only weakly active (Doherty *et al.*, 1999; M. Engelhardt *et al.*, unpublished results), and a trace amount of $*rSA_5$ was allowed to bind to the ribozyme. P5abc was then added in *trans* with G, and cleavage of $*rSA_5$ was followed (Figure 6). The observed rate constant was linearly dependent on P5abc concentration to >1 μM P5abc, giving an apparent second-order rate constant of 5×10^6 M^{-1} min^{-1}. This most likely represents the rate constant for P5abc binding to $E^{\Delta P5abc}$, and is in good agreement with pulse-chase experiments using native polyacrylamide gel electrophoresis to separate bound and unbound P5abc (E.A. Doherty, R.R., M. Engelhardt, D.H. & J.A. Doudna, unpublished results). The maximal value of k_{obs} of 10 min^{-1} represents a lower limit for adoption of the native structure upon P5abc binding (k^{late}_{P5abc} in Figure 5). This is considerably larger than the rate constant for folding of the wild-type of 1 min^{-1}, which would be expected to limit the onset of activity if $E^{\Delta P5abc}$ were required to globally unfold before productively binding P5abc. It remains possible that some local unfolding of $E^{\Delta P5abc}$ occurs before P5abc binds, but this experiment demonstrates that the folded $E^{\Delta P5abc}$ can bind P5abc to form the native structure without unfolding past the rate-limiting folding step.

Discussion

It was previously shown that the peripheral element P5abc acquires tertiary structure more quickly than any other structural element in the *Tetrahymena* ribozyme. This observation, along with considerations of the overall tertiary architecture, led to a model in which P5abc could serve as a nucleation site for additional tertiary structure, with a cluster of five Mg ions in P5abc possibly functioning analogously to the hydrophobic core of proteins (Cate *et al.*, 1997; Sclavi *et al.*, 1998). Here, we have used the oligonucleotide cleavage activity of the ribozyme as a readout for native structure formation and have examined the effect of removing P5abc on the kinetics of folding to the native state. Preventing formation of P5abc tertiary structure by removing the domain has only a small effect on the overall rate of folding to an active structure under standard *in vitro* conditions. Additionally, the rapid onset of full activity of the $E^{\Delta P5abc}\cdot$P5abc complex upon addition of P5abc to the folded $E^{\Delta P5abc}$ shows that an intermediate with the global tertiary fold but lacking P5abc structure is competent for folding (see below). These results strongly suggest that a pathway in which P5abc forms late exists for folding of the wild-type ribozyme. It is likely that there are additional pathways in which P5abc folds at intermediate stages in the overall folding process, as illustrated in Figure 5(a) (broken arrow). However, in folding of the wild-type ribozyme under typical *in vitro* conditions, P5abc folds first in most molecules. This is simply because the P5abc element can fold fast and because it forms a stable structure in the absence of other surrounding tertiary structural elements, not because the overall folding process is dependent on the presence of a structured P5abc.

An intermediate lacking a structured P5abc is competent for folding

It has been proposed that RNA folds through multiple pathways (Thirumalai & Woodson, 1996; see below). As noted above, the observation that the $E^{\Delta P5abc}$ variant forms an active structure nearly as rapidly as the wild-type suggests that an intermediate with the global fold but lacking a structured P5abc lies along an alternative folding pathway for the wild-type (Scheme 1(a); U refers to the unfolded state, N refers to the native state, and I refers to the intermediate that lacks P5abc structure). However, this evidence alone leaves open the possibility that the intermediate lacking P5abc represents an off-pathway intermediate for the wild-type that must fully or partially unfold back to an structure that is on the pathway before reaching the native state (Scheme 1(b), U or Scheme 1(c), X. X represents a folding intermediate that is on the pathway and is at the branchpoint for correct *versus* incorrect folding). That is, the intermediate that has achieved the global fold but lacks P5abc structure could be a dead-end that branches from a linear pathway rather than an intermediate that lies along an alternative pathway.



Whether the partially active, folded $E^{\Delta P5abc}$ is on an alternative folding pathway (Scheme 1(a)) or off the folding pathway (Scheme 1(b) or (c)) was distinguished by adding P5abc in *trans* to the folded $E^{\Delta P5abc}$. The observation that $E^{\Delta P5abc}$ is able to productively bind P5abc much faster than the wild-type ribozyme folds rules out the possibility that $E^{\Delta P5abc}$ must unfold completely (Scheme 1(b)), because after unfolding and binding P5abc the ribozyme would be required to fold along the "normal" wild-type pathway with a rate constant of 1 min^{-1}.

Scheme 1(c) is ruled out as follows. Previous studies showed that P5abc forms tertiary structure prior to the step that is rate-limiting for folding (Zarrinkar & Williamson, 1994; Sclavi *et al.*, 1997, 1998). Thus, any alternative pathway in which P5abc is formed late must branch from the normal pathway prior to the rate-limiting step. If the active $E^{\Delta P5abc}$ were a dead-end folding intermediate that branched from the normal pathway, to reach the native state it would have to return to the pathway at an intermediate in which P5abc is unfolded (X in Scheme 1(c)). This would, however, require transit through the slow step in overall folding, giving a rate constant for activation of 1 min^{-1}. Thus, the observed activation of the folded $E^{\Delta P5abc}$ upon addition of P5abc with a rate constant of >10 min^{-1} strongly suggests that the ribozyme can follow a folding pathway in which P5abc folds last rather than first.

It was recently suggested that folding of the *Tetrahymena* pre-rRNA self-splicing precursor, from which the ribozyme is derived, also proceeds through multiple pathways analogous to Scheme 1(a), with a portion folding fast and a portion folding slowly due to the presence of a kinetic trap involving an alternative secondary pairing (Thirumalai & Woodson, 1996; Pan *et al.*, 1997; Pan & Woodson, 1998). These studies convincingly

demonstrate the existence of a kinetic trap in the folding pathway, but they do not establish the relationship of the trapped species to the overall folding pathway. As suggested by Pan *et al*. (1997), the kinetic trap could represent an intermediate on a pathway that reaches the native state parallel to the direct pathway (Scheme 1(a)). However, it is also possible that to reach the native structure the kinetically trapped intermediate must first fully or partially unfold, returning to a linear pathway from which it can again partition between productive and non-productive folding, analogous to Scheme 1(b) or (c).

Preferred pathways in ribozyme folding

The results herein demonstrate that the native state can be achieved with nearly the same overall kinetics *via* two distinct pathways. Nevertheless, it should be emphasized that the pathway in which P5abc folds early will be greatly preferred under normal folding circumstances. This preference is expected to arise simply from kinetic competition between the two pathways. The rate constant for P5abc tertiary structure formation and assembly of P4-P6 is ~50 min^{-1} (Sclavi *et al*., 1998), much larger than the rate constant for overall folding in the absence of P5abc of 0.5 min^{-1} (Table 1 and Figure 5). This difference leads to the expectation of a 100:1 (50 min^{-1}/0.5 min^{-1}) preference for the pathway with early P5abc formation. This preference is expected because once P4-P6 forms it is stable (Murphy & Cech, 1993; Zarrinkar & Williamson, 1994; Sclavi *et al*., 1998), ensuring that no more than a small fraction of the intermediate with P4-P6 folded will subsequently unfold P5abc to follow the pathway in which P5abc folds last. The early formation of P5abc is an example of what will be referred to as ''independent'' order in folding. The formation of P5abc has little effect on the overall folding rate, but it forms early most of the time simply because it is able to form quickly.

In an extreme model, P5abc folding can be viewed as entirely independent from folding of the rest of the ribozyme. In this model, P5abc is equally likely to fold at any point in the overall folding process. Furthermore, its presence as a folded element has no influence on the folding rates of any other structural elements. This view is consistent with a structural model of the ribozyme (Lehnert *et al*., 1996), in which P5abc lies on the surface of the molecule and can be imagined to fold and form tertiary contacts with the rest of P4-P6 throughout the folding process without influencing folding of other domains.

However, it remains possible that tertiary structure formation of P5abc does increase the rates of some folding steps without substantially affecting the overall folding rate. More generally, there may be steps in folding that occur in a preferred order because formation of one structural element is facilitated by the prior formation of other structure, referred to herein as ''dependent'' order (Zarrinkar & Williamson, 1994, 1996; Brion & Westhof, 1997). Even for the $E^{\Delta P5abc}$ variant, formation of P4-$P6_{\Delta P5abc}$ structure (P4-P6 without P5abc) likely precedes P3-P7-P8 tertiary structure formation. This view is supported by the crystal structure of a 247-nucleotide derivative of the ribozyme, which has shown that the P3-P7-P8 domain wraps around the face of P4-P6 opposite from P5abc, suggesting that formation of P3-P7-P8 tertiary structure is essentially dependent on the prior formation of P4-P6 (Golden *et al*., 1998). Additionally, P3-P7-P8 does not form stable tertiary structure in the absence of P4-P6 (Doherty & Doudna, 1997), consistent with the idea that P3-P7-P8 formation is dependent on P4-P6. It is possible that deletion of P5abc slows folding of P4-P6 considerably. A decrease of up to ~50-fold in the rate of formation of P4-P6 caused by the deletion of P5abc would not have been detected in these experiments because P4-$P6_{\Delta P5abc}$ formation would remain faster than the subsequent rate-limiting step. Oligonucleotide or hydroxyl radical protection approaches will be required to address this question by comparing the rate of formation of P4-$P6_{\Delta P5abc}$ in folding of the $E^{\Delta P5abc}$ ribozyme with the rate of formation of P4-P6 in folding of the wild-type ribozyme.

Finally, it is possible that order in folding that arises in an independent fashion under one set of conditions has a dependent origin under another set of conditions. Herein, the early formation of P5abc represents independent order. However, at physiological concentrations of Mg^{2+} (0.5-1 mM) it is likely that P5abc formation increases the overall folding rate; that is, its early formation is an example of dependent order. This is because the P4-P6 domain appears to be unstable in the absence of P5abc with physiological concentrations of Mg^{2+} (Szewczak *et al*., 1998). If the formation of P4-P6 structure facilitates subseqent steps such as formation of P3-P7-P8, then by stabilizing this otherwise unstable P4-P6 intermediate, P5abc folding would increase the overall folding rate.

Additional new pathways and RNA folding landscapes

A misfolded species is populated

Surprisingly, we found that under typical *in vitro* folding conditions (37 °C, 10 mM Mg^{2+}) only 25 % of the ribozyme population reached the native state quickly, folding with a single rate constant of ~1 min^{-1}. The remaining 75 % failed to reach the native state within an hour at 37 °C. This problem was even more severe at lower temperature, with only 10 % reaching the native state in the initial phase of folding at 25 °C (Table 1). These results suggest that upon initiation of tertiary folding with

the addition of Mg^{2+}, the ribozyme population partitions between pathways leading to the native structure and one or more pathways to at least one stable inactive conformation (see also Walstrum & Uhlenbeck, 1990). This partitioning was observed under conditions identical to those used in previous folding studies (see Materials and Methods), suggesting that these earlier studies monitored folding primarily to an inactive conformation†. This finding illustrates the value of using chemical reactivity to follow RNA folding in conjunction with direct physical techniques. The two approaches are complementary because the former can distinguish native state formation from more general tertiary structure formation, whereas the latter can detect and probe the structures of intermediates that accumulate during folding.

"Fast-track" pathways are not significantly populated

It has been suggested that proteins can fold through direct, fast-track pathways, avoiding kinetically trapped intermediates that slow folding. Partitioning between fast and slow pathways has been observed for the hen egg white lysozyme protein, with a fraction acquiring stable tertiary structure much faster than the rest of the population (Radford *et al*., 1992; Kiefhaber, 1995). Moreover, several small, single-domain proteins have been observed to fold on the microsecond to millisecond timescale with two-state kinetics, suggesting that these proteins fold without accumulating any folding intermediates (Jackson & Fersht, 1991; Huang & Oas, 1995; Schindler *et al*., 1995). Fast-track pathways have also been proposed for the folding of RNAs (Thirumalai & Woodson, 1996; Pan *et al*., 1997). However, direct folding to the native state is expected to be less likely for large, multi-domain proteins and RNAs than for small ones (Thirumalai & Woodson, 1996; Dill & Chan, 1997). More generally, although one can readily imagine folding processes that are continuously downhill in potential energy, free energy barriers for structure formation may nevertheless exist, giving barriers to folding of various magnitudes. Additionally, the extent to which specific intermediates facilitate or slow folding processes has not been resolved (see Baldwin, 1995).

Herein, folding of the *Tetrahymena* ribozyme was observed to proceed without a detectable burst, indicating that a population of "fast-track" folding molecules, if present, must represent <1 % of the total. This is consistent with previous folding studies on both the *Tetrahymena* ribozyme and RNase P RNA in which no rapid bursts of global structure or activity were detected (Zarrinkar & Williamson, 1994; Zarrinkar *et al*., 1996; Pan & Sosnick, 1997).

Although a direct folding pathway has been suggested to exist for the *Tetrahymena* self-splicing intron from which the ribozyme is derived (Pan *et al*., 1997), the presence of a significant fraction of the intron folding along such a pathway is not firmly established. This proposal was based on the inference that roughly 8 % of the intron population acquired native-like tertiary structure within a few seconds, while the remainder was slowed by kinetic traps and folded on the timescale of minutes. However, the results can also be explained without invoking a fast-folding fraction, as the data from which a fast-folding fraction was inferred are fit reasonably well by an equation for a single exponential that does not include a burst (Figure 3(b) of Pan *et al.* (1997)). Our results and those of others suggest that for large RNAs like the *Tetrahymena* ribozyme, folding pathways that lead directly to the native state by avoiding slow steps are not significantly populated.

Why might fast-track folding pathways not be significantly populated for large RNAs like the *Tetrahymena* ribozyme? It may be that evasion of the slow folding steps for large RNAs requires an exceedingly unlikely event early in folding. This unlikely event could be the fortuitous pre-positioning of two domains relative to one another, avoiding the necessity of a slow re-orientation step (Zarrinkar & Williamson, 1996). Another possibility is that the formation of tertiary contacts in a strongly non-preferred order avoids the formation of a stable interaction that must subsequently be broken for folding to proceed along the populated folding pathway (Treiber *et al*., 1998; Pan & Woodson, 1998; Pan & Sosnick, 1997). Finally, as noted above, there may in general be free energy barriers to complex folding processes.

Implications for an RNA folding landscape

Because P5abc is able to acquire tertiary structure quickly, its formation is observed to be the first step in folding of the ribozyme. However, experiments herein have exposed an alternative pathway on the folding landscape in which P5abc forms last rather than first. This pathway had previously been obscured because only a very small fraction of the population, estimated to be 1 %, use it. Although the preference against it rendered the alternative pathway undetectable, the energetic preference against it is only ~3 kcal/mol. This amount of energetic stabilization can be achieved by the formation of as few as two hydrogen bonds, suggesting that elimination of only a few contacts that are formed in the transition state ensemble for P5abc folding could dramatically restructure the

† The protection from oligonucleotides or hydroxyl radicals observed in previous studies exceeded 25 % (Zarrinkar & Williamson, 1994; Sclavi *et al*., 1998), the fraction shown herein to fold correctly, suggesting that there is also protection in the inactive conformation. The rate constants observed previously for overall folding are similar those observed here for folding specifically to the native state, suggesting that the native and inactive conformations are formed with similar rate constants, a conclusion that is supported by additional experiments (R.R. & D.H., unpublished results).

folding landscape such that the pathway in which P5abc folds last is favored.

More generally, it is likely that the number of pathways that are observed in macromolecular folding experiments will be dependent on the sensitivity of the approach. Observation of a folding pathway typically requires that a large fraction of the RNA or protein population traverse it; thus, less-traveled pathways remain unnoticed. For example, the lack of detection of a ''fast-track'' pathway herein does not rule out the existence of such a pathway; it merely sets an upper limit of 1 % on the fraction of ribozyme molecules that fold along it. By removing P5abc, and thereby forcing the ribozyme to fold by an alternative pathway, we were able to increase the experimental sensitivity for detection of this specific alternative folding pathway. It is expected that additional folding pathways for the ribozyme could similarly be revealed by blocking specific steps along the preferred folding pathway. The identification of additional folding pathways, coupled with quantitative determination of the partitioning between these pathways, should allow probing of the molecular features underlying the ''choices'' made by an RNA population as it partitions between the multiple pathways that make up a folding landscape.

Materials and Methods

Materials

To prepare P5abc RNA, a pUC19 derivative carrying P5abc DNA was linearized with *Bsa*I (Doherty *et al.*, 1999). P5abc, and the L-21 *Sca*I (E) and $E^{\Delta P5abc}$ ribozymes, were prepared by *in vitro* transcription as described (Zaug *et al.*, 1988). All RNAs were purified using a Qiagen RNeasy column following the manufacturer's instructions; this technique has been shown to yield ribozyme of identical activity to that obtained by gel purification (Russell & Herschlag, 1999). Oligonucleotide substrates were synthesized using standard solid-phase methods by the Protein and Nucleic Acid Facility at Stanford, and were HPLC purified as described (Russell & Herschlag, 1999). Concentrations of RNAs were determined spectrophotometrically. The following extinction coefficients (260 nm) were used: L-21 *Sca*I ribozyme, 3.9×10^6 M^{-1} cm^{-1}; $E^{\Delta P5abc}$, 3.2×10^6 M^{-1} cm^{-1}; P5abc, 7.0×10^5 M^{-1} cm^{-1}; rSA_5, 1.3×10^5 M^{-1} cm^{-1}; $CUCUA_5$, 1.1×10^5 M^{-1} cm^{-1}. Oligonucleotide substrates were 5′ end-labeled with [γ-^{32}P]ATP using T4 polynucleotide kinase and purified by non-denaturing polyacrylamide gel electrophoresis as described (Zaug *et al.*, 1988; Herschlag *et al.*, 1993).

General kinetics methods

For all reactions, the wild-type and $E^{\Delta P5abc}$ ribozymes were heated to 95 °C in the absence of Mg^{2+} for one minute to remove secondary structure, then cooled to the desired temperature. In experiments following the activity of pre-folded ribozyme, the ribozyme was subsequently incubated for 30 minutes at 50 °C in the presence of 10 mM Mg^{2+}. Additional Mg^{2+} was then added to achieve the desired concentration. Reaction conditions were 50 mM Na·Mops (pH 7.0) (determined at 25 °C), with the indicated concentration of Mg^{2+} added as $MgCl_2$. Reactions were followed by separating the ^{32}P-labeled reaction product from the labeled substrate using 20 % (w/v) polyacrylamide/8 M urea gel electrophoresis, and the fraction of label present as product was determined using a Molecular Dynamics Phosphorimager. For slow reactions, time courses were obtained by removing aliquots from a reaction and quenching them with two volumes of 20 mM EDTA in 90 % (v/v) formamide with 0.005 % (w/v) xylene cyanol, and 0.01 % (w/v) bromophenol blue. Reactions containing higher concentrations of Mg^{2+} were stopped with a quench solution containing 100 mM EDTA. To follow reactions at times less than ten seconds, quench solution was added directly to the reaction, such that each time point represents a separate reaction.

Multiple turnover kinetics

To measure formation of the native state for the wild-type ribozyme, a small excess of the substrate rSA_5 was used (200 nM ribozyme and 500-1200 nM rSA_5, including ~0.1 nM (5′-^{32}P)-labeled rSA_5, abbreviated *rSA_5. All concentrations given, here and elsewhere, are final concentrations). Reactions were initiated by adding Mg^{2+} and rSA_5 to a solution containing the ribozyme and 1 mM G. Progress curves displayed a burst of product formation, followed by a slower linear increase. The data were adequately described by equation (2), in which $frac_t^P$ is the fraction of label present as product at a given time (t):

$$frac_t^P = A(1 - e^{-k_{obs}t}) + st \quad (2)$$

A is the amplitude of the burst, k_{obs} is the observed rate constant for oligonucleotide cleavage which is rate-limited by native state formation ($k_{obs} = k_{fold}$), and s is the slope of the slower phase, limited by product release (see Results).

This approach measures the rate constant for folding in the presence of bound rSA_5, because the substrate is completely bound by the ribozyme simply by base-pairing under these conditions with a rate constant much faster than that for folding (Herschlag & Cech, 1990). A technique that uses ribozyme activity to measure folding in the absence of bound substrate by adding substrate at various times after initiating folding with Mg^{2+} has also been developed (Zarrinkar & Williamson, 1994; see also Pan & Sosnick, 1997; Pan *et al.*, 1999). The two methods yielded identical values of k_{fold} for the wild-type ribozyme at 37 °C (Figure 3(a) and data not shown), confirming previous results (Zarrinkar & Williamson, 1994) and suggesting that bound rSA_5 does not affect the overall rate of folding.

To compare directly the results herein with those from previous folding studies (Zarrinkar & Williamson, 1994; Sclavi *et al.*, 1998), some folding reactions were performed under identical conditions to those used previously. A reaction performed in the presence of 50 mM Tris-HCl (pH 8.1), 10 mM NaCl, 0.1 mM EDTA, 1 mM dithiothreitol, 10 mM Mg^{2+} at 37 °C (Zarrinkar & Williamson, 1994; $k_{fold} = 0.72$ min^{-1}) gave a value for k_{fold} of 1.7 min^{-1} and a fraction of 0.24 folded to the native state. Similarly, a reaction performed in the presence of 10 mM sodium cacodylate (pH 7.1), 10 mM Mg^{2+} at 42 °C (Sclavi *et al.*, 1998; $k_{fold} = 1$ min^{-1}) gave a value for k_{fold} of 1.7 min^{-1} and fraction of 0.28 folded to the native state.

Steady-state lag kinetics to measure folding

Formation of the native state for the $E^{\Delta P5abc}$ and wild-type ribozymes was monitored by the approach to steady-state using the substrate $CUCUA_5$ (20-50 μM including ~0.1 nM $*CUCUA_5$) or rSA_5 (2.5 μM including ~0.1 nM $*rSA_5$) in excess of the ribozyme (0.1-1 μM). Reactions were initiated with the addition of Mg^{2+} and the oligonucleotide substrate to solutions containing the ribozyme and 1 mM G or 0.5 mM CUCG, and progress curves displayed a lag followed by a linear increase of product. The data were adequately described by equation (3), in which $\mathrm{frac}_t^{\mathrm{P}}$ is the fraction of label present as product at a given time (t):

$$\mathrm{frac}_t^{\mathrm{P}} = s\left[t - \frac{1}{k_{\mathrm{obs}}} + \frac{1}{k_{\mathrm{obs}}}(e^{-k_{\mathrm{obs}}t})\right] \quad (3)$$

s is the slope of the linear portion, and k_{obs} is the rate constant describing the lag, which gives the folding rate constant. Equation (3) can be obtained by integrating equation (4):

$$\frac{[\mathrm{d}(\mathrm{frac}_t^{\mathrm{P}})]}{\mathrm{d}t} = s(1 - e^{-k_{\mathrm{obs}}t}) \quad (4)$$

which describes an exponential approach to steady-state.

Onset of enhanced activity upon P5abc binding

The $E^{\Delta P5abc}$ ribozyme (20-200 nM) was pre-folded for 30 minutes at 50 °C in the presence of 10 mM Mg^{2+}. A trace amount of $*rSA_5$ was added and incubated two minutes, sufficient time to completely bind the ribozyme (Herschlag & Cech, 1990; M. Engelhardt *et al.*, unpublished results). P5abc (50 nM-1 μM) and 1 mM G were added to initiate reactions. A first-order rate equation was fit to the data to give k_{obs}, the rate constant describing the onset of activity. Reactions followed good first-order kinetics, except at concentrations of P5abc sufficient to give $k_{\mathrm{obs}} > 0.4\ \mathrm{min}^{-1}$, in which case a slower phase was observed (~0.3 min^{-1}, amplitude ~40 % of input $*rSA_5$). Adding excess unlabeled rSA_5 with P5abc eliminated the slow phase without affecting the fast phase (data not shown), suggesting that the slow phase represents dissociation of $*rSA_5$ from an inactive form of the ribozyme and re-binding to the active fraction. Additional experiments have suggested that a fraction of $E^{\Delta P5abc}$ exists as an inactive species that is in equilibrium with the native state, but exchanges only very slowly in the presence of bound P5abc (R.R. & D.H., unpublished results).

Acknowledgements

We thank Alan Lambowitz for the generous gift of the plasmid encoding $E^{\Delta P5abc}$, and Liz Doherty and Jennifer Doudna for the gift of the plasmid encoding P5abc. Thanks to Buzz Baldwin and members of the Herschlag lab for helpful discussions and comments on the manuscript. This work was supported by NIH grant GM49243 (to D.H.). R.R. was supported in part by the Janet M. Shamberger Fellowship Fund, and by an NIH postdoctoral fellowship.

References

Baldwin, R. L. (1995). The nature of protein folding pathways: the classical versus the new view. *J. Biomol. NMR*, **5**, 103-109.

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.

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.

Brion, P. & Westhof, E. (1997). Hierarchy and dynamics of RNA folding. *Annu. Rev. Biophys. Biomol. Struct.* **26**, 113-137.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R. & Doudna, J. A. (1996a). Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science*, **273**, 1678-1685.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. R. & Doudna, J. A. (1996b). RNA tertiary structure mediation by adenosine platforms. *Science*, **273**, 1696-1699.

Cate, J. H., Hanna, R. L. & Doudna, J. A. (1997). A magnesium ion core at the heart of a ribozyme domain. *Nature Struct. Biol.* **4**, 553-558.

Cech, T. R. & Herschlag, D. (1997). Group I ribozymes: substrate recognition, catalytic strategies, and comparative mechanistic analysis. In *Catalytic RNA* (Eckstein, F. & Lilley, D. M. J., eds), vol. 10, pp. 1-17, Springer-Verlag, Berlin.

Cech, T. R., Damberger, S. H. & Gutell, R. R. (1994). Representation of the secondary and tertiary structure of group I introns. *Nature Struct. Biol.* **1**, 273-280.

Celander, D. W. & Cech, T. R. (1991). Visualizing the higher order folding of a catalytic RNA molecule. *Science*, **251**, 401-407.

Dill, K. A. & Chan, H. S. (1997). From Levinthal to pathways to funnels. *Nature Struct. Biol.* **4**, 10-19.

Doherty, E. A. & Doudna, J. A. (1997). The P4-P6 domain directs higher order folding of the *Tetrahymena* ribozyme core. *Biochemistry*, **36**, 3159-3169.

Doherty, E. A., Herschlag, D. & Doudna, J. A. (1999). Assembly of an exceptionally stable RNA tertiary interface in a Group I ribozyme. *Biochemistry*, **38**, 2982-2990.

Doudna, J. A. & Cate, J. H. (1997). RNA structure: crystal clear? *Curr. Opin. Struct. Biol.* **7**, 310-316.

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.

Herschlag, D. & Cech, T. R. (1990). Catalysis of RNA cleavage by the *Tetrahymena thermophila* ribozyme. 1. Kinetic description of the reaction of an RNA substrate complementary to the active site. *Biochemistry*, **29**, 10159-10171.

Herschlag, D. & Khosla, M. (1994). Comparison of pH dependencies of the *Tetrahymena* ribozyme reactions with RNA 2′-substituted and phosphorothioate substrates reveals a rate-limiting conformational step. *Biochemistry*, **33**, 5291-5297.

Herschlag, D., Eckstein, F. & Cech, T. R. (1993). Contributions of 2′-hydroxyl groups of the RNA substrate to binding and catalysis by the *Tetrahymena* ribo-

zyme. An energetic picture of an active site composed of RNA. *Biochemistry,* **32**, 8299-8311.

Huang, G. S. & Oas, T. G. (1995). Submillisecond folding of monomeric lambda repressor. *Proc. Natl Acad. Sci. USA,* **92**, 6878-6882.

Jackson, S. E. & Fersht, A. R. (1991). Folding of chymotrypsin inhibitor 2.1. Evidence for a two-state transition. *Biochemistry,* **30**, 10428-10435.

Joyce, G. F. & Inoue, T. (1987). Structure of the catalytic core of the *Tetrahymena* ribozyme as indicated by reactive abbreviated forms of the molecule. *Nucl. Acids Res.* **15**, 9825-9840.

Joyce, G. F., van der Horst, G. & Inoue, T. (1989). Catalytic activity is retained in the *Tetrahymena* group I intron despite removal of the large extension of element P5. *Nucl. Acids Res.* **17**, 7879-7889.

Kiefhaber, T. (1995). Kinetic traps in lysozyme folding. *Proc. Natl Acad. Sci. USA,* **92**, 9029-9033.

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.

Murphy, F. L. & Cech, T. R. (1993). An independently folding domain of RNA tertiary structure within the *Tetrahymena* ribozyme. *Biochemistry,* **32**, 5291-5300.

Narlikar, G. J. & Herschlag, D. (1997). Mechanistic aspects of enzymatic catalysis: lessons from comparison of RNA and protein enzymes. *Annu. Rev. Biochem.* **66**, 19-59.

Pan, J. & Woodson, S. A. (1998). Folding intermediates of a self-splicing RNA: mispairing of the catalytic core. *J. Mol. Biol.* **280**, 597-609.

Pan, J., Thirumalai, D. & Woodson, S. A. (1997). Folding of RNA involves parallel pathways. *J. Mol. Biol.* **273**, 7-13.

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.

Pan, T., Fang, X. & Sosnick, T. (1999). Pathway modulation, circular permutation and rapid RNA folding under kinetic control. *J. Mol. Biol.* **286**, 721-731.

Radford, S. E., Dobson, C. M. & Evans, P. A. (1992). The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature,* **358**, 302-307.

Russell, R. & Herschlag, D. (1999). Specificity from steric restrictions in the guanosine binding pocket of a group I ribozyme. *RNA,* **5**, 158-166.

Schindler, T., Herrler, M., Marahiel, M. A. & Schmid, F. X. (1995). Extremely rapid protein folding in the absence of intermediates. *Nature Struct. Biol.* **2**, 663-673.

Sclavi, B., Woodson, S., Sullivan, M., Chance, M. R. & Brenowitz, M. (1997). Time-resolved synchrotron X-ray ''footprinting'', a new approach to the study of nucleic acid structure and function: application to protein-DNA interactions and RNA folding. *J. Mol. Biol.* **266**, 144-159.

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.

Strobel, S. A. & Doudna, J. A. (1997). RNA seeing double: close-packing of helices in RNA tertiary structure. *Trends Biochem. Sci.* **22**, 262-266.

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

Thirumalai, D. & Woodson, S. A. (1996). Kinetics of folding of proteins and RNA. *Acc. Chem. Res.* **29**, 433-439.

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.

Walstrum, S. A. & Uhlenbeck, O. C. (1990). The self-splicing RNA of *Tetrahymena* is trapped in a less active conformation by gel purification. *Biochemistry,* **29**, 10573-10576.

Zarrinkar, P. P. & Williamson, J. R. (1994). Kinetic intermediates in RNA folding. *Science,* **265**, 918-924.

Zarrinkar, P. P. & Williamson, J. R. (1996). The kinetic folding pathway of the *Tetrahymena* ribozyme reveals possible similarities between RNA and protein folding. *Nature Struct. Biol.* **3**, 432-438.

Zarrinkar, P. P., Wang, J. & Williamson, J. R. (1996). Slow folding kinetics of RNase P RNA. *RNA,* **2**, 564-573.

Zaug, A. J., Been, M. D. & Cech, T. R. (1986). The *Tetrahymena* ribozyme acts like an RNA restriction endonuclease. *Nature,* **324**, 429-433.

Zaug, A. J., Grosshans, C. A. & Cech, T. R. (1988). Sequence-specific endoribonuclease activity of the *Tetrahymena* ribozyme: enhanced cleavage of certain oligonucleotide substrates that form mismatched ribozyme-substrate complexes. *Biochemistry,* **27**, 8924-8931.

Edited by D. E. Draper

(Received 5 May 1999; received in revised form 12 July 1999; accepted 13 July 1999)