The rate-limiting step in the folding of a large ribozyme without kinetic traps

X.-W. Fang*, P. Thiyagarajan†, T. R. Sosnick*‡§, and T. Pan*§

*Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637; †Argonne National Laboratory, Argonne, IL 60439; and ‡Institute for Biophysical Dynamics, University of Chicago, Chicago, IL 60637

Communicated by S. Walter Englander, University of Pennsylvania School of Medicine, Swathmore, PA, May 14, 2002 (received for review March 1, 2002)

A fundamental question in RNA folding is the nature of the rate-limiting step. Folding of large RNAs often is trapped by the need to undo misfolded structures, which precludes the study of the other, potentially more interesting aspects in the rate-limiting step, such as conformational search, metal ion binding, and the role of productive intermediates. The catalytic domain of the *Bacillus subtilis* RNase P RNA folds without a kinetic trap, thereby providing an ideal system to elucidate these steps. We analyzed the folding kinetics by using fluorescence and absorbance spectroscopies, catalytic activity, and synchrotron small-angle x-ray scattering. Folding begins with the rapid formation of early intermediates wherein the majority of conformational search occurs, followed by the slower formation of subsequent intermediates. Before the rate-limiting step, more than 98% of the total structure has formed. The rate-limiting step is a small-scale structural rearrangement involving prebound metal ions.

Understanding the nature of the rate-limiting step in tertiary RNA folding is crucial to our ability to design, modify, and elucidate the evolutionary relationship of functional RNA structures. Early *in vitro* studies showed that tRNA often folds slowly because of the need to undo nonnative structures (1–3). Recent studies on RNA folding have focused on large ribozymes (4–24), such as the group I and II ribozymes and the bacterial RNase P RNA. The rate-limiting step in the folding of the *Tetrahymena* group I ribozyme is the disruption of an alternate helix, AltP3, which takes minutes at 37°C. Under certain conditions, AltP3 does not form, and the *Tetrahymena* group I ribozyme folds to the native state (N) in seconds or less (12). Folding of the ai5γ group II ribozyme is limited by need to form key long-range tertiary interactions (24). Folding of the *Bacillus subtilis* RNase P RNA is limited by the need to undo a nonnative structure that forms between its two domains (14).

The transitory, nonnative structure in the folding of the full-length *B. subtilis* P RNA is eliminated when both domains are separated (15). The catalytic domain (C domain, Fig. 1*A*) folds in less than 0.2 s at 37°C and 30 s at 10°C, more than 100 times faster than the folding of the full-length P RNA. Furthermore, the presence of the denaturant urea can decrease the folding rate of the C domain, whereas urea increases the folding rate of RNAs known to be limited by error-correction steps. In conjunction with a detailed analysis of the metal ion dependence of the folding kinetics, we concluded previously that the folding of C-domain is not limited by misfolding events.

In the absence of misfold/reorganization barriers, what process is rate-limiting? Local secondary structures such as hairpin-loops can rapidly form in microseconds in the absence of tertiary structure. The formation of long-range helices and loop–loop interactions may require an extensive conformational search and is likely to be much slower (21, 24). Alternatively, the limiting process may involve interactions between the RNA and the divalent metal ions required for the formation of the native structure. Metal ion binding sites are crucial for maintaining the tertiary architecture of large RNAs (25–30) and their formation is an event that cannot be avoided.

In this work, we demonstrate that the limiting step in C-domain folding occurs after nearly all conformational search. This conclusion is derived from chevron analysis of the metal ion series Mg^{2+}, Ca^{2+}, Sr^{2+}, and Ba^{2+} and the temperature dependence of the folding rate on these metal ions. Furthermore, this step involves small-scale consolidation of RNA structure around one or more prebound metal ions, even though the limiting step does not involve uptake of additional metal ions.

Materials and Methods

RNA Synthesis. The 5′-fluorescein labeled C-domain was prepared by *in vitro* transcription using standard methods in 14 mM $MgCl_2$, with 1 mM each ATP, CTP, and UTP, 0.5 mM GTP, and 0.75 mM 5′ fluorescein-G, and purified by denaturing gel electrophoresis. The unfolded state (U_{aq}) was obtained by heating the previously denatured C-domain in Me^{2+}-free Tris·HCl buffer (pH 8.1) at 85°C for 2 min, followed by incubation at room temperature for 3 min. The completely unfolded state (U_{urea}) was obtained by the addition of urea and EDTA to the unfolded state at a final concentration of 8 M and 1 mM, respectively.

Folding Kinetics Monitored by Fluorescence Spectroscopy (0.2–400 s). Folding was initiated on injecting concentrated $MeCl_2$ by a computer-interfaced Hamilton titrator into 1 ml of a solution of 0.02–0.03 μM C domain in Tris·HCl, pH 8.1. For unfolding kinetics, prefolded C domain was injected into 1 ml of buffer containing the appropriate amount of $MeCl_2$. Folding and unfolding kinetics were monitored by using $\lambda_{excitation} = 490 \pm 10$ nm and $\lambda_{emission} > 500$ nm. The dead-time of the experiment was 0.2 s.

Folding Kinetics Monitored by Absorbance Spectroscopy (0.002–1 s). Rapid mixing absorbance experiments used a Biologic SFM4 stopped-flow apparatus. C domain (2 μM) in 20 mM Tris·HCl (pH 8.1) with or without 8 M urea and 1 mM EDTA was diluted 10-fold with 20 mM Tris·HCl (pH 8.1) containing the appropriate amount of $MgCl_2$. Folding was monitored by the absorbance change at 267 nm.

Folding Kinetics Monitored by Synchrotron Small-Angle X-Ray Scattering (SAXS). SAXS experiments were carried out at the BESSRC ID-12 beam line of the Argonne National Laboratory (31). Folding of 0.3 mg/ml C-domain (3.8 μM) at 20 mM Tris·HCl, pH 8.1/0.3 mM $MgCl_2$ (I_{eq} state) was initiated on the addition of $MgCl_2$ to a final concentration of 10 mM. In comparison to the spectroscopic measurements, higher Mg^{2+} concentration was needed to fully populate I_{eq} for the SAXS studies because of the higher RNA concentration used (31). Data were acquired for 1 s each at 3, 6, 10, 25, 35, and 75 s after the initiation of folding.

Abbreviations: C domain, catalytic domain; N, native state; R_g, radius of gyration; SAXS, small-angle X-ray scattering; U, unfolded state.

§To whom reprint requests may be addressed. E-mail: trsosnic@midway.uchicago.edu or taopan@midway.uchicago.edu.

 www.pnas.org/cgi/doi/10.1073/pnas.142288399

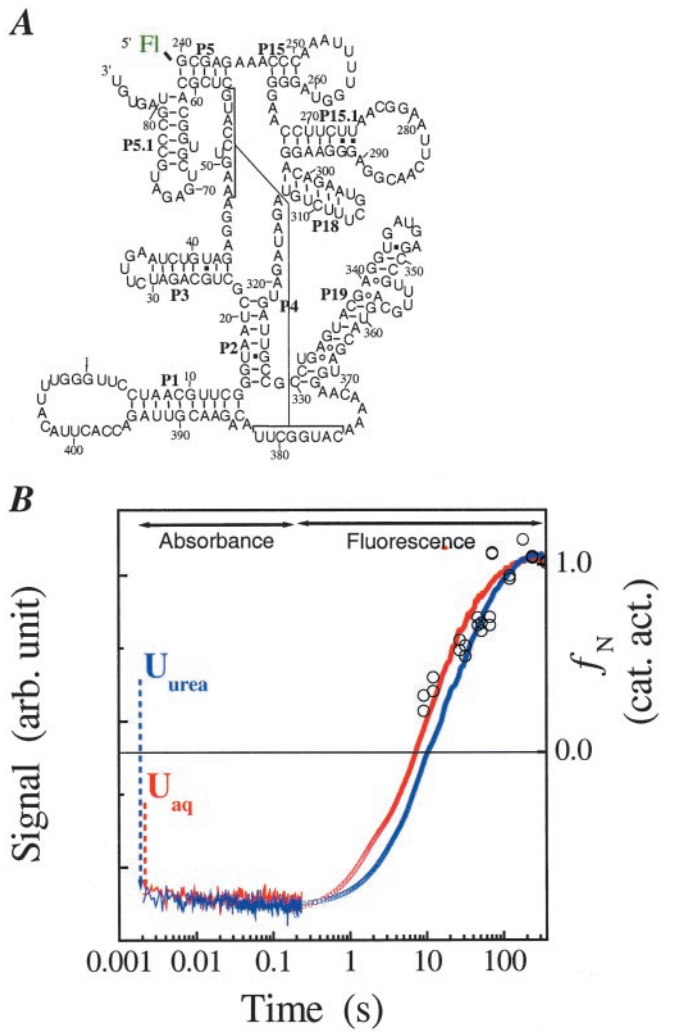


Fig. 1. (*A*) The *B. subtilis* C-domain (47). A fluorescein is covalently attached to the 5′ end that allows monitoring the folding kinetics by fluorescence spectroscopy. (*B*) Folding kinetics at 10°C starting from either the U_{aq} state (20 mM Tris·HCl, no Me^{2+}) or the U_{urea} state (20 mM Tris·HCl, 8 M urea, 1 mM EDTA, no Me^{2+}). Folding is monitored by absorbance from 0.002 to 0.2 s, and the dead-time change in signal is indicated with dashed arrows. Folding is monitored by fluorescence between 0.2 to 360 s. The fluorescence signal is fit with a two-exponential function. The slow phase has the same rate as that obtained by catalytic activity (open circles, right axis) and directly reports the rate-limiting step. The fast phase corresponds to the signal change of the I_{eq}-to-I_1^k transition and is ≈6 times faster than the slow phase.

Chevron Analysis. The dependence of relaxation rates of the I_{eq}-to-N transition on ion concentration (15) was fit assuming two sequentially populated kinetic intermediates, I_1^k and I_2^k, according to

$$k_{obs} = \frac{[Me^{2+}]^{n_1}}{[Me^{2+}]^{n_1} + (K_1)^{n_1}} \times k_f + \frac{(K_2)^{n_2}}{[Me^{2+}]^{n_2} + (K_2)^{n_2}} \times k_u, \qquad [1]$$

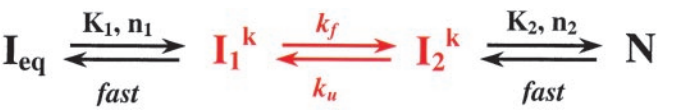
where K_1, n_1 and K_2, n_2 are the Me^{2+} concentration at the midpoint and the Hill constant of the I_{eq}-to-I_1^k and I_2^k-to-N transition, respectively; k_f and k_u are the forward and backward rates of the I_1^k-to-I_2^k transition. Both the I_{eq}-to-I_1^k and I_2^k-to-N transitions are fast enough to be in equilibrium relative to the limiting I_1^k-to-I_2^k transition.

Results

Folding to the Equilibrium Intermediate, I_{eq}. As described (31, 32), the equilibrium folding pathway of the C domain consists of two transitions: U-to-I_{eq} and I_{eq}-to-N. The unfolded state contains varying amount of structure, depending on the solution condition and temperature. Previous SAXS measurements at 37°C show that the radius of gyration (R_g) of the unfolded state is 80 Å in the aqueous solvent (U_{aq}), whereas in the presence of 8 M urea and 1 mM EDTA (U_{urea}), R_g increases to 180 Å (31). The formation of I_{eq} results in a dramatic collapse as indicated by the R_g change from 80–180 Å to 46 Å (native value is 38 Å). The intermediate is highly structured, and forms at ≈5 times lower Mg^{2+} concentration needed to fold the native structure (15, 31–33). The I_{eq} species contains the native amount of secondary structure as well as some tertiary structure, according to absorbance measurements, an analysis of urea sensitivity, and a comparative folding study with a thermophilic homologue (34). Regardless of the solution condition, however, I_{eq} has the same size, shape, and structure.

Folding to I_{eq} starting from U_{aq} and U_{urea} at 10°C (Fig. 1*B*) and 37°C (data not shown) is investigated by using stopped-flow absorbance spectroscopy. The UV absorbance significantly increases on the addition of urea (different starting level of absorbance of unfolded state at zero time in Fig. 1*B*), consistent with a loss of residual structure. In all measurements, I_{eq} is formed in less than 1 ms. Essentially all of the change in UV absorbance for the entire folding reaction to the native state occurs in this early step. Hence, an extensive conformational search occurs very quickly and results in the formation of this highly structured and compact intermediate.

Folding from I_{eq} to the Native Structure. The I_{eq}-to-N transition is at least four orders of magnitude slower than the U-to-I_{eq} transition (Fig. 1*B*). Our previous study showed that C domain folds at the same rate whether folding was initiated from either U_{aq} or I_{eq}, indicating that the I_{eq}-to-N transition is always rate-limiting. We also demonstrated that the I_{eq}-to-N transition includes the transient formation of two kinetic intermediates, I_1^k and I_2^k, and the rate-limiting step is the conversion between these two species:



Scheme 1.

In this scheme, K_1, K_2 and n_1, n_2 are the Mg^{2+} midpoints and the Hill coefficients that define the relative populations of the Mg^{2+}-dependent I_{eq}-to-I_1^k and the I_2^k-to-N transitions, respectively; k_f and k_u are the folding and unfolding rates, respectively, of the Mg^{2+}-independent I_1^k-to-I_2^k transition. Both the I_{eq}-to-I_1^k and I_2^k-to-N transitions are in fast equilibrium relative to the I_1^k-to-I_2^k transition.

This kinetic folding pathway was deduced from the "Mg^{2+} chevron" (15), which reports the dependence of the folding and unfolding rates on the Mg^{2+} concentration, as measured by fluorescence spectroscopy (Fig. 2*A*). For C-domain folding, the Mg^{2+} dependence of the observed rate, k_{obs}, produces a V or "chevron" shape. The vertex, where k_{obs} is the slowest, generally represents the condition when the folding rate equals to the unfolding rate. Folding is faster than unfolding when Mg^{2+} concentration is above K_{Mg} (right side of chevron), whereas unfolding is faster than folding when Mg^{2+} concentration is below K_{Mg} (left side of chevron).

On the folding side, k_{obs} becomes saturated when the Mg^{2+} concentration is above 3 mM (Fig. 2*A*), indicating a Mg^{2+}-independent rate-limiting process in the I_{eq}-to-N transition

BIOCHEMISTRY

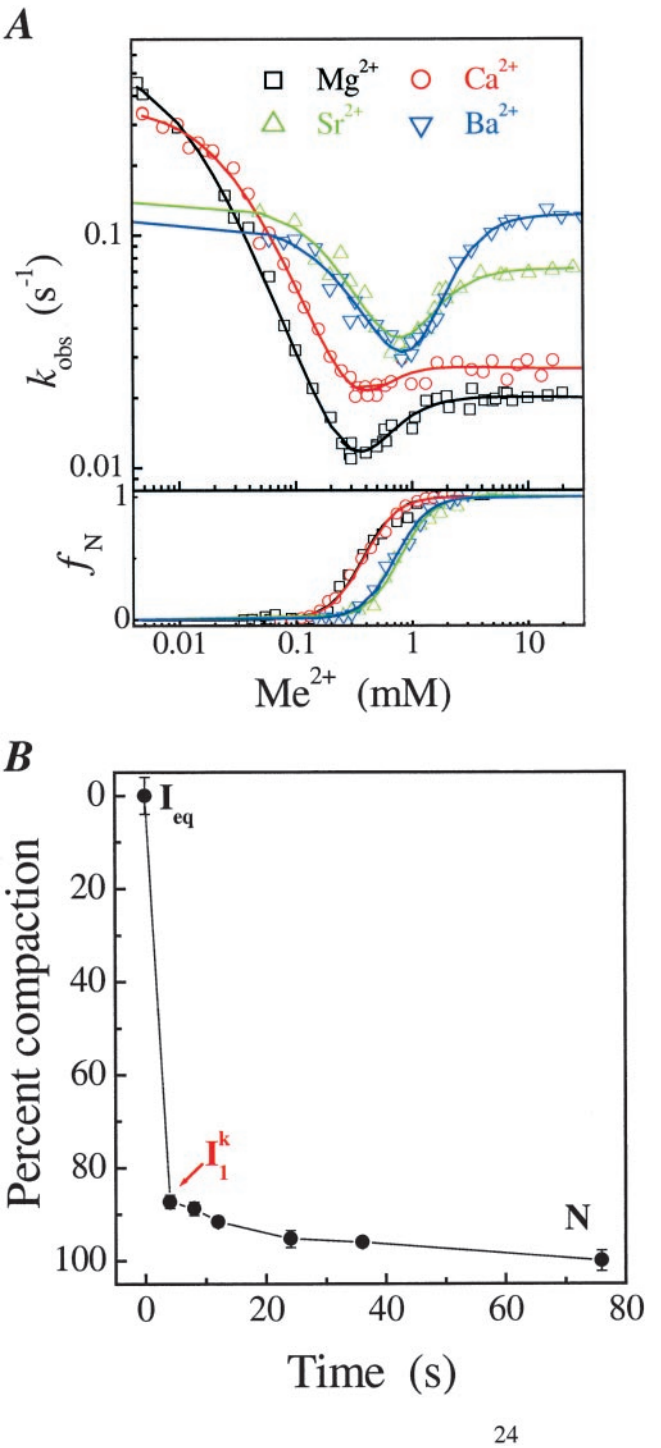


Fig. 2. (*A*) Metal-ion chevron analysis using fluorescence spectroscopy (*Upper*). In the chevron analysis, rate saturation at the high or low [Me^{2+}] indicates the formation of a folding and unfolding intermediate, I_1^k and I_2^k, respectively. [Me^{2+}] at the vertex of the chevron is approximately the same as the midpoint of the equilibrium I_{eq}-to-N transition measured independently (*Lower*). (*B*) Folding kinetics of C-domain from I_{eq} at 10 mM $MgCl_2$, 10°C, monitored by SAXS. The percentage of compaction is calculated from radius gyration. This result indicates that I_1^k is highly compact and forms quickly.

(Scheme 1). This "rollover" in a chevron plot occurs because at high [Mg^{2+}], the kinetic intermediate I_1^k is fully populated, as opposed to folding at lower [Mg^{2+}] where the k_{obs} is the product of k_f and the population fraction of I_1^k. On the unfolding side, k_{obs} becomes saturated at [Mg^{2+}] below 0.005 mM because the kinetic intermediate I_2^k is fully populated at very low [Mg^{2+}], as opposed to higher [Mg^{2+}] where the k_{obs} is the product of k_u and the population fraction of I_2^k.

Because k_f and k_u, the interconversion rates of the I_1^k-to-I_2^k transition, do not depend on the Mg^{2+} concentration, the rate-limiting step in C-domain folding does not involve a change in the net number of metal ions bound. Furthermore, the interconversion between these two intermediates is accompanied by minimal surface burial, as indicated in our previous study by the identical folding rate at high Mg^{2+} concentration in the absence and presence of urea (15).

SAXS provides direct evidence on the transient accumulation of I_1^k (Fig. 2*B*). Beginning from I_{eq}, folding occurs in two discrete steps according to SAXS. On the addition of 10 mM Mg^{2+}, C domain collapses rapidly to I_1^k. Subsequent acquisition of the native state occurs >5-fold slower. The I_{eq}-to-I_1^k transition accounts for ≈85% of the total compaction between I_{eq} and N.

Combining the current results with our previous studies, we conclude that before the limiting step, C domain forms nearly all of the native structure. The high degree of compaction and structure formation is reflected in the SAXS data as well as in the surface burial measured previously by urea dependence (15). The urea *m* value, a valuable diagnostic of surface burial, for the I_1^k-to-N transition is only ≈0.2 kcal·mol^{-1}·M^{-1}, whereas the total *m* value for the U_{urea}-to-N transition is estimated to be ≈9 kcal·mol^{-1}·M^{-1} (35). Similarly, the difference in R_g between I_1^k and N is only ≈1 Å, whereas the difference between U_{urea} and N is ≈140 Å (31). Therefore, more than 98% of the total surface burial and compaction has taken place on the formation of I_1^k, relative to the fully unfolded state, U_{urea}. Furthermore, the folding rate of the limiting I_1^k-to-I_2^k transition is insensitive to urea, indicating that most of the remaining 2% of the structure forms after the limiting step, in the I_2^k-to-N transition (15). Therefore, the rate-limiting step is a very small-scale event that involves only a minimal change in the RNA structure.

Metal Ion Dependence of the Limiting Step. To further reveal the molecular origin of the rate-limiting step, we carried out metal ion chevron analysis using four group IIA ions, Mg^{2+}, Ca^{2+}, Sr^{2+}, and Ba^{2+} (Figs. 2*A* and 3, Table 1). C domain folds to its native structure in the presence of each ion as confirmed by catalytic activity (data not shown). The chevron plots for all four metal ions have a similar characteristic shape and can be analyzed according to the same kinetic scheme, I_{eq}-to-I_1^k-to-I_2^k-to-N. The chevron analysis for all metal ions yields the same Hill constant for the I_{eq}-to-I_1^k transition ($n_1 \approx 2$, folding side) and for the I_2^k-to-N transition ($n_2 \approx 1$, unfolding side). Folding and unfolding rates (k_f and k_u) saturate at high and low metal ion concentration, respectively. This behavior demonstrates that folding induced by all four metal ions goes through the same two kinetic intermediates (I_1^k and I_2^k) and has the same rate-limiting I_1^k-to-I_2^k transition.

The Hill constant and the midpoint of the equilibrium I_{eq}-to-N transition can be determined from the kinetic parameters obtained from the chevron analysis. These kinetically determined values agree well with the independently measured equilibrium parameters (Fig. 2*A Lower*, Table 1). Hence, the free energy and the metal-ion binding of the equilibrium I_{eq}-to-N transition can be accounted for in the same, minimal kinetic scheme for each of the four metal ions.

For all metal ions, both k_f and k_u saturate at high and low ion concentrations, respectively. This result confirms that the rate-limiting I_1^k-to-I_2^k step does not involve bimolecular binding of a free metal ion to the RNA. Rather, the limiting step is a unimolecular event. On the other hand, k_f and k_u do depend on the particular divalent metal ion used to fold the RNA. Therefore, the rate-limiting step does involve at least one specific metal ion. These metal ions are already bound in the I_1^k intermediate. Because the

Table 1. Folding parameters of C-domain

Metal ion	Mg^{2+}	Ca^{2+}	Sr^{2+}	Ba^{2+}
k_f, s^{-1}	0.021 ± 0.002	0.027 ± 0.003	0.071 ± 0.006	0.12 ± 0.01
k_u, s^{-1}	0.55 ± 0.01	0.36 ± 0.02	0.14 ± 0.02	0.11 ± 0.01
$-RT \ln (k_f/k_u)$, kcal/mol	1.85 ± 0.05	1.47 ± 0.07	0.38 ± 0.09	−0.05 ± 0.08
K_1, mM	0.5 ± 0.1	0.5 ± 0.1	1.4 ± 0.2	2.4 ± 0.2
n_1	2.1 ± 0.3	2.3 ± 1.2	2.2 ± 0.5	2.2 ± 0.3
K_2, mM	0.011 ± 0.004	0.028 ± 0.002	0.23 ± 0.04	0.27 ± 0.05
n_2	1.3 ± 0.2	1.3 ± 0.1	1.4 ± 0.3	1.3 ± 0.3
K_{Me}, mM*	0.40 ± 0.02	0.41 ± 0.03	0.79 ± 0.03	0.72 ± 0.03
$\Delta H^{\ddagger}$, kcal/mol	23.6 ± 1.0	21.2 ± 1.8	18.4 ± 1.3	11.4 ± 1.1

All measurements are in 20 mM Tris·HCl (pH 8.1), 10°C, except for $\Delta H^{\ddagger}$ measurements, which are conducted over a temperature range of 2–30°C.

*This value is the midpoint of the equilibrium I_{eq}-to-N transition and is also approximately the vertex of the chevron trace. R, gas constant; T, temperature.

rate-limiting step involves a metal ion, but is accompanied by little change in the RNA structure, we conclude that the rate-limiting step is a small-scale conformational change, or structural consolidation involving prebound metal ions.

The metal ion dependence of k_f and k_u has a linear dependence on the reciprocal of the ionic radius. As the ionic radius increases ($Mg^{2+} < Ca^{2+} < Sr^{2+} < Ba^{2+}$), the folding rate increases, whereas the unfolding rate decreases by almost equal measures. Therefore, the consolidation step is more destabilizing for the smaller metal ion [$\Delta G = -RT \ln (k_f/k_u)$, Table 1]. For the smallest ion, Mg^{2+}, the consolidation requires 1.9 kcal·mol^{-1} of free energy, whereas for the largest ion, Ba^{2+}, the consolidation is energetically neutral.

The binding affinity of a metal ion to the RNA (K_1, K_2, and K_{Me}, Table 1) is stronger for the ion with smaller radius. In the initial bimolecular binding event, smaller ions are preferred. But, in the subsequent unimolecular rate-limiting step, larger ions are favored. In short, the stronger the metal ion binds to the RNA, the slower is the folding rate and the faster is the unfolding rate.

Further indication that prebound metal ions are involved in the rate-limiting step is derived from the temperature dependence of the folding rate, k_f (Fig. 3*A*). At a first glance, the activation enthalpy appears to correlate with the binding affinity of the metal ion to the RNA—the stronger the binding affinity, the higher is the activation enthalpy. A correlation also exists between the activation enthalpies and those of dehydration for these metal ions (ref. 36, Fig. 3*B*), potentially indicative of a change in the hydration state of the metal ions at the limiting step.

Discussion

How a large RNA folds into its biologically active structure has been intensely studied in the last decade. Treiber and Williamson (37) have proposed that the general RNA-folding mechanisms can be summarized into three classes. The first class includes fast folding RNAs with sequentially forming and productive intermediates. The second class includes slow folding RNAs with kinetic traps. The third class includes slow folding RNAs that result in unstable tertiary structures, which are subsequently captured by another molecule, e.g., a protein (21). Recently, Pyle and coworkers (24) proposed a fourth class wherein folding is trap-free, but still slow because of a conformational search for a key tertiary interaction early in the pathway. The *B. subtilis* C domain belongs to the first class of RNAs that fold with sequentially populated and productive intermediates. Because the C domain folds without kinetic traps, several important aspects of the RNA folding problem can be directly approached. In this work, we examined the speed and the extent of conformational search, the role of the productive intermediates, and the nature of the rate-limiting step (Fig. 4).

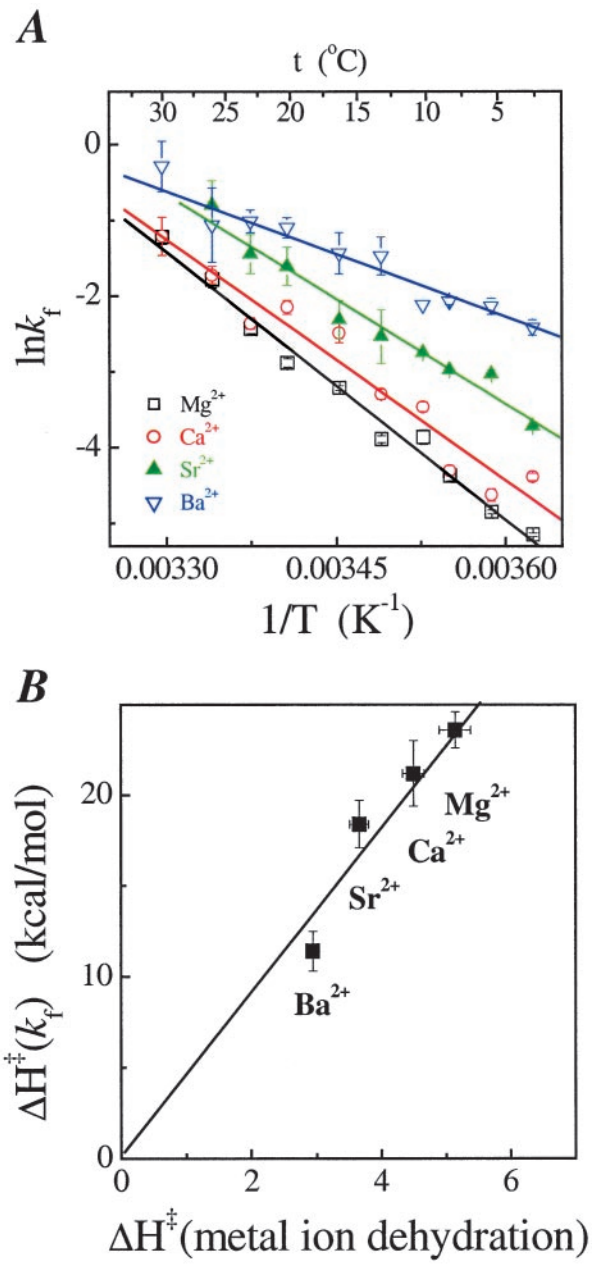
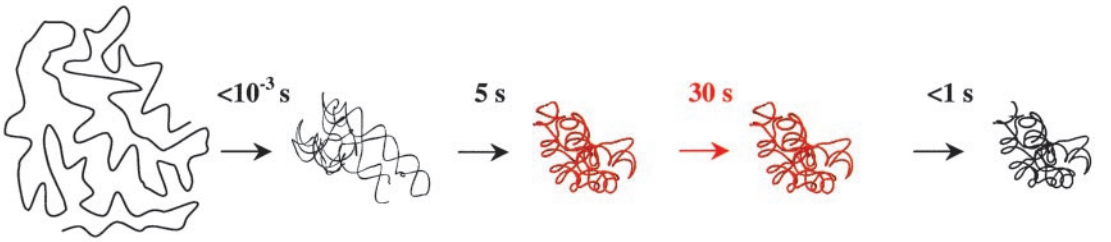


Fig. 3. (*A*) Temperature dependence of the folding rate. (*B*) Correlation between the activation enthalpy for folding and the dehydration of the metal ions measured by NMR (36).

BIOCHEMISTRY



	U_{urea}	I_{eq}	I_1^k	I_2^k	N
% compaction (R_g)	0%	94%	99%		100%
%m-value change (est.)	0%	89%	98%		100%

Fig. 4. Summary of C-domain folding. The rate-limiting step is shown in red. The folding half-lives at 10°C are shown, in seconds, above the arrows. The starting point is the U_{urea} state (C domain in 8 M urea and 1 mM EDTA), which has little RNA structure. I_{eq} contains the native amount of secondary structure and some tertiary structure. I_1^k contains >98% of the native structure. In less than 1 ms, the majority of the conformational search occurs with the formation of I_{eq}. All folding half-lives are measured in this work. Percent compaction is calculated from the data shown in Fig. 2*B* and from our previous work (31). Percent *m*-value change is calculated from published data (15).

In C-domain folding, nearly all conformational search occurs before the rate-limiting step. Regardless of the residual structure, the addition of sufficient amount of Mg^{2+} allows the C-domain to fold into the first well-defined intermediate, I_{eq}, in less than 1 ms. When starting from the fully denatured state, U_{urea}, formation of I_{eq} compacts the RNA to ≈90% of the native dimension. The subsequent formation of the second intermediate, I_1^k, further compacts the RNA to ≈98% of the native dimension. In each step, the large-scale compaction is accompanied by an extensive conformational search.

The present results, along with those from the group I ribozyme (38), indicate that when the RNA is permitted to overcome the electrostatic repulsion of its negatively charged backbone, it can compact extremely rapidly. Folding of the full-length RNase P RNA also involves very rapid compaction from the unfolded state to an intermediate state at the same time scale as in the folding of C domain (8). In the case of the full-length P RNA, however, a nonnative structure forms in the intermediate state and this significant nonnative structure has to be disrupted for the folding to proceed. The primary reason for the rapid, trap-free folding of the C domain to the native state is likely the early conformational search resulting in the formation of correct structures in I_{eq} and in I_1^k.

Three intermediates can be clearly identified in the folding kinetics of C-domain carried out in the four different metal ions, Mg^{2+}, Ca^{2+}, Sr^{2+}, and Ba^{2+}. The first intermediate, I_{eq}, is the same as the thermodynamic intermediate and it significantly populates at low Me^{2+} concentrations. The second and third intermediates, I_1^k and I_2^k, only populate transiently, with a dwell time of ≈30 s for I_1^k and ≈1 s for I_2^k when folding in Mg^{2+}.

All three intermediates appear to be productive in folding. We have not been able to find conditions that permit the C domain to bypass any of them, and an increase in their stability enhances relaxation rates. Because the I_{eq} and the native state have the same UV absorbance, we suggested previously that all secondary structures of the C domain are already present in I_{eq} (32). On the basis of comparative studies with a thermophilic homologue (34), we also suggested that some tertiary structures are present in I_{eq} as well. Compared with I_{eq}, I_1^k clearly has more structure and binds more metal ions (the Hill coefficient of the I_{eq}-to-I_1^k transition is ≈2). At this time, we have little information on the precise structure of these intermediates.

The rate-limiting step in C-domain folding is the conversion of two intermediates, from I_1^k to I_2^k. We have shown that this step has no dependence on the Mg^{2+} and urea concentration (15). We demonstrate here that this step is always limiting when folding is carried out in other divalent metal ions. More importantly, the metal-ion chevron analysis demonstrates that the limiting step still involves metal ions that are prebound in I_1^k. The hydration properties of these metal ions may play a role as well. The simplest interpretation we can envision is that the limiting step represents a small-scale consolidation of RNA structure around prebound metal ions. Specifically bound divalent metal ions are present in the structure of all large RNAs. These metal ions are located in sites of high electrostatic potential, which stabilize interactions between different structural motifs, and hence, are crucial for maintaining the tertiary architecture (39–41). It is likely that the formation of one or more of these binding sites is rate-limiting, although concrete experimental evidence is still lacking.

Models for Consolidation. The dependence of the rate-limiting step on the size of the prebound metal ion suggests that structural changes occur in the rate-limiting step. There is a trend in the folding and unfolding rate regarding the four group IIA metal ions—the higher is the binding affinity of the metal ion to the RNA, the slower is the folding rate and the faster is the unfolding rate. Several models involving specific metal ion binding can be proposed to explain this trend.

(*i*) An interaction between an RNA ligand and metal ion is broken to allow a nucleotide base to flip, for example, and then the same or a different metal ion–RNA interaction forms. Allowing the same interaction to form would predict an increase in folding and unfolding rates with larger metal ions. Although increased folding rates are observed, the opposite trend is observed for the unfolding rate. Hence, the metal ion–RNA interaction formed after the rate-limiting step has to be different from that formed before the rate-limiting step, and be weaker for the smaller ions.

(*ii*) An RNA-RNA structural rearrangement, e.g., switching of a sugar pucker, occurs without the loss of RNA-metal ion interactions. This structural rearrangement would be retarded because of the increased stabilization of the I_1^k state through tighter binding with smaller ions. Again, for this model to be consistent with the unfolding rates, the RNA structure formed after the rate-limiting step has to have lower affinity for smaller metal ions.

(*iii*) The rate-limiting step involves the release of a prebound metal ion at a high affinity site in I_1^k and rebinding at a low affinity site in I_2^k. This model can rationalize the observed trend

for the folding rate, as a higher affinity ion would be released more slowly from the high-affinity site. As before, in order for this model to be consistent with the unfolding rates, the RNA structure formed after the rate-limiting step has to exhibit a lower affinity for smaller metal ions.

(*iv*) Consolidation of the RNA around smaller ions is more difficult because of steric and/or electrostatic crowding of the negatively charged RNA backbone. A more detailed model, based on the crystallographic observation of multiple inner-sphere metal-phosphate coordination, envisions that the I_1^k to I_2^k transition involves an increase in the number of RNA ligands around the prebound divalent metal ions. This model rationalizes the slower k_f and the faster k_u, as well as the free energy difference between I_2^k and I_1^k. Within an extremely constrained environment, a smaller ion shuns the bulkier and charged RNA ligand. The overall binding affinity for the particular metal ions in the rate-limiting step, however, may still be higher for the smaller ions. This increase in affinity is caused by the more favorable metal ion binding for the smaller ions on the formation of I_1^k (K_1 values in Table 1). The overall binding affinity would represent the combined free energy of the favorable binding step and the unfavorable, increased electrostatic and steric crowding after the rate-limiting step.

C-domain folding occurs many orders of magnitude slower than what one would expect for a simple ligand-exchange reaction. This discrepancy may be explained in part by the state of the metal ion that is already bound to the negatively charged RNA ligand, as well as a term related to RNA conformational change. The prebound metal ion probably is much less electrophilic. Thus, the rate of the ligand exchange on the C domain should be significantly slower than the rate of ligand exchange on a metal-ion free in solution. In the folding of our RNA, the conformational component must be consistent with the negligible amount of surface burial and compaction in the rate-limiting step. This component could represent the frequency of productive contacts in a small-scale conformational search.

Conformational Search in Protein and RNA Folding. The folding of protein and RNA poses a considerable conformational search problem. However, the mode of search and the limiting steps have significant differences. In protein folding, bigger search units (e.g., α-helices) may form before the transition state, but these units are not stable, and the same stretch of amino acid residues may transiently adopt different conformations during the search. Such structural flexibility of protein secondary structure entails that the intrinsic rate-limiting step is a global conformational search involving all residues. The key element is the formation of the native topology of the backbone (42, 43), which often is accompanied by the formation of a hydrophobic core with long-range contacts along with the hydrogen-bond network needed to support this association (44).

In RNA folding, secondary structure often is independently stable and can form on the microsecond time-scale. A tertiary structure involving a three-way helical junction can form in a millisecond (45). Error-free folding can occur if the formation of stable nonnative structure is avoided, which depends on the sequence (21, 24, 37, 46) and connectivity (14). The rapid formation of stable secondary structure allows the search process to be conducted among preformed structural units whose number is significantly smaller than the total number of nucleotide residues. These units in turn assemble into a tertiary structure through nonlocal contacts mediated by loops, packed helices, and bridging metal ions. Two scenarios have been observed in tertiary RNA folding. In some cases, conformational search is rate-limiting because of the difficulty of forming key tertiary interactions (24). For the C domain, structural consolidation around prebound metal ions is rate-limiting. This event occurs after nearly the entire conformational search, and serves to lock down the tertiary structure at a late stage in folding.

We thank Drs. P. Cluzel, J. Piccirilli, J. R. Williamson, D. Herschlag, S. Chu, and J. Staley and the reviewers for helpful discussions and comments on the manuscript.This work was supported by National Institutes of Health Grant GM57880 (to T.P. and T.R.S.), and the U. S. Department of Energy, BES-Materials Science under Contract W-31-109-ENG-38.

1. Cole, P. E. & Crothers, D. M. (1972) *Biochemistry* **11,** 4368–4374.
2. Webb, P. K. & Fresco, J. R. (1973) *J. Mol. Biol.* **74,** 387–402.
3. Lynch, D. C. & Schimmel, P. R. (1974) *Biochemistry* **13,** 1841–1852.
4. Zarrinkar, P. P. & Williamson, J. R. (1994) *Science* **265,** 918–924.
5. Banerjee, A. R. & Turner, D. H. (1995) *Biochemistry* **34,** 6504–6512.
6. Zarrinkar, P. P. & Williamson, J. R. (1996) *Nat. Struct. Biol.* **3,** 432–438.
7. Zarrinkar, P. P., Wang, J. & Williamson, J. R. (1996) *RNA* **2,** 564–573.
8. Pan, T. & Sosnick, T. R. (1997) *Nat. Struct. Biol.* **4,** 931–938.
9. Pan, J., Thirumalai, D. & Woodson, S. A. (1997) *J. Mol. Biol.* **273,** 7–13.
10. Treiber, D. K., Rook, M. S., Zarrinkar, P. P. & Williamson, J. R. (1998) *Science* **279,** 1943–1946.
11. Sclavi, B., Sullivan, M., Chance, M. R., Brenowitz, M. & Woodson, S. A. (1998) *Science* **279,** 1940–1943.
12. Pan, J. & Woodson, S. A. (1998) *J. Mol. Biol.* **280,** 597–609.
13. Rook, M. S., Treiber, D. K. & Williamson, J. R. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 12471–12476.
14. Pan, T., Fang, X. & Sosnick, T. R. (1999) *J. Mol. Biol.* **286,** 721–731.
15. Fang, X., Pan, T. & Sosnick, T. R. (1999) *Nat. Struct. Biol.* **6,** 1091–1095.
16. Silverman, S. K., Deras, M. L., Woodson, S. A., Scaringe, S. A. & Cech, T. R. (2000) *Biochemistry* **39,** 12465–12475.
17. Deras, M. L., Brenowitz, M., Ralston, C. Y., Chance, M. R. & Woodson, S. A. (2000) *Biochemistry* **39,** 10975–10985.
18. Pan, J., Deras, M. L. & Woodson, S. A. (2000) *J. Mol. Biol.* **296,** 133–144.
19. Zhuang, X., Bartley, L. E., Babcock, H. P., Russell, R., Ha, T., Herschlag, D. & Chu, S. (2000) *Science* **288,** 2048–2051.
20. Kent, O., Chaulk, S. G. & MacMillan, A. M. (2000) *J. Mol. Biol.* **304,** 699–705.
21. Buchmueller, K. L., Webb, A. E., Richardson, D. A. & Weeks, K. M. (2000) *Nat. Struct. Biol.* **7,** 362–366.
22. Heilman-Miller, S. L., Pan, J., Thirumalai, D. & Woodson, S. A. (2001) *J. Mol. Biol.* **309,** 57–68.
23. Russell, R. & Herschlag, D. (2001) *J. Mol. Biol.* **308,** 839–851.
24. Swisher, J. F., Su, L. J., Brenowitz, M., Anderson, V. E. & Pyle, A. M. (2002) *J. Mol. Biol.* **315,** 297–310.
25. Cate, J. H., Hanna, R. L. & Doudna, J. A. (1997) *Nat. Struct. Biol.* **4,** 553–558.
26. Ferre-D'Amare, A. R., Zhou, K. & Doudna, J. A. (1998) *Nature (London)* **395,** 567–574.
27. Shi, H. & Moore, P. B. (2000) *RNA* **6,** 1091–1105.
28. Jovine, L., Djordjevic, S. & Rhodes, D. (2000) *J. Mol. Biol.* **301,** 401–414.
29. Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. (2000) *Science* **289,** 905–920.
30. Rupert, P. B. & Ferre-D'Amare, A. R. (2001) *Nature (London)* **410,** 780–786.
31. Fang, X., Littrell, K., Yang, X., Henderson, S. J., Siefert, S., Thiyagarajan, P., Pan, T. & Sosnick, T. R. (2000) *Biochemistry* **39,** 11107–11113.
32. Fang, X., Pan, T. & Sosnick, T. R. (1999) *Biochemistry* **38,** 16840–16846.
33. Seifert, S., Winans, R. E., Tiede, D. M. & Thiyagarajan, P. (2000) *J. Appl. Crystallogr.* **33,** 782–784.
34. Fang, X. W., Golden, B. L., Littrell, K., Shelton, V., Thiyagarajan, P., Pan, T. & Sosnick, T. R. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 4355–4360.
35. Shelton, V. M., Sosnick, T. R. & Pan, T. (1999) *Biochemistry* **38,** 16831–16839.
36. Helm, L. & Hertz, H. G. (1981) *Z. Physik. Chem.* **127,** 23–44.
37. Treiber, D. K. & Williamson, J. R. (2001) *Curr. Opin. Struct. Biol.* **11,** 309–314.
38. Russell, R., Millett, I. S., Tate, M. W., Kwok, L. W., Nakatani, B., Gruner, S. M., Mochrie, S. G., Pande, V., Doniach, S., Herschlag, D. & Pollack, L. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 4266–4271.
39. Hermann, T. & Westhof, E. (1998) *Structure (London)* **6,** 1303–1314.
40. Chin, K., Sharp, K. A., Honig, B. & Pyle, A. M. (1999) *Nat. Struct. Biol.* **6,** 1055–1061.
41. Misra, V. K. & Draper, D. E. (2000) *J. Mol. Biol.* **299,** 813–825.
42. Sosnick, T. R., Mayne, L. & Englander, S. W. (1996) *Proteins* **24,** 413–426.
43. Plaxco, K. W., Simons, K. T. & Baker, D. (1998) *J. Mol. Biol.* **277,** 985–994.
44. Krantz, B. A., Srivastava, A. K., Nauli, S., Baker, D., Sauer, R. T. & Sosnick, T. R. (2002) *Nat. Struct. Biol.*, in press.
45. Kim, H. D., Nienhaus, G. U., Ha, T., Orr, J. W., Williamson, J. R. & Chu, S. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 4284–4289.
46. Woodson, S. A. (2000) *Cell Mol. Life Sci.* **57,** 796–808.
47. Brown, J. W., Nolan, J. M., Haas, E. S., Rubio, M. A., Major, F. & Pace, N. R. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 3001–3006.