6674–6687 *Nucleic Acids Research, 2005, Vol. 33, No. 21*
doi:10.1093/nar/gki973

An obligate intermediate along the slow folding pathway of a group II intron ribozyme

Linhui Julie Su[1], Christina Waldsich[1] and Anna Marie Pyle[1,2,*]

[1]Department of Molecular Biophysics and Biochemistry and [2]Howard Hughes Medical Institute, 266 Whitney Avenue, Box 208114, Yale University, New Haven, CT 06520, USA

Received August 31, 2005; Revised and Accepted November 4, 2005

ABSTRACT

Most RNA molecules collapse rapidly and reach the native state through a pathway that contains numerous traps and unproductive intermediates. The D135 group II intron ribozyme is unusual in that it can fold slowly and directly to the native state, despite its large size and structural complexity. Here we use hydroxyl radical footprinting and native gel analysis to monitor the timescale of tertiary structure collapse and to detect the presence of obligate intermediates along the folding pathway of D135. We find that structural collapse and native folding of Domain 1 precede assembly of the entire ribozyme, indicating that D1 contains an on-pathway intermediate to folding of the D135 ribozyme. Subsequent docking of Domains 3 and 5, for which D1 provides a preorganized scaffold, appears to be very fast and independent of one another. In contrast to other RNAs, the D135 ribozyme undergoes slow tertiary collapse to a compacted state, with a rate constant that is also limited by the formation D1. These findings provide a new paradigm for RNA folding and they underscore the diversity of RNA biophysical behaviors.

INTRODUCTION

RNA molecules navigate a variety of different folding pathways to the functional, native state. In the most common pathway, an RNA molecule traverses a rough energetic landscape that contains numerous opportunities for the molecule to misfold into stable, nonfunctional conformations or to become kinetically trapped in native interactions (1,2). The structural rearrangement of these kinetically trapped species is often the rate-limiting step in the folding pathway of these RNA. However, certain RNA molecules appear to traverse a less treacherous landscape, folding directly to the native state without adopting undesirable conformations from which it is difficult to emerge (3). In several of these cases, folding proceeds through a series of on-pathway, or obligate, intermediates (4–7). These productive intermediates have been reported to form very rapidly (5,6); however, the structural configurations of most of those folding intermediates and their role in folding remain enigmatic.

RNA folding pathways are not rigid paradigms, as it is becoming increasingly evident that *in vitro* folding pathways are strongly affected by ionic conditions, temperature, proteins and even the procedures used in handling the RNA (8–12). This makes comparisons between studies and paradigms difficult, but it also opens up new possibilities for studying RNA molecules and their transitions to the native state. One group of studies focuses exclusively on tertiary folding of large RNAs. In these cases, the RNA is preincubated in monovalent cations and allowed to form secondary structure prior to the addition of magnesium ions (8,13–16). This allows one to focus exclusively on tertiary structure transitions and to investigate pathways from a defined structural starting point. The folding of group II intron RNAs has been studied exclusively in this manner (17–20).

Group II introns are autocatalytic RNA molecules that promote their own splicing from RNA and retrotransposition into DNA (21). They can be subdivided into three main families (groups IIA, B and C), although RNA folding studies have been conducted exclusively on the *Sc*. ai5γ intron, which is a prototypical member of subgroup IIB (22,23). In addition to their characteristic splicing mechanism, which involves a branching reaction that circularizes the intron RNA (24–26), group II introns are distinguished by their large size (typically 500–1000 nt without including open reading frames), a lack of phylogenetic sequence conservation, and

*To whom correspondence should be addressed. Tel: +1 203 432 5733; Fax: +1 203 432 5316; Email: anna.pyle@yale.edu
Present address:
Linhui Julie Su, Whitehead Institute for Biomedical Research, Cambridge, MA, USA

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

the remarkable modularity of their defined structural domains (21,27). This modularity is underscored by the fact that the minimal group II intron ribozyme is composed exclusively of Domains 1 and 5, which can cleave small oligonucleotides with multiple turnover (28–30). The chemical rate of reaction is greatly facilitated by Domain 3 and other substructures that are believed to complete and stabilize the active site of the molecule (31,32).

A particularly useful model system for studying group II intron folding is a ribozyme construct known as D135, which is derived from the *Sc.* ai5γ intron (18). This large RNA (618 nt) contains all substructures required for efficient catalysis and it folds homogeneously to a single, active conformer (18,33). Initial studies utilized hydroxyl radical footprinting to gauge solvent accessibility of D135 and to assess which portions of the molecule were internalized upon folding (18). Using this approach, together with circular dichroism and ultracentrifugation the Mg^{2+} dependence and kinetics of D135 folding were also examined (17–19). These studies revealed an unusual folding pathway for D135 in which the molecule traversed a slow, but direct path to the native state. Not only did the molecule require high Mg^{2+} concentrations for both local and global native folding, but it required high Mg^{2+} for molecular collapse.

To determine if an obligate intermediate exists along the D135 folding pathway, we reasoned that it might be confined to a particular domain of the intron. We therefore created deletion constructs of the intron and, for each case, we examined the thermodynamics and kinetics for molecular collapse and native folding. Our studies revealed that Domain 1, which is an independent folding unit, adopts its native structure in the absence of Domains 3 and 5. Importantly, interaction sites for Domain 5 and for the exonic substrate are preorganized, implying that Domain 1 serves as a structural scaffold. In line with the architectural importance of Domain 1, we were also able to establish that this domain is responsible for the high salt requirements of this intron and that folding of D1 is the rate-limiting step in folding of the entire D135 ribozyme. Subsequent docking of Domains 3 and 5 is fast and independent of one another. Therefore, we conclude that formation of Domain 1 constitutes an obligate albeit transient intermediate on the pathway to the native conformation. The molecular collapse of this intermediate occurs on a comparable timescale and has similar Mg^{2+} requirements (both within 2-fold) as folding to the native state of D135.

MATERIALS AND METHODS

Construct design

Plasmids pLJS1 (pBS$^-$ΔT7D1) and pLJS2 (pBS$^-$ΔT7D13) were constructed by PCR amplification of D1 and D13, respectively, from plasmid pQL71 (18). Plasmid pEL31 (pBS$^-$ΔT7D15) was constructed by PCR amplification of D15 from plasmid pBSΔT7exD15. Forward primer design for all constructs includes a T7 RNA polymerase promoter sequence and an EcoRI restriction site. The reverse primer for D1 contains a HindIII restriction site located 15 nt downstream from terminus of D1 sequence. Reverse primer for D13 contains a HindIII restriction site located 17 nt downstream from the terminus of D3 sequence. Plasmid pBS$^-$ΔT7D5TM (ζ′ mutant) was constructed from pQL71 (18) via site-directed mutagenesis (34).

The D135 construct contains elements spanning Domains 1–5 of the *Saccharomyces cerevisiae* mitochondrial intron coxI/ai5γ. The construct begins at the first nucleotide of the intron and terminates with a 36 nt of unrelated 3′ tail that lies immediately downstream of D5 (the tail sequence is derived from the vector). D2 and D4 were shortened to hairpins that are capped by UUCG loops (18). For D135 ζ′ mutant, the tetraloop sequence in Domain 5 is modified from GAAA to UUCG. The D1 construct includes the nucleotides in J1/2 but not D2. The D15 construct contains Domains 1 and 5, with D2 and D4 shortened and capped as described above, and D3 is entirely removed (linkers J2/3 and J3/4 were maintained). Lastly, the D13 construct contains Domains 1 and 3, whereas D2 is shortened and capped as described above.

RNA preparation

Following standard *in vitro* transcription procedures (35), all constructs were transcribed from HindIII-linearized plasmids. RNA transcripts were gel-purified and subsequently eluted in a buffer of 40 mM MOPS, pH 6.0 and 10mM EDTA. RNA stock solutions were stored at −80°C in 10 mM MOPS, pH 6.0 and 1 mM EDTA (ME) in order to reduce non-specific cleavage.

The RNA substrate S17/7 (5′-CGUGGUGGGACAUU-UUCGAGCGGU-3′), representing the last 17 nt of the 5′ exon and the first 7 nt of the intron (18), was synthesized on an ABI 392 synthesizer and deprotected according to standard methods (36).

Monitoring folding by hydroxyl radical footprinting

The RNA was labeled either at the 5′ end with [γ-^{32}P]ATP and T4-polynucleotide kinase (NEB), or at the 3′ end with [α-^{32}P]dCTP, a complementary oligonucleotide, and DNA polymerase I large fragment [Klenow; NEB; (37)]. Labeled RNA samples (4 nM) were incubated in a buffer of 80 mM MOPS, pH 6.0 and 500 mM KCl at 95°C for 1 min. Samples were then cooled to 42°C, mixed with $MgCl_2$ (100 mM f.c.), and folded at 42°C for 10 min prior to addition of peroxonitrous acid footprinting reagent (30 mM f.c.; reagent kindly supplied by V. E. Anderson).

Footprinting reactions were allowed to proceed for 5 s prior to addition of sodium acetate (300 mM f.c.) and 3 vol of ethanol for precipitation. Products were resolved on denaturing 5–20% polyacrylamide gels. The gels were analyzed using a Molecular Dynamics Storm 840 Phosphorimager. The quantification of footprinting gels was carried out using the software ImageQuant (Molecular Dynamics) as described previously (18). In general, a region is considered 'protected' if ⩾1.7-fold degree of protection is observed. Footprinted regions were summarized for all constructs in Table 1; however, a few small and in other constructs non-recurring footprints were not included in that table, but listed as follows. *D1 construct*: 136, 272, 348, 351–352, 366–367; *D13 construct*: 248, 260, 350, 637; *D15 construct*: 226, 271, 273–274, 346–349, 364; *D135ζ′ mutant*: 206, 226, 261, 307, 362–363, 365–366, 602–603, 629, 631, 805–806.

To monitor the magnesium dependence of folding by footprinting, 4 μl of a 5× $MgCl_2$ stock was added to each 4 nM labeled RNA sample to initiate the folding reaction, which was

Table 1. Internalization of structural elements is compared for different constructs derived from the Sc. ai5γ group II intron

Region	Positions	D135	D1	D13	D15	D135ζ′
1	1–10	●	○	◉	◉	◉
2	13–16	●	○	○	●	●
3a	23–28	●	◉	●	●	●
3b	34–36	○	○	◉	●	●
4	45–47	●	●	●	●	●
5	55–60	●	●	●	●	●
6a	66–71	●	●	●	●	●
6b	72–78	●	●	●	●	●
7	84–86	●	○	●	●	●
8	96–101	●	○	○	○	○
9	110–119	●	○	●	●	●
10a	122–124	●	●	●	●	●
10b	125–127	○	◉	●	●	●
11	128–134	●	●	●	●	●
12a	142–146	●	◉	◉	◉	●
12b	148–151	○	○	○	●	●
13a	170–176	●	◉	◉	●	●
13b	179–190	●	●	●	●	●
14	198–204	●	●	●	●	●
15	211–214	●	○	○	◉	●
16a	220–223	●	●	●	●	●
16b	228–236	●	●	●	●	●
16c	237–242	○	◉	●	◉	●
16d	250–253	○	●	○	○	●
16e	254–256	○	●	●	●	●
16f	264–267	●	●	●	●	●
17a	314–324	●	●	●	●	●
17b	325–329	○	●	●	●	●
18	330–342	●	●	●	●	●
19	355–359	●	◉	◉	◉	◉
20	371–373	●	●	●	●	●
21	380–385	●	●	●	●	●
22a	393–395	●	○	○	○	○
22b	396–398	○	◉	◉	●	●
23	399–401	●	●	●	●	●
24a	407–409	○	○	○	●	●
24b	410–414	●	○	○	●	●
25a	415–419	○	○	○	◉	●
25b	421–424	●	○	○	○	◉
25c	580–583	○		○	●	●
26	584–590	●		○	●	●
27	596–601	●		○		○
28	607–610	●		○		○
29	620–621	●		○		○
30	659–665	●		◉		◉
31	674–676	●		○	○	○
32a	807–827	●			◉	○
32b	828–833	●			●	○
32c	834–848	●			◉	○

The wt D135 ribozyme was compared with a mutant D135ζ′ ribozyme, which is catalytically inactive, as well as to diverse deletion constructs, in which entire intronic domains were deleted (for description see Materials and Methods). The protected regions are numbered (1–32 with subdivisons) and the residues constituting a region are specified as well. The color of the dots implicates the percentage of residues that are protected in a particular footprinting region. White dots indicate that 0–32% of the nucleotides are protected within this defined region, while gray refers to a range of 33–65% and black implies a protection level of 66–100%. Footprints that consist of only two residues are not included in the table but are listed in the Materials and Methods.
Open circles, 0–32%; Gray circles, 33–65%; Closed circles, 66–100%

allowed to proceed at 42°C for 10 min prior to the addition of peroxonitous acid footprinting reagent. The concentration of $MgCl_2$ stock solutions ranged from 5 to 750 mM in these experiments (resulting in 1–150 mM $MgCl_2$ f.c.).

The extent of protection, Y, for each region was plotted with respect to increasing $[Mg^{2+}]$ and fit to a binding isotherm as described previously (19):

$$Y = \frac{[\mathrm{Mg}]^n}{K_{\mathrm{Mg}}^n + [\mathrm{Mg}]^n} \quad \mathbf{1}$$

Time-resolved footprinting experiments were initiated with the addition of $MgCl_2$ to end labeled D1, resulting in a f.c. of 80 mM MOPS, pH 6.0, 500 mM KCl and 100 mM $MgCl_2$. Folding was allowed to proceed at 42°C for a specified time (5 to ∼600 s) prior to the addition of peroxonitrous acid footprinting reagent. The products of the footprinting reactions were resolved as described above and quantification of footprinting gels were carried out as described previously (19). The extent of protection, Y, for each region was plotted with respect to time and fit to a first-order rate equation as follows (19):

$$Y = 1 - e^{-kt} \quad \mathbf{2}$$

where Y = fractional saturation of a given site at time, t, and k is the first-order rate constant (min^{-1}).

Monitoring D1 compaction by sedimentation velocity

D1 RNA (17 nM f.c.) was denatured by incubation at 95°C for 1 min in the presence of 500 mM KCl and 80 mM MOPS, pH 7.0. After cooling down the samples to 42°C, $MgCl_2$ (100 mM f.c.) was added to initiate folding, which was then allowed to proceed at 42°C for 10 min. In case of the unfolded state the RNA was incubated for 10 min at 42°C in the presence of 0.5 M KCl only. In addition, reference samples without RNA were prepared. The RNA samples as well as the corresponding reference solutions were loaded into assembled cells and then subjected to analytical ultracentrifugation as it has been described previously for the analysis of D135 RNA (17). Scans were analyzed as described previously to obtain the hydrodynamic parameters $S_{20,w}$ and $D_{20,w}$, to calculate Stokes radius (R_H), frictional coefficient (f/f_o), and the axial ratio (a/b) for a prolate ellipsoid (17).

Single-turnover kinetics

Substrate S17/7 was 5′-end labeled with [γ^{32}P]ATP and T4-polynucleotide kinase (NEB). Cleavage reactions (20 µl) were carried out at 42°C in 80 mM MOPS, pH 7.0, 100 mM $MgCl_2$ and 500 mM KCl. The ribozyme (100 nM f.c.) and the substrate S17/7 (1 nM f.c.) were incubated separately at 95°C for 1 min in 80 mM MOPS, pH 7.0 and 500 mM KCl. Samples were cooled to 42°C, $MgCl_2$ was added (100 mM f.c.), and RNA was allowed to fold for 10 min. Cleavage reactions were initiated upon the addition of the substrate to the ribozyme. Aliquots (1 µl) were removed from the reaction volume at specified timepoints, mixed with 3 µl of quench buffer, and subjected to 20% polyacrylamide gels. Gels were quantified as described above and fit to the following equation to obtain the reaction rate constant:

$$frac[\mathrm{Precursor}] = A_1 + A_2 e^{-kt}, \quad \mathbf{3}$$

where A_1 represents fraction of unreacted precursor, A_2 the fraction of reacted precursor substrate, and k is the first-order rate constant (in min^{-1}).

Native gel electrophoresis

The 5′-labeled RNA (0.1 nM, f.c.) was denatured by heating the sample for 1 min at 95°C in the presence of 500 mM KCl and 80 mM MOPS, pH 7.0. After cooling the samples to 42°C, $MgCl_2$ was added to initiate folding. For Mg^{2+} titration experiments a range of 0.1–200 mM $MgCl_2$ (f.c.) was added and folding was allowed to proceed at 42°C for 10 min followed by loading of the samples on the native gel. Folding kinetics were performed in the presence of 100 mM $MgCl_2$ (f.c.) and aliquots were taken at the given time points (0–10 min). Native gel electrophoresis was performed according to published methods (38–40).

Quantification was performed as described above (Monitoring folding by hydroxyl radical footprinting). The radioactivity per lane reflects $f_U + f_N = 1$. The fraction of the native RNA population was plotted versus the $MgCl_2$ concentration and fit to a Hill model (Equation 1). *Y* represents the fraction of the native population, K_{Mg} is the midpoint of transition and *n* is the Hill constant. For the folding kinetics the fraction of the native RNA population was plotted versus time and the data were fit to a single-exponential Equation 3. A_1 represents the fraction of the unfolded RNA, while A_2 is the fraction of the native population and *k* is the folding rate constant.

RESULTS

D1 folds independently and properly in isolation

Although previous studies have indicated that D1 can adopt a folded form in isolation (20), the experiments were not designed to determine if the isolated D1 fold is structurally similar to that of the intact D135 ribozyme. To address this issue, we conducted comparative hydroxyl radical footprinting studies on D135 and D1 RNA molecules. This method allows us to determine the relative solvent accessibility of each D1 nucleotide in both contexts and to determine if the same D1 residues participate in formation of the functional, compacted state.

Upon comparison of protection patterns in folded forms of both molecules, it is apparent that most of the nucleotides that are protected in D1 of the D135 ribozyme are also protected in the isolated D1 RNA (Table 1). These include the D1b and D1c stems (regions 5 to 6b and 13b), the first section of D1d (regions 14 and 21) and the region that contains EBS1 (regions 17a to 19; see red squares in Figure 1A and gels in Figure 1B). This strongly suggests that the folded form of isolated D1 is relevant for catalytic function. The observed footprints (Figure 1B) increase in intensity in the presence of other intron domains, indicating that these domains significantly contribute to the global stability of the entire structure. Importantly, the sections of D1 that are protected only in the context of the active D135 ribozyme are substructures that are known to participate in active-site formation together with catalytic subdomains D5 and D3 [vida infra; (41–44)]. These include the first 10 nt of the intron (region 1), the D1c1 extension that contains ε and λ (regions 7 and 9), and parts of the extended D5 receptor in the D1d stem (region 15; see blue squares in Figure 1A and gels in Figure 1C). The few nucleotides that internalize only in the context of isolated D1 (regions 10b, 16c, 17b and 22b) are primarily short extensions of substructures that are protected in both constructs (see green squares, Figure 1A). It would, in fact, be surprising if there were not at least some small portions of D1 (regions 16d and e) that did not undergo rearrangement in the presence of other domains. Taken together, the hydroxyl radical footprinting patterns indicate that the structural core of D1 is folded independently of other domains and that it does not undergo large conformational rearrangements upon docking of additional intronic components to form the active ribozyme.

While hydroxyl radical footprinting provides information on local structural accessibility, it does not provide information on global folding or molecular compaction. To better understand global features of the D1 architecture, analytical ultracentrifugation was used to compare the folded and unfolded states of the D1 and D135 molecules. It is noteworthy that, in this study, the unfolded state is defined as the conformation of D1 or D135 in the presence of 0.5 M KCl (17). Modification interference studies of the KCl state indicate that it lacks all tertiary structure, although duplex regions are formed (C. Waldsich and A. M. Pyle, unpublished data). Consistent with formation of a defined secondary structure, addition of KCl to D135 results in significant compaction, which most probably reflects the electrostatic relaxation of the molecule (R_H changes from 146 to 80 Å). A large degree of additional compaction occurs when Mg^{2+} is added [to 61Å; (17)], representing the transition to a D135 native state that contains requisite tertiary structural elements.

Ultracentrifugation studies of D1 indicate that the sedimentation coefficient ($S_{20,w}$) increases substantially in the folded form (Table 2), and that the folded state of D1 has become significantly compacted relative to the unfolded state. Remarkably, the degree of compaction is similar for the D1 and D135 molecules (24% versus 23%, respectively) as judged from R_H values from Table 2 and (17).

The high Mg^{2+} requirements for group II intron function involve D1 folding

All group II introns that have been characterized *in vitro* require conditions of high ionic strength for self-splicing (45–48). In the case of the *Sc.* ai5γ intron (and probably others), this is due to a high [Mg^{2+}] requirement for folding. The apparent two-state behavior of D135 folding is consistent with one of two models: (i) a substructure with a high affinity constant for Mg^{2+} (a weak metal binding site) is required for a late and concerted transition that brings together the entire intron; (ii) there is an obligate intermediate early in the folding pathway that requires high concentrations of Mg^{2+} to form (i.e. the intermediate may have a weak affinity constant for Mg^{2+} that results from a slow k_{on} for Mg^{2+} binding or a fast k_{off}). To determine if this potential intermediate lies within a substructure of the intron, we examined D1 and D135 folding as a function of Mg^{2+}.

To examine the Mg^{2+} requirements for all regions of D1 simultaneously, and to determine if it contains any regions of unusually high stability, we treated D1 and D135 with varying concentrations of Mg^{2+} and then subjected them to hydroxyl radical footprinting [Supplementary Figure 1A and (19)]. As reported previously for D135, all regions of D1 fold with similar, relatively high values for $[Mg^{2+}]_{1/2}$ (Figure 2A). Strikingly, these values (K_{Mg} of 24 ± 2.6 to 33 ± 3.1 mM) are very

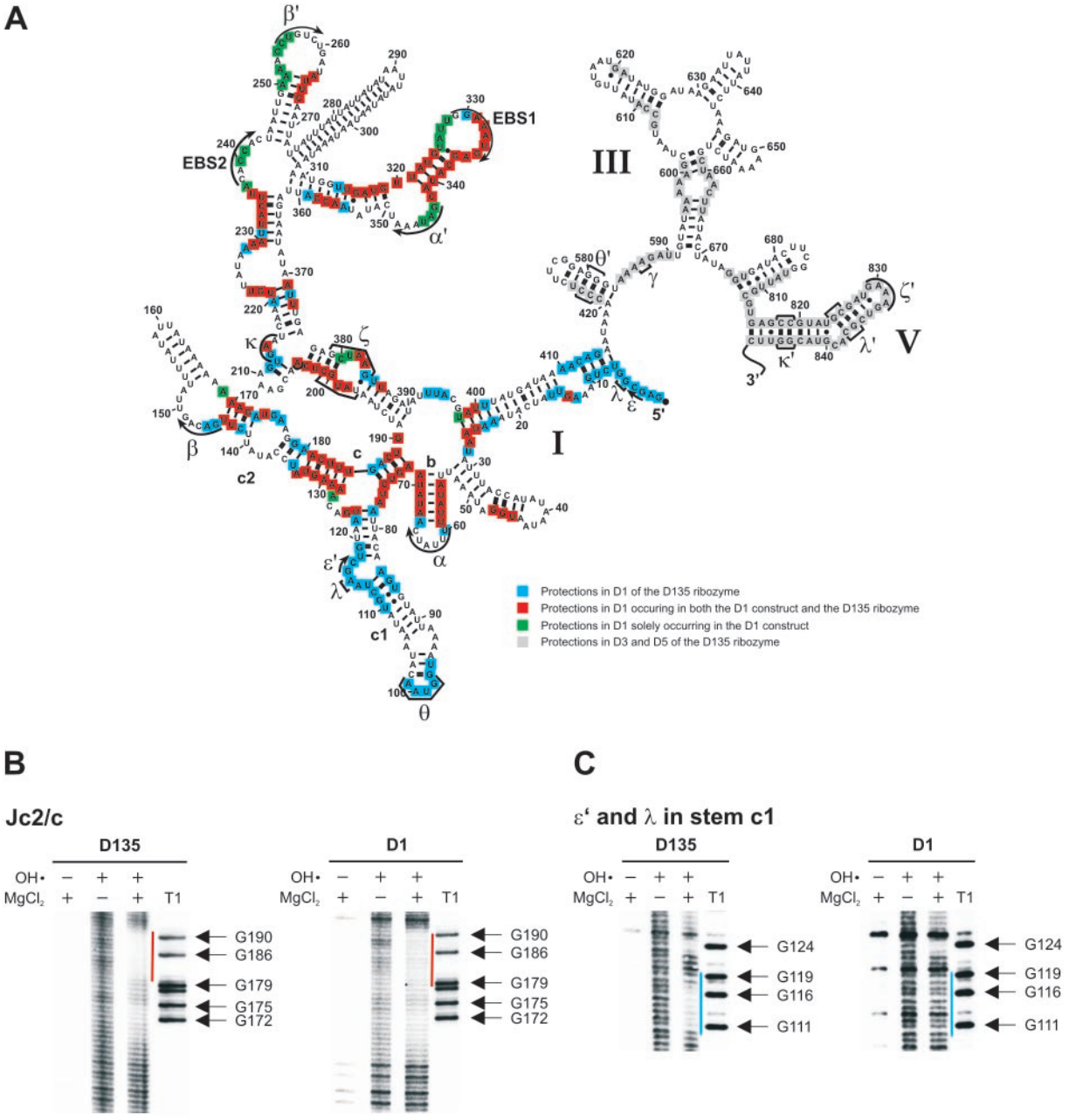

Figure 1. D1 is an independent folding unit. (**A**) Secondary structure map of the *Sc*. ai5γ group II D135 ribozyme, which summarizes the hydroxyl radical footprinting pattern obtained for the D135 ribozyme and a D1 construct. Protections that only occur within D1 of the D135 ribozyme are highlighted in blue; those that occur in D1 of both the D1 construct and the D135 ribozyme are colored in red, while protections unique to the D1 construct are depicted in green. Footprints within the D135 ribozyme that are not part of D1 are shown in gray. (**B**) Representative gels are shown for a structural element within D1 that is protected in both the D135 ribozyme and D1 in isolation. Red bars next to the gels indicate the region of interest. (**C**) Representative gels are shown for a structural element within D1 that is protected in the D135 ribozyme but not in D1 in isolation. Blue bars next to the gels indicate the region of interest.

similar to those reported previously for D135 [K_{Mg} of 18 ± 1.3 to 47 ± 3.7 mM; (19)], indicating that the folding requirements for the complete ribozyme are due specifically to the requirements of D1.

While hydroxyl radical footprinting provides insight into local Mg^{2+} requirements, it is informative to examine the global Mg^{2+} dependence of D1 and to compare it with that of D135. Gel-mobility shift experiments, like sedimentation velocity experiments, provide a metric of both global folding and molecular compaction. We therefore examined the Mg^{2+} dependence of global compaction for both D1 and D135 using a gel-shift method that we adapted recently for study of group II introns (Figure 2B). The results for the two molecules were virtually identical (K_{Mg} of 13.9 ± 1.4 and 12.3 ± 0.7 mM), indicating that D1 and D135 have the same ionic requirements for global folding and molecular compaction (Figure 2C and Supplementary Table 1). Importantly, the profile of Mg^{2+} dependence for D135 determined by this new gel-shift method matches previously published values reported from sedimentation velocity experiments [K_{Mg} of 15 ± 2 mM; (17)].

Furthermore, as reported previously for D135, the $[Mg^{2+}]_{1/2}$ for compaction of D1 is unusually high, which indicates that an RNA substructure with a weak Mg^{2+}-binding constant is involved in the earliest stages of D1 (and hence *Sc*. ai5γ group II intron) assembly.

Folding of D1 is the rate-limiting step for compaction and native assembly of D135

Although Mg^{2+} titrations provide information on the relative stability of D1 and its constituent substructures, they do not provide direct information on the kinetic pathway for folding of the molecule. To understand the kinetic behavior of native D1 and D135 folding and compaction, we used hydroxyl radical footprinting and native gel electrophoresis, respectively.

By monitoring the appearance of hydroxyl radical protection patterns as a function of time, we observe that all regions of D1 internalize simultaneously, as reported previously for D135 [Supplementary Figure 1B and Figure 3A; (17,19)]. Furthermore, the rate constants for internalization (0.7 ± 0.2 to 1.3 ± 0.2 min^{-1}) match the rate constant for folding of D135 determined by hydroxyl radical footprinting [0.9 ± 0.2 to 2.8 ± 0.9 min^{-1}; (19)] and by ribozyme activity assays [~1 min^{-1}; (18)]. These data indicate that D1 reaches the native state with the same rate constant as D135, and therefore the rate-limiting step for folding of the *Sc*. ai5γ intron is folding of D1.

Many RNA molecules have been reported to reach the native state slowly despite a rapid early compaction event. To determine if D1 and D135 compact with the same rate constant, and whether this rate constant is faster than that for reaching the native state, we monitored collapse of the two molecules by native gel electrophoresis (Figure 3B). The resultant time courses for D1 and D135 compaction are almost

Table 2. Comparison of sedimentation values and hydrodynamic parameters for D1 and D135 constructs

Parameters	Unfolded D1	Folded Dl	Unfolded D135[a]	Folded D135[a]
$S_{20,w}$	8.6 ± 0.02	11 ± 0.2	10 ± 0.04	14 ± 0.3
$D_{20,w}$	3.8 ± 0.3	4.4 ± 0.1	0.2 ± 0.1	4.2 ± 0.5
f/f_0	2.2 ± 0.0	1.7 ± 0.03	2.3 ± 0.0	1.8 ± 0.0
R_H(Å)	67 ± 0.2	51 ± 0.9	80 ± 0.3	61 ± 1.5
a/b	13 ± 0.1	5.6 ± 0.3	15 ± 0.0	7.1 ± 0.0

RNAs were folded in the presence of 0.1 M $MgCl_2$ and 0.5 M KCl, while the RNA remains unfolded in 0.5 M KCl only. Average values and deviation was obtained from two independent experiments in case of the D1 construct.
[a]Data taken from Su *et al*. (17)

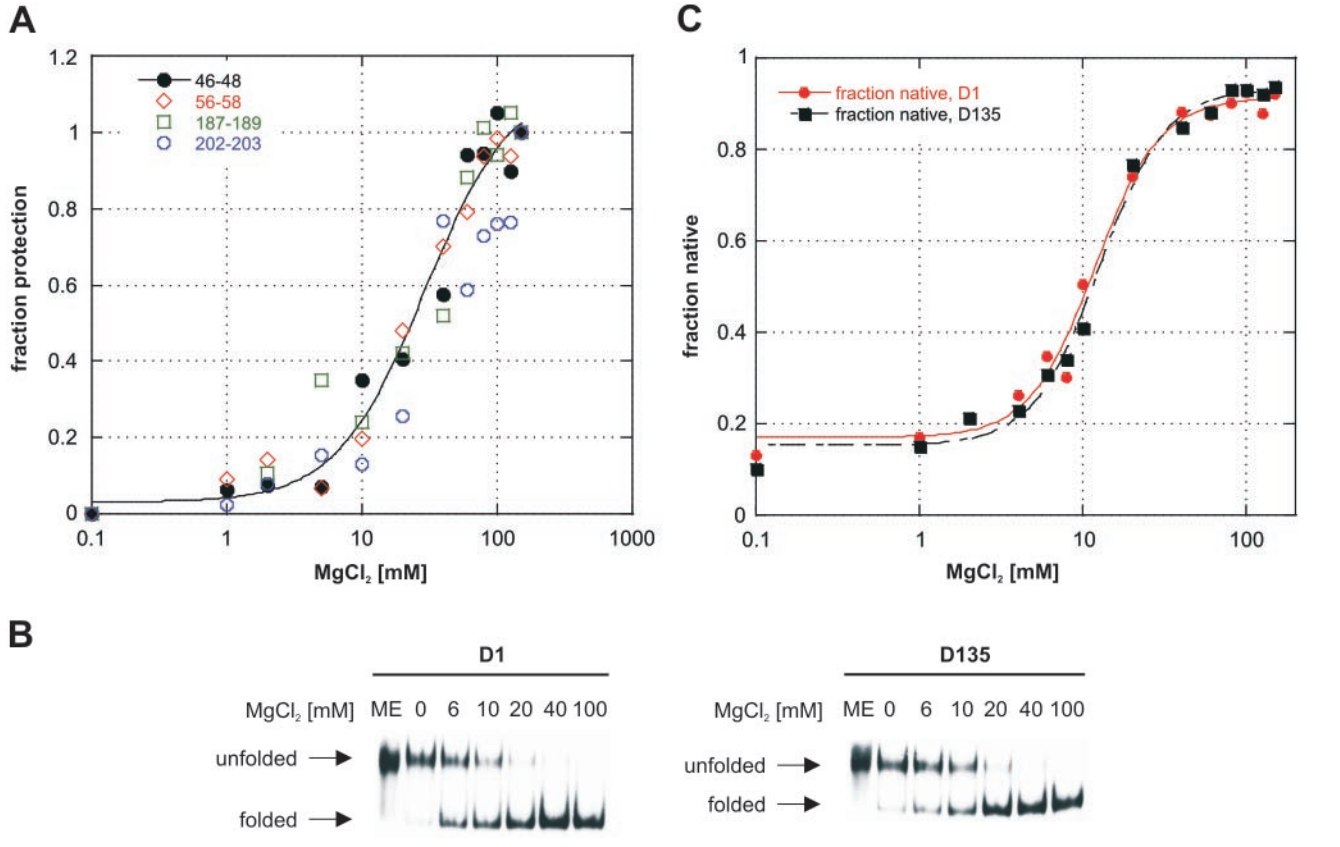

Figure 2. Folding of D1 is the high salt-requiring step in formation of the native D135 fold. (**A**) Representative titration curves from nt 46–48 (closed circles), 56–58 (open diamond), 187–189 (open square) and 202–203 (open circle) in D1 RNA, which become protected from hydroxyl radical cleavage with increasing Mg^{2+} concentration. K_{Mg} for these regions ranges from 24 ± 2.6 to 33 ± 3.1 mM, with n_H values from 1.5 ± 0.21 to 2.3 ± 0.24. (**B**) Mg^{2+}-dependent folding of the D1 construct and the D135 ribozyme monitored by native gel analysis. (**C**) Representative titration curves deduced from native gels for the D1 construct and the D135 ribozyme are shown in red (closed circle) and black (closed square). K_{Mg} for D1 in isolation and for D135 are 13.9 ± 1.4 and 12.3 ± 0.7 mM, with n_H values of 1.69 ± 0.29 and 1.72 ± 0.34.

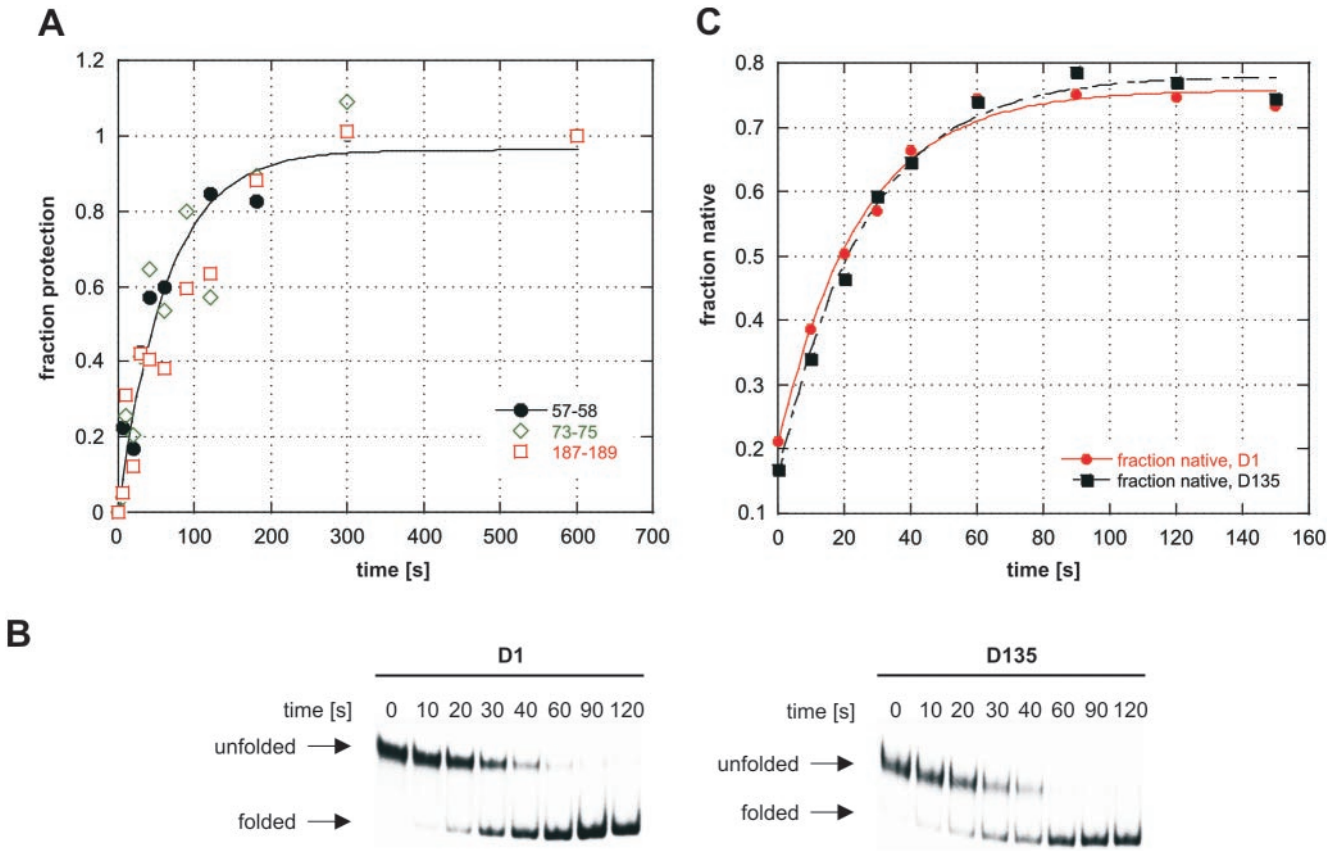

Figure 3. Folding of D1 is the rate-limiting step in formation of the native D135 fold. (**A**) Kinetic progress curves for protected regions of D1, which become increasingly protected from hydroxyl radical cleavage over time. Representative regions shown are nt 57–58 (closed circle), 73–75 (open diamond) and 187–189 (open square). Observed rate constants for these regions range from 0.7 ± 0.2 to 1.3 ± 0.2 min^{-1}. (**B**) Time-dependent folding of the D1 construct and the D135 ribozyme monitored by native gel analysis. (**C**) Kinetic progress curves deduced from native gels for the D1 construct and the D135 ribozyme are shown in red (closed circle) and black (closed square). The k_{fold} was found to be 1.9 (±0.1 or 0.3) min^{-1} for both D1 in isolation and the D135 ribozyme.

superimposable, and the rate constants (1.9 ± 0.1 and 1.9 ± 0.3 min^{-1}, Figure 3C and Supplementary Table 1) are very similar (within 2-fold) to those reported for native state formation, as measured by all other methods of analysis [~1 min^{-1}, Figure 3B; (17,18)]. Thus, D135 not only reaches the native state slowly, but it compacts slowly as well. Furthermore, the rate-limiting step for compaction and native state assembly lies with an event involved in D1 folding.

It is important to note that the $[Mg^{2+}]_{1/2}$ for folding to the native state [determined by footprinting of D135; see (19)] and for molecular collapse of the intron (as studied using a D12356 construct; data not shown) is not affected by the presence of 0.5 M KCl in the reaction. This indicates that the high Mg^{2+} requirements observed here and in previous work (17–19) are not owing to a competition between the K^+ and Mg^{2+} ions for binding to the RNA. This assumption is further supported by the fact that the footprinting patterns remain unaltered in the absence of KCl. Footprints are more defined and have higher intensity in the presence of KCl (18), an observation which correlates with the fact that homogeneity of the folded population increases upon preincubation with KCl (18,19). However, addition of KCl during the preincubation step (to stimulate secondary structural collapse) does not increase subsequent requirements for Mg^{2+} in this case.

Domain 1 is a preorganized scaffold for intron assembly

The previous experiments demonstrate that D1 folds independently into a conformation that is biologically relevant. Together with published work on functional substrate binding by D1 in isolation (20), this suggests that D1 may have a high level of preorganization and that it acts as a scaffold for subsequent assembly of other intronic domains. We were therefore interested in exploring the level of functional D1 preorganization and its role in interdomain binding.

The κ and ζ regions of D1 are known to interact with the 'binding face' of D5 (42,43,49). The resultant contacts form a strong interface between D1 and D5, thereby positioning the opposite face of D5 for function in catalysis. It was therefore of interest to determine whether the κ and ζ elements (regions 14, 15 and 21) of D1 are somewhat preorganized or 'folded' prior to D5 binding. Hydroxyl radical protection patterns were examined for the κ and ζ regions of three different RNA constructs: D1 in isolation, the D15 construct (Materials and Methods) and D135. Remarkably, significant amounts of κ and ζ protection are observed even for D1 in isolation (Figure 4A). Notably, the ζ region exhibits a very distinct degree of protection, while that of the κ element is less pronounced (Figure 4A). For both regions these protections expand and grow more intense in the presence of D5, in the D15 as well as the D135 constructs. In the latter case,

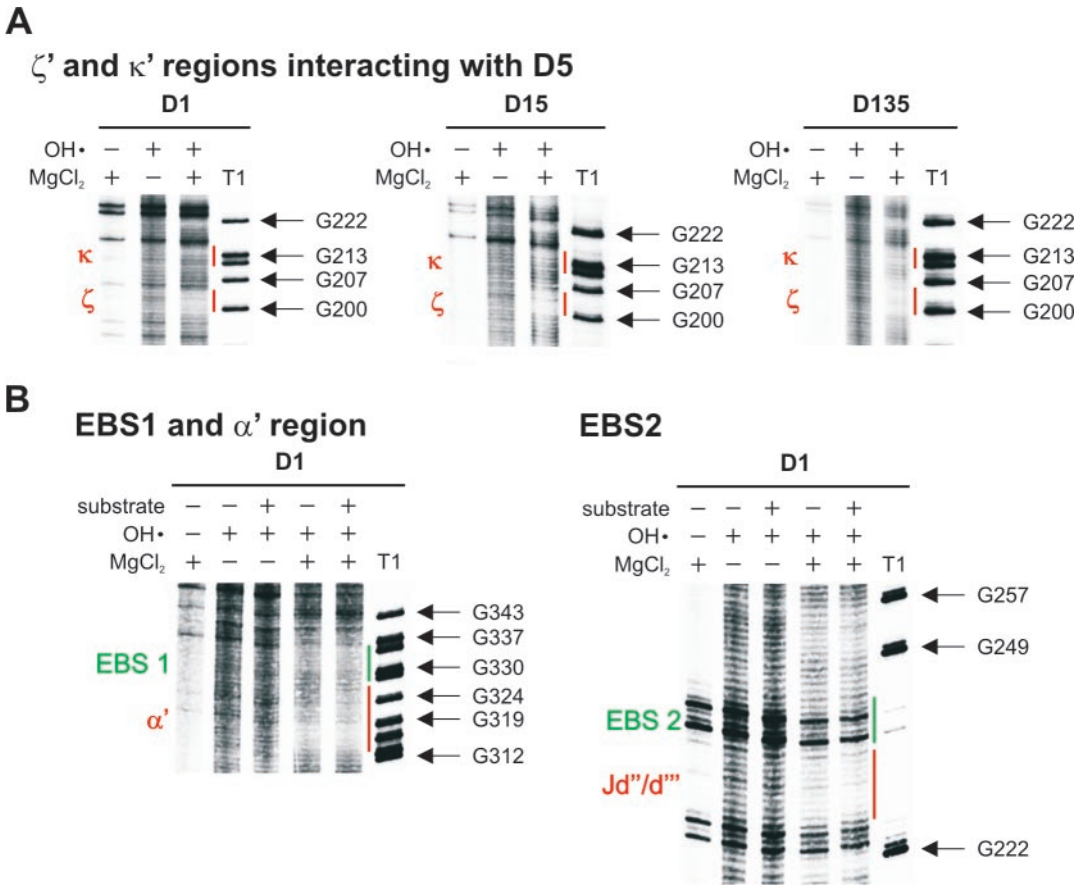

Figure 4. D1 serves as a scaffold for docking of other intron domains and provides preorganized binding sites for other domains. (**A**) Representative gels display the κ and ζ regions in D1, with which D5 interacts. Red bars next to the gels indicate the regions of interest. (**B**) In the hydroxyl radical footprinting gels structural elements within D1 (in isolation) that are involved in substrate binding are compared in the absence and presence of the substrate S17/7. Green bars next to the gels indicate EBS1 and EBS2, respectively, while the red bars mark footprints close to those regions.

where both D3 and D5 are present, the greatest amounts of protection are observed, which is consistent with the expected greater degree of water exclusion when all catalytic domains are assembled as well as an increased global stability (Figure 4A and Table 1). Thus, while protections in D1 are certainly enhanced upon D5 binding (especially the footprint at the κ element), the receptor site ζ is already highly organized and protected from solvent prior to binding of D5 itself.

The EBS regions of D1 form two sets of base pairs with the 5′-exon, and the α′ region that is close to EBS1 is involved in a long-range tertiary contact within D1 (22,50,51). Owing to the functional importance of these regions, we used hydroxyl radical footprinting to assess their level of preorganization in the absence of 5′-exon (or analogous substrate) binding. We find that the EBS regions are only weakly protected both in the presence and absence of exonic substrate. The relative solvent accessibility of the EBS region is therefore consistent with binding of the exons along the folded intron surface (52). Importantly, the adjacent structural elements α–α′ [which is necessary for the proper substrate configuration; (51,53)] and Jd″/d‴ are both highly internalized in the presence of the substrate (Figure 4B). Given that the protection patterns of the EBS regions and of their surrounding tertiary structures do not change upon substrate binding, it is probable that this region is preorganized prior to association of the 5′-exon.

D5 and D3 dock independently of one another into D1

Having established that D1 is a scaffold and that it folds independently, it was of interest to determine if there is a hierarchy in the subsequent docking of additional domains. As a first step towards understanding this, we sought to learn whether D3 and D5 can dock independently of one another as suggested by previous kinetic and binding data (28,32,54).

First, it was of interest to determine whether the presence of D3 is required for internalization of the D5. In the construct D135, one of the most protected regions is D5, which is essentially buried within the folded RNA [Figure 5A; (18)]. In the absence of D3 (construct D15), the D5 region is still highly protected from solvent, thereby indicating that D5 can interact independently with D1. As expected for a more complete construct, footprints are stronger and more extensive in the case of D135, but they are nonetheless prominent for D15. An important control for D5 docking is the D135ζ′ mutant, in which D5 contains a mutant tetraloop that can no longer dock into the ζ–ζ′ receptor site (Figure 5A). After treating D135ζ′ under conditions that normally promote folding of the

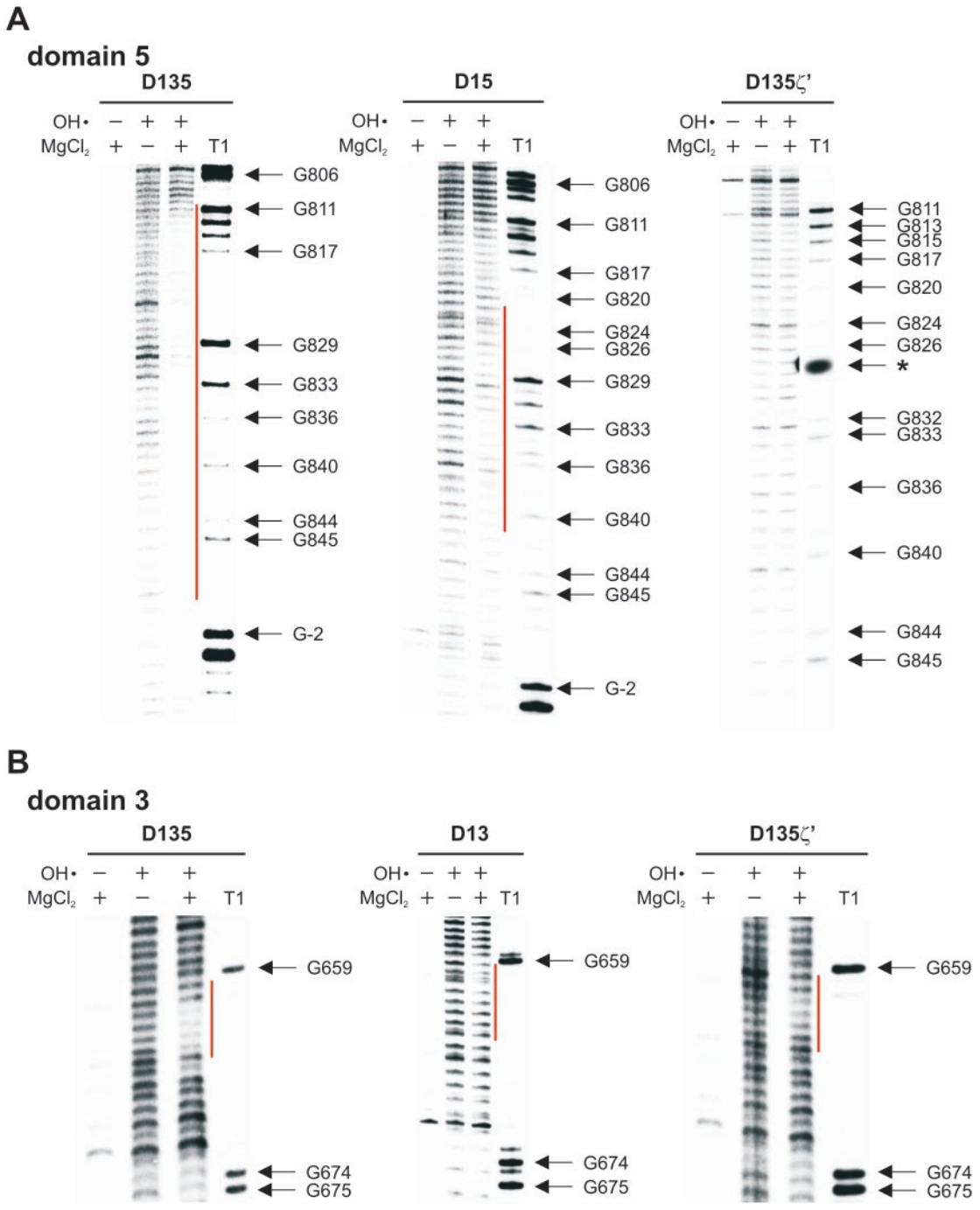

Figure 5. D3 and D5 dock independently of each other onto D1. (**A**) Representative footprinting gels depict changes within the protection pattern of D5 in the absence and presence of D3. The asterisk indicates an aberrant break of the labeled RNA. (**B**) Changes within the protection pattern of D3 in the absence and presence of D5. Red bars next to all gels indicate the regions of interest.

intron, D5 is not internalized. Thus, protection of D5 requires functional tertiary contacts with D1 but does not require D3.

Second, it was of interest to determine whether the presence of D5 is required for footprinting of the Domain 3. The internal loop at the base of D3 is highly conserved in IIB introns and is among the most highly protected regions of D135 [Figure 5B and (18)]. This important region (nt 660–665) is still footprinted in the D13 construct that lacks D5 and in the D135ζ′ mutant that lacks a D5 capable of proper assembly. Thus, D5 is not required for docking of an important D3 element into D1.

Consistent with the comparative studies on D1 and D135, addition of a single additional domain (i.e. either as construct D13 or D15) does not affect the stability (as reflected by the $[Mg^{2+}]_{1/2}$) or the rate constants for folding of the RNA (Table 1 and Supplementary Figure 2). Therefore, these domains provide no individual contribution to the folding pathway and are probable to act independently of one another. Nonetheless,

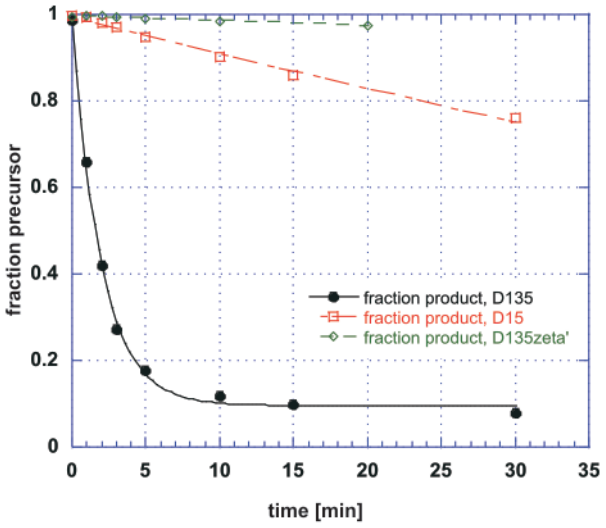

Figure 6. Catalytic capability of the D15, D135 and D135ζ′ ribozymes. Single-turnover cleavage assays of a substrate by D135 (closed circle) and D15 (open square) ribozymes as well as by the D135ζ′ mutant ribozyme (open diamond) are shown. Observed rate constants are 0.5 ± 0.05, 0.013 ± 0.004 and 0.0018 ± 0.0006 min^{-1}, respectively.

their role in catalysis is underscored by the fact that removal of D3 strongly inhibits the reaction and elimination of D5 docking abolishes catalysis altogether [Figure 6; reviewed in (28)].

In order to extend this analysis to folding of a functionally intact intron, we examined compaction of the D12356 construct (in which only D4 was shortened to a small UUCG capped hairpin). Preliminary analysis indicates that D12356 has Mg^{2+} requirements (K_{Mg}) that are similar to D1 and D135 and it displays a comparable timescale for molecular collapse (data not shown). Thus, D1 is clearly the key player in folding of the D135 and the D12356 ribozymes.

DISCUSSION

Domain 1 folding precedes subsequent intron assembly

Here we show that D1 and D135 collapse and adopt the native state with the same rate constants and with the same Mg^{2+} requirements for stability. This indicates that, despite its apparent two-state character, folding of the *Sc.* ai5γ ribozyme (D135) is hierarchical and wholly dependent on formation of the most upstream domain.

Given that D1 is the first region of the intron to be transcribed, this is a reasonable biological strategy. One can envision that, immediately after transcription (or possibly during transcription), D1 forms a stable structure that is protected from degradation and which is preorganized to nucleate subsequent domain assembly. Rapid docking of D3 and D5 into preassembled D1 pockets ensures that downstream regions of the intron are also protected from the cellular environment after synthesis and that these critical active-site components do not interact improperly, misfold or create a ribozyme with alternative specificities. By assembling downstream core components (i.e. D3 and D5) late in the folding process, ribozyme activation is not achieved until all players are present and docked in their proper configuration. This may be important for a self-splicing/transposing molecule that requires faithful selection of reaction sites in order to ensure proper biological function.

Care must be taken, however, in extending any interpretations of our data to the situation *in vivo*, where *Sc*.ai5γ splices at much lower concentrations of Mg^{2+} and where a specific protein has been implicated in function. Indeed, the function of *Sc*.ai5γ *in vivo* (along with that many other RNAs) is strongly dependent on expression of the yeast protein Mss116p (55). Mss116p and other proteins may stimulate splicing through one of many different mechanisms, all of which have precedent in the literature. Proteins have been shown to lower Mg^{2+} requirements by binding to obligate intermediates or to the native state in an RNA folding pathway (11). Proteins can provide selective electrostatic stabilization of an RNA substructure, thereby alleviating the need for high concentrations of magnesium. Alternatively, proteins with mechanical properties (such as RNA stimulated ATPases) can actively rearrange RNA structures through unwinding and/or annealing (56–58). The Mss116p protein may function through any of these mechanisms, and it will be an interesting area for further exploration. It is worth noting, however, that core folding of *Sc*.ai5γ (k_{fold} = 1–2 min^{-1}) is far more rapid than the fastest reported rate constants for self-splicing of this intron [<0.2 min^{-1} *in vitro*; (59)], so proteins and other cellular factors may contribute at multiple stages during folding and splicing of *Sc*.ai5γ. Protamines and polyamines, which are highly abundant in the cell, also have strong effects on the activity and ionic requirements of *Sc*.ai5γ (59), so any consideration of *in vivo* activities must take these into account, as well.

Although D1 folds first, group II intron assembly is only partially hierarchical. Ordered assembly of subsequent domains does not appear to be required, at least in the D135 ribozyme construct. D3 and D5 can dock independently into D1, which suggests that interactions with D1 drive intron assembly rather than interactions between domains themselves. The findings also help clarify the role of D3 as an 'enhancer' of chemical catalysis (31,32). Consistent with previous binding studies, D3 does not stimulate catalysis by helping D5 to interact with D1. Rather, D3 may function by tightening up the folded structure, by excluding additional water, or by direct contribution of catalytic functional groups.

D135 follows a unique pathway for RNA folding and provides a novel paradigm

Most large RNA molecules that have been studied to date do not fold directly to the native state but are caught in misfolded conformations, or kinetic traps, from which they must escape to reach the native conformation (1,2). Furthermore, many of these RNA molecules are observed to compact rapidly into non-native states that reorganize slowly into active forms of the molecule (60–62). Even RNA molecules that have been reported to fold directly to the native state typically undergo rapid initial collapse at magnesium concentrations well below those required for native folding (5,7,14,63,64). Where obligate intermediates are observed along the folding pathway, their formation is not the rate-limiting event during folding. An exception is the RNAse P C-domain, which folds rapidly through a series of on-pathway intermediates, one of which involves a rate-limiting transition to the native state (5).

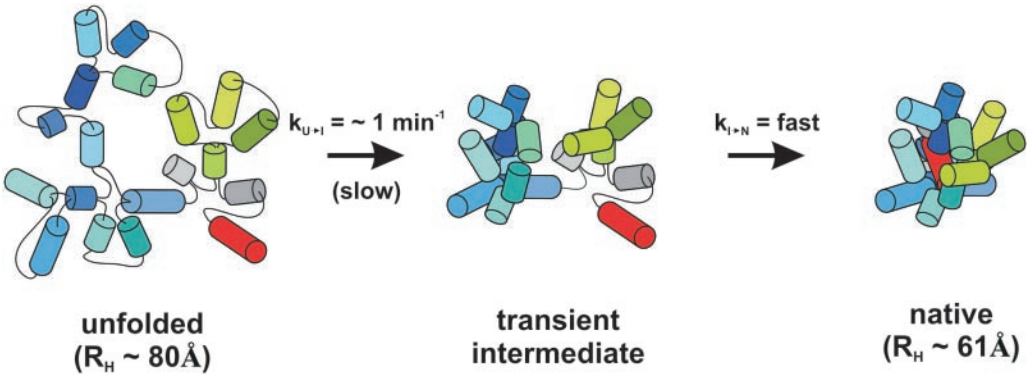

Figure 7. The folding pathway of the Sc. ai5γ D135 ribozyme. Structural elements within Domain 1 are colored in blue shades, while Domains 3 and 5 are highlighted in green shades and red, respectively. Domains 2 and 4 are shown in gray. In the unfolded state only the secondary structure is formed, while in case of the intermediate state Domain 1 compacts and forms tertiary structure thereby providing the scaffold for docking of Domains 3 and 5, which completes folding to the native conformation.

The D135 ribozyme deviates from established RNA folding pathways in almost all significant respects. As reported previously, D135 does not become caught in kinetic traps along the folding pathway. Folding of the ribozyme is slow, direct and it proceeds through an apparent two-state mechanism (17,19). Both compaction and native folding of D135 require high concentrations of Mg^{2+} and there are no subdomains of exceptionally high stability.

The results reported here provide additional support for the unique folding pathway of D135, while providing more detailed insights into this unusual paradigm. As suggested in previous work (17–19), an apparent two-state pathway would mask the presence of an obligate folding intermediate, particularly if the formation of this intermediate was the rate-limiting event for assembly of the entire molecule. Here we show that D1 contains an obligate intermediate, the formation of which is rate limiting for D135 folding (Figure 7). To our knowledge D1 represents only the second known intermediate that is both rate-limiting and on the pathway to a native RNA tertiary structure (5). In comparison with other RNAs, native folding and tertiary collapse of D1 (and hence of the entire ribozyme) are extraordinarily slow, thereby contradicting the notion that early assembly events are universally fast.

That D1 is specifically an obligate intermediate along the folding pathway is based on multiple lines of evidence. (i) D1 collapses and folds prior to any other events along the pathway. (ii) The actual footprinting patterns show that D1 has a similar structure in isolation to that in larger contexts such as D135, indicating that it does not require additional domains to fold. (iii) Any conformational rearrangement within D1 that might take place upon docking of Domains 3 and 5 is necessarily minor and local because one does not observe a faster rate constant for the formation of D1 versus the D13, D15 or D135 constructs. (iv) The mechanism of D135 folding is inconsistent with the presence of an off-pathway intermediate (i.e. a kinetic trap), based on previous studies (17,19). That said, it remains possible that local alterations in D1 structure occur when other domains dock, although this would not affect the basic mechanism presented here.

Although previous studies had shown that native folding of D1 and D135 is slow and direct, they did not rule out an early rapid Mg^{2+}-induced collapse into a precursor native-like state, much like other RNAs examined to date (17,19). Indeed, previous studies on D135 lacked a critical piece of information: kinetic information on the timescale for initial collapse. In this work, we provide the first kinetic analysis of early tertiary compaction by D1 and D135. We find that the rate constants for Mg^{2+}-induced collapse are identical for the two molecules and, perhaps most significantly, that the collapse itself is very slow compared with any other RNA studied to date. These findings significantly widen the range of timescales at which specific RNA folding events are expected to occur.

Although the D1 collapse is very slow, it does not occur with exactly the same rate constant as formation of the D1 native state. The rate constant for collapse is 1.9 min^{-1}, while that for native folding is no higher than 1 min^{-1} (18). Intriguingly, the K_{Mg} for compaction is ~2-fold lower than that for native state formation. These 2-fold differences in kinetic and thermodynamic behavior would be easy to dismiss, as they are relatively small. However, these deviations are highly reproducible, they are observed with a diversity of methods (e.g. compare the K_{Mg} values obtained by sedimentation velocity, gel-shift and hydroxyl radical footprinting), and they were detected during the course of two separate studies [see also (17)]. Taken together, these preliminary findings suggest that the early compacted state of D1 and the native state of D1 are not completely identical. This is suggested by the fact that footprints are barely detectable at 10 mM Mg^{2+} compared with 100 mM Mg^{2+} (Supplementary Figure 1A). Indeed, the data are consistent with the presence of an on-pathway intermediate that precedes formation of the D1 native state. Analysis of a putative metastable 'near-native state' (14,63–65) along the folding pathway of the D135 ribozyme will be an interesting area for future investigations.

In order to put this work into context so that it can be interpreted in light of other RNA folding studies, it is important to emphasize that we have focused on a specific stage in the molecular collapse of group II intron ribozymes. Specifically, we have analyzed the Mg^{2+}-dependent transition from a pool of electrostatically relaxed RNA molecules that had been treated with monovalent cations (which we are calling the 'unfolded state') to the fully compact native conformation. We adopted this strategy for two reasons: (i) we are interested primarily in the formation of tertiary interactions, which occur

only during the Mg^{2+}-dependent transition; (ii) we think that interpretation of folding studies is clearer if both the starting state and the native state are well-defined. Like any other RNA, there are several stages in the collapse of D135 from a completely random polymer. In the absence of any cations, sedimentation velocity experiments show that D135 is maximally extended [146 Å; (17)] and that it compacts significantly when monovalent cations are added [to 80 Å in 0.5 M KCl; (17)]. The Mg^{2+}-induced phase of compaction results in an even greater reduction in molecular size [61 Å; (17)], and this is the event we monitor here. It is worth noting that, despite the significant compaction that occurs upon addition of KCl to D135, the resultant state displays no protection from hydroxyl radicals. This means that our 'unfolded state' has no regions of internalization that protect it from solvent and that the RNA has not adopted a tertiary fold. DMS chemical probing data on this relaxed state show that major elements of secondary structure, but no tertiary contacts, are formed in the KCl-induced relaxed state (C. Waldsich and A. M. Pyle, unpublished data). This is in striking contrast to other ribozymes. For example, the *Tetrahymena* ribozyme adopts a native-like structure in the presence of Na^+ alone [1 M; (15,66)].

One can compare the compaction and native folding of D135 to other ribozymes by focusing on the first phase of molecular compaction that is entirely Mg^{2+} dependent. If one examines this transition for systems that have been studied, it is clear that the tertiary structure collapse of D135 is slow compared with other large ribozymes, such as the RNAse P C-domain and the *Tetrahymena* group I intron. For example, during the second phase of compaction by the *Tetrahymena* ribozyme, electrostatically relaxed molecules collapse in a Mg^{2+}-dependent manner on a 100 ms timescale (16,60). This transition involves RNA motifs that are ultimately required for mediating native long-range tertiary contacts (16). Similarly, RNAse P RNA C-domain forms the collapsed intermediate I_{eq}, which contains secondary and some tertiary structure, on a 1 ms timescale (5,64). If one compares the timescales for the first Mg^{2+}-induced change in compaction of these large ribozymes, it is evident that D135 does not undergo a comparably fast Mg^{2+}-induced phase. Nonetheless, it is essential to note that the *Tetrahymena* ribozyme, the RNAse P RNA C-domain and *Si*. ai5γ D135 ribozyme traverse very different types of free-energy folding landscapes (1,3) and that each RNA has specific optimal conditions for reaching the native state. Owing to these differences in folding behavior and sensitivity to reaction conditions, any comparisons remain difficult and preliminary.

It is probable that D1 contains an even smaller substructure that represents the minimal unit for folding with the kinetic and thermodynamic signatures described here. It will be worthwhile to define this element more precisely and to determine the exact nature of its constituent motifs. This substructure will be particularly valuable because it will represent the only known RNA folding intermediate that limits the rate of the entire folding reaction. If its components are known, it will be uniquely amenable to direct analysis and structural manipulation, making it possible to dissect the molecular and environmental determinants that actually drive native folding. Given the central role of group II introns in eukaryotic evolution, the results will provide important insights into the assembly of numerous macromolecular machines that are now essential for gene expression.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Olga Fedorova for many helpful discussions and for critically evaluating the manuscript. We also thank Eric Westhof for suggesting that we create the D135ζ′ mutant for examining independence of D1 folding. Vernon E. Anderson is acknowledged for the gift of peroxonitrous acid and Olga Fedorova for the synthesis of RNA oligos. This work was supported by NIH grant GM50313 to A.M.P and by a Schrödinger fellowship (J2332) from the Austrian Science Foundation to C.W. (FWF). A.M.P. is an investigator of the Howard Hughes Medical Institute. Funding to pay the Open Access publication charges for this article was provided by Howard Hughes Medical Institute.

Conflict of interest statement. None declared.

REFERENCES

1. Treiber,D.K. and Williamson,J.R. (1999) Exposing the kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.*, **9**, 339–345.
2. Treiber,D.K. and Williamson,J.R. (2001) Beyond kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.*, **11**, 309–314.
3. Sosnick,T.R. and Pan,T. (2003) RNA folding: models and perspectives. *Curr. Opin. Struct. Biol.*, **13**, 309–316.
4. Fang,X.W., Pan,T. and Sosnick,T.R. (1999) Mg^{2+}-dependent folding of a large ribozyme without kinetic traps. *Nature Struct. Biol.*, **6**, 1091–1095.
5. Fang,X., Thiyagarajan,P., Sosnick,T. and Pan,T. (2002) The rate-limiting step in the folding of a large ribozyme without kinetic traps. *Proc. Natl Acad. Sci. USA*, **99**, 8518–8523.
6. Rangan,P., Masquida,B., Westhof,E. and Woodson,S.A. (2003) Assembly of core helices and rapid tertiary folding of a small bacterial group I ribozyme. *Proc. Natl Acad. Sci. USA*, **100**, 1574–1579.
7. Buchmueller,K.L. and Weeks,K.M. (2003) Near native structure in an RNA collapsed state. *Biochemistry*, **42**, 13869–13878.
8. Russell,R., Zhuang,X., Babcock,H.P., Millett,I.S., Doniach,S., Chu,S. and Herschlag,D. (2002) Exploring the folding landscape of a structured RNA. *Proc. Natl Acad. Sci. USA*, **99**, 155–160.
9. Webb,A.E., Rose,M.A., Westhof,E. and Weeks,K.M. (2001) Protein-dependent transition states for ribonucleoprotein assembly. *J. Mol. Biol.*, **309**, 1087–1100.
10. Woodson,S.A. (2005) Metal ions and RNA folding: a highly charged topic with a dynamic future. *Curr. Opin. Chem. Biol.*, **9**, 104–109.
11. Schroeder,R., Barta,A. and Semrad,K. (2004) Strategies for RNA folding and assembly. *Nature Rev. Mol. Cell Biol.*, **5**, 908–919.
12. Das,R., Travers,K.J., Bai,Y. and Herschlag,D. (2005) Determining the Mg(2+) stoichiometry for folding an RNA metal ion core. *J. Am. Chem. Soc.*, **127**, 8272–8273.
13. Pan,J. and Woodson,S.A. (1998) Folding intermediates of a self-splicing RNA: mispairing of the catalytic core. *J. Mol. Biol.*, **280**, 597–609.
14. Buchmueller,K.L., Webb,A.E., Richardson,D.A. and Weeks,K.M. (2000) A collapsed non-native RNA folding state. *Nature Struct. Biol.*, **7**, 362–366.
15. Takamoto,K., Das,R., He,Q., Doniach,S., Brenowitz,M., Herschlag,D. and Chance,M.R. (2004) Principles of RNA compaction: insights from the equilibrium folding pathway of the P4-P6 RNA domain in monovalent cations. *J. Mol. Biol.*, **343**, 1195–1206.
16. Das,R., Kwok,L.W., Millett,I.S., Bai,Y., Mills,T.T., Jacob,J., Maskel,G.S., Seifert,S., Mochrie,S.G., Thiyagarajan,P. *et al.* (2003) The fastest global events in RNA folding: electrostatic relaxation and

tertiary collapse of the *Tetrahymena* ribozyme. *J. Mol. Biol.*, **332**, 311–319.
17. Su,L.J., Brenowitz,M. and Pyle,A.M. (2003) An alternative route for the folding of large RNAs: apparent two-state folding by a group II intron ribozyme. *J. Mol. Biol.*, **334**, 639–652.
18. Swisher,J., Duarte,C., Su,L. and Pyle,A. (2001) Visualizing the solvent-inaccessible core of a group II intron ribozyme. *EMBO J.*, **20**, 2051–2061.
19. Swisher,J., Su,L., Brenowitz,M., Anderson,V. and Pyle,A. (2002) Productive folding to the native state by a Group II intron ribozyme. *J. Mol. Biol.*, **315**, 297–310.
20. Qin,P.Z. and Pyle,A.M. (1997) Stopped-flow fluorescence spectroscopy of a group II intron ribozyme reveals that domain 1 is an independent folding unit with a requirement for specific Mg^{2+} ions in the tertiary structure. *Biochemistry*, **36**, 4718–4730.
21. Lehmann,K. and Schmidt,U. (2003) Group II introns: structure and catalytic versatility of large natural ribozymes. *Crit. Rev. Biochem. Mol. Biol.*, **38**, 249–303.
22. Michel,F., Umesono,K. and Ozeki,H. (1989) Comparative and functional anatomy of group II catalytic introns—a review. *Gene*, **82**, 5–30.
23. Toor,N., Hausner,G. and Zimmerly,S. (2001) Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. *RNA*, **7**, 1142–1152.
24. Peebles,C.L., Perlman,P.S., Mecklenburg,K.L., Petrillo,M.L., Tabor,J.H., Jarrell,K.A. and Cheng,H.-L. (1986) A self-splicing RNA excises an intron lariat. *Cell*, **44**, 213–223.
25. Schmelzer,C. and Schweyen,R.J. (1986) Self-splicing of group II introns *in vitro*: mapping of the branch point and mutational inhibition of lariat formation. *Cell*, **46**, 557–565.
26. van der Veen,R., Arnberg,A.C., van der Horst,G., Bonen,L., Tabak,H.F. and Grivell,L.A. (1986) Excised Group II introns in yeast mitochondria are lariats and can be formed by self-splicing *in vitro*. *Cell*, **44**, 225–234.
27. Michel,F. and Ferat,J.-L. (1995) Structure and activities of group II introns. *Annu. Rev. Biochem.*, **64**, 435–461.
28. Fedorova,O., Julie Su,L. and Pyle,A.M. (2002) Group II introns: highly specific endonucleases with modular structures and diverse catalytic functions. *Methods*, **28**, 323–335.
29. Michels,W.J.Jr and , Pyle,A.M. (1995) Conversion of a group II intron into a new multiple-turnover ribozyme that selectively cleaves oligonucleotides: elucidation of reaction mechanism and structure/function relationships. *Biochemistry*, **34**, 2965–2977.
30. Koch,J.L., Boulanger,S.C., Dib-Hajj,S.D., Hebbar,S.K. and Perlman,P.S. (1992) Group II introns deleted for multiple substructures retain self-splicing activity. *Mol. Cell. Biol.*, **12**, 1950–1958.
31. Podar,M., Dib-Hajj,S. and Perlman,P.S. (1995) A UV-induced Mg^{2+}-dependent cross-link traps an active form of Domain 3 of a self-splicing group II intron. *RNA*, **1**, 828–840.
32. Fedorova,O., Mitros,T. and Pyle,A.M. (2003) Domains 2 and 3 interact to form critical elements of the group II intron active site. *J. Mol. Biol.*, **330**, 197–209.
33. Griffin,E.A.Jr, Qin,Z., Michels,W.J.Jr and Pyle,A.M. (1995) Group II intron ribozymes that cleave DNA and RNA linkages with similar efficiency, and lack contacts with substrate 2′-hydroxyl groups. *Chem. Biol.*, **2**, 761–770.
34. Kunkel,T.A., Bebenek,K. and McClary,J. (1991) Efficient site-directed mutagenesis using uracil-containing DNA. *Methods Enzymol.*, **204**, 125–139.
35. Pyle,A.M. and Green,J.B. (1994) Building a kinetic framework for group II intron ribozyme activity: quantitation of interdomain binding and reaction rate. *Biochemistry*, **33**, 2716–2725.
36. Wincott,F., DiRenzo,A., Shaffer,C., Grimm,S., Tracz,D., Workman,C., Sweedler,D., Gonzalez,C., Scaringe,S. and Usman,N. (1995) Synthesis, deprotection, analysis and purification of RNA and ribozymes. *Nucleic Acids Res.*, **23**, 2677–2684.
37. Huang,Z. and Szostak,J.W. (1996) A simple method for 3′-labeling of RNA. *Nucleic Acids Res.*, **24**, 4360–4361.
38. Pyle,A.M., McSwiggen,J.A. and Cech,T.R. (1990) Direct measurement of oligonucleotide substrate binding to wild-type and mutant ribozymes from *Tetrahymena*. *Proc. Natl Acad. Sci. USA*, **87**, 8187–8191.
39. Emerick,V.L. and Woodson,S.A. (1994) Fingerprinting the folding of a group I precursor RNA. *Proc. Natl Acad. Sci. USA*, **91**, 9675–9679.
40. Pan,J. and Woodson,S.A. (1998) Folding intermediates of a self-splicing RNA: mispairing of the catalytic core. *J. Mol. Biol.*, **280**, 597–609.
41. Boudvillain,M. and Pyle,A.M. (1998) Defining functional groups, core structural features and inter-domain tertiary contacts essential for group II intron self-splicing: a NAIM analysis. *EMBO J.*, **17**, 7091–7104.
42. Boudvillain,M., Delencastre,A. and Pyle,A.M. (2000) A new RNA tertiary interaction that links active-site domains of a group II intron and anchors them at the site of catalysis. *Nature*, **406**, 315–318.
43. Costa,M. and Michel,F. (1995) Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J.*, **14**, 1276–1285.
44. Fedorova,O. and Pyle,A.M. (2005) Linking the group II intron catalytic domains: tertiary contacts and structural features of domain 3. *EMBO J.*, in press.
45. Qin,P.Z. and Pyle,A.M. (1998) The architectural organization and mechanistic function of group II intron structural elements. *Curr. Opin. Struct. Biol.*, **8**, 301–308.
46. Adamidi,C., Fedorova,O. and Pyle,A.M. (2003) A group II intron inserted into a bacterial heat-shock operon shows autocatalytic activity and unusual thermostability. *Biochemistry*, **42**, 3409–3418.
47. Matsuura,M., Saldanha,R., Ma,H., Wank,H., Yang,J., Mohr,G., Cavanagh,S., Dunny,G.M., Belfort,M. and Lambowitz,A.M. (1997) A bacterial group II intron encoding reverse transcriptase, maturase, and DNA endonuclease activities: biochemical demonstration of maturase activity and insertion of new genetic information within the intron. *Genes Dev.*, **11**, 2910–2924.
48. Jarrell,K.A., Peebles,C.L., Dietrich,R.C., Romiti,S.L. and Perlman,P.S. (1988) Group II intron self-splicing: alternative reaction conditions yield novel products. *J. Biol. Chem.*, **263**, 3432–3439.
49. Konforti,B.B., Abramovitz,D.L., Duarte,C.M., Karpeisky,A., Beigelman,L. and Pyle,A.M. (1998) Ribozyme catalysis from the major groove of group II intron domain 5. *Mol. Cell*, **1**, 433–441.
50. Jacquier,A. and Michel,F. (1987) Multiple exon-binding sites in class II self-splicing introns. *Cell*, **50**, 17–29.
51. Harris-Kerr,C.L., Zhang,M. and Peebles,C.L. (1993) The phylogenetically predicted base-pairing interaction between α and α′ is required for group II splicing *in vitro*. *Proc. Natl Acad. Sci. USA*, **90**, 10658–10662.
52. DeLencastre,A., Hamill,S. and Pyle,A.M. (2005) A single active-site region for a group II intron. *Nature Struct. Mol. Biol.*, **12**, 626–627.
53. Qin,P.Z. and Pyle,A.M. (1999) Antagonistic substrate binding by a group II intron ribozyme. *J. Mol. Biol.*, **291**, 15–27.
54. Konforti,B.B., Liu,Q. and Pyle,A.M. (1998) A map of the binding site for catalytic domain 5 in the core of a group II intron ribozyme. *EMBO J.*, **17**, 7105–7117.
55. Huang,H.R., Rowe,C.E., Mohr,S., Jiang,Y., Lambowitz,A.M. and Perlman,P.S. (2005) The splicing of yeast mitochondrial group I and group II introns requires a DEAD-box protein with RNA chaperone function. *Proc. Natl Acad. Sci. USA*, **102**, 163–168.
56. Yang,Q. and Jankowsky,E. (2005) ATP- and ADP-dependent modulation of RNA unwinding and strand annealing activities by the DEAD-box protein DED1. *Biochemistry*, **44**, 13591–13601.
57. Fairman,M.E., Maroney,P.A., Wang,W., Bowers,H.A., Gollnick,P., Nilsen,T.W. and Jankowsky,E. (2004) Protein displacement by DExH/D 'RNA helicases' without duplex unwinding. *Science*, **304**, 730–734.
58. Jankowsky,E., Gross,C.H., Shuman,S. and Pyle,A.M. (2000) The DExH protein NPH-II is a processive and directional motor for unwinding RNA. *Nature*, **403**, 447–451.
59. Daniels,D.L., Michels,W.J.Jr and Pyle,A.M. (1996) Two competing pathways for self-splicing by group II introns: a quantitative analysis of *in vitro* reaction rates and products. *J. Mol. Biol.*, **256**, 31–49.
60. 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. *et al.* (2002) Rapid compaction during RNA folding. *Proc. Natl Acad. Sci. USA*, **99**, 4266–4271.
61. Pan,T. and 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.
62. Rangan,P., Masquida,B., Westhof,E. and Woodson,S.A. (2004) Architecture and folding mechanism of the *Azoarcus* Group I Pre-tRNA. *J. Mol. Biol.*, **339**, 41–51.
63. Perez-Salas,U.A., Rangan,P., Krueger,S., Briber,R.M., Thirumalai,D. and Woodson,S.A. (2004) Compaction of a bacterial group I

ribozyme coincides with the assembly of core helices. *Biochemistry*, **43**, 1746–1753.
64. Fang,X., Littrell,K., Yang,X.J., Henderson,S.J., Siefert,S., Thiyagarajan,P., Pan,T. and Sosnick,T.R. (2000) Mg^{2+}-dependent compaction and folding of yeast tRNAPhe and the catalytic domain of the *B.subtilis* RNase P RNA determined by small-angle X-ray scattering. *Biochemistry*, **39**, 11107–11113.
65. Russell,R., Millett,I.S., Doniach,S. and Herschlag,D. (2000) Small angle X-ray scattering reveals a compact intermediate in RNA folding. *Nature Struct. Biol.*, **7**, 367–370.
66. Shcherbakova,I., Gupta,S., Chance,M.R. and Brenowitz,M. (2004) Monovalent ion-mediated folding of the *Tetrahymena thermophila* ribozyme. *J. Mol. Biol.*, **342**, 1431–1442.