Structure of a Folding Intermediate Reveals the Interplay Between Core and Peripheral Elements in RNA Folding

Nathan J. Baird¹,², Eric Westhof³, Hong Qin⁴, Tao Pan⁴* and Tobin R. Sosnick²,⁴*

¹Department of Chemistry, University of Chicago, Chicago IL 60637, USA
²Institute for Biophysical Dynamics, University of Chicago, Chicago, IL 60637 USA
³Institut de biologie moléculaire et cellulaire, UPR9002, CNRS Université Louis Pasteur Strasbourg Cedex, F-67084 France
⁴Department of Biochemistry and Molecular Biology, University of Chicago, Chicago IL 60637, USA

*Corresponding authors

Keywords: modeling; P RNA; S-domain

Introduction

As a result of the hierarchical nature of RNA structure, intermediates populate at significant levels in RNA folding.¹,² The folding transition to the native structure often occurs in a highly cooperative manner, which can be identified by monitoring structural or spectroscopic features that change as a function of Mg²⁺ concentration.³,⁴ A general perception is that the folding intermediates are rich in secondary structure (e.g. Watson–Crick helices) but lack extensive amounts of tertiary contacts. In this view, the transition to the native structure would involve only the organization of the pre-formed helical domains into a compact architecture. This notion of “secondary structure first, tertiary structure later” is applicable to the Mg²⁺-dependent, equilibrium folding of small tertiary RNAs such as yeast tRNA⁵ and the hammerhead and hairpin ribozymes.⁶,⁷

The folding of larger RNAs is likely to be more complex. Structural information of key folding intermediates at the nucleotide level of resolution is required to investigate the interplay of secondary and tertiary structure formation in these large systems. Such structural detail, however, is sparse, as compared to the tremendous amount of information available on native RNA structures. Because partially folded structures are dynamic and unlikely to be crystallized, other strategies must be employed to obtain the detailed structural information for better understanding of the forces mediating the assembly and stability of large RNAs.

Here, we investigate the equilibrium folding of the specificity domain, or S-domain, of the Bacillus subtilis RNase P RNA. The crystal structure contains four tertiary structural modules: a rigid core (core); a four-way junction; a GAAA tetraloop (TL)–receptor; and an unusual motif involving two tertiary interacting loops (J11/12.
and J12/11) (Figure 1). In the thermodynamic folding pathway, an intermediate, termed \( I_{eq} \), populates. A priori, it is unclear which of the native secondary and tertiary modules are present in the folding intermediate.

We derive a nucleotide-level model of the intermediate using chemical and nuclease mapping, circular dichroism (CD) spectroscopy, small-angle X-ray scattering (SAXS) and molecular modeling. The intermediate contains two of the four native tertiary modules. It lacks the core, which results in many of the structural elements extending outwards. Folding from this intermediate to the native structure involves the formation of the core in concert with large-scale motions that bring two peripheral helices into proximity to form the tetraloop–receptor interaction. The coupling of core and peripheral structure formation rationalizes the high degree of cooperativity observed in the folding of this RNA.

**Results**

**The S-domain architecture**

The core of the S-domain is composed of the U-shaped, T-loop motif\(^8,9\), located at the end of the P10.1 helix, two coaxially stacked short helices (P10 and P11) and two bulged adenosine bases: A229 intercalates between G133 and C134, while A230 stacks on A130 of P9 (Figure 1). The four-way junction is composed of helices P7–P10. The J11/12 module is composed of two large internal loops, termed J11/12 and J12/11, located between helices P11 and P12. Each of the two loops forms a T-loop motif, and the nucleotide bases within these loops interact extensively through non-Watson–Crick base-pairs. The TL–receptor module is composed of a GAAA tetraloop (P12) interacting with a canonical tetraloop receptor embedded in P10.1.

**Mg\(^{2+}\)**-dependent folding transition

Equilibrium folding is monitored as a function of increasing concentration of Mg\(^{2+}\) using CD at 260 nm and 275 nm (Figure 2). The signal at 260 nm first increases and then decreases. This behavior is indicative of a sequential, three-state folding transition (unfolded-to-intermediate-to-native, \( U \)-to-\( I_{eq} \)-to-\( N \)), similar to the folding behavior of other tertiary RNAs.\(^10\) At 275 nm, only the \( I_{eq} \)-to-\( N \) transition is observed. This transition corresponds to the folding to the native structure, as confirmed by structural mapping (see below). The midpoint of the \( I_{eq} \)-to-\( N \) transition, \( K_{Mg} \), is 0.23(±0.03) mM with a Hill constant of 4.8±0.3 (Figure 2(b)).

The thermodynamic \( I_{eq} \) state represents the ensemble of RNA structures present at the beginning of the \( I_{eq} \)-to-\( N \) transition, under the following conditions; 20 mM Tris–HCl (pH 8), 0.1 mM MgCl\(_2\) and <0.5 \( \mu \)M S-domain RNA at 37 °C. This intermediate has properties distinct from those of both the unfolded and the native populations. The CD signal at 260 nm for the intermediate under the above conditions is 4.0 cm\(^2\) mmol\(^{-1}\), whereas the CD signals of both the unfolded and native
states are lower, 2.9 cm$^2$ mmol$^{-1}$ and 3.3 cm$^2$ mmol$^{-1}$, respectively. No combination of unfolded and native signals can produce the signal of the intermediate. Therefore, the thermodynamic intermediate is not a combination of the unfolded and native states.

To determine the surface area buried in the Ieq-to-N folding transition, Mg$^{2+}$ titrations are carried out in the presence of 0–3 M urea. The accompanying decrease in stability (Figure 2(b)) has a linear dependence on the concentration of urea:

$$DG([\text{urea}]) = DG(0) + m[\text{urea}]$$

where the slope ($m$-value) is proportional to the amount of surface area buried in the transition. $K_{M_{\text{G}}} and n$

were used to calculate the free energy of the N relative to Ieq according to:

$$DG_{\text{IN}}([\text{Mg}^{2+}]) = -nRT \ln([\text{Mg}^{2+}]/K_{M_{\text{G}}})$$

The $m$-value for this transition is 1.5(±0.1) kcal mol$^{-1}$ M$^{-1}$ (Figure 2(b)), which corresponds to the amount of surface area buried in a 22 base-pair RNA duplex.$^4$

Characterization of Ieq

The extent of solvent exposure of nucleotide bases and ribosephosphate backbone as a function of [Mg$^{2+}$] is determined using three chemicals and two nucleases to obtain structural information at the nucleotide level (Figure 3).$^{12,13}$ Dimethyl sulfate (DMS) modifies N1 of adenosine and N3 of cytidine; kethoxal (KE) modifies N1 and N2 of guanosine; and diethyl pyrocarbonate (DEPC) modifies N7 of adenosine.$^{14}$ Nuclease T$_1$ cleaves the backbone to guanosine and nuclease V$_1$ cleaves the backbone of stacked nucleotides, particularly those in RNA helices.$^{12}$

The extent of reactivity for selected residues monitored by structural mapping agrees well with the overall folding transition monitored by CD (Figure 4(a)). Some nucleotides change structure only in the U-to-Ieq transition, while other residues change structure only in the Ieq-to-N transition. Additional residues exhibit changes in protection in both transitions (Figure 3).

The overall protection pattern indicates that two structural modules are formed largely in the Ieq intermediate: the helices involved in the four-way junction (P7–P10) and the tertiary interactions of the structure involving J11/12 and J12/11 loops (residues 185–197 and 216–225). The helices in the four-way junction are formed largely in the U-to-Ieq transition, although P9 undergoes a slight rearrangement in the native transition. The extent of nuclease protection for the J11/12 and J12/11 loops reveals that this non-canonical module is formed in the Ieq intermediate (Figure 4(b)). The core of the S-domain, however, does not form until the Ieq-to-N transition. The formation of the basal T-loop of P10.1 (U175–A179) in the core is particularly striking in the Ieq-to-N transition (Figure 4(c)). Lastly, while the P12 and P10.1 helices are present in Ieq, the TL–receptor interaction between them is not made until the Ieq-to-N transition (Figure 4(d)).

The size and shape of the native and Ieq intermediate structures are determined by SAXS.$^{15}$ The size is characterized by the radius of gyration ($R_g$), whereas the global shape is derived from the pair-distance distribution function, $P(r)$, which is the sum of all mass pair distances in the molecule. Both the $R_g$ and the $P(r)$ functions for the native structure obtained from SAXS are in good agreement with those calculated from the crystal structure (Figure 4(e), inset). $R_g$ for the Ieq intermediate is 41.1(±1.1) Å, which is 8.8(±1.2) Å.
larger than the $R_g$ of the experimentally measured native structure. $P(r)$ analysis indicates that $I_{eq}$ has a more extended shape, with a maximum dimension of 135 Å, versus 110 Å for the native structure (Figure 4(e), continuous line). As expected, the $U$ state is much more extended than either of the other two states (Figure 4(e)).

**Molecular modeling**

The site-resolved structural mapping patterns along with the size and shape information were applied as constraints to generate molecular models of the $I_{eq}$ intermediate (Figure 5) using the modeling software MANIP. The modeling process included both global and local arrangements of modules and specific nucleotides. The crystal structure of the S-domain serves as a starting point for developing a model of the intermediate.

In the first stage, the structural mapping data are used to identify the modules that may be rearranged and those that should be left largely...
intact. Accordingly, the J11/12 module is kept native-like along with most of the helices in the four-way junction, except P9, which undergoes a local rearrangement in the Ieq-to-N transition. The helices involved in the TL–receptor interaction are formed largely in Ieq, although the TL–receptor interaction itself is absent.

In the second stage, the native structure is divided into modular pieces near the core to allow for movement and adjustment of the individual helices using the modeling program. In order to generate a structure with an increased \( R_g \) value \( \Delta R_g \) that approaches the experimentally determined \( \Delta R_g \) value between the native and the Ieq structure \((8.8(\pm 1.2) \text{ Å} \text{ or } 7.6–10.0 \text{ Å})\), selected tertiary interactions are removed:

1. The TL–receptor interaction between helices P10.1 and P12 is broken; \( R_g \) increases by 1.6 Å (Figures 5 and 6(a)).
2. P10.1 is further rotated while leaving the rest of the molecule in its native conformation; \( R_g \) increases by a maximum of 4.7 Å (Figures 5 and 6(a)).
3. To obtain an additional increase in \( R_g \), the four-way junction and the arm containing P11-J11/12-P12 (P12 arm) are adjusted. In the four-way junction, P9 is sterically unhindered but

Figure 4. Structural information for Ieq. (a) Ieq-to-N transition monitored by KE modification at nucleotide 111 (○), 112 (●), 113 (∇), or 116 (◇) and by DMS modification at nucleotide 114 (□) or 117 (◮) (upper panel). Transition monitored by digestion with nuclease V1 (lower panel) at nucleotide 132 (●) or 137 (◇). The overlay of the CD fit at 275 nm (—) in both panels demonstrates the agreement of the spectroscopic and biochemical probes. (b) T1 cleavage of the J11/12 region in the U, Ieq, and N states shows protection in the U-to-Ieq transition. (c) DMS (upper) and KE (lower) modification of the T-loop in the core reveal strong protections only in N compared to U and I. (d) DEPC modification of the GAAA tetraloop shows dramatic protection in the native state. Reference lanes for (b)–(d) are: —, no nuclease; OH, partial alkaline hydrolysis; U, A, C, G, sequencing ladder. [MgCl₂] for U, I, and N in (b)–(d): U (0 mM), I (0.5 × \( K_{Mg} \)), N (>5 × \( K_{Mg} \)). (e) Pair-distribution functions determined by SAXS (⋯, N; —,Ieq; •–•–•–•–•–•, U) reveal compaction of Ieq state. \( R_g \) calculated from the Ieq model (—) is shown. The inset shows the overlay of native state \( P(r) \) functions from SAXS (⋯) and from the crystal structure (—).

Figure 5. Sequential Ieq modeling using \( R_g \) as a constraint. (i) The TL–receptor interaction was disrupted and P10.1 was rotated away from its native position. This rotation achieved an increase in \( R_g \) of 1.6 Å, as compared to the native structure. (ii) P10.1 pivoted to a larger angle increased \( R_g \) to 4.7 Å. (iii) The P12 arm was adjusted to a \( \Delta R_g \) value similar to that obtained via SAXS experiments for the Ieq-to-N transition. The four-way junction was rotated commensurately to accommodate this movement. The refined structure of Ieq is the furthest on the right and has \( \Delta R_g = 7.6 \text{ Å} \) as compared to the native structure. Prominent features include the extended P10.1 and P12 helices, the folded non-canonical J11/12 module, and the expanded nucleotides U175–A179 in the core.
its rotation leaves $R_g$ unchanged (Figure 6(a)). Therefore, further modeling does not include any adjustments of P9. The position of P10.1 alongside P7 and P8 of the four-way junction precludes the movement of most of the junction. The P12 arm, however, is adjustable, and its movement increases $\Delta R_g$ to 9.5 Å, within the experimental range for $\Delta R_g$ (8.8(±1.2) Å).

In summary, in order to achieve a $\Delta R_g$ value that approaches the experimental value, modular rearrangements in accordance with the mapping data require that both P10.1 and P12 extend in opposite directions. Following adjustment of the backbone dihedral angles to values observed in crystal structures, the structure has $\Delta R_g = 7.6$ Å (Figure 5), which is at the lower end of the experimental values.

The possible alternative structures with this value of $\Delta R_g$ are very limited. The only plausible option involves increasing the amount of single-stranded regions at the expense of helical structures, for example, P10.1. Nuclease protection patterns argued against this possibility. Hence, the global arrangement with P10.1 and P12 splayed out in opposite directions is the major structural event consistent with the SAXS and structural mapping data.

The reorientation of the P10.1 and J11/12-P12 regions from their native positions results in rearrangement of other areas. Various linking sequences are adjusted manually and oriented appropriately in accordance with structural mapping data. Particular emphasis is devoted to nucleotides U175–A179 at the base of P10.1, and A231–C234 mediating the P10/P11 stack.

On the basis of the arrangement of these linking regions and the proximity of various nucleotide bases, two non-native base-pairs are introduced into the model, G176-C134 and A135-U175, consistent with structural mapping data. Additionally, the movement of P11 disrupts the A130/A230 stacking interaction and the intercalation of A229 between G133 and C134, which occur in the native structure. Due to the structural arrangement in the Ieq, we model A229 as stacking inside the P11 helix between A231 and G182. Minor adjustments are made in the orientation of the four-way junction to accommodate the rearrangement of linking regions with the core in order to relieve steric constraint.

**Validation of the model**

The Ieq model is examined in detail for further consistency against the global and site-specific experimental information including shape (SAXS), surface area burial (urea-dependence), and site-resolved information (nuclease and chemical mapping). The extended nature of both P10.1 and P12 is necessary to obtain the experimental value of $\Delta R_g$ (Figures 5 and 6(a)). Comparing $P(r)$ functions of the Ieq calculated from the model and from SAXS (Figure 4(e)) indicates that the shape of the model approximates that measured by experiment.

The slight dearth of long vectors for the model suggests that the model could be improved by adjustments, e.g. straightening P12 out even further. This adjustment would increase the $R_g$ by ~1 Å. Perturbations around the present orientation of P10.1 result in only a slight change in $R_g$ as the angle is varied (Figure 6(b)). Although the Ieq model is quite extended, various regions may be flexible enough to allow for further increase in $R_g$. Such regions may include the loop E motif in the middle of P10.1 (residues 139–142 and 166–170), which has a strong dependence on Mg$^{2+}$ for stabilization,11,17 and the area surrounding the bulged nucleotides in P11. In both of these regions, only a minimal amount of structural mapping information could be resolved. Hence, our model may be best viewed as a lower bound on the extended nature of the intermediate.

Another global validation of the model is the urea $m$-value, which correlates with the change in
accessible surface area in the \(I_{eq}\)-to-N transition. The total amount of surface area buried is roughly proportional to the number of nucleotides in the RNA, giving a total \(m\)-value for the S-domain of 5.4 \(\text{kcal mol}^{-1} \text{M}^{-1}\). The \(I_{eq}\) model has \(\sim 70\%\) of secondary and tertiary contacts present in the native structure. Hence, 30\% of the surface area burial occurs in the \(I_{eq}\)-to-N transition, giving a predicted \(m\)-value (1.6 \(\text{kcal mol}^{-1} \text{M}^{-1}\)) similar to that determined experimentally (1.5 \(\text{kcal mol}^{-1} \text{M}^{-1}\)).

Site-specific protection patterns also support the structure of the intermediate. The T-loop (U175–A179) at the base of P10.1 is extended in the \(I_{eq}\) structure (Figure 7(a)). This open conformation allows for increased modification by chemical probes as compared to the compact, H-bonded native structure. A177 and A178 are largely solvent-exposed in the \(I_{eq}\) structure but in the native conformation, they interact with base-pairs G90-C235 and G132-C234 via an A-minor motif. Consistently, DMS and KE modify this region strongly in both the U and I states but not in the native structure (Figure 4(c)).

**Discussion**

This work describes the identification of the extensive tertiary interactions in the major thermodynamic folding intermediate. The majority of the S-domain molecules adopt this structure at a solution condition located between two distinct folding transitions, U-to-\(I_{eq}\) and \(I_{eq}\)-to-N. The same folding transitions were observed when monitored by chemical and nuclease mapping, CD spectroscopy and SAXS.

**Features of the \(I_{eq}\) structure and the \(I_{eq}\)-to-N transition**

A general expectation is that RNA folding intermediates largely lack tertiary structures. However, we find that the J11/12-J12/11 junction is formed in \(I_{eq}\) despite the fact that this structure is composed entirely of tertiary interactions. This region yielded strong protection factors when probed with nucleases \(T_1\) and \(V_1\) in the U-to-\(I_{eq}\) transition, but protection is largely absent from the \(I_{eq}\)-to-N transition. Despite the presence of extensive tertiary interactions, however, this loop–loop interaction can still be considered to be local because the loops are the exit strands of the same hairpin. This result suggests that the topology and connectivity of the tertiary modules may dictate the order of structure formation in the equilibrium folding of an RNA.

A striking implication of the model is that the P10.1 helix must significantly reorient upon folding to the native structure. This reorientation occurs upon the formation of the U-shaped T-loop, which is fully extended in the intermediate (Figure 7(a) and (b)). This T-loop, formed by U175–A179, is present at the junction connecting two tertiary structure modules, the central core and the TL–receptor. The \(I_{eq}\)-to-N transition involves the rearrangement of this extended five nucleotide region to a U-shaped turn (Figure 7). Formation of the turn swings the P10.1 helix into a native-like position.

The T-loop region plays a pivotal role in the folding transition through the formation of the U175–A179 reversed Watson–Crick base-pair. The formation of this base-pair enables A177 and A178 to form an A-minor motif with the P7/P10 stack in the core. This docking event, in concert with the formation of the U175–A179 pair, orients the P10.1 helix properly for its TL–receptor (P12-P10.1) interaction. Additionally, the bulged G168 in the P10.1 helix is positioned for stacking on A194 of J11/12. The TL–receptor and the interaction of these bulged, stacked bases serve as latches holding this upper module in the native conformation.

Our model also implies structural changes in the
four-way junction and the core occur in the I_{eq}-to-N transition. In the I_{eq} structure, the P8/P9 stack is slightly tilted away from the junction where P7–P10 converge. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orients P8 closer to the inverted P10.1, in agreement with the increase in chemical convergence. This tilt orient...
non-local events, is likely to be a major part of the origin of the folding cooperativity observed for tertiary RNAs.

The action of metal ion binding in a core region and stabilizing a structural adjustment of other regions appears to be general, being observed in the folding of the hammerhead ribozyme and the four-way junction of the hairpin ribozymes.\textsuperscript{7,22} In each of these ribozymes, a metal ion binds near the core junction of the helices, thus promoting the correct orientation of the peripheral elements so that their apical loops are positioned for enzymatic activity.

The detailed structural model of the S-domain can serve as a foundation for investigating other questions of the I\textsubscript{eq}-to-N transition. Folding kinetic studies from U to N indicate the appearance of the same I\textsubscript{eq} state, followed by two additional intermediates (N.B., T.P. & T.R.S., unpublished results). The most conspicuous structural elements that are absent from I\textsubscript{eq} are the peripheral TL–receptor and the T-loop in the core. Modeling both kinetic intermediates that follow I\textsubscript{eq} will provide structural insights into the movements involved in the rate-limiting step. One possibility is that the formation of the core precedes, and is held in place by the TL–receptor interaction.

**Concluding remarks**

The equilibrium folding of the S-domain from the intermediate to the native state is a multifaceted process that involves movements of the four-way junction, the core, the TL–receptor, and base-stacking interactions. The structure of the intermediate hints at the impending intricate network of interactions to be formed in the folding transition to the native structure. While full atomic-level simulations of RNA folding processes for large RNAs are currently unfeasible, a more reasonable aim is to understand the thermodynamic and kinetic properties of the folding transition that leads to the native structure.\textsuperscript{23} However, to date, the necessary high-resolution models of intermediate states have not been available to achieve these goals.\textsuperscript{24} Our nucleotide-level model provides detailed structural information required for future studies, such as calculating the stability of the native structure using methods such as non-linear Poisson–Boltzmann formalism.\textsuperscript{23,24}

**Materials and Methods**

**Sample preparation**

The S-domain was transcribed by \textit{in vitro} transcription using phage T7 RNA polymerase,\textsuperscript{25} purified on denaturing polyacrylamide gels, and stored in water at \(-20^\circ\text{C}\). The unfolded (U) state was prepared by heating in 20 mM Tris–HCl (pH 8.1) at 85 °C for 2 min followed by 5 min at room temperature.

**Mg\textsuperscript{2+}-titration monitored by circular dichroism**

To RNA (0.3 mM) in the U state, Mg\textsuperscript{2+} solutions were added using a Hamilton titrator connected to a Jasco J715 spectrometer.\textsuperscript{11} The waiting time was up to 60 min due to slow folding kinetics (unpublished results), and the data acquisition time was 30 s.

The two Mg\textsuperscript{2+}-dependent transitions observed at CD\textsubscript{270} were described according to a cooperative binding model:

\[
U \overset{K_{UI}}{\leftrightarrow} I_{eq} \overset{K_{IN}}{\leftrightarrow} N
\]

where \( n \) and \( K \) are the Hill coefficient and Mg\textsuperscript{2+}-midpoint of the transitions. The signal, \( S \), is fit for simplicity with \( K_{IN} \) and \( n \) denoted as \( K_{eq} \) and \( n \) according to:

\[
S = \frac{S_u + S_{eq}((\text{Mg}^{2+})/K_{eq})^{n_u} + S_N((\text{Mg}^{2+})/K_{eq})^{n_u}((\text{Mg}^{2+})/K_{eq})^{n_N}}{1 + ((\text{Mg}^{2+})/K_{eq})^{n_u} + ((\text{Mg}^{2+})/K_{eq})^{n_N}}
\]

(1)

The CD\textsubscript{275} trace has one transition and was fit to:

\[
S = \frac{S_u + S_{eq}((\text{Mg}^{2+})/K_{eq})^{n_u}}{1 + ((\text{Mg}^{2+})/K_{eq})^{n_u}}
\]

(2)

**Mg\textsuperscript{2+}-titration monitored by partial digestion by nucleases T\textsubscript{1} and V\textsubscript{1}**

A sample of 0.3 mM (5′\textsuperscript{32P})-labeled S-domain in the U state was mixed with MgCl\textsubscript{2} in individual tubes, each with a different [Mg\textsuperscript{2+}]. The mixture was incubated at 50 °C for 5 min to allow a rapid approach to equilibrium. Samples were then returned to room temperature for ~5 min before the addition of nuclease. Nuclease reaction was performed at 37 °C for 5 min at 0.1 unit/μl of T\textsubscript{1} and 0.32 units/μl of V\textsubscript{1}. The solutions were mixed with an equal volume of 9 M urea, 100 mM EDTA and loaded immediately onto denaturing polyacrylamide gels. After electrophoresis, the gels were quantified using a phosphorimager. For nuclease V\textsubscript{1}, increased cleavage was interpreted as an increase in structure formation. Decreases in cleavages by V\textsubscript{1} could be a result of either structural disorganizations or inaccessibility of the nuclease to the nucleotides due to structural steric.

**Mg\textsuperscript{2+}-titration monitored by chemical modification**

A sample of 0.3 mM RNA in the U state was mixed with MgCl\textsubscript{2} in individual tubes, each at a different [Mg\textsuperscript{2+}]. The mixture was incubated at 37 °C for 10 min. Reactions were initiated by adding 1 μl of DMS (from 400× concentrated stock) or 1 μl of kethoxal (KE) (from 400× concentrated stock) and performed at 37 °C for 5 min. Reactions were quenched by adding 40 μl of stopping mixture (for DMS: 62.5 mM β-mercaptoethanol, 1 μg of \textit{Escherichia coli} tRNA, 45 mM potassium acetate, 180 mM KCl; for KE: 75 mM sodium acetate (pH 6.0), 6.25 mM boric acid, 1 μg of \textit{E. coli} tRNA, 45 mM potassium acetate, 180 mM KCl). RNA was then precipitated twice with ethanol, and stored in water. Modification was detected by reverse transcription using AMV reverse transcriptase and a (5′,32P)-labeled primer complementary to nucleotides 221–239 of the S-domain (5′ GCGAGGGGTTTACCGCGTT). Modification by DEPC was carried out as described.\textsuperscript{26}

Protection factors for chemical and nuclease mapping experiments were calculated from data obtained from the
phosphorimager. Based on CD titration data, certain lanes were chosen where populations of U, Ieq, and N states would dominate. Protected regions are defined as having a protection factor of >1.5 among U and Ieq or Ieq and N. Exposures are defined as regions having a protection factor of <0.67.

**SAXS**

Experiments were performed at the BioCAT beamline at the Advanced Photon Source at the Argonne National Laboratory. Samples (0.3 μM) at various [Mg²⁺] (20 mM Tris–HCl, pH 8.1) were incubated at 50 °C for 10 min, followed by incubation at room temperature for 5 min. The sample was loaded into a capillary tube controlled by a Hamilton titrator thermostatically controlled at 37 °C. Data collection was performed under continuous-flow conditions to avoid radiation damage.

The equilibrium SAXS measurements covered a range of [Mg²⁺] from 0–10 mM. The 0.4 mM MgCl₂ condition best correlated with the Ieq state, as indicated by nuclease mapping at this [RNA] (data not shown). This condition has [Mg²⁺] about fourfold higher than that obtained from CD or structural mapping, due to the higher [RNA] used in SAXS (0.3 μM versus 0.3 μM). The observed shift, 0.3 mM MgCl₂ at 6.3 μM S-domain, was within the expected range (0.23–0.46 mM, assuming every two to four ribosephosphate molecules bind an extra magnesium ion²⁷–²⁹).

**Molecular modeling**

To model Ieq, secondary structure modules as well as individual nucleotide bases were manipulated using MANIP;¹⁶ and changes were refined with the restraint refinement program NUCLIN-NUCLSQ. Adjustment of the backbone dihedral angles to values observed in crystal structures is done automatically by NUCLIN-NUCLSQ. The process was followed as described.¹⁶,³¹ Calculations of P(r) functions from crystal structures and predicted models were conducted using the program XTL.⁵²

**Acknowledgements**

This work was supported by a NIH grant (GM57880). We thank the reviewers for their insightful comments. We thank Dr Jaby Jacob for dedicated assistance in collection of the SAXS data. Use of the Advanced Photon Source was supported by the US Department of Energy, Basic Energy Sciences, Office of Science, under contract no. W-31-109-ENG-38. BioCAT is a National Institutes of Health-supported Research Center RR-08630.

**References**


*Edited by J. Doudna*

(Received 17 May 2005; received in revised form 27 June 2005; accepted 5 July 2005)  
Available online 25 July 2005