

Modular Construction of a Tertiary RNA Structure: The Specificity Domain of the *Bacillus subtilis* RNase P RNA[†]

Hong Qin,[‡] Tobin R. Sosnick,^{‡,§} and Tao Pan^{*,‡}

Department of Biochemistry and Molecular Biology and Institute for Biological Dynamics, University of Chicago, Chicago, Illinois 60637

Received January 11, 2001; Revised Manuscript Received July 24, 2001

ABSTRACT: The structure of the specificity domain (S-domain) of the *Bacillus subtilis* RNase P RNA has been proposed to be composed of a core and a buttress module, analogous to the bipartite structure of the P4–P6 domain of the *Tetrahymena* group I ribozyme. The core module is the functional unit of the S-domain and contains the binding site for the T stem–loop of a tRNA. The buttress module provides structural stability to the core module and consists of a GA₃ tetraloop and its receptor. To explicitly test the hypothesis that modular construction can describe the structure of the S-domain and is a useful RNA design strategy, we analyzed the equilibrium folding and substrate binding of three classes of S-domain mutants. Addition or deletion of a base pair in the helical linker region between the modules only modestly destabilizes the tertiary structure. tRNA binding selectivity is affected in one but not in two other mutants of this class. Elimination of the GA₃ tetraloop–receptor interactions significantly destabilizes the core module and results in the loss of tRNA binding selectivity. Replacing the buttress module with that of a homologous RNase P RNA maintains the tRNA binding selectivity. Overall, we have observed that the linker regions between the two modules can tolerate moderate structural changes and that the buttress modules can be shuffled between homologous S-domains. These results suggest that it is feasible to design an RNA using a buttress module to stabilize a functional module.

The complexity of the RNA structure reflects the diverse function of cellular RNAs. The structure of large ribozymes often can be subdivided into domains. For example, the *Tetrahymena* group I ribozyme is composed of two structurally complex domains, P4–P6 and P3–P9 (1–3). The P4–P6 domain can fold independently from the remainder of the molecule (3). Similarly, bacterial RNase P RNAs are composed of two independently folding domains (4).

Can the structure of an RNA folding domain be considered to be composed of smaller structural modules? The structure of the P4–P6 domain of the *Tetrahymena* group I intron ribozyme has been proposed to have a bipartite scheme with a core module and a buttress module (5; Figure 1A). The core module binds the substrate P1 helix and forms a part of the active site of the group I ribozyme (5). The buttress module, proposed to be critical for the integrity of the structure of the core module (3), consists of a GA₃ tetraloop and its canonical [CCUAAG-UAUGG] receptor (Figure 1A). Although the two modules do not physically interact, the core module is no longer formed upon substitution of the GA₃ tetraloop with a UUCG tetraloop, a mutation that disrupts the tetraloop–receptor interactions (3). Hence, the tetraloop and its receptor constitute a structural module that buttresses the core module.

A similar modular construction can be envisioned for each of the folding domains of bacterial RNase P RNAs (4), a ribozyme required to produce the mature 5' end of all tRNAs (6, 7). Of the two domains of the *Bacillus subtilis* RNase P RNA (P RNA),¹ the specificity domain (S-domain) most closely resembles the P4–P6 domain of the *Tetrahymena* ribozyme (Figure 1B). In the S-domain, the core module contains a functional region that binds the T stem–loop of a pre-tRNA substrate. The potential buttress module consists of a GA₃ tetraloop and a tetraloop receptor (2, 8). Just as with the P4–P6 domain, the modules in the S-domain are connected through two linker regions (Figure 1C).

To test the general design principle that RNA domains are composed of a functional module stabilized by a buttress module, we have analyzed the equilibrium folding and the tRNA binding specificity of three classes of S-domain mutants (Figure 2). The first class may either alter the distance and the orientation between the modules or introduce conformational stress to the S-domain structure upon changes in a helical linker region between the modules. The second class disrupts the tetraloop–receptor interactions in the buttress module. The third class replaces the buttress module of the *B. subtilis* S-domain with a homologous buttress module from the S-domains of either *Bacillus brevis* or *Escherichia coli*. Our results indicate that the S-domain is

[†] This work was supported by a grant from the NIH (GM52993).

* Corresponding author. Tel: (773) 702-4179. Fax: (773) 702-0439.

E-mail: taopan@midway.uchicago.edu.

[‡] Department of Biochemistry and Molecular Biology, University of Chicago.

[§] Institute for Biological Dynamics, University of Chicago.

¹ Abbreviations: CD, circular dichroism; DEPC, diethyl pyrocarbonate; P RNA, the catalytic RNA component of the *Bacillus subtilis* RNase P; S-domain, the specificity domain of the P RNA consisting of nucleotides 86–239.

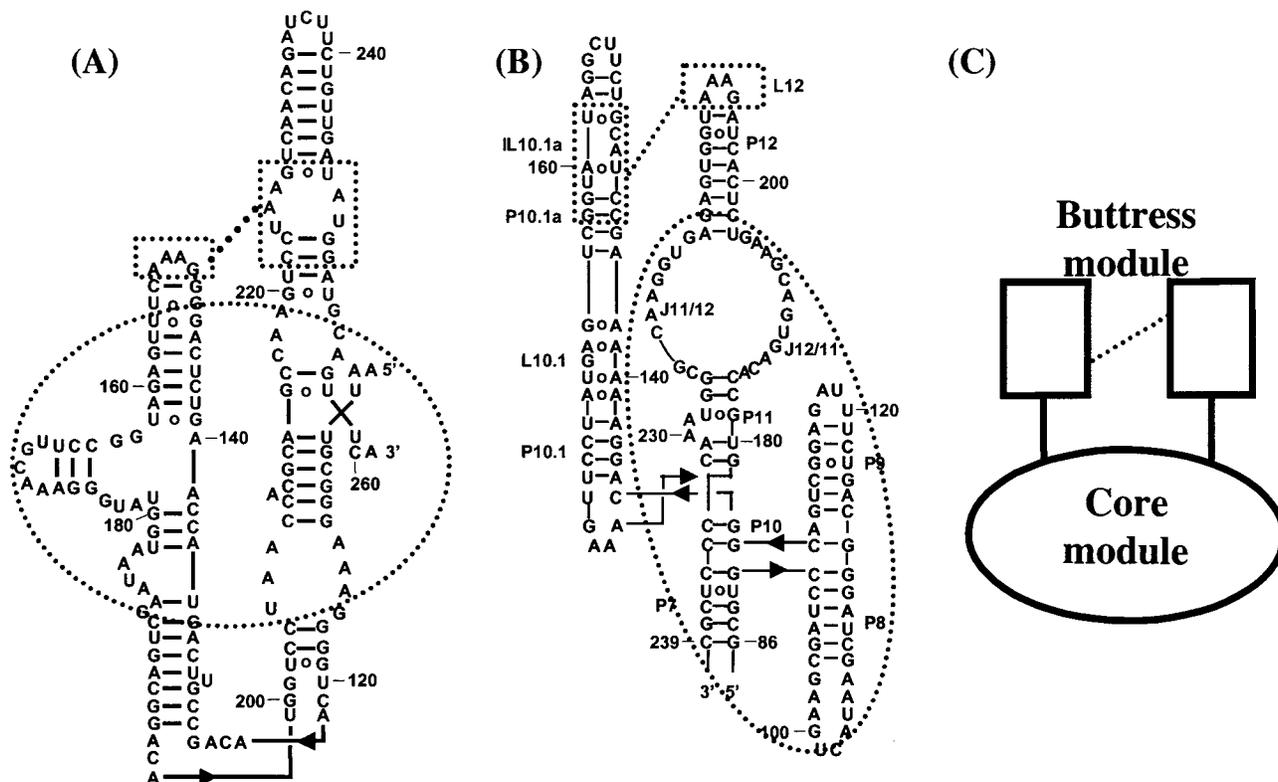


FIGURE 1: Modular construction model for a tertiary RNA structure. (A) The P4–P6 domain of the *Tetrahymena* group I ribozyme (5). (B) The specificity domain of the *B. subtilis* P RNA (13). The proposed butress module contains the L12 tetraloop and the internal loop within the P10.1a helix. (C) Schematic presentation.

functionally robust to modification of the linker region and the stabilizing butress module, but details of variation can affect the precise conformation of the core module and tRNA binding. Hence, modular construction appears to be a valid strategy in the design of a tertiary RNA domain.

MATERIALS AND METHODS

Plasmid Construction of the S-Domain Mutants. The wild-type S-domain (residues 86–239 of *B. subtilis* P RNA) with a T7 RNA polymerase promoter at the 5′ side and a *FokI* restriction site at the 3′ side was cloned into a pUC18 vector as previously described (4). The U86–A239 base pair was changed to G86–C239 to facilitate transcription. The mutants were generated by PCR amplification using the appropriate primers.

Equilibrium Folding Monitored by Circular Dichroism. The Mg^{2+} -dependent folding of the S-domain was quantitatively analyzed by circular dichroism (CD) as described (9). Briefly, RNA samples in 20 mM Tris-HCl, pH 8.1, were heated at 85 °C for 2 min and then left at room temperature for 3 min. A total amount of 450 pmol of denatured RNA was added to 1.5 mL of 20 mM Tris-HCl, pH 8.1, in a quartz cuvette. Mg^{2+} titration was controlled by a computer-interfaced, automated Hamilton titrator using a low and a high concentrated Mg^{2+} stock to ensure even distribution of data points. For each Mg^{2+} concentration, CD and UV absorbance at 260 and 278 nm were collected. The data points, CD signal versus $[Mg^{2+}]$, were adjusted for dilution.

Equilibrium Folding Monitored by DEPC Modification. Upon carboxymethylation of the N⁷ group at neutral pH,

adenosines become cleavage sites for aniline-induced scission (10). The 5′ ³²P-labeled S-domain variant at 0.3 μM was heated in 20 mM sodium cacodylate (pH 6.6) at 85 °C for 2 min, followed by incubation at 24 °C for 3 min. After the addition of various concentrations of $MgCl_2$, the RNA was incubated at 37 °C for 15 min, followed by incubation at 24 °C for another 15 min. Modification was initiated by addition of 1/20 volume of DEPC and terminated after 1 h by the addition of 100 μL of cold stop solution (4 μg of *E. coli* tRNA in 50 mM potassium acetate and 200 mM KCl). The samples were twice precipitated by ethanol and then vacuum-dried. The dried samples were suspended in 1 M aniline in 0.6 M sodium acetate, pH 4.5. The scission reaction was carried out by incubating the samples at 60 °C in the dark for 20 min and terminated by the addition of a cold solution of 50 mM potassium acetate and 200 mM KCl. The samples were then ethanol precipitated and dried. Finally, the samples were suspended in gel-loading buffer (9 M urea and 100 mM EDTA) and separated on denaturing polyacrylamide gels containing 7 M urea. The gels were scanned by a phosphor-imager (Molecular Dynamics) and the radioactive bands quantitated using ImageQuant software.

Data Analysis for CD and DEPC Modifications. We used a Hill-type analysis to quantitatively describe the Mg^{2+} -dependent folding of an RNA (11, 12). Folding of the S-domain was described as a three-state process with two folding transitions: U to I to N. Each transition could be observed by CD and had two associated parameters: the Mg^{2+} concentration at the midpoint of the transition, K_{Mg} , and the Hill coefficient, n . CD data were fit to linked

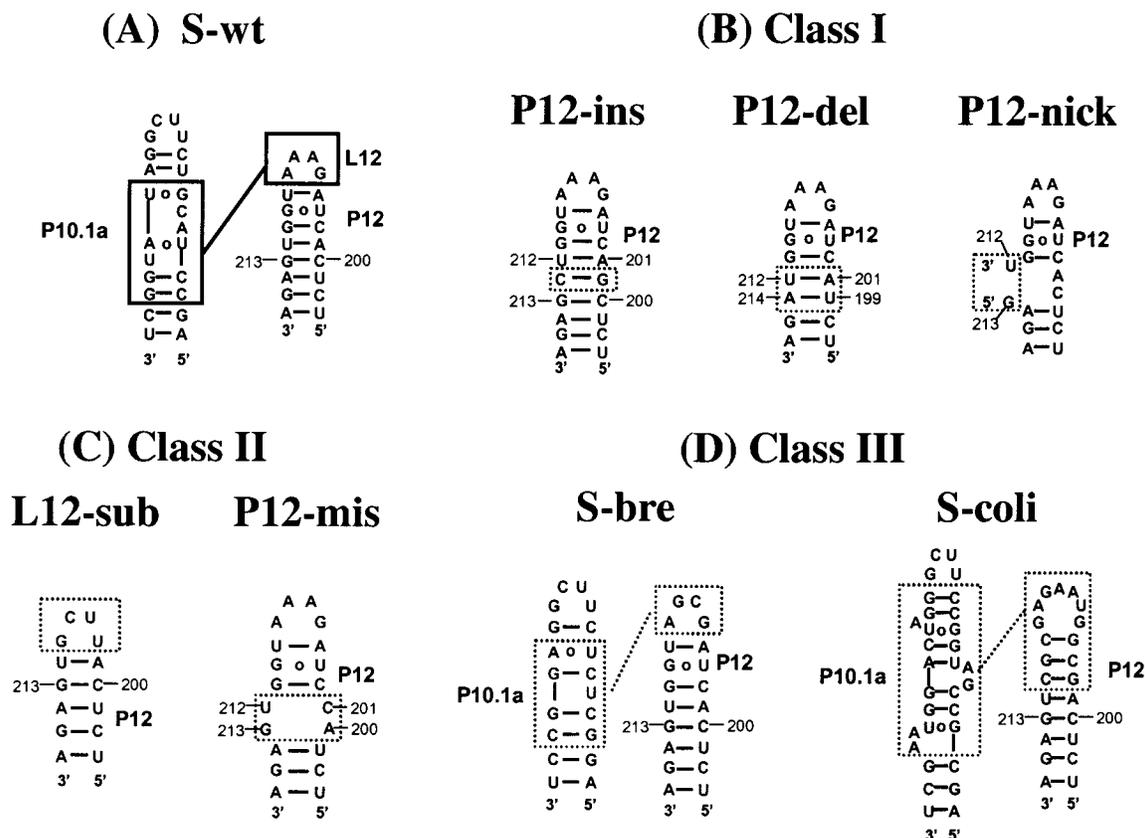


FIGURE 2: Design of the S-domain mutants. Only the buttress module of the S-domain is shown. The mutations are located within the dotted boxes. (A) The wild-type S-domain. The GA₃ tetraloop and its receptor are boxed. (B) Class I mutants. P12-ins has one base pair inserted in the P12 helix between the base pairs C200/G213 and A201/U212. P12-del has the C200/G213 base pair deleted. P12-nick is a circularly permuted S-domain, corresponding to a backbone break in the P12 helix. The 5' and 3' ends of the wild-type S-domain at G86 and C239 are linked by a UUCG loop. (C) Class II mutants. In L12-sub, the GA₃ tetraloop is replaced by a UUCG tetraloop, and the adjacent three base pairs are also deleted. In P12-mis, two adjacent nucleotides at 200 and 201 are switched to generate an internal loop. (D) Class III mutants. In S-bre and in S-coli, the L12 and IL10.1a from *B. subtilis* are replaced by their counterparts from *B. brevis* or *E. coli* RNase P RNAs, respectively.

equilibria:

$$\Delta\epsilon_{\text{obs}} = \frac{\Delta\epsilon_U + \Delta\epsilon_I \left(\frac{[\text{Mg}^{2+}]^{n_1}}{K_{\text{Mg1}}} \right) + \Delta\epsilon_N \left(\frac{[\text{Mg}^{2+}]^{n_1}}{K_{\text{Mg1}}} \right) \left(\frac{[\text{Mg}^{2+}]^{n_2}}{K_{\text{Mg2}}} \right)}{1 + \left(\frac{[\text{Mg}^{2+}]^{n_1}}{K_{\text{Mg1}}} \right) + \left(\frac{[\text{Mg}^{2+}]^{n_1}}{K_{\text{Mg1}}} \right) \left(\frac{[\text{Mg}^{2+}]^{n_2}}{K_{\text{Mg2}}} \right)} \quad (1)$$

where $\Delta\epsilon_{\text{obs}}$ is the observed signal intensity, $\Delta\epsilon_U$, $\Delta\epsilon_I$, and $\Delta\epsilon_N$ are the spectroscopic signals for the U, I, and N states, K_{Mg1} and K_{Mg2} are the Mg^{2+} concentrations at 50% change in CD signal for the transitions, and n_1 and n_2 are the Hill coefficients of the U to I and I to N transitions, respectively. Only the I to N transition could be unambiguously observed in DEPC modification. Hence, the DEPC modification was fit according to

$$\text{fraction protected} = \frac{a_I \left(\frac{[\text{Mg}^{2+}]^{n_1}}{K_{\text{Mg1}}} \right) + \left(\frac{[\text{Mg}^{2+}]^{n_1}}{K_{\text{Mg1}}} \right) \left(\frac{[\text{Mg}^{2+}]^{n_2}}{K_{\text{Mg2}}} \right)}{1 + \left(\frac{[\text{Mg}^{2+}]^{n_1}}{K_{\text{Mg1}}} \right) + \left(\frac{[\text{Mg}^{2+}]^{n_1}}{K_{\text{Mg1}}} \right) \left(\frac{[\text{Mg}^{2+}]^{n_2}}{K_{\text{Mg2}}} \right)} \quad (2)$$

where K_{Mg1} and n_1 are the parameters obtained from the CD

measurement and a_I is the contribution of the I state to the protection. The contributions of the U and N states to the protection extent are 0 and 1, respectively.

Binding of a Pre-tRNA Substrate by the S-Domain. Binding of a pre-tRNA^{Phe} substrate by the S-domain and mutants was determined using an inhibition assay. The reaction mixture contained a trace amount of a pre-tRNA substrate (~1 nM), 25 nM full-length P RNA, and varying amounts of the S-domain. The S-domain bound the pre-tRNA substrate and acted as a competitive inhibitor to cleavage by the full-length P RNA. Because the P RNA concentration was significantly below K_m under this condition (~0.2 μM), a simplified competitive inhibition equation was used to describe the normalized substrate cleavage rate as a function of the concentration of the S-domain and mutants:

$$k_{\text{norm}} = \frac{1}{1 + \frac{[\text{S-domain}]}{K_I}} \quad (3)$$

where k_{norm} is the first-order rate constant for substrate cleavage normalized to the rate constant without inhibitor and K_I is the inhibition constant of the S-domain for the pre-tRNA substrate used in the assay.

RESULTS

Design of the S-Domain Mutants. Bacterial RNase P RNAs can be divided into two types as represented by the *E. coli* RNase P RNA (A type) and the *B. subtilis* RNase P RNA (B type) (13). The S-domains from both types of RNase P RNAs show a conserved core structure but diversified peripheral regions (13). The peripheral regions in the B-type S-domain generally contain the L12 GA₃ tetraloop and the P10.1a region tetraloop receptor (Figure 1B). The peripheral regions in the A-type S-domain contain the L13 loop, generally a seven nucleotide loop of conserved sequence, and the internal loop in the P12 region, proposed to be the matching structure that interacts with the L13 loop. On the basis of phylogeny and hydroxyl radical protection data, we proposed that the S-domain of both A- and B-type P RNAs has a bipartite structural organization composed of a core module and a buttress module (Figure 1C; 13–15). The core module is composed of the conserved portion, while the buttress module is composed of the L12 and P10.1a regions in the B-type P RNAs and the L13 and P12 regions in the A-type P RNAs.

The S-domain of the *B. subtilis* P RNA, residues 86–239, has been shown to contact the T stem–loop of a pre-tRNA substrate (14, 16). The buttress module is connected to the core module through two linker regions, the P12 helix and an internal loop within the P10.1a region (IL 10.1a). In the B-type P RNAs, P12 always has eight base pairs, suggesting that the length of this helix may be critical for the distance and the orientation of the core and buttress modules.

The first class of mutants is designed to examine the importance of the invariant length and flexibility of the P12 linker region (Figure 2B). One base pair is either added (P12-ins) or deleted (P12-del) in P12 to change the number of base pairs from eight to nine or seven. The base pair insertion/deletion can change the distance by ~ 3.6 Å and the orientation by $\sim 34^\circ$ between the modules. However, if the GA₃ tetraloop and its receptor maintain their interaction, insertion and deletion of one base pair are likely to introduce conformational stress at P12 that could be passed on to other regions in the S-domain. The third mutant introduces a backbone break in P12 (P12-nick) in the context of a circularly permuted molecule. The termini in the wild-type S-domain at C239 and G86 are linked through a UUCG loop. This particular nick may allow more flexibility in the helical structure.

The second class of mutants is designed to eliminate the tetraloop–receptor interaction (Figure 2C). The GA₃ tetraloop is substituted with a UUCG tetraloop (L12-sub), a substitution that significantly destabilized the tertiary structure of the *Tetrahymena* P4–P6 domain (2). To completely abolish the interaction between the L12 tetraloop and IL10.1a region, three adjacent base pairs in P12 are also deleted in L12-sub. The two middle base pairs within the helix P12 are mismatched (P12-mis) to reduce the helical structure supporting the GA₃ tetraloop from eight to three base pairs. This mutation may have two potential effects: the mismatch either increases the flexibility of the P12 helix or destabilizes the P12 helix to the extent that it, and the adjoining tetraloop, would not form.

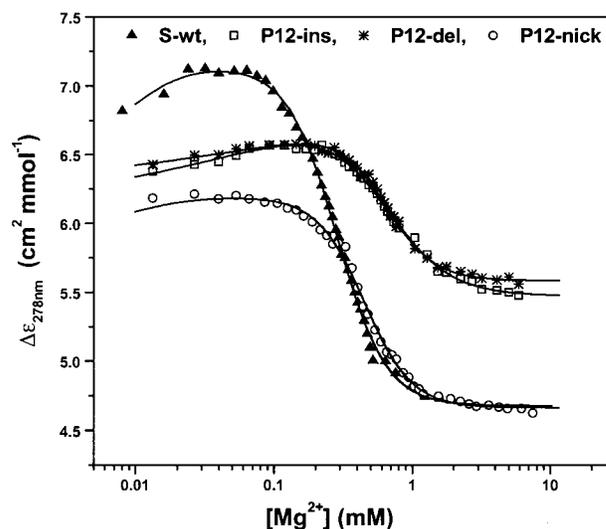


FIGURE 3: Equilibrium folding of class I mutants monitored by CD.

The third class of mutants is designed to examine the context dependence of the buttress module from various RNase P RNAs (Figure 2D). The *B. subtilis* buttress module is replaced with a different buttress module from a homologous B-type P RNA (S-bre) or with the buttress module from an A-type P RNA (S-coli). The *B. brevis* S-domain contains an unusual GCGA tetraloop and a corresponding receptor that is different from that of the *B. subtilis* S-domain. The *E. coli* S-domain contains a seven-nucleotide loop, 5'GUAA-GAG, and its corresponding binding site, an internal loop (13).

Equilibrium Folding. Two assays are used to determine the effect of these mutations on the equilibrium folding of the S-domain. Circular dichroism provides a quantitative measurement on the folding of the overall structure (Figure 3), whereas DEPC modification provides residue-specific information on folding (Figure 4). Like several other RNAs analyzed to date (11, 12, 17, 18), two folding transitions are observed by CD. These transitions can be described by a simple two-step folding scheme, U to I to N. Using a Hill-type analysis, each folding transition is characterized by a Mg²⁺ midpoint, K_{Mg} , and a Hill coefficient, n . If the I to N transitions of the wild type and the mutants have identical Hill constants, the change in the stability of the mutants correlates with the changes in K_{Mg} and $\Delta\Delta G = -nRT \ln [K_{Mg}(\text{wild-type})/K_{Mg}(\text{mutant})]$ (9).

According to the CD data, the Mg²⁺-dependent folding of the class I mutants where a base pair in the stem is inserted or deleted has the identical Hill constant and approximately 2-fold higher K_{Mg} compared to the folding of the wild-type S-domain (Table 1, Figure 3). These results indicate that the class I mutants fold into structures similar to that of the wild-type S-domain and the overall structure is minimally destabilized (< 1.2 kcal/mol).

The CD-monitored folding of the class II mutants where the tetraloop–receptor interaction is eliminated requires significantly higher Mg²⁺ concentration (Table 1). Furthermore, these mutants may not even have an overall structure similar to that of the wild-type S-domain at 10 mM Mg²⁺ (Figure 4C). The disruption of the buttress module destabilizes the S-domain structure by at least 5.2 kcal/mol, and perhaps the core module even adopts a different structure.

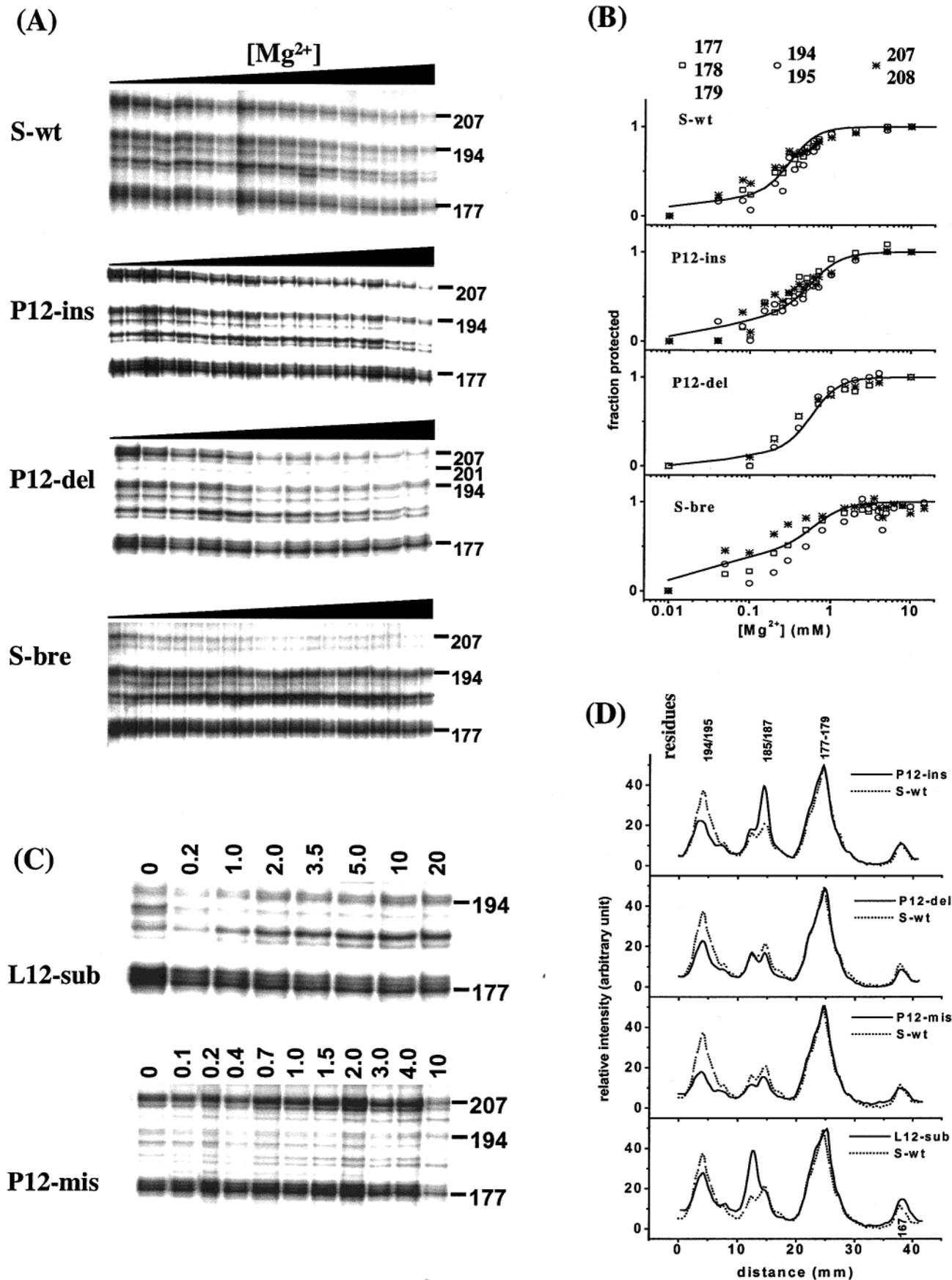


FIGURE 4: Equilibrium folding monitored by DEPC modification at nucleotides 177–179, 194/195, and 207/208. (A) In S-wt, P12-ins, P12-del, and S-bre, protection occurs as a function of Mg^{2+} concentration. Modification at A201 occurs in P12-del but not in P12-ins. (B) Fraction folded as a function of Mg^{2+} concentration for S-wt, P12-ins, P12-del, and S-bre. (C) L12-sub and P12-mis exhibit minimal protection as a function of Mg^{2+} concentration. The protection at the tetraloop (207/208) is clearly lost in P12-mis. (D) Comparison of the relative modification of various mutants to the wild-type S-domain.

Table 1: Equilibrium Folding of the S-Domain and Mutants

| class | RNA | method ^a | K_{Mg} (mM) | n | $\Delta\Delta G$ (kcal/mol) ^b |
|----------|---------|---------------------|-----------------|-----------------|--|
| I | S-wt | CD | 0.30 ± 0.02 | 2.6 ± 0.1 | |
| | | DEPC | 0.3 ± 0.1 | 2.5 ± 0.4 | |
| | P12-ins | CD | 0.68 ± 0.02 | 2.5 ± 0.1 | 1.1 ± 0.2 |
| | | DEPC | 0.6 ± 0.2 | 2.2 ± 0.6 | |
| | P12-del | CD | 0.68 ± 0.02 | 2.5 ± 0.1 | 1.2 ± 0.2 |
| | | DEPC | 0.7 ± 0.3 | 2.3 ± 0.5 | |
| P12-nick | CD | 0.46 ± 0.01 | 2.6 ± 0.1 | 0.5 ± 0.2 | |
| | DEPC | 0.4 ± 0.2 | 2.0 ± 1.0 | | |
| II | L12-sub | CD | >10 | ND ^c | >5.2 |
| | | DEPC | >10 | ND | |
| | P12-mis | CD | >10 | ND | >5.2 |
| DEPC | | >10 | ND | | |
| III | S-bre | CD | 0.7 ± 0.1 | 2.0 ± 0.3 | 1.2 ± 0.2 |
| | | DEPC | 0.6 ± 0.2 | 1.9 ± 0.9 | |
| | S-coli | CD | 0.58 ± 0.04 | 2.5 ± 0.6 | 1.2 ± 0.2 |
| | | DEPC | 0.6 ± 0.3 | 2.4 ± 0.6 | |

^a Conditions: CD, 20 mM Tris-HCl, pH 8, 24 °C; DEPC, 20 mM sodium cacodylate, pH 6.6, 24 °C. CD data are averaged from changes at 260 and 278 nm. ^b $\Delta\Delta G = -nRT \ln[K_{Mg}(\text{wild type})/K_{Mg}(\text{mutant})]$. The average values from CD and DEPC ($n = 2.5$) are presented. ^c ND: not determined.

Compared to the wild-type S-domain, the equilibrium folding of the class III mutants where the buttress module has been substituted has ~ 2 -fold higher K_{Mg} and a similar Hill constant (Table 1). This result suggests that the buttress modules from homologous S-domains can substitute for each other without a significant loss in the stability of the overall RNA structure.

Folding of the S-domain and the mutants monitored by DEPC modification strongly depends on Mg^{2+} concentration (Figure 4A,B). In the wild-type S-domain, modifications of residues A177–179 and A194/195 in the core module and A207/208 in the buttress module are especially sensitive to Mg^{2+} . They are protected against chemical modification at the higher Mg^{2+} concentrations, and the extent of modification can be fit with a Hill-type equation to obtain K_{Mg} and n for the I to N transition (Figure 4B, Table 1). These two parameters obtained from DEPC modification are the same as those for the I to N transition measured by CD. This consistency of equilibrium folding parameters obtained from two different probes supports our interpretation that these two probes are monitoring the formation of the same structure.

Modification of the class II mutants, where the tetraloop interaction is eliminated, is insensitive to the change of Mg^{2+} concentration up to 10–20 mM (Figure 4C), indicating that these mutants are not correctly folded, consistent with the CD results.

Close inspection of the DEPC modification pattern for the less disruptive S-domain mutants at high Mg^{2+} concentration reveals subtle differences within the core module (Figure 4D). For example, residues A185 and A187 are modified equally in the wild-type S-domain but not in P12-ins and S-bre. Residue A201 is only significantly modified in P12-del, even though this residue is in the putative P12 helix. This particular modification suggests that the structure of the P12 region in P12-del may not be the same as that in other class I mutants. Because base mismatches in the P12 helix destroy S-domain folding, the subtle changes in this region may implicate that P12-del is structurally different from the other class I mutants (see below).

Binding of tRNA by the S-Domain and Mutants. An inhibition assay is used to determine the effect of these mutations on this known S-domain function. The assay tests the ability of the S-domain to compete with the full-length P RNA in the binding of a pre-tRNA substrate. The inhibition assay can only be performed at high ionic conditions due to the inherently weak binding between P RNA and the pre-tRNA substrate in the absence of the RNase P protein. The inhibition constant of the wild-type S-domain determined in this way is $1.4 \mu\text{M}$, ~ 500 -fold less than the binding constant of the full-length P RNA to a tRNA product under identical conditions (16, 19, 20).

To demonstrate that the inhibition constant correlates with the binding affinity of the S-domain to the pre-tRNA substrate, the inhibition of the S-domain to the cleavage of two mutant pre-tRNA substrates is measured (Figure 5). The conserved G19-C56 base pair in the wild-type tRNA is a direct contact site with the P RNA in the absence of the P protein (16, 19). Mutation of this base pair to either a19u56 or c19 g56 reduces the affinity of full-length P RNA binding by 11- and 19-fold, respectively (16). The wild-type S-domain has 4- and 13-fold weaker inhibition constants for these pre-tRNA mutants, indicating that the wild-type S-domain maintains nearly the same capability to distinguish different T stem–loop structures. This mutation study also suggests that, like its counterpart in the full-length P RNA, the wild-type S-domain directly interacts with the T stem–loop portion of the pre-tRNA.

Two of the three class I mutants, P12-ins and P12-nick, bind the pre-tRNA only 2-fold or less well compared to pre-tRNA binding by the wild-type S-domain. These two mutants show similar reduction in binding affinity ($\Delta\Delta G_{\text{bind}}$) to each of the two mutant pre-tRNA^{Phe} substrates, a19u56 and c19g56 (Table 2). These results indicate that inserting a base pair or nicking the P12 helix has little effect on S-domain function.

In contrast, the third class I mutant, P12-del, shows a different behavior in pre-tRNA binding in the inhibition assay (Table 2). Binding of the P12-del to the wild-type pre-tRNA decreases by 5–6-fold. Furthermore, mutations of the G19-C56 base pair in the substrate have little effect on the binding affinity of P12-del. These results suggest that deleting a base pair in the P12 helix produces a functionally different S-domain as compared to inserting a base pair or nicking the P12 helix.

Although the class II mutants are not able to fold at 10–20 mM Mg^{2+} , these mutants also inhibit the cleavage reaction by the full-length P RNA. Both the inhibition constant and the $\Delta\Delta G_{\text{bind}}$ of the class II mutants are strikingly similar to those of P12-del (Table 2). There are at least two possible explanations for the observed inhibition. First, it is possible that the class II mutants can fold to a functional structure at the high ionic conditions used in our assay (100 mM $MgCl_2$, 0.6 M KCl). In support of this proposal, many deletion mutants in the full-length P RNA were found to have catalytic activity similar to that of the wild-type P RNA, but only at elevated ionic conditions (21, 22). Second, an unstructured S-domain mutant may bind the pre-tRNA elsewhere to inhibit the cleavage reaction by the full-length P RNA. For example, it is possible that a region in this ~ 150 -residue RNA can interact with the single-stranded 5' leader and/or the 3' CCA region in pre-tRNA. This potential mutant S-domain–5' leader interaction would not respond to pre-

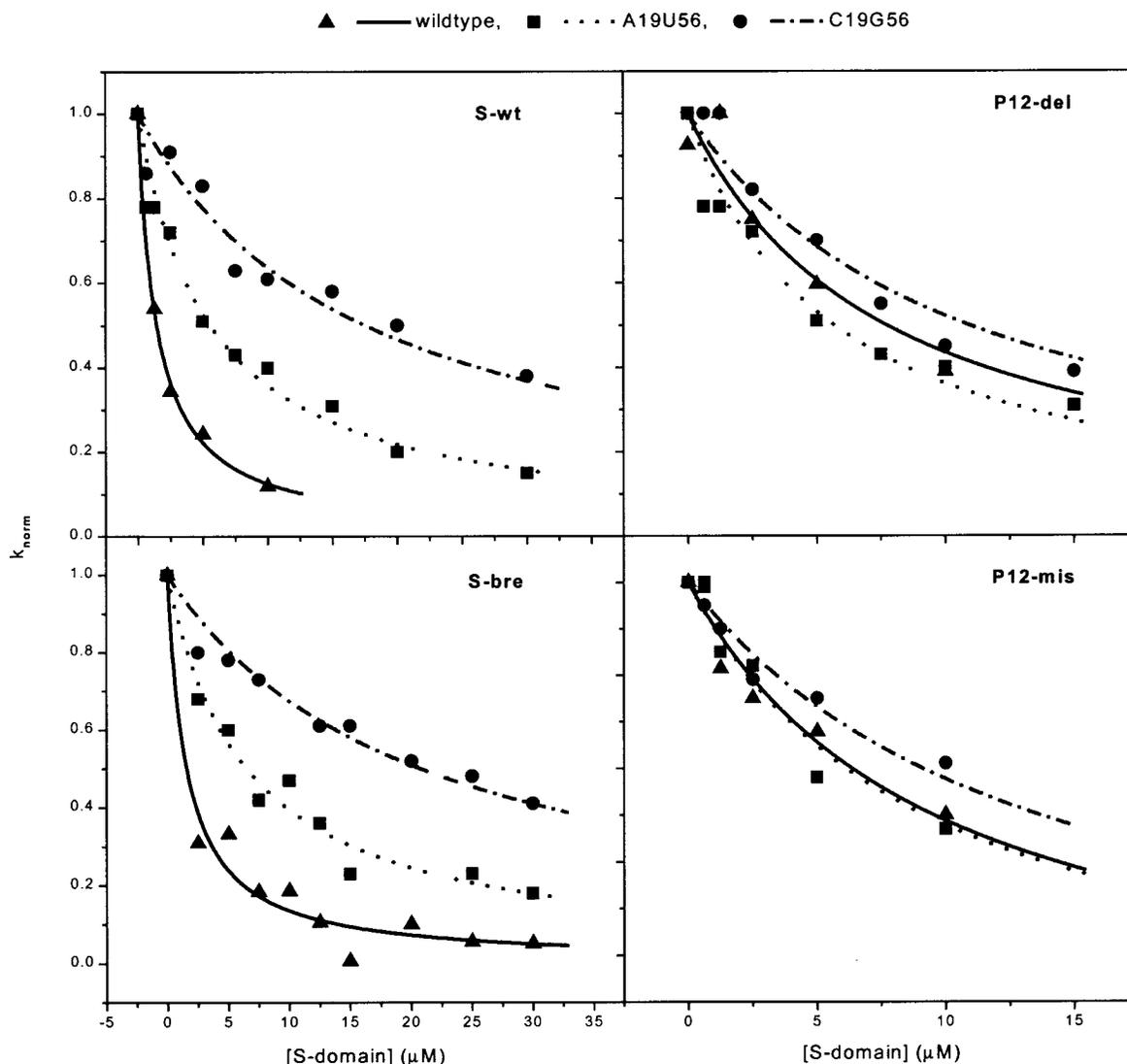


FIGURE 5: Binding of the S-domain to a pre-tRNA^{Phe} substrate monitored by the inhibition of P RNA cleavage. Three pre-tRNA^{Phe} substrates, wild type (G19C56) and the a19u56 and c19g56 mutants, are used to assess the selectivity of S-domain interaction with the pre-tRNA. k_{norm} is the ratio of the observed rate constant divided by the rate constant in the absence of the S-domain. The data are fit according to eq 3.

tRNA mutations at G19C56. Regardless of the detailed interpretation, these results do suggest that the class II mutants are impaired in tRNA binding.

Among the class III buttress module substitution mutants, S-bre, but not S-coli, binds pre-tRNA with an affinity and specificity similar to that of the wild-type S-domain, despite both mutants having similar stability (Figure 5 and Table 2). S-coli binds pre-tRNA much less well, even though the selectivity for the pre-tRNA mutants is still somewhat maintained. To determine whether class III mutants are also sensitive to changes in the helical linker region, a single base pair is inserted or deleted in the P12 helix of S-bre. Unlike the wild-type S-domain, both insertion and deletion mutants in S-bre have decreased affinity for pre-tRNA by >10-fold. Furthermore, these mutants show no difference in binding for the a19u56 and c19g56 mutations in the pre-tRNA (data not shown). Insertion or deletion of a single base pair in the P12 helix of S-coli does not increase the affinity for pre-tRNA binding (data not shown). These results suggest that a type B buttress module can be substituted with another type B module and, to some extent, with the type A module. However, within the limit of the inhibition assay, the

juxtaposition of the modules has a significant context dependence for the function of the S-domain.

DISCUSSION

Modular Construction of an RNA Domain. The present equilibrium folding and functional study confirms that modular construction is a useful strategy to build complex RNA structures. The key component for this scheme is the utilization of simple modules, e.g., a tetraloop and its receptor, placed in appropriate context, to stabilize a more complex functional structure. This strategy, initially suggested for the stabilization of rRNAs (23) and group I (2, 5, 24, 25) and group II ribozymes (26–28), is now extended to include a domain of a bacterial RNase P RNA.

Uniquely, the present work addresses the following issues: How crucial is the role of the linker region between a buttress module and a core module? Can different types of buttress modules substitute and stabilize the same core module? The linker region is altered by making a nick in the conserved P12 helix or changing the number of base pairs from 8 to 9 or 7. Nicking the P12 helix has little effect on the stability and the functionality of the S-domain. The

Table 2: Pre-tRNA Binding by the S-Domain and Mutants Determined by Inhibition

| class | S-domain | substrate ^a | K_i (μ M) | $\Delta\Delta G_{\text{bind}}$ (kcal/mol) ^b | $\Delta\Delta G_{\text{bind}}$ (kcal/mol) ^c | | |
|----------|----------|------------------------|------------------|--|--|-----|-----|
| I | S-wt | G19C56 | 1.4 \pm 0.1 | | | | |
| | | a19u56 | 5.6 \pm 0.5 | 0.8 | | | |
| | | c19g56 | 18 \pm 2 | 1.5 | | | |
| | P12-ins | G19C56 | 3.3 \pm 0.4 | | | 0.5 | |
| | | a19u56 | 7.2 \pm 0.9 | 0.5 | | | |
| | | c19g56 | 18 \pm 2 | 1.0 | | | |
| | | P12-del | G19C56 | 7.8 \pm 1.6 | | | 1.0 |
| | | | a19u56 | 5.2 \pm 0.7 | -0.2 | | |
| | | | c19g56 | 11 \pm 1 | 0.2 | | |
| P12-nick | G19C56 | 1.9 \pm 0.2 | | | 0.2 | | |
| | a19u56 | 4.7 \pm 0.2 | 0.5 | | | | |
| | c19g56 | 13 \pm 2 | 1.2 | | | | |
| II | L12-sub | G19C56 | 8 \pm 2 | | | 1.0 | |
| | | a19u56 | 8 \pm 1 | 0.0 | | | |
| | | c19g56 | 14 \pm 2 | 0.3 | | | |
| | P12-mis | G19C56 | 9.1 \pm 1.0 | | | 1.1 | |
| | | a19u56 | 9.4 \pm 1.6 | 0.0 | | | |
| | | c19g56 | 14 \pm 3 | 0.3 | | | |
| | III | S-bre | G19C56 | 1.6 \pm 0.2 | | | 0.1 |
| | | | a19u56 | 6.5 \pm 0.5 | 0.8 | | |
| | | | c19g56 | 21 \pm 1 | 1.5 | | |
| S-coli | | G19C56 | 21 \pm 2 | | | 1.6 | |
| | | a19u56 | 48 \pm 7 | 0.5 | | | |
| | | c19g56 | >60 | >0.6 | | | |

^a The wild-type pre-tRNA has G19-C56 interaction. This base pair is mutated to a19u56 or c19g56. ^b $\Delta\Delta G = -RT \ln[K_i(\text{wild-type substrate})/K_i(\text{mutant substrate})]$ for the same S-domain mutant. ^c $\Delta\Delta G = -RT \ln[K_i(\text{wild-type S-domain})/K_i(\text{mutant S-domain})]$ for the wild-type pre-tRNA substrate.

quantitative destabilization of the overall S-domain structure is similar in magnitude to the effect of making a nick in a contiguous helix (29).

We had expected that changing the number of base pairs in the P12 helix would have a large effect, because one base pair added or deleted corresponds to ~ 3.4 Å change in distance and $\sim 34^\circ$ in orientation between the two ends of a standard RNA helix. However, both deletion and insertion mutants destabilize the overall structure by only ~ 1 kcal/mol, much smaller than the destabilization caused by the elimination of the tetraloop interaction (>5 kcal/mol). Apparently, the S-domain structure is robust enough to tolerate the change of one base pair in the P12 helix.

We also demonstrate that the buttress module from homologous S-domains can substitute for each other in both folding and function. This result resembles a study of a group I ribozyme in which several matching tetraloop-receptor pairs are functionally equivalent (24, 25). In our limited sampling of two B-type modules and one A-type module, it appears that the B-type module from *B. subtilis* can be easily substituted by another B-type module from *B. brevis*. Substitution with an A-type module from *E. coli*, however, is less effective in this context. Because the A-type module has a significantly different secondary structure than a B-type module, the ineffectiveness of substituting the A-type module for a B-type S-domain may be due to our incomplete understanding of what exactly constitutes a buttress module in the A-type S-domain.

It is plausible that the sequence of an individual buttress module has been optimized by the evolutionary process to fit a specific S-domain context. We tried unsuccessfully to improve the stability and function of S-bre and S-coli through

addition and deletion of one base pair in the P12 linker. This attempt exposes our limited understanding on the detailed interaction between the buttress and the core module. At the moment, successful optimization may still require an in vitro selection approach using either a buttress module library with fixed linker regions or a linker library with a fixed buttress module.

Our inhibition assay indicates that the S-domain shows a moderate inhibition effect on the P RNA reaction. Possibly, there is more than one mode of interaction between the S-domain and the pre-tRNA substrates. Beside the T stem-loop, the S-domain may interact with other regions in the pre-tRNA structure such as the acceptor stem, the single-strand 5' leader, and/or the 3' CCA sequence. However, the interaction between the S-domain and the T stem-loop of the wild-type pre-tRNA appears to be the major interaction for the wild-type S-domain, P12-ins, P12-nick, and S-bre. The interaction between the S-domain and T stem is diminished in P12-del, L12-sub, P12-mis, and S-coli.

Although the buttress module acts as its name indicates, this nomenclature belies our observation that this module can influence the functionality of a folded core module. Conversely, it is unclear whether an isolated core module, if adequately stabilized (e.g., in 1 M Mg^{2+}), could fold and function by itself. Hence, the rigorous definition that the buttress module only serves to stabilize the core module is an oversimplification, albeit a useful one.

Stabilization of Complex RNA Structures by the Peripheral Regions. Large RNAs generally contain a conserved core structure and diversified peripheral regions. The peripheral regions of the group I ribozyme have been considered as a series of buttresses that stabilize the catalytic core. One of the peripheral structures, P5abc, facilitates the formation of the overall structure and is critical for the high cleavage activity of the ribozyme (30–35). The extensive studies on the group I and other ribozymes illustrate that an RNA enzyme with a stable structure is more effective than one with a dynamic structure, an idea first proposed for protein enzymes (reviewed in ref 36). This notion can account for the large size of many protein enzymes and some ribozymes. Many different peripheral structures can play the same stabilization roles, thus accounting for the diversity of the peripheral structures in biological RNAs.

Also consistent with the correlation of stability and catalytic efficiency, the extremely stable hepatitis delta ribozyme is much more efficient in catalysis than similarly small but flexible catalytic RNAs, such as the hammerhead ribozyme (37). The most active form of hammerhead enzyme is thought to be thermodynamically unstable and is only a transient structure in solution (37–41).

Propagation of Conformational Change and Implications for Design. Modular construction entails building a complex RNA structure with a core module, stabilized by one or more buttress modules. What is not straightforward, however, is how to link the buttress module(s) to the core module to achieve and to maximize stability.

We demonstrate here that the structure of the linker region can accommodate some moderate deformation. The length of the P12 helix in the B-type RNase P RNAs always has eight base pairs (42). A nick and even insertion of an extra base pair in P12 only modestly destabilizes the overall structure and has minimal effect on the function of the

S-domain. A deletion in this region also has modest effect on the stability of the S-domain. Therefore, some changes in the linker region are tolerated. The details of toleration, however, depend on the precise composition of the buttress module, because insertion of an extra base pair in S-bre severely affects tRNA binding.

One unanticipated result is the observed differential binding to pre-tRNA by the deletion and the insertion S-domain mutants. In the deletion mutant P12-del, P12 is overwound by 14%, or $\sim 5^\circ$ per base pair, while in the insertion mutant P12-ins, P12 is underwound by 11%, or $\sim 4^\circ$ per base pair. A typical plasmid DNA has a supercoiling density, σ , of -0.03 to -0.09 (43), and the base pair deletion and insertion are equivalent to σ values of $+0.14$ and -0.11 for the modified P12 helix, respectively. This simple calculation shows that both deletion and insertion introduce stress in the P12 helix. To maintain the helical structure, this stress must be passed on to other parts of the molecule.

The results from DNA studies (44) suggest that the stress generated by deletion and insertion in P12 is likely to be passed on to loop regions in the S-domain. Two loop regions in the S-domain may be particularly suited to absorb this stress: the internal loop, L10.1, which is a portion of the second linker region between the modules, and the J11/12 or the J12/11 loop, which is a portion of the core module.

Cut-and-Replace or Plug-and-Play Modularity. The organization of many RNA tertiary structures is governed by the arrangement of their helices (45). Tetraloops are always positioned at the end of a helix while tetraloop receptors are usually an integral part to a helix. Hence, the tetraloop-receptor interaction is a very efficient way to compact helices. A GNRA tetraloop and its receptor can be viewed as a plug and socket in the construction of an RNA tertiary structure, as indicated by their pervasiveness in RNA sequences (25).

Loops are predominant elements in biological functions of RNA. This work demonstrates that the helical length between the loops influences their conformation. Therefore, considerable attention should be given to the correct number of base pairs in a helix. The conformational stress generated within the helix is most likely absorbed by loops that may reside at functional sites of designed RNAs.

ACKNOWLEDGMENT

We thank Drs. C. Correll, T. Steck, and members of the Pan laboratory for helpful discussions. We also thank the reviewers for insightful comments on the manuscript.

REFERENCES

- Laggerbauer, B., Murphy, F. L., and Cech, T. R. (1994) *EMBO J.* 13, 2669–2676.
- Murphy, F. L., and Cech, T. R. (1994) *J. Mol. Biol.* 236, 49–63.
- Murphy, F. L., and Cech, T. R. (1993) *Biochemistry* 32, 5291–5300.
- Loria, A., and Pan, T. (1996) *RNA* 2, 551–563.
- Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science* 273, 1678–1685.
- Altman, S., and Kirsebom, L. (1999) in *The RNA World* (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) pp 351–380, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Frank, D. N., and Pace, N. R. (1998) *Annu. Rev. Biochem.* 67, 153–180.
- Jaeger, L., Michel, F., and Westhof, E. (1994) *J. Mol. Biol.* 236, 1271–1276.
- Sosnick, T. R., Fang, X., and Shelton, V. M. (2000) *Methods Enzymol.* 317, 393–409.
- Ehrenberg, L., Fedorcsak, I., and Solymosy, F. (1976) *Prog. Nucleic Acid Res. Mol. Biol.* 16, 189–262.
- Pan, T., and Sosnick, T. R. (1997) *Nat. Struct. Biol.* 4, 931–938.
- Fang, X., Pan, T., and Sosnick, T. R. (1999) *Biochemistry* 38, 16840–16846.
- Massire, C., Jaeger, L., and Westhof, E. (1998) *J. Mol. Biol.* 279, 773–793.
- Odell, L., Huang, V., Jakacka, M., and Pan, T. (1998) *Nucleic Acids Res.* 26, 3717–3723.
- Pan, T. (1995) *Biochemistry* 34, 902–909.
- Loria, A., and Pan, T. (1997) *Biochemistry* 36, 6317–6325.
- Shelton, V. M., Sosnick, T. R., and Pan, T. (1999) *Biochemistry* 38, 16831–16819.
- Pan, T., Fang, X., and Sosnick, T. (1999) *J. Mol. Biol.* 286, 721–731.
- Loria, A., and Pan, T. (1999) *Biochemistry* 38, 8612–8620.
- Beebe, J. A., and Fierke, C. A. (1994) *Biochemistry* 33, 10294–10304.
- Waugh, D. S., Green, C. J., and Pace, N. R. (1989) *Science* 244, 1569–1571.
- Darr, S. C., Zito, K., Smith, D., and Pace, N. R. (1992) *Biochemistry* 31, 328–333.
- Woese, C. R., Winker, S., and Gutell, R. R. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 8467–8471.
- Costa, M., and Michel, F. (1997) *EMBO J.* 16, 3289–3302.
- Costa, M., and Michel, F. (1995) *EMBO J.* 14, 1276–1285.
- Abramovitz, D. L., and Pyle, A. M. (1997) *J. Mol. Biol.* 266, 493–506.
- Costa, M., Deme, E., Jacquier, A., and Michel, F. (1997) *J. Mol. Biol.* 267, 520–536.
- Jestin, J. L., Deme, E., and Jacquier, A. (1997) *EMBO J.* 16, 2945–2954.
- Walter, A. E., Turner, D. H., Kim, J., Lyttle, M. H., Muller, P., Mathews, D. H., and Zuker, M. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91, 9218–9222.
- Engelhardt, M. A., Doherty, E. A., Knitt, D. S., Doudna, J. A., and Herschlag, D. (2000) *Biochemistry* 39, 2639–2651.
- Doherty, E. A., and Doudna, J. A. (1997) *Biochemistry* 36, 3159–3169.
- Williams, K. P., Fujimoto, D. N., and Inoue, T. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89, 10400–10404.
- Zheng, M., Wu, M., and Tinoco, I., Jr. (2001) *Proc. Natl. Acad. Sci. U.S.A.* 98, 3695–3700.
- Treiber, D. K., and Williamson, J. R. (2001) *J. Mol. Biol.* 305, 11–21.
- Mohr, G., Caprara, M. G., Guo, Q., and Lambowitz, A. M. (1994) *Nature* 370, 147–150.
- Narlikar, G. J., and Herschlag, D. (1997) *Annu. Rev. Biochem.* 66, 19–59.
- Doherty, E. A., and Doudna, J. A. (2000) *Annu. Rev. Biochem.* 69, 597–615.
- Wedekind, J. E., and McKay, D. B. (1998) *Annu. Rev. Biophys. Biomol. Struct.* 27, 475–502.
- Hertel, K. J., Peracchi, A., Uhlenbeck, O. C., and Herschlag, D. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94, 8497–8502.
- Scott, W. G., Finch, J. T., and Klug, A. (1995) *Cell* 81, 991–1002.
- Pley, H. W., Flaherty, K. M., and McKay, D. B. (1994) *Nature* 372, 68–74.
- Pace, N. R., and Brown, J. W. (1995) *J. Bacteriol.* 177, 1919–1928.
- Bauer, W. R. (1978) *Annu. Rev. Biophys. Bioeng.* 7, 287–313.
- Strick, T. R., Allemand, J. F., Bensimon, D., and Croquette, V. (1998) *Biophys J.* 74, 2016–2028.
- Hermann, T., and Patel, D. J. (1999) *J. Mol. Biol.* 294, 829–849.