The EMBO Journal Vol.17 No.5 pp.1515–1525, 1998

Comparative photocross-linking analysis of the tertiary structures of *Escherichia coli* and *Bacillus subtilis* RNase P RNAs

Jiunn-Liang Chen, James M.Nolan[1], Michael E.Harris[2] and Norman R.Pace[3]

Departments of Plant and Microbial Biology, and Molecular and Cell Biology, 111 Koshland Hall, University of California, Berkeley, CA 94720-3102, USA

[1]Present address: Department of Biochemistry, Tulane University Medical Center, New Orleans, LA 70112-2699, USA
[2]Present address: Department of Molecular Biology and Microbiology, School of Medicine, Case Western Reserve University, Cleveland, OH 44106-4960, USA

[3]Corresponding author
e-mail: nrpace@nature.berkeley.edu

Bacterial ribonuclease P contains a catalytic RNA subunit that cleaves precursor sequences from the 5′ ends of pre-tRNAs. The RNase P RNAs from *Bacillus subtilis* and *Escherichia coli* each contain several unique secondary structural elements not present in the other. To understand better how these phylogenetically variable elements affect the global architecture of the ribozyme, photoaffinity cross-linking studies were carried out. Photolysis of photoagents attached at homologous sites in the two RNAs results in nearly identical cross-linking patterns, consistent with the homology of the RNAs and indicating that these RNAs contain a common, core tertiary structure. Distance constraints were used to derive tertiary structure models using a molecular mechanics-based modeling protocol. The resulting superimposition of large sets of equivalent models provides a low resolution (5–10 Å) structure for each RNA. Comparison of these structure models shows that the conserved core helices occupy similar positions in space. Variably present helical elements that may play a role in global structural stability are found at the periphery of the core structure. The P5.1 and P15.1 helical elements, unique to the *B.subtilis* RNase P RNA, and the P6/16/17 helices, unique to the *E.coli* RNA, occupy similar positions in the structure models and, therefore, may have analogous structural function.

Keywords: *Bacillus subtilis*/photoaffinity cross-linking/ribozyme/RNase P/RNA structure

Introduction

Ribonuclease P (RNase P) is the site-specific endonuclease that processes precursor tRNAs (pre-tRNAs) to produce the mature 5′ ends of tRNAs. RNase P is a ribonucleoprotein consisting of a small protein subunit (~120 amino acids) and a large RNA subunit (~400 nucleotides) (Altman *et al*., 1993; Pace and Brown, 1995). The enzymatic activity of the bacterial RNase P resides in its RNA component, hence it is a ribozyme. The RNase P RNA requires either high ionic strength (*in vitro*) or its protein subunit (*in vivo*) for activity (Guerrier-Takada *et al*., 1983).

The specificity and catalytic activity of enzymes depend upon their three-dimensional structures. Consequently, we need to know the structure of RNase P RNA in order to understand its function fully. Since no crystal structure has yet been reported for the RNase P ribozyme, biochemical studies must be done to understand its three-dimensional structure better. We therefore used comparative site-specific photocross-linking and computer modeling to further our knowledge of RNase P RNA structure. Phylogenetic comparative sequence analysis provides a wealth of secondary and some tertiary structural information (Haas *et al*., 1994; Gautheret *et al*., 1995; Pace and Brown, 1995). Additional three-dimensional distance and positional data come from photoaffinity cross-linking experiments (Burgin and Pace, 1990; Nolan *et al*., 1993; Harris *et al*., 1994; Oh and Pace, 1994). This large data set provides constraints for modeling the tertiary structure of the RNase P RNA–tRNA complex.

The RNase P RNAs from different organisms are remarkably variable in structure (Darr *et al*., 1992; Pace and Brown, 1995). Some elements of sequence and secondary structure are absolutely conserved and are considered to be important for catalytic activity. Other structural elements vary in length or even occurrence. These phylogenetically variable elements seem to participate mainly in stabilizing the global structures of the RNAs (Siegel *et al*., 1996). In spite of structural differences, however, all RNase P RNAs are ‘homologous’ (of common ancestry) and consequently are expected to have a core of common architecture and function (Chen and Pace, 1997). Previous intermolecular cross-linking studies using a photoagent-conjugated tRNA revealed a similar pattern of cross-linking sites among the RNase P RNAs from *Bacillus subtilis*, *Chromatium vinosum* and *Escherichia coli* (Burgin and Pace, 1990; Oh and Pace, 1994). Additionally, chemical footprint analysis of tRNA bound to different bacterial RNase P RNAs gave similar results (LaGrandeur *et al*., 1994). These data indicate that nucleotides involved in substrate binding and catalysis consist primarily of regions that are conserved among different bacterial RNase P RNAs. Thus, any tertiary structure model should be applicable to different bacterial RNase P RNAs.

In this study, we extend the previous library of cross-linking constraints and use a phylogenetic perspective to enhance the interpretation of the data. To understand better the general three-dimensional structure of bacterial RNase P RNA and how it varies in evolution, we have subjected the *B.subtilis* RNA to intramolecular photocross-linking analysis. The *B.subtilis* and related RNase P RNAs have the most disparate secondary structure compared with the *E.coli* type among all bacterial RNase P RNAs known (Haas *et al*., 1996b). The cross-linking results complement

as well as extend the information obtained in the studies with *E.coli* RNA. The results are used to refine the working model of the tertiary structure of bacterial RNase P RNA.

Results

Catalytic properties of circularly permuted RNAs

We used circularly permuted (cp) RNase P RNAs, as described previously (Nolan *et al*., 1993), to manipulate the positions of 5′ ends for the specific attachment of the photoaffinity cross-linking agent azidophenacyl bromide. Fifteen different cpRNase P RNAs were made by PCR of a tandemly repeated RNase P RNA gene (Figure 1). CpRNAs retain the native RNA sequence but have an artificial loop that connects the native 5′ and 3′ ends, generating new 5′ and 3′ ends located elsewhere in the structure. Discontinuities in the phosphodiester backbone of RNase P RNA generally have little effect on substrate binding and catalytic activity (Waugh and Pace, 1993; Harris *et al*., 1994; Pan and Zhong, 1994), indicating that cpRNAs retain the native structure.

Instead of the native 5′ and 3′ ends, each cpRNA has a new 5′ end located on the nucleotide designated for attachment of the photoagent. For example, cpBs34 represents the circularly permuted *B.subtilis* RNA with its 5′ end at nucleotide 34 and its 3′ end at nucleotide 33. The endpoints for all 15 cpRNase P RNAs in this study were chosen for their strategic and homologous locations (Figure 1). Some endpoints were chosen to coincide with homologous positions in the *E.coli* RNA for which cross-linking data had been obtained (Harris *et al*., 1994). Other cpRNAs were made to gain structural information for phylogenetically variable regions such as helices P1 and P9, or species-specific regions, such as helices P5.1, P15.1, P10.1, P12 and P19, which occur in the *B.subtilis* RNA, but not in *E.coli* RNA. The 5′ end of each cpRNA was chosen to begin with one or more guanosine residues, the preferred initiation sequence of phage T7 RNA polymerase (Milligan and Uhlenbeck, 1989).

As shown in Table I, cpRNAs used in this study retain near-native activity under conditions of high ionic strength (3 M NH_4OAc). The cpRNAs require higher ionic strength for optimum activity than the native optimum of 1–2 M NH_4OAc (Chen, 1997). The nicked backbones of the cpRNAs conceivably result in greater flexibility, which can be overcome by increasing the ionic strength (Gardiner *et al*., 1985). The native-like K_M values for the cpRNase P RNAs suggest that they fold accurately into the native conformation under the assay conditions used. Similarly, cross-linking reactions were carried out at high concentrations of monovalent (3 M NH_4OAc) and divalent (75 mM $MgCl_2$) salts.

Intramolecular photoaffinity cross-linking

The cpRNAs were synthesized by transcription in the presence of guanosine 5′-monophosphorothioate (GMPS),

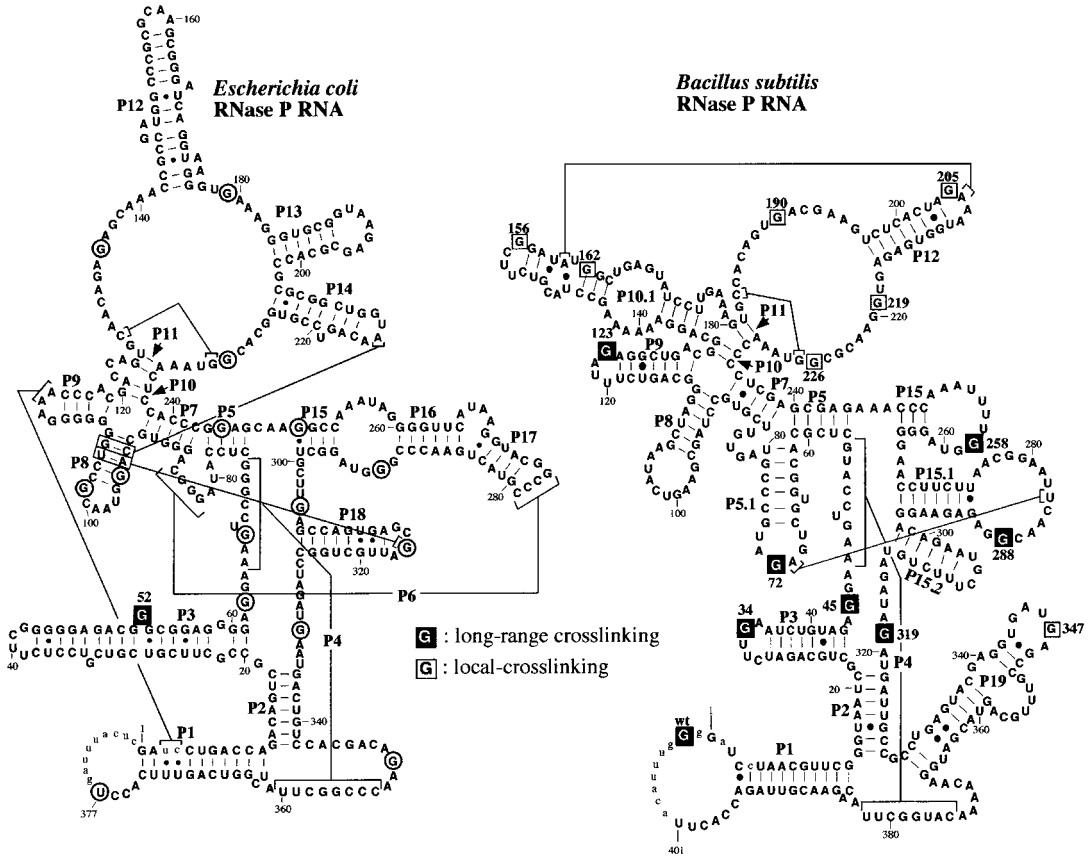

Fig. 1. Secondary structures and 5′ ends of cpRNAs. The numbered Gs in square boxes denote the 5′ ends of cpRNAs analyzed in this study. Photoagent attachment sites described previously (Harris *et al.*, 1994, 1997) are shown in open circles. Photoagents attached to Gs in filled boxes engage in long-range cross-linking, whereas photoagents attached to Gs in open boxes cross-link to local regions, the 3′ end or near the 3′ end of the cpRNAs. The artificial loops of cpRNAs between native 5′ and 3′ ends of *E.coli* and *B.subtilis* RNAs are shown in lower case letters. Known tertiary interactions are indicated by brackets connected by lines. Filled dots represent non-Watson–Crick base pairs.

which is incorporated during *in vitro* transcription only as the initiating nucleotide. The photoaffinity agent azidophenacyl bromide was then attached to the GMPS-initiated transcripts at the unique 5′-phosphorothioate (Burgin and Pace, 1990). After cross-linking under 302 nm UV irradiation, cpRNAs form circular molecules and lariats. Circular RNAs result from cross-linking between the 5′ end and a site near the 3′ end of the same molecule. The lariat RNAs result from cross-linking to sites more distant from the 3′ end (Harris *et al*., 1997). Both circular and lariat RNAs were separated from uncross-linked RNAs by denaturing polyacrylamide gel electrophoresis (Figure 2A). The concentration independence of the extent of cross-linking at different RNA concentrations (500 or 50 nM) indicates that the cross-linked species result from intramolecular cross-linking, rather than intermolecular interaction. The efficiencies of cross-linking generally range from 1 to 14% (Table II). Cross-link sites of lariats were identified by primer extension as previously described (Figure 2B), and results are compiled in Table II. Eight cpRNAs and native *B.subtilis* RNA formed lariats after cross-linking. Among them, only cpBs258 and cpBs319 RNAs cross-link to tRNA (Chen, 1997), which is consistent with the results of homologous *E.coli* cpRNAs (Harris *et al.*, 1994).

Overall, the results seen with the *B.subtilis* and *E.coli* cpRNase P RNAs are similar. The corroborative results, therefore, lend validity to cross-links previously established for the *E.coli* RNA. Results with the *B.subtilis* RNA also provide information that was not detected with the *E.coli* RNA analysis. For instance, as shown in

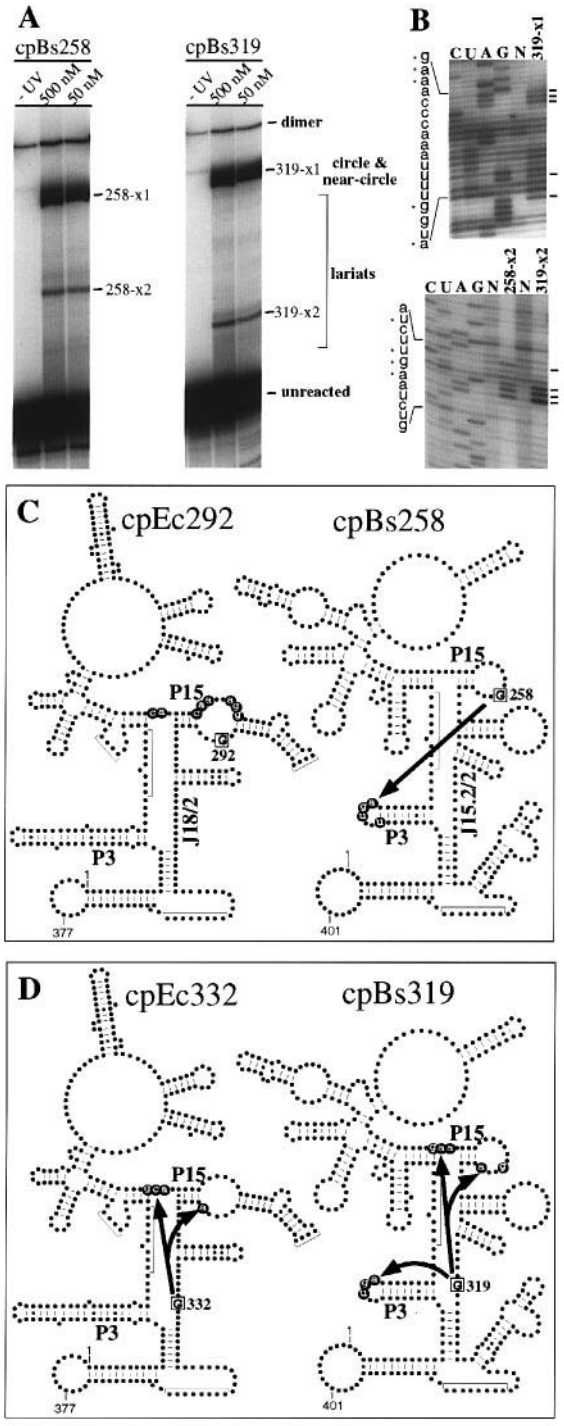

Table I. Kinetic characteristics of circularly permuted RNAs[a]

cpRNA	K_M (nM)	k_{cat} (per min)
Native Bs	25.4 (±10.5)	0.29 (± 0.10)
cpBs34	20.8 (± 0.4)	0.32 (± 0.10)
cpBs45	25.5 (± 9.5)	0.10 (± 0.03)
cpBs72	29.0 (± 7.6)	0.21 (± 0.04)
cpBs123	19.9	0.41
cpBs156	22.0	0.24
cpBs190	26.4 (± 5.6)	0.16 (± 0.07)
cpBs219	93.5 (± 2.3)	0.79 (± 0.23)
cpBs226	24.0 (± 7.6)	0.25 (± 0.09)
cpBs258	27.4 (± 5.3)	0.48 (± 0.23)
cpBs288	13.7	0.26
cpBs319	19.6 (± 3.4)	0.14 (±0.06)
cpBs347	16.1	0.17
cpEc52	19.9 (± 0.9)[b]	0.22 (± 0.06)
cpT18	83.0	0.15

[a]Kinetic assays were done in 3 M NH_4OAc and 75 mM Mg^{2+}.
[b]The K_M for native *E.coli* RNase P RNA is 26.1 (± 6.2) nM and the k_{cat} is 0.2 (± 0.04)/min.

Fig. 2. (**A**) Cross-linking analysis of cpBs258 and cpBs319 RNAs. Uniformly ^{32}P-labeled photoagent-containing cpRNAs with (UV +) or without (UV –) 302 nm UV irradiation were analyzed by gel electrophoresis and autoradiography. The cross-linking reactions contained 50 or 500 nM photoagent-containing RNAs. The dimer RNA band of each lane resulted from formation of UV-independent disulfide bonds between GMPS-initiated RNAs. The circle, lariat and unreacted RNAs of preparative cross-linking reactions were gel purified for primer extension analysis. (**B**) Primer extension mapping of cross-linking sites. The primer extension of cross-linked RNAs, cpBs319-x1, cpBs258-x2 and cpBs319-x2 were carried out with 5′-^{32}P-labeled oligonucleotides BS319R and BS80R. Lanes C, U, A and G correspond to sequencing reactions with non-cross-linked RNA template, and lane N is the control primer extension without dideoxynucleotides. The primer extension reactions using the cross-linked RNA species are indicated above each lane. The termination sites of primer extension, at which reverse transcriptase pauses or terminates because of the cross-linked residue, not structure in the RNA template, are indicated to the right of each gel. The RNA sequences are shown alongside each gel. Asterisks denote the actual cross-link sites which are one nucleotide 5′ to the primer extension stop sites. (**C**) Comparison of cross-linking sites between cpEc292 (Harris *et al.*, 1994) and cpBs258 RNAs. (**D**) Comparison of cross-linking sites between cpEc332 (Harris *et al.*, 1994) and cpBs319 RNAs. The square-boxed Gs indicate the photoagent attachment sites, located at the 5′ end of each cpRNA. The highlighted letters represent their corresponding cross-linking sites. Arrowheads indicate the direction of long-range cross-links.

Table II. Cross-linking analysis of circularly permuted RNAs[a]

RNA	Cross-linked RNA species	Cross-linking efficiency (%)	Cross-linking sites
Native Bs	wt-x1	3	ND
	wt-x2	1	122–124
cpBs34	34-x1	8	ND
	34-x2	1	317–318
	34-x3	1	258–261
	34-x4	1	252–253
cpBs45	45-x1	2	384–385
	45-x2	12	370–374
	45-x3	5	313–315; 317–321
cpBs72	72-x1	13	ND
	72-x2	3	280–281
cpBs123	123-x1	7	ND
	123-x2	2	399–400, 402–405
cpBs156	156-x1	10	ND
cpBs162	162-x1	13	ND
cpBs190	190-x1	7	ND
cpBs205	205-x1	7	ND
cpBs219	219-x1	7	ND
cpBs226	226-x1	8	ND
cpBs258	258-x1	11	ND
	258-x2	2	30; 33–35
cpBs288	288-x1	5	ND
	288-x2	1	30; 33–35
cpBs319	319-x1	12	244–246; 258; 261
	319-x2	2	33–35
cpBs347	347-x1	14	ND
cpEc52	52-x1	9	28–29
	52-x2	1	332
	52-x3	1	292–293; 295
	52-x4	1	254
cpT18	18-x1	1	233–234
	18-x2	7	233–234

[a]The cross-linking reactions were done in 3 M NH_4OAc and 75 mM Mg^{2+}.
ND (not determined): according to the gel mobility of the cross-linked RNAs (Harris *et al.*, 1997), these cross-link sites are near the 3′ end of the cpRNAs, and hence provide no significant constraint on the structure model.

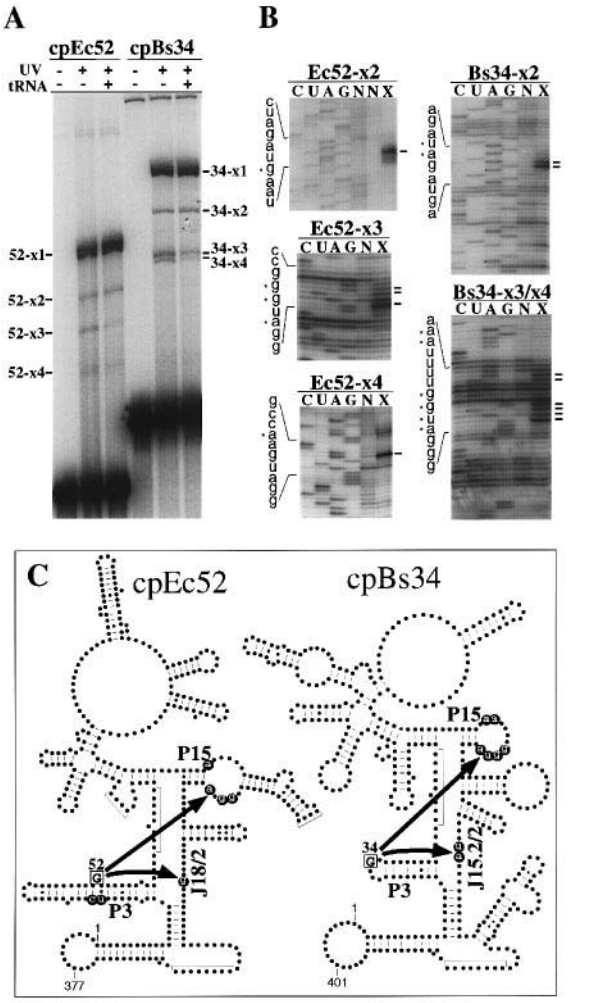

Fig. 3. (**A**) Cross-linking analysis of cpEc52 and cpBs34 RNAs. Photoagent-containing cpRNAs were analyzed as described in Figure 2A. Irradiation with 302 nm UV was carried out for 30 min in either the presence (+) or absence (–) of 500 nM pre-tRNA. (**B**) Primer extension analysis of cross-linked RNAs. CpEc52-x2, cpEc52-x3, cpEc52-x4, cpBs34-x2 and cpBs34-x3/x4 cross-linked RNAs were purified from preparative gels and analyzed using primer extension with appropriate 5′ ^{32}P-labeled primers, BS52R, BS319R, BS350R and BS382R. Lanes C, U, A and G denote sequencing reactions on the non-cross-linked RNAs, and lane N indicates the same primer extension without adding dideoxynucleotides. The primer extension reactions using the cross-linked RNA species are indicated above each lane. The termination sites of primer extension at which reverse transcriptase pauses or terminates because of the cross-linked residue, not structure in the RNA template, are indicated to the right of each gel. The RNA sequences are shown alongside each gel. Asterisks denote the actual cross-link sites which are one nucleotide 5′ to the primer extension stop sites. (**C**) Comparison of cross-link sites between cpEc52 and cpBs34 RNAs. The square-boxed Gs indicate the photoagent attachment sites, located at the 5′ end of each cpRNA. The highlighted letters represent the corresponding cross-linked sites. Arrowheads indicate the direction of long-range cross-links.

Figure 2C, the *B.subtilis* RNA with the photoagent at nucleotide 258 cross-linked to L3 (the loop of the P3 helix), which was not seen with the cross-linking assay of the homologous cpEc292 RNA. Similarly, cpBs319 RNA cross-linked to the L3 region, but comparable cross-linking with P3 was not seen in the analysis of cpEc332 RNA (Figure 2D). To verify the validity of P15–P3 and J18/2–P3 cross-links, cpEc52 and cpBs34 RNAs, with the photoagents placed at homologous positions in the P3/L3 region of *E.coli* and *B.subtilis* RNAs, were made and assayed (Figure 1). Reciprocal cross-links from P3 back to the P15 loop and J18/2 region (*B.subtilis* J15.2/2) are seen in both the *E.coli* and *B.subtilis* cpRNAs (Figure 3). The cross-linking efficiencies for the RNA species Ec52-x3, -x4, Bs34-x3 and -x4 are reduced in the presence of pre-tRNA (Figure 3A), consistent with the fact that pre-tRNA interacts with L15 and J18/2 (Burgin and Pace, 1990; Oh and Pace, 1994). These similar cross-linking results with the cpBs and cpEc RNAs are also consistent with the same tertiary structure.

Long-range interactions involving stem–loop structures that dock elsewhere in an RNA are emerging as important architectural themes in the structures of large RNAs. Such features seem to serve as trusses to bolster the otherwise more tenuous tertiary structure (Brown *et al.*, 1996). Several of these bridging interactions have been identified in RNase P RNA and provide important criteria for the global model. As documented in Figure 4A, reciprocal cross-linking occurs between the ends of P1 and P9, demonstrating that P1 and P9 are close to, and potentially interact with, one another. Additionally, we detected cross-linking between L9 and the artificial loop of P1 in cpBs162, cpBs190 and cpBs219 RNAs using 254 nm UV

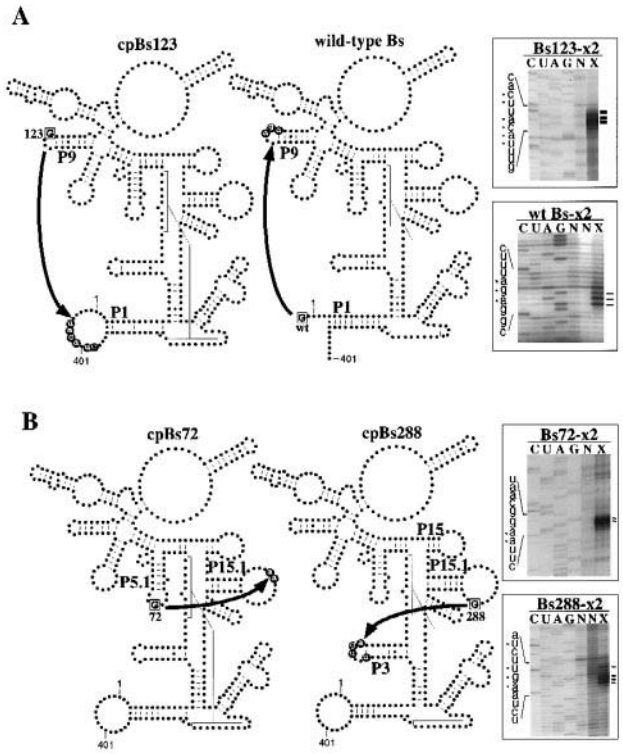

Fig. 4. (**A**) Comparison of cross-linking sites between cpBs123 and wild-type (wt) *B.subtilis* RNase P RNAs. (**B**) Comparison of cross-linking sites between cpBs72 and cpBs288 RNAs. The square-boxed Gs indicate the photoagent attachment sites located at the 5′ end of each cpRNA, and the highlighted letters correspond to cross-linked sites. Arrowheads indicate the direction of long-range cross-links. Primer extension analyses of cross-linked RNAs, cpBs123-x2 and wild-type Bs-x2 were performed using 5′-^{32}P-labeled primers, BS123R and BS156R, respectively. Primer extension analyses of cross-linked RNAs, cpBs72-x2 and cpBs288-x2, were carried out using primers BS45R and BS123R. Lanes C, U, A and G are sequencing reactions on the unreacted RNA, and lane N is the same reaction without adding dideoxynucleotides. The termination sites of primer extension at which reverse transcriptase pauses or terminates because of the cross-linked residue, not structure in the RNA template, are indicated to the right of each gel. The RNA sequences are shown alongside each gel. Asterisks denote the actual cross-link sites which are one nucleotide 5′ to the primer extension stop sites.

irradiation in the absence of photoagent (Chen, 1997). The primer extension analysis indicates that the three uridine bases of L9 may form pyrimidine dimers with the three uridine bases of the artificial loop of P1 (Chen, 1997). Furthermore, a photoagent at the 3′ end, but not the 5′ end, of native *E.coli* RNA also cross-linked to L9 (Harris *et al*., 1997). Finally, a potential tertiary interaction between these two elements has also been suggested by phylogenetic correlation (Massire *et al*., 1997). Taken together, these results strongly indicate that P1 and loop L9 are adjacent and probably interact with one another in both *B.subtilis* and *E.coli* RNAs.

Helices P5.1 and P15.1 are present in the RNase P RNAs of representatives of the low G + C Gram-positive phylogenetic group of bacteria, but do not occur in other types of RNase P RNAs, such as that of *E.coli*. To examine the relative orientation and position of these two phylogenetically variable elements, cpBs72 and cpBs288 RNAs were constructed and assayed. As shown in Figure 4B, the photoagent at G72 in the loop of P5.1 cross-links to A280 and A281 in the loop of P15.1. Additionally, a potential tertiary interaction between nucleotide A71 in the loop of P5.1 and nucleotide U283 in the loop of P15.1 has been identified by phylogenetic sequence comparison (Haas *et al*., 1996a). This evidence points to a tertiary interaction between the loops of P5.1 and P15.1 in the *B.subtilis* RNA. The photoagent on the loop of P15.1 does not reciprocally cross-link back to P5.1, however, but to L3 (Figure 4B). This provides another important constraint, between P3, P15 and P15.1, for molecular modeling.

CpBS156, cpBS162, cpBS190, cpBs205, cpBs219, cpBs226 and cpBs347 displayed only local cross-link patterns (Figure 1 and Table II), resulting from the 5′ photoagent reaction with the 3′ end of the RNA. Local cross-linking provides no useful structural information at the resolution of the cross-linking agent (nominally 9 Å). The 5′ ends of the cpRNAs with only local cross-links are located within P10.1, P12 and P19, and the single-stranded regions, J11/12 and J12/11 (Figure 1). The lack of long-range cross-linking by these cpRNAs may be due to the lack of proximal reactive groups and/or a poor orientation of the photoagent.

Computer modeling of E.coli and B.subtilis RNase P RNAs

In order to survey the global arrangement and orientation of helices for both *E.coli* and *B.subtilis* RNase P RNAs, the available data were utilized as distance constraints for computer modeling. Tertiary structure models of RNase P RNA were constructed by using the modeling program YAMMP-RNA, in which helices or nucleotides are treated as ‘pseudoatoms’ and relative positions are calculated using an energy minimization approach to satisfy the distance constraints optimally (Malhotra *et al.*, 1994). The initial computer modeling used only the intramolecular constraints of secondary structure and intramolecular cross-links to construct an RNase P RNA-only structure without the influence of intermolecular constraints established between tRNA and RNase P RNA.

One hundred initially random models were minimized with the input constraints using YAMMP-RNA, and helix axes represented as cylinders were superimposed (Figure 5A and B) to assess the correspondence of the models. The collections of models of the two types of RNA closely correspond in distribution and orientation of helices. While no single model is specifically accurate, the superimposition of 100 distinct calculated models reveals an internally consistent arrangement of helices, with standard deviations of the positions of helix midpoints in the superimposed YAMMP models (Figure 5) of ~5 Å for most helices in the core structure (Figure 5C). The significant degree of similarity between the *E.coli* and *B.subtilis* structure models agrees with the homology of the RNAs.

Previous YAMMP modeling using available intermolecular constraints, between RNase P RNA and tRNA, did not position the tRNA definitively in the complex. Consequently, in previously published models, the tRNA was positioned manually onto the ribozyme (Harris *et al.*, 1994, 1997). To characterize better the orientation of the tRNA in the complex, additional constraints were required. The 5′ and 3′ ends of mature tRNA have been shown to be in the vicinity of J18/2 (*B.subtilis* J15.2/2) and L15/16 (*B.subtilis* L15), respectively (Burgin and Pace, 1990;

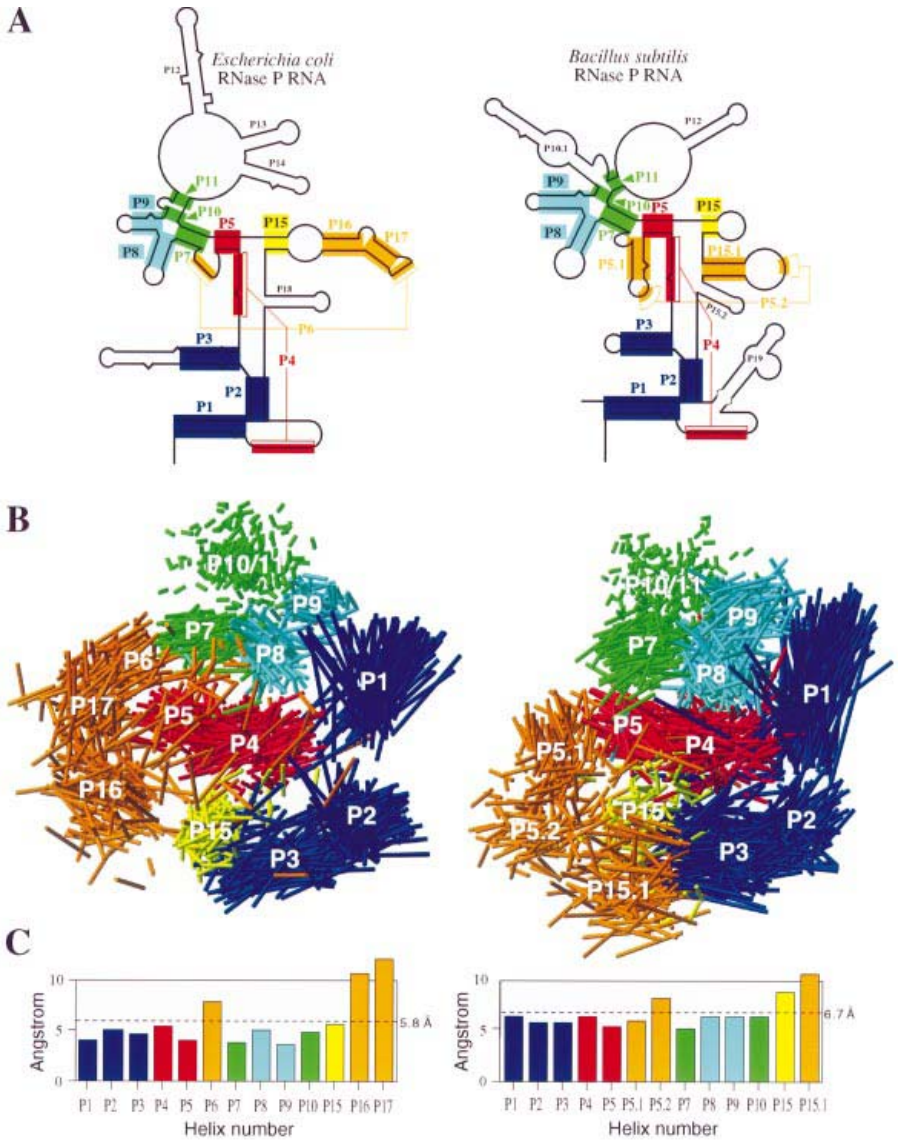

Fig. 5. Superimpositions of 100 YAMMP models of *E.coli* and *B.subtilis* RNAs. (**A**) Secondary structures of *E.coli* and *B.subtilis* RNase P RNAs colored according to the helical domains. Nucleotides 70–72 (L5.1) and 282–284 (L15.1) of *B.subtilis* RNA were modeled as a base-paired region, P5.2. (**B**) Superimpositions of helices of *E.coli* and *B.subtilis* RNase P RNAs. Helices in each model are represented by a thin cylinder that is 5% of the width of a normal A-form RNA helix. Colored elements correspond to the secondary structure shown in (A). Each superimposition consists of 100 distinct models derived by molecular mechanics. (**C**) The standard deviations of each helix in the superimposed model set were calculated from the position of helix midpoints. The standard deviation, in Å, is shown by columns. The dashed line represents the mean of standard deviations of the helices shown, which is 5.8 Å for *E.coli* RNA and 6.7 Å for *B.subtilis* RNA. The collection of YAMMP helical models is internally consistent, evidenced by their low average potential energies, 0.56 ± 0.51 and 0.69 ± 0.52 kcal/mol for the *E.coli* and *B.subtilis* RNAs, respectively.

Kirsebom and Svard, 1994; Oh and Pace, 1994). Intermolecular cross-links from the cpBs258 and cpBs319 RNAs to pre-tRNA (Chen, 1997) also indicate the close proximity between nucleotides 258 and 319 of *B.subtilis* RNA and its substrate pre-tRNA. These constraints position the 5′ and 3′ ends of tRNA at the active site of RNase P RNA, which consists mainly of P4 helix and its flanking regions, J5/15, J18/2 (*B.subtilis* J15.2/2) and J2/4 (*B.subtilis* J19/4). To obtain additional lateral constraints, experiments involving intermolecular cross-linking with a new circularly permuted tRNA, cptRNA18, were carried out. Photoagent attached to the 5′ end of cptRNA18 (at nucleotide 18) cross-links to nucleotides A233 and A234 of *E.coli* RNase P RNA, adjacent to helix P11 (Figure 6). Furthermore, a photoagent attached to the 5′ end of cptRNA53, located in the T stem–loop, cross-links to the bulged A of helix P9 in both *E.coli* and *B.subtilis* RNase P RNAs (Nolan *et al.*, 1993 and data not shown). Nucleotide 62 of the tRNA has been reported to interact with A230 of *B.subtilis* RNase P RNA through ribose 2′-hydroxyl-base hydrogen bonding (Pan *et al.*, 1995). YAMMP modeling that includes all available data generates a satisfactory model with tRNA juxtaposed to the active site of RNase P RNA. The position and orientation of the tRNA within the new YAMMP model agrees with that of the previous model (Harris *et al.*, 1997). Some details of the RNase P–tRNA complexes of *E.coli* and *B.subtilis* are discussed below.

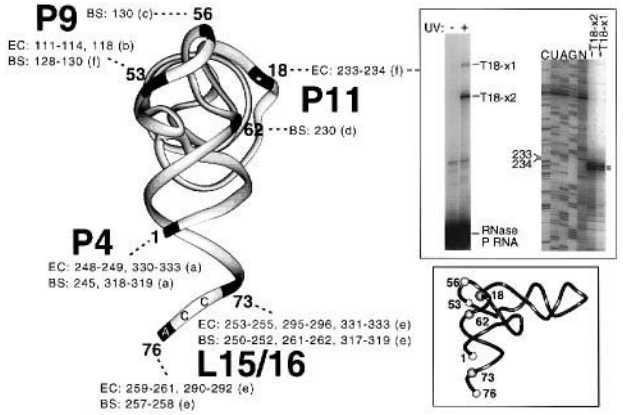

Fig. 6. Intermolecular constraints from tRNA to RNase P RNAs of *E.coli* and *B.subtilis*. The top inset shows cross-linked species that result from intermolecular cross-linking between photoagent-attached cpT18 RNA and uniformly ^{32}P-labeled *E.coli* RNase P RNA upon irradiation with 302 nm UV. The sites of cross-linking in the T18-x1 and -x2 RNAs were mapped by primer extension analysis. The nucleotide numbers of cross-linked sites are indicated alongside the primer extension gel. The bottom inset shows the side view of the tRNA three-dimensional structure, in which the informative nucleotides are shown as small spheres with their nucleotide numbers marked. Structural constraints are shown as dashed lines between the numbered nucleotides of tRNA (1, 18, 53, 56, 62, 73 and 76) and the corresponding nucleotide numbers of *E.coli* or *B.subtilis* RNase P RNAs. The secondary structural elements of RNase P RNA that are proximal to tRNA are also indicated. References for each intermolecular constraint are: (a) Burgin and Pace (1990); (b) Nolan *et al.* (1993); (c) Loria and Pan (1997); (d) Pan *et al.* (1995); (e) Oh and Pace (1994); (f) this study.

Discussion

All structure models of macromolecules are based on best fits to experimental data such as NMR-derived distance constraints, electron densities or, as used in this study, cross-linking and base-pairing constraints. For accuracy, it is important that the collection of constraints is validated properly. The validity of cross-linking constraints in the RNase P data set has been verified repeatedly by: (i) catalytic activity of the cross-linked RNA (Harris *et al.*, 1997); (ii) near-native kinetic parameters of the cpRNAs (Table I and Harris *et al.*, 1994); (iii) reciprocal cross-linking (Figure 4); and (4) homologous cross-links in the *B.subtilis* and *E.coli* RNAs (Figures 2 and 3). These new data, together with the large set of structural constraints obtained in previous analyses of both *E.coli* and *B.subtilis* RNAs, were used to derive internally consistent structure models using a molecular mechanics program. The models of the *B.subtilis* and *E.coli* RNAs exhibit a high degree of similarity and suggest some interesting instances of phylogenetic compensation.

The cross-linking and comparative data provide the orientation of helices in the global structure (Figure 5). Moreover, two-thirds of the sequence lengths of the RNAs are occupied by secondary structure proven by comparative analysis. We expect that the elements of secondary structure adopt A-form atomic parameters. Helical elements are joined by 'single-strand' sequences, usually short and mainly with undetermined structures. Nonetheless, positional constraints of helices also limit the possible structures of the joining sequences. Consequently, we were able to use predicted structural information and manual refinement to formulate global working models of the *E.coli* and *B.subtilis* RNase P RNAs (Figure 7). The current model of the *E.coli* RNA is refined beyond the previously published model (Harris *et al.*, 1997). This refined *E.coli* model exhibits the same core tertiary structure as the newly derived *B.subtilis* model, even though each contains different secondary structural elements not present in the other RNA. We discuss the model in the context of domains which individually possess distinct structure and perhaps function: (i) the P1–P3 domain; (ii) the P4–P5 domain (catalytic domain); (iii) the P7–P11 domain (cruciform domain); (iv) the P15–P18 domain which includes P15, P15.1, P15.2 and P5.1 in *B.subtilis* RNA (CCA-binding domain); and (v) the P12–P14 domain, which includes P10.1 and P12 in the *B.subtilis* RNA. The latter structural component is not included in the current global model because of insufficient constraints (Table II).

The P1–P2–P3 domain occurs in all RNase P RNAs (Chen and Pace, 1997). In both *E.coli* and *B.subtilis* RNAs, the position of the P1–P2–P3 domain is well grounded by inter-domain cross-links, for example P1–P9 (Figure 4A) and P3–P15 (Figure 3C), and comparative data (Massire *et al.*, 1997). The two cross-links were seen with both *B.subtilis* and *E.coli* RNAs. In the formulation of the global model, these two constraints are decisive because they lock down both ends of the P1–P3 domain, to the P7–P9 domain and the P15 domain, respectively.

Helices P4 and P5 together with other 'single-stranded' regions, J3/4, J5/15, J18/2 (or J15.2/2) and J2/4 (or J19/4), are highly conserved and form the catalytic core of RNase P. Many nucleotides in this catalytic domain are universally conserved, not only in bacterial, but also in archaeal and eukaryal RNAs (Chen and Pace, 1997). The functional importance of these conserved regions is supported by chemical modification–interference (Harris and Pace, 1995), and mutagenesis experiments (M.A.T. Rubio and N.R.Pace, unpublished data). The proximity of this functional domain to the cleavage site of the substrate, pre-tRNA, is indicated by high efficiency intermolecular cross-linking from the 5′ end of mature tRNA, the substrate phosphate (Burgin and Pace, 1990). The highly conserved single-stranded sequences probably are packed with helix P4 to form the catalytic site of RNase P. All conserved regions in the models are confined within this catalytic core (Figure 7B).

The P15–P16–P17 domain of the *E.coli* RNA is a series of short helices and internal bulges (J15/16 and J16/17) that docks elsewhere in the RNA to form helix P6. This irregular helical structure has at least two roles: (i) P15 and the J15/16 loop, which together bind the 3′-terminal CCA moiety of the tRNA substrate (Kirsebom and Svard, 1994; LaGrandeur *et al.*, 1994), fit the substrate into the active site; and (ii) the P15–P16–P17–P6 element serves as a structural truss across the RNA. This truss stabilizes the structure of the RNA; removal results in instability of the global structure (Darr *et al.*, 1992). In the *B.subtilis* RNA, these two functions are provided by independent structures. The P15/L15 structure that binds to the substrate remains in the *B.subtilis* RNA, but the structurally stabilizing element P16–P17–P6 in the *E.coli* RNA is replaced in the *B.subtilis* RNA by helices P5.1 and P15.1. Interaction between the loop of P5.1 and P15.1 is strongly supported by both phylogenetic evidence and photocross-linking data

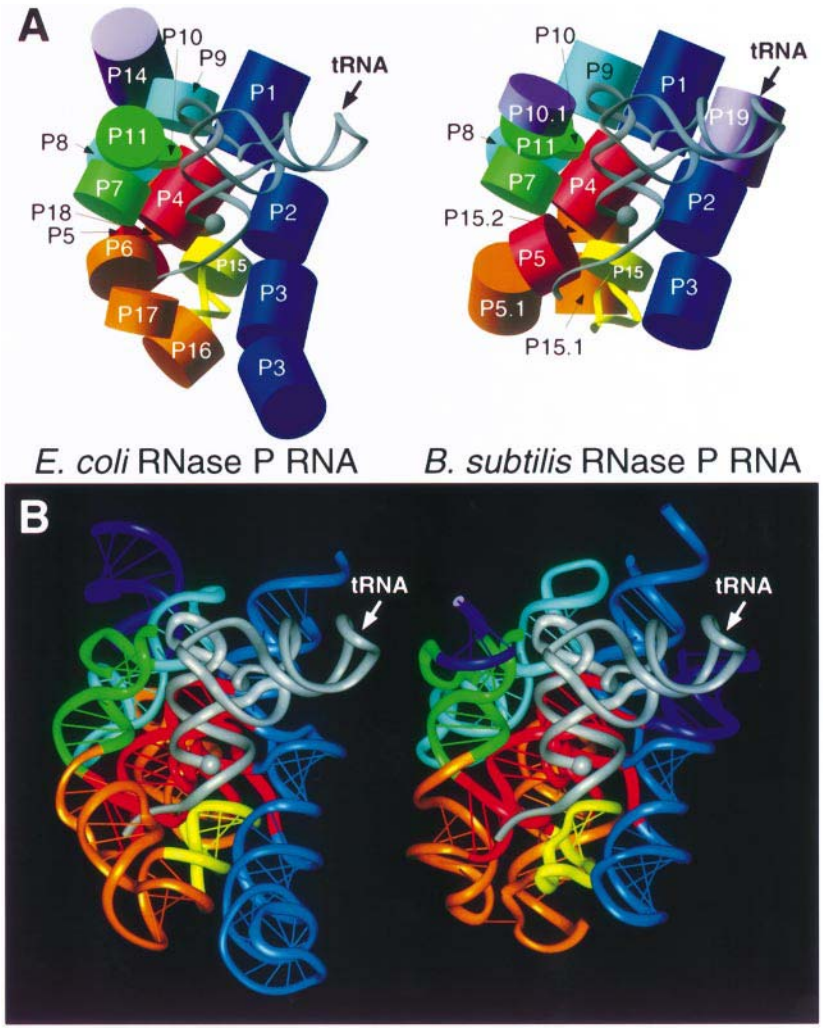

Fig. 7. (**A**) Helix barrel models of *E.coli* and *B.subtilis* RNase P RNA–tRNA complexes. Each cylinder represents the position and orientation of an A-form helix of appropriate length. tRNA is displayed as a ribbon in which the 5′-phosphate (the RNase P substrate site) is shown as a gray sphere. The CCA-binding elements, J15/16 of *E.coli* RNA and L15 of *B.subtilis* RNA, are shown as yellow ribbons. Elements are colored and numbered according to the secondary structure shown in Figure 5A, except helices in purple, which were excluded from YAMMP modeling. (**B**) Ribbon models of *E.coli* and *B.subtilis* RNase P RNA–tRNA complexes. The backbones of RNA are shown as ribbons and colored as in (A). tRNA is shown as a gray ribbon with the 5′-phosphate indicated as a sphere.

(Haas *et al.*, 1996a; Figure 4B and Table II). The structural constraints establish that the P5.1–P15.1 unit in the *B.subtilis* RNA and the P16–P17–P6 unit in the *E.coli* RNA occupy approximately the same relative positions in the superstructures of the two RNAs, but with different configurations (Figures 5 and 7). Based on our model, then, the interaction between P5.1 and P15.1 could be a functional equivalent to helices P16–P17–P6 of the *E.coli* RNA, which stabilizes the global structure of that RNA (Darr *et al.*, 1992). It is clear from studies of a broad diversity of RNase P RNAs that the *E.coli* type of RNA is ancestral (Haas *et al.*, 1996b). Thus, in the course of evolution, the *B.subtilis* (low G + C Gram-positive phylogenetic group) RNA presumably lost the P16–P17–P6 element and gained the P5.1–P15.1 structures in compensation.

The P7–P11 cruciform domain is present in all bacterial RNase P RNAs, with some length variation. Cross-linking, chemical modification and mutagenesis data all implicate this domain in binding the TψC stem–loop of the tRNA substrate (Nolan *et al.*, 1993; Pan *et al.*, 1995; Loria and Pan, 1997). Although the helical elements are well positioned by cross-linking constraints (Figures 5, 6 and 7), the structure of the intersection of the four helices cannot be resolved with these methods. The structure of this four-way junction is likely to be critical to the function of the domain.

Approximately 30% of the molecular length of the *E.coli* (nucleotides 127–234, including P12, P13 and P14) and *B.subtilis* RNase P RNAs (nucleotides 134–230, including P10.1 and P12) are not included in the YAMMP analysis (Figure 5) and the current structure model (Figure 7) because of insufficient constraints. Most photoaffinity agents positioned in this region cross-link only locally, so few long-range positioning constraints are available. In the case of the *E.coli* RNA, helix P14 is indicated by comparative analysis to dock into helix P8 (Brown *et al.*, 1996). The *B.subtilis* RNA lacks P14, but contains a novel element, P10.1, that is proposed to function as a docking site for P12 (Tanner and Cech, 1995). The P10.1–P12 interaction has not been confirmed by cross-linking results, however (Figure 1 and Table II). Further work will clarify

Table III. Oligonucleotides used for constructing cpRNA genes

	Sequence
BS34F	5′-GGGGTACC**TAATACGACTCACTATA**GAATCTGTAGAGGAAAGTCCA-3′
BS34R	5′-CCCGACGCGTCAAGATCTGCAGCGATTACC-3′
BS45F	5′-TTGTCGAGCTC**TAATACGACTCACTATA**GGGGAAAGTCCATGCTCGCAC-3′
BS45R	5′-TCTACAGATTCAAGATCT-3′
BS72F	5′-TGTCGAGCTC**TAATACGACTCACTATA**GATGCCCGTAGTGTTCG-3′
BS72R	5′-TCAGCACCGTGCGAGCA-3′
BS123F	5′-GGGGTACC**TAATACGACTCACTATA**GAGGCTGACGGCAGGAAAA-3′
BS123R	5′-GGGGTACCTAAAGACTGCCCTAGCTTAT-3′
BS156F	5′-GGGGTACC**TAATACGACTCACTATA**GGATATGGCTGAGTATCCTT-3′
BS156R	5′-GGGGTACCGAAGACGTAGGCTTTTTTCC-3′
BS162F	5′-GGGGTACC**TAATACGACTCACTATA**GGCTGAGTATCCTTGAAAG-3′
BS162R	5′-GGGGTACCATATCCGAAGACGTAGGCTT-3′
BS190F	5′-TGTCGAGCTC**TAATACGACTCACTATA**GACGAAGTCTCACTAGAAATG-3′
BS190R	5′-ACTGTGGCACTTTCAAGG-3′
BS205F	5′-GGGGTACC**TAATACGACTCACTATA**GAAATGGTGAGAGTGGAAC-3′
BS205R	5′-GGGGTACCTAGTGAGACTTCGTCACTGT-3′
BS219F	5′-**TAATACGACTCACTATA**GGAACGCGGTAAACCCCTC-3′
BS219R	5′-ACTCTCACCATTTCTAGT-3′
BS226F	5′-TGTCGAGCTC**TAATACGACTCACTATA**GGTAAACCCCTCGAGCG-3′
BS226R	5′-GCGTTCCACTCTCACCA-3′
BS258F	5′-TGTCGAGCTC**TAATACGACTCACTATA**GGTAGGGGAACCTTCTTA-3′
BS258R	5′-AAAATTTGGGTTTCTCGCT-3′
BS288F	5′-GGGGTACC**TAATACGACTCACTATA**GGAGAGAAGGACAGAATGCT-3′
BS288R	5′-GGGGTACCGTTGAATTCCGTTAAGAAGGTTC-3′
BS319F	5′-TGTCGAGCTC**TAATACGACTCACTATA**GATGATTGCCGCCTGAG-3′
BS319R	5′-TATCTACAGAAAGCATTC-3′
BS347F	5′-GGGGTACC**TAATACGACTCACTATA**GAGCCGTTTGCAGTACGA-3′
BS347R	5′-GGGGTACCATCACCTCGTACTCAGGCG-3′
EC52F	5′-GGAATTC**TAATACGACTCACTATA**GGCGGAGGGGAGGAAAGT-3′
EC52R	5′-CCCGACGCGTCCGTCTCCCCCGAAGAGG-3′
T18F	5′-**TAATACGACTCACTATA**GGTTAGAATGCCTGCCTG-3′
T18R	5′-GGGGTACCAACTGAACTACCGGACCATT-3′

The promoter sequence for phage T7 RNA polymerase is shown in bold; restriction sites are underlined.

the structure of the portion of the RNase P RNA omitted from YAMMP analysis. Although universally conserved sequences are located in the region omitted from the model, this region is dispensable for catalysis; synthetic RNase P RNAs that lack this region are catalytically active (Waugh and Pace, 1993; Green *et al.*, 1996).

This comparative analysis of the tertiary structures of *E.coli* and *B.subtilis* RNase P RNAs shows how different peripheral structural elements in the two types of RNA contribute to the global stability of these large RNAs. The occurrence of peripheral, modular, stabilizing structural elements also has been suggested from structural studies with group I introns (Murphy and Cech, 1994; Lehnert *et al.*, 1996; Zarrinkar and Williamson, 1996). The combination of phylogenetic comparative and cross-linking analyses, coupled with molecular mechanics-based modeling, provides increasingly reliable, helix-level models of the *B.subtilis* and *E.coli* RNase P RNAs. The consistency of these two models validates our view of the structure and reveals one way in which evolution maintains tertiary structure required for catalytic function even in the face of substantial architectural reorganization. The experimental methods of comparative and cross-linking studies are applicable to the structures of other large RNAs.

Materials and methods

Oligonucleotides

Oligonucleotides used for constructing cpRNA genes are listed in Table III.

Construction of cpRNA genes

Genes for transcription of cpRNase P RNAs were constructed by PCR with a DNA template that contains tandemly repeated genes as described previously (Nolan *et al.*, 1993; Harris *et al.*, 1994). The p153Bamtan plasmid DNA was used as template for amplification of the gene for cptRNAs (Nolan *et al.*, 1993). CpBs45, 72, 190, 219, 226, 258 and 319 RNA genes were cloned into the vector pSBF as described (Harris *et al.*, 1994). The PCR products of other cpRNA genes with *Kpn*I linkers at both ends were digested with *Kpn*I and ligated into the *Kpn*I-linearized vector, pJLC. In this pUC19-derived vector, a *Fok*I site was positioned strategically downstream from the *Kpn*I site. The *Fok*I site was used to generate a 3′ terminus at the precise 3′ end of cpRNA genes for run-off transcription. All cpRNA genes sequences were verified by dideoxy-terminated sequencing (Sanger *et al.*, 1977).

Preparation of RNAs

RNase P RNAs were prepared by run-off *in vitro* transcription with T7 RNA polymerase (a gift from Dr B.Pace) and *Fok*I-digested plasmid DNA. RNAs containing 5′-thiophosphate were generated by *in vitro* transcription in the presence of GMPS as described (Burgin and Pace, 1990), except that a ratio of GMPS:GTP of 6:1 was used. GMPS was synthesized as described previously (Burgin and Pace, 1990). The substrate, pre-tRNA^{Asp} from *B.subtilis*, was ^{32}P-labeled uniformly by *in vitro* transcription using T7 RNA polymerase and *Bst*NI-linearized plasmid pDW152DNA, as previously described (Reich *et al.*, 1986). Transcripts were resolved by electrophoresis through 4% polyacrylamide–8 M urea gels and visualized by UV shadowing. RNAs were eluted passively from excised gel slices in elution buffer (0.3 M NaOAc, 10 mM Tris–HCl pH 8.0, 1 mM EDTA, 0.1% SDS) at 4°C overnight. For 5′-GMPS-containing RNAs, 10 mM dithiothreitol (DTT) was included in the elution buffer to eliminate the formation of disulfide bonds between RNAs. The eluted RNAs were ethanol precipitated twice and resuspended in water.

Enzymatic activity assays

RNase P assays were carried out in reactions containing 1 nM RNase P RNA, 15–100 nM pre-tRNA^{Asp}, 3 M NH_4OAc, 75 mM $MgCl_2$, 0.05%

NP-40 and 50 mM Tris–HCl pH 8.0. Reactions were incubated at 37°C for 15 min and stopped by adding 3 vols of ethanol. RNAs were precipitated and suspended in 10 µl of loading buffer (8 M urea, 10 mM Tris–HCl pH 8.0, 20 mM EDTA, 0.1% SDS, 0.1% xylene cyanol, 0.1% bromophenol blue). Products and uncleaved substrate were separated on 6% polyacrylamide–8 M urea gels. Gels were dried and products were analyzed using a Molecular Dynamics PhosphorImager. The kinetic parameters k_{cat} and K_M were extracted by plotting velocity versus velocity/[substrate] (Eadie–Hofstee plots).

Photoaffinity cross-linking analysis

The photoaffinity cross-linking agent azidophenacyl bromide (Pierce) was specifically attached to the 5′-thiophosphate of the gel-purified 5′-GMPS-containing RNAs as described (Burgin and Pace, 1990). Photoagent-conjugated RNAs were ethanol precipitated and resuspended in 3 M NH_4OAc, 75 mM $MgCl_2$, 0.05% NP-40 and 50 mM Tris–HCl pH 8.0. After pre-incubation at 50°C for 3 min, 37°C for 5 min and 0°C for 2 min, the cross-linking reaction was carried out by UV irradiation at 302 nm at room temperature for 30 min. The cross-linking reaction products were ethanol precipitated and resolved on 4% polyacrylamide–8 M urea gel at 30 W for 5–6 h. For analytical gels, the cross-linked products were detected by autoradiography and Molecular Dynamics PhosphorImager. For preparative gels, the cross-linked RNAs were excised from ethidium bromide-stained gels, eluted into elution buffer at 4°C overnight and ethanol precipitated.

Primer extension analysis of cross-linked RNAs

The gel-purified, cross-linked RNase P RNAs were analyzed by primer extension with appropriate oligonucleotide primers complementary to the 3′ end sequences of each cross-linked RNA. Primers were 5′-end-labeled with [γ-^{32}P]ATP (Amersham) and T4 polynucleotide kinase (New England Biolabs). End-labeled primers were annealed to cross-linked RNAs in 25 mM KCl, 50 mM Tris–HCl pH 8.3 by heating at 85°C for 2 min and slowly cooling to 45°C. The extension reactions were carried out with avian myeloblastosis virus (AMV) reverse transcriptase (Seikagaku) in 50 mM Tris–HCl pH 8.5, 50 mM KCl, 10 mM DTT, 0.5 mM spermidine, 10 mM $MgCl_2$ and 400 µM of each dNTP at 45°C for 30 min, and terminated by adding an equal volume of 2× denaturing loading buffer (80% formamide, 0.1% xylene cyanol, 0.1% bromophenol blue, 20 mM EDTA). The labeled extension products were resolved by electrophoresis in 6 or 8% polyacrylamide–8 M urea–1× TBE gels and detected by autoradiography (Sambrook *et al.*, 1989). Positions of the primer extension stops (one nucleotide before the cross-linking sites) were determined by direct comparison with dideoxynucleotide chain termination sequencing. The dideoxynucleotide-terminated sequencing reactions were carried out with the same ^{32}P-labeled primer on non-cross-linked RNA in four different termination reactions containing 10 µM each of ddGTP, ddATP, ddTTP and ddCTP.

Computer modeling

The molecular mechanics computer program YAMMP was used to develop RNA tertiary structure models (Malhotra *et al.*, 1990, 1994; Tan and Harvey, 1993; Malhotra and Harvey, 1994). Previous and current cross-linking data were used as distance constraints to calculate 100 RNase P RNA models from random coil structures which were refined further through energy minimization with constraints of secondary structure and cross-linking data. The set of refined models was superimposed to indicate the average position of helix axes in the global structure. The graphics of superimposed models were generated using Ribbons (Carson, 1997). Based on the helix barrel structures of the *E.coli* and *B.subtilis* RNAs, pseudo-atom models were constructed in which nucleotides are represented by pseudo-atoms. Helices and single-stranded regions of RNase P RNA were constructed and minimized individually using Biopolymer and Discover modules of the Insight II program (Biosym/Molecular Simulations). The CCA-binding elements, P15/16 and P15/L15 of *E.coli* and *B.subtilis* RNA, respectively, have been modeled (Easterwood and Harvey, 1997) and were incorporated into our modeling. Individual helices and single-stranded linkers were integrated manually into an RNase P RNA global structure model corresponding to the superimposed helical structures generated by YAMMP. The RNA structure models were then refined further with both steepest descent and conjugate gradient energy minimization using the AMBER forcefield. The crystal structure of tRNA was positioned manually into the refined RNase P RNA model according to intermolecular constraints established in cross-linking experiments.

Acknowledgements

We are grateful to Drs A.Malhotra and S.Harvey for helpful discussions on YAMMP. We thank Dr Dan Frank for programing assistance. We also thank lab members and Dr Brian Thomas for careful reading of the manuscript. This work was supported by NIH grant GM34527 (N.R.P.).

References

Altman,S., Kirsebom,L. and Talbot,S. (1993) Recent studies of ribonuclease P. *FASEB J.*, **7**, 7–14.

Brown,J.W., Nolan,J.M., Haas,E.S., Rubio,M.A.T., Major,F. and Pace,N.R. (1996) Comparative analysis of ribonuclease P RNA using gene sequences from natural microbial populations reveals tertiary structural elements. *Proc. Natl Acad. Sci. USA*, **93**, 3001–3006.

Burgin,A.B. and Pace,N.R. (1990) Mapping the active site of ribonuclease P RNA using a substrate containing a photoaffinity agent. *EMBO J.*, **9**, 4111–4118.

Carson,M. (1997) Ribbons. *Methods Enzymol.*, **277**, 493–505.

Chen,J.-L. (1997) *Structure of Ribonuclease P RNA.* Indiana University, Bloomington.

Chen,J.-L. and Pace,N.R. (1997) Identification of the universally conserved core of ribonuclease P RNA. *RNA*, **3**, 557–560.

Darr,S.C., Zito,K., Smith,D. and Pace,N.R. (1992) Contributions of phylogenetically variable structural elements to the function of the ribozyme ribonuclease P. *Biochemistry*, **31**, 328–333.

Easterwood,T.R. and Harvey,S.C. (1997) Ribonuclease P RNA: models of the 15/16 bulge from *Escherichia coli* and the P15 stem loop of *Bacillus subtilis*. *RNA*, **3**, 577–585.

Gardiner,K.J., Marsh,T.L. and Pace,N.R. (1985) Ion dependence of the *Bacillus subtilis* RNase P reaction. *J. Biol. Chem.*, **260**, 5415–5419.

Gautheret,D., Damberger,S.H. and Gutell,R.R. (1995) Identification of base-triples in RNA using comparative sequence analysis. *J. Mol. Biol.*, **248**, 27–43.

Green,C.J., Rivera-Leon,R. and Vold,B.S. (1996) The catalytic core of RNase P. *Nucleic Acids Res.*, **24**, 1497–1503.

Guerrier-Takada,C., Gardiner,K., Marsh,T., Pace,N. and Altman,S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, **35**, 849–857.

Haas,E.S., Brown,J.W., Pitulle,C. and Pace,N.R. (1994) Further perspective on the catalytic core and secondary structure of ribonuclease P RNA. *Proc. Natl Acad. Sci. USA*, **91**, 2527–2531.

Haas,E.S., Armbruster,D.W., Vucson,B.M., Daniels,C.J. and Brown,J.W. (1996a) Comparative analysis of ribonuclease P RNA structure in Archaea. *Nucleic Acids Res.*, **24**, 1252–1259.

Haas,E.S., Banta,A.B., Harris,J.K., Pace,N.R. and Brown,J.W. (1996b) Structure and evolution of ribonuclease P RNA in Gram-positive bacteria. *Nucleic Acids Res.*, **24**, 4775–4782.

Harris,M.E. and Pace,N.R. (1995) Identification of phosphates involved in catalysis by the ribozyme RNase P RNA. *RNA*, **1**, 210–218.

Harris,M.E., Nolan,J.M., Malhotra,A., Brown,J.W., Harvey,S.C. and Pace,N.R. (1994) Use of photoaffinity crosslinking and molecular modeling to analyze the global architecture of ribonuclease P RNA. *EMBO J.*, **13**, 3953–3963.

Harris,M.E., Kazantsev,A.V., Chen,J.-L. and Pace,N.R. (1997) Analysis of the tertiary structure of the ribonuclease P ribozyme–substrate complex by site-specific photoaffinity crosslinking. *RNA*, **3**, 561–576.

Kirsebom,L.A. and Svard,S.G. (1994) Base pairing between *Escherichia coli* RNase P RNA and its substrate. *EMBO J.*, **13**, 4870–4876.

LaGrandeur,T.E., Hüttenhofer,A., Noller,H.F. and Pace,N.R. (1994) Phylogenetic comparative chemical footprint analysis of the interaction between ribonuclease P RNA and tRNA. *EMBO J.*, **13**, 3945–3952.

Lehnert,V., Jaeger,L., Michel,F. and Westhof,E. (1996) New loop–loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the *Tetrahymena thermophila* ribozyme. *Chem Biol*, **3**, 993–1009.

Loria,A. and Pan,T. (1997) Recognition of the T stem–loop of a pre-tRNA substrate by the ribozyme from *Bacillus subtilis* ribonuclease P. *Biochemistry*, **36**, 6317–6325.

Malhotra,A. and Harvey,S.C. (1994) A quantitative model of the *Escherichia coli* 16S RNA in the 30S ribosomal subunit. *J. Mol. Biol.*, **240**, 308–340.

Malhotra,A., Tan,R.K. and Harvey,S.C. (1990) Prediction of the three-dimensional structure of *Escherichia coli* 30S ribosomal subunit: a molecular mechanics approach. *Proc. Natl Acad. Sci. USA*, **87**, 1950–1954.

Malhotra,A., Tan,R. and Harvey,S.C. (1994) Modeling large RNAs and ribonucleoprotein particles using molecular mechanics techniques. *Biophys. J.*, **66**, 1777–1795.

Massire,C., Jaeger,L. and Westhof,E. (1997) Phylogenetic evidence for a new tertiary interaction in bacterial RNase P RNAs. *RNA*, **3**, 553–556.

Milligan,J.F. and Uhlenbeck,O.C. (1989) Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.*, **164**, 51–62.

Murphy,F.L. and Cech,T.R. (1994) GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. *J. Mol. Biol.*, **236**, 49–63.

Nolan,J.M., Burke,D.H. and Pace,N.R. (1993) Circularly permuted tRNAs as specific photoaffinity probes of ribonuclease P RNA structure. *Science*, **261**, 762–765.

Oh,B.-K. and Pace,N.R. (1994) Interaction of the 3′-end of tRNA with ribonuclease P RNA. *Nucleic Acids Res.*, **22**, 4087–4094.

Pace,N.R. and Brown,J.W. (1995) Evolutionary perspective on the structure and function of ribonuclease P, a ribozyme. *J. Bacteriol.*, **177**, 1919–1928.

Pan,T. and Zhong,K. (1994) Selection of circularly permuted ribozymes from *Bacillus subtilis* RNase P by substrate binding. *Biochemistry*, **33**, 14207–14212.

Pan,T., Loria,A. and Zhong,K. (1995) Probing of tertiary interactions in RNA: 2′-hydroxyl-base contacts between the RNase P RNA and pre-tRNA. *Proc. Natl Acad. Sci. USA*, **92**, 12510–12514.

Reich,C., Gardiner,K.J., Olsen,G.J., Pace,B., Marsh,T.L. and Pace,N.R. (1986) The RNA component of the *Bacillus subtilis* RNase P: sequence, activity, and partial secondary structure. *J. Biol. Chem.*, **261**, 7888–7893.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual.* 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.

Siegel,R.W., Banta,A.B., Haas,E.S., Brown,J.W. and Pace,N.R. (1996) *Mycoplasma fermentans* simplifies our view of the catalytic core of ribonuclease P RNA. *RNA*, **2**, 452–462.

Tan,R.K. and Harvey,S.C. (1993) YAMMP: development of a molecular mechanics program using the modular programming method. *J. Comput. Chem.*, **14**, 455–470.

Tanner,M.A. and Cech,T.R. (1995) An important RNA tertiary interaction of group I and group II introns is implicated in Gram-positive RNase P RNAs. *RNA*, **1**, 349–350.

Waugh,D.S. and Pace,N.R. (1993) Gap-scan deletion analysis of *Bacillus subtilis* RNase P RNA. *FASEB J.*, **7**, 188–195.

Zarrinkar,P.P. and Williamson,J.R. (1996) The P9.1–P9.2 peripheral extension helps guide folding of the *Tetrahymena* ribozyme. *Nucleic Acids Res.*, **24**, 854–858.

Received October 29, 1997; revised and accepted 7 January, 1998