Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 15212–15217, December 1998
Biochemistry

Protein component of the ribozyme ribonuclease P alters substrate recognition by directly contacting precursor tRNA

(catalysis/crosslinking/binding/cleavage)

S. NIRANJANAKUMARI*, TRAVIS STAMS†, SHARON M. CRARY*, DAVID W. CHRISTIANSON†, AND CAROL A. FIERKE*‡

*Department of Biochemistry, Box 3711, Duke University Medical Center, Durham, NC 27710; and †Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104-6323

Communicated by Gordon G. Hammes, Duke University Medical Center, Durham, NC, October 29, 1998 (received for review September 22, 1998)

ABSTRACT **The protein component of ribonuclease P (RNase P) binds to the RNA subunit, forming a functional ribonucleoprotein complex *in vivo* and enhancing the affinity of the precursor tRNA (pre-tRNA) substrate. Photocrosslinking experiments with pre-tRNA bound to RNase P reconstituted with the protein component of *Bacillus subtilis* ribonuclease P (P protein) site specifically modified with a crosslinking reagent indicate that: (*i*) the central cleft of P protein directly interacts with the single-stranded 5′ leader sequence of pre-tRNA, and (*ii*) the orientation and register of the pre-tRNA leader sequence in the central cleft places the protein component in close proximity to the active site. This unique mode of interaction suggests that the catalytic active site in RNase P occurs near the interface of RNA and protein. In contrast to other ribonucleoprotein complexes where the protein mainly stabilizes the active tertiary fold of the RNA, a critical function of the protein component of RNase P is to alter substrate specificity and enhance catalytic efficiency.**

RNA molecules play a variety of indispensable structural, catalytic, and regulatory roles within the cell, often in association with specific protein cofactors (1). The ability of RNAs to fold into complex three-dimensional structures provides a diverse array of shapes and spatial arrangements of functional groups for recognition by proteins (2). Because of this complexity, the rules governing RNA–protein recognition are not well understood. Furthermore, the catalytic function of the RNA and protein components in ribonucleoprotein (RNP) complexes has not been clearly identified in most cases. This task is daunting for large RNP complexes because 30–100 proteins and several RNA species are involved in spliceosome assembly and function (3), while more than 50 proteins and several RNAs are present in the ribosome (4). Nonetheless, steady progress is being made in the functional dissection of these large RNP complexes. Investigation of simpler systems in which a single RNA subunit and a single protein subunit associate to form a catalytic heterodimeric RNP enzyme provides fundamental insights. One such system is ribonuclease P (RNase P), which is composed of a large RNA component of ≈400 nt and a small basic protein of ≈120 aa and catalyzes the 5′ maturation of precursor tRNA (pre-tRNA) (5, 6). Eubacterial RNase P RNA (the RNA component of *Bacillus subtilis* RNase P) alone is catalytically active *in vitro* in the presence of high concentrations of monovalent and divalent cations (7, 8); however, the high catalytic efficiency under physiological conditions is accomplished by specific interaction with its protein cofactor.

Although catalytic RNA molecules function in the absence of protein factors *in vitro* (7–10), in many cases there is compelling genetic and biochemical evidence for the involvement of protein factors for catalytic function *in vivo* (11–14). Protein components in many RNA-catalyzed reactions have been implicated in either stabilizing the active tertiary structure of the RNA component (15, 16) or functioning as RNA chaperones (17). In RNase P, the protein component recognizes the folded structure of the catalytic RNA (18–20) but is not essential for forming the tertiary fold (21). Moreover, the protein component alters the substrate specificity of RNase P (22, 23) and enhances the affinity of RNase P for pre-tRNA by interactions with nucleotides at the −2 and −5 positions in the 5′ leader sequence (24, 25). This recognition of the leader sequence could be the result of direct contacts between the substrate and the protein and/or a protein-induced conformational change in the RNA. To address the proximity of the protein component to the substrate and active site of RNase P, we have used site-specific photocrosslinking and primer extension analyses. These crosslinking experiments, interpreted in light of the recent crystal structure of the protein component (26), indicate that a central cleft in the RNase P protein interacts directly with nucleotides in the 5′ leader of pre-tRNA. Additionally these data position the protein component within a few nucleotides of the active site. Therefore, both the RNA and protein components of RNase P play fundamental roles in substrate recognition and catalysis.

MATERIALS AND METHODS

Mutagenesis. Variants of P protein (protein component of *B. subtilis* RNase P) with cysteine substituted at positions 7, 16, 20, 26, 27, 32, 34, 46, 49, 65, 70, 86, 92, and 102 were generated by megaprimer PCR amplification (27) using the plasmid encoding wild-type P protein, pPWT1, as template (28). In the first round of PCR, a DNA fragment was amplified by using a mutagenic primer and a T7 primer (5′-TAATACGACTCACTATAGG-3′), and in the second round the full-length gene was generated by using the amplified fragment from the first round and a 3′ primer (5′-CATGGGATCCTTATTTGGAGGAAGATTTCTTATATAATGAAG-3′, *Bam*HI site underlined). The amplified fragment was digested with *Nco*I and *Bam*HI and cloned into pET 28b (Novagen). P protein variants were overexpressed in BL21(DE3)pLysS cells and purified to homogeneity as described (28).

Derivatization of P Protein Cysteine Variants. The unique cysteine in a P protein variant (100 μM) was derivatized by incubation with azidophenacyl (AzP) bromide (700 μM), a sulfhydryl-specific photoaffinity reagent, at room temperature for 3 h in the dark in 50 mM Tris·HCl, pH 7.5/100 mM NaCl/5% glycerol/1 mM tris(2-carboxyethyl)phosphine. The

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: RNase P, ribonuclease P; RNP, ribonucleoprotein; P RNA, RNA component of *B. subtilis* RNase P; P protein, protein component of *B. subtilis* RNase P; AzP, azidophenacyl; AzP S49C P protein, P protein modified with an AzP moiety on a unique cysteine engineered at position 49; pre-tRNA, precursor tRNA.

‡To whom reprint requests should be addressed.

derivatized protein was separated from the excess reagent by passing through a PD-10 gel filtration column (Pharmacia). The extent of modification was ≥ 95%, as determined by reaction of the residual free thiol group with 5,5′-dithio-bis(2-nitrobenzoic acid) (29).

Crosslinking Assays. RNase P holoenzyme was reconstituted by incubating equimolar RNase P RNA and derivatized P protein (0.4 μM each) in 50 mM Mes/50 mM Tris, pH 6.0/5 mM $CaCl_2$/100 mM NH_4Cl (28). Pre-tRNA^{Asp} substrates, varying in 5′ leader length from 2 to 14 nt, were generated as described (25) and designated as pTR2, pTR3, pTR4, pTR5, pTR6, pTR7, pTR10, and pTR14, where the number indicates the 5′ leader length. Pre-tRNA, tRNA, and 5′ leader RNA were radiolabeled at the 5′ end by incubation with γ-^{32}P-ATP and T4 polynucleotide kinase and purified by urea-PAGE (30). Derivatized RNase P and substrates were preincubated separately at 37°C for 20 min, mixed, and then incubated at 37°C for an additional 20 min. Samples were irradiated for 1 min on ice by using a 312-nm hand torch at ≈5-cm distance with a polystyrene filter. Samples were denatured in SDS-sample buffer (50 mM Tris·Cl, pH 6.8/0.25% SDS/0.1 M DTT/10% glycerol) by heating at 85°C for 5 min, and then the crosslinked pre-tRNA was separated from pre-tRNA on a 15% SDS/PAGE gel (31). The self-cleavage activity of the crosslinked pre-tRNA/protein/P RNA complex was measured by the addition of 10 μM unlabeled pTR14 (to inhibit the activity of the uncrosslinked holoenzyme) followed by the addition of $MgCl_2$ to a final concentration of 10 mM and incubation at 37°C. At a defined time, 5 μl of the sample was quenched with SDS sample buffer and analyzed by 15% SDS/PAGE.

Competition Experiments. To establish the specificity of crosslinking, the derivatized holoenzyme (0.4 μM) was incubated with radiolabeled pTR14 substrate in the presence of excess cold competitors such as mature tRNA, 5′ leader and pTR14 at 5 μM, and poly G and poly C RNA (Sigma) at 250 μg/ml. UV irradiation and processing of the sample were done as described above.

Primer Extension Analyses. To map the position of the crosslink in pre-tRNA, the crosslinked complex was eluted from the gel by passive diffusion and digested with proteinase K (Boehringer Mannheim) for 30 min at 37°C, followed by phenol/chloroform extraction and ethanol precipitation. In a typical primer extension experiment, 1 ng of 5′ end-labeled primer complementary to the 3′ end of pTR14 (5′-TGGCGGTCCGGACGGGAC, TR-Rev) was hybridized to the pre-tRNA/P protein conjugate by incubating at 90°C for 3 min in 50 mM Tris·Cl, pH 8.3/60 mM NaCl/10 mM DTT and then slowly cooling to room temperature. The oligonucleotide was extended with avian myeloblastosis virus reverse transcriptase (Amersham) in 50 mM Tris·Cl, pH 8.3/60 mM NaCl/10 mM DTT/6 mM $MgCl_2$/400 μM each dNTP at 60°C for 15 min (32). Dideoxynucleotide-terminated sequencing reactions using unmodified pre-tRNA as template (see Fig. 4, lanes G, A, T, and C) were prepared as controls and analyzed in parallel with the primer extension products on an 8% polyacrylamide-urea gel, followed by quantitation of the products by using a PhosphorImager (Molecular Dynamics).

RESULTS

Recent kinetic studies indicate that the protein component of RNase P specifically enhances the affinity of pre-tRNA^{Asp} (24, 25). To ascertain whether this enhancement is the result of direct interactions between P protein and pre-tRNA, we carried out photocrosslinking experiments where the crosslinking reagent was covalently attached to the P protein at specific sites. The P protein of *B. subtilis* RNase P is well suited for this study because of the absence of cysteine residues; hence, a unique cysteine could be engineered at a desired position for site-specific modification with the photoagent. Site-specific crosslinking, a powerful method to investigate protein/RNA contacts, is used here to position the protein component in the holoenzyme·pre-tRNA complex and to probe for direct protein/substrate contacts. This approach has been used previously to understand the structural landscape in RNA–RNA (33, 34), RNA–protein (35), and DNA–protein complexes (36). Phenylazide-mediated photocrosslinking occurs through photogeneration of a phenyl nitrene, followed by rearrangement of this nitrene to an electrophilic species, which then reacts with a nearby nucleophile (36). Crosslink formation requires close contact between the photogenerated reactive species and its target; however, the crosslinked region can be as far away as 11 Å because of the length of the phenyl nitrene.

In our first round of experiments, we introduced a unique cysteine into the P protein at position 7, 46, 49, 65, 70, or 92 by using site-directed mutagenesis and then derivatized the single sulfhydryl group with the photoagent AzP bromide (37). Then the holoenzyme complex was reconstituted by incubating equimolar concentrations of *B. subtilis* RNase P RNA and P protein variant derivatized with AzP bromide. In each case, holoenzyme was successfully prepared as indicated by a ≈100-fold enhancement of the turnover rate for pre-tRNA^{Asp} hydrolysis compared with P RNA alone in low salt (10 mM Tris·Cl, pH 7.5/5 mM $MgCl_2$) (data not shown; ref. 28). To stabilize the enzyme·pre-tRNA^{Asp} complex, the concentration of holoenzyme was significantly higher than the substrate K_D for wild-type RNase P, and $MgCl_2$ was replaced with $CaCl_2$, which reduces the hydrolytic rate constant by 10^4-fold but still allows P RNA to fold and bind substrate (24, 38). The derivatized holoenzyme·pre-tRNA complex was irradiated with UV light (312 nm); crosslinking was analyzed by SDS/PAGE and quantitated by using a PhosphorImager. Irradiation of the RNase P·substrate complex reconstituted with AzP S49C P protein (a P protein modified with an AzP moiety on a unique cysteine engineered at position 49) generated a pre-tRNA species that migrated more slowly in the gel than pre-tRNA, consistent with the formation of a protein–RNA crosslink with ≈20% efficiency (Fig. 1*A*, lane 6). However, no P protein/pre-tRNA crosslinks were observed when using labeled R7C, V46C, R65C, A70C, or S92C P protein variants (data not shown). Both AzP-labeled protein and UV irradiation are required for the formation of the putative crosslinked RNA-protein complex (Fig. 1*A*, compare lanes 4–6). Additionally, no crosslinking is observed after irradiation of holoenzyme reconstituted with either wild-type (lane 3) or unmodified S49C (lane 4) P protein. Furthermore, proteinase K digestion of the irradiated complex results in a RNA species with mobility similar to pre-tRNA (lane 7) confirming that the protein component is essential for the altered mobility of the crosslinked complex. Because the slowly migrating pre-tRNA species is unaffected by denaturant (SDS) or high temperature (85°C), we conclude that a covalent bond is formed between pre-tRNA and the derivatized P protein variant when exposed to UV light.

Crosslinking Efficiency Depends on the Length of the 5′ Leader Sequence. Previously we have demonstrated that the protein component of RNase P enhances interactions with the −2 and −5 nucleotides of the substrate leader sequence (25). To determine the length of the leader sequence required to observe crosslinking, RNase P holoenzyme prepared with AzP-S49C P protein was incubated with radiolabeled pre-tRNA substrates of different leader lengths (pTR0, pTR2, pTR3, pTR4, pTR5, pTR6, pTR7, pTR10, and pTR14; the number indicates the length of the 5′ leader sequence) and then irradiated with UV light. Specific P protein/pre-tRNA crosslinks were observed with substrates containing a leader sequence of 4 nt and longer (Fig. 1*B*, lanes 4–9) but not with mature tRNA (lane 1) or substrates with a 2- or 3-nt leader (lanes 2 and 3). These results indicate that the S49C residue in P protein is located near (within 11 Å) the leader sequence of

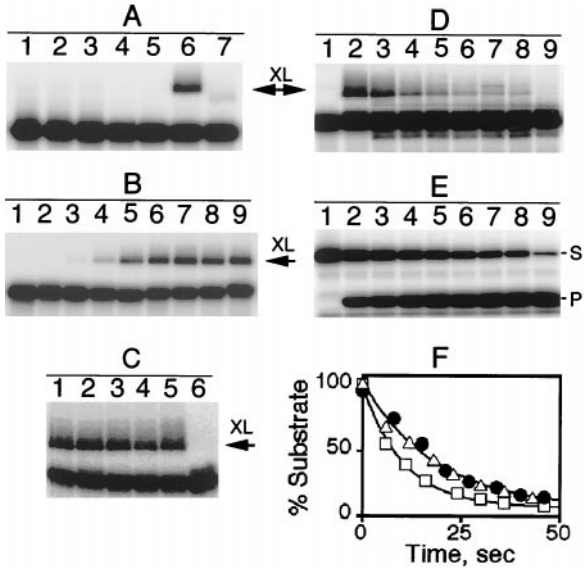

FIG. 1. Crosslinking of pre-tRNA and RNase P holoenzyme reconstituted with AzP-labeled S49C P protein. The RNase P·pre-tRNA complex was formed by incubating ^{32}P-labeled pre-tRNA (5–10 nM pTR14 RNA) and 0.4 μM holoenzyme at 37°C in 50 mM Mes-Tris buffer, pH 6.0, 5 mM $CaCl_2$. Samples were exposed to UV light (312 nm) for 1 min. The crosslinked (XL) products were separated on 15% SDS/PAGE and visualized by using a PhosphorImager. (*A*) Lane 1, pre-tRNA alone; lane 2, pre-tRNA and P RNA; lane 3, pre-tRNA and holoenzyme; lane 4, pre-tRNA and holoenzyme reconstituted with S49C P protein; lanes 5–7, pre-tRNA and holoenzyme reconstituted with AzP-labeled S49C P protein, no irradiation (lane 5), 1-min irradiation (lane 6), and irradiation followed by proteinase K digestion (five units at 37°C for 20 min) (lane 7). (*B*) Crosslinking depends on the length of the leader. Lane 1, mature tRNA; lanes 2–9, pre-tRNA substrates with a leader sequence of 2, 3, 4, 5, 6, 7, 10, and 14 nt, respectively. (*C*) RNA specificity of crosslinks. pTR14 and holoenzyme reconstituted with AzP-S49C P protein was irradiated in the absence of competitor (lane 1), or in the presence of 5 μM mature tRNA (lane 2), 5 μM 5′ leader RNA (lane 3), 250 μg/ml of poly G RNA (lane 4), 250 μg/ml of poly C RNA (lane 5), and 2 μM pTR14 RNA (lane 6) as competitors. Only the substrate effectively inhibits formation of the crosslink. (*D*) Cleavage of the crosslinked pTR14/holoenzyme complex. The cleavage of pTR14 substrate crosslinked to AzP-S49C RNase P was initiated by the addition of $MgCl_2$ (10 mM final concentration) and unlabeled pTR14 substrate (10 μM final concentration) to inhibit the activity of the uncrosslinked RNase P. The reaction was incubated at 37°C, and a 5-μl aliquot was quenched at defined times by the addition of SDS sample buffer. The crosslinked pre-tRNA (XL) and free pre-tRNA were separated on a 15% SDS/PAGE gel and quantitated by using a PhosphorImager. Lanes 1 and 2, ^{32}P-labeled pTR14 and holoenzyme reconstituted with AzP-S49C P protein before and after UV irradiation; lanes 3–9, 8, 15, 21, 27, 34, 40, and 300 sec, respectively, after the addition of $MgCl_2$. (*E*) Single turnover cleavage of pTR14 catalyzed by RNase P reconstituted with unmodified S49C P protein. ^{32}P-labeled pre-tRNA (10 nM pTR14) was incubated with a saturating concentration (0.4 μM) of RNase P at 37°C in 50 mM Mes, 50 mM Tris, pH 6.0, 5 mM $CaCl_2$, and the cleavage reaction was initiated by the addition of $MgCl_2$ (10 mM final concentration). Pre-tRNA (S) and the 5′ leader sequence (P) were separated on a 15% SDS/PAGE gel and quantitated by using a PhosphorImager. Lane 1, ^{32}P-labeled pTR14 and holoenzyme before addition of $MgCl_2$; lanes 2–9, 6, 12, 18, 23, 30, 36, 43, and 314 sec, respectively, after the addition of $MgCl_2$. (*F*) The percent substrate as a function of time is plotted for the single turnover cleavage of pTR14 catalyzed by wild-type RNase P (□), RNase P reconstituted with S49C P protein (△), and pTR14 crosslinked to RNase P reconstituted with AzP-S49C P protein (●). The data are fit to a single exponential decay by using the Kaleidagraph (Synergy software) curve-fitting program; the fitted rate constant and endpoint are 0.09 ± 0.01 s^{-1} and 6 ± 2% for wild-type RNase P and 0.05 ± 0.01 s^{-1} and 5 ± 2% for either RNase P reconstituted with S49C P protein or pTR14 crosslinked to RNase P reconstituted with AzP-S49C P protein.

the bound pre-tRNA and suggest that the protein may directly contact the substrate.

The specificity of the P protein/pre-tRNA crosslink was established by competition experiments. As shown in Fig. 1*C*, the formation of a crosslink between holoenzyme reconstituted with AzP-S49C P protein and pTR14 substrate is not inhibited by the addition of either unlabeled mature tRNA (lane 2), 5′ leader (lane 3), poly G (lane 4), or poly C (lane 5). However, addition of unlabeled pTR14 (lane 6) did inhibit the formation of this crosslink. Therefore, the formation of a specific RNA–protein complex between pre-tRNA and P protein is necessary for photocrosslinking. The lack of inhibition by mature tRNA, despite a concentration that is several-fold higher than the dissociation constant (25), suggests that the crosslinked pre-tRNA species is kinetically trapped in an enzyme substrate complex during irradiation.

Pre-tRNA Crosslinked to RNase P Is Rapidly Cleaved. To demonstrate that the crosslinked species reflects a catalytically relevant binding mode, $MgCl_2$ was added to the crosslinked enzyme-substrate complex and cleavage of the substrate was monitored (Fig. 1*D*) in the presence of excess unlabeled pTR14 to inhibit the cleavage activity of the uncrosslinked holoenzyme. The crosslinked complex is cleaved more than 90% within 1 min under these conditions, whereas the concentration of free pre-tRNA decreases $< 10\%$ (Fig. 1*D*, lanes 2–8). The disappearance of the crosslinked band is well described by a single exponential decay (Fig. 1*F*) with a rate constant of 0.05 ± 0.01 s^{-1}. This value is slightly slower than the single turnover cleavage rate constant for wild-type RNase P holoenzyme under the same conditions (0.09 ± 0.01 s^{-1}) but identical to the cleavage rate constant measured for holoenzyme reconstituted with either unlabeled or AzP S49C P protein (Fig. 1 *E* and *F*). The wild-type cleavage rate constant is decreased ≈3.5-fold compared with the previous measurements (25), likely the result of the presence of $CaCl_2$ in the assay. This experiment indicates that the majority of the crosslinked pre-tRNA is positioned such that specific cleavage can be efficiently catalyzed.

Central Cleft Residues Crosslink with Pre-tRNA. The x-ray crystal structure of P protein reveals a central cleft formed by helix A and the face of the four-stranded β-sheet (Fig. 2) (26). The bottom of the cleft is formed by three aromatic residues, Phe-16 and Phe-20 on helix A and Tyr-34 on β-strand 2, spaced ≈4.5 Å apart. Additional residues lining the cleft are Ser-49 on β-strand 3, Ile-86 on β-strand 4, and Val-32 on β-strand 2. The observation of a crosslink between AzP-derivatized S49C P protein and pre-tRNA suggests that the central cleft binds the leader sequence of pre-tRNA. To test this hypothesis, a second round of experiments was conducted in which a unique cysteine was engineered into the cleft by a single-point mutation at residue 16, 20, 32, 34, or 86. These variants were derivatized with AzP bromide and irradiated in the presence of radiolabeled pre-tRNA substrates. As shown in Fig. 3, all of these AzP-derivatized P protein variants efficiently form a crosslink to pre-tRNA with a leader length of 4 nt or more, but not to the mature tRNA or the substrates with shorter leader sequences (compare *A-E*, lanes 1–9). These results suggest that all of the central cleft residues are in close proximity to the precursor region of pre-tRNA. Additionally, for each variant the formation of the crosslink is inhibited only by cold pre-tRNA, and the crosslinked substrate can be efficiently cleaved upon addition of $MgCl_2$ (data not shown). Furthermore, a comparison of the intensities of the crosslinks as a function of the length of the leader with respect to the position of the side chain in the cleft suggests that the leader passes from left to right of the central cleft as oriented in Fig. 2 (3′ → 5′).

To confirm this orientation of the leader of pre-tRNA bound to the central cleft of P protein, a third round of crosslinking experiments was conducted by using variants with mutations at three additional positions (V26C, A27C, and

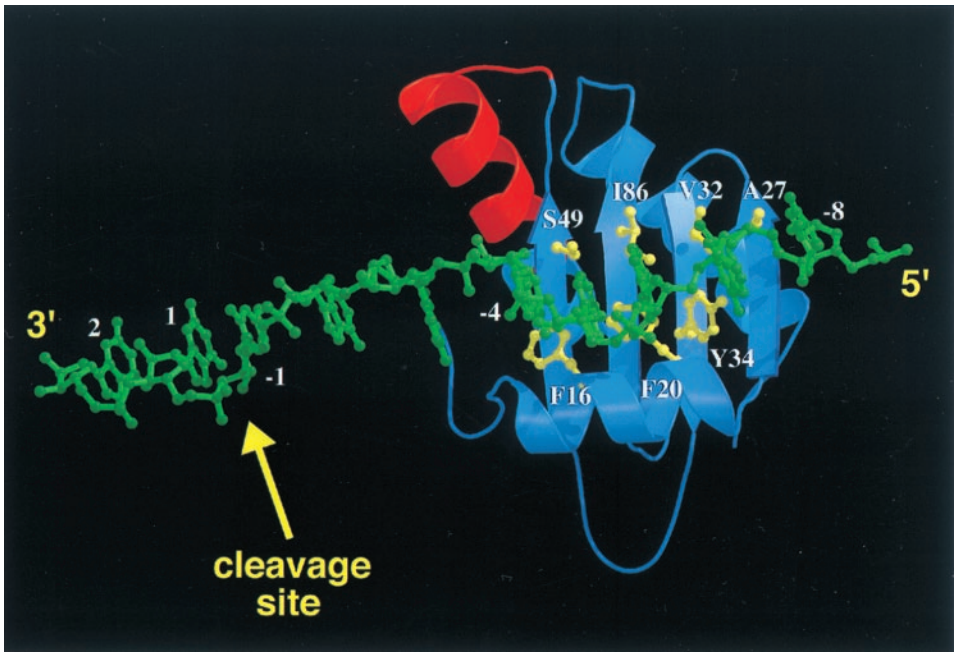

FIG. 2. Molecular structures are generated with MOLSCRIPT (54) and RASTER3D (55, 56). Single-stranded pre-tRNA leader sequence (green) modeled into the central cleft of P protein, oriented, and registered according to the crosslinking and primer extension mapping data discussed in the text. Residues yielding positive crosslinking results are indicated in yellow. The RNR motif of the protein is red; this is the only significant sequence segment conserved among all bacterial P proteins (39). The orientation and register of the substrate leader sequence in the central cleft is such that the RNR motif is 10–15 Å away from the cleavage site (between nucleotides 1 and −1). Therefore, the RNR motif is likely near the active site of the ribozyme, where it may possibly interact with conserved bases.

S102C). Residues V26 and A27 are located in β-strand 1 with V26 facing away from the cleft whereas A27 faces toward the cleft. Residue S102 lies on helix C and points away from the cleft. As shown in Fig. 3, RNase P reconstituted with AzP-A27C P protein (*G*) crosslinks to pre-tRNA substrates containing a leader of ≥ 4 nt whereas RNase P containing derivatized V26C and S102C P protein (*F* and *H*) crosslink only to substrates with longer leaders (pTR10 and pTR14). These data confirm that the single-stranded leader sequence of pre-tRNA binds in the cleft so as to orient the cleavage site on the side of the β-sheet near S49 and F16 and suggests that the longer leader sequence also may interact with helix C. Strikingly, this interaction orients the RNR motif, an 18-residue

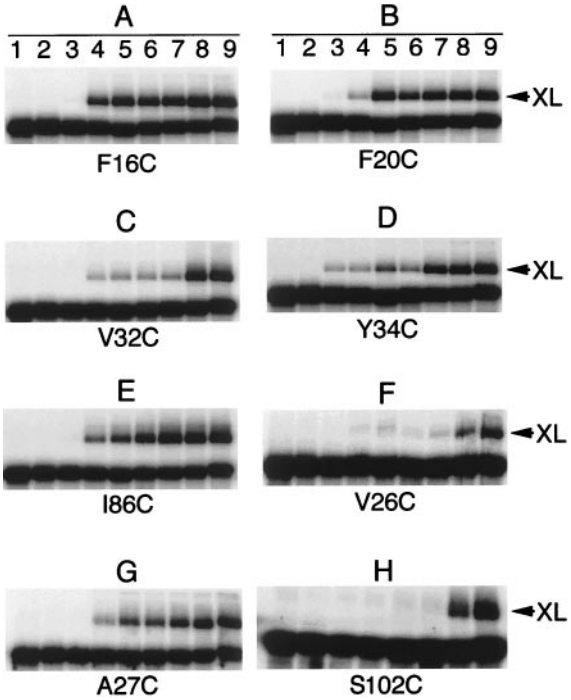

FIG. 3. Orientation of pre-tRNA in the central cleft of P protein. RNA/protein crosslinks were formed by irradiating truncated pre-tRNA substrates (see Fig. 1*B*) bound to RNase P holoenzyme reconstituted by using AzP-labeled P protein with the following unique cysteine substitutions: *A*, F16C; *B*, F20C; *C*, V32C; *D*, Y34C; *E*, I86C; *F*, V26C, *G*, A27C; and *H*, S102C.

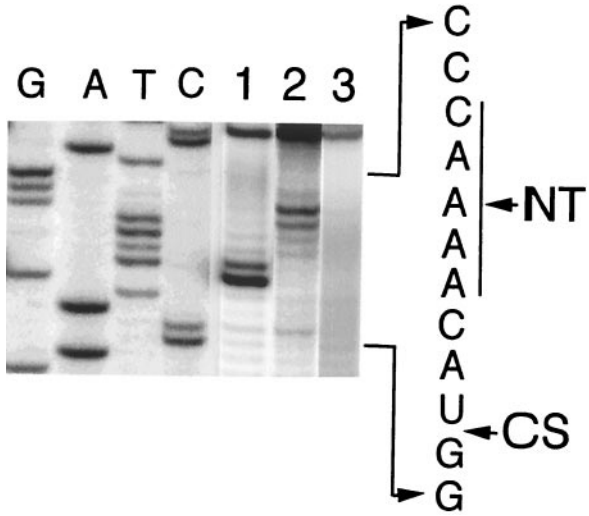

FIG. 4. Primer extension to map the crosslink sites in pTR14. Lanes G, A, T, and C represent dideoxynucleotide-terminated sequencing reactions by using unmodified pTR14 RNA as template. Lane 1, pTR14 crosslinked to holoenzyme reconstituted with AzP-S49C P protein; lane 2, pTR14 crosslinked to holoenzyme reconstituted with AzP-V26C P protein; and lane 3, uncrosslinked pTR14. The primer extension products are separated on an 8% polyacrylamide-urea gel. The corresponding nucleotide sequence of the pre-tRNA leader is shown. NT and CS represent the nucleotides in the cleft and cleavage site, respectively.

segment conserved among all known bacterial P proteins (39), toward the active site of the RNase P holoenzyme (Fig. 2).

Primer Extension Analysis Confirms Substrate Binding and Orientation. To determine the position of the crosslinks in pre-tRNA, we used reverse transcriptase to catalyze extension of an end-labeled primer complementary to the 3′ terminal sequence of pre-tRNA (32). We mapped the position of the crosslinks in pre-tRNA formed with RNase P reconstituted with AzP-derivatized S49C or V26C P protein where the former residue is located at one edge of the cleft and the latter residue is at the back side of the cleft. To identify the sites of crosslinking, dideoxynucleotide-terminated reactions were done by using unmodified pre-tRNA as template (Fig. 4, lanes G, A, T, and C) and analyzed in parallel with the primer extension products. Premature termination of the transcripts, compared with the same reaction using unmodified pre-tRNA, was observed in the crosslinked complex with S49C (Fig. 4, lane 1) and V26C (Fig. 4, lane 2) P protein at nucleotide positions −3 and −7 from the cleavage site, respectively. Reverse transcriptase cannot polymerize past the modifications in the RNA template so the enzyme terminates DNA synthesis at the nucleotide that immediately precedes a crosslink site (32, 33). Therefore, we conclude that the crosslinking sites are at nucleotides −4 and −8, respectively. These results further confirm the orientation and register of the single-stranded leader sequence bound in the central cleft of P protein.

DISCUSSION

The RNA component of RNase P has been the subject of extensive study that has provided a strong foundation for understanding ribozyme catalysis. The tertiary structure of the pre-tRNA substrate is believed to be the primary recognition determinant because the tRNA precursors do not share significant sequence similarity (40). The studies on substrate recognition by the RNA component of RNase P have demonstrated that several elements of pre-tRNA contribute to the specific recognition and efficient cleavage, including: the conserved -GTΨCR- sequence in the T loop (41, 42), the primary structure and length of the amino acid acceptor stem (43–45), the identity of the nucleotides at and around the cleavage site (5, 45, 46), and the 3′ terminal CCA sequence (45–47).

Although our mechanistic understanding of pre-tRNA recognition and cleavage catalyzed by the RNase P holoenzyme under physiological conditions is relatively less well studied, recent experiments help to clarify the functional role of the protein component in the RNP complex. Most notably, substrate recognition in the RNase P holoenzyme is altered in several respects: (*i*) the holoenzyme cleaves the precursors of tRNA, the *Escherichia coli* 4.5S RNA and the C4 repressor RNA of coliphage P1 and P7, with similar efficiencies, whereas P RNA catalyzes cleavage of pre-tRNA most efficiently (5, 22, 48); (*ii*) the holoenzyme binds pre-tRNA 10^4-fold more tightly than tRNA whereas the RNA component binds product more tightly (24); (*iii*) interactions of RNase P with the leader sequence (−1 to −5 nucleotides) and the T stem of pre-tRNA are altered (25, 49); and (*iv*) the protein component lowers the concentration of magnesium required for efficient catalysis but does not enormously enhance the cleavage rate constant (8, 24, 50). These functional changes in the holoenzyme could be the result of direct contacts between the substrate and the protein and/or to a protein-induced conformational change in the RNA. Previous chemical protection and hydroxyl radical footprinting studies in *E. coli* RNase P (in which the RNA subunit is designated M1 and the protein subunit is designated C5) suggested that the C5 protein binding site on M1 RNA was not near either the active site or the substrate binding site of the ribozyme but suggested that the conformation of the RNA altered upon binding the protein component (18–20).

Our results with *B. subtilis* RNase P contrast with those acquired from *E. coli* RNase P. Although no crosslinks between pre-tRNA and P protein were observed in the N-terminal region (R7), the conserved RNR motif (R65 and A70), or the loop preceding helix C (S92) of P protein, strong photocrosslinks observed between the single-stranded 5′ leader sequence of pre-tRNA and holoenzyme reconstituted with P protein containing AzP-derivatized cysteine at position 16, 20 (helix A), 27 (β-strand 1), 32, 34 (β-strand 2), 49 (β-strand 3), or 86 (β-strand 4) clearly implicate the central cleft of P protein in substrate binding (Figs. 1*B* and 4). Moreover, the orientation and register of the precursor, with nucleotides −4 to −8 passing from left to right in the central cleft as illustrated in Fig. 2, has been identified by crosslinking experiments with the truncated pre-tRNA substrates and further confirmed by primer extension analyses (Figs. 1*B*, 4, and 5). These data are consistent with binding data indicating that P protein/pre-tRNA interaction occurs at the −5 nucleotide (25). The strong crosslinks observed between the central cleft of P protein and the precursor part of pre-tRNA, but not to the mature tRNA, suggest that the protein component is an essential determinant for distinguishing the substrate and product. The central cleft of P protein may act as a clamp to hold the leader sequence, thereby facilitating the proper docking of the substrate to the active site.

The central cleft of P protein, approximately 20 Å long and 10 Å wide, is formed by α-helix A and the face of the β-sheet. These dimensions are comparable to those of the RNA binding clefts of the U1A and U2B′′ spliceosomal proteins, each of which binds RNA with high affinity and specificity; the crystal structures of these proteins complexed with their cognate RNA hairpins reveal intricate and extensive protein–RNA hydrogen bond networks and base-stacking interactions with aromatic residues in their central clefts (51, 52). However, the central cleft of P protein cannot bind RNA with sequence specificity like that of the spliceosomal proteins: this cleft must accommodate and orient many different substrate leader sequences lacking any similarity. Nevertheless, it is possible that many different substrate leader sequences can bind in this cleft, with purine or pyrimidine bases engaging equally well in base-stacking interactions with Phe-16, Phe-20, and Tyr-34. Two of these exposed aromatic residues (Phe-16 and Phe-20) are located on α-helix A, and Tyr-34 is the sole aromatic residue on the β-sheet. Induced fit conformational changes accompanying RNA binding involving side-chain torsion angles, as well as subtle changes in the helix-sheet interaction, may facilitate different geometries of base-stacking interactions, thereby ensuring broad specificity in RNA recognition. Alternatively, other sequence-nonspecific RNA binding modes are possible, e.g., as observed in an RNA complex with the methyltransferase VP39 (53).

An additional factor that may contribute to nonspecific recognition of variable 5′ pre-tRNA leader sequences may be the number of potential hydrogen bonding atoms in the central cleft of P protein. Intriguingly, in comparison with the spliceosomal proteins, the central cleft of P protein contains only half as many side-chain atoms capable of hydrogen bonding. Given that extensive, alternating protein-RNA hydrogen bond networks account for discrimination between cognate and noncognate RNA hairpins by U1A and U2B" (51, 52), a minimal number of polar residues deep within the central cleft of P protein will allow for some protein–RNA hydrogen bonds but will limit the formation of extensive, alternating protein–RNA hydrogen bond networks that otherwise could result in enhanced base discrimination and specificity. Of course, hypotheses regarding the structural basis of nonspecific RNA recognition by P protein in the RNase P holoenzyme remain to be examined by x-ray crystallographic studies.

The most exciting implication of the current study is that not only does the protein component directly contact the substrate, but it also is in close proximity to the active site because the cleavage site is only 3–4 nt from the central cleft. Furthermore, this orientation places the conserved RNR motif on helix B close to the catalytic site, suggesting that the protein component may participate more directly in the architecture of the active site than previously thought—-perhaps this motif interacts with conserved RNA bases in the RNase P active site. Given the close proximity of the P protein to the cleavage site (estimated at 10–15 Å), these data suggest that the active site in this RNP complex is close to the interface of RNA and protein, and the protein is not an innocent bystander in catalysis. Rather, the protein component is essential for altering substrate specificity and enhancing catalytic efficiency under physiological conditions. Interestingly, the proximity of the P protein to the site of catalytic function in RNase P may reflect an efficient mechanism of evolution from the RNA world to the protein world. These findings may have general implications for the function of other RNP complexes, such as the ribosome and spliceosome—-some of the protein components in these complexes similarly may enhance catalysis and substrate affinity rather than simply acting as inert, molecular "keystones" to stabilize the active RNA conformation.

We thank Dr. Norman Pace for giving us helpful plasmids. We thank David Cox, Jeffrey Kurz, and Dr. M. Puttaraju for advice and thoughtful reading of this manuscript. This work was supported by National Institutes of Health Grant GM 55387.

1. Arnez, J. G. & Cavarelli, J. (1997) *Q. Rev. Biophys.* **30,** 195–240.
2. Tinoco, I., Jr., Puglisi, J. D. & Wyatt, J. R. (1990) *Nucleic Acids. Mol. Biol.* **4,** 205–226.
3. Staley, J. P. & Guthrie, C. (1998) *Cell* **92,** 315–326.
4. Wilson, K. S. & Noller, H. F. (1998) *Cell* **92,** 337–349.
5. Altman, S., Kirsebom, L. & Talbot, S. (1993) *FASEB J.* **7,** 7–14.
6. Frank, D. N. & Pace, N. R. (1998) *Annu. Rev. Biochem.* **67,** 153–180.
7. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. (1983) *Cell* **35,** 849–857.
8. Guerrier-Takada, C. & Altman, S. (1984) *Science* **223,** 285–286.
9. Cech, T. R. (1993) *Gene* **135,** 33–36.
10. Nitta, I., Kamada, Y., Noda, H., Ueda, T. & Watanabe, K. (1998) *Science* **281,** 666–669.
11. Schedl, P. & Primakoff, P. (1973) *Proc. Natl. Acad. Sci. USA* **70,** 2091–2095.
12. Benne, R. (1988) *Trends Genet.* **4,** 181–182.
13. Burke, J. M. (1988) *Gene* **73,** 273–294.
14. Weeks, K. M. & Cech, T. R. (1995) *Cell* **82,** 221–230.
15. Caprara, M. G., Mohr, G. & Lambowitz, A. M. (1996) *J. Mol. Biol.* **257,** 512–531.
16. Weeks, K. M. & Cech, T. R. (1996) *Science* **271,** 345–348.
17. Herschlag, D., Khosla, M., Tsuchihashi, Z. & Karpel, R. L. (1994) *EMBO J.* **13,** 2913–2924.
18. Talbot, S. J. & Altman, S. (1994) *Biochemistry* **33,** 1406–1411.
19. Westhof, E., Wesolowski, D. & Altman, S. (1996) *J. Mol. Biol.* **258,** 600–613.
20. Kim, J. J., Kilani, A. F., Zhan, X., Altman, S. & Liu, F. (1997) *RNA* **3,** 613–623.
21. Pan, T. (1995) *Biochemistry* **34,** 902–909.
22. Peck-Miller, K. A. & Altman, S. (1991) *J. Mol. Biol.* **221,** 1–5.
23. Liu, F.-Y. & Altman, S. (1994) *Cell* **77,** 1093–1100.
24. Kurz, J. C., Niranjanakumari, S. & Fierke, C. A. (1998) *Biochemistry* **37,** 2393–2400.
25. Crary, S. M., Niranjanakumari, S. & Fierke, C. A. (1998) *Biochemistry* **37,** 9409–9416.
26. Stams, T., Niranjanakumari, S., Fierke, C. A. & Christianson, D. W. (1998) *Science* **280,** 752–755.
27. Barik, S. (1996) *Methods Mol. Biol.* **57,** 203–215.
28. Niranjanakumari, S., Kurz, J. C. & Fierke, C. A. (1998) *Nucleic Acids Res.* **26,** 3084–3090.
29. Ellman, G. L. (1959) *Arch. Biochem. Biophys.* **82,** 70–77.
30. Beebe, J. A. & Fierke, C. A. (1994) *Biochemistry* **33,** 10294–10304.
31. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
32. Inoue, T. & Cech, T. R. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 648–652.
33. Burgin, A. B. & Pace, N. R. (1990) *EMBO J.* **9,** 4111–4118.
34. Harris, M. E., Kazantsev, A. V., Chen, J. L. & Pace, N. R. (1997) *RNA* **3,** 561–576.
35. Rinke-Appel, J., Junke, N., Stade, K. & Brimacombe, R. (1991) *EMBO J.* **10,** 2195–2202.
36. Chen, Y. & Ebright, R. H. (1993) *J. Mol. Biol.* **230,** 453–460.
37. Hixson, S. H. & Hixson, S. S. (1975) *Biochemistry* **14,** 4251–4254.
38. Smith, D., Burgin, A. B., Haas, E. S. & Pace, N. R. (1992) *J. Biol. Chem.* **267,** 2429–2436.
39. Brown, J. W. (1998) *Nucleic Acids Res.* **26,** 351–352.
40. Komine, Y., Adachi, T., Inokuchi, H. & Ozeki, H. (1990) *J. Mol. Biol.* **212,** 579–598.
41. Steinberg, S., Misch, A. & Sprinzl, M. (1993) *Nucleic Acids Res.* **21,** 3011–3015.
42. Thurlow, D. L., Shilowski, D. & Marsh, T. L. (1991) *Nucleic Acids Res.* **19,** 885–891.
43. Holm, P. S. & Krupp, G. (1992) *Nucleic Acids Res.* **20,** 421–423.
44. Kirsebom, L. A. & Svärd, S. G. (1992) *Nucleic Acids Res.* **20,** 425–432.
45. Kirsebom, L. A. & Svärd, S. G. (1993) *J. Mol. Biol.* **231,** 594–604.
46. Kirsebom, L. A. & Svärd, S. G. (1994) *EMBO J.* **13,** 4870–4876.
47. Kufel, J. & Kirsebom, L. A. (1994) *J. Mol. Biol.* **244,** 511–521.
48. Hartmann, R. K., Heinrich, J., Schlegl, J. & Schuster, H. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 5822–5826.
49. Loria, A., Niranjanakumari, S., Fierke, C. A. & Pan, T. (1998) *Biochemistry,* **37,** 15466–15473.
50. Reich, C., Olsen, G. J., Pace, B. & Pace, N. R. (1988) *Science* **239,** 178–181.
51. Oubridge, C., Ito, N., Evans, P. R., Teo, C.-H. & Nagai, K. (1994) *Nature (London)* **372,** 432–438.
52. Price, S. R., Evans, P. R. & Nagai, K. (1998) *Nature (London)* **394,** 645–650.
53. Hodel, A. E., Gershon, P. D. & Quiocho, F. A. (1998) *Mol. Cell* **1,** 443–447.
54. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950.
55. Bacon, D. & Anderson, W. P. (1988) *J. Mol. Graphics* **6,** 219–220.
56. Merritt, E. A. & Murphy, M. E. P. (1994) *Acta Crystallogr. D* **50,** 869–873.