PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

STRUCTURE NOTE

Structural analysis of *Escherichia coli* C5 protein

Jae-Sun Shin,[1,2] Kwang-Sun Kim,[3,4] Kyoung-Seok Ryu,[5] Kook Han,[1] Younghoon Lee,[1] and Byong-Seok Choi[1]*

[1] Department of Chemistry, KAIST, 373-1 Guseong-dong, Yuseong-gu, Daejeon 305-701, Korea

[2] Biomedical Proteomics Research Center, KRIBB, 125 Gwahak-ro, Yuseong-gu, Daejeon 305-806, Korea

[3] Systems and Synthetic Biology Research Center, KRIBB, 125 Gwahak-ro, Yuseong-gu, Daejeon 305-806, Korea

[4] Biosystems and Bioengineering Program, University of Science and Technology (UST), Daejeon 305-350, Korea

[5] Division of Magnetic Resonance Research, Korea Basic Science Institute, Yangcheong-Ri 804-1, Ochang-Eup, Cheongwon-Gun, Chungbuk 363-883, Korea

Key words: ribonuclease P; C5 protein; *E. coli*; NMR; structure determination.

INTRODUCTION

RNase P is an enzyme that matures precursor tRNA (pre-tRNA) by hydrolysis of a specific phosphodiester bond. RNase Ps in bacterial species are holoenzymes that consist of a large RNA (350–400 nt; 100–130 kDa) and a protein component (~120 aa; 12–13 kDa).[1] The catalytic activity is conducted by the RNA, earning these proteins the title of ribozymes. The catalytic activity of the RNA component is maintained in the absence of the protein moiety in vitro at high ionic strength. However, the protein component is essential for pre-tRNA processing at physiological ionic strength and in vivo. The function of the bacterial RNase P protein is not entirely clear, but high affinity for tRNA at physiological ionic strength suggests that the holoenzyme is stabilized by the protein subunit via relief of electrostatic repulsion.

According to phylogenetic comparative analysis, the RNA subunit of RNase P is divided into two distinct structural classes: Ancestral type (A-type) and Bacillus type (B-type).[2] The A-type class, a group containing *E. coli*, is phylogenetically predominant, and the B-type class is found only in low-GC Gram-positive bacteria. In addition, secondary structures of A- and B-type RNA subunits are substantially different and the sequence similarities among RNase P proteins are low.[2] Nonetheless, protein subunits from both A- and B-type holoenzymes can stabilize the opposite type of RNA *in vitro*,[3] thus binding site of RNase P RNA and the protein are expected as similar and common features in the tertiary structure of RNase P protein are also expected. Structural studies of RNase P proteins have been limited to B-type protein, although RNase P protein in *E. coli* (C5 protein) is mostly well investigated through bacteria. Here, we firstly report the structure features determination of the C5 protein [Gene Bank ID: 948215] by nuclear magnetic resonance (NMR) in solution and explain why substrate recognition by RNase P is different in *E. coli* and *B. subtilis*.

MATERIALS AND METHODS

Cloning, expression, and purification of C5 protein

BL21(DE3) pLysS cells were transformed with pSSC5[4] plasmids and grown at 37°C in minimal M9 media containing 1 g/L ^{15}N ammonium chloride and 2 g/L ^{13}C glucose as the nitrogen and carbon sources, respectively, supplemented with 50 μg/L ampicillin and 30 μg/L chloramphenicol. Protein expression was induced by addition of 0.4 m*M* isopropyl β-D-thiogalactoside (IPTG) when

Grant sponsor: National Research Foundation of Korea (NRF; funded by the Korean government (MEST)); Grant numbers: 2009-0092818, KRF-2009-220-C00036, 2011-0020322; Grant sponsor: High Field NMR Research Program of Korea Basic Science Institute (NMR facility); Grant number: T3122E

*Correspondence to: B.-S. Choi, Department of Chemistry, KAIST, 373-1 Guseong-dong, Yuseong-gu, Daejeon 305-701, Korea. E-mail: byongseok.choi@kaist.ac.kr

Received 24 September 2011; Revised 6 November 2011; Accepted 8 November 2011

Published online 16 November 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.23253

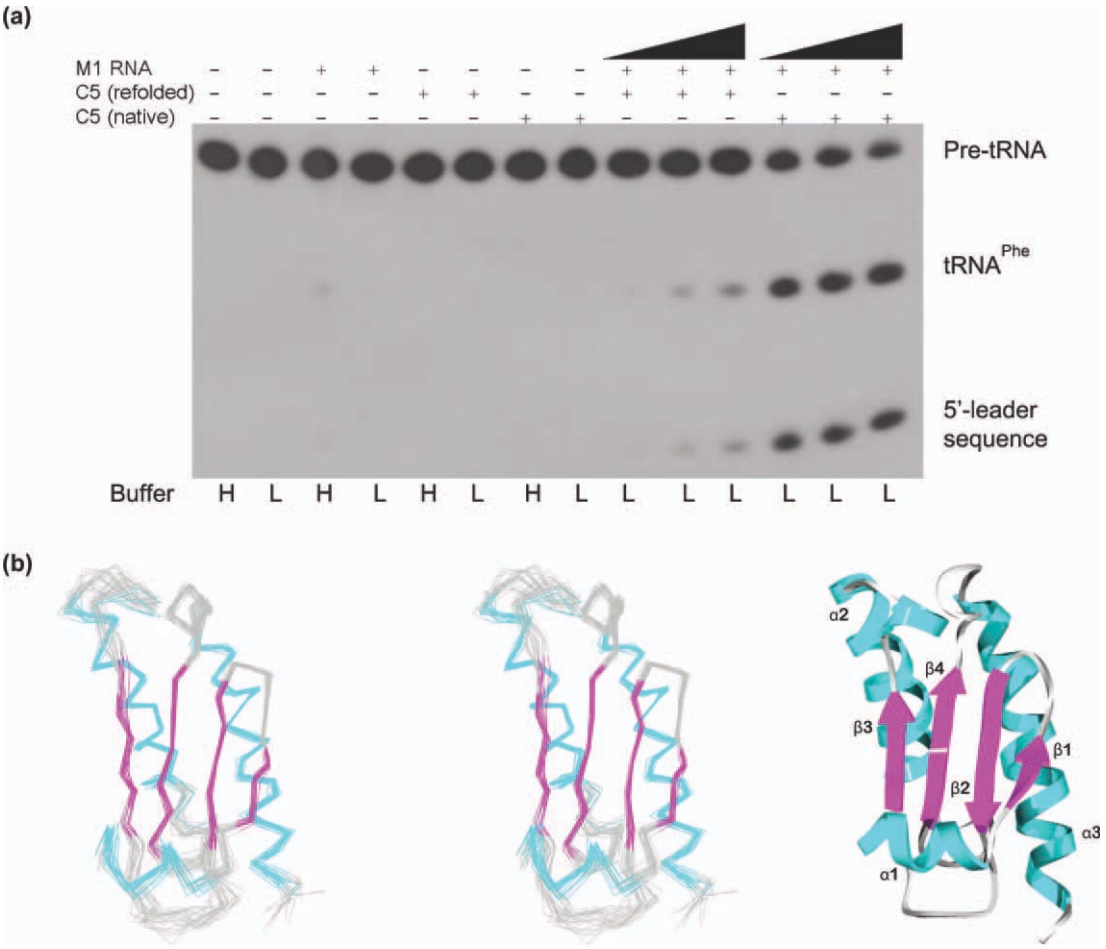

Figure 1

a: Cleavage reaction : RNase P cleavage of the precursor of $tRNA^{Phe}$. The catalytic reactions were carried out as described in "Materials and Methods." The concentration of M1 RNA (lanes 3 and 4) and C5 protein (lane 5–8) is fixed to 6 n*M* and 60 n*M*, respectively. C5 protein concentrations in the two triads of lanes (9–11 and 12–14) were 6, 30, and 60 n*M*. The lengths of pre-$tRNA^{Phe}$, $tRNA^{Phe}$, and 5′-leader sequence are 118, 76, and 42 nucleotides, respectively. **b**: NMR structure of C5 protein: (left) Superposition of the final 20 energy-minimized C5 protein structures. The structures are superimposed for minimal mutual deviation of the backbone atoms. The secondary structures are colored as follows: helices, cyan; β sheets, purple. Structures are shown in stereo view. Residues 17–119 are displayed. (right) The ribbon representation of the lowest energy conformer structure.

the cells reached $OD_{600} \sim 0.7$, harvested after 3 h by centrifugation (20 min at 2700*g*, 4°C) and stored at −20°C before further use. Thawed cells were suspended in 30 mL of buffer A (25 m*M* sodium phosphate, pH 7.2/150 m*M* NaCl/1 m*M* DTT) and lysed on ice by sonication, and cell debris was removed by centrifugation (30 min at 29,000*g*, 4°C). C5 protein was precipitated by 30% (w/v) ammonium sulfate and harvested by centrifugation (15 min at 20,000*g*, 4°C). Precipitated C5 protein was resuspended in 30 mL of buffer A and purified by ion-exchange chromatography with gradient elution from 150 m*M* to 2*M*, followed by gel filtration chromatography. Purified C5 protein was concentrated and dialyzed against buffer B (25 m*M* sodium phosphate, pH 7.2/200 m*M* NaCl/1 m*M* DTT).

NMR measurements and structure determination

NMR experiments were performed on Varian INOVA600 and Bruker Avance 800 MHz spectrometers equipped with a pulse-field gradient triple-resonance probe. All NMR spectra were obtained at pH 7.2 and 20°C, using a solution of ~0.8 m*M* ^{15}N or $^{15}N/^{13}C$-labeled C5 protein in 90% H_2O/10% ^2H_2O or 100% ^2H_2O. NMR data were processed with NMRPipe,[5] and the spectra were analyzed with Sparky.[6] The backbone resonance assignments of C5 protein were obtained from the HNCACB, CBCA(CO)NH, HNCO, and HN(CA)CO spectra. Side-chain proton resonances were assigned with HCCH-TOCY, HCCH-COSY, and (H)CCH-TOCSY spectra. 3D ^{15}N-edited NOESY and ^{13}C-edited NOESY spectra were

acquired for NOE constraints. For structure calculation, the NOE peak assignment was obtained by a manual peak picking procedure. ψ, φ-angle constraints for the protein backbone were generated from C^{α} secondary chemical shifts using the program TALOS.[7] Hydrogen bond constraints were introduced throughout the calculation, based on hydrogen/deuterium exchange experiments. NOE, dihedral angle, and hydrogen bond restraints were used for structure calculation using CYANA 2.1.[8] Calculated structures were refined with a water refinement protocol using CNS,[9] and PROCHECK-NMR,[10] and Chimera[11] programs were used to analyze the final structures. Chemical shifts of C5 protein have been deposited in the BioMagResBank (accession code 17952 and coordinates have been submitted to the Protein Data Bank (PDB accession code 2ljp).

Cleavage activity of holoenzyme RNase P

To compare protein activity of native and refolded C5 protein, both native and refolded C5 protein were assayed for their ability to allow the M1 RNA to cleave the 5′-leader sequences of an *E. coli* pre-$tRNA^{Phe}$. Reactions were carried out as described previously[12] in low ionic strength buffer (50 m*M* Tris-HCl, pH 8.0/10 m*M* $MgCl_2$/100 m*M* NH_4Cl) or high ionic strength buffer (50 m*M* Tris-HCl, pH 8.0/100 m*M* $MgCl_2$/100 m*M* NH_4Cl/5% polyethylene glycol).

RESULTS AND DISCUSSION

Expression of soluble C5 protein

Although many biochemical studies have been done on *E. coli* C5 protein, its structural features have been inaccessible because of its insolubility.[13] A preliminary structure of a mutant construct was reported, but this might not be the same as the wild-type structure.[13] Until now, C5 protein has been obtained only by refolding or as a tagged protein (MBP-C5) because intact C5 protein became aggregated during concentration in presence of cellular RNAs.[13] To overcome this problem, we developed an expression and purification protocol that produces soluble C5 protein. Using it, we repeatedly expressed soluble and intact C5 protein from a pSSC5 plasmid, which suppresses leaky expression of protein before induction.[4] Cellular RNAs were removed by ammonium sulfate precipitation and anion exchange chromatography to prevent protein aggregation during concentration and C5 protein was purified with cation exchange chromatography and gel filtration chromatography.

Cleavage of pre-tRNA by RNase P

We measured the protein activity of native and refolded C5 protein, prepared as previously described.[4]

Table I
Structural Statistics for C5 Protein

Distance restraints	
All	1422
Intraresidue	653
Sequential ($\|i-j\|=1$)	355
Medium range ($2\leq\|i-j\|\leq4$)	172
Long range ($\|i-j\|>4$)	242
Hydrogen bonds	42
Dihedral angle restraints	
All	168
Φ	84
Ψ	84
Structure statistics	
Violations (mean and s.d.)	
Distance violation (>0.3 Å)	0
Dihedral angle violation (>5°)	0
Energies	
Total energy(kcal/mol)	−3895.58 ± 93.10
Residual target function (Å^2)	0.43 ± 0.09
Average pairwise RMSD[a]	
Backbone atoms	0.61 ± 0.07 Å
All heavy atoms	1.39 ± 0.14 Å
Ramachandran plot (%)	
Residues in most favored regions	83.3
Residues in additional allowed regions	16.0
Residues in generously allowed regions	0.6
Residues in disallowed regions	0.1

[a]RMSD is calculated for residues 17–117.

Figure 1(a) shows that RNase P activity was over 10-fold higher with holoenzyme containing native protein [Fig. 1(a), lanes 12–14] compared with the holoenzyme containing refolded protein [Fig. 1(a), lanes 9–11]. No 5′-leader sequence cleavage product was observed with C5 protein or M1 RNA alone in low ionic strength buffer. These data suggest that C5 protein refolded *in vitro* is not properly folded and the resulting RNase P protein is less active than that containing native C5 protein. Buck et al.[3] reported previously that *E. coli* RNase P RNA folding is induced by the RNase P protein subunit, so it follows that *E. coli* RNase P holoenzyme may also be formed by induced fit, rendering the enzyme's activity dependent on a correct protein fold.

Structural features of C5 protein

The NMR structure of C5 protein was determined based on the restraints in Table I. The ensemble of the 20 lowest-energy structures is displayed in Figure 1(b). The RMSD of 20 representative conformers was 0.61 ± 0.07Å for the backbone atoms and 1.39 ± 0.14Å for heavy atoms in residues 17 to 117. C5 protein shares the α/β-fold common to bacterial RNase P proteins.[14] The structure of C5 protein consists of four β strands and three α helices which are arranged as α1-β1-β2-β4-α2-β3-α3, similar to other bacterial RNase P proteins such as *B. subtilis*,[14] *S. aureus*,[15] and *T. maritima*.[16] One face, β4-α2-β3, contains the "RNR" sequence highly

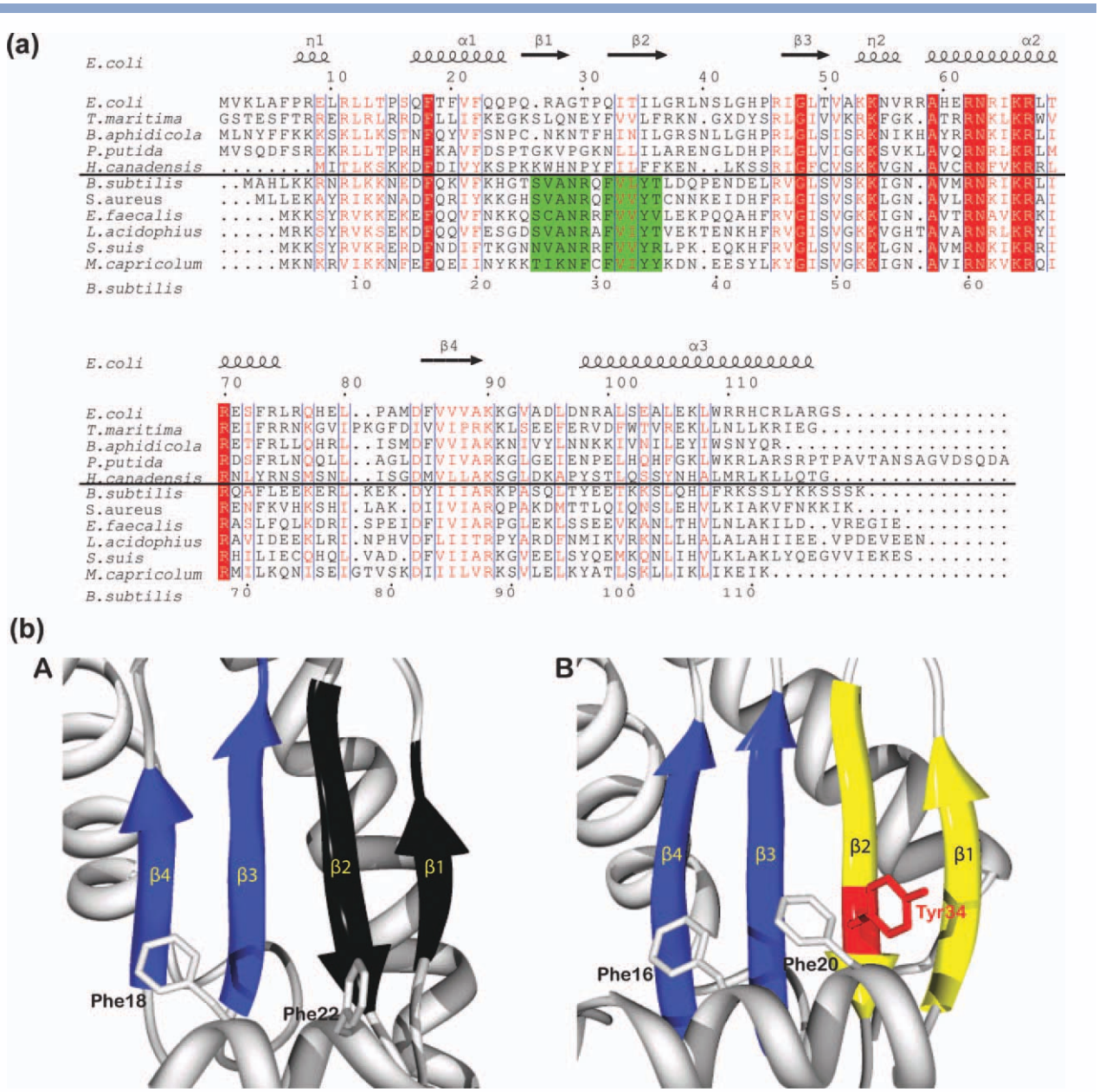

Figure 2

a: Sequence alignment of RNase P proteins: Sequence alignment of RNaseP proteins among bacteria. Highly conserved residues are colored red. Conserved residues of in the central cleft of the B-type RNase P proteins are colored green. **b**: Structural comparison between *E. coli* and *B. subtilis* RNase P proteins: Residues of β1, β2 strands are not conserved in (A) *E. coli* RNase P protein (black) but are conserved in (B) *B. subtilis* RNase P protein (yellow). Tyr34 of *B. subtilis* RNaseP protein is colored red.

conserved among bacterial RNase P proteins. The other conserved face, all four β strands with α1, forms a hydrophobic central cleft which has been suggested by photoaffinity crosslinking experiments to interact with the 5′-leader sequence of precursor tRNAs.[17]

The results of many previously reported mutagenesis studies on *E. coli* C5 protein can now be understood in light of the solution structure presented here. The R46H mutant of C5 protein was reported to be temperature-sensitive.[18] Based on the structure, this behavior is the consequence of a salt bridge between R46 and D84 that is responsible for maintaining protein stability. Also, F18A and F22A mutations were shown to dramatically decrease the cleavage activity of RNase P.[19] The solution structure of C5 protein shows that F18 and F22 are located in the central cleft (consisting of α1 and β4) that is bound to 5′-leader sequence of pre-tRNA.

The similarity of tertiary structures among bacterial RNase P proteins was established by RNase P RNA–protein swapping experiments between *B. subtilis* and

E. coli.[3] This similarity of protein structure suggests that binding interface of the M1 RNA and C5 protein components is conserved among bacteria and that the role of RNase P proteins is similar across species. However, Buck et al.[3] reported that substrate recognition by RNase P proteins is different between *E. coli* (A-type) and *B. subtilis* (B-type). A sequence alignment of bacterial RNase P proteins [Fig. 2(a)] provided insight into the differences in substrate recognition between A- and B-type RNase Ps. There is a major difference at the β1 and β2 strands of the central cleft region. The "SxANR" sequence at β1 and the "FVxY" sequence at β2 are highly conserved among B-type RNase Ps, but these sequences are not present in A-type RNase Ps [Fig. 2(a)]. The —OH group of the β2 tyrosine is known to be involved in a hydrogen bond with the 5′-leader sequence in *B. subtilis* RNase P,[20] but β2 residues are not conserved in *E. coli* RNase P. The difference of conservation suggests that the mechanism of substrate recognition in *E. coli* RNase P is distinct from that of the *B. subtilis* protein.

In summary, we determined the high resolution structure of C5 protein from *E. coli* and give a specific structural interpretation of the C5 protein's activity and specificity by comparison with *B. subtilis*. These results provide a critical structural basis on which to understand RNA–protein interactions of RNase P.

ACKNOWLEDGMENTS

Melissa Stauffer, PhD, of Scientific Editing Solutions, edited the manuscript.

REFERENCES

1. Frank DN, Pace NR. Ribonuclease P: unity and diversity in a tRNA processing ribozyme. Annu Rev Biochem 1998;67:153–180.
2. Haas ES, Banta AB, Harris JK, Pace NR, Brown JW. Structure and evolution of ribonuclease P RNA in Gram-positive bacteria. Nucleic Acids Res 1996;24:4775–4782.
3. Buck AH, Dalby AB, Poole AW, Kazantsev AV, Pace NR. Protein activation of a ribozyme: the role of bacterial RNase P protein. EMBO J 2005;24:3360–3368.
4. Park BH, Choi YN, Park JW, Sim S, Gil MC, Kim S, Kim M, Lee Y. Expression of C5 protein, the protein component of *Escherichia coli* RNase P, from the tac promoter. Mol Cells 1998;8:96–100.
5. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 1995;6:277–293.
6. Goddad TD, Kneller DG. SPARKY3, 2004.
7. Cornilescu G, Delaglio F, Bax A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR 1999;13:289–302.
8. Herrmann T, Guntert P, Wuthrich K. Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. J Biomol NMR 2002;24:171–189.
9. Linge JP, Williams MA, Spronk CA, Bonvin AM, Nilges M. Refinement of protein structures in explicit solvent. Proteins 2003;50:496–506.
10. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 1996; 8:477–486.
11. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem 2004;25:1605–1612.
12. Vioque A, Altman S. Affinity chromatography with an immobilized RNA enzyme. Proc Natl Acad Sci USA 1986;83:5904–5908.
13. Choe HW, Jeong DG, Park JH, Schlesinger R, Labahn J, Hofmann KP, Buldt G. Preliminary X-ray characterization of the ribonuclease P (C5 protein) from *Escherichia coli*: expression, crystallization and cryoconditions. Acta Crystallogr D Biol Crystallogr 2003;59 (Part 2):350–352.
14. Stams T, Niranjanakumari S, Fierke CA, Christianson DW. Ribonuclease P protein structure: evolutionary origins in the translational apparatus. Science 1998;280:752–755.
15. Spitzfaden C, Nicholson N, Jones JJ, Guth S, Lehr R, Prescott CD, Hegg LA, Eggleston DS. The structure of ribonuclease P protein from *Staphylococcus aureus* reveals a unique binding site for single-stranded RNA. J Mol Biol 2000;295:105–115.
16. Kazantsev AV, Krivenko AA, Harrington DJ, Carter RJ, Holbrook SR, Adams PD, Pace NR. High-resolution structure of RNase P protein from *Thermotoga maritima*. Proc Natl Acad Sci USA 2003;100: 7497–7502.
17. Niranjanakumari S, Stams T, Crary SM, Christianson DW, Fierke CA. Protein component of the ribozyme ribonuclease P alters substrate recognition by directly contacting precursor tRNA. Proc Natl Acad Sci USA 1998;95:15212–15217.
18. Baer MF, Wesolowski D, Altman S. Characterization in vitro of the defect in a temperature-sensitive mutant of the protein subunit of RNase P from *Escherichia coli*. J Bacteriol 1989;171:6862–6866.
19. Gopalan V, Baxevanis AD, Landsman D, Altman S. Analysis of the functional role of conserved residues in the protein subunit of ribonuclease P from *Escherichia coli*. J Mol Biol 1997;267:818–829.
20. Koutmou KS, Zahler NH, Kurz JC, Campbell FE, Harris ME, Fierke CA. Protein-precursor tRNA contact leads to sequence-specific recognition of 5′ leaders by bacterial ribonuclease P. J Mol Biol 2010; 396:195–208.