Crystal Structure of a Ribonuclease P Protein Ph1601p from *Pyrococcus horikoshii* OT3: An Archaeal Homologue of Human Nuclear Ribonuclease P Protein Rpp21[†,‡]

Yoshimitsu Kakuta,[§] Ikuko Ishimatsu,[§] Tomoyuki Numata,[§] Kazumi Kimura,[§] Min Yao,[||] Isao Tanaka,[||] and Makoto Kimura*[,§]

Laboratory of Biochemistry, Department of Bioscience and Biotechnology, Faculty of Agriculture, Graduate School, Kyushu University, Fukuoka 812-8581, Japan, and Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan

Received April 21, 2005; Revised Manuscript Received July 14, 2005

ABSTRACT: Ribonuclease P (RNase P) is a ribonucleoprotein complex involved in the removal of 5′ leader sequences from tRNA precursors (pre-tRNA). The human protein Rpp21 is essential for human RNase P activity in tRNA processing in vitro. The crystal structure of Ph1601p from the hyperthermophilic archaeon *Pyrococcus horikoshii* OT3, the archaeal homologue of Rpp21, was determined using the multiple anomalous dispersion (MAD) method with the aid of anomalous scattering in zinc and selenium at 1.6 Å resolution. Ph1601p comprises an N-terminal domain (residues 1−55), a central linker domain (residues 56−79), and a C-terminal domain (residues 80−120), forming an L-shaped structure. The N-terminal domain consists of two long α-helices, while the central and C-terminal domains fold in a zinc ribbon domain. The electrostatic potential representation indicates the presence of positively charged clusters along the L arms, suggesting a possible role in RNA binding. A single zinc ion binds the well-ordered binding site that consists of four Cys residues (Cys68, Cys71, Cys97, and Cys100) and appears to stabilize the relative positions of the N- and C-domains. Mutations of Cys68 and Cys71 or Cys97 and Cys100 to Ser destabilize the protein structure, which results in inactivation of the RNase P activity. In addition, site-directed mutagenesis suggests that Lys69 at the central loop and Arg86 and Arg105 at the zinc ribbon domain are strongly involved in the functional activity, while Arg22, Tyr44, Arg65, and Arg84 play a modest role in the activity.

Ribonuclease P (RNase P),[1] a ubiquitous cellular ribonucleoprotein complex, catalyzes the removal of 5′ leader sequences from tRNA precursors (pre-tRNA) (*1*, *2*). Since Altman and co-workers discovered that the *Escherichia coli* RNae P RNA itself can hydrolyze pre-tRNA in vitro (*3*), biochemical and structural studies on RNase P have so far been mainly focused on eubacterial RNase Ps (for a review, see ref *4*). These studies identified crucial nucleotides at helix P4 in RNase P RNA (*5*−*8*); three-dimensional structures of the eubacterial protein subunit were established (*9*−*11*), and a three-dimensional model for the RNase P holoenzyme has been proposed on the basis of combining footprinting data with information from biochemical and biophysical studies (*12*). In addition, high-resolution crystal structures of specificity domains of *Bacillus subtilis* and *Thermus thermophilus* RNase P RNAs were established (*13*, *14*). Recent findings indicated that a main function of the protein component in eubacterial RNase P is to interact with the 5′ leader sequence which enhances the affinity of both pre-tRNA and specific magnesium ions bound to the RNase P−pre-tRNA complex (*15*).

In contrast, only a few studies have thus far been carried out with eukaryotic and archaeal RNase Ps because it is difficult to obtain a large amount and there is complexity in their subunit components. Eukaryotic RNase Ps, such as nuclear RNase Ps from humans and *Saccharomyces cerevisiae*, are composed of at least nine protein subunits associated with an RNA subunit, and in the absence of protein subunits, the RNA subunit does not alone exhibit enzyme activity in vitro (*16*−*18*). These findings suggest that the eukaryotic RNA and protein subunits work in an interdependent manner and fulfill the requirements for substrate recognition and/or catalysis. Recently, Mann et al. reported that human proteins Rpp21 and Rpp29, together with the RNA subunit H1 RNA, are sufficient to obtain weak cleavage of 5′ leader sequence of pre-tRNA in vitro (*19*). In addition, a spatial organization was reported on human and yeast RNase Ps (*20*−*22*). However, a common arrangement of protein subunits in the eukaryotic RNase P has yet to emerge.

As for archaeal RNase P, we earlier found in reconstitution experiments that RNase P RNA and proteins Ph1481p,

[†] This work was supported in part by a grant from the National Project on Protein Structural and Functional Analyses from the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

[‡] The atomic coordinates and structure factors have been deposited in the Protein Data Bank as entry 1XOT.

* To whom correspondence should be addressed: Laboratory of Biochemistry, Department of Bioscience and Biotechnology, Faculty of Agriculture, Graduate School, Kyushu University, Hakozaki 6-10-1, Higashi-ku, Fukuoka 812-8581, Japan. Telephone and fax: +81-92-642-2853. E-mail: mkimura@agr.kyushu-u.ac.jp.

[§] Kyushu University.

[||] Hokkaido University.

[1] Abbreviations: CD, circular dichroism; MAD, multiple anomalous dispersion; pre-tRNA, precursor transfer RNA; RNase P, ribonuclease P; Rpb9, *Thermococcus celer* RNA polymerase II subunit 9; Se-Met, selenomethionine.

10.1021/bi050738z CCC: $30.25
Published on Web 08/18/2005

Table 1: Summary of Data Collection

	native	remote	Zn peak	Zn edge	Se peak	Se edge
space group			$P2_12_12_1$			
resolution range (Å)	20−1.5	20−2.6	20−2.6	20−2.5	35−2.7	35−2.5
unit cell parameters (Å)	$a = 42.59$		$a = 43.24$			
	$b = 52.69$		$b = 49.14$			
	$c = 62.67$		$c = 63.58$			
wavelength (Å)	1.0000	0.9000	1.2818	1.2825	0.9789	0.9792
no. of observations	159773	30348	30607	34129	27207	31972
no. of unique reflections	23409	4463	4501	5019	4001	4772
redundancy	6.8	6.8	6.8	6.8	6.8	6.7
data completeness (%)[a]	98.8 (87.7)	98.8 (100.0)	99.0 (100.0)	98.4 (98.8)	99.2 (100.0)	93.8 (98.8)
$\langle I/\sigma(I)\rangle$[a]	23.2 (2.5)	15.8 (6.3)	11.3 (7.9)	12.4 (5.1)	16.3 (6.7)	13.7 (6.0)
R_{sym} (%)[a,b]	5.3 (51.5)	8.7 (34.6)	9.7 (31.1)	9.8 (36.6)	9.7 (32.3)	8.2 (31.6)

[a] Values in parentheses correspond to the highest-resolution shell. [b] $R_{sym} = \sum|I - I_{av}|/\sum I$, where the summation is over all equivalent reflections.

Ph1601p, and Ph1771p are minimal components for the RNase P activity of the hyperthermophilic archaeon *Pyrococcus horikoshii* OT3, and that addition of the fourth protein, Ph1877p, strongly stimulated the activity (*23*). Subsequently, the fifth protein, Ph1496p (an archaeal homologue for human Rpp38), was found to be involved in elevation of the optimum temperature of the reconstituted RNase P (H. Fukuhara et al., unpublished data). Thus, the *P. horikoshii* RNase P RNA, like eukaryotic counterparts, is deficient in functions and cooperatively functions with five protein subunits in substrate recognition and/or catalysis. It is thus of interest to understand how protein subunits in eukaryotic and archaeal RNase Ps exert their functions in the catalytic activity. These investigations will aid in establishing structure−function relationships of the ribonucleoprotein ribozyme, and the resulting information will ultimately shed light on the transition from the proposed RNA world to the modern protein world.

We have been studying structures of individual proteins from the *P. horikoshii* RNase P, and we determined the crystal structures of Ph1771p and Ph1877p at 2.0 and 1.8 Å resolution, respectively (*24*, *25*). Furthermore, the essential amino acid residues in Ph1877p for the *P. horikoshii* RNase P activity were assigned by site-directed mutagenesis (*25*). In the study presented here, we extended this structural study to Ph1601p, an archaeal homologue of the human Rpp21 (*26*). Ph1601p consists of 120 amino acid residues with a calculated molecular mass of 14 589 Da, and its sequence is 27.5 and 17.5% identical to those of eukaryotic RNase P proteins Rpp21 and Rpr2 from human and *S. cerevisiae*, respectively. We report here the crystal structure of Ph1601p determined at 1.6 Å resolution with the aid of anomalous signals from selenomethionines (Se-Met) and zinc ion. Furthermore, amino acids Tyr44, Lys69, Arg86, Arg105, and four Cys residues (Cys68, Cys71, Cys97, and Cys100) were thought to be essential amino acids for the catalytic activity of RNase P, as determined by site-directed mutagenesis.

MATERIALS AND METHODS

Protein Preparation and Crystallization. Expression, purification, and crystallization of Ph1601p were essentially carried out as previously described (*23*−*25*). The expression and purification of Se-Met-derivatized Ph1601p were carried out in a manner identical to those described for the native protein, except that *E. coli* B834 strain (Novagen) grown in minimal medium was used as the host. The purified Se-Met-derivatized Ph1601p was concentrated to 10 mg/mL and crystallized by vapor diffusion against 0.2 M tri-ammonium citrate (pH 7.0), 20% (w/v) polyethylene glycol 3350, 10 mM $ZnCl_2$, and 10% (v/v) ethanol. Using the hanging drop vapor diffusion method, small crystal nuclei from these crystals were used to microseed fresh drops of the protein and reservoir solution, and then the drops were equilibrated at 20 °C. Finally, crystals suitable for a high-resolution X-ray diffraction measurement were obtained using a reservoir solution of 25.5% (w/v) polyethylene glycol 4000, 85 mM tri-sodium citrate (pH 5.6), 0.17 M ammonium acetate, 10 mM $ZnCl_2$, and 15% (v/v) glycerol.

Data Collection and Refinement. The crystal mounted on a cryoloop was flash-cooled in a nitrogen gas stream at −180 °C. X-ray diffraction data were collected with an ADSC system and synchrotron radiation at beamline BL38B1 in SPring-8. X-ray diffraction data for the initial structure determination were collected to ∼2.6 Å resolution from crystals containing Se-Met-derivatized Ph1601p with zinc ion. Five data sets were collected on and around the selenium K absorption edge and zinc K absorption edge to determine the structure of the protein, using the multiple anomalous dispersion (MAD) method. A high-resolution native data set was collected to 1.5 Å resolution from the crystal containing the native Ph1601p with zinc ion. All data were processed using HKL 2000 (*27*). Positions for two selenium atoms and one zinc atom were determined using CNS (*28*). CNS was then used for refinement of the selenium sites and zinc sites, and solvent flattening and histogram matching were carried out. The model was built using O (*29*) and ARP/wARP (*30*). The stereochemistry of the refined model was verified using PROCHECK (*31*). Data collection, phase calculation, and refinement statistics are summarized in Tables 1, 2, and 3, respectively. The coordinates have been deposited in the Protein Data Bank as entry 1XOT.

Mutant Proteins. The site-directed mutagenesis of Ph1601p was carried out using a Quick mutagenesis kit (Stratagene), and the resulting mutants were purified using the same protocol that was used for wild-type Ph1601p (*23*). The RNase P activity of the reconstituted RNase P containing the Ph1601p mutants was evaluated as we previously described, using as a substrate the pre-$tRNA^{tyr}$ from *P. horikoshii* OT3 (*23*). Circular dichroism (CD) spectra in the far-ultraviolet range, from 200 to 250 nm, were recorded at room temperature on a Jasco J-720 spectropolarimeter. Proteins were dissolved in 20 mM Tris-HCl (pH 7.0) containing 200 mM NaCl, and signal averaging during the accumulation of four scans was done automatically.

Table 2: Summary of Phase Calculation

	remote	Zn peak	Zn edge	Se peak	Se edge
resolution range (Å)			20−2.6		
no. of sites (Se/Zn)			2/1		
R_{cullis}					
(centric/acentric)	−	0.45/0.56	0.37/0.51	0.52/0.65	0.45/0.58
Fried (centric/acentric)	−/0.55	0.47/0.47	0.38/0.47	0.52/0.54	0.46/0.49
phasing power					
(centric/acentric)	−	2.32/2.07	2.67/2.39	1.69/1.64	2.02/1.93
Fried (centric/acentric)	2.17/2.02	2.84/2.61	2.87/2.62	2.26/2.20	2.59/2.46
figure of merit					
(centric/acentric)	−	0.69/0.34	0.75/0.38	0.68/0.33	0.66/0.33
Fried (centric/acentric)	−/0.31	0.71/0.40	0.76/0.40	0.69/0.39	0.67/0.39
final (centric/acentric)	0.95/0.81				

Table 3: Final Refinement Statistics

resolution range (Å)	50−1.6
rmsd for bonds (Å)	0.004
rmsd for angles (deg)	1.07
R-factor (R_{free})	21.7 (2.24)
quality of Ramachandran plot	
percentage of residues in most favored regions	92.8
percentage of residues in additional allowed regions	6.2
percentage of residues in generously allowed regions	1.0
percentage of residues in disallowed regions	0.0
no. of atoms	
protein	906
water	65
Zn	1

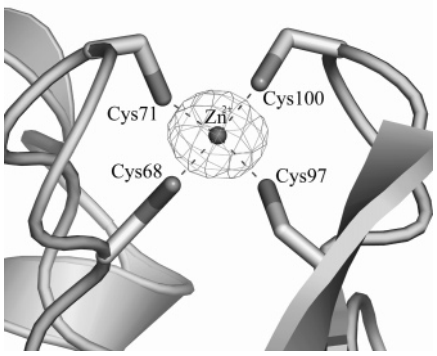

FIGURE 1: Electron density map for a single zinc ion coordinated by four Cys residues. A portion of the final $F_o - F_c$ electron density map around the zinc ion binding site constituted by four Cys residues, contoured at 4σ, superimposed on the refined model. All figures were prepared using Pymol (http://www.pymol.org).

RESULTS

Structure Determination. Protein Ph1601p was overproduced and purified as previously described in ref *23*. The protein Ph1601p gave a few bands that differed in mobility on SDS−PAGE, despite an apparent homogeneity as judged by the N-terminal sequencing. Since Ph1601p contains a zinc-binding Cys cluster at the C-terminal region, the heterogeneity observed on SDS−PAGE is probably caused by distinct conformations due to the presence or absence of zinc ion at the binding site. Indeed, addition of zinc ion in the reservoir solution gave crystals that are suitable for crystallographic analysis. The Ph1601p crystals belong to space group $P2_12_12_1$ with the following unit cell parameters: $a = 42.59$ Å, $b = 52.69$ Å, and $c = 62.67$ Å. The crystal contains one molecule in the asymmetric unit, with a solvent content of 47.6% and a V_M value of 2.4 Å^3 Da^{-1}, which is in the range of most common protein crystals. The crystal structure was determined from an electron density map that was calculated using data collected from the crystals of Ph1601p containing Se-Met and zinc, and phased by the five-wavelength MAD method. The initial phasing provided a partial electron density map for C-terminal β-strands, including a single zinc ion bound to four Cys residues (Cys68, Cys71, Cys97, and Cys100) (Figure 1). After multiple cycles of refinement using CNS, the *R*-factor and R_{free} were refined to 37.8 and 42.6%, respectively. An interpretable density for the N-terminal region was obtained using a high-resolution native data to 1.5 Å resolution which was phased by molecular replacement using the initial C-terminal structure as a search model. The current refinement model of Ph1601p consists of 107 amino acids (residues 4−110), one zinc ion, and 65 water molecules. The working and free *R*-factors at 1.6 Å resolution are 21.7 and 24.5%, respectively (Table 3). The refined structure has 99.0% of its residues in allowed and additionally allowed regions of the Ramachandran plot.

Overall Structure. Figure 2 shows a ribbon diagram of the α-carbon tracing of the Ph1601p protein. The secondary structure of Ph1601p, as defined by PROCHECK, is given in Figure 3. Ph1601p forms an L-shaped structure, with approximate dimensions of 49 Å × 39 Å × 14 Å. Ph1601p consists of three distinct domains, an N-domain (residues 1−50), a central domain (residues 51−79), and a C-domain (residues 80−120). The N-domain has two long α-helices (α1 and α2), interacting with each other through hydrophobic residues. They include Ala19, Ile20, Ile23, Leu26, Phe27, Leu29, Ala39, and Val33 on α1 and Leu40, Ala41, Tyr44, Val45, Leu47, Ala48, Ile51, and Ala55 on α2. The central domain has an unstructured loop followed by the C-domain which adopts a zinc ribbon domain (*32*). The zinc ribbon domain consists of a three-stranded antiparallel β-sheet, and the single zinc ion is coordinated by the side chain sulfate atoms of four Cys residues (Cys68, Cys71, Cys97, and Cys100). The zinc ion appears to play a role in stabilization of the conformation of the linker peptide between N- and C-domains, which results in fixing a relative orientation of the N- and C-domains, which do not make extensive hydrophobic contacts with one another.

Figure 3 shows the sequence alignment of its homologous proteins, including three archaeal proteins as well as two eukaryotic proteins. This comparison shows that there are 17 absolutely conserved residues which are distributed throughout the molecules. The hydrophobic residues at the N-terminal helices are highly conserved in homologous proteins. In addition, Cys residues involved in coordination

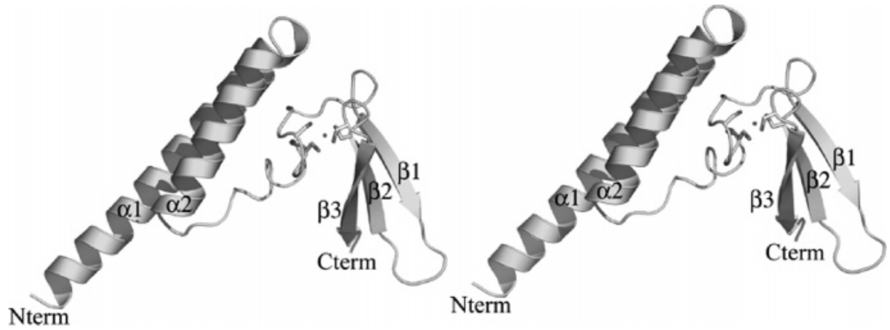

FIGURE 2: Stereoscopic drawing of Ph1601p from *P. horikoshii*. The molecular comprises an N-terminal domain (residues 1−55), a central linker domain (residues 56−79), and a C-terminal domain (residues 80−120), forming an L-shaped structure.

```
                    α1                           α2
              10        20        30            40        50
PH      MVDIVKRRDWEKKEKKKIAIERIDTLFTLAERVARYSPD---LAKRYVELALEIQKKAKV
PAB     ----MKKVSWEKREKKKVAIERIDTLFTLAEKVVKYSPD---LARRYVELALEIQKKSKV
RPR2    -MCKVESPRSGVLLLRRIALERAERLLRLARTVYYEDPD---RARRYVELARRIAMKARV
MTH     -----MRRGKRPRWMLKIAEERIDILFRMADREFSANPH---RSHRYTELARNIAMKYRV
RPP21   ---------MAGPVKDREAFQRLSFLCQAAHCVLSQNPENQALARFYCHTEKTIAKRLVL
                          *  *   *   *       *        *      *     :

                                     β1         β2       β3
          60        70        80         90        100       110
PH      KIPRKWKRRYCKRCHTFLIPGVNARVRLRTKR-MPHVVITCLECGYIMRYPYLREVKQ--
PAB     KLPRKWKRRYCKRCHAFLVPGFNARVRLRTDR-MPHVVITCLECGHIMRYPYLREVKE--
RPR2    KLPKHLKRSFCKRCNTPLIPGVTARVRLRQNR-MPHVSVTCLECGYIYRYPYLREVKE--
MTH     RIPREWRRRFCRKCYSFLKPGANCTVRIADG----KVNFRCHECGHIMRFPYIREKKD--
RPP21   RQDPSVKRTLCRSCSSLLIPGLTCTQRQRRRKGQRWTVQTCLTCQRSQRFLNDPKHLLWG
              :*  *: * : * **     *             *  *    *:

                      120
PH      -----------------KRKKAT---------------------------
PAB     -----------------KRKRKKD--------------------------
RPR2    -----------------RRRRHMEGVKDRDAG------------------
MTH     -----------------RRRNKIESHTTKEGTDEQITVGAHNKCGESQSDR
RPP21   DRPE---------AQLENQADINPSEPLPNIADLPKENIQTQALNTSE---
                          :
```

FIGURE 3: Sequence comparison of the Ph1601p protein from *P. horikoshii* with homologous proteins from other organisms. PH, PAB, RPR2, MTH, and RPP21 indicate amino acid sequences of Ph1601p from *P. horikoshii*, PAB0385 from *Pyrococcus abyssi*, Rpr2p from *S. cerevisiae*, Mth1618p from *M. thermoautorophicus*, and Rpp21 from human, respectively. Secondary structure elements determined in this work using PROCHECK (*31*) are also given. The numbering for residues refers to Ph1601p from *P. horikoshii*. The residues marked by asterisks represent complete conservation and those by colons conservative mutation.

of zinc ion are completely conserved in the protein family. It is thus assumed that the Ph1601p structure appears to be common with the Ph1601p protein family, and its entire molecule may play an essential role in the structure and/or function of RNase P.

To identify a surface feature of Ph1601p that may be important for the interaction with other molecules, we analyzed the molecular surface of the protein in terms of the distribution of charged amino acids. Basic amino acids are clustered around the N-terminal region at α1 (Lys6, Arg7, Arg8, Lys12, Lys13, Lys15, Lys16, Lys17, and Arg22), the central loop (Lys58, Arg61, Lys62, Lys64, Arg65, Arg66, Lys69, and Arg70), and the zinc ribbon domain at β-strand β1 (Arg82, Arg84, Arg86, Lys88, and Arg89). Because of this characteristic distribution, the electrostatic potential surface representation revealed that positively charged residues are concentrated on one face along the L arms, as depicted in Figure 4.

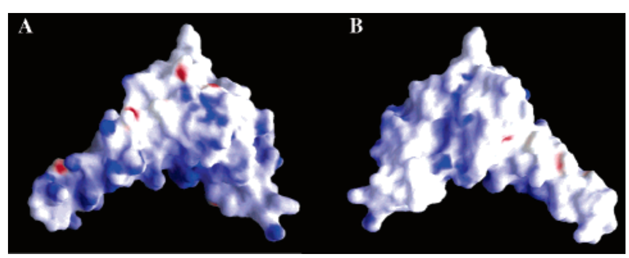

FIGURE 4: Molecular surface rendering of the Ph1601p molecule. (A) Surfaces are colored according to their electrostatic potentials as calculated by GRASP (*42*). The surface potential is displayed as a color gradient from red (negative) to blue (positive), showing the relative strong electropositive character of the putative RNA-binding site. (B) View after a 180° rotation from panel A.

Structural Similarity. The three-dimensional structures of three eubacterial RNase P proteins from *B. subtilis* (*9*), *Staphylococcus aureus* (*10*), and *Thermotoga maritima* (*11*) have been determined, and their structure similarity was

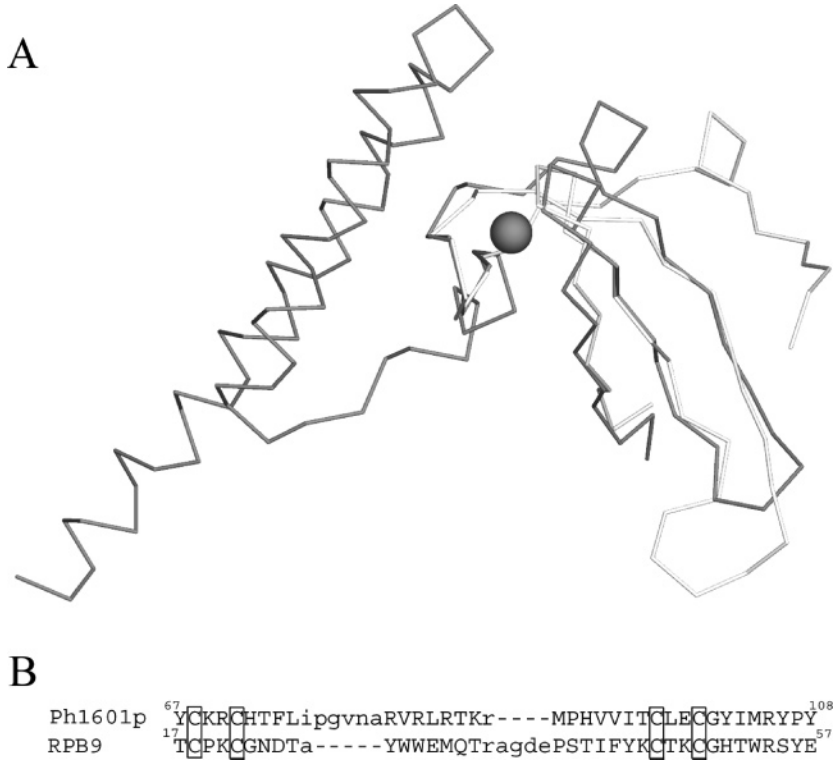

FIGURE 5: Comparison of the zinc ribbon domain of Ph1601p with that of Rpb9 from *T. celeri*. (A) Superimposition of the backbone atoms of Ph1601p (thick line) and RPB9 (thin line). (B) Structure-based sequence alignment of Ph1601p from *P. horikoshii* with the sequence of the Rpb9 from *T. celeri*. Identical residues are boxed. Structurally identical residues are shown in capital letters. Four Cys residues involved in coordination of zinc ion are boxed.

reported. However, neither marked sequence homology nor structure similarity was detected between eubacterial proteins and Ph1601p. To search for known protein structures that are homologous to Ph1601p, the coordinates were submitted to the VAST server. Although the overall Ph1601p fold represents a novel protein architecture, the C-terminal domain (residues 48−120) is structurally similar to proteins having the zinc ribbon domain found in a growing class of proteins that interact with nucleic acids, such as RNA polymerase subunits and transcriptional factors. Among them, the Ph1601 zinc ribbon domain shows the highest degree of structural similarity to archaea *Thermococcus* celer RNA polymerase II subunit 9 (Rpb9), which is involved in regulation of start-site selection and elongational arrest (*33*). The main chain for residues 64−106 in Ph1601p can be perfectly superimposed with that for the corresponding residues 17−57 in the Rpb9 with a root-mean-square (rms) deviation of 1.1 Å^2 (Figure 5A). The only positional differences are observed around the central loop and the $\beta1-\beta2$ loop in Ph1601p, where Ph1601p has a five-amino acid insertion and a four-amino acid deletion, respectively, as compared with the corresponding loop in Rpb9 and has a difference in Cα positions of up to 3 Å (residue 75 Cα). Despite the significant structural similarity, when the sequence of the zinc ribbon domain of Ph1601p is aligned according to its structure with the sequence of the zinc ribbon domain in Rpb9, it shares only a few residues, including Cys residues with the corresponding domain, and an insertion and a deletion are included for the alignment (Figure 5B). Similar observations that the structural resemblance is not reflected in an obvious homology at the sequence level have been noted in other zinc ribbon domains. It is thus likely that Ph1601p and Rpb9 have not diverged from a common ancestral protein; rather, a convergent evolution has resulted in the same structural solution for providing a platform for interaction with either proteins or nucleic acids.

Functional Residues. As shown in the sequence alignment, the amino acid residues conserved in the Ph1601p protein family are distributed throughout the molecule. To consider roles of these amino acids in the RNase P activity, we made use of Ala-scanning site-directed mutagenesis. The residues that were replaced were Arg22, Tyr44, Arg65, Lys69, Arg84, Arg86, and Arg105. Figure 6A shows the main chain folding with side chains of the amino acid residues analyzed in this study. In addition, Cys mutants C68S/C71S and C97S/C100S, in which Cys68 and Cys71 or Cys97 and Cys100 were replaced with Ser, were prepared to evaluate the importance of the zinc ion in the RNase P activity.

All mutants were expressed in *E. coli* BL21(DE3) RIL cells using the expression vector pET-22b and were purified by ion-exchange chromatography on SP-Sepharose, as described for the wild type (*23*). Enzymatic analyses of the reconstituted RNase P containing the mutant proteins and the relative values obtained in this analysis are given in panels B and C of Figure 6. The Cys mutations had a significant effect on the enzymatic activity; mutant proteins C68S/C71S and C97S/C100S exhibited little enzymatic activity (Figure 6B,C). The result demonstrated a strong involvement of the zinc ion in the RNase P activity. Additionally, mutations of Lys69 at the central loop and Arg86 and Arg105 at the zinc ribbon domain also had a significant effect on the enzymatic activity (Figure 6B). In particular, the mutation of Arg105 to Ala caused a significant reduction of the reconstituted RNase P activity as compared with that reconstituted by wild-type Ph1601p (Figure 6B). In contrast, mutations of Arg22 and Tyr44 at helices, Arg65

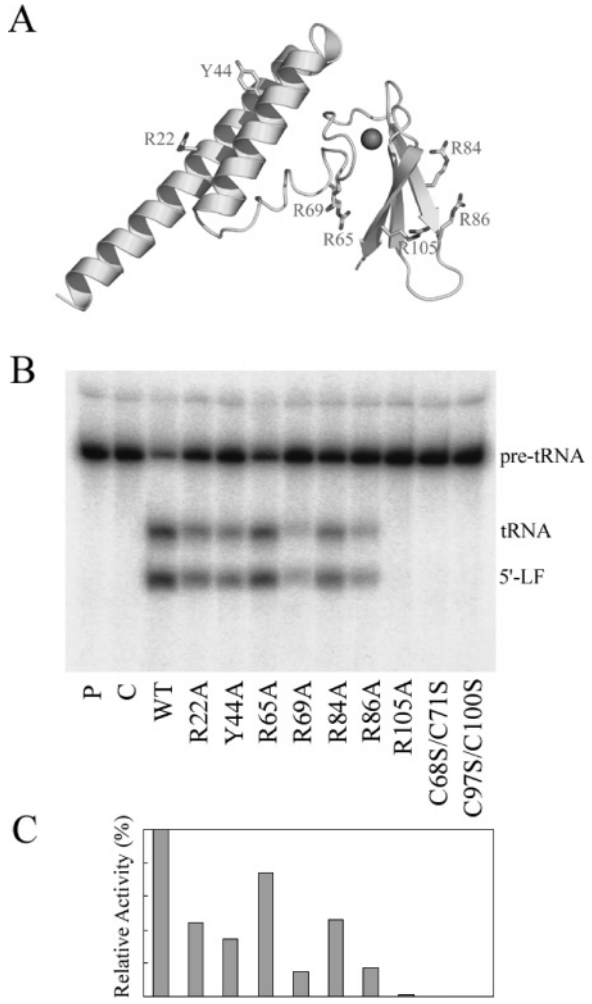

FIGURE 6: Amino acid residues involved in Ph1601p function. (A) Amino acid residues examined using site-directed mutagenesis. Amino acid residues conserved in homologous proteins were chosen for site-directed mutagenesis in this study. (B) Characterization of the enzymatic activity of the in vitro-reconstituted RNase P containing either wild-type Ph1601p or it mutant proteins. RNase P activities of the in vitro-reconstituted mixtures were assayed at 45 °C for 30 min using ^{32}P-labeled *P. horikoshii* pre-tRNA^{Tyr} as a substrate, as described in ref *23*. Lane P contained pre-tRNA^{Tyr}, and lane C contained the reaction product of the pre-tRNA^{Tyr} digested with the in vitro-reconstituted *P. horikoshii* RNase P without wild-type Ph1601p. The remaining lanes indicate the reaction products of the pre-tRNA^{Tyr} digested with the in vitro-reconstituted *P. horikoshii* RNase P containing wild-type Ph1601p and its mutant proteins. Pre-tRNA, tRNA, and 5′-LF at the right side indicate positions of pre-tRNA, tRNA, and the 5′ leader fragment. (C) The relative enzymatic activities by Ph1601p and its mutants were expressed as that of the reconstituted RNase P containing wild-type Ph1601p as 100. The experiments were carried out in at least triplicate, and the mean values are presented.

at the central loop, and Arg84 at the zinc ribbon domain had a modest effect on the enzymatic activity of RNase P, showing that the reconstituted particle containing the mutants exhibits ~40−75% of the enzymatic activity of wild-type Ph1601p.

To examine the structural integrity of mutants with the reduced activity, their CD spectra in the short wavelength region (200−250 nm) were compared with that of the wild type. This analysis showed that the CD spectra of mutants, in which Arg22, Tyr44, Lys69, Arg86, or Arg105 was replaced with Ala, were essentially the same as that of the

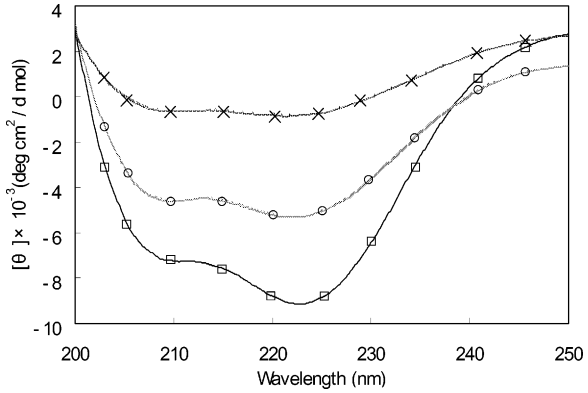

FIGURE 7: CD spectra of Ph1601p and its mutants in the far-ultraviolet region. The CD spectra of Ph1601p and its mutant proteins C68S/C71S and C97S/C100S were recorded on a Jasco J-720 spectropolarimeter at 25 °C. The protein concentration was 100 mg/mL in 20 mM Tris-HCl (pH 7.0) containing 200 mM NaCl. CD spectra of Ph1601p (□), C68S/C71S (○), and C97S/C100S (×).

wild type (data not shown); it is therefore indicated that replacements of four residues with Ala did not cause a structural change in the proteins. It is thus likely that these four amino acids play essential roles in the function of Ph1601p. In contrast, it was shown that mutations of Cys residues resulted in structural changes of proteins C68S/C71S and C97S/C100S (Figure 7). This finding demonstrates that the Cys residues may not be involved in a direct function but rather play an important role in structural stability of Ph1601p, possibly through coordination of zinc ion.

DISCUSSION

This study establishes the crystal structure of Ph1601p. A characteristic feature of the Ph1601p structure is that the C-terminal domain folds in the zinc ribbon domain. To date, the structures of several distinct zinc ribbon domains have been determined mainly by NMR (*33*−*38*). This fold consists of a rubredoxin knuckle containing the first CXXC motif followed by a β-strand of variable length connected to an antiparallel strand by a flexible loop. The second β-strand is followed by a second CXXC motif containing a rubredoxin knuckle, which connects to a third very short antiparallel β-strand. The zinc ribbon domain in Ph1601p consists of a three-stranded antiparallel β-sheet, with residues Cys68, Cys71, Cys97, and Cys100 forming a tetrahedral coordination site that binds one atom of zinc. Mutations that abolish zinc binding result in destabilization of the mutant structures, as judged from CD spectra (data not shown). This finding supports the proposal that the zinc ion plays a role in stabilization of the relative conformation of the N- and C-domains in Ph1601p.

Despite the structural similarity of the zinc ribbon domain, proteins containing this mini domain have very diverse amino acid sequence and function. Examples include the RNA polymerase II initiation factor TFIIB (*34*, *35*), elongation factors TFIIS (*36*), and eukaryotic translational initiation factor 2γ (*37*, *38*). The only conserved sequence features of this fold appear to be the CXXC (H) and CXXC motifs. It has been proposed that the great sequence diversity of zinc ribbon domains is due to the planar nature of this fold, along with zinc coordination, which eliminates the need for a hydrophobic core. Examining the amino acid residues in the

zinc ribbon domain in Ph1601p, we find they are variable as compared with the corresponding residues in other zinc ribbon domains (Figure 5B). The zinc ribbon domain of Ph1601p has a cluster of basic amino acids (Arg82, Arg84, Arg86, Lys88, Arg89, His92, Arg105, and Arg110) and displays a highly positive charged molecular surface electrostatic potential representation, as shown in Figure 4. These amino acids are replaced with other amino acids in other zinc ribbon domains, but conserved in the Ph1601p protein family (Figure 3). Therefore, these amino acids may play a specific function for the Ph1601p protein family. Indeed, the mutations of Arg86 and Arg105 abolished the RNase P activity. It is therefore suggested that these basic amino acids would be involved in a direct interaction with RNA or protein partners.

The recent study using the yeast two-hybrid system showed that Ph1601p interacts with both Ph1771p and Ph1481p: Ph1601p is thus suggested to act as a mediator for proteins Ph1481p and Ph1771p (*39*). A similar spatial organization was noted on the *Methanothermobacter thermoautotrophicus* (*40*) and yeast (*22*) RNase Ps, where Mth1618 and Rpr2, respectively, interact with Mth11 (*M. thermoautotrophicus* homologue of Ph1771p) and Pop4 (yeast homologue of Ph1771p). Likewise, Rpp21 (human homologue of Ph1601p) seems to interact with Rpp29 (human homologue of Ph1771p) as well as other proteins (*20*). Thus, it is assumed that Ph1601p plays roles not only in the recognition of substrates but also in the assembly of RNase P, by bridging protein partners Ph1481p and Ph1771p. In this regard, the structural feature wherein Ph1601p forms the L-shaped structure has led to speculation that both arms may be responsible for interaction with Ph1481p and Ph1771p in the *P. horikoshii* RNase P. In the crystal structure of Ph1771p, strand β7 located at the C-terminal region forms an antiparallel intersubunit β-sheet by interacting with strand β4 in a symmetry-related molecule (*24*). Furthermore, Sm proteins, exhibiting a structural resemblance to Ph1771p, oligomerize to form a homo- or hetero-oligomeric complex through β-strand pairing (*41*). These findings suggest that either β4 or β7 may be involved in interaction with Ph1601p through β-strand pairing. The C-terminal β-strands folded in the zinc ribbon domain of Ph1601p would therefore be a good candidate for the interaction site with Ph1771p. On the other hand, the N-terminal helices would be predicted to be involved in interaction with the protein Ph1481p. The recently determined structure of Ph1481p has extended α-helices at the C-terminal region (S. Kawano et al., unpublished data). It is tempting to speculate that the N-terminal helices in Ph1601p would interact with those of Ph1481p by forming an α-helix bundle structure.

It is clear that a detailed understanding of the mechanism involved requires a knowledge of the high-resolution structure of RNase P. However, progress on the structure determination of RNase P has been retarded mainly by the difficulty of obtaining significant quantities of pure RNase P needed for crystallographic studies. An alternative approach that is now on the advance is to obtain high-resolution structures of individual RNase P proteins by X-ray crystallography, and then obtain the structure of complexes of specific proteins with fragments of RNase P RNA. By this piecemeal approach, it will be possible to understand the structure of the RNase P in much greater detail than is possible now, especially when the high-resolution information is combined with a knowledge of the overall arrangement of the components of the RNase P. The present structural analysis of Ph1601p, together with our previous structural analyses of Ph1771p (*24*) and Ph1877p (*25*), provides structural information about three *P. horikoshii* RNase P proteins. In addition, we recently established the crystal structures of Ph1481p and Ph1496p which correspond to human hpop5 and Rpp38, respectively. These structures will be described in detail elsewhere (S. Kawano, H. Fukuhara, and M. Kimura, unpublished results). A completion of three-dimensional structures of the *P. horikoshii* RNase P protein subunits will provide a framework for understanding structure−function relationships of archaeal and eukaryotic RNase Ps.

ACKNOWLEDGMENT

The synchrotron radiation experiments were performed at beamlines BL38B1 and BL41XU in the SPring-8 with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (Proposal 2004B0873-NL1-np-P3k, 2004-B0874-NL1-np-P3k). We thank beamline staff (Dr. K. Hasegawa) at SPring-8 for kind help with data collection and M. Ohara (Fukuoka) for helpful comments on the manuscript.

REFERENCES

1. Frank, D. N., and Pace, N. R. (1998) Ribonuclase P: Unity and diversity in tRNA processing enzyme, *Annu. Rev. Biochem. 67*, 153−180.
2. Altman, S., Kirsebom, L., and Ribonuclase, P. (1999) in *The RNA world* (Gesteland, R. F., Cech, T., and Atkins, J. F., Eds.) pp 351−380, Cold Spring Harbor Laboratory Press, Plainview, NY.
3. 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.
4. Hsieh, J., Andrews, A. J., and Fierke, C. A. (2004) Roles of protein subunits in RNA-protein complexes: Lessons from ribonuclease P, *Biopolymers 73*, 79−89.
5. Harris, M. E., and Pace, N. R. (1995) Identification of phosphates involved in catalysis by the ribozyme RNase P RNA, *RNA 1*, 210−218.
6. Kazantsev, A. V., and Pace, N. R. (1998) Identification by modification-interference of purine N-7 and ribose 2′-OH groups critical for catalysis by bacterial ribonuclease P, *RNA 4*, 937−947.
7. Christian, E., Kaye, N. M., and Harris, M. E. (2000) Helix P4 is a divalent metal ion briding site in the conserved core of the ribonuclease P ribozyme, *RNA 6*, 511−519.
8. Crary, S. M., Curz, J. C., and Fierke, C. A. (2002) Specific phosphorothioate substitutions probe the active site of *Bacillus subtilis* ribobuclease P, *RNA 8*, 933−947.
9. Stams, T., Niranjanakumari, S., Fierke, C. A., and Christianson, D. W. (1998) Ribonuclease P Protein Structure: Evolutionary Origins in the Translational Apparatus, *Science 280*, 752−755.
10. Spitzfaden, C., Nicholson, N., Jones, J. J., Guth, S., Lehr, R., Prescott, C. D., Hegg, L. A., and Eggleston, D. S. (2000) The structure of ribonuclease P protein from *Staphylococcus aureus* reveals a unique binding site for single-stranded RNA, *J. Mol. Biol. 295*, 105−115.
11. Kazantsev, A. V., Krivenko, A. A., Harrington, D. J., Carter, R. J., Holbrook, S. R., Adams, P. D., and Pace, N. R. (2003) High-resolution structure of RNase P protein from *Thermotoga maritima, Proc. Natl. Acad. Sci. U.S.A. 100*, 7497−7502.
12. Tsai, H.-T., Masquida, B., Biswas, R., Westhof, E., and Gopalan, V. (2003) Molecular modeling of the three-dimensional structure of the bacterial RNase P holoenzyme, *J. Mol. Biol. 325*, 661−675.

13. Krasilnikov, A. S., Yang, T., Pan, T., and Mondragon, A. (2003) Crystal structure of the specificity domain of ribonuclease P, *Nature 421*, 760−764.
14. Krasilnikov, A. S., Xiao, Y., Pan, T., and Mondragon, A. (2004) Basis for structural diversity in homologous RNAs, *Science 306*, 104−107.
15. Kurz, J. C., and Fierke, C. A. (2002) The affinity of magnesium binding sites in the *Bacillus subtilis* RNase P Pre-tRNA complex is enhanced by the protein subunit, *Biochemistry 41*, 9545−9558.
16. Bartkiewicz, M., Gold, H., and Altman, S. (1989) Identification and characterization of an RNA molecule that copurifies with RNase P activity from HeLa cells, *Genes Dev. 3*, 488−499.
17. Yuan, Y., and Altman, S. (1995) Substrate recognition by human RNase P: Identification of small, model substrates for the enzyme, *EMBO J. 14*, 159−168.
18. Xiao, S., Houser-Scott, F., and Engelke, D. R. (2001) Eukaryotic ribonuclease P: Increased complexity to cope with the nuclear pre-tRNA pathway, *J. Cell. Physiol. 187*, 11−20.
19. Mann, H., Ben-Asouli, Y., Schein, A., Moussa, S., and Jarrous, N. (2003) Eukaryotic RNase P: Role of RNA and protein subunits of a primordial catalytic ribonucleo-protein in RNA-based catalysis, *Mol. Cell 12*, 925−935.
20. Jiang, T., and Altman, S. (2001) Protein−protein interactions with subunits of human nuclear RNase P, *Proc. Natl. Acad. Sci. U.S.A. 98*, 920−925.
21. Jiang, T., Guerrier-Takada, C., and Altman, S. (2001) Protein-RNA interactions in the subunits of human nuclear RNase P, *RNA 7*, 937−941.
22. Houser-Scott, F., Xiao, S., Millikin, C., Zengel, J. M., Lindahl, L., and Engelke, D. R. (2002) Interactions among the protein and RNA subunits of *Saccharomyces cerevisiae* nuclear RNase P, *Proc. Natl. Acad. Sci. U.S.A. 99*, 2684−2689.
23. Kouzuma, Y., Mizoguchi, M., Takagi, H., Fukuhara, H., Tsukamoto, M., Numata, T., and Kimura, M. (2003) Reconstitution of archaeal ribonuclease P from RNA and four protein components, *Biochem. Biophys. Res. Commun. 306*, 666−673.
24. Numata, T., Ishimatsu, I., Kakuta, Y., Tanaka, I., and Kimura, M. (2004) Crystal structure of archaeal ribonuclease P protein Ph1771p from *Pyrococcus horikoshii* OT3: An archaeal homolog of eukaryotic ribonuclease P protein Rpp29, *RNA 10*, 1423−1432.
25. Takagi, H., Watanabe, M., Kakuta, Y., Kamachi, R., Numata, T., Tanaka, I., and Kimura, M. (2004) Crystal structure of the ribonuclease P protein Ph1877p from hyperthermophilic archaeon *Pyrococcus horikoshii* OT3, *Biochem. Biophys. Res. Commun. 319*, 787−794.
26. Jarrous, N., Reiner, R., Wesolowski, D., Mann, H., Guerrier-Takada, C., and Altman, S. (2001) Function and subnuclear distribution of Rpp21, a protein subunit of the human ribonucleoprotein ribonuclease P, *RNA 7*, 1153−1164.
27. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
28. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
29. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.
30. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement, *Nat. Struct. Biol. 6*, 458−463.
31. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: A program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
32. Qian, X., Gozani, S. N., Yoon, H., Jeon, C. J., Agarwal, K., and Weiss, M. A. (1993) Novel zinc motif in the basal transcriptional machinery: Three-dimensional NMR studies of the nucleic acid binding domain of transcriptional elongation factor TFIIS, *Biochemistry 32*, 9944−9959.
33. Wang, B., Jones, D. N., Kaine, B. P., and Weiss, M. A. (1998) High-resolution structural motif in eukaryotic RNA polymerases, *Structure 6*, 555−569.
34. Hahn, S., and Roberts, S. (2000) The zinc ribbon domains of the general transcription factors TFIIB and Brf: Conserved functional surfaces but different roles in transcription initiation, *Genes Dev. 14*, 719−730.
35. Zhu, W., Zeng, Q., Colangelo, C. M., Lewis, M., Summers, M. F., and Scott, R. A. (1996) The N-terminal domain of TFIIB from *Pyrococcus furiosus* forms a zinc ribbon, *Nat. Struct. Biol. 3*, 122−124.
36. Olmsted, V. K., Awrey, D. E., Koth, C., Shan, X., Morin, P. E., Kazanis, S., Edwards, A. M., and Arrowsmith, C. H. (1998) Yeast transcript elongation factor (TFIIS), structure and function. I: NMR structural analysis of the minimal transcriptionally active region, *J. Biol. Chem. 273*, 22589−22594.
37. Roll-Mecak, A., Alone, P., Cao, C., Dever, T. E., and Burley, S. K. (2004) X-ray structure of translation initiation factor eIF2g. Implications for tRNA and eIF2a binding, *J. Biol. Chem. 279*, 10639−10642.
38. Schmitt, E., Blanquet, S., and Mechulam, Y. (2002) The large subunit of initiation factor aIF2 is a close structural homologue of elongation factors, *EMBO J. 21*, 1821−1832.
39. Kifusa, M., Fukuhara, H., Hayashi, T., and Kimura, M. (2005) Protein−protein interactions in the subunits of ribonuclease P in hyperthermophilic archaeon *Pyrococcus horikoshii* OT3, *Biosci., Biotechnol., Biochem. 69*, 1209−1212.
40. Hall, T. A., and Brown, J. W. (2004) Interactions between RNase P protein subunit in archaea, *Archaea 1*, 247−253.
41. Kambach, C., Walke, S., Young, R., Avis, J. M., de la Fortelle, E., Raker, V. A., Luhrmann, R., Li, J., and Nagai, K. (1999) Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs, *Cell 96*, 375−387.
42. Nicholls, A., Sharp, K., and Honing, B. (1991) Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins 11*, 281−296.

BI050738Z