Structure of *Pfu* Pop5, an archaeal RNase P protein

Ross C. Wilson**, Christopher J. Bohlen†, Mark P. Foster**, and Charles E. Bell*†

*Ohio State Biochemistry Program, Departments of †Biochemistry and §Molecular and Cellular Biochemistry, Ohio State University, Columbus, OH 43210

Edited by Norman R. Pace, University of Colorado, Boulder, CO, and approved November 30, 2005 (received for review September 15, 2005)

We have used NMR spectroscopy and x-ray crystallography to determine the three-dimensional structure of PF1378 (*Pfu* Pop5), one of four protein subunits of archaeal RNase P that shares a homolog in the eukaryotic enzyme. RNase P is an essential and ubiquitous ribonucleoprotein enzyme required for maturation of tRNA. In bacteria, the enzyme's RNA subunit is responsible for cleaving the single-stranded 5′ leader sequence of precursor tRNA molecules (pre-tRNA), whereas the protein subunit assists in substrate binding. Although in bacteria the RNase P holoenzyme consists of one large catalytic RNA and one small protein subunit, in archaea and eukarya the enzyme contains several (≥4) protein subunits, each of which lacks sequence similarity to the bacterial protein. The functional role of the proteins is poorly understood, as is the increased complexity in comparison to the bacterial enzyme.*Pfu* Pop5 has been directly implicated in catalysis by the observation that it pairs with PF1914 (*Pfu* Rpp30) to functionally reconstitute the catalytic domain of the RNA subunit. The protein adopts an α-β sandwich fold highly homologous to the single-stranded RNA binding RRM domain. Furthermore, the three-dimensional arrangement of *Pfu* Pop5's structural elements is remarkably similar to that of the bacterial protein subunit. NMR spectra have been used to map the interaction of Pop5 with *Pfu* Rpp30. The data presented permit tantalizing hypotheses regarding the role of this protein subunit shared by archaeal and eukaryotic RNase P.

R Nase P is a ubiquitous and essential enzyme in all domains of life responsible for cleaving the single-stranded 5′ leader sequence of precursor tRNA (pre-tRNA), a vital reaction in the maturation of tRNA (1, 2). RNase P is a ribonucleoprotein complex that, in bacteria, consists of a large RNA (P RNA, ~400 nt) responsible for mediating catalysis (3) and a small protein (~120 residues) that is required in vivo (4) and appears to function by stabilizing the RNA subunit and/or increasing the enzyme’s affinity for the substrate by binding to its typically single-stranded 5′ leader sequence (5–8). The bacterial RNA subunit is capable of cleaving its pre-tRNA substrate in vitro in the absence of its protein moiety at high concentrations of Mg2++; thus, RNase P is a ribozyme. The recent crystal structures of the bacterial RNA subunits from *Thermotoga maritima* (9) and *Bacillus stearothermophilus* (10) have provided a rich framework for understanding the function of this enzyme, while validating predictions based on biochemical and phylogenetic data (11, 12).

In eukaryotes RNase P is more complex, consisting of a smaller RNA and at least nine protein subunits (13). Although sequence and predicted secondary structural elements are conserved between the RNA subunits of eukaryal and bacterial RNase P (14), the eukaryal RNA has not been shown to be catalytically active in vitro. Despite similarity between the RNA components in Archaea and Eukarya, genome sequence analysis has failed to identify any eukaryotic homolog of the bacterial protein. Furthermore, the functional roles of the multiple protein subunits in eukaryotic RNase P remain to be elucidated.

Evidence suggests that the RNase P enzyme from archaea represents a mosaic of features of the eukaryal and bacterial enzymes. The archaeal enzyme contains an RNA that is more similar to the bacterial than to the eukaryal RNA (15), and RNA subunits from some archaea have been shown to be catalytically active in vitro under high salt conditions (16). On the other hand, archaeal RNase P is similar to the eukaryal enzyme in that it is comprised of multiple protein subunits (four; perhaps more), each of which has marked sequence similarity to protein subunits of eukaryal RNase P (17): Rpp21 (18), Rpp29 (19), Rpp30 (20), and Pop5 (21) [also known as Rrp2, Pop4 (22), Rpp1 (23), and hPop5, respectively]. The four archaeal proteins can be recombinitely expressed and combined in vitro with the corresponding in vitro-transcribed RNA subunit to yield an active enzyme, as demonstrated for the enzymes from *Methanothermobacter thermautotrophicus* (24), *Pyrococcus furiosus* (H.-Y. Tsai and V. Gopalak, personal communication), and *Pyrococcus horikoshii* (25). Although little is known about the arrangement of the proteins in the archaeal holoenzyme, yeast two-hybrid experiments using the protein subunits from *Methanothermobacter thermautotrophicus* and *Pyrococcus horikoshii* have detected strong interactions between two pairs of proteins, Pop5+Rpp30 and Rpp21+Rpp29 (26, 27), mirroring similar findings with the eukaryotic proteins (28, 29).

In vitro reconstitution experiments using *Pyrococcus furiosus* components (H.-Y. Tsai and V. Gopalak, personal communication) provide more evidence that the proteins function in pairs, because the Pop5+Rpp30 and Rpp21+Rpp29 subsets of proteins are each capable of forming an active enzyme in the presence of the intact RNA subunit. Furthermore, Pop5+Rpp30 can be functionally reconstituted with the catalytic domain of the RNA subunit. Detailed structural data on the RNase P components will greatly increase our understanding of how this ribonucleoprotein complex assembles and performs catalysis.

The three-dimensional structures of three archaeal RNase P proteins have been determined, each of which adopts a different fold: Rpp29 contains a beta-barrel fold (24, 30–32) reminiscent of archaeal Sm proteins (32); Rpp30 is a TIM barrel (33); Rpp21 contains a zinc ribbon (34). Here we report the crystal structure of *Pyrococcus furiosus* Pop5 (**Pfu** Pop5), a 120-residue protein conserved both within and between Archaea and Eukarya (Fig. 1), and present as a paralog in multicellular eukaryotes as Rpp14 (35). NMR chemical shift perturbations have allowed us to identify the site of interaction with Rpp30. **Pfu** Pop5 is unique among the archaeal RNase P proteins in that its α–β sandwich bears structural similarity to the bacterial RNase P protein. With the observation that the **Pfu** Pop5 structure is homologous to the ubiquitous single-stranded nucleic acid binding RNA recognition motif (RRM) domain (reviewed in ref. 36), these data provide clues regarding the evolution and function of RNase P.

Results

Characterization. The **Pfu** Pop5 construct used consists of 120 wild-type residues (Fig. 1) with a single Cys to Ser point mutation at residue 72 that improved solution behavior as assayed by NMR.

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: RRM, RNA recognition motif; **Pfu**, *Pyrococcus furiosus*.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID 2AV5). The NMR chemical shifts have been deposited in the BioMagResBank, www.bmrb.wisc.edu (accession no. 6917).

*To whom correspondence should be addressed. E-mail: foster.281@osu.edu.*

© 2006 by The National Academy of Sciences of the USA
Caenorhabditis elegans (coccus jannaschii code NP 126368), Methanothermobacter thermoautotrophicus (Mth, NCBI code NP 275830), Archaeoglobus fulgidus (Afu, NCBI code NP 069325), Methanococcus jannaschii (Mja, NCBI code NP 247470), Saccharomyces cerevisiae (Sce, NCBI code NP 009369), Arabidopsis thaliana (Ath, NCBI code NP 683274), Caenorhabditis elegans (Cel, NCBI code NP 499477), Drosophila melanogaster (Dme, NCBI code NP 648955), Mus musculus (Mmu, NCBI code NP 080674), and Homo sapiens (Has, NCBI code NP 057002). Alignment was generated with CLUSTALW (64) and colored according to similarity using a Risler scoring matrix (65); shaded residues indicate identity, and boxed residues indicate a global similarity score >0.7. Similarity scores within archaea were calculated separately. Secondary structure elements observed in the crystal structure as assigned by DSSP (66) and elements predicted by chemical shift index (CSI) analysis (37) are indicated. Arrows indicate sites of trypsin cleavage as identified by digestion followed by ESI mass spectrometry (expected mass, 10,611 Da; observed, 10,610 Da). The figure was generated by using ESPRIT (67).

Fig. 1. Sequence conservation and secondary structure of Pfu Pop5. Sequence alignment of Pfu Pop5 with sequences from Archaea (black) and Eukarya (gray). Sequences are from Pyrococcus furiosis (Pfu, NCBI code accession no. NP 579107), Pyrococcus horikoshii (Pho, NCBI code NP 143344), Pyrococcus abyssi (Pab, NCBI code NP 126368), Methanothermobacter thermoautotrophicus (Mth, NCBI code NP 275830), Archaeoglobus fulgidus (Afu, NCBI code NP 069325), Methanococcus jannaschii (Mja, NCBI code NP 247470), Saccharomyces cerevisiae (Sce, NCBI code NP 009369), Arabidopsis thaliana (Ath, NCBI code NP 683274), Caenorhabditis elegans (Cel, NCBI code NP 499477), Drosophila melanogaster (Dme, NCBI code NP 648955), Mus musculus (Mmu, NCBI code NP 080674), and Homo sapiens (Has, NCBI code NP 057002). Alignment was generated with CLUSTALW (64) and colored according to similarity using a Risler scoring matrix (65); shaded residues indicate identity, and boxed residues indicate a global similarity score >0.7. Similarity scores within archaea were calculated separately. Secondary structure elements observed in the crystal structure as assigned by DSSP (66) and elements predicted by chemical shift index (CSI) analysis (37) are indicated. Arrows indicate sites of trypsin cleavage as identified by digestion followed by ESI mass spectrometry (expected mass, 10,611 Da; observed, 10,610 Da). The figure was generated by using ESPRIT (67).

and in vitro reconstitution experiments (H.-Y. Tsai, unpublished results). The N-terminal Met is removed during expression in E. coli as confirmed by ESI mass spectrometry (expected mass, 13,708 Da; observed, 13,707 Da). Dynamic light scattering data indicate that even the mutant protein tends to aggregate to approximately an octamer at 1 mM concentration, but behaves as a monomer at 300 μM concentration. Limited trypsin proteolysis experiments indicate that 12 and 13 residues at the N and C termini, respectively, are removed within 30 min, whereas the remaining core was not sensitive to digestion. Mass spectrometric analysis of the trypsin fragments revealed a mass of 10,610 Da, consistent with the 10,611 Da expected for residues 14–107, as indicated in Fig. 1.

NMR Spectroscopy and Resonance Assignments. Two-dimensional 1H-15N-correlated NMR spectra recorded at 22°C exhibited poor signal-to-noise and dispersion, indicating that the protein is poorly folded or aggregated at this temperature. Despite good dispersion and spectral quality in data collected at 55°C (see supporting information, which is published on the PNAS web site), relatively broad lines and dynamic light scattering experiments at elevated concentrations suggested that under those conditions Pfu Pop5 was aggregating, with an average mass of 117 kDa, not the expected 13.7 kDa monomer. Spectra of wild-type and C72S constructs were nearly identical, confirming that the overall fold of the mutant was unperturbed (data not shown). Backbone resonance assignments were obtained for 90% of the nonproline residues using standard triple resonance techniques. Only five residues could be assigned for the N-terminal 16 residues; heteronuclear 15N-{1H} NOE data suggested that the N terminus was dynamically disordered on time scales faster than overall tumbling, whereas the rest of the protein appeared to be well ordered, with an average NOE value of 0.81 (see supporting information); NOE values dipped slightly in the loops and the C-terminal helix. Analysis of C, N, H3, Cα, Cβ, and Hα chemical shifts (37, 38) predicted a βαβαβα secondary structure pattern for residues 18–115 (Fig. 1).

Crystal Structure. Crystals of Pfu Pop5 belong to the P41212 space group and have five monomers in the asymmetric unit. The crystal structure (Fig. 2) was determined by single isomorphous replacement from data collected on a crystal soaked with ethylmercuric phosphate, combined with solvent flattening and fivefold noncrystallographic symmetry averaging. The secondary structure pattern obtained from NMR analysis was used as a guide in tracing the backbone; there was excellent agreement between this predicted pattern and the observed fold, and a model was readily built into the electron density map by using O (39) (Fig. 2e; see supporting information). For each of the five molecules, interpretable density could be observed for residues 15–120. The absence of density at the N terminus is apparently due to local disorder, which is corroborated by the heteronuclear NOE and limited trypsin proteolysis data. The Pfu Pop5 structure was refined to 3.15 Å with Rfree and Rwork values of 22.90% and 25.66%, respectively.

Pfu Pop5 adopts an α−β sandwich fold comprised of a central, four-stranded antiparallel β-sheet surrounded by four α-helices (Fig. 2b). One face of the β-sheet packs with helices α1, α2, and α3 to form the hydrophobic core. The other face is more exposed to solvent, but features three exposed apolar side chains that loosely pack with six apolar side chains from helix α4 (Fig. 2c). At the end of helix α3, residues 114–116 adopt a 3-10 helix (η1) conformation followed by a short loop, allowing the side chains of Phe-117 and Trp-119 to be partially packed against the apolar side chains of the β-sheet (Fig. 2c). β-bulges perturb strands β3 and β4 from ideal β-strand geometry at residues 58–59 and
observed near helix all monomers in the asymmetric unit, another dimer interface is dihedral angles unfavored for non-glycine residues (according to secondary structure as assigned by DSSP (66). (H.-Y. Tsai and V. Gopalan, personal communication). Large chemical shift perturbations for specific residues in Pfu Pop5 (see Fig. 4a) and complementary perturbations in the spectra of Pfu Rpp30 (unpublished data) are indicative of a high-affinity specific interaction between the two proteins. Although the details of the intermolecular interface remain to be elucidated, the most strongly perturbed amide resonances belonged to residues located principally in helix $a3$, strand $b3$, and the end of strand $b1$ (Fig. 4b).

### Discussion

**Pop5 Resembles the Bacterial RNase P Protein Subunit.** Remarkably, the three dimensional arrangement of the structural elements that compose the $\alpha$-$\beta$ sandwich fold of Pfu Pop5 is similar to that of the bacterial RNase P proteins from Bacillus subtilis, Staphylococcus aureus, and Thermotoga maritima (40–42) (Fig. 3a). The similarity is particularly evident upon aligning helix $a1$ of Pop5 with helix $a2$ of the bacterial protein, which contains the conserved “RNR motif” and has been shown by tethered Fe-EDTA footprinting to be in close proximity to the bacterial catalytic RNA subunit (12). With this alignment, it can be seen that the principal difference in the three-dimensional organization of their structural elements is in the orientation of helix $a4$ of Pop5 and its corresponding helix $a1$ in the bacterial proteins; notably, this helix is poorly packed against the central sheet suggesting it could be readily rearranged. The high degree of structural similarity is particularly remarkable considering the yeast two-hybrid (26, 27) and partial reconstitution experiments (H.-Y. Tsai and V. Gopalan, personal communication). Large chemical shift perturbations for specific residues in Pfu Pop5 (see Fig. 4a) and complementary perturbations in the spectra of Pfu Rpp30 (unpublished data) are indicative of a high-affinity specific interaction between the two proteins. Although the details of the intermolecular interface remain to be elucidated, the most strongly perturbed amide resonances belonged to residues located principally in helix $a3$, strand $b3$, and the end of strand $b1$ (Fig. 4b).

**Interaction with Pfu Rpp30.** Two-dimensional $^1$H-$^15$N correlated NMR spectra of Pop5 in complex with Rpp30 were recorded to examine the interaction between these two proteins implied by

![Image](c)
proteins’ differing primary sequence and secondary structural topology, \( \alpha\beta\beta\alpha\beta \) in Pop5 versus \( \alpha\beta\beta\alpha\beta \) in the bacterial RNase P proteins, suggesting that the two proteins have different evolutionary origins.

**Pfu Pop5 Adopts the RRM Fold.** Comparative fold analysis using DALI (43) revealed that Pfu Pop5 is a structural homolog of one of the most abundant protein domains in eukarya, the RRM, also known as the ribonucleoprotein domain (RRN) or the RNA-binding domain. RRM domain-containing proteins are prevalent in all posttranscriptional events, including RNA processing, splicing, and editing (reviewed in ref. 36). DALI identified the N-terminal RRM domain of U1A (44) as a close structural homolog of Pfu Pop5 (Fig. 3b), with an rms deviation of 2.7 Å over 76 equivalent Ca atoms (Z score 5.9). In addition to a central \( \beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4\) fold, the RRM domain is identified by conserved residues in the central strands \( \beta_1 \) and \( \beta_3 \), termed RNP1 and RNP2, respectively (45–47). The RRM \( \alpha-\beta \) sandwich fold is often accompanied by N- and C-terminal extensions that are disordered in the free protein, yet capable of aiding nucleic acid binding (36). Residues 17–103 of Pfu Pop5 embody a classic RRM domain, with the exception of the insertion of the short helix \( \alpha_2 \) and the presence of longer \( \beta \)-strands than those of a typical RRM. Despite tertiary structure similarity, sequence analysis using BLAST (48) did not identify Pfu Pop5 as a putative RRM; indeed, when canonical residues in strands \( \beta_1 \) and \( \beta_3 \) are compared to the corresponding RNP1 and RNP2 consensus sequences, there is little similarity (see supporting information). Other RRM domains, such as the one in ALY (49) and the C-terminal RRM domain of La (50), show considerable divergence from the consensus sequences while maintaining typical RRM architecture.

**A Putative Functional Role for Pop5.** Based on its membership to the RRM family of single-stranded RNA-binding domain and structural similarity to bacterial RNase P protein subunit, we propose that Pop5 may function in precursor tRNA substrate binding. Single-stranded RNA binding by the RRM motif and the bacterial RNase P protein is accomplished by using aromatic and apolar side chains exposed from the central \( \beta \)-sheet (see supporting information) (6, 36, 41). Indeed, the exposed face of the \( \beta \)-sheet of Pfu Pop5 does form a tract populated with aromatic (Tyr-18, Trp-57) and hydrophobic residues (Ile-59, Ile-69) that compose a comparable RNA–RNA-binding surface (Fig. 3d). In the crystal, however, this putative RNA-binding surface is partially obscured by the C-terminal helix \( \alpha_2 \) (Fig. 2b), which would have to rearrange to allow RNA binding. Such conformational changes in an RRM domain are not without precedent, as the C-terminal residues of U1A RRM1 reorient in an induced fit manner upon binding to its RNA substrate (36, 51). Although the heteronuclear NOE values for residues in the C-terminal helix (see supporting information) indicate the helix is not highly flexible, propensity for conformational change by this helix is suggested by the observation that it, like the flexible N terminus, is readily severed at Lys-107 by limited trypsin proteolysis. A conserved glycin preceding this helix (Gly-104, Fig. 1) could serve as a hinge, allowing the helix to move, exposing the hydrophobic residues on the face of the \( \beta \)-sheet. Given this precedent, it is not unreasonable to imagine that, upon holoenzyme assembly, the C-terminal helix of Pfu Pop5 could adopt a position analogous to that of helix \( \alpha_2 \) of the bacterial RNase P protein (Fig. 3d), where the pre-tRNA substrate binds to functionally important hydrophobic residues in the central \( \beta \) sheet (5). Moreover, such a reorientation would position three aromatic side chains (Tyr-18, Phe-113, and Phe-117) similarly to those known to be involved in RNA binding in the *Bsu* RNase P protein (Tyr-34, Phe-20, and Phe-16, respectively) (Fig. 3d) (6, 7, 40).

Of the known archaeal RNase P proteins, only Pop5 bears recognizable similarity to the highly conserved bacterial protein subunit, and yet the bacterial protein structure is highly conserved, even between organisms with highly divergent RNase P enzymes (40–42). Furthermore, the bacterial protein has been shown capable of hetero-reconstituting the activity of archaeal RNase P RNA (16). The highly conserved three-dimensional structures of the bacterial protein, despite variations in the RNA subunits, suggest that the conserved shape is dictated by its functional role of binding both the catalytic RNA and the substrate precursor-tRNA. Consequently, the global similarity between Pop5 and the bacterial protein allows us to postulate that, in archaeal and eukaryotic RNase P, Pop5 is fulfilling a role similar to that of its bacterial counterpart.

The structural similarity of Pfu Pop5 to the bacterial protein subunit suggests it functions in an analogous, if imperfect, manner. The divergent secondary structure topologies of the proteins, \( \alpha\beta\beta\alpha\beta \) in Pfu Pop5 versus \( \alpha\beta\beta\alpha\beta \) in the bacterial RNase P proteins, suggest that the archaeal and bacterial proteins have different evolutionary origins. Two plausible evolutionary scenarios are worth considering: in one, ancestral catalytic RNA subunits in bacteria and archaea could have independently recruited similarly shaped \( \alpha-\beta \) sandwich proteins; alternatively, a Pop5-like protein with sufficiently similar shape may have been able to usurp the role of the single bacterial protein, promoting RNase P catalysis and proliferating throughout the archaeal and eukaryal domains. Plausible functional roles for the other protein subunits can be proposed by hypothesizing that the evolutionary increase in RNase P protein content from bacteria to archaea to eukarya is that the smaller RNA subunits found in archaea and eukarya depend on a concomitant increase in protein content to stabilize a catalytically active conformation.

**Pop5, One of Several Protein Subunits of Archaeal and Eukaryotic RNase P.** Yeast two-hybrid experiments with eukaryotic and archaeal homologs suggested that Pop5 interacts with Rpp30 (26–29). Furthermore, *Pfu* Pop5 and Rpp30 form a functional pair that together can partially reconstitute *Pfu* RNase P (in the absence of the other proteins, Rpp21 and Rpp29) (H.-Y. Tsai and V. Gopalan, personal communication). The finding that a truncated version of the RNase P RNA missing the specificity domain (P11, P12, and the loop regions) can also be reconstituted with this pair of proteins suggests that these proteins are associating with the catalytic portion of the RNase P RNA and/or are involved in substrate binding. A crosslinking study performed by using the homologous human components presents compelling evidence that Rpp30 binds the catalytic RNA subunit (52), although Pop5 did not crosslink to the central RNA when tested in the same study.

NMR chemical shift perturbations (Fig. 4a) indicate that the Rpp30-contacting surface can be located to a patch of hydrophobic residues constituted largely by helix \( \alpha_3 \), strand \( \beta_3 \), and the end of strand \( \beta_1 \), on the nearly opposite side of Pop5 as the putative single-stranded RNA-binding face (Fig. 4bc). This is not an unlikely binding interface; a cluster of solvent-exposed hydrophobic side chains (Tyr-78, Ile-81, Phe-82, Leu-86, Ile-96, and Val-99) on this surface of Pop5 might reasonably be expected to recognize a similar hydrophobic patch of Rpp30 in binding (Fig. 4c). This finding is consistent with the structural similarity of Pop5 to the bacterial proteins and their mode of interaction with the RNA subunit and substrate (6, 12, 40, 41). That is, assuming the single-stranded 5′ leader of the precursor tRNA is bound in the opened \( \beta \) sheet of Pop5 as in the bacterial protein, and that helix \( \alpha_1 \) contacts the catalytic RNA subunit (as is done by the analogous helix \( \alpha_3 \) in the bacterial protein), the only available surface for Rpp30 binding is that in which the largest chemical shift perturbations are observed in the spectra of Pop5 when Rpp30 binds to it.
Summary. With the elucidation of the Pfu Pop5 structure, a functional role for the protein can be hypothesized based on its structural similarity to the bacterial RNase P protein and the RRM. Additionally, NMR chemical shift perturbations have revealed the surface of Pfu Pop5 used in binding to its protein partner Pfu Rpp30. These data provide clues regarding the role of Pop5 in the archaeal and eukaryal RNase P enzymes, although further work will be necessary to determine how all of the protein subunits assemble onto the RNase P RNA and help it perform catalysis, ultimately providing insight into the structural and functional connections between the three domains of life of this conserved yet compositionally variable enzyme.

Materials and Methods

Sample Preparation. The Pfu Pop5 gene (PF1378 from Pyrococcus furiosus DSM 3638) was amplified by PCR with Pfu genomic DNA as the template and gene-specific primers and cloned into pET33b. Clone authenticity was established by using automated DNA sequencing. Initial structural work with the wild-type protein was hindered by precipitation at ~1 mM concentration. Mutation of cysteine residues 42 and 72 to serine singly and as a double mutant using a protocol modified from the QuikChange Site-Directed Mutagenesis kit (Stratagene) resulted in three mutants with improved solubility. The C72S mutant was used for all experiments and is referred to as “Pfu Pop5”; the wild-type protein is referred to explicitly when applicable. The protein was expressed in Escherichia coli Rosetta (DE3) cells (Novagen) and purified from both the soluble and insoluble fractions by selective denaturation and ion exchange chromatography (see supporting information). Trypsin (Sigma) was used to probe the protein for regions accessible to proteolysis; 10% (wt/wt) trypsin was added to 700 μM Pfu Pop5 at room temperature, and the reaction was terminated at 30 min by addition of 1 mg/ml Pefabloc (Roche) and frozen until analyzed by ESI mass spectrometry (Q-Tof II, Micromass, Manchester, U.K.). The diffusion coefficient of Pop5 was assayed at differing concentrations by using a DynaPro-801 dynamic light scattering/molecular sizing instrument (Protein Solutions).

Pfu Rpp30 (PF1914) was prepared by subcloning the protein-coding gene from pET33b into the pHEM vector for bacterial expression as a fusion to His6-tagged maltose binding protein (53), purified by immobilized metal-affinity chromatography and processed with tobacco etch virus protease (54) (see supporting information). The Pop5-Rpp30 complex was formed by titrating Pop5 into Rpp30 in 10 mM sodium acetate, pH 5.0/0.5 mM KCl to final protein concentrations of 260 μM for Pop5 and 280 μM (a slight excess) for Rpp30.

NMR Spectroscopy. NMR spectra were recorded at 55°C on 600- and 800-MHz Bruker (Billericia, MA) Avance DMX and DRX spectrometers equipped with triple-resonance probes and pulsed-field gradients. NMR data processing and analysis were performed with NMRPipe (55) and NMRView (56). NMR samples typically contained 300–600 μM Pfu Pop5 and 20 mM sodium acetate (pH 5.0), 0.5 mM KCl, 50 μM Na3P, and 95% H2O/5% D2O. Backbone resonance assignments were made by using standard triple resonance spectra [HNCACB, CBCA(CO)NH, and HNCO] (57, 58) recorded at 600 MHz. Chemical shift perturbations were analyzed by comparing peak positions in two-dimensional heteronuclear 1H-15N correlated and three-dimensional HNCO spectra of Pop5 in the absence and presence of Rpp30: the peaks in highly overlapped regions or peaks with weak intensity in the Rpp30-sample were not included in analysis; the peaks corresponding to residues 97, 99, and 100 unambiguously vanished upon formation of the complex and were assigned the largest perturbation value. Signals were referenced to the residual water signal and adjusted for statistical compliance (59).

Crystallization. Crystals of Pfu Pop5 C72S were grown at room temperature by hanging drop vapor diffusion. The reservoir solution consisted of 250 mM imidazole (pH 6.5) and the hanging drop was prepared by mixing 1 μl of reservoir solution with 1 μl of 28 mg/ml protein in 10 mM sodium acetate (pH 5.0). A mercury derivative was prepared by adding 0.25 mM HgCl2 and 1 mM ethylmercuric phosphate to a 1-M solution of reservoir protein. Crystallization was performed at 15°C by hanging drop vapor diffusion. The reservoir solution consisted of 250 mM imidazole (pH 6.5), 20 mM sodium acetate, and 10% (vol/vol) polyethylene glycol 4000, 1 μl of reservoir solution with 1 μl of 1-M protein. Crystals were mounted in 0.7-mm glass capillaries (Hampton Research) and soaked in well solution (100% reservoir solution with 1 M mercury derivative) and then 100% reservoir solution for 10 min. Crystals of Pop5-Rpp30 were grown in a similar manner.

X-Ray Data Collection, Structure Determination, and Refinement. Crystals were mounted in 0.7-μl glass capillaries (Hampton Research). X-ray diffraction data for the native and mercury-soaked crystals were collected at room temperature using a Rigaku rotating anode generator and an R-AXIS-IV++ image plate detector. Reflection intensities were integrated and scaled using CrystalClear (Molecular Structure), and processed further by using the truncate module of the CCP4 program suite (60, 61) (see supporting information). Five heavy atom positions were located by using a direct methods heavy atom search procedure (RANTAN), and initial phases were calculated with MLPHARE. The phases were improved at 3.8-Å resolution by using the solvent flipping mode of SOLOMON with a solvent fraction of 65%, and the resulting electron density map allowed a polyalanine trace to be built using the program O (39). After forming the polyalanine model for the first monomer to the other four positions in the asymmetric unit, the phases were...
Tight fivefold noncrystallographic symmetry restraints were using PROCHECK (63).


R.C.W. and C.J.B. acknowledge support from the National Institutes of Health Foundation through Research Experiences for Undergraduates Grant MCB-0092962 (to M.P.F. and Venkat Gopalan). R.C.W. and C.J.B. acknowledge support from National Institutes of Health Grant GM070862, C.B.E. acknowledges support from National Institutes of Health Grant GM067947.