Chapter 3:

The oligomerization and ligand-binding properties of Sm-like archaeal proteins (SmAPs)
Abstract

Intron splicing is one example of the many types of RNA processing that directly utilize small nuclear ribonucleoprotein (snRNP) complexes. Sm proteins form the cores of most snRNPs, so to further elucidate structural principles of snRNP assembly, we have characterized the oligomerization and ligand-binding properties of Sm-like archaeal proteins (SmAPs) from *Pyrobaculum aerophilum* (*Pae*) and *Methanobacterium thermautotrophicum* (*Mth*). Ultracentrifugation shows that *Mth* SmAP1 is exclusively heptameric in solution, whereas *Pae* SmAP1 forms either disulfide-bonded 14-mers or sub-heptameric states (depending on the redox potential). By electron microscopy, we show that *Pae* and *Mth* SmAP1 polymerize into sheets composed of well-ordered polar fibers that are formed by head-to-tail stacking of heptamers. The crystallographic results reported here corroborate these findings by showing heptamers and tetradecamers of both *Mth* and *Pae* SmAP1 in several new crystal forms. The 1.9-Å resolution structure of *Mth* SmAP1 bound to uridine-5’-monophosphate (UMP) reveals conserved ligand-binding sites. The likely RNA binding site in *Mth* agrees with that determined for *Archaeoglobus fulgidus* (*Afu*) SmAP. Finally, we find that both *Pae* and *Mth* SmAP1 gel-shift negatively supercoiled DNA. These results distinguish SmAPs from eukaryotic Sm proteins, and suggest possible differences in their functions.

**Abbreviations:** snRNP, small nuclear ribonucleoprotein; SmAP, Sm-like archaeal protein; *Pae*, *Pyrobaculum aerophilum*; *Mth*, *Methanobacterium thermautotrophicum*; *Afu*, *Archaeoglobus fulgidus*; UMP, uridine-5’-monophosphate; MPD, 2-methyl-2,4-pentanediol; EM, electron microscopy; NCS, non-crystallographic symmetry; wt, wild type; nt, nucleotide; DTT, dithiothreitol; ss(D/R)NA, single-stranded (D/R)NA; OB-fold, oligosaccharide/oligonucleotide-binding fold
Introduction

Excision of non-coding regions (introns) is one of the most vital steps in the maturation of precursor mRNAs. Most eukaryotic protein-coding genes contain multiple introns, so high-fidelity pre-mRNA processing is essential to ensure a mature mRNA with correctly registered exons. A transiently stable assembly of five small nuclear ribonucleoproteins (snRNPs) catalyzes the simultaneous excision of introns and splicing of exons in eukaryotic pre-mRNA. This large assembly of uridine-rich snRNPs (U snRNPs) is known as the spliceosome. It contains five small nuclear RNAs (snRNAs) and at least 80 proteins, making it roughly the same size as the ribosome (sedimentation coefficient of $\approx 60S$). At various stages in its catalytic cycle, the spliceosome consists of the U1, U2, U4/U6, and U5 snRNPs. The recent isolation of a novel U1•U2•U4/U6•U5 “penta-snRNP” devoid of mRNA suggests needed modification of the long-held belief that spliceosome assembly requires pre-mRNA.

Extensive biochemical and genetic data have shown that stepwise binding of seven cytoplasmic Sm proteins to exported snRNAs is a key step in snRNP biogenesis (recently reviewed in ref. 6). Each U snRNP complex is composed of a $\approx 110-180$ nucleotide uridine-rich snRNA and two classes of proteins: (i) snRNP-specific proteins that provide snRNP-specific functions (e.g., U1A protein of U1 snRNPs), and (ii) a set of Sm or Sm-like (Lsm) proteins that are common to each snRNP core. The snRNA component contains a single Sm or Lsm binding site with the uridine-rich consensus sequence PuAU$^{n=4-6}$GPu (Pu = purine). However, specificity for this sequence is not very stringent and there can be redundancy in Sm•snRNA binding. The Sm sites are predicted
to be single-stranded RNA (ssRNA) regions flanked by stem-loop structures (reviewed in refs. 2, 4). Sm binding is highly sensitive to modifications of the flanking stem-loops and the Sm site of a given snRNA, and varies from one snRNA to another.\(^9\) Sm•snRNA binding may be modulated by recently discovered interactions between some Sm proteins and the survival of motor neurons (SMN) protein complex,\(^10\) and by symmetric dimethylation of arginine residues in some of the RG dipeptide repeats of Sm\(^11,12,13\) and Lsm\(^14\) proteins by a putative “methylosome”.\(^15\) In eukaryotes, the Sm D1•D2 and E•F•G heteromers simultaneously bind to snRNA to yield a “subcore” snRNP complex.\(^6,16,17\)

The final component to join the Sm complex (the B/B’•D3 heterodimer) triggers hypermethylation of the 5’ m\(^7\)G cap of snRNA to a trimethylated cap (m\(^3\)G). The m\(^3\)G cap and the snRNA•Sm core complex form a bipartite nuclear localization signal that results in transit of the snRNP core to the nucleus, where association of various snRNP-specific proteins completes the assembly process.

Aside from their roles in assembly and nuclear import of snRNP cores, the primary structural or catalytic function of Sm proteins is not known. Forming the core of snRNPs, Sm proteins probably mediate critical RNA-RNA, RNA-protein, and protein-protein interactions, and may recruit both snRNP and non-snRNP proteins to the assembling spliceosome. The vast network of protein-protein and protein-RNA interactions in which Sm proteins probably engage was recently reviewed by Will and Lührmann\(^7,6\) and experimentally verified by genome-wide two-hybrid screens of yeast Lsm proteins.\(^18\) An example of such interactions is the discovery that the extended and highly charged C-terminal tails of Sm B, D1, and D3 are involved in sequence-
independent binding to pre-mRNA substrate, and may stabilize U1 snRNP•pre-mRNA interactions.\textsuperscript{19,20} Chemical cross-linking of the U1 snRNP 70K protein to the Sm B and D2 proteins provides suggests a possible role of snRNP Sm cores in recruiting other proteins to the assembling spliceosome.\textsuperscript{21} The importance of Sm proteins in RNP assemblies is underscored by their phylogenetic distribution: in addition to the canonical Sm and Lsm proteins found in eukaryotes ranging from yeast to humans, the Sm-like archaeal protein (“SmAP”) family has been discovered.\textsuperscript{22,23} The recent demonstration that the \textit{E. coli} bacteriophage host factor \textit{Hfq} is an Sm-like protein provides the first example of a eubacterial Sm protein.\textsuperscript{24,25} These results imply fundamental roles for Sm proteins in the early evolution of RNA metabolism.

Sm proteins have a remarkable tendency to associate into ring-shaped oligomers. Prompted by biochemical and genetic data, electron microscopic (EM) investigations of U snRNP particles revealed the “doughnut-shaped” ultrastructure of Sm and Lsm cores.\textsuperscript{26,27} The realization that Sm and Lsm proteins occur in groups of at least seven paralogs within the genome of a given organism suggests that snRNP cores are formed from Sm heteroheptamers, and two recent results verify this. First, Stark \textit{et al.} reconstructed a 10-Å resolution map of the U1 snRNP by cryo-EM and found that a model of the Sm heptamer could be docked into the ring-shaped body of the snRNP.\textsuperscript{28} Next, the \textit{in vivo} stoichiometry of Sm proteins in yeast spliceosomal snRNPs was determined by a differential tag/pull-down assay, showing that the snRNP core domain contains a single copy of each of the seven Sm proteins.\textsuperscript{29} Intriguingly, stable subheptameric Sm complexes have been suggested as intermediates along the snRNP core
assembly pathway (e.g., a D1•D2•E•F•G complex capable of binding snRNA), and ultracentrifugation and EM show that some of these oligomers (e.g., a (E•F•G)\(_2\) hexamer) can form ring-like structures that resemble intact, heptameric snRNP cores. Such findings emphasize the importance of the Sm heptamer at the snRNP core, and suggest the possibility of other oligomeric states.

There is no atomic-resolution structure of a eukaryotic snRNP core; however, the crystal structures of Sm-like archaeal proteins from \(Afu\), \(^{31}\) \(Pae\), \(^{23}\) and \(Mth\)\(^{32}\) reveal a ring-shaped Sm heptamer and provide a model for snRNA binding in the snRNP core. Sm monomers fold as strongly bent, five-stranded antiparallel \(\beta\)-sheets, \(^{33}\) and form toroidal heptamers that are perforated by a conserved cationic pore. The inner surface of this pore appears to be the oligouridine binding site. The structural similarity between the SmAP1 monomers and dimers and the nearly identical human Sm D1•D2 and B•D3 heterodimers\(^{34}\) justifies SmAP1-based models for the heptameric snRNP core. In order to elucidate further structural principles of snRNP assembly, we have characterized the oligomerization and ligand-binding properties of \(Pae\) and \(Mth\) SmAP1. Many of our results distinguish these two Sm-like archaeal proteins from eukaryotic Sm proteins, and suggest that the functions of archaeal Sm proteins may be quite different from the snRNP-based roles of eukaryotic Sm proteins.
Materials and Methods

Cloning, expression, and purification of Pae and Mth SmAP1s

A genomic phosmid clone that contains the Pae SmAP1 open reading frame (ORF) was kindly provided by the laboratory of Jeffrey H. Miller (UCLA), and genomic DNA containing the Mth (strain ΔH) SmAP1 ORF was kindly provided by the laboratory of John Reeve (Ohio State University). Based on the ORF DNA sequences, we used these primers for PCR amplification with Deep Vent<sub>R</sub> polymerase (New England Biolabs):

(Pae sense) 5’ CCATATGGCCTCGGATATATCT 3’
(Pae antisense) 5’ AAGCTTTCCCCGTCCTGGTACT 3’
(Mth sense) 5’ CCATATGATAGATGTGAGTTCAC 3’
(Mth antisense) 5’ AAGCTTTCCCCGGATATGTA 3’

Blunt-end PCR products were cloned into a pET-22b(+) expression vector (Novagen) via intermediate subcloning into the pCR-Blunt vector (Invitrogen). Ligation products were directly transformed into chemically competent NovaBlue E. coli (Novagen), and plasmids from overnight cultures of positive transformants (as assayed by PCR screening of colonies) were mini-prepped (Qiagen). DNA sequencing (David Sequencing) of these plasmids verified that the expressed proteins would contain a C-terminal His-tag after a 10-residue serine protease-sensitive linker. That is, the constructs were designed as: wild type (wt) SmAP1 + GR*GKLAAALEHHHHH (single letter amino acid codes, * indicates intended protease site). Recombinant proteins were over-expressed in BL21(DE3) E. coli at 37°C by standard protocols using 1 mM isopropyl-β-D-thiogalactoside induction of the T7lac-based promoter. At least 120 mg of soluble protein was expressed per liter of cell culture. The Cys8→Ser mutant of Pae SmAP1 was created...
in a similar manner, except that site-directed mutagenesis was achieved via overlap-extension PCR with an additional pair of primers that contained the mutant site.

Harvested cells (stored at –20°C overnight) were thawed and re-suspended in a high salt concentration buffer (20 mM NaHEPES pH 7.8, 1.5 M NaCl, 0.5% v/v Triton X100, 30 mM PMSF). Cells were lysed by a combination of lysozyme treatment (0.3 mg/ml chicken egg white lysozyme) and French-press (1000 psig). Initial purification of the thermostable proteins was achieved by heating the cleared supernatant to ≈ 80°C, followed by high-speed centrifugation (37,000 g) to remove the bulk of denatured *E. coli* proteins. The SmAP1-His6x proteins were further purified by affinity chromatography on a Ni²⁺-charged iminodiacetic acid-sepharose column (both proteins eluted as broad peaks over the range 170-400 mM imidazole). Both *Mth* and *Pae* SmAP1 were >99% pure by this point (as determined by SDS-PAGE and MALDI-TOF mass spectrometry). Since the His6x tag prevents heptamer formation for some SmAPs (unpublished data, Mura & Eisenberg), the next step was proteolytic removal of the C-terminal tag for both *Pae* and *Mth* SmAP1 (*wt Mth* SmAP1 is 81-amino acid residues, and has a MW of 9,029 Da; *wt Pae* SmAP1 is 80-amino acid residues, and has a MW of 8,800 Da).

The His-tag and most of the linker were removed by limited proteolysis with trypsin (since thrombin was ineffective). The peak fractions from the Ni²⁺-column that were judged as pure by SDS-PAGE were pooled and dialyzed at room temperature into phosphate-buffered saline (PBS) supplemented with 15 mM EDTA (to prevent His-tag mediated aggregation). The EDTA concentration was gradually reduced to zero over 2-3 buffer exchanges. Porcine trypsin was added to the SmAP1 (at ≈ 1 mg trypsin per 100 mg
SmAP1), and complete removal of the tag occurred after \( \approx 4 \) h at 37°C (the extent of proteolytic digestion was assayed by MALDI-TOF spectra of time points). Transfer of the protein to 4°C and addition of a protease inhibitor (50 mM PMSF) terminated the reaction. The amino acid composition of \( Mth \) and \( Pae \) SmAP1 led to calculated isoelectric points of \( \approx 5.2 \) and 5.8, respectively; therefore, anion exchange chromatography was used to separate cut (\( i.e., \) wt) SmAP1 from trypsin, uncut protein, and any other contaminants. In preparation for anion exchange chromatography on a quaternary ammonium matrix (UNO-Q6, BioRad), \( Pae \) SmAP1 was dialyzed against 20 mM Tris, pH 8.55. \( Mth \) SmAP1 was insoluble at 4°C or in the Tris-alone buffer, and had to be dialyzed versus 20 mM Tris pH 8.55, 30 mM EDTA pH 8.0 at room temperature (EDTA did not interfere with chromatography). Both SmAP1s eluted at \( \approx 80 \) mM NaCl in the salt concentration gradient. Pure fractions (assayed by SDS-PAGE and MALDI-TOF) were pooled and dialyzed into a buffer for crystallization.

**Crystallization of Pae SmAP1 and Mth SmAP1**

For \( Pae \) SmAP1 crystallization, the protein buffer was “XB” (10 mM Tris pH 7.8, 5 mM EDTA pH 8.0), and for \( Mth \) (which requires higher ionic strength buffers for solubility) it was “XB6β” (10 mM Tris pH 7.8, 5 mM EDTA pH 8.0, and 0.1 M NaCl). Protein concentrations in these buffers were increased to various values for crystallization (noted below) by using 3 kDa molecular weight cutoff Centripreps to reduce sample volume. Initial sparse matrix screening of crystallization conditions utilized the commercially available kits from Hampton Research and Emerald Biosystems, Inc. Final, optimized \( Pae \) SmAP1 crystals of the \( C222_1 \) form were grown by the hanging-drop vapor
diffusion method in 24-well Linbro trays. An 11 μl drop [4 μl well buffer + 5 μl wt 29.6 mg/ml *Pae* SmAP1 + 1 μl 0.1 M dithiothreitol (DTT) + 1 ml 0.1 M uridine-5’-monophosphate (UMP)] was equilibrated against a 800 μl well [0.1 M sodium acetate pH 8.20, 0.1 M ammonium acetate, 8.6% w/v PEG-4000, and 23.8% v/v glycerol] at room temperature (≈ 19.8°C). Orthorhombic crystals reached maximum dimensions of 0.1 × 0.1 × 0.3 mm within 5 days. Hanging drops contained a mixture of the new *C222*₁ crystals and the previously reported *C2* form (used to solve the original *Pae* SmAP1 structure).  

Three forms of *Mth* SmAP1 crystals were obtained under three conditions by hanging-drop vapor diffusion at room temperature. For the *P1* form, *Mth* SmAP1 was at 56 mg/ml in buffer XB6β. The drop was 4 μl of protein + 4 μl of well buffer. The well was 600 μl of [0.1 M sodium citrate pH 5.60, 15% w/v PEG-4000, 0.2 M ammonium acetate]. Crystals grew to maximum dimensions of ≈ 0.1 × 0.1 × 0.25 mm within 7 days. For the *P2*₁₂₂₁ form, *Mth* SmAP1 was at 42 mg/ml in buffer XB6β. The drop was 3 μl of protein + 3 μl of well buffer. The well was 600 μl of [0.1 M Tris pH 8.50, 10% v/v isopropanol]. Crystals grew to maximum dimensions of ≈ 0.3 × 0.3 × 0.6 mm within 3 days. For the *P2*₁ form, *Mth* SmAP1 was at 30.3 mg/ml in a modified form of buffer XB6β that contained a 26-nt single-stranded DNA [10 mM Tris pH ≈ 7.7, 3 mM EDTA pH 8.0, 55 mM NaCl, 0.6 mM ssDNA]. The drop was 2.5 μl of protein/ssDNA + 2.5 μl of well buffer + 1 μl of 0.1 M UMP. The 600 μl well contained 55 μl of 1.0 M sodium citrate (pH 5.6), 5 μl of 1.0 M sodium citrate (pH 8.0), 60 μl of 2.0 M ammonium acetate,
180 μl of neat MPD and 300 μl of sterile dH2O (interestingly, 2.5 M 1,6-hexanediol could be substituted for MPD). Crystals grew to maximum dimensions of ≈ 0.15 × 0.15 × 0.25 mm within 7 days.

**Cryoprotection and data collection**

The C2221 Pae SmAP1 and P21 Mth SmAP1 crystals did not require the addition of a cryosolvent, due to the 23.8% v/v glycerol or 30% v/v MPD in those drops, respectively. The other two Mth SmAP1 crystal forms had to be cryoprotected as follows: (i) for the P1 form, ethylene glycol was added directly to the drop to a final concentration of ≈ 20% v/v, and crystals were allowed to soak for 20 sec prior to mounting in a cryo loop; (ii) for the fragile P212121 crystals, the cryoprotectant was ethylene glycol (mixed with well buffer), and had to be introduced gradually over several hours (in ≈ 5% v/v increments). The P212121 crystals were allowed to soak for only a very short time (2-3 sec) at the final ethylene glycol concentration (30% v/v). Diffraction data were collected either at the synchrotron (P1 and P21 form Mth xtns) or in-house (P212121 Mth and C2221 Pae crystals) on an ADSC Quantum-4 charge-coupled device (CCD) detector. All crystals were mounted in a cryogenic nitrogen stream at -168°C for data collection.

After autoindexing, all images were indexed/integrated/reduced in DENZO, and reflections were scaled and merged in SCALEPACK. Complete data sets were collected from single crystals (Table 3.1). Unit cell dimensions for the Pae C2221 form are a = 91.83, b = 113.76, c = 126.59 Å; for the Mth crystals they are: a = 45.07, b = 54.08, c = 62.35 Å, α = 87.58°, β = 72.86°, γ = 81.45° (P1); a = 65.25, b = 109.96, c = 83.76 Å, β = 95.81° (P21); a = 40.37, b = 114.70, c = 238.60 Å (P212121). The large unit cell edge of
the \textit{Mth} \textit{P2_12_12_1} crystals ($c = 238.60$ Å) led to spot overlap for high-resolution reflections ($d < 3$ Å), so multiple data sets were collected at two $2\theta$ values (0°, -12°) for two crystal alignments (related by a 45° azimuthal rotation).

\textit{Structure determination, refinement, and validation}

Initial phases for the \textit{C222_1} \textit{Pae} SmAP1 structure were determined by the evolutionarily-programmed molecular replacement algorithm (EPMR).\textsuperscript{36} The most reasonable Matthews coefficient ($V_M = 2.58$ Å\textsuperscript{3}/Da) corresponded to a heptamer in the asymmetric unit (a.u.), so the search model was the identical \textit{Pae} SmAP1 heptamer from the \textit{C2} crystal form.\textsuperscript{23} The EPMR solution was used for manual model building in the program \textit{O},\textsuperscript{37} and model refinement in \textit{CNS}.\textsuperscript{38} Refinement in \textit{CNS} proceeded by standard protocols, using the maximum-likelihood target function for amplitudes (mlf), bulk solvent correction, and anisotropic B-factor correction terms. Seven-fold non-crystallographic symmetry (NCS) was determined by calculation of a locked self-rotation function, but NCS restraints were not imposed at any time. Solvent molecules were added as necessary (water, glycerol, acetate). Refinement of individual atomic positions, isotropic temperature factors, and simulated annealing torsion angle dynamics was performed in most rounds. Each refinement round ended with inspection of the agreement between the model and $\sigma_A$-weighted $2F_o - F_c$, $F_o - F_c$, and simulated annealing omit maps (the latter only as necessary).

Determination of the \textit{Mth} \textit{P1} structure proceeded in two steps. First, a homology model of the \textit{Mth} SmAP1 heptamer was built from the \textit{Pae} SmAP1 structure using an in-house script (unpublished, Mura & Eisenberg), and was used as a search model for
molecular replacement with EPMR ($V_M = 2.29 \, \text{Å}^3/\text{Da}$ for a single heptamer in the $P1$ cell). Next, the unambiguous EPMR solution was converted to a polyalanine model and subjected to free-atom model refinement with the ARP/wARP program in the “molrep” mode (side chains from the $Mth$ sequence were built in the final wARP stage). This $Mth$ $P1$ structure was refined in the usual manner with CNS, as described above for the $Pae$ $C222_1$ structure. The $P2_1$ and $P2_12_12_1$ $Mth$ structures were solved by molecular replacement (EPMR) with the refined $P1$ $Mth$ model. Self-rotation functions and $|F_o|^2$ Patterson maps were calculated to deduce the NCS between heptamers in the $P2_1$ and $P2_12_12_1$ forms (each of which contains 14 monomers per a.u.). Solvent was added as necessary for all structures (see Table 3.1), and no NCS restraints were enforced at any point in the refinements.

Refinement statistics for the single $Pae$ and three $Mth$ structures are shown in Table 3.1. Each of the four protein models is complete, except for anywhere from 6-11 missing N-terminal residues in various models (see PDB files). The stereochemistry and geometry of each SmAP1 monomer was validated with the programs PROCHECK\textsuperscript{39} and ERRAT,\textsuperscript{40} and found to be acceptable (e.g., no residues in the disallowed region of $\phi, \psi$ space for the $Pae$ $C222_1$ model). Final model coordinates and diffraction intensity data were submitted to the PDB, with ID codes 1JBM, 1LOJ, 1JRI, and 1LNX (see Table 3.1).

**Analytical ultracentrifugation**

The $wt$ $Pae$ protein in 75 mM NaCl, 10 mM Tris, pH 7.8, was examined by sedimentation velocity in a Beckman Optima XL-A analytical ultracentrifuge at 52,000 rpm and 20°C using absorption optics at 273 nm and a 12 mm pathlength double sector.
cell. The sedimentation coefficient distribution was determined from a \( g(s) \) plot using the Beckman Origin-based software (Version 3.01). The peak sedimentation coefficient was corrected for density and viscosity to an \( S_{20,\text{wat}} \) value by using a value for the partial specific volume at 20°C of 0.743 (calculated from the amino acid composition\(^\text{41}\) and corrected to 20°C\(^\text{42}\)).

Sedimentation equilibrium runs were performed on all three proteins – \textit{wt Mth}, \textit{wt Pae}, and the \textit{Pae} C8S mutant – in 150 mM NaCl, 10 mM Tris, pH 7.8, again using a Beckman Optima XL-A analytical ultracentrifuge. Each protein was examined at three different concentrations and four speeds, using 12 mm pathlength six-sector cells. Protein concentrations used were 3.4, 0.69 and 0.19 mg/ml for \textit{wt Pae}; 5.9, 1.26 and 0.32 mg/ml for the C8S mutant of \textit{Pae}; and 4.1, 0.85 and 0.22 mg/ml for \textit{wt Mth}. Rotor speeds were 8,000, 10,000, 12,500 and 14,500 rpm. Protein concentration was monitored by absorption at 280 nm and, for the lowest protein concentrations, at 232 nm. A partial specific volume of 0.743, calculated as described above, was used for all three proteins. Individual scans were analyzed using the Beckman Origin-based software (Version 3.01) to perform a nonlinear least-squares exponential fit for a single ideal species, thus giving the weight-averaged molecular weight for each protein.

\textit{Transmission electron microscopy}

The following protein samples were prepared for electron microscopy: (1) 0.5 mg/ml wild-type \textit{Mth} SmAP1 in 10 mM Tris pH 7.5, 60 mM NaCl, (2) 1.2 mg/ml wild-type \textit{Pae} SmAP1 in 25 mM Tris pH 7.5, 30 mM NaCl, (3) 1.1 mg/ml C8S mutant \textit{Pae} SmAP1 in the same buffer as the \textit{wt} protein, and (4) 1.2 mg/ml wild-type \textit{Pae} SmAP1 in
reductant buffer (25 mM Tris pH 7.5, 30 mM NaCl, 10 mM DTT). Carbon-coated parlodion support films mounted on copper grids were made hydrophilic immediately before use by high voltage, alternating current glow-discharge. Protein samples were applied directly onto the grids and allowed to adhere for 2 min. Grids were rinsed with distilled water and negatively stained with 1% w/v uranyl acetate. Specimens were examined in a Hitachi H-7000 electron microscope at an accelerating voltage of 75 kV.

**Gel-shift assays**

For gel-shift experiments, negatively supercoiled plasmid DNA was prepared by transforming the plasmid into *E. coli* BL21(DE3) cells and mini-prepping (Qiagen) it from spun down cells that had reached stationary phase. Several different plasmids were tested, including ones derived from pUC18, pACYC, pET-22b(+) (Novagen), and pCR-Blunt (Invitrogen). Titration of plasmids with ethidium bromide was used to verify the negative superhelicity of the DNA via electrophoretic mobility changes in agarose gels. Single-stranded DNAs of various lengths and sequences were synthesized by Integrated DNA Technologies, Inc., and were re-hydrated in 10 mM Tris pH 7.8 (e.g., the 26-mer in Fig. 3.8(b) with the following sequence: \(5'\text{CGGATCCTCAGTAAAAAGTGCGGAAA}3'\)). Stock solutions of protein were wt *Pae* at 5.6 mg/ml in buffer XB (see above) or wt *Mth* at 5.6 mg/ml in buffer XB6β (see above). Except as noted, buffer, DNA, and protein samples were mixed to produce 25- or 50-μl reactions that were incubated at room temperature (generally for 30-60 min). Gel-shift of the DNA was assayed by electrophoresis at a constant voltage (120V) in 1.3% or 1.5% w/v TAE/agarose gels. Examples of typical reactions are shown in Fig. 3.8.
Results

Crystallization and determination of the Pae and Mth SmAP1 structures

As with many proteins, crystallization of Pae SmAP1 was not straightforward. Wild-type (wt) Pae SmAP1 crystallized only in the presence of dithiothreitol (DTT), as described in the Methods section. Identical crystallization buffers that lacked DTT failed to produce crystals, and presumably this additive is required because it reduces the seven disulfide bonds that form between Cys8 residues in the Pae tetradecamer (which can therefore be thought of as a dimer of heptamers rather than as a heptamer of dimers). We found that other reductants (e.g., β-mercaptoethanol) can substitute for DTT to yield crystals, although such crystals are of poorer quality than the DTT-based condition. Apparently, reduction of the disulfides frees heptamers to crystallize independently in orientations that relax crystal lattice strain, even when the 14-mer persists in the crystal (as in the $C_{2221}$ form reported here). The only other notable (but unnecessary for crystallization) additive to the Pae crystallization condition was uridine-5′-monophosphate (UMP).

As shown in Table 3.1, diffraction data extended to at least 2.05-Å resolution for the $C_{2221}$-form Pae SmAP1 crystals. Previously, we determined the crystal structure of Pae SmAP1 in spacegroup C2 by multiple-wavelength anomalous dispersion phasing. Thus, the $C_{2221}$ structure reported here was solved by the stochastic molecular replacement method in the EPMR program, using the C2 heptamer as a search model (the Matthews coefficient and 7-fold NCS in the locked self-rotation function suggested a heptamer in the $C_{2221}$ asymmetric unit). However, NCS restraints were not imposed
During crystallographic refinement. As discussed in detail below, only the uridine fragment of UMP was built into the final refined model, and atomic occupancies ($q$) were refined only for uridine (not for any other ligand or protein atoms). Partial occupancies for uridine atoms were restricted to a reasonable range ($0.2 < q < 1.5$). The final structure was refined to an $R/R_{\text{free}}$ of 18.2%/22.6%, with reasonable model geometry (Table 3.1) and no outliers in a Ramachandran plot.

Crystallization of $Mth$ SmAP1 was relatively simple, and in fact this protein could be crystallized in three forms ($P1, P2_1, P2_{1121}$) under three dissimilar conditions (also, a fourth form was crystallized by Collins et al.\textsuperscript{32}). The most intriguing result is that $Mth$ SmAP1 crystallized in the $P2_1$ form only in the presence of a single-stranded DNA (ssDNA) to which it was thought to bind, even though ssDNA was not found in the crystal structure. Since diffraction data were obtained from the $P1$ form before the first $Mth$ SmAP1 structure was reported by Collins et al., we solved the $P1$ $Mth$ structure by a combination of molecular replacement and free-atom model refinement (in ARP/wARP). Briefly, a homology model of $Mth$ SmAP1 was built from the $Pae$ SmAP1 structure. An unambiguous molecular replacement solution was found for this search model against the $Mth$ $P1$ data (using EPMR). In order to reduce $Pae$ model bias, this solution was converted to polyalanine and phases from this initial model were used to autobuild a completely new model with the ARP/wARP program. Initial phases for the $P2_1$ and $P2_{1121}$ $Mth$ data were obtained by molecular replacement with the refined $P1$ model (as summarized in Table 3.1). No NCS restraints were applied in refinement of any of the $Mth$ structures, and various non-protein molecules were built into electron density as
appropriate (based on the crystallization condition and $F_o - F_c$ density >+3σ in strength). Electron density for the UMP-binding sites was more interpretable in *Mth* SmAP1 than in the *Pae* structure, and permitted model building of six complete UMPs (only uridine fragments were built for the other eight UMPs in the *Mth* 14-mer). As with the *Pae*•UMP model, partial occupancies of UMP atoms were refined (0.2 < q < 1.5). All three *Mth* structures were refined to reasonable values of $R/R_{\text{free}}$ and model geometries (Table 3.1).

**Comparisons of known SmAP monomer, dimer, and heptamer structures**

Several structures of Sm proteins and SmAPs are now available, and make possible the comparative structural analyses of these proteins. The previously reported *Mth* SmAP1 heptamer structure\(^{32}\) is virtually identical to the *Mth* structures reported here (0.65 Å RMSD for superimposition of the *P1* heptamer using mainchain atoms). The results from pairwise comparisons of the *Pae, Mth, and Afu* SmAP1s are shown in Fig. 3.1 and Table 3.2, and show that the compact, ≈ 80-amino acid SmAP monomer structures are nearly identical. The most similar monomer structures are the *Afu/Mth* pair (0.51 Å RMSD), and the most dissimilar are *Mth/Pae* SmAP1 (1.02 Å RMSD). These values do not correlate to pairwise sequence similarities. The increase in pairwise RMSDs in going from monomer alignments to dimers and heptamer alignments (Table 3.2) suggests that there are slight rigid-body variations in the monomer orientations in the higher-order oligomers (i.e., slight variations in interfaces cause the RMSDs to propagate when comparing heptamers to dimers and dimers to monomers).

The absolute conservation of the dimer interface in three different SmAP1s is emphasized by the view in Fig. 3.1. Sequence conservation is low for interfacial residues
relative to the rest of the SmAP monomer sequence; this is probably because the interface is largely formed by hydrogen bonding between mainchain atoms of the β4 strand of one monomer and the β5 strand of an adjacent monomer. The interface structure is also conserved between SmAPs and human Sm heterodimers.23 The main structural difference in the three SmAP1 heptamers is the width of the pore: \( \approx 8-9 \, \text{Å} \) diameter for \( \text{Pae} \) versus \( \approx 12-15 \, \text{Å} \) in \( \text{Mth} \) and \( \text{Afu} \). Such variation largely arises from differences in the structures of the pore-forming loops L2 and L4 (Fig. 3.1). Mapping of the phylogenetic conservation of SmAP residues onto the \( \text{Pae}, \text{Mth}, \) or \( \text{Afu} \) heptamer structures shows that most of the conserved residues cluster about the pore region (data not shown, Mura & Eisenberg). One of the least conserved features of the SmAP1 heptamer structures is the calculated electrostatic potential of the surfaces: the \( \text{Pae} \) and \( \text{Mth} \) heptamers display a strongly acidic L4 face, while the surface of the \( \text{Afu} \) heptamer is much more basic (ref. 23 for \( \text{Pae} \), ref. 31 for \( \text{Afu} \), and unpublished data for \( \text{Mth} \), Mura & Eisenberg).

**Various oligomeric states of SmAP1, including sub-heptamers and 14-mers**

Biophysical characterization of \( \text{Pae} \) and \( \text{Mth} \) SmAP1 by a variety of methods reveals peculiar oligomerization properties. These methods include: mass spectrometry, size exclusion HPLC, native polyacrylamide gel electrophoresis (PAGE), and analytical ultracentrifugation. Sedimentation velocity ultracentrifugation revealed that \( \text{wt Pae} \): (i) is monodisperse in solution; (ii) has a symmetric and narrow Gaussian-shaped distribution of sedimentation coefficients, with a coefficient at 20°C of \( S_{20,w} = 6.49 \, \text{S} \); and (iii) has a frictional coefficient ratio close to one (\( ff_o = 1.2 \), where \( f = \) experimentally derived frictional coefficient and \( f_o = \) ideal frictional coefficient for a sphere with the MW of
SmAP1). These preliminary results suggested a roughly spherical, high-order \textit{Pae} oligomer \((\text{SmAP1}_n)\) with \(n \approx 12 \pm 2\) (data not shown, Mura, Phillips, & Eisenberg).

The results of equilibrium sedimentation analyses of \textit{wt Mth}, \textit{wt Pae}, and the C8S mutant of \textit{Pae} SmAP1 are shown in Fig. 3.2. Molecular weights were estimated by fitting experimental curves to single exponential models, and include a roughly 2-3\% error (residuals are shown in the top panels). The calculated molecular weight of \textit{wt Pae} suggests that it exists as a tetradecamer. Other data suggested a disulfide-bonded 14-mer (see Discussion), so the single cysteine of \textit{Pae} SmAP1 was mutated to serine to give the C8S mutant of \textit{Pae} SmAP1. Sedimentation results with this mutant can be fit only by species with molecular weights much less than that of a heptamer (\textit{e.g.}, the 46.7 kDa species shown in Fig. 3.2(b)), suggesting pentameric or hexameric states \((n = 5\) gives a MW of \(\approx 45\) kDa). The monodispersity of the data in Fig. 3.2(b) suggests a single, stable sub-heptameric complex. In contrast to \textit{Pae}, sedimentation equilibrium data for \textit{Mth} SmAP1 show that it only forms a stable, monodisperse heptamer (Fig. 3.2(c)). The concentration dependence of the experimentally calculated MWs (not shown), as well as the slight upward concavity of the residuals in Fig. 3.2(b) and 2(c), provide additional evidence for \textit{Pae} and \textit{Mth} SmAP1 monomer \(\leftrightarrow\) oligomer association reactions.

\textit{Polymerization of SmAP1 into polar fibers}

The polymerization of both \textit{Pae} and \textit{Mth} SmAP1 into well-ordered fibers was an unexpected result, and is shown in the transmission electron micrographs (EM) of Fig. 3.3. Protein samples were in standard buffers (\textit{e.g.}, 25 mM Tris pH 7.5, 30 mM NaCl for \textit{Pae} SmAP1), and reproducibly formed the striated sheets of fibers seen in these EMs.
Measurement of the sheet and fiber dimensions, together with the diameters of SmAP1 heptamers from crystal structures (≈ 70-75 Å), suggests a model in which the fibers are formed by head-to-tail stacking of heptamers, with the SmAP1 7-fold axis roughly parallel to the fiber axis (see white arrows in Fig. 3.3(b)). Several fibers may associate laterally to form sheets, such as those seen most clearly in Figs. 3(a) and (b).

In order to test this head-to-tail stacking model, we assayed fiber formation by wt Pae and the C8S mutant. Under oxidative conditions, wt Pae SmAP1 forms disulfide-bonded 14-mers in which the highly acidic L4 faces are exposed at either end of the barrel-shaped structure (see the Pae 14-mer in Fig. 3.4(c)). Such a 14-mer would be constrained to form only head-to-head interfaces (i.e., loop L4 face-to-loop L4 face) in a fiber, and would probably not do so because of the unfavorable electrostatic cost of closely apposing these anionic faces (at least not at the neutral pHs or low ionic strength conditions in which the SmAP1s were buffered). As expected, wt Pae forms only ring-shaped structures under oxidative conditions (Fig. 3.3(c)). However, when the seven disulfide bonds that link heptamers into 14-mers are eliminated, Pae SmAP1 assembles into fibers with roughly similar morphologies as Mth fibers. Polymerization can be achieved either by addition of a reducing agent (as in Fig. 3.3(d)) or by mutation of the cysteine (C8S mutant in Fig. 3.3(e)). Such fiber formation has been hitherto unreported for Sm proteins.

**Packing of Mth and Pae SmAP1 heptamers in four crystal forms**

Crystallization of Mth and Pae SmAP1 in several forms is a fortuitous result, since different packing geometries of SmAP1 heptamers in these various crystal forms
shed light on the oligomerization results described above. The \textit{Pae} SmAP1 \textit{C222}_{1} structure differs from the original \textit{C2} form in that heptamers pack face-face in the orthorhombic lattice to give a 14-mer with 72-point group symmetry. The crystal packing is shown in Fig. 3.4\textit{(a)} and interacting surfaces are shown in Fig. 3.4\textit{(b)}. This 14-mer is likely to be significant because: (i) it is consistent with the oligomerization results described above from biophysical characterization, (ii) it persists in the \textit{C222}_{1} lattice despite the requirement of DTT for crystallization (the sulfhydryls in Fig. 3.4\textit{(c)} are separated by >8-9 Å), (iii) the heptamer-heptamer interface occludes 7,550 Å\textsuperscript{2} of surface area, and (iv) it is corroborated by an \textit{Mth} 14-mer in the asymmetric unit of the \textit{P2}_{1} form. The total buried surface area in the heptamer interface of the \textit{P2}_{1} \textit{Mth} 14-mer is probably significant (3,000 Å\textsuperscript{2}), although less than half as much as the \textit{Pae} interface.

In the \textit{Mth} \textit{P1} and \textit{P2}_{1}\textit{2}_{1}\textit{2}_{1} lattices, SmAP1 heptamers form quasi-hexagonal layers that stack upon one another to give a crystal. In the \textit{P1} form these layers are staggered; however, in the \textit{P2}_{1}\textit{2}_{1}\textit{2}_{1} form these layers are in register. Fig. 3.4\textit{(c)} shows how the head-to-tail stacking of SmAP1 heptamers in this crystal form produces cylindrical tubes. A slight tilt of each heptamer (≈15\textdegree) with respect to the tube axis results in the SmAP1 7-fold axes being parallel, but not coaxial. Since they are formed by head-to-tail stacking of asymmetric heptamers, these tubes have a defined polarity, and, when rendered as molecular surfaces, they bear a striking resemblance to the EM fibers shown in Fig. 3.3. The tubes are also consistent with EM fiber dimensions. Therefore, the \textit{P2}_{1}\textit{2}_{1}\textit{2}_{1} crystal structure provides a model for the atomic structure of SmAP1 fibers.
addition to providing insights into oligomerization states, two of the crystal forms (Pae C222_1 and Mth P2_1) were used to investigate the ligand-binding properties of SmAP1s.

**Crystal structures of Mth and Pae SmAP1 bound to various ligands**

The 1.90-Å resolution crystal structure of Mth SmAP1 bound to uridine-5’-monophosphate (UMP) is shown in Fig. 3.5. The protein was co-crystallized with this ribonucleotide in an effort to determine its likely RNA-binding site (co-crystallization efforts were unsuccessful with single-stranded DNA or RNA oligonucleotides). As shown in Fig. 3.5(a), SmAP1 binds UMP with a 1:1 stoichiometry, so that 14 UMPs are bound to the 14-mer near the pore region. The orthogonal view in Fig. 3.5(a) shows that the UMPs are bound near the flat face of the heptamer, opposite the highly acidic loop L4 face. The structure of the SmAP1•UMP complex is shown in more detail in Fig. 3.5(b), where it can be seen that the binding site is well defined by electron density. The uracil ring intercalates between the guanidinium group of Arg72 and the imidazole ring of His46 (both of these residues are highly conserved in SmAPs). The planes of these three moieties are spaced ≈ 3.6 Å apart, as expected for energetically favorable stacking interactions between these conjugated π-systems. Individual protein-UMP contacts are discussed in greater detail below.

In addition to the expected UMP binding site, we found that each Mth SmAP1 monomer binds a molecule of MPD. The MPD binding site is somewhat solvent-exposed, near the periphery of the SmAP1 ring (Fig. 3.5(a)). Protein-MPD recognition is the same in each of the 14-monomers, and is shown in detail in Fig. 3.5(c). The primary contact is hydrogen bonding between the Ser21 hydroxyl and MPD, and there are several water-
mediated SmAP···H₂O···MPD contacts. The cryoprotectant in the P₁ and P₂₁₂₁ Mth SmAP₁ structures was ethylene glycol (Table 3.1), and in these structures some of the SmAP₁ monomers bind ethylene glycol in the same site as MPD.

A UMP binding site was found in the Pae SmAP₁•UMP co-crystal structure as well, but it is not as clearly defined in electron density as for Mth SmAP₁•UMP. Fig. 3.6 shows the Pae•UMP structure, which was refined to a resolution of 2.05-Å. UMPs bind to the same face of the heptamer as in Mth (i.e., the “flat face” opposite L₄), but are much more distant from the pore. As shown by the 2F₀ – Fc maps in Fig. 3.6(b), only the planar uracil fragment of UMP is clearly defined in electron density. Protein-UMP contacts are scarce in this binding site. Asn₄₆···UMP distances are shown by dashed lines in Fig. 3.6(b) only for the sake of completeness – the geometries of these interactions do not satisfy standard hydrogen bond criteria (in terms of both distances and angles), and favorable interactions probably do not exist between the UMP O₄ oxygen and the amide nitrogen of the Asn₄₆ side chain or between the UMP N₃ nitrogen and the amide oxygen of Asn₄₆. Also, there are no aromatic side chains in this region to participate in π-stacking interactions with the uracil base. As with Mth SmAP₁, additional small-molecule binding sites exist in Pae SmAP₁: many of the modeled glycerol molecules are bound identically near the loop L₄ faces (see Fig. 3.6(a)). The significance of such binding sites is unknown.

The structure of an Afu SmAP₁•U₃ complex was determined recently by Törö et al.,³¹ and permits a comparison of the mode of uridine recognition in Afu and Mth SmAP₁. The UMP binding site and SmAP₁···UMP interactions clearly differ in Mth and
Pae SmAP1, and, since the binding site was poorly resolved in the Pae•UMP complex, this structure was not included in the comparative analysis shown in Fig. 3.7. In the Mth and Afu structures, the aromatic pyrimidine ring intercalates between the side chains of the highly conserved Arg/His pair, and specific uracil recognition is achieved by hydrogen bonding of the uracil ring to the side chain of a strictly conserved asparagine residue (Asn48_{Mth}). The main chain amide nitrogen of a highly conserved aspartate (Asp74_{Mth}) also participates in hydrogen bonding to a uracil carbonyl oxygen. The pattern of hydrogen bond donors/acceptrs in the Asn48/Asp74_{Mth} pair makes binding specific for a uracil (if RNA) or thymine (if DNA) base. Additional specificity for uracil may be achieved by two means: (i) recognition of the 2’ hydroxyl of the ribose (RNA versus DNA discrimination) and (ii) the C5 carbon of the pyrimidine ring of uracil is only 3.8 Å from the backbone carbonyl oxygen of Leu45_{Mth} from an adjacent monomer, thus providing steric and polar discrimination against the methyl on the C5 carbon of thymine.

We crystallized Pae and Mth SmAP1 in the presence of various other nucleoside monophosphates (e.g., AMP, CMP, GMP), but there was no evidence for binding of these non-uridine NMPs (data not shown, Mura & Eisenberg). The only significant differences in uridine recognition by Mth and Afu SmAP1 are highlighted by green arrows in Fig. 3.7(b). These are: (i) hydrogen bonding of an Mth Arg72 side chain from an adjacent monomer to the 2’ hydroxyl of the ribose, and (ii) hydrogen bonding between a phosphate oxygen and an imidazole nitrogen from the His46 residue of an adjacent monomer. Overall, it appears that the mode of uridine recognition is conserved in the SmAP family.
Pae and Mth SmAP1 gel-shift negatively supercoiled DNA

In our initial attempts to determine the biochemical function of Pae SmAP1, we inadvertently found that this protein gel-shifts supercoiled plasmid DNA. This activity was further investigated for both the Mth and Pae SmAP1s, and examples of it are shown in Fig. 3.8. Migration of the negatively supercoiled plasmid “p5L1c1” is severely retarded by incubation with μM amounts of Mth heptamer in Fig. 3.8(a). Interestingly, the extent of gel-shift increases at higher concentrations of Mth SmAP1, until saturation of the effect occurs at ≈ 60 μM (compare lanes 7 and 8). A similar gel-shift occurs to supercoiled DNA when it is incubated with wt Pae SmAP1, as shown in lane 4 of Fig. 3.8(b). This experiment also shows that the gel-shift can be eliminated by incubation with a 26-nucleotide single stranded DNA (ssDNA). Inhibition of the gel-shift activity is titratable, and at higher concentrations of ssDNA there is no gel-shift (lane 8).

Similar DNA gel shift assays and control experiments have revealed that: (i) the Pae activity is specific for supercoiled (sup) plasmid DNA, whereas Mth SmAP1 gel-shifts both sup and linearized (lin) plasmids; (ii) Pae activity is eliminated by MgSO₄, whereas the dependence of Mth activity on divalent metals such as Ca²⁺, Mg²⁺, and Mn²⁺ is not as straightforward; (iii) ssDNA of any sequence and length >≈ 20-nt inhibits the gel-shift activity of Pae and Mth in a concentration dependent manner; (iv) both Pae and Mth activities are nonspecific with respect to the sup DNA; (v) Pae and Mth are not linearizing or otherwise cutting both strands of the sup DNA; (vi) Mth gel shift activity is not temperature-dependent at and above room temperature, whereas the extent of Pae-induced gel shift abruptly increases at ≈ 55-60°C. All of these results come from

66
experiments in which the migration of a large (>4,000 nt) plasmid DNA is assayed in agarose gels. Binding of Mth SmAP1 to any one of the ssDNAs that inhibit the sup DNA gel shift (e.g., Fig 8(b)) has been assayed in preliminary native PAGE experiments; these results suggest that ssDNA inhibits the sup DNA gel shift by directly binding to SmAP1.
Discussion

Comparative structural analysis of Sm proteins and SmAPs

Recent work has shown that SmAPs form a phylogenetically well-conserved family of proteins whose sequences are similar to eukaryotic Sm and Lsm proteins, and whose structures are nearly identical to the human Sm D3•B and D1•D2 heterodimer structures.\(^\text{32,23,31}\) The Sm and SmAP monomers form antiparallel, five-stranded $\beta$-sheets capped by a short N-terminal $\alpha$-helix. The Sm $\beta$-sheet is highly bent, into a $\beta$-barrel like structure that closely resembles proteins of the oligosaccharide/oligonucleotide binding (OB) fold family.\(^\text{43}\) As shown by Fig. 3.1 and Table 3.2, the $\textit{Afu}$, $\textit{Mth}$, and $\textit{Pae}$ SmAP1 monomers and homodimers are nearly identical to one another. Besides the N- and C-termini, the only significant deviations in Sm and SmAP monomer structures occur in the loops: the L2 and L4 loops are the most structurally variant regions in Fig. 3.1, and several eukaryotic Sm proteins have insertions of up to 30 amino acids in loop L4. Work in progress with another SmAP homolog that contains a similar L4 insertion ($\textit{Pae}$ SmAP3) shows that it too forms heptamers. The recent solution structure of the SMN Tudor domain, which interacts with Sm proteins to form snRNP cores, has provided an unexpected result: the SMN and Sm monomers have the same fold, and nearly identical structures.\(^\text{10}\) This raises the intriguing possibility that the SMN protein interacts with the Sm complex by forming mixed heteromers. Overall, there is a high degree of phylogenetic and structural conservation of the SmAP domain from archaea to eukaryotes.
Afu, Mth, and Pae SmAP1s form heptamers with remarkably similar structures, primarily because the structure of the homodimer interface is extremely well conserved (Fig. 3.1). Greater RMSDs for heptamers compared to dimers (and dimers compared to monomers) shows that a large fraction of the structural variation in higher-order SmAP oligomers (Table 3.2) is due to rigid-body displacements of monomers with respect to one another. A feature of the SmAP1 heptamers that is highly conserved in terms of sequence and overall structure is the central cationic pore. The largest difference between SmAP monomers (L2, L4 loops) results in the largest difference between SmAP heptamers: variation in the width of the pore in Pae (≈ 8-9 Å diameter) versus Afu and Mth (≈ 12-15 Å) is due to main chain and side chain rotamer variations in the L2 and L4 loops. The other significant difference between Afu, Mth, and Pae SmAP1s is the calculated electrostatic potential of the heptamer surface: the L4 face of the Afu surface is very basic, while these Pae and Mth faces are intensely acidic. Such differences are likely to be important for modulating putative SmAP-RNA interactions. Overall, the near identity of the human Sm D3•B and D1•D2 heterodimers to Afu, Mth, and Pae SmAP homodimers qualifies the SmAP1 homoheptamer as an accurate model for the Sm heteroheptamer of eukaryotic snRNP cores.

The oligomerization properties of SmAPs

Like the Lsm (but not Sm) proteins, Pae, Mth, and Afu SmAP1 form heptamers in the absence of RNA. We also found that SmAP1 exhibits complex self-association properties that result in 14-mers and sub-heptameric oligomers, in addition to the expected heptamers. Various oligomeric states were characterized in vitro (primarily by
ultracentrifugation, Fig. 3.2), revealing roughly spherical disulfide-bonded *Pae* SmAP1 14-mers and a monodisperse population of *Mth* SmAP1 heptamers. Additionally, we created a cysteine-free point mutant of *Pae* SmAP (C8S), and found that it forms sub-heptameric states (most likely pentamers). Interestingly, similar plasticity of oligomerization behavior has been reported for human Sm proteins. Lührmann *et al.* found that a human Sm E•F•G complex forms a stable oligomer – most likely a (E•F•G)$_2$ hexamer – whose ring-shaped structure resembles intact Sm heteroheptamers by EM.$^{17,30}$ One of these studies also found that stable, sub-heptameric complexes of human Sm proteins (e.g., a D1•D2•E•F•G pentamer) may be intermediates in the Sm-RNA assembly pathway.$^{17}$ Recently, another *Afu* SmAP was reported to form hexamers (personal communication cited in ref. 29), and in the structure of the human Sm D3•B the heterodimers pack as (D3•B)$_3$ hexamers in the asymmetric unit of the crystal.

We found that *Pae* and *Mth* SmAP1 reproducibly oligomerize into 14-mers, either *in vitro* (*Pae*) or in various crystal forms (*Pae* and *Mth*). The highly acidic L4 faces are exposed in the barrel-shaped 14-mers (Fig. 3.4), as expected from electrostatics. The heptamer-heptamer interface buries a large amount of surface area in both *Pae* (7,550 Å$^2$) and *Mth* (3,005 Å$^2$), suggesting the significance of these oligomers. Preliminary crystallographic data from another SmAP homolog (*Pae* SmAP3) show that it also forms 14-mers in the asymmetric unit (unpublished results, Mura & Eisenberg). The propensity of ring-shaped SmAPs to crystallize as head-head oligomers with dihedral symmetry is shared by another single-stranded RNA binding protein: the *trp* RNA-binding attenuation
protein (TRAP) forms a toroidal 11-mer that forms both head-head and head-tail 22-mers in the crystal\cite{44} (in fact, the structures of Sm and TRAP monomers are quite similar).

Perhaps the most novel property of SmAP1s is their polymerization into extremely well-ordered fibers. *Pae* and *Mth* SmAP1 samples at physiological conditions form these fibers, which we observe by EM. Three lines of evidence suggest that these fibers form by the head-to-tail stacking of heptamers (Fig. 3.3): differential fiber formation by C8S and *wt Pae* SmAP1, comparison of measured fiber dimensions with SmAP1 heptamer dimensions, and electrostatic considerations for the packing of highly charged heptameric disks. The packing of *Mth* SmAP1 heptamers in the *P*2_{1}2_{1}2_{1} lattice supports our head-to-tail polymerization model, and provides an atomic-resolution structure for the fibers (Fig. 3.4). Such peculiar oligomerization properties have not been reported for eukaryotic Sm proteins, and the biological significance of SmAP1 14-mers and homogeneous, fibrillar polymers is not yet known.

**Comparison and analysis of the ligand-binding properties of SmAPs**

Comparison of the structures of *Mth* SmAP1 bound to UMP and *Afu* SmAP1 bound to oligouridine (U₃) suggest a highly conserved mode of RNA recognition in SmAPs. UMP binds near the 7-fold axis, suggesting the pore as a putative RNA binding site. Diagrams of SmAP1···UMP interactions show that these proteins specifically bind the uracil base by a combination of π-stacking and hydrogen bond interactions with strictly conserved SmAP residues (Fig. 3.7). Differences between UMP binding in *Mth* and *Afu* are limited to interactions with the ribophosphate moiety, and may not be significant since *Mth* SmAP1 was co-crystallized with free UMP nucleotide, whereas *Afu*
SmAP1 was crystallized with a U₃ oligouridine. The oligo(U) specificity of RNA binding to *Afu* SmAP1 mimics the substrate specificity of eukaryotic Sm proteins. The geometry of binding of several of the UMPs in *Mth* SmAP1 allows them to be strung together into a hypothetical oligouridine that may mimic biologically relevant RNA binding in the Sm core of snRNPs. Failure of other NMPs to co-crystallize with *Mth* or *Pae* SmAP1 supports the specificity of uridine binding that we infer from the crystal structures.

In addition to the *Mth* UMP-binding site, several other ligand-binding sites exist in *Mth* and *Pae* SmAP1. Unlike the well-defined UMP site in *Mth*, the uridine-binding site in *Pae* is distant from the pore and not easily interpretable in electron density maps, suggesting low affinity binding at this site (Fig. 3.6). Also, the *Pae* SmAP1 residues in the region of this uridine are not very conserved. If all SmAPs specifically bind to an oligouridine site in RNA *in vivo*, then geometric considerations require any such RNA to bind near the 7-fold symmetry axis (i.e., the pore), and therefore the *Pae* binding pocket described here cannot be biologically relevant. Presumably, breaking of 7-fold symmetry in eukaryotic Sm heteroheptamers is reconcilable with RNA binding away from the pore (although there is no evidence for this). We note that the UMP-binding site in *Afu* and *Mth* exists in *Pae* SmAP1, and that UMP can be docked into this putative *Pae* binding site with only minimal changes required for side chain rotamers. We also found other sites in *Mth* and *Pae* occupied in each monomer by MPD, ethylene glycol, or glycerol. Though these ligands are clearly defined by electron density and many of the residues in
these binding sites are phylogenetically conserved, any biological significance of these additional sites is not yet clear.

The gel-shift activity of Mth and Pae SmAP1 on negatively supercoiled DNA substrates (Fig. 3.8) is especially interesting given the similarity of SmAP monomers to the OB fold. We found that SmAP1s non-specifically gel-shift a variety of supercoiled DNA substrates and that ssDNA oligonucleotides of >20-nt inhibit the gel-shift (Fig. 3.8(b)), possibly by direct binding to SmAP1. Since eukaryotic Sm proteins bind to ssRNA, and since SmAP homoheptamers probably do not function identically to Sm heteroheptamers, we propose that SmAPs may have a generic single-stranded nucleic acid binding activity (e.g., as a nucleic acid chaperone). The striking resemblance of the SmAP and OB folds corroborates this idea, given that several OB-fold proteins bind to ssDNA non-specifically. The following recently determined structures are all very similar (and in some cases nearly identical) to the Sm fold: the single-stranded DNA-binding domain of replication factor A, the S1 RNA-binding domain, the single-stranded telomeric DNA binding protein, and the Streptococcus pneumoniae SP14.3 protein (which is fused to a domain that is homologous to ribosomal protein S3).

**Emerging differences between SmAPs and canonical Sm proteins**

Eukaryotic Sm and Lsm proteins and their archaeal homologs, which we term Sm-like archaeal proteins, share a number of structural and functional features. Most significant is the similarity in Sm and SmAP 3D and quaternary structures: the monomers are nearly identical, and the SmAP homoheptamer parallels the Sm heteroheptamer that forms snRNP cores. Also, both sets of proteins apparently bind specifically to
oligouridine-containing RNA. However, several differences are emerging between SmAPs and the canonical Sm proteins. The results presented here show that SmAPs associate into many oligomeric states besides the standard heptamer (e.g., 14-mers and sub-heptamers), and can polymerize into homogeneous fibers. Such behavior is unreported for eukaryotic Sm proteins. No structural information is available for Sm proteins bound to RNA (or any other ligand), so it is difficult to evaluate the similarity of uridine binding by eukaryotic Sm proteins and SmAPs. Cross-linking experiments corroborate RNA binding near the pore in human Sm heptamers. The near identity of the Sm and SmAP dimer structures, as well as the strictly conserved mode of uridine recognition between *Afu* and *Mth* SmAP1, suggest that the SmAP1 UMP-binding site is an accurate model for RNA binding in the snRNP core. In this model, snRNA wraps around the circumference of the pore, but does not thread through it. Further elucidation of the similarities and differences between archaeal SmAP complexes and the Sm cores of eukaryotic snRNPs will be the aim of future experiments, and will provide insight into the structure and assembly of snRNPs.
References


12. Friesen, W. J., Paushkin, S., Wyce, A., Massenet, S., Pesiridis, G. S., Van Duyne, G., 
containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins. Mol 
Cell Biol 21(24), 8289-300.

Methylation of Sm proteins by a complex containing PRMT5 and the putative U 

Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B' and 
the Sm-like protein LSm4, and their interaction with the SMN protein. Rna 7(11), 
1531-42.


snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in 
a specific and thermodynamically stable manner. Mol Cell Biol 19(10), 6554-65.

pathway: identification of stable core protein heteromeric complexes and an snRNP 

18. Fromont-Racine, M., Mayes, A. E., Brunet-Simon, A., Rain, J. C., Colley, A., Dix, I., 
protein interaction screens reveal functional networks involving Sm-like proteins. 
Yeast 17(2), 95-110.

complex. Mol Cell 7(2), 319-29.


association of the U1-specific 70K and C proteins with U1 snRNPs is mediated in 


Figure and Table legends

Table 3.1: Statistics for several crystal forms. Crystallographic statistics are given for the Mth and Pae SmAP1 structures in different spacegroups (with various packing geometries) and with bound ligands (UMP, MPD, etc.). Data were collected either in-house (λ = 1.54 Å) or at the synchrotron (λ = 1.10 Å). Statistics for the highest resolution shell are given in [square brackets]. $R_{\text{cryst}} = \Sigma_{hkl} ||F_{\text{obs}}|| - ||F_{\text{calc}}|| / \Sigma_{hkl} |F_{\text{obs}}|$, and $R_{\text{free}}$ was computed identically, except that 5% of the reflections were omitted as a test set. Non-protein molecules were added based on the chemical composition of the crystallization condition and sufficiently strong $F_o - F_c$ density (>3σ).

Table 3.2: Pairwise RMSDs between Pae, Mth, and Afu SmAP1. RMSDs are shown for pairwise 3D alignments of Pae, Mth, and Afu SmAP1 monomers, dimers, and heptamers (using mainchain atoms only). The Afu-Mth pair superimposes best, while the Pae-Mth monomer structures are most dissimilar.

Figure 3.1: 3D structural alignment of Pae, Mth, and Afu SmAP1 dimers. A depth-cued stereoview is shown of the Cα trace for aligned Pae (red hues), Mth (blue hues), and Afu (green hues) SmAP1 dimers. N- and C-termini, as well as loops L2 and L4 are indicated. The greatest structural variation is in the positions of these two pore-forming loops, and the dimer interface is strictly conserved (asterisks). The large difference in the width of the heptameric pores in Pae (∼8-9 Å diameter) versus Afu and Mth (∼12-15 Å diameter) is due to two structural features: (i) side chain variation: the position of the positively-charged, pore-lining side chain R29 is extended into the pore in Pae, but K31 extends along the direction of the pore in Mth, and (ii) backbone variation: the distance
between identical Cα atoms in loop L2 of the seven monomers is greater in Mth than in Pae, i.e., the backbone protrudes further into the pore in the Pae heptamer.

**Figure 3.2: The oligomeric states of Pae and Mth SmAP1s in solution.**

Representative sedimentation results for analytical ultracentrifugation of wt Pae SmAP1 (a), the C8S mutant of Pae SmAP1 (b), and wt Mth SmAP1 (c) are shown. Data were collected at 20°C, at a rotor speed of 12,500 rpm, with absorbance measured at 280 nm. Protein concentrations were 0.69 mg/ml (a), 1.26 mg/ml (b), and 0.85 mg/ml (c). Weight-average molecular weights (given in kDa) were determined by fitting experimental data (circles) with a single exponential (solid line), and include roughly 2-3% error (residuals are in top panels); note that the protein samples are monodisperse. The molecular weight of the wt Pae protein suggests that it exists as a 14-mer, while the wt Mth data closely fits a heptamer. The molecular weight of the C8S mutant is significantly less than that of a heptamer, suggesting that it may exist in lower oligomerization states (4-, 5-, or 6-mers). Such heptamer "subcomplexes" have been detected for eukaryotic Sm proteins (see text for details).

**Figure 3.3: Polymerization of SmAP1s into polar fibers.** Transmission EMs are shown for wt Mth (a, b), wt Pae (c oxidized, d reduced), and the C8S mutant of Pae SmAP1 (e). The scale bar represents 10 nm for panel (c), and 50 nm for all other panels. The striated sheets formed by Mth SmAP1 (a, b) and non-disulfide bonded Pae SmAP1 (d, e) are extremely well ordered. The distance between the inner arrow tips in (b) corresponds to ≈ 8.3 nm (in agreement with the heptamer diameters from crystal structures), and suggests that the fiber axis is parallel to the heptameric 7-fold. The ≈ 50 nm distance between the
outer white arrows in (b) corresponds closely to six heptamer widths. Together with heptamer packings in various crystal forms, these EMs suggest that SmAP1 fibers form by head-to-tail stacking of heptamers (see Fig. 3.4). Doughnut-shaped SmAP1s are visible in the background of these EMs (most clearly for the wt Pae sample in panel (c)).

**Figure 3.4: Various crystalline oligomers of Pae and Mth SmAP1.** Panel (a) provides orthogonal views of the quasi-hexagonal packing of Mth SmAP1 heptamers in the $P2_12_12_1$ crystal form. Heptamers stack upon one another to form cylindrical tubes, thus providing a model for the structure of the EM fibrils (see text for explanation). The head-to-tail association of heptamers gives the tubes a defined polarity (colored arrows). Molecular surfaces show that the lateral packing of tubes in the crystal may generate the striated sheets seen by EM. A unit cell of the Pae SmAP1 $C222_1$ crystal form is shown in (b), along with examples of crystallographic 2-fold and $2_1$ screw axes. The asymmetric unit is a heptamer (shown as Cα traces in red or blue), and a Pae SmAP1 14-mer is formed from adjacent asymmetric units as shown in (c). Interaction surfaces and cysteines are illustrated. The 14-mer has 72-point group symmetry, with a 2-fold axis coinciding with a crystallographic 2-fold, and buries 7,547 Å$^2$ of surface area at the heptamer-heptamer interface.

**Figure 3.5: Ligand-binding sites in the structure of the Mth 14-mer bound to UMP.** Orthogonal views are shown for the two Mth heptamers (red, blue) in the asymmetric unit of the $P2_1$ form (a). A single molecule of MPD binds identically to each monomer, and is shown in space-filling (colored by atom type, yellow carbons). Space-filling models of the 14 UMP ligands show that they bind in the pore region (colored by atom type, gray
Electron densities for the UMP and MPD binding sites are shown in (b) and (c), respectively. The $2F_o - F_c$ density is contoured at $+1.2\sigma$ (green) and $F_o - F_c$ maps are contoured at $-3.2\sigma$ (red) or $+3.2\sigma$ (blue). Conserved residues that form these ligand-binding sites are labeled, and residues from different monomers are distinguished by primes. Hydrogen-bond distances are not shown in (b) for the sake of clarity (see Fig. 3.7 and the text for details of the SmAP1-ligand interactions).

**Figure 3.6: Ligand-binding sites in the structure of the Pae 14-mer bound to UMP.** Orthogonal views are shown in (a) for the Pae SmAP1 14-mer that is found in the $C222_1$ lattice (heptamer per asymmetric unit). Ten glycerol molecules bind to each heptamer (shown in space-filling, green-colored carbons), and seven of them occupy identical sites. The uridine fragments of UMP were modeled, and are shown in space-filling (gray-colored carbons). Electron density for one of the UMP-binding sites is shown in (b), contoured at $+1.2\sigma$ for $2F_o - F_c$ density (green) and at $+/–3.2\sigma$ for $F_o - F_c$ density (blue/red). While electron density for the uracil moiety is clearly defined, this ligand-binding site may not be a biologically relevant RNA-binding site (see text for discussion).

**Figure 3.7: Conserved mode of uridine recognition by Mth and Afu SmAP1.** Interactions between SmAP1 and uridine are diagrammed for Afu (a) and Mth (b). The remainder of the $U_3$ oligouridine from the Afu structure (indicated by a U2~~) has been omitted in (a) for the sake of clarity. Parenthesized letters after residue labels denote individual monomers. In both structures, the aromatic uracil base intercalates between a highly conserved pair of Arg/His side chains – e.g., the guanidinium of Arg72 and
imidazole of His46 for *Mth* SmAP1. Specific interactions and differences between *Afu* and *Mth* are discussed in the text. This figure was derived from a LIGPLOT\(^5\) output.

**Figure 3.8: Gel-shift of supercoiled DNA by *Mth* and *Pae* SmAP1.** The ability of *Mth* and *Pae* SmAP1 to shift the electrophoretic mobility of supercoiled plasmid DNA is shown in the agarose gels of (a) and (b), respectively. In (a), increasing concentrations of *Mth* SmAP1 were incubated with a negatively supercoiled plasmid (“p5L1c1”). The first onset of gel-shift is apparent at the lowest concentration of *Mth* (1.1 μM heptamer, arrow in lane 3), and saturates by the highest concentration (60 μM, lane 8). The ability of a 26-nucleotide ssDNA to inhibit the gel-shift induced by *Pae* SmAP1 is shown in (b). Lane 1 provides a DNA ladder, lanes 2 and 3 serve as negative controls, and the arrow in lane 4 shows the maximal gel-shift in the absence of ssDNA (which may inhibit the gel-shift by binding to SmAP1).
Table 3.1: Statistics for several crystal forms.

<table>
<thead>
<tr>
<th>Crystal form</th>
<th>$P1$ (Mth)</th>
<th>$P1$ (Mth with UMP)</th>
<th>$P2\times2_1$ (Mth)</th>
<th>$C2\times2_1$ (Pae with UMP)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X-ray wavelength (Å)</td>
<td>1.1000</td>
<td>1.1000</td>
<td>1.5418</td>
<td>1.5418</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>90.0 – 1.85</td>
<td>100.0 – 1.90</td>
<td>100.0 – 2.80</td>
<td>100.0 – 2.05</td>
</tr>
<tr>
<td>Compleness (%)</td>
<td>93.8 [67.4]</td>
<td>97.0 [92.4]</td>
<td>99.0 [92.5]</td>
<td>97.4 [95.5]</td>
</tr>
<tr>
<td>$I / \sigma(I)$</td>
<td>20.4 [4.1]</td>
<td>25.8 [2.3]</td>
<td>19.3 [3.3]</td>
<td>17.9 [4.1]</td>
</tr>
<tr>
<td>$R_{merge}$ (%)</td>
<td>5.4 [25.5]</td>
<td>4.8 [56.5]</td>
<td>11.2 [38.7]</td>
<td>11.5 [50.9]</td>
</tr>
<tr>
<td><strong>Molecular replacement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Search model</td>
<td>Homology model of Pae SmAP1 structure (1I8F)</td>
<td>Refined $P1$ Mth SmAP1 structure (1JBM)</td>
<td>Refined $P1$ Mth SmAP1 structure (1JBM)</td>
<td>Refined $C2$ Pae SmAP1 structure (1I8F)</td>
</tr>
<tr>
<td>Search method</td>
<td>Search w/ EPMR, then ARP/ wARP of polyalanine-ized version (to reduce model bias)</td>
<td>EPMR</td>
<td>EPMR</td>
<td>EPMR</td>
</tr>
<tr>
<td>Crystal packing</td>
<td>Heptamer per a.u.</td>
<td>Face-face tetradecamer per a.u. (pseudo-72 point group symmetry)</td>
<td>Edge-edge tetradecamer per a.u.</td>
<td>Heptamer per a.u.; face-face tetradecamer in crystal (72 point group symmetry)</td>
</tr>
<tr>
<td><strong>Model refinement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>20.0 – 1.85</td>
<td>20.0 – 1.90</td>
<td>15.0 – 2.80</td>
<td>20.0 – 2.05</td>
</tr>
<tr>
<td>&lt;B-factor&gt; (prot / water, Å²)</td>
<td>27.9 / 37.2</td>
<td>41.7 / 48.6</td>
<td>50.9 / 40.1</td>
<td>25.1 / 34.8</td>
</tr>
<tr>
<td>Number of solvent or ligand molecules included in model:</td>
<td>water 273</td>
<td>water 387</td>
<td>water 86</td>
<td>water 325</td>
</tr>
<tr>
<td></td>
<td>ethylene glycol 11</td>
<td>MPD 14</td>
<td>ethylene glycol 13</td>
<td>glycerol 10</td>
</tr>
<tr>
<td></td>
<td>acetate 5</td>
<td>UMP 14</td>
<td>chloride 3</td>
<td>acetate 2</td>
</tr>
<tr>
<td>RMSDs:</td>
<td>bonds (Å)</td>
<td>0.021</td>
<td>0.014</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>angles (°)</td>
<td>2.06</td>
<td>1.85</td>
<td>1.44</td>
</tr>
<tr>
<td>PDB code for submission</td>
<td>1JBM</td>
<td>1LOJ</td>
<td>1JRI</td>
<td>1LNX</td>
</tr>
<tr>
<td>$R_{cryst} / R_{free}$ (%)</td>
<td>19.6 / 23.8</td>
<td>20.7 / 25.0</td>
<td>19.9 / 29.0</td>
<td>18.2 / 22.6</td>
</tr>
</tbody>
</table>
Figure 3.1: 3D structural alignment of *Pae*, *Mth*, and *Afu* SmAP1 dimers.

Table 3.2: Pairwise RMSDs between *Pae*, *Mth*, and *Afu* SmAP1.

<table>
<thead>
<tr>
<th></th>
<th><em>Afu</em></th>
<th><em>Mth</em></th>
<th><em>Pae</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Afu</em></td>
<td></td>
<td>0.51 Å (monomer)</td>
<td>0.90 Å (monomer)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.61 Å (dimer)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.81 Å (heptamer)</td>
<td></td>
</tr>
<tr>
<td><em>Mth</em></td>
<td></td>
<td>1.02 Å (monomer)</td>
<td>1.19 Å (dimer)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pae</em></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 3.2: The oligomeric states of *Pae* and *Mth* SmAP1 in solution.

(a) wt Pae SmAP1   (b) C8S mutant of Pae SmAP1   (c) wt Mth SmAP1

MW=110000   MW=46700   MW=58000
Figure 3.3: Polymerization of SmAP1s into polar fibers.
Figure 3.4: Various crystalline oligomers of *Pae* and *Mth* SmAP1.
Figure 3.5: Ligand-binding sites in the structure of the *Mth* 14-mer bound to UMP.
Figure 3.6: Ligand-binding sites in the structure of the *Pae* 14-mer bound to UMP.
Figure 3.7: Conserved mode of uridine recognition by *Mth* and *Afu* SmAP1.

(a) *Afu* SmAP1

(b) *Mth* SmAP1

**Key**
- Ligand bond
- Non-ligand bond
- Hydrogen bond and its length
- Additional interactions found in *Mth*-UMP
- Non-ligand residues involved in hydrophobic contact(s)
- Corresponding atoms involved in hydrophobic contact(s)
Figure 3.8: Gel-shift of supercoiled DNA by *Mth* and *Pae* SmAP1.

(a) Lane 1: p5L1c1 +, Lane 2: wt Mth −, Lane 3: p9L2c1 −,
Lane 4: wt Pae −, Lane 5: ssDNA −

(b) Lane 1: p5L1c1 +, Lane 2: wt Mth −, Lane 3: p9L2c1 −,
Lane 4: wt Pae −, Lane 5: ssDNA −