The Euryarchaeota, Nature’s Medium for Engineering of Single-stranded DNA-binding Proteins*

Received for publication, November 15, 2004
Published, JBC Papers in Press, January 24, 2005, DOI 10.1074/jbc.M412870200

Justin B. Robbins,a,b,c Mary C. McKinney,a,b,d,e Claudia E. Guzman,a,f Binjon Sriratana,a,f
Sorel Fitz-Gibbon,a Taekjip Ha,a,b,h,i and Isaac K. O. Canna,b,i,j

From the aDepartment of Animal Sciences, the bDepartment of Physics, and the cInstitute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801 and the dCenter for Astrobiology, Institute of Geophysics and Planetary Physics, UCLA, Los Angeles, California 90095-1567

The architecture of single-stranded DNA-binding proteins, which play key roles in DNA metabolism, is based on different combinations of the oligonucleotide/oligosaccharide binding (OB) fold. Whereas the polypeptide serving this function in bacteria contains one OB fold, the eukaryotic functional homolog comprises a complex of three proteins, each harboring at least one OB fold. Here we show that unlike these groups of organisms, the Euryarchaeota has exploited the potential in the OB fold to re-invent single-stranded DNA-binding proteins many times. However, the most common form is a protein with two OB folds and one zinc finger domain. We created several deletion mutants of this protein based on its conserved motifs, and from these structures functional chimeras were synthesized, supporting the hypothesis that gene duplication and recombination could lead to novel functional forms of single-stranded DNA-binding proteins. Biophysical studies showed that the orthologs of the two OB fold/one zinc finger replication protein A in Methanosarcina acetivorans and Methanopyrus kandleri exhibit two binding modes, wrapping and stretching of DNA. However, the ortholog in Ferroplasma acidarmanus possessed only the stretching mode. Most interestingly, a second single-stranded DNA-binding protein, FaeRPA2, in this archaeon exhibited the wrapping mode. Domain analysis of this protein, which contains a single OB fold, showed that its architecture is similar to the functional homologs thought to be unique to the Crenarchaeotes. Most unexpectedly, genes coding for similar proteins were found in the genomes of eukaryotes, including humans. Although the diversity shown by archaeal single-stranded DNA-binding proteins is unparalleled, the presence of their simplest form in many organisms across all domains of life is of greater evolutionary consequence.

Single-stranded DNA (ssDNA)1-binding proteins are ubiquitous in Archaea (1), Bacteria (2), and Eukarya (3), which constitute the three domains of life (4). The bacterial and eukaryotic proteins are commonly designated single-stranded DNA-binding proteins (SSBs) and replication protein A (RPA), respectively, and due to the high similarity between the archaeal and eukaryotic proteins, the archaeal counterparts are also termed RPAs (5). Aside from the major domains of life, ssDNA-binding proteins have been isolated from viruses, such as adenovirus (6), and bacteriophages (7). Indeed, the first example of this important protein family was described in T4 phage (7). The widespread distribution of ssDNA-binding proteins in nature is due to their key roles in DNA transactions, especially in replication (8), recombination (9), and repair (10, 11). These proteins bind to stabilize and protect transiently exposed ssDNAs from nucleases during DNA replication and repair (3). In addition, they prevent formation of secondary structures in exposed ssDNA, and several members of this protein family are reported to stimulate their cognate DNA polymerases during DNA synthesis (1, 12, 13), as well as promote strand exchange in homologous recombination (14–16).

Structural analyses have shown that SSBs (17, 18) and RPAs (19, 20) contain ssDNA binding domains with a fold reminiscent of an OB (oligonucleotide/oligosaccharide binding) fold, a five-stranded β-sheet coiled to form a closed β-barrel (21). In ssDNA-binding proteins, the OB folds typically comprise ~100 amino acids (21), and although the majority of bacterial SSB genes code for products that contain a single OB fold, their active forms are homotetramers (22–24). Most interestingly, in Deinococcus radiodurans (25) and Thermus aquaticus (26), genes encoding SSBs with two OB folds have been described recently. The D. radiodurans SSB, however, dimerizes to mimic its homotetrameric counterparts (27).

In contrast to bacterial SSBs, eukaryotic RPA is a heterotrimERIC complex composed of proteins of ~70, 30, and 14 kDa (3), whereas in the archaea, the organization of RPAs is more complex, as demonstrated in this study. Currently, culturable archaea are grouped into two main subdomains, namely Crenarchaeota and Euryarchaeota (4). Whereas to date the crenarchaeotes have been shown to harbor a single gene encoding a single OB fold RPA, the euryarchaeotes on the other hand display an unusual diversity of these proteins. An interesting aspect of the crenarchaeal RPA is that it has an eukaryotic/
archaeal type OB fold; however, the domain organization is very similar to bacterial SSBs (28, 29). The striking homologies among the individual OB folds of the eukaryotic RPAs and the significant structural similarity they share with the ssDNA binding domain of bacterial SSBs have led to the postulate that RPAs evolved from gene duplication events in an ancestral SSB gene (30, 31). Here we show that no group of organisms has exploited the potential in the ssDNA-binding module for engineering of novel RPAs as the archaeal subdomain Euryarchaeota. Members of this group, which includes hyperthermophiles, mesophiles, halophiles, sulfate reducers, acidophiles, and methanogens, have re-invented RPAs many times. However, the most widespread RPA in these ecologically different organisms is a protein with two OB folds and a zinc finger domain. To further our understanding of this novel RPA, we used deletion, biochemical, and biophysical analyses to investigate the structure and function of naturally occurring candidates and their artificially synthesized counterparts. In addition, we showed that an RPA type that was originally thought to be exclusive to crenarchaeotes is indeed found in some eur- yarchaeotes and in many eukaryotes, including humans, plants, and a nematode.

**EXPERIMENTAL PROCEDURES**

Cloning of macrpa3, mkarpa, facrpa1, and facrpa2 Genes—The cloning of macrpa3 was described previously (1). A single gene (GenBank accession number NP_614724.1), encoding the only RPA-like protein to be exclusive to crenarchaeotes is indeed found in some eur- yarchaeotes, phyla, mesophiles, halophiles, sulfate reducers, acidophiles, and methanogens, have re-invented RPAs many times. However, the most widespread RPA in these ecologically different organisms is a protein with two OB folds and a zinc finger domain. To further our understanding of this novel RPA, we used deletion, biochemical, and biophysical analyses to investigate the structure and function of naturally occurring candidates and their artificially synthesized counterparts. In addition, we showed that an RPA type that was originally thought to be exclusive to crenarchaeotes is indeed found in some euryarchaeotes and in many eukaryotes, including humans, plants, and a nematode.

**EXPERIMENTAL PROCEDURES**

Cloning of macrpa3, mkarpa, facrpa1, and facrpa2 Genes—The cloning of macrpa3 was described previously (1). A single gene (GenBank accession number NP_614724.1), encoding the only RPA-like protein to be exclusive to crenarchaeotes is indeed found in some euryarchaeotes, phyla, mesophiles, halophiles, sulfate reducers, acidophiles, and methanogens, have re-invented RPAs many times. However, the most widespread RPA in these ecologically different organisms is a protein with two OB folds and a zinc finger domain. To further our understanding of this novel RPA, we used deletion, biochemical, and biophysical analyses to investigate the structure and function of naturally occurring candidates and their artificially synthesized counterparts. In addition, we showed that an RPA type that was originally thought to be exclusive to crenarchaeotes is indeed found in some euryarchaeotes and in many eukaryotes, including humans, plants, and a nematode.
dialyzed against buffer B (50 mM Tris-HCl, pH 8.0; 50 mM NaCl), and the dialysate was applied to an anion exchange column (HiTrap Q, 5 ml; Amersham Biosciences) fitted to a high pressure liquid chromatography apparatus (AKTA Explorer 10; Amersham Biosciences). After washing with 4 column volumes of buffer B, the bound proteins were eluted with a linear gradient (0–1.0M) of buffer C (50 mM Tris-HCl, pH 8.0; 1.0M NaCl) at a flow rate of 1 ml/min, and fractions of 0.5 ml in volume were collected. Aliquots of fractions were analyzed by SDS-PAGE to examine their purity. Highly purified proteins were dialyzed against buffer B supplemented with dithiothreitol (DTT) at 0.5 mM and stored at 4 °C. Recombinant *M. acetivorans* PolBI and *M. kandleri* PolBI were produced as described for the RPA proteins. However, because the plasmid used in constructing these two vectors contained the kanamycin resistance gene instead of ampicillin, all cultures were supplemented with kanamycin at 30 μg/ml and chloramphenicol at 50 μg/ml.

**Gel Filtration Analysis**—To determine the subunit organization of the purified full-length RPA proteins and the truncated derivatives, each protein was subjected to gel filtration analysis on a Superdex 200 HR 10/30 column (Amersham Biosciences). The column was fitted to the HPLC apparatus described above, and the resin was pre-equilibrated with a buffer made up of 50 mM sodium phosphate (pH 7.0) and 150 mM NaCl. Each protein was dialyzed against this buffer and injected at 100 μg. The chromatography was developed with the same buffer at a flow rate of 0.5 ml/min, and fractions of volume 0.2 ml were collected. To
calibrate the column, a set of protein standards (ferritin, 440 kDa; catalase, 232 kDa; aldolase, 158 kDa; albumin, 67 kDa; ovalbumin, 43 kDa; and ribonuclease A, 13.7 kDa; Amersham Biosciences) was analyzed with the same column under the same conditions. The fractions were examined by SDS-PAGE.

**Gel Retardation Assay**—The full-length RPAs from *M. kandleri* (MkaRPA) and *F. acidarmanus* (FacRPA1 and FacRPA2) and the chimeric proteins were used in gel retardation assays to determine whether these proteins can discriminate between ssDNA and double-stranded DNA (dsDNA). For the experiments with MkaRPA and FacRPA1, 5 pmol of protein was incubated with 1 pmol of a ϕX174-RF-end-labeled 42-mer oligonucleotide (Table I, gel retardation assay) in a buffer containing 20 mM Tris-HCl, pH 8.0, 15 mM MgCl₂, 2 mM DTT, and 0.05 mg/ml bovine serum albumin. The final volume of each reaction was 20 μl, and the FacRPA1 reaction was incubated at 25 °C, whereas that of the hyperthermophilic MkaRPA was incubated at 65 °C. The same buffer and temperature conditions were used for the gel retardation assay for FacRPA2, except for the amount of protein (7.5 pmol) added to the reaction mixture. For the chimeric proteins, except for RPA2-N/RPA3-C (5 pmol/reaction), a higher amount of protein (10 pmol/reaction) was necessary to clearly visualize the retarded products.

**Thermostability of MkaRPA**—The MkaRPA comes from an organism that grows at temperatures above 100 °C. Therefore, we tested its thermostability by heating at 85, 95, and 100 °C for 30 min followed by centrifugation to pellet any denatured protein. The protein remaining in solution was then tested for its ability to bind to ssDNA.

**Fluorescence Polarization Anisotropy (FPA)**—To obtain a measure of the binding of the various RPAs and the truncated derivatives to ssDNA, we conducted FPA experiments. FPA measures the rotational freedom of a fluorophore attached to a biomolecule by determining the intensity of the polarized fluorescence as the polarized excitation changed (1, 32, 33). If the dye rotates very quickly within its fluorescent lifetime, there will be little correlation between the excitation and emission polarization, resulting in an anisotropy value close to 0. However, if the dye is very rigidly attached, the emission polarization will be similar to the excitation polarization resulting in an anisotropy value of 0.4. The FPA measurements were carried out with an HPLC-purified 18-mer ssDNA tagged with fluorescein at the 5′ end (Table I, FPA) in a reaction buffer containing 20 mM Tris, pH 8.0, 15 mM MgCl₂, and 2 mM DTT. The reactions were performed at 23 ± 1 °C in a fluorometer (Cary Eclipse, Varian Inc.) by exciting the fluorescent dye attached to a biomolecule (1, 34–36). The energy transfer efficiency can be defined as $E = 1 - (1 + (R_0/R_0))$, where $R_0$ is the distance between a donor and an acceptor fluorophore, and $R_0$ is a characteristic constant depending on the particular dyes used (37). The dyes used in our experiments were attached to a DNA construct consisting of 18 bp of dsDNA followed by a 3′ single-stranded DNA tail of thymines of varying length known as Tₘ, where $N$ is the number of thymines in the tail. The donor dye, Cy3, was attached at the 3′ end of the tail, and the acceptor dye, Cy5, was attached at the 5′ end of the complementary strand at the junction of the dsDNA and ssDNA (Fig. 2). Thus, the FRET measured will be dependent only on the end-to-end distance of the ssDNA tail. By measuring the fluorescence intensity of both the donor and acceptor, one can approximate $E$ by $I_0/I_0 + I_0$, where $I_0$ is the acceptor emission intensity, and $I_0$ is the donor emission intensity, and we thereby infer the relative distance between the two dyes. Thymine was chosen to minimize DNA secondary structures that may interfere with the FRET measurements (38). The donor (Cy3)-labeled strand was annealed to the acceptor (Cy5)-labeled strand (Fig. 2A) in an annealing buffer of 400 mM NaCl, 20 mM Tris, pH 8.0, at a 1:1.4 ratio to a final concentration of 10–20 μM. Fluorescence measurement were performed in a fluorometer with reaction buffer as described under “Fluorescence Polarization Anisotropy.” Cy3 was excited at 532 nm, and fluorescence emission was collected at 585 nm (Cy3) and 665 nm (Cy5). Titration of the proteins were carried out until FRET ceased to change.

**Prime Extension Analysis**—The effect of the various RPA proteins on the primer extension capacity of the family B DNA polymerase (MacPolBI) found in *M. acetivorans* was investigated as described previously (1). Essentially, an oligonucleotide that is complementary to positions 6205–6234 of the M13mp18 genome (39) was end-labeled with [γ-³²P]ATP and T4 polynucleotide kinase (Roche Applied Science). To prepare the substrate, 1 pmol of the ³²P-labeled primer was annealed to 1.0 μg of M13mp18 ssDNA by heating in DNA polymerase buffer (20 mM Tris-HCl, pH 8.8; 100 mM NaCl, 5 mM MgCl₂, and 2 mM β-mercaptoethanol) to 95 °C for 5 min followed by gentle cooling to room temperature. The primer extension reaction was initiated by adding 250 μM of each dNTP followed by 0.5 μg of MacPolBI. The effect of individual RPAs or truncated derivatives on primer extension was investigated by adding each protein to the reaction mixture at 1.5 pmol/μl (1). The primer extension reaction was carried out at 25 °C for 30 min and terminated with 6 μl of stop solution (98% formamide, 1 μM EDTA, 0.1% xylene cyanol, 0.1% bromphenol blue). The products were then analyzed on 1% alkali agarose gel as described previously (1). In experiments that investigated the effect of RPAs on the hyperthermophilic MacPolBI, the primer extension reaction was carried out at 65 °C.

**Protein Concentration**—A commercial kit based on the Bradford method (Bio-Rad) was used in the estimation of protein concentrations, and bovine serum albumin (New England Biolabs) was used to derive the standard curve.

**Phylogenetic Analysis**—Protein sequences were gathered from the NCBI (www.ncbi.nlm.nih.gov), split manually into OB fold domains using BLAST (40), aligned using ClustalW (41), and manually edited using Se-Al (42). 3D-PSSM (43) was used to confirm and align remote homologs via secondary structure prediction, and PAUP (44) was used for phylogenetic analysis.

**RESULTS**

A Novel Group of RPA Conserved in Different Lineages of the Euryarchaeota—We recently discovered three functional RPA proteins in the mesophilic archaeon *M. acetivorans* (1), although...
purification. MacRPA3 was purified as described previously (England Biolabs); chromatography, the protein was loaded onto an anion exchange column (HiTrap Q), and a more purified protein was obtained in the flow-through fraction (Fig. 3A, lane 3). Because the MkaRPA is from a hyperthermophile, the E. coli cell extract containing the recombinant protein was first heat-treated at 85 °C for 10 min and centrifuged to remove the denatured host proteins. The supernatant was then applied to the affinity column. A highly purified protein, as judged by SDS-PAGE (Fig. 3, lane 5), was obtained by elution with 150 mM imidazole. FacRPA2 was purified by a single step through immobilization on the affinity column (Fig. 3A, lane 4). The estimated molecular masses of FacRPA1, FacRPA2, and MkaRPA were 41.9, 20.2, and 50.0 kDa, respectively, and these corresponded well with their molecular masses as determined by SDS-PAGE (Fig. 3A, lanes 3–5, respectively). We estimated previously that MacRPA3 (Fig. 3A, lane 2), a protein with a similar domain organization to those of FacRPA1 and MkaRPA, existed as a homodimer in solution (1). Therefore, to investigate whether this subunit organization was common to FacRPA1 and MkaRPA, each protein was subjected to gel filtration analysis individually. FacRPA2 alone and in a mixture with FacRPA1 were analyzed to determine whether the products of the two F. acidarmanus RPA genes formed a complex. The elution volume of FacRPA1 and FacRPA2 (Fig. 4A, FacRPA1 and FacRPA2) suggested a relative molecular mass of 73.1 ± 0.5 kDa for the former and 23.6 ± 0.3 kDa for the latter. These values suggested that FacRPA1 exists as a homodimer in solution and FacRPA2 as a monomer. The FacRPA1/FacRPA2 mixture eluted as two peaks (data not shown) representing the relative molecular masses of the individual components. It is thus likely that the two RPA proteins in F. acidarmanus do not interact in vivo. Most surprisingly, unlike the other two OB fold/zinc finger RPAs analyzed so far, the MkaRPA eluted as a very large protein. The estimated mass of 154.4 ± 1.4 kDa suggested that MkaRPA exists as a homotrimer in solution, which is very different from its homologs described to date (Fig. 4A, MkaRPA).
To further understand the structure and function of the two OB fold RPAs, we used a PCR method to create Mac/H9004 N57, a MacRPA3 with an N-terminal truncation, and also two C-terminally truncated derivatives (Fig. 1B, Mac/H9004 C1 and Mac/H9004 C2). In addition, Mka/H9004 C1, a C-terminally truncated MkaRPA, was made. The truncations are described under “Experimental Procedures,” and for ease of purification, each truncated protein was also produced with an N-terminal His6 tag.

In the gel filtration analysis of Mac/H9004 N57, the protein eluted as a single peak of relative molecular mass 104.0 ± 0.1 kDa (data not shown). Because the estimated molecular mass of a monomer of Mac/N57 is 42.8 kDa, the result suggested either a homodimeric or homotrimeric protein in solution. Furthermore, MacC1, which lacked the zinc finger domain and its downstream sequences, also eluted as a single peak and was estimated to be 47.5 ± 0.3 kDa. The predicted molecular mass of a monomer of this protein is 32.6 kDa, and thus this analysis also suggested a size representing either a monomer or homodimer in solution (data not shown). Deletion of the second OB fold and its downstream sequence, thus leaving a protein with a single OB fold (MacC2), resulted in a suggested size of 29.1 ± 0.9 kDa by gel filtration (data not shown). Because the predicted size of MacC2 is 20.2 kDa, the result suggested either a monomer or a dimer in solution. From the foregoing

---

**Fig. 4. Gel filtration analysis of RPA proteins.** The native molecular masses of the RPA proteins were estimated by gel filtration on a Superdex 200 HR 10/30 column. The arrows indicate the elution positions of protein markers run as standards to calibrate the column. Arrows 1–6 represent ferritin (440 kDa), catalase (232 kDa), aldolase (158 kDa), albumin (67 kDa), ovalbumin (43 kDa), and ribonuclease A (13.7 kDa), respectively. The corresponding RPA proteins and their peak elution volumes are indicated in each panel. The absorbance at 280 nm is in milliabsorbance units. A, full-length polypeptides. B, artificially synthesized proteins.
observations, the effects of deleting the first 57 amino acids or the zinc finger domain and its downstream sequence or the last OB fold together with its downstream sequence on oligomerization of MacRPA3 were not clearly evident. In an attempt to gain further insight into the contribution of the zinc finger domain and its downstream sequence to protein oligomerization, an MkaC1 was made. The estimated molecular mass of this protein was 34.4 kDa, and its gel filtration analysis resulted in two peaks that were not concentration-dependent. The elution volumes of the two peaks suggested relative molecular masses of 106.8 and 47.5 kDa, which suggested a homotrimer for the first peak and a monomer for the second peak.

Chimeric MacRPAs—In a previous study (1), we reported that MacRPA2, a protein whose two OB folds are about 60% identical, bound to ssDNA with less affinity than MacRPA3, whose two OB folds were only 25% identical (Fig. 1B). To investigate whether the level of similarity between the two OB folds in this type of RPA influences the affinity for ssDNA, we created three chimeric RPAs. The chimeric RPAs were also to help us understand whether shuffling of different OB folds will also result in functional ssDNA-binding proteins. The OB folds of two of the chimeric RPAs were selected so that they would share 20% identity, whereas the OB folds in the third chimeras were >50% identical. Thus the chimeras with dissimilar (±20%) identity between their OB folds mimicked MacRPA3, whereas the one exhibiting >50% identity mimicked MacRPA2. All of the proteins were produced as highly soluble proteins (Fig. 3B), and each was subjected to gel filtration analysis. The elution volume of RPA2-N/RPA3-C and RPA3-N/RPA2-C suggested a mass of 84.3 ± 0.8 and 75.4 ± 0.3 kDa, respectively. Because the estimated molecular masses of RPA2-N/RPA3-C and RPA3-N/RPA2-C were 53.2 and 47.3 kDa, respectively, each of these two chimeric proteins with dissimilar OB folds appeared to form homodimers in solution (Fig. 4B, RPA2-N/RPA3-C and RPA3-N/RPA2-C) as was observed with the parental proteins (1). Most surprisingly, RPA3/RPA1-OB, which was a chimera made by replacing the two OB folds of MacRPA3 with the two highly similar OB folds (59% identity) of MacRPA1 (Fig. 1A, and B), exhibited a different subunit organization in solution. Three different peaks corresponding to different subunit organizations were detected by gel filtration analysis (Fig. 4B). The difference between the subunit organization of this chimera and those of its counterparts that have dissimilar OB folds made it unfeasible to compare their binding characteristics to determine whether the similarities in OB folds influence affinity for ssDNA.

Determination of DNA Binding Properties by Gel Retardation Assay—The recombinant FacRPA1, FacRPA2, and MkaRPA were tested for their ability to discriminate ssDNA from dsDNA by using a gel retardation assay. Each protein bound to ssDNA (Fig. 5A–C, lane 2), and adding 10 times (Fig. 5A–C, lane 3) or 50 times (Fig. 5A–C, lane 4) excess cold ssDNA resulted in the labeled ssDNA being outcompeted by the cold
ssDNA. Next, the labeled ssDNA was incubated with RPA proteins in the presence of excess cold dsDNA. In all cases, the ssDNA was not outcompeted by dsDNA added at 10 times the concentration of the labeled ssDNA (Fig. 5, A–C, lane 5). In contrast, binding to labeled ssDNA by MkaRPA was completely outcompeted in the presence of 50 times cold dsDNA, and under the same conditions FacRPA1 and FacRPA2 still bound to the labeled ssDNA (Fig. 5, A–C, lane 6). Note, however, that a low amount of labeled ssDNA remained free in the case of FacRPA1. In addition, the temperature (65 °C) at which the MkaRPA was tested was far below the 80–110 °C at which this archaeon grows. The results, however, showed that all three proteins preferentially bind to ssDNA. The three chimeric RPAs were also tested for their ability to bind preferentially to ssDNA, and except for RPA3-N/RPA2-C, which showed a tendency to bind less of the labeled ssDNA in the presence of cold dsDNA, all three chimeras clearly showed a preference for ssDNA (Fig. 5, D–F, lanes 5 and 6). RPA2-N/RPA3-C exhibited two binding steps that may represent a different number of molecules bound to the ssDNA (Fig. 5E). A similar observation was made for RPA3/RPA1-OB. This protein exists in multiple oligomerization states, and this may explain the different gel-retarded products.

**Determination of ssDNA Binding by Fluorescence Polarization Anisotropy—**FPA is a widely used method to measure the binding or mobility of molecules that are tagged with a fluorescent dye (1, 32). The anisotropy of the ssDNA alone is a rather low 0.1, but as protein is added to the solution and it binds to the DNA, the anisotropy increases because of the larger size of the DNA-protein complex (Fig. 6). The ssDNA used in these experiments was 18 bases and was tagged with a fluorescein at the 5’ end (FL-18 in Table I). As more protein is added, more DNA is bound by the protein, and the anisotropy continues to increase until the DNA is completely saturated with protein. The FPA results for MacRPA3, MkaRPA, FacRPA1, and FacRPA2 are shown in Fig. 6A, and the results of the Hill model fits to these data are shown in Table II. As reported previously (1), MacRPA3 has a strong affinity for ssDNA having a $K_d$ value of $16.1 \pm 0.7$ nM. MkaRPA, however, binds even more readily with a $K_d$ value approximately half that of MacRPA3. However, neither of the *F. acidarmanus* RPAs bound with particularly strong affinity with $K_d$ values of $33.3 \pm 3.2$ nM for FacRPA1 and $81.9 \pm 18.0$ nM for FacRPA2. The large error associated with the FacRPA2 fit indicates that a more complicated binding process that may not fit the requirements of the Hill binding model may be displayed by this protein. Also revealed in the fit to the FPA data is the Hill coefficient, $n$, that indicates cooperativity of the protein binding to the ssDNA. MacRPA3, MkaRPA, and FacRPA1 all show positive cooperativity with Hill coefficients greater than 1, whereas FacRPA2 has a Hill coefficient of approximately 1, indicating each protein binds independently to ssDNA. A buffer containing 50 mM NaCl instead of MgCl$_2$ was also tested, and for most proteins the change made little difference on the $K_d$ and $n$ values (data not shown). However, the FacRPA2 binding was greatly affected by the change of salt from MgCl$_2$ to NaCl, and it did not bind at all when the NaCl concentration was increased, whereas the MacRPAs and MkaRPA could still bind ssDNA at high NaCl concentrations.

The ssDNA binding properties of the truncated RPA proteins were also tested with FPA (Fig. 6B). The small N-terminal truncation of MacRPA3 (MacΔN57) increased the affinity of the protein for ssDNA, and the Hill coefficient indicated that this protein still binds cooperatively to the ssDNA, but only with a coefficient of $-2$ instead of $-3$ for the native protein. In all cases, C-terminal truncations of the protein resulted in poorer binding of the protein as indicated by the $K_d$ values (Table II). The MacRPA3 that had been truncated to leave only one OB fold, MacΔC2, bound very weakly to the ssDNA, and increasing the concentration of the protein to 800 nM did not increase the anisotropy appreciably (data not shown).
The oligonucleotides under rpa genes were used for amplifying the genes encoding RPAs in M. acetivorans (Mac), M. kandleri (Mka), and F. acidarmanus (Fac). The oligonucleotides under chimeric genes were used in creating truncated derivatives of MacRPA3 and MkaRPA. The oligonucleotides under chimeric genes were used in creating chimeric MacRPA genes. The mkapolBI oligonucleotides were used to amplify MkaPolBI; the gel retardation assay (GRA) oligonucleotide was the probe for gel retardation assay; the FPA oligonucleotides were used for fluorescence polarization anisotropy experiments; the FRET oligonucleotides were annealed together to measure fluorescence resonance energy transfer as protein bound to the ssDNA region; and the primer extension oligonucleotide was annealed to the bacteriophage M13mp18 single-stranded DNA in the primer extension assay. Restriction sites are underlined.

### Table I

**Oligonucleotides used in this study**

The oligonucleotides under rpa genes were used for amplifying the genes encoding RPAs in M. acetivorans (Mac), M. kandleri (Mka), and F. acidarmanus (Fac). The oligonucleotides under chimeric genes were used in creating truncated derivatives of MacRPA3 and MkaRPA. The oligonucleotides under chimeric genes were used in creating chimeric MacRPA genes. The mkapolBI oligonucleotides were used to amplify MkaPolBI; the gel retardation assay (GRA) oligonucleotide was the probe for gel retardation assay; the FPA oligonucleotides were used for fluorescence polarization anisotropy experiments; the FRET oligonucleotides were annealed together to measure fluorescence resonance energy transfer as protein bound to the ssDNA region; and the primer extension oligonucleotide was annealed to the bacteriophage M13mp18 single-stranded DNA in the primer extension assay. Restriction sites are underlined.

| Experiment | Name | Nucleotide sequence | Restriction site |
|------------|------|---------------------|-----------------|
| **rpa genes** | | | |
| MacRPA2F | 5’-CATATGAAACAGATGTGAGAATCCATTAGAG | NdeI-tagged |
| MacRPA2R | 5’-TTCGAGTTACTCTCACTCAGGAAAAACCAAC |
| MacRPA3F | 5’-CATATGGAAGAAGAATACTCGCTGAGGAG |
| MacRPA3R | 5’-TTCGAGTTAAGCTAATACCTCCAGGAAAA |
| MkaRPAF | 5’-CATATGGGAACTGTTACATTACAC |
| MkaRPAR | 5’-TTCGAGTCATTCGGAAAGTTCGGAAAG |
| FacRPA1F | 5’-CATATGGCCAGCCACAGGAAAAACCATTA |
| FacRPA1R | 5’-TTCGAGTCATTTGTTATGTCATCCT |
| FacRPA2F | 5’-CATATGGAAGTCAATATTTGCAAGT |
| FacRPA2R | 5’-TTCGAGTTACTCTCTCCTGAGG |
| **Truncated genes** | | | |
| MacAN57F | 5’-CATATGGTAAAAAGGCTTCCCTCCCTGAGAAG |
| MacAC1R | 5’-TTCGAGTCAGGATGATAGTACATGTTACG |
| MacAC2R | 5’-TTCGAGTCATTGGCTAAGTATAGTACG |
| MacAC3R | 5’-TTCGAGTCATTCCAGGAAAAACCATTA |
| Chimeric genes | | | |
| RPA2-ChimR2 | 5’-GCCGGCGGAATGGGAATTGGAATTTCTCCTCAGGAAACGGTTTTTA |
| RPA3-ChimF2 | 5’-TAAAACACGTTCTCCTAGAGAAAGTTCCTCAATCTCCATTCCGGCC |
| RPA3-ChimR2 | 5’-CTACTCTCAAGGGTTGCGAAGATGATAGTACG |
| RPA3-ChimF2 | 5’-GAAGCTGTGTTAATGATAGTACG |
| Chim3-R2 | 5’-CGGGCTACGGGAAAAACACCTCCGGAAGATGATAGTACG |
| Chim3-F3 | 5’-GGGAAAAATTTCCGAGGTTTGTCGGTCAGGCAGG |
| mkapolBI gene | | | |
| MkaPolBF | 5’-CATATGGCTCGTGACATGGGGATGATAGTACG |
| MkaPolBR | 5’-GCCGGCGCTACGGGAAAAACCATTA |
| GRA | MacMC-R | 5’-TTTTCTCGAGGTATGGCAACAGGAATTTTATAGTACG |
| FPA | FL-18 | 5’-Fluorescein-GCCTCGCTGCCGTCGCCA |
| FRET | FL-34 | 5’-Fluorescein-AAATGGAACAGCTGCTGAGG |
| Primer extension | (M13 6205–6234) | 5’-ATTCTGTAATCTGATGGGTACAG |

### Table II

**Dissociation constants (K_d) and Hill coefficients (n) determined by FPA and estimated oligomerization states of RPA proteins**

The changes in anisotropy of an 18-base long oligonucleotide tagged with a fluorescent dye were recorded as the amount of the respective RPA protein was increased. The data were fitted to a Hill binding model to calculate the Hill coefficients and dissociation constants. Gel filtration chromatography was used to estimate the oligomerization state of each protein.

| Protein | K_d (nM) | Hill coefficient (n) | Oligomerization state |
|---------|----------|----------------------|----------------------|
| Full length | | | |
| MacRPA3 | 16.1 ± 0.7 | 3.2 ± 0.3 | Dimer |
| MkaRPA | 7.4 ± 0.4 | 1.6 ± 0.1 | Trimer |
| FacRPA1 | 33.3 ± 3.2 | 2.1 ± 0.3 | Dimer |
| FacRPA2 | 81.9 ± 18.0 | 0.9 ± 0.1 | Monomer |
| **Truncated derivative** | | | |
| MacAC1 | 6.8 ± 0.3 | 2.1 ± 0.2 | Trimer |
| MacAC2 | 21.1 ± 0.8 | 1.0 ± 0.1 | Dimer |
| MkaAC1 | 879.0 ± 624.0 | 1.1 ± 0.2 | Monomer |
| **Chimeras** | | | |
| RPA3-N/RPA2-C | 28.4 ± 10.2 | 1.1 ± 0.3 | Dimer |
| RPA2-N/RPA3-C | 5.3 ± 0.2 | 1.3 ± 0.1 | Dimer |
| RPA3/RPA1-OB | 11.6 ± 0.6 | 1.6 ± 0.1 | Dimer |

The chimeric RPAs were also tested by FPA for their ssDNA binding properties. Chimera RPA3-N/RPA2-C bound to ssDNA with a K_d value of 28.4 ± 10.2 nM. This indicated that this artificial RPA binds more weakly than the wild type MacRPA3; however, it is much better than the wild type MacRPA2, which did not fit the Hill model because it binds to ssDNA very weakly (1). Chimera RPA2-N/RPA3-C bound to ssDNA with a K_d of 5.3 ± 0.2, making it the protein with the highest affinity for ssDNA among all of the proteins tested. Thus, this chimera possessed better binding affinity than its parental RPAs (Mac-
RPA2 and MacRPA3) from which it was derived. The Hill coefficients of these two chimeric proteins were approximately 1, suggesting that in each case a dimer of the protein binds independently to the ssDNA. The last chimera, RPA3/RPA1-OB, was also tested for ssDNA binding properties. From the gel retardation assay (Fig. 5F) and the graph in Fig. 6C, this protein binds fairly well to ssDNA, albeit not as well as the native MacRPA3 or MacRPA1 (1). As suggested by gel filtration analysis, this chimera existed in three states (monomers, dimers, and trimers), but we chose to analyze the protein as a dimer to facilitate comparison to the other chimeras, and this resulted in a \( K_d \) of 11.6 \( \pm \) 0.6 and \( n \) of 1.6 \( \pm \) 0.1.

**Determination of ssDNA Binding by Fluorescence Resonance Energy Transfer**—To reveal more about the binding conformation of the RPAs to ssDNA, we used FRET. FRET efficiency was approximated by the ratio of the acceptor emission intensity to the total emission intensity. Fig. 7A shows the binding of recombiant MacRPA3, MkaRPA, FacRPA1, and FacRPA2 to a T\(_{40}\) DNA. In all FRET experiments, as in the FPA experiments, the protein concentration is plotted with respect to the oligomerization state of each particular protein. The gel retardation experiments showed that these proteins have a very low affinity to dsDNA as compared with ssDNA, so the 18 dsDNA bases in our FRET probe (Fig. 2) will not appreciably affect the experiment, and the length of DNA available for binding will be the length of the tail; here it is 40 bases. The initial FRET value for the T\(_{40}\) DNA in the experimental buffer is \(-0.45\). As protein is added, the FRET value will indicate how the protein binds to the ssDNA as well as the extent of binding. The FRET measurements for MacRPA3 and MkaRPA show interesting behavior as the protein concentration is increased. Initially, the FRET value increases (Fig. 7A), indicating the end of the ssDNA tail is being brought closer to the other end (Fig. 2, B and C). However, as the protein concentration is increased, the FRET value begins to decrease, indicating a stretching of the tail (Fig. 2D). Each of the *F. acidarmatus* RPAs behaved quite differently. The FacRPA1 always decreases FRET (stretches the ssDNA), whereas the FacRPA2 always increases FRET (compacts the ssDNA). At very high concentrations of protein, FacRPA2 can also start to slightly decrease FRET, but it requires over 25 times protein to DNA concentration for this to become apparent (data not shown).

The truncated RPAs were also tested with FRET to see if the binding behavior changes with deletions of certain sequence motifs/domains in the protein. The N-terminal deletion from MacRPA3 did not affect the two binding modes (high and low FRET). However, each of the C-terminal deletions removed the high FRET binding mode, and only the low FRET binding mode was seen (Fig. 7B). Contrasting this behavior, the C-terminal truncation from the MkaRPA did not remove the high FRET binding mode. In fact, the highest FRET value for this truncated MkaRPA occurred at a lower concentration of protein, indicating the stretching mode takes over more quickly than for the native RPA (Fig. 7B). FRET experiments were also performed on the chimeras. Each chimera only lowered FRET (Fig. 7C), and there was no indication of wrapping or compacting of the ssDNA with these RPAs.

In an attempt to further understand the two behaviors (stretching and compacting the ssDNA) of MacRPA3, the same titrations of the protein were performed on various lengths of ssDNA tails: T\(_{23}\), T\(_{40}\), and T\(_{70}\) (Fig. 7D). Although this protein is capable of binding to shorter ssDNAs (15 bases), the most noticeable changes in FRET occur at the intermediate lengths, 23 or 40 bases. In Fig. 7D, the FRET data is scaled so that it is forced to start at 0 and reach 1 at its peak, which is the concentration of protein needed to move to the second mode of binding. To clarify, scaled FRET = (FRET – initial value)/ (maxFRET – initial value). For T\(_{23}\), the maximum FRET value occurred at about 12 nm, with the concentration of ssDNA of 15 nm, and this was about one dimer per ssDNA. For a longer tail, 40 bases, the maximum FRET value occurred at about 22 nm, and for the longest tested, 70 bases, the maximum shifted even higher to about 41 nm.

**FacRPA Binding Properties**—To gain further insight into the binding properties of the FacRPAs, additional FPA and FRET measurements were performed by using both proteins compet-
FacRPA1 was added until saturation of the FRET value, and FacRPA2 was then added to compete for DNA binding. The experiment was then repeated with FL-34 in a FPA experiment. In the second group of experiments, FacRPA2 was added until saturation, and FacRPA1 was then added to compete for DNA binding. The experiment was then repeated with FL-34 in an FPA experiment.

**FIG. 8.** Competition between FacRPA1 and FacRPA2. A, FRET experiments were performed on T₄₀ DNA; FPA experiments were performed on FL-34 (Table I). Protein concentrations for FacRPA1 are as a dimer and FacRPA2 as a monomer. In the first experiment, FacRPA1 was added to T₄₀ until saturation of the FRET value, and FacRPA2 was then added to compete for DNA binding. The experiment was then repeated with FL-34 in an FPA experiment. B, in the second group of experiments, FacRPA2 was added until saturation, and FacRPA1 was then added to compete for DNA binding. The experiment was then repeated with FL-34 in an FPA experiment.

Results showing for the ssDNA. First, FacRPA1 was added to the solution with T₄₀ DNA and FRET decreased (Fig. 8A). After a saturating condition was achieved at ~200 nM FacRPA1, FacRPA2 was then added to the same sample in small increments. FRET increased with the addition of the protein almost back to the starting point of ssDNA after it had been saturated with FacRPA1. In contrast, only two times of FacRPA1 is required to bring FRET back to the starting point of ssDNA after it had been saturated with FacRPA2, and further addition can continue to decrease FRET. Because FacRPA1 has a lower $K_d$ value, it is less likely to dissociate from the ssDNA, hence requiring an overwhelming amount of FacRPA2 to make noticeable changes in FRET.

**Methanosarcina DNA Polymerase B1 Is Stimulated by Various Archaeal RPAs**—Human replication protein A (hRPA) has been shown to stimulate human pol α (12), and recently we showed that the primer extension capacity of MacPolB1, a pol α-like DNA polymerase in *M. acetivorans*, was also enhanced by each of three different RPAs found in this archaeon. The hRPA also stimulated human pol δ in the presence of other accessory factors, including proliferating cell nuclear antigen. However, unlike human pol α, this stimulatory effect on pol δ was nonspecific because substitution of hRPA with other ssDNA-binding proteins, such as *E. coli* SSB or T₄ gp32, yielded similar results (12). We were therefore interested in determining whether RPA proteins other than those native to *M. acetivorans* will stimulate primer extension by MacPolB1. Recombinant MacPolB1 alone synthesized a product of ~0.5 kb in size (Fig. 9A, lane 8), and as reported previously (1), addition of MacRPA3 to the reaction mixture stimulated primer extension, and full-length products (~7.2 kb) were synthesized (Fig. 9A, lane 9). The deletion of the C-terminal half of MacRPA3, leaving the two OB folds intact, did not abolish ssDNA binding, although this protein (MacAC1) exhibited a lower affinity for ssDNA (Table II). Most interestingly, MacAC1 also stimulated primer extension by MacPolB1, although a fairly visible amount of the product was observed at 0.5 kb (Fig. 9A, lane 10). The primer extension capacity of MacPolB1 was then tested in the presence of other archaean RPAs. As shown in Fig. 9A, lane 11, FacRPA1, a protein with similar domain architecture and subunit organization as MacRPA3, stimulated primer extension by MacPolB1. On the other hand, MkaRPA, a hyperthermophilic protein that also has a similar architecture as MacRPA3 and FacRPA1 but a different subunit organization, only slightly stimulated primer extension by MacPolB1, although a fairly visible amount of the product was observed at 0.5 kb (Fig. 9A, lane 12). The primer extension capacity of MacPolB1 was then tested in the presence of other archaean RPAs. As shown in Fig. 9A, lane 13, similar FRET results were obtained. At the same concentrations as those of MacRPA3 and FacRPA1, both MkaRPA and FacRPA2 failed to relieve the salient pause by the DNA polymerase at ~0.5 kb, and removing the zinc finger motif together with its downstream sequences from MkaRPA failed to rectify this condition (Fig. 9, lane 14). MkaRPA is a thermostable protein, and heating it at 100 °C for 30 min did not affect its ssDNA binding activity. The 25 °C at which the primer extension reaction was carried out was far below its temperature optimum. However, because a mesophilic DNA polymerase was being investigated, an appropriate temperature for MkaRPA could not be used. FacRPA2 binds ssDNA differentially from the two OB fold/zinc finger RPAs, and in addition it binds ssDNA weakly. These properties may explain why its effect was different from MacRPA3 and FacRPA1.

**Effect of MkaRPA on MkaPolB1**—The RPA of the thermophilic archaeon *Methanothermobacter thermoautotrophicus* (Mth) was shown to inhibit DNA synthesis by its cognate DNA polymerase MthPolB1 (45). From the above experiments, we did not detect an inhibitory effect of MkaRPA on *M. acetivorans* PolB1. However, the reaction was carried out at a temperature
Fig. 9. Effect of various RPAs on the primer extension capacity of MacPolBI and MkaPolBI. A, primer extension was compared in the presence and absence of individual RPA proteins (MacRPA3, MacΔC1, FacRPA1, FacRPA2, MkaRPA, and MkaΔC1). The reaction mixtures were incubated at 37 °C for 15 min, and the products were analyzed by 1% alkali agarose gel electrophoresis followed by visualization using autoradiography. B, primer extension by MkaPolBI was compared in the presence and absence of MkaRPA and MkaΔC1. The reaction mixtures were incubated at 65 °C for 15 min, and the products were analyzed by 1% alkali-agarose gel electrophoresis followed by visualization using autoradiography. The ΔC1 proteins lack their zinc finger domains and downstream amino acid sequences.
that was far below the 80—110 °C at which M. kandleri grows (46). To determine whether MkaRPA has an inhibitory effect on its cognate DNA polymerase, we produced recombinant MkaPolBI (Fig. 2, lane 7), and we investigated its primer extension capacity in the presence of increasing concentrations of MkaRPA. As observed for MthPolBI in the presence of MthRPA, primer extension by MkaPolBI was inhibited by MkaRPA in a concentration-dependent manner (Fig. 9, lanes 6 and 7). Most interestingly, the MkaC1, which lacks the zinc finger domain and its downstream sequence, did not exhibit this inhibitory property, suggesting that this region may influence the inhibitory property of MkaRPA.

Diversity of RPAs in Euryarchaeota—Unlike in the Crenarchaeota, Bacteria, and Eukarya, several RPAs of different domain organization have been described in the Euryarchaeota. In the hyperthermophilic methanogen Methanococcus jannaschii, a large RPA with four OB folds and a zinc finger domain in a single protein (Fig. 1A) has been biochemically characterized (5). In the thermophilic methanogen M. thermocautrophilus, the RPA has five OB folds and a zinc finger domain (Fig. 1A) (45), whereas in Pyrococcus furiosus the RPA comprises three different subunits (Fig. 1A) that form a heterotrimeric RPA as in eukaryotes (16). The zinc finger domain of P. furiosus RPA is located in the large subunit as observed in the eukaryotic RPAs. Recently, we described a very unique RPA in the Methanosarcinaceae (1). This large RPA has four OB folds. However, unlike all previously described euryarchaeal RPAs, MacRPA1 lacks a zinc finger motif. From the same organism, we also discovered and biochemically characterized two copies of a new form of RPA that is composed of two OB folds and a novel zinc finger motif. In the present work, we have shown that the two OB fold/one zinc finger motif RPA is more widely distributed than any other known archaeal RPA. Furthermore, we found orthologs of an RPA that has been designated the crenarchaeal RPA (28, 29) in the Thermoplasma and Ferroplasma, all members of the Euryarchaeota. Further searches in the publicly available databases showed that, although yet to be biochemically characterized, indeed genes coding for RPA-like proteins with similar architecture to the crenarchaeal RPA and FacRPA2 are also found in eukaryotes such as Homo sapiens, Arabidopsis thaliana, Oryza sativa, and Caenorhabditis elegans (Fig. 10). Thus, it is clear that the single OB fold RPA, originally thought to be unique to the crenarchaeotes, is more widespread in the archaeal/eukaryotic lineage.

DISCUSSION

In bacteria and eukaryotes, the proteins that function as ssDNA-binding proteins are highly conserved (2, 3), and in
The Euryarchaeae and Evolution of RPAs

archaea this observation is true for the crenarchaeal subdomain from which only a single type of RPA has been described (28, 29). In contrast, the Euryarchaeota, which represents the other cultivable archaeal subdomain, display an unusual diversity of RPAs (Fig. 1A), and some of the candidates have been suggested as representing evolutionary intermediates between bacterial SSB and eukaryotic RPAs (30). Indeed, the domain organization of the crenarchaeal RPA is very similar to that of bacterial SSB, in that the protein has a single ssDNA binding domain and a flexible C-terminal tail containing a number of acidic residues (28, 29). On the other hand, the RPA of the euryarchaeote *F. purisus* is a heterotrimeric protein complex that is reminiscent of eukaryotic RPAs (3). Despite the diverse RPAs found in the euryarchaeotes, our analysis as shown in Fig. 1A suggests that the most widespread of the euryarchaeal RPAs is a protein with two OB folds in the N-terminal half and a conserved zinc finger domain within the C-terminal half. In this report, we investigated several orthologs of this group of RPA, and in addition we analyzed some candidates for the contribution of their different domains to their structure and function.

The euryarchaeotes that harbor the two OB fold/one zinc finger RPA, except for *M. kandleri*, also contain other RPA-like sequences (Fig. 1A). In our previous report (1), we demonstrated that the different RPA proteins from *F. acetivorans* do not interact in vitro, and in the present work we further showed that the two RPA-like proteins from *F. acidarmanus* do not form a heterocomplex. Using both FRET and FPA analysis, we demonstrated that indeed in solution FacRPA1 and FacRPA2 tend to displace each other in a concentration-dependent manner (Fig. 8). To further our understanding of the two OB fold/one zinc finger RPAs, we made deletions in both FacRPA3 and MkaRPA to determine how different regions in the polypeptide contribute to oligomerization, ssDNA binding, and the ability to stimulate primer extension. By using gel filtration analysis, we determined that wild type MacRPA3 eluted with a relative molecular mass of 104.0 kDa (1), and in the present report deletion of the first 57 amino acids resulted in a protein with a relative molecular mass of 104.0 ± 0.1 kDa and thus suggested a trimeric protein. Therefore, the N-terminal region appears to be essential for the proper dimerization of MacRPA3. The deletion of the zinc finger and its downstream sequence in MacRPA3 and MkaRPA yielded different results. Although the oligomerization state of MacAC1 was not clear, MacAC1 resulted in a protein that existed in two oligomerization states. Unlike its mesophilic counterparts that were estimated to be homodimeric, the MkaRPA was estimated to be a homotrimeric protein (Fig. 4A). Despite the similarity in domain organization of the mesophilic and hyperthermophilic RPAs, the OB folds of MkaRPA are very different, and as shown in Fig. 1A, their phylogenetic positions could not be determined. Thus, it seems that although different OB folds in different RPAs may show some similarities in amino acid sequences, they have been optimized through the evolutionary process to achieve polypeptide folds or unique tertiary structures required for their function. This became very evident, as discussed below, in our construction of chimeric RPAs using different OB folds from RPAs exhibiting different oligomerization states. Despite these differences, MacAC1 and MkaAC1 bound to ssDNA with high affinity, implying that the two OB folds were enough for binding to ssDNA. Contrasting these results, MacAC2 was almost devoid of ssDNA binding activity (Table I).

The three proteins MacRPA3, MkaRPA, and FacRPA1 exhibited similar domain architectures. Each contains two OB folds and a zinc finger. However, they differed in binding properties when examined by FRET. MacRPA3 and MkaRPA both show two binding modes, one with high FRET and the other with low FRET. Most interestingly, the MkaRPA mutant lacking the zinc finger still showed both binding modes, whereas the MacRPA3 required the zinc finger for both binding modes. Because these experiments were all carried out on T4SS DNA and no changes were made to the OB folds, it seems that the zinc finger plays an important role in the stability of the high FRET state. We propose that the wrapping mode, which brings the ends of the ssDNA in close proximity, is the more favorable conformation when one protein (a dimer in the case of MacRPA3) is bound to the DNA and the zinc finger stabilizes this conformation. If the zinc finger is deleted, this binding mode is still able to exist but less stably, as shown as there is enough protein in solution, the stretching mode takes over. For MacAC1, the wrapping conformation is so unstable that only the stretching conformation is seen. We found no evidence that suggested that FacRPA1, which has the same domain architecture as MacRPA3 and MkaRPA, also exhibits a wrapping mode. However, this organism has a second RPA that wraps ssDNA. Thus, if a process in the cell would require an RPA that wrapped ssDNA, FacRPA2 could accomplish this function.

As reported in our previous work (1), the binding site size of MacRPA3 is ~20 bases. To investigate this further, various lengths of ssDNA tail were titrated with MacRPA3 to determine the effect of the wrapping versus stretching of the ssDNA on longer tail lengths. The FRET value initially increased for ssDNA of lengths ~20 bases or longer (data not shown); however, the protein concentration at which the FRET began to decrease was proportional to the tail length. For a 23-base tail, the peak value of FRET occurred at about a 1:1 ratio of dimer with ssDNA. For 40 bases there are about 1.5 dimers per ssDNA, and for 70 bases there are about 3 dimers per ssDNA. This increase in the concentration of protein at the maximum FRET value indicates that more proteins are binding to the longer tails in the wrapping mode before it starts to stretch the ssDNA. The schematic in Fig. 2 illustrates a possible binding process for MacRPA3. For shorter ssDNA, one RPA dimer would cause the ssDNA to wrap (Fig. 2B), but if there are enough RPAs present in solution, two dimers could bind with each dimer not completely wrapping the DNA, so the overall DNA geometry is more stretched than the free ssDNA. For longer ssDNA, multiple dimers can bind in the wrapping mode (Fig. 2C), and even more RPA would need to be present in the solution in order for the stretching mode to appear (Fig. 2D). It is interesting to note that with the truncation, MacAN57 still shows the same FRET pattern, and the peak is also at about a 1:1.5 ratio of DNA:dimer. This implies that the deletion of the N-terminal 57 amino acids, which seemed to alter the oligomerization of the protein, did not interfere with the wrapping/stretching behavior of the protein. It is known that *E. coli* SSB can bind to ssDNA in two different conformations as follows: the (SSB)$_{35}$ binding mode that occludes 35 bases of ssDNA per tetramer, and the (SSB)$_{65}$ mode that occludes 65 bases of ssDNA per tetramer (18). The presence of each mode is known to be dependent on the salt concentration and the SSB concentration. It was proposed that the different binding modes are functionally relevant in the cell by either obscuring or making available two of the four binding domains thus inhibiting or allowing the SSB to be directly transferred during lagging strand DNA replication (47). There may be a similar function for these archaeal RPAs. Thus, the two modes of binding may serve to regulate various functions in the cell that depend on the ssDNA to be in a particular conformation or the RPA to have certain sites available for interactions with other proteins or DNA. Most interestingly, the FacRPAs behaved differently from the other RPAs when examined by FRET. One RPA wraps...
the ssDNA with a lower affinity, whereas the other stretches the ssDNA but has a higher affinity. It may be that FacRPA1 and FacRPA2 are involved in different cellular processes that require different conformations of the ssDNA in order to recruit another replication, transcription, or recombination proteins. In contrast, in MkaRPA and MacRPA3, a single RPA protein has acquired both the wrapping and stretching functions.

Complex RPAs are hypothesized to have developed from gene duplications and recombination events (30). We reasoned that the two OB fold/one zinc finger RPAs could have evolved from a fusion of two genes coding for products such as a single OB fold protein and a protein with one OB fold and a zinc finger, and evidence for such genes is found in Fig. 1A (Afu and Hsp). To test this hypothesis, we made several chimeric RPAs that fused an anterior OB fold from one protein with a posterior OB fold and its zinc finger domain from another protein. The chimeric proteins were made to mimic RPAs with dissimilar OB folds and those with more similar OB folds. All chimeras bound to ssDNA specifically, thus supporting the hypothesis that different RPAs can evolve through fusions orchestrated by gene duplications and recombination events. The different subunit organization seen in the chimera RPA3/RPA1-OB (Fig. 4B) suggested that the OB folds in the two OB fold/zinc finger RPAs have been optimized in nature, and replacing them with OB folds from MacRPA1 (Fig. 1B), which belongs to a different class of RPAs, resulted in a different fold with a concomitant change in subunit organization.

The primer extension studies showed that other mesophilic orthologs (FacRPA1) of MacRPA3 can stimulate DNA synthesis by its mesophilic counterparts. As seen in other thermophilic RPAs, at 65 °C MkaRPA inhibited DNA synthesis by its mesophilic counterparts. We propose that fused an anterior OB fold from one protein with a posterior zinc finger, and evidence for such genes is found in Fig. 1A (Afu and Hsp). To test this hypothesis, we made several chimeric RPAs with dissimilar OB folds and those with more similar OB folds. All chimeras bound to ssDNA specifically, thus supporting the hypothesis that different RPAs can evolve through fusions orchestrated by gene duplications and recombination events. The different subunit organization seen in the chimera RPA3/RPA1-OB (Fig. 4B) suggested that the OB folds in the two OB fold/zinc finger RPAs have been optimized in nature, and replacing them with OB folds from MacRPA1 (Fig. 1B), which belongs to a different class of RPAs, resulted in a different fold with a concomitant change in subunit organization.

The primer extension studies showed that other mesophilic orthologs (FacRPA1) of MacRPA3 can stimulate DNA synthesis by its mesophilic counterparts. As seen in other thermophilic RPAs, at 65 °C MkaRPA inhibited DNA synthesis by its mesophilic counterparts. We propose that fused an anterior OB fold from one protein with a posterior zinc finger, and evidence for such genes is found in Fig. 1A (Afu and Hsp). To test this hypothesis, we made several chimeric RPAs with dissimilar OB folds and those with more similar OB folds. All chimeras bound to ssDNA specifically, thus supporting the hypothesis that different RPAs can evolve through fusions orchestrated by gene duplications and recombination events. The different subunit organization seen in the chimera RPA3/RPA1-OB (Fig. 4B) suggested that the OB folds in the two OB fold/zinc finger RPAs have been optimized in nature, and replacing them with OB folds from MacRPA1 (Fig. 1B), which belongs to a different class of RPAs, resulted in a different fold with a concomitant change in subunit organization.

The primer extension studies showed that other mesophilic orthologs (FacRPA1) of MacRPA3 can stimulate DNA synthesis by its mesophilic counterparts. As seen in other thermophilic RPAs, at 65 °C MkaRPA inhibited DNA synthesis by its mesophilic counterparts. We propose that fused an anterior OB fold from one protein with a posterior zinc finger, and evidence for such genes is found in Fig. 1A (Afu and Hsp). To test this hypothesis, we made several chimeric RPAs with dissimilar OB folds and those with more similar OB folds. All chimeras bound to ssDNA specifically, thus supporting the hypothesis that different RPAs can evolve through fusions orchestrated by gene duplications and recombination events. The different subunit organization seen in the chimera RPA3/RPA1-OB (Fig. 4B) suggested that the OB folds in the two OB fold/zinc finger RPAs have been optimized in nature, and replacing them with OB folds from MacRPA1 (Fig. 1B), which belongs to a different class of RPAs, resulted in a different fold with a concomitant change in subunit organization.

The primer extension studies showed that other mesophilic orthologs (FacRPA1) of MacRPA3 can stimulate DNA synthesis by its mesophilic counterparts. As seen in other thermophilic RPAs, at 65 °C MkaRPA inhibited DNA synthesis by its mesophilic counterparts. We propose that fused an anterior OB fold from one protein with a posterior zinc finger, and evidence for such genes is found in Fig. 1A (Afu and Hsp). To test this hypothesis, we made several chimeric RPAs with dissimilar OB folds and those with more similar OB folds. All chimeras bound to ssDNA specifically, thus supporting the hypothesis that different RPAs can evolve through fusions orchestrated by gene duplications and recombination events. The different subunit organization seen in the chimera RPA3/RPA1-OB (Fig. 4B) suggested that the OB folds in the two OB fold/zinc finger RPAs have been optimized in nature, and replacing them with OB folds from MacRPA1 (Fig. 1B), which belongs to a different class of RPAs, resulted in a different fold with a concomitant change in subunit organization.

The primer extension studies showed that other mesophilic orthologs (FacRPA1) of MacRPA3 can stimulate DNA synthesis by its mesophilic counterparts. As seen in other thermophilic RPAs, at 65 °C MkaRPA inhibited DNA synthesis by its mesophilic counterparts. We propose that fused an anterior OB fold from one protein with a posterior zinc finger, and evidence for such genes is found in Fig. 1A (Afu and Hsp). To test this hypothesis, we made several chimeric RPAs with dissimilar OB folds and those with more similar OB folds. All chimeras bound to ssDNA specifically, thus supporting the hypothesis that different RPAs can evolve through fusions orchestrated by gene duplications and recombination events. The different subunit organization seen in the chimera RPA3/RPA1-OB (Fig. 4B) suggested that the OB folds in the two OB fold/zinc finger RPAs have been optimized in nature, and replacing them with OB folds from MacRPA1 (Fig. 1B), which belongs to a different class of RPAs, resulted in a different fold with a concomitant change in subunit organization.

The primer extension studies showed that other mesophilic orthologs (FacRPA1) of MacRPA3 can stimulate DNA synthesis by its mesophilic counterparts. As seen in other thermophilic RPAs, at 65 °C MkaRPA inhibited DNA synthesis by its mesophilic counterparts. We propose that fused an anterior OB fold from one protein with a posterior zinc finger, and evidence for such genes is found in Fig. 1A (Afu and Hsp). To test this hypothesis, we made several chimeric RPAs with dissimilar OB folds and those with more similar OB folds. All chimeras bound to ssDNA specifically, thus supporting the hypothesis that different RPAs can evolve through fusions orchestrated by gene duplications and recombination events. The different subunit organization seen in the chimera RPA3/RPA1-OB (Fig. 4B) suggested that the OB folds in the two OB fold/zinc finger RPAs have been optimized in nature, and replacing them with OB folds from MacRPA1 (Fig. 1B), which belongs to a different class of RPAs, resulted in a different fold with a concomitant change in subunit organization.

The primer extension studies showed that other mesophilic orthologs (FacRPA1) of MacRPA3 can stimulate DNA synthesis by its mesophilic counterparts. As seen in other thermophilic RPAs, at 65 °C MkaRPA inhibited DNA synthesis by its mesophilic counterparts. We propose that fused an anterior OB fold from one protein with a posterior zinc finger, and evidence for such genes is found in Fig. 1A (Afu and Hsp). To test this hypothesis, we made several chimeric RPAs with dissimilar OB folds and those with more similar OB folds. All chimeras bound to ssDNA specifically, thus supporting the hypothesis that different RPAs can evolve through fusions orchestrated by gene duplications and recombination events. The different subunit organization seen in the chimera RPA3/RPA1-OB (Fig. 4B) suggested that the OB folds in the two OB fold/zinc finger RPAs have been optimized in nature, and replacing them with OB folds from MacRPA1 (Fig. 1B), which belongs to a different class of RPAs, resulted in a different fold with a concomitant change in subunit organization.

The primer extension studies showed that other mesophilic orthologs (FacRPA1) of MacRPA3 can stimulate DNA synthesis by its mesophilic counterparts. As seen in other thermophilic RPAs, at 65 °C MkaRPA inhibited DNA synthesis by its mesophilic counterparts. We propose that fused an anterior OB fold from one protein with a posterior zinc finger, and evidence for such genes is found in Fig. 1A (Afu and Hsp). To test this hypothesis, we made several chimeric RPAs with dissimilar OB folds and those with more similar OB folds. All chimeras bound to ssDNA specifically, thus supporting the hypothesis that different RPAs can evolve through fusions orchestrated by gene duplications and recombination events. The different subunit organization seen in the chimera RPA3/RPA1-OB (Fig. 4B) suggested that the OB folds in the two OB fold/zinc finger RPAs have been optimized in nature, and replacing them with OB folds from MacRPA1 (Fig. 1B), which belongs to a different class of RPAs, resulted in a different fold with a concomitant change in subunit organization.