doi:10.1006/jmbi.2001.5213 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2001) **314**, 1077–1085

JMB

Archaeal RadA Protein Binds DNA as Both Helical Filaments and Octameric Rings

Shixin Yang[1], Xiong Yu[1], Erica M. Seitz[2], Stephen C. Kowalczykowski[2] and Edward H. Egelman[1]*

[1]*Department of Biochemistry and Molecular Genetics University of Virginia Health Sciences Center, Charlottesville VA 22908-0733, USA*

[2]*Division of Biological Sciences Sections of Microbiology and of Molecular and Cellular Biology University of California, Davis CA 95616-8665, USA*

**Corresponding author*

The *Escherichia coli* RecA protein has been a model for understanding homologous eukaryotic recombination proteins such as Rad51. The active form of both RecA and Rad51 appear to be helical filaments polymerized on DNA, in which an unusual helical structure is induced in the DNA. Surprisingly, the human meiosis-specific homolog of RecA, Dmc1, has thus far only been observed to bind DNA as an octameric ring. Sequence analysis and biochemical studies have shown that archaeal RadA proteins are more closely related to Rad51 and Dmc1 than the bacterial RecA proteins. We find that the *Sulfolobus solfataricus* RadA protein binds DNA in the absence of nucleotide cofactor as an octameric ring and in the presence of ATP as a helical filament. Since it is likely that RadA is closely related to a common ancestral protein of both Rad51 and Dmc1, the two DNA-binding forms of RadA may provide insight into the divergence that has taken place between Rad51 and Dmc1.

Keywords: recombination; electron microscopy; image analysis; protein-DNA complexes; ring proteins

Introduction

Homologous genetic recombination is the process wherein two similar DNA molecules exchange stretches of sequence. This process is important to the generation of genetic diversity, but is also an important mechanism for the repair of DNA and the maintenance of genome stability. Much of our detailed knowledge about protein-mediated recombination has come from studies of the *Escherichia coli* RecA protein.[1] In response to massive DNA damage, RecA can be one of the most abundant proteins present in the cell. *In vitro*, RecA can catalyze a strand exchange reaction between homologous DNA molecules.[2,3] The active form of RecA is believed to be an unsual nucleoprotein filament, in which the DNA is stretched by ~50%[4] and untwisted to ~18.6 base-pairs per turn,[5] resulting in a change of the pitch of DNA from ~ 36 Å in *B*-form to ~ 95 Å in the complex with RecA.

The eukaryotic Rad51 proteins are homologs of RecA. While bacterial *recA* and yeast *RAD51* are not essential genes (although loss of *recA* function results in a large percentage of inviable cells), it has been shown that *RAD51* knockouts are lethal in both chicken and mammalian cell lines.[6–8] The yeast[9] and human[10,11] Rad51 proteins form helical filaments on DNA that are very similar in helical parameters to the RecA filaments. Dmc1 is another eukaryotic RecA homolog but, unlike Rad51, Dmc1 expression is meiosis-specific.[12] It was expected that Dmc1 would also bind DNA as a helical filament, but thus far has only been observed *in vitro* to bind DNA as octameric rings.[13,14]

With the identification of the RadA proteins in Archaea,[15] it has become clear that RecA-like proteins exist in all three phylogenetic domains of life. However, the RadA proteins are more similar to the Rad51 proteins (~40% identity at the amino acid level) than they are to the RecA proteins (~20% amino acid residue identity)[15] suggesting that the RadA proteins may provide an even better model for understanding human Rad51 and Dmc1 than does RecA. It has been shown that RadA protein is capable of catalyzing *in vitro* strand exchange reactions in a similar manner to Rad51[16–18] and that it forms a nucleoprotein filament on DNA.[16]

We have used electron microscopy and image analysis to characterize the archaeal *Sulfolobus sofa-*

Abbreviations used: ss, single-stranded; ds, double-stranded; IHRSR, iterative helical real space reconstruction; AFM, atomic force microscopy.

E-mail address of the corresponding author: egelman@virginia.edu

0022-2836/01/051077–9 $35.00/0

taricus RadA protein. We find that the protein exists in two different oligomeric states, octameric rings and helical filaments, and that it can bind DNA in both forms. This provides some insight into how the filamentous Rad51 and octameric Dmc1 forms might have diverged from a common ancestor.

Results

RadA exists as rings and helical filaments

In the absence of DNA and nucleotide cofactor, RadA protein is observed to exist mainly as ring structures (Figure 1(a)). Although it is difficult to quantify the relative abundance of different oligomeric species in the electron microscope, most of the protein appears to exist in this ring state. Some of the rings show a clear stain-filled central channel, while others appear more spherical. The differences in the appearance of the rings can be due to different orientations of the rings on the grid, and this random orientation can be utilized for three-dimensional reconstructions (see below). In the presence of DNA and the absence of nucleotide cofactor, rings of RadA can be seen bound to single-stranded (ss)DNA (Figure 1(c)). Similar stacks of RadA rings were observed on double-stranded (ds)DNA. The binding of the rings to DNA occurred after incubations at either 37 °C or 65 °C. The nearly coaxial arrangement of these rings strongly suggests that DNA is bound within the central channel, as was shown for rings of the hexameric phage T7 gp4 helicase bound along ssDNA,[19] rings of hexameric RuvB bound to dsDNA[20] and octameric rings of human Dmc1 bound along dsDNA.[13] In contrast, rings of DNA-binding proteins that do not stack coaxially on DNA, such as Rad52,[21] the β protein from bacteriophage λ,[22] or the bacteriophage P22 erf protein (Y.Y., Chen, Poteete and E.H.E., unpublished results) appear to bind DNA on the outside of the rings.

In the presence of DNA and an ATP analog, helical filaments of RadA were observed on DNA (Figure 1(b) and (d)). The formation of these helical filaments required DNA, as RadA protein in the presence of either ADP or ATP but the absence of DNA existed mainly in the ring form. The appearance of the filaments that were formed depended upon both temperature and the nucleotide cofactor, and the pitch within individual filaments could range from more than 100 Å (which we call extended) to less than 70 Å (which we call compressed). After incubations at either 37 °C or 65 °C with ATPγS used as an ATP analog (Figure 1(d)), the filaments formed on both ssDNA and dsDNA were quite compressed with respect to the helical pitch. An average of 5274 segments of such RadA-ATPγS-dsDNA filaments (Figure 1(d), inset)

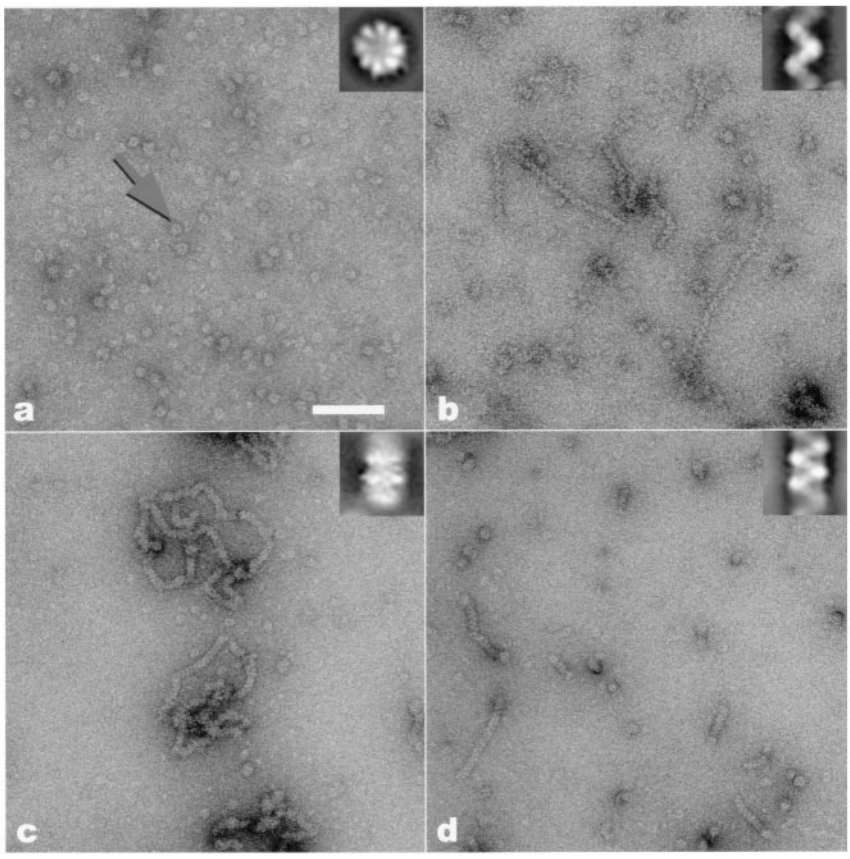

Figure 1. Electron micrographs of three states of the RadA protein: (a) rings in the absence of DNA and nucleotide cofactor; (b) extended helical filaments on dsDNA in the presence of ATP and aluminum fluoride; (c) rings bound to circular ssDNA molecules in the absence of nucleotide cofactor; and (d) compressed helical filaments on dsDNA in the presence of ATPγS. The arrow in (a) indicates a ring that is oriented on the grid so that the 8-fold axis is approximately perpendicular to the plane of projection. A stain-filled central channel can be seen. Insets show averages of these different states. The inset in (a) is from 14,035 images of rings, the inset in (c) is from 2083 images of rings on DNA (each image containing ~three rings), the inset in (b) is from 1924 helical segments selected as having a pitch near 100 Å, and the inset in (d) is from 5274 helical segments. A comparison between the stacked rings (c) and the compressed filaments (d) shows that they are readily distinguished. The scale bar represents 1000 Å.

showed a helical pitch of ~75 Å, and these segments could be sorted into subsets with pitch values from 65 to 80 Å. Due to the RadA-catalyzed hydrolysis of ATP, we would not expect that ATP itself would be useful for structural studies. In fact, after incubations with ATP and DNA no extended filaments were seen, and the helical pitch was quite compressed. The filaments formed on ssDNA after incubations at 65 °C using AMP-PNP, Br-ATP or ADP were also quite compressed. Using aluminum fluoride to stabilize ATP filaments in either an ATP or ADP-P_i state (by forming a stable ADP-aluminium fluoride complex following ATP hydrolysis), both compressed and extended filaments were seen on ssDNA after incubations at either 65 °C or 80 °C. With ATP, aluminum fluoride and dsDNA, on the other hand, only extended filaments were seen after incubations at either 65 °C or 80 °C (Figure 1(b)).

Three-dimensional reconstruction of rings

Single-particle image analysis[23] was used to further examine the structure of the rings in the absence of DNA. A global average of 14,035 ring images (Figure 2(a)) strongly suggests an octameric organization of the protein, but is problematic due to the fact that many different orientations of the ring are being averaged together. We have therefore used a three-dimensional reconstruction of the human Dmc1 octameric ring[13] as a starting point for sorting the images into groups based upon their common orientation. Averages (Figure 2(b) and (c)) from subsets of images selected as having a more homogeneous orientation show the 8-fold symmetry of these rings even more clearly, and no imposition of such symmetry has been made in generating these averages. An iterative approach was used to generate projections of a three-dimensional volume to use for the classification of the images by multi-reference alignment, with back projection from the aligned images generating a new three-dimensional volume.[24] The final three-dimensional reconstruction (Figure 2(d) and (e)) was generated using 12,600 images, rejecting those that displayed a poor correlation with the reference projections. The iterative method was found to be stable for ~15 additional cycles, before the structure began to diverge due to increasing misalignments.

A concern reasonably exists that the starting model (Dmc1) might have biased the outcome of the entire process. First, the correctness of the solution can be seen by the fact that class averages of images generated independently of the reconstruction process look nearly identical to corresponding projections of the three-dimensional reconstruction (data not shown). Second, the structure shown (Figure 2(d) and (e)) looks significantly different from the starting Dmc1 ring. Third, the resolution of the RadA reconstruction is judged to be ~16 Å, significantly better than the resolution of the starting Dmc1 octamer. If artifacts existed due to the

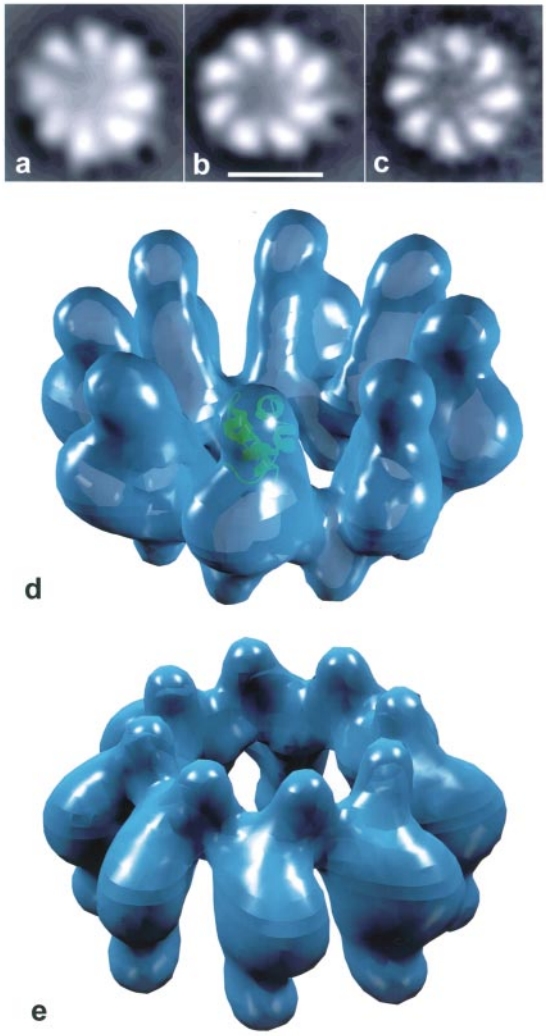

Figure 2. Octameric ring of RadA in the absence of DNA. A reference-free average [24] of all 14,035 images is shown in (a). This average is complicated by the fact that projections of particles having many different orientations in three dimensions are being averaged together. Nevertheless, eight subunits can easily be seen. The 8-fold symmetry of the ring becomes much clearer if rings with only limited tilts of the 8-fold axis from normal to the plane of projection are averaged together. Using the assignments of angles generated by the iterative alignment of the images against projections of the 3D reconstruction, 5823 images were found to correspond to rings which had tilts of the symmetry axis away from the normal to the plane of projection by $\leqslant 40°$. These images were aligned by the reference-free method, and their average, with a clear 8-fold symmetry, is shown in (b). This average still contains images of rings with significantly different tilt angles, so a reference-free average of 752 images corresponding to projections of rings tilted by 20° is shown in (c). Two views are shown ((d) and (e)) of the surface of a three-dimensional reconstruction of the RadA ring, generated from images of 12,600 particles. The atomic structure of the homologous human Rad51 N-terminal domain[27] has been inserted into one of the lobes of the RadA subunit in (d), showing that this feature is the appropriate size and shape to be due to the N-terminal domain of RadA.

fact that the starting model imposed a bias, we would expect that the final resolution would be worse, and not better, than the starting model.

The ring reconstruction is ~140 Å in diameter and 80 Å in height, with a large central channel that is ~50 Å in diameter at its narrowest point. The reconstruction of the ring displays a domain structure for the protein subunit, with a relatively globular central core, and two unequal protrusions at opposite ends of the core. It has been shown that the N-terminal extensions (missing in the bacterial RecA proteins) in RadA and human Rad51 are quite homologous.[15,25,26] We have therefore tried to dock a structure of the N-terminal domain of the human Rad51 protein[27] into the ring reconstruction. It can be seen (Figure 2(d)) that the fit is very good when the N-terminal domain is docked into the larger of the two projections. We cannot exclude the possibility that the N-terminal domain is actually part of the globular core. However, the fact that the N-terminal regions of both the human Rad51 protein[27] and the archaeal *Pyrococcus furiosus* RadA protein[26] have been shown to be independently folding domains suggests that it is more likely that they will be seen as protrusions than as part of the globular core. In addition, the fact that DNA binds to this lobe (below) is consistent with the observation that the homologous N-terminal domain of the human Rad51 protein binds DNA.[27]

Reconstruction of rings on DNA

We have used the reconstruction of the RadA ring in the absence of DNA (Figure 2(d)) as a starting model for reconstructing the rings bound to circular ssDNA. Images from 12,912 such rings were used but, unlike the rings in the absence of DNA, the imposition of symmetry is problematic. Since subunits in the ring are identical and related by an 8-fold symmetry, the binding of DNA must break the symmetry of the ring as all subunits cannot bind the DNA equivalently. In fact, the simplest model is that only one subunit of the eight is binding DNA, as described for the T7 gp4 hexamer.[28] Imposition of 8-fold symmetry generated a reconstruction with a density in the central channel along the symmetry axis (data not shown), consistent with the location of the DNA in the central channel. However, this imposition of symmetry obscured completely the binding of DNA to an individual subunit or subunits. We therefore generated a reconstruction treating the ring as a completely asymmetric object. This greatly lowered the signal-to-noise ratio, and as a result the process began to diverge after approximately four cycles, rather than after approximately 15 cycles when symmetry was used for the rings in the absence of DNA. The reconstruction generated after the fourth cycle of this procedure (Figure 3) shows density in the central channel that is offset from the symmetry axis. The simplest interpretation of this additional density is that it is due to the DNA. Most importantly, two points of contact can be seen between this density and the protein ring. There is a strong contact with the globular core, as well as with the domain emerging from the core that we have suggested (Figure 2(d)) is the N-terminal domain.

Helical filaments on DNA

As noted above, large variations in the structure of the filaments formed by RadA on DNA were observed as a function of the nucleotide cofactor used and the temperature of the incubation. We have generated a number of reconstructions from filaments formed during incubations at 65 °C in the presence of ATP and aluminum fluoride. We think that these reflect filament formation at a temperature close to physiological, and with a very good analog for ATP. These filaments were thus formed under similar conditions to those used by Seitz *et al.* in their observation of RadA filaments by atomic force microscopy (AFM).[16] Reconstructions of the RadA helical filaments under these conditions were still complicated by the large variability in helical pitch. We sorted filament segments into groups based upon cross-correlations with refer-

Figure 3. Three-dimensional reconstruction of the RadA ring bound to ssDNA (generated from 12,912 images), shown in three different orientations. Due to the fact that no symmetry has been imposed upon this reconstruction, it is considerably more noisy and less reliable than the reconstruction in Figure 2(d) and (e). Nevertheless, a clear density in the central channel (colored yellow) arises that must be due to the DNA, and this density is offset from the axis of the ring. There are two strong regions of contact between this density and the protein ring (indicated by arrows). One of these regions is the lobe that we have suggested is due to the N-terminal domain (Figure 2(d)) while the other region is the globular core of the protein.

ence filaments having a pitch from 85-110 Å in 5 Å steps, and then used these subsets for helical reconstructions with a new approach based upon real space refinement of the helical geometry (iterative helical real space reconstruction, or IHRSR).[11] Reconstructions were generated from groups that were initially sorted as having helical pitches close to 95, 100, 105 or 110 Å. These subsets contained 1516, 1924, 1742, and 2758 segments, respectively. The actual pitches found for these four groups, after ~200 cycles of iteration, were 100.3 Å, 104.8 Å (Figure 4(a)), 105.6 Å and 109.6 Å, respectively. For comparison, the average pitch reported by AFM measurements was 109(±19) Å.[16] The ''twist'' of these subsets, as measured by the number of subunits per turn, was much less variable, and was about 6.6 subunits/turn for each set (actual values were 6.59. 6.55, 6.52 and 6.62, respectively). The extended state of the human Rad51 filament has ~6.4 subunits/turn, the yeast Rad51 filament has ~6.3 subunits/turn, the bacteriophage T4 UvsX filament has ~6.2 subunits/turn[29] and the *E. coli* RecA filament has ~6.2 subunits/turn.[30] The reconstructions show that the twist of the RadA filaments and the surface features remain fairly constant as the pitch changes by ~10 %.

Surprisingly, the RadA-ATP-aluminum fluoride filament reconstructions (Figure 4(a)) do not display the prominent lobes, seen in the human Rad51 filaments (Figure 4(c) and (d)), that we have interpreted as due to the N-terminal domain.[30] This is surprising not just because this domain is homologous between the two proteins, but because the octameric RadA ring (Figures 2 and 3) displays prominent smaller lobes into which the human Rad51 N-terminal domain can be docked (Figure 2(d)). In contrast, reconstructions from the compressed RadA-ATPγS filaments do show these lobes. A reconstruction from segments selected as having a helical pitch near 70 Å is shown in Figure 4(b). This reconstruction was generated from 1019 segments, and the average pitch was found to be 71.4 Å with 6.30 units/turn after 100 cycles of IHRSR. The best fit between the extended RadA filament and the compressed RadA filament is shown in Figure 5. It can be seen that even though the helical parameters are incommensurate, a single subunit can be matched quite well with the exception of the pendulous lobes that are seen in the compressed state (blue arrow, Figure 5).

We have considered four possible explanations for the failure to see this putative N-terminal lobe in the extended RadA helical filaments. One is that through some form of proteolysis, this domain is removed during or after filament formation. Gel electrophoresis of the protein, either before or after filament formation, does not show any appreciable proteolysis (data not shown), so this explanation can be completely excluded. Another possibility is that the subunit undergoes a major conformational change, such that the N-terminal domain is no longer seen at low resolution as a discrete domain, as it is seen in both the RadA ring and the compressed RadA filament formed on DNA. A third possibility is that the packing of subunits is quite different in the extended RadA filament than it is in the compressed RadA filament, the extended human Rad51 filament, and the compressed human Rad51 filament, so that the N-terminal domain is no longer a protruding lobe in the extended RadA filament. The fourth possibility is that this domain is extremely disordered in the extended state, so it is not visualized in the recon-

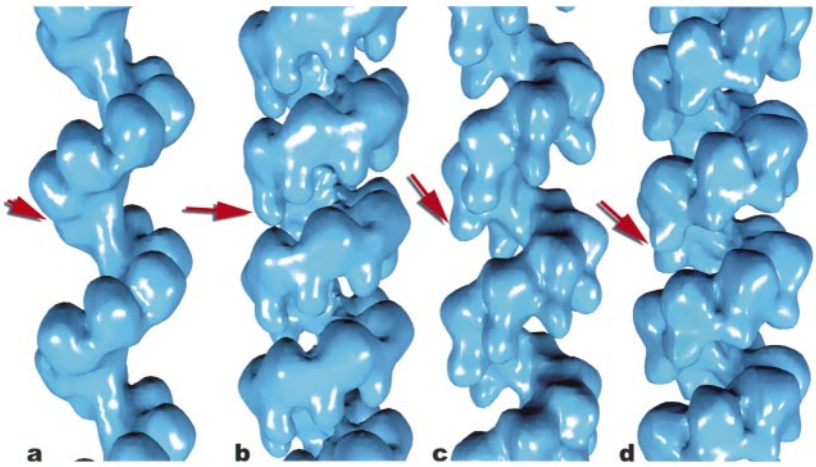

Figure 4. A comparison between (a) and (b) two states of the RadA-DNA filament and (c) and (d) two states of the human Rad51-DNA filament.[14] The extended filaments in (a) and (c) were formed in the presence of ATP and aluminum fluoride, while the compressed filaments in (b) and (d) were prepared in the presence of ATPγS. The best fit (Figure 5) of this (a) extended filament to the (b) compressed state is with the polarity shown. The arrows in (b), (c) and (d) indicate the subunit lobes that have been interpreted as being due to the N-terminal domain.[14] In contrast, these lobes are not visualized in the extended RadA filament, and the arrow (a) indicates where this domain would be located based upon the alignment with the compressed filament in (b).

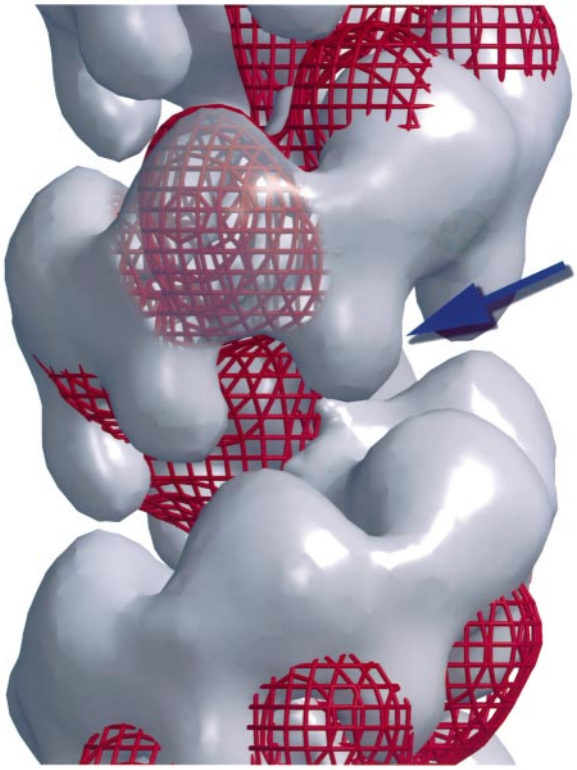

Figure 5. A superposition of the compressed RadA-DNA-ATPγS filament (continuous grey surface) and the extended RadA-DNA-ATP-aluminum fluoride filament (red mesh surface). The region of one subunit has been made transparent in the compressed filament surface, showing that the helical ''backbone'' of both filaments, containing a globular core region, superimposes very well. In contrast, the pendulous lobes are not seen in the extended filament. The simplest interpretation is that these subunit lobes (blue arrow) are not seen in the extended state due to either a large mobility of this compact domain or an unfolding of this region.

structions. If this last possibility is correct, one should be able to superimpose the extended and compressed RadA reconstructions, and the difference should be almost entirely limited to the additional N-terminal lobe. Such a superposition is shown in Figure 5, and it can be seen that the main difference is due to this protruding lobe. Additional considerations (discussed below) from observations of both human and yeast Rad51 filaments suggest that a large mobility or disorder in the N-terminal domain is the most reasonable explanation for the failure to see this lobe in the extended RadA filaments. However, alternative possibilities can still not be excluded.

Discussion

We have shown that an archaeal RadA protein exists in two different oligomeric states, octameric rings and helical filaments, and that it binds DNA in both states. The helical RadA nucleoprotein filament formed on DNA in the presence of ATP analogs is similar to the filaments formed by *E. coli* RecA, bacteriophage T4 UvsX,[31] and the eukaryotic Rad51 proteins. Rings of both *E. coli* RecA[32] and human Rad51[33] have been described, and we have shown that these rings are hexameric for *Thermophilus aquaticus* RecA[34] and octameric for human Rad51 (X.Y. and E.H.E., unpublished results). But no binding of RecA or Rad51 rings to DNA has been observed. In contrast, it has been shown that octameric rings of human Dmc1 bind DNA,[13] and no helical filaments have been observed thus far for Dmc1. Since the archaeal RadA proteins have been postulated to be orthologous to the common ancestor of Rad51 and Dmc1,[15] it is tempting to speculate that two functional forms of RadA, the helical filament and the octameric ring, differentiated over evolution into two distinctly different proteins, helical Rad51 and octameric Dmc1. Greater insight into this possibility will come from understanding the different functional roles of the two RadA oligomeric states.

Like RecA,[35] human Rad51[14] and bacteriophage T4 UvsX[29] proteins, the RadA helical filaments exists in two different states, ''compressed'' and ''extended''. We have shown for these other proteins that the two states are structurally distinct, even though the helical pitch in these two states can actually overlap for the RecA protein.[14] Thus, the distinction is more than just helical pitch. All evidence suggests that the compressed filament is the ADP state, the extended filament is the ATP state, and the active presynaptic filament in recombination reactions is the extended ATP filament. We have found that, like UvsX and human Rad51, the two helical states of RadA do not overlap in pitch and are therefore easily distinguishable. In addition, the compressed filaments (formed on DNA in the presence of ATPγS) are easily distinguished from the stacks of rings that form on DNA in the absence of nucleotide cofactor. This is in part due to the fact that there are eight subunits in each ring, while the filaments have ~six subunits per helical turn, and provides a further argument that there is unlikely to be a simple transition between the rings formed by the RecA-like proteins and the filaments.[34] In the case of human Dmc1, the stacks of rings formed by this protein could never interconvert into polar helical filaments, since the stacks contain rings of alternating polarity.[13] Such a situation is similar to the stacked disk aggregate of tobacco mosaic virus, where it has been shown that these structures are bipolar and therefore cannot interconvert to the helical form where all subunits have the same polarity.[36,37]

The slowly hydrolyzable ATP analog ATPγS does not induce the extended filamentous state of RadA, as it does for RecA. This is consistent with the fact that in the presence of ATPγS, RadA protein does not promote DNA strand exchange *in vitro* (E.M.S. and S.C.K., unpublished results),

but this ATP analog does allow for strand exchange with the RecA protein.[38]

We have been surprised to find that the extended RadA filaments differ significantly in appearance from the human Rad51 filaments, despite overall sequence similarity. The main difference is that the prominent lobes seen in hRad51 filaments are not seen in the extended RadA filaments, even though both rings of RadA and compressed filaments of RadA show these lobes. Several possibilities were considered to explain this difference. We have been able to eliminate the possibility of proteolytic removal of these lobes during or after filament formation. The simplest remaining explanation is that the N-terminal domain in RadA is extremely disordered in the extended helical filaments, and thus not visualized after extensive averaging. This possibility is consistent with the fact that the highly homologous N-terminal region of hRad51 has been shown to exist as an independently folding domain in solution,[27] and thus might be connected by a flexible linker region in RadA to the conserved core that forms the helical backbone. The disordered N-terminal hypothesis is also consistent with the fact that the backbones of the extended and compressed RadA filaments superimpose very well (Figure 5).

A comparison between the compressed and extended human Rad51 filaments showed that the main change involved a large shift of the putative N-terminal domain.[14] The difference between the compressed and extended RadA filaments, on the other hand, involves the ''disappearance'' of this same lobe. Thus, in human Rad51 the difference between the two filament states involves a shift of this domain, while in RadA it may involve a destabilization of this domain. Another protein that polymerizes, bacterial flagellin, has been shown to contain both N and C-terminal domains that are largely disordered in solution,[39,40] suggesting that it may be difficult to distinguish between unfolding of a domain and large mobility of a domain when this domain is not visualized by structural methods.

While the core structures of Dmc1 and UvsX have been predicted to be similar to that of RecA,[41] we have little detailed structural information about the quaternary structure of RecA-like recombination filaments. The only high-resolution information came from the helical filament seen in an *E. coli* RecA crystal in the absence of ATP and DNA.[42] Electron microscopic and biochemical studies have suggested that there are large conformational changes that occur between RecA self-polymers and RecA-DNA-ATP filaments, but that the subunit-subunit interface may be relatively conserved between the two forms.[14,35,43,44] The more recent crystal structure of *Mycobacterium tuberculosis* RecA in the presence of an ATP analog shows a highly conserved subunit-subunit interface in comparison with the *E. coli* structure, which is not surprising given the 62% identity between these proteins.

However, several lines of additional evidence suggest that the packing of subunits is not conserved between the RecA/UvsX filaments, on the one hand, and the RadA/Rad51 filaments, on the other hand. When the crystal structure of the RecA subunit is fit into the hRad51-DNA filament, a different subunit-subunit interface from that found in the RecA filament is predicted; when the same RecA crystal structure is fit into UvsX-DNA filaments, the same subunit-subunit interface that is seen in RecA is predicted.[29] It was noted that the blocks of sequence similarity between RecA and the Rad proteins (RadA and Rad51) on the one hand, and between RecA and UvsX, on the other hand, are totally disjoint, and involve no overlap.[45] Remarkably, these blocks involve the predicted subunit-subunit interfaces in both UvsX and Rad51.[29]

A different packing raises questions. The same unusual DNA structure induced by the RecA, UvsX and Rad51 helical filaments gave rise to the suggestion that the conservation of this DNA form was the main force in evolution to constrain the divergence of the quaternary protein organization.[9] This would explain why the helical parameters of the bacterial RecA filament (~6.2 subunits per turn of a 92 Å pitch helix) are similar to those of the human Rad51 protein (~6.4 subunits per turn of a 99 Å pitch helix), if the role of both filaments was to induce a similar structure in DNA. But a different packing of subunits between RecA and UvsX on the one hand, and RadA and Rad51 on the other, would raise an alternate and opposite possibility: the helical parameters of the recombination filaments have been conserved due to an intrinsic structural property of DNA, in which a helical pitch of ~95 Å with ~18 base-pairs per turn is favored. An aspect of this model, that RecA binds to a stretched state of DNA that spontaneously occurs by thermal fluctuations, has been proposed.[46] A prediction of such a model is that there would exist a coupling between the helical pitch and twist of such protein-DNA filaments over evolution that would be due to the intrinsic properties of this unusual DNA conformation. The extended RadA filaments have the longest average pitch (~105 Å) of any of the RecA-like filaments thus far examined. Interestingly, they also have more subunits per turn (~6.6) than in the filaments formed by any of the orthologous proteins. This suggests a coupling between these two parameters, and an analysis of these parameters from RadA, RecA, Rad51 and UvsX extended filaments shows that such a coupling does exist.[47] In this new view, the higher-order assembly of similar helical filaments by RadA, Rad51, UvsX and RecA proteins may likely be the result of convergent evolution, even though the subunits diverged from a common ancestor. Higher-resolution structural studies of UvsX, Rad51, RadA and RecA filaments will be essential for exploring this possibility.

Materials and Methods

Formation of complexes

RadA protein was prepared as described.[16] RadA rings were observed after incubating 2 μM RadA in 25 mM triethanolamine-HCl (Fisher) buffer (pH 7.2) with 2 mM magnesium acetate (Sigma) at 37 °C for 15 minutes. RadA-ssDNA ring complexes were formed with a RadA concentration of 4 μM in 25 mM triethanolamine-HCl (Fisher) buffer (pH 7.2) and 2 mM magnesium acetate (Sigma) and incubated at 65 °C for 15 minutes, with RadA to M13 DNA (Sigma) ratio of 40:1 (w/w). RadA-dsDNA-ATPγS filaments were formed by incubating 4 μM RadA in 25 mM triethanolamine-HCl (Fisher) buffer (pH 7.2) at 65 °C for 20 minutes, with a RadA to linearized ϕX174 dsDNA (Sigma) ratio of 20:1 (w/w), 2.5 mM ATPγS (Roche Diagnostics), and 2 mM magnesium acetate (Sigma). RadA-dsDNA-ATP-aluminum fluoride filaments were formed with a RadA concentration of 4 μM in 25 mM triethanolamine-HCl (Fisher) buffer (pH 7.2), 2.5 mM ATP (Sigma), and 2 mM magnesium acetate (Sigma), and incubated at 65 °C for 15 minutes, with a RadA to linearized ϕX174 dsDNA (Gibco BRL) ratio of 20:1 (w/w). Then NaF (Aldrich) and $Al(NO_3)_3$ (Aldrich) were added to a final concentration of 2.5 mM and incubated at 65 °C for an additional 15 minutes. Circular ϕX174 dsDNA was linearized as described.[35]

Electron microscopy and image analysis

Samples were applied to carbon-coated grids and negatively stained with 1% (w/v) uranyl acetate. Specimens were examined in a Philips Tecnai 12 electron microscope at an accelerating voltage of 80 keV and a nominal magnification of 30,000×. Negatives were densitometered with a Leaf 45 scanner, using a raster of 3.9 Å/pixel. The SPIDER software package[23] was used for single particle image analysis. Helical reconstructions were generated within this package using the method of IHRSR.[11] Resolution was determined by randomly dividing the images into two equal sets, and generating two independent reconstructions. The resolution was taken as the point at which the coefficient of correlation between shells in the Fourier transform fell below 0.5.

Acknowledgments

This work was supported by NIH grants GM35269 (to E.H.E.) and by GM62653 (to S.C.K.). We thank Margaret Van Loock for helpful discussions.

References

1. Roca, A. I. & Cox, M. M. (1997). RecA protein: structure, function, and role in recombinational DNA repair. *Prog. Nucl. Acid Res. Mol. Biol.* **56**, 129-223.
2. Kowalczykowski, S. C. & Eggleston, A. K. (1994). Homologous pairing and DNA strand-exchange proteins. *Annu. Rev. Biochem.* **63**, 991-1043.
3. Dunderdale, H. J. & West, S. C. (1994). Recombination genes and proteins. *Curr. Opin. Genet. Dev.* **4**, 221-228.
4. Stasiak, A., DiCapua, E. & Koller, T. (1981). Elongation of duplex DNA by RecA protein. *J. Mol. Biol.* **151**, 557-564.
5. Stasiak, A. & DiCapua, E. (1982). The helicity of DNA in complexes with RecA protein. *Nature,* **229**, 185-186.
6. Lim, D. S. & Hasty, P. (1996). A mutation in mouse rad51 results in an early embryonic lethal that is suppressed by a mutation in p53. *Mol. Cell Biol.* **16**, 7133-7143.
7. Tsuzuki, T., Fujii, Y., Sakumi, K., Tominaga, Y., Nakao, K., Sekiguchi, M., Matsushiro, A., Yoshimura, Y. & Morita, T. (1996). Targeted disruption of the RAD51 gene leads to lethality in embryonic mice. *Proc. Natl Acad. Sci. USA,* **93**, 6236-6240.
8. Sonoda, E., Sasaki, M. S., Buerstedde, J. M., Bezzubova, O., Shinohara, A., Ogawa, H. *et al.* (1998). Rad51-deficient vertebrate cells accumulate chromosomal breaks prior to cell death. *EMBO J.* **17**, 598-608.
9. Ogawa, T., Yu, X., Shinohara, A. & Egelman, E. H. (1993). Similarity of the yeast RAD51 filament to the bacterial RecA filament. *Science,* **259**, 1896-1899.
10. Benson, F. E., Stasiak, A. & West, S. C. (1994). Purification and characterization of the human Rad51 protein, an analogue of *E. coli* RecA. *EMBO J.* **13**, 5764-5771.
11. Egelman, E. H. (2000). A robust algorithm for the reconstruction of helical filaments using single-particle methods. *Ultramicroscopy,* **85**, 225-234.
12. Bishop, D. K., Park, D., Xu, L. & Kleckner, N. (1992). DMC1: a meiosis-specific yeast homolog of E. coli recA required for recombination, synaptonemal complex formation, and cell cycle progression. *Cell,* **69**, 439-456.
13. Passy, S. I., Yu, X., Li, Z., Radding, C. M., Masson, J. Y., West, S. C. & Egelman, E. H. (1999). Human Dmc1 protein binds DNA as an octameric ring. *Proc. Natl Acad. Sci. USA,* **96**, 10684-10688.
14. Yu, X., Jacobs, S. A., West, S. C., Ogawa, T. & Egelman, E. H. (2001). Domain structure and dynamics in the helical filaments formed by RecA and Rad51 on DNA. *Proc. Natl Acad. Sci. USA,* **98**, 8419-8424.
15. Sandler, S. J., Satin, L. H., Samra, H. S. & Clark, A. J. (1996). recA-like genes from three archaean species with putative protein products similar to Rad51 and Dmc1 proteins of the yeast *Saccharomyces cerevisiae*. *Nucl. Acids Res.* **24**, 2125-2132.
16. Seitz, E. M., Brockman, J. P., Sandler, S. J., Clark, A. J. & Kowalczykowski, S. C. (1998). RadA protein is an archaeal RecA protein homolog that catalyzes DNA strand exchange. *Genes Dev.* **12**, 1248-1253.
17. Spies, M., Kil, Y., Masui, R., Kato, R., Kujo, C., Ohshima, T., Kuramitsu, S. & Lanzov, V. (2000). The RadA protein from a hyperthermophilic archaeon *Pyrobaculum islandicum* is a DNA-dependent ATPase that exhibits two disparate catalytic modes, with a transition temperature at 75 degrees C. *Eur. J. Biochem.* **267**, 1125-1137.
18. Miyata, T., Yamada, K., Iwasaki, H., Shinagawa, H., Morikawa, K. & Mayanagi, K. (2000). Two different oligomeric states of the RuvB branch migration motor protein as revealed by electron microscopy. *J. Struct. Biol.* **131**, 83-89.
19. Egelman, E. H., Yu, X., Wild, R., Hingorani, M. M. & Patel, S. S. (1995). Bacteriophage T7 helicase/primase proteins form rings around single-stranded

DNA that suggest a general structure for hexameric helicases. *Proc. Natl Acad. Sci. USA,* **92**, 3869-3873.
20. Stasiak, A., Tsaneva, I. R., West, S. C., Benson, C. J. B., Yu, X. & Egelman, E. H. (1994). The *Escherichia coli* RuvB branch migration protein forms double hexameric rings around DNA. *Proc. Natl Acad. Sci. USA,* **91**, 7618-7622.
21. Kagawa, W., Kurumizaka, H., Ikawa, S., Yokoyama, S. & Shibata, T. (2001). Homologous pairing promoted by the human Rad52 protein. *J. Biol. Chem.* **276**, 35201-35208.
22. Passy, S. I., Yu, X., Li, Z., Radding, C. M. & Egelman, E. H. (1999). Rings and filaments of beta protein from bacteriophage lambda suggest a superfamily of recombination proteins. *Proc. Natl Acad. Sci. USA,* **96**, 4279-4284.
23. Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M. & Leith, A. (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol.* **116**, 190-199.
24. Penczek, P. A., Radermacher, M. & Frank, J. (1992). Three-dimensional reconstruction of single particles embedded in ice. *Ultramicroscopy,* **40**, 33-53.
25. Brendel, V., Brocchieri, L., Sandler, S. J., Clark, A. J. & Karlin, S. (1997). Evolutionary comparisons of RecA-like proteins across all major kingdoms of living organisms. *J. Mol. Evol.* **44**, 528-541.
26. Brocchieri, L. & Karlin, S. (1998). A symmetric-iterated multiple alignment of protein sequences. *J. Mol. Biol.* **276**, 249-264.
27. Aihara, H., Ito, Y., Kurumizaka, H., Yokoyama, S. & Shibata, T. (1999). The N-terminal domain of the human Rad51 protein binds DNA: structure and a DNA binding surface as revealed by NMR. *J. Mol. Biol.* **290**, 495-504.
28. Yu, X., Hingorani, M. M., Patel, S. S. & Egelman, E. H. (1996). DNA is bound within the central hole to one or two of the six subunits of the T7 DNA helicase. *Nature Struct. Biol.* **3**, 740-743.
29. Yang, S., VanLoock, M. S., Yu, X. & Egelman, E. H. (2001). Comparison of Bacteriophage T4 UvsX and human Rad51 filaments suggests that RecA-like polymers may have evolved independently. *J. Mol. Biol.* **312**, 999-1009.
30. Karlin, S. & Brocchieri, L. (1996). Evolutionary conservation of RecA genes in relation to protein structure and function. *J. Bacteriol.* **178**, 1881-1894.
31. Yu, X. & Egelman, E. H. (1993). DNA conformation induced by the bacteriophage T4 UvsX protein appears identical to the conformation induced by the *Escherichia coli* RecA protein. *J. Mol. Biol.* **232**, 1-4.
32. Brenner, S. L., Zlotnick, A. & Griffith, J. D. (1988). RecA protein self-assembly. Multiple discrete aggregation states. *J. Mol. Biol.* **204**, 959-972.
33. Baumann, P., Benson, F. E., Hajibagheri, N. & West, S. C. (1997). Purification of human Rad51 protein by selective spermidine precipitation. *Mutat. Res.* **384**, 65-72.
34. Yu, X. & Egelman, E. H. (1997). The RecA hexamer is a structural homologue of ring helicases. *Nature Struct. Biol.* **4**, 101-104.
35. Yu, X. & Egelman, E. H. (1992). Structural data suggest that the active and inactive forms of the RecA filament are not simply interconvertible. *J. Mol. Biol.* **227**, 334-346.
36. Diaz-Avalos, R. & Caspar, D. L. (1998). Structure of the stacked disk aggregate of tobacco mosaic virus protein. *Biophys. J.* **74**, 595-603.
37. Dore, I., Ruhlmann, C., Oudet, P., Cahoon, M., Caspar, D. L. & Van Regenmortel, M. H. (1990). Polarity of binding of monoclonal antibodies to tobacco mosaic virus rods and stacked disks. *Virology,* **176**, 25-29.
38. Menetski, J. P., Bear, D. G. & Kowalczykowski, S. C. (1990). Stable DNA heteroduplex formation catalyzed by the *Escherichia coli* RecA protein in the absence of ATP hydrolysis. *Proc. Natl Acad. Sci. USA,* **87**, 21-25.
39. Vonderviszt, F., Aizawa, S. & Namba, K. (1991). Role of the disordered terminal regions of flagellin in filament formation and stability. *J. Mol. Biol.* **221**, 1461-1474.
40. Vonderviszt, F., Kanto, S., Aizawa, S. & Namba, K. (1989). Terminal regions of flagellin are disordered in solution. *J. Mol. Biol.* **209**, 127-133.
41. Story, R. M., Bishop, D. K., Kleckner, N. & Steitz, T. A. (1993). Structural relationship of bacterial RecA proteins to recombination proteins from bacteriophage T4 and yeast. *Science,* **259**, 1892-1896.
42. Story, R. M., Weber, I. T. & Steitz, T. A. (1992). The Structure of the *E. coli* recA protein monomer and polymer. *Nature,* **355**, 318-325.
43. Yu, X., Shibata, T. & Egelman, E. H. (1998). Identification of a defined epitope on the surface of the active RecA-DNA filament using a monoclonal antibody and three-dimensional reconstruction. *J. Mol. Biol.* **283**, 985-992.
44. Skiba, M. C. & Knight, K. L. (1994). Functionally important residues at a subunit interface site in the RecA protein from *Escherichia coli*. *J. Biol. Chem.* **269**, 3823-3828.
45. Karlin, S., Weinstock, G. M. & Brendel, V. (1995). Bacterial classifications derived from recA protein sequence comparisons. *J. Bacteriol.* **177**, 6881-6893.
46. Leger, J. F., Robert, J., Bourdieu, L., Chatenay, D. & Marko, J. F. (1998). RecA binding to a single double-stranded DNA molecule: a possible role of DNA conformational fluctuations. *Proc. Natl Acad. Sci. USA,* **95**, 12295-12299.
47. Egelman, E. H. (2001). Does a stretched DNA structure dictate the helical geometry of RecA-like filaments? *J. Mol. Biol.* **309**, 539-542.

Edited by M. Belfort

(Received 5 June 2001; received in revised form 4 September 2001; accepted 25 October 2001)