Structural Polymorphism of the RecA Protein from the Thermophilic Bacterium *Thermus aquaticus*

Xiong Yu,* Evelina Angov,‡ R. Daniel Camerini-Otero,‡ and Edward H. Egelman*
*Department of Cell Biology and Neuroanatomy, University of Minnesota Medical School, Minneapolis, Minnesota 55455, and ‡Genetics and Biochemistry Branch, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, Maryland 20892, USA

ABSTRACT The *Escherichia coli* RecA protein has served as a model for understanding protein-catalyzed homologous recombination, both in vitro and in vivo. Although RecA proteins have now been sequenced from over 60 different bacteria, almost all of our structural knowledge about RecA has come from studies of the *E. coli* protein. We have used electron microscopy and image analysis to examine three different structures formed by the RecA protein from the thermophilic bacterium *Thermus aquaticus*. This protein has previously been shown to catalyze an in vitro strand exchange reaction at an optimal temperature of about 60°C. We show that the active filament formed by the *T. aquaticus* RecA on DNA in the presence of a nucleotide cofactor is extremely similar to the filament formed by the *E. coli* protein, including the extension of DNA to a 5.1-Å rise per base pair within this filament. This parameter appears highly conserved through evolution, as it has been observed for the eukaryotic RecA analogs as well. We have also characterized bundles of filaments formed by the *T. aquaticus* RecA in the absence of both DNA and nucleotide cofactor, as well as hexameric rings of the protein formed under all conditions examined. The bundles display a very large plasticity of mass within the RecA filament, as well as showing a polymorphism in filament-filament contacts that may be important to understanding mutations that affect surface residues on the RecA filament.

INTRODUCTION

The *Escherichia coli* RecA protein has served as a model system for understanding protein-catalyzed homologous genetic recombination (see Kowalczykowski and Eggleston, 1994; West, 1992; Rao et al., 1995, for reviews). For over 15 years the biochemical properties of the *E. coli* RecA protein have been characterized, and the functions of the protein have been studied by means of both genetics and in vitro strand exchange reactions. Two lines of evidence suggest that the insight gained from studying RecA will be more generally applicable than only to recombination and repair in *E. coli*. On the one hand, RecA-like proteins have now been sequenced from over 60 different prokaryotic species (Roca and Cox, 1995), and all are seen to have a well-conserved primary structure. On the other hand, eukaryotic analogs of the RecA protein, such as the yeast (Shinohara et al., 1992) and human Rad51 protein (Shinohara et al., 1993), have now been identified.

One of the most important features of the *E. coli* RecA protein is that all activity, including the cleavage of repressor molecules, resides in an extended polymer that is formed by RecA on DNA in the presence of ATP. This polymer is a ternary complex, because it has been shown that RecA's affinity for DNA depends upon the presence of ATP, and RecA's affinity for ATP depends upon the presence of DNA (Weinstock et al., 1981). Within this polymer, the DNA exists in a highly unusual form: it is stretched to an axial rise of 5.1 Å per base or base pair (Stasiak et al., 1981), and untwisted to 18.6 base pairs per turn (Stasiak and DiCapua, 1982). This contrasts with B-form DNA, which has an axial rise of about 3.4 Å per base pair and about 10.5 base pairs per turn (Saenger, 1984). Thus, the *E. coli* RecA protein changes the pitch of DNA from about 36 Å (3.4 Å/bp × 10.5 bp/turn) to about 95 Å (5.1 Å/bp × 18.6 bp/turn), which, not coincidentally, is the pitch of the RecA helix. It has now been shown that the yeast Rad51 protein (Ogawa et al., 1993) and the human Rad51 protein (Benson et al., 1994) induce similar changes in the DNA contained within their polymers. It has also been shown that the bacteriophage T4 UvsX protein, which shares a weaker homology with RecA, induces a very similar conformation in DNA (Yu and Egelman, 1993). Thus, the unusual DNA conformation described for the *E. coli* RecA-DNA filament is found for three other filaments formed by RecA-like proteins from phage, yeast, and humans.

Although a large number of RecA-like proteins have now been identified, very few experimental studies have been made on the proteins from organisms other than *E. coli*. An important question is whether the structural parameters that have been described for these four different protein polymers are actually universal, or whether there exist large variations in these parameters. An extreme conservation of the degree of extension and untwisting of DNA induced by the *E. coli* RecA protein would suggest that this particular conformation may be an essential element in the mechanics of homologous pairing and strand exchange. Toward this end, we have investigated the structures formed by the RecA protein from the thermophilic bacterium *T. aquaticus*. It has been shown that this protein catalyzes an in vitro

Received for publication 8 June 1995 and in final form 22 August 1995.
Address reprint requests to Dr. Edward H. Egelman, Department of Cell Biology and Neuroanatomy, University of Minnesota Medical School, 321 Church Street SE, Minneapolis, MN 55455-0303. Tel.: 612-626-0100; Fax: 612-624-8118; E-mail: egelman@egel1.med.umn.edu.

0006-3495/95/12/2728/11 $2.00

strand exchange reaction with a peak efficiency at about 60°C (Angov and Camerini-Otero, 1994), a temperature that is actually fairly close to the t_m for thermal denaturation of dsDNA (Voet and Voet, 1990). To better understand the mechanism of RecA-mediated recombination, it is relevant to determine whether the *T. aquaticus* protein induces the same structures at these higher temperatures as the *E. coli* protein forms at much lower temperatures.

MATERIALS AND METHODS

Specimen preparation and electron microscopy

The *T. aquaticus* RecA protein was purified as previously described (Angov and Camerini-Otero, 1994). Unless otherwise noted, all reactions of *T. aquaticus* RecA and DNA were carried out in a 25 mM triethanolamine-HCl (Fisher) buffer (pH 7.2) at 65°C for 10 min, with a RecA concentration of 3 μM, 25 mM KCl, 1.25 mM ATPγS (Boehringer), and φX174 dsDNA (Gibco BRL), with a RecA to DNA ratio of 40:1 (w/w). The reaction tube was transferred to an ice bath immediately after incubation. For reactions with linearized dsDNA, circular φX174 dsDNA (Gibco BRL) was linearized as described in Yu and Egelman (1992).

A JEOL 1200EXII, operating at an accelerating voltage of 80 kV, was used for all microscopy. The defocus was typically less than 0.5 μm. Images used for averaging or reconstructions were recorded under minimal dose conditions, with no exposure to the beam at high magnification before recording. Images were recorded at approximately 30,000× magnification, with the exact magnification determined from the 23 Å pitch helix of tobacco mosaic virus particles coprepared with the RecA.

Image analysis

Electron micrographs were digitized with a 12-bit Eikonix photodiode camera, with a sampling of approximately 3 Å/pixel or 6 Å/pixel. Filament images were corrected for curvature (Egelman, 1986), padded into 128 × 2048 pixel arrays, and Fourier transformed. All transforms were performed on a Sky Warrior array processor, running on a VAX 3200. Layer lines were extracted from the transforms, and the near and far sides were averaged together. A single layer line set, chosen from the filaments being averaged, was used as a starting reference in searching for the translation, rotation, and polarity of each filament layer line set. After an initial average was calculated, this averaged layer line set was used as the new reference for the next cycle of alignment. This procedure was iterated three or four times, until there was no shift of individual filaments with respect to the average.

Contour length measurements were made using prints placed on a digitizing tablet, using the known magnification of the electron microscope and the magnification of the print as a conversion factor. Pitch measurements were determined from Fourier transforms of straightened filaments, using tobacco mosaic virus particles as an internal standard.

The single particle analysis of the ring structures was performed using the SPIDER software package (Frank et al., 1981). A reference-free alignment algorithm (Penczek et al., 1992) was used to search for the rotation and translation necessary to bring each individual image into an optimal alignment with all the other images.

RESULTS

The active filament

The active form of *E. coli* RecA has been shown to be a helical polymer of the protein formed on DNA in the presence of a nucleoside triphosphate (reviewed in Egelman, 1993). We have therefore used similar conditions for looking at the *T. aquaticus* RecA protein. Fig. 1 *a* shows the filaments that are seen after the *T. aquaticus* RecA protein is incubated with linear dsDNA and the slowly hydrolyzable ATP analog ATP-γ-S for 10 min at 65°C. Incubations at 22–24°C produced filaments, but the filaments were much shorter. Increasing the incubation time did not significantly increase the length of the filaments, suggesting that the effect of temperature was more than simply on the total rate of polymerization. Incubation at 37°C produced longer filaments, but they were still shorter than the ones formed at 65°C. The total amount of polymeriza-

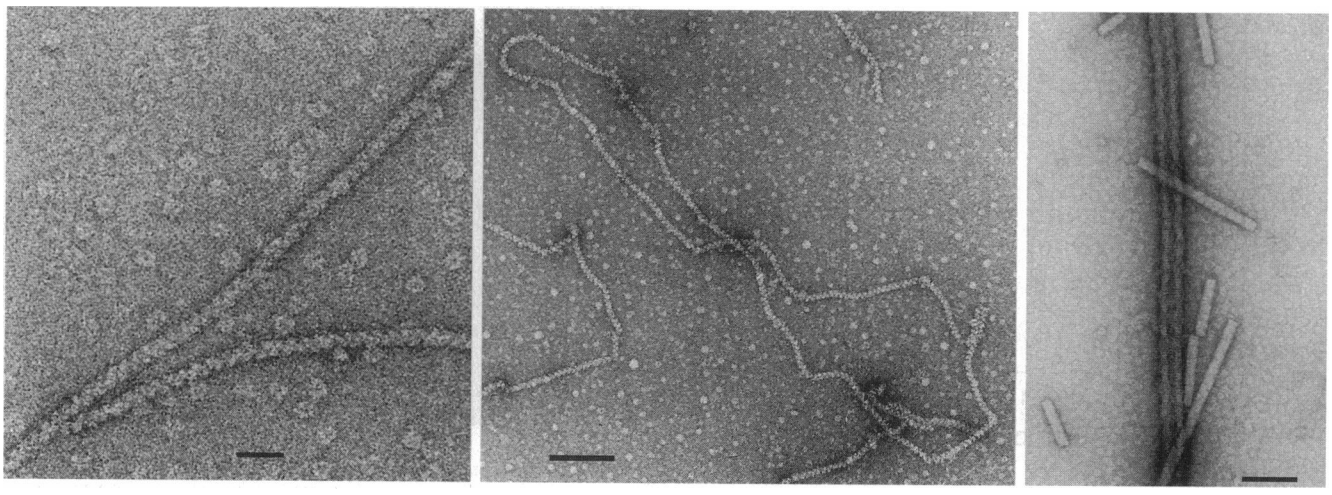

FIGURE 1 Electron micrographs of three structural states of the *T. aquaticus* RecA protein. In *a* and *b*, the active filament formed by the protein on DNA in the presence of ATP-γ-S can be seen. The two filaments in *a* are on linear dsDNA, and the closed circle in *b* contains a nicked circular φX174 dsDNA molecule (5, 386 base pairs). The black circles in *a* indicate the ~130 Å diameter ring-like structures formed by the *T. aquaticus* RecA protein. These same structures can be seen at lower magnification in the background in both *b* and *c*. In *c*, no DNA or nucleotide cofactor has been added, and both rings and a large bundle can be observed. Tobacco mosaic virus particles are also seen in *c*, and these have been added as an internal magnification standard. The space bar in *a* is 300 Å, and the space bars in *b* and *c* are each 1000 Å.

tion appeared to be smaller at the lower temperatures. Interestingly, filament formation did not depend upon the presence of magnesium, as all filaments used for study in this paper were prepared in the absence of magnesium. In the presence of 6 mM magnesium similar filaments were formed (data not shown). Under all conditions used for these studies (in the presence and absence of nucleotide cofactor, and the presence and absence of DNA), the protein formed ring-like structures that can be seen as a background in Fig. 1, *a–c*. These are indicated with circles in Fig. 1 *a*; the analysis of these stuctures will be discussed later.

A remarkable feature of the *E. coli* RecA protein (Stasiak et al., 1981), the yeast Rad51 protein (Ogawa et al., 1993), and the bacteriophage UvsX protein (Yu and Egelman, 1993b) is that they all stretch dsDNA to about a 5.1 Å average rise per base pair from 3.4 Å for B-DNA. This extension can be easily measured by saturating circular relaxed dsDNA circles with the protein. Fig. 1 *b* shows a ϕX174 DNA (5, 386 bp) molecule covered with the *T. aquaticus* RecA protein in the presence of ATP-γ-S. Fig. 2 shows the histogram of contour lengths for 32 such circular molecules. We expect that a number of factors will contribute to many such DNA molecules not being fully saturated by the protein, such as multiple nucleation points on a circle, DNA secondary structure, etc. We therefore suggest that the most extended molecules correspond to fully saturated circles, and the arrow in Fig. 2 indicates the extension that would correspond to an average axial rise per base pair of 5.1 Å. This same behavior of incomplete extension for many covered DNA molecules can be seen for the yeast Rad51 protein (Fig. 2; Ogawa et al., 1993) and the bacteriophage T4 UvsX protein (Fig. 3; Yu and Egelman, 1993b). Thus, the *T. aquaticus* RecA protein induces the same characteristic extension of dsDNA at 65°C as the other proteins do at 37°C.

Fig. 3 shows a computed Fourier transform of a *T. aquaticus* RecA–dsDNA–ATP-γ-S filament. The layer line labeled as $l = 6$ arises from the strong one-start helix relating each subunit, and the axial spacing of this layer line can be used to measure the pitch of this helix. Fig. 4 shows a histogram of the distribution of pitch for 72 such filaments. The mean pitch, 88.6 Å ± 0.3 Å (SEM), is somewhat less than the approximately 92 Å pitch found for the *E. coli* RecA protein prepared under similar conditions (Yu and Egelman, 1990). Interestingly, when filaments formed by an incubation at 37°C were examined, their pitch was significantly smaller. Measurements of Fourier transforms from 10 such filaments yielded a mean of 82.0 ± 2.5 (SD) Å.

The ratio of axial spacings between the strong layer line labeled as $l = 6$ and the weaker layer line labeled as $l = 7$ in Fig. 3 can be used as a sensitive measure of the number of subunits per turn along the one-start helix. This computation requires a knowledge of the indexing of the transform, which is the order of the Bessel function or functions

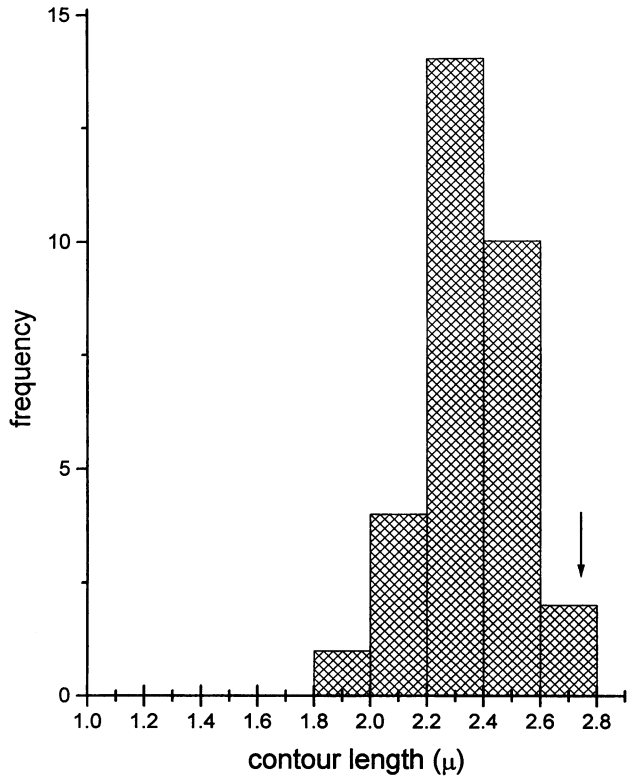

FIGURE 2 The contour lengths of 31 nicked circular ϕX174 dsDNA molecules covered by the *T. aquaticus* RecA protein in the presence of ATP-γ-S (as in Fig. 1 *b*). The arrow indicates the extension that would correspond to an axial rise per base pair of 5.1 Å. This maximal extension appears to be found when the dsDNA is completely saturated with the protein, whereas the shorter extensions observed in the histogram are consistent with incomplete coverage by the *T. aquaticus* RecA protein.

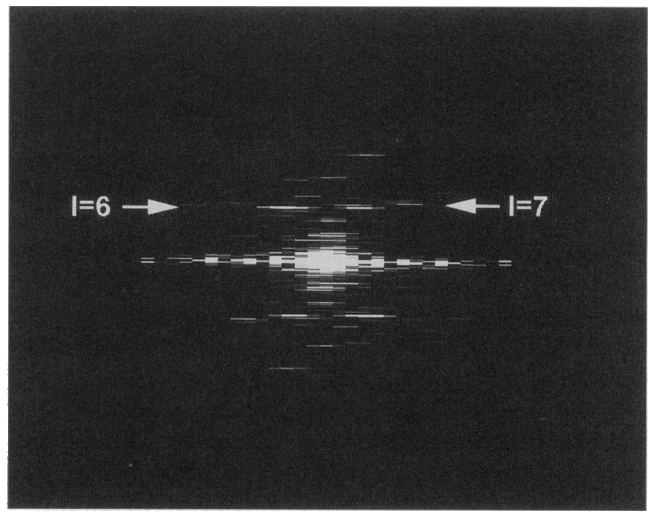

FIGURE 3 A computed Fourier transform of a single filament of the *T. aquaticus* RecA protein on linear dsDNA in the presence of ATP-γ-S (as in Fig. 1 *a*). The strongest layer line, labeled $l = 6$, arises from the approximately 90-Å pitch, right-handed one-start helix. The layer line labeled as $l = 7$ arises from a left-handed 5-start helix. For purposes of simplicity, the layer line numbering scheme from the *E. coli* RecA protein has been used, even though the average ratio of the spacing of the "first" to the "sixth" layer line is closer to 1:9 (see Fig. 5).

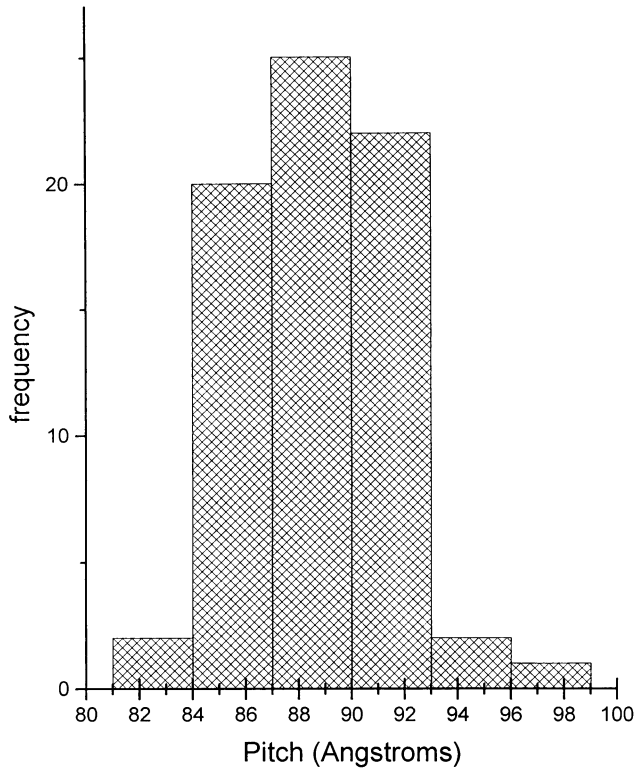

FIGURE 4 The pitch of the right-handed one-start helix of the filament formed by the *T. aquaticus* RecA protein on linear dsDNA in the presence of ATP-γ-S (as in Fig. 1 *a*). The pitch has been measured for 72 filaments using the position of the layer line labeled as $l = 6$ in Fig. 3. The mean is 88.6 Å ± 2.8 Å (SD). Tobacco mosaic virus particles coprepared with the filaments have been used to determine the absolute scale.

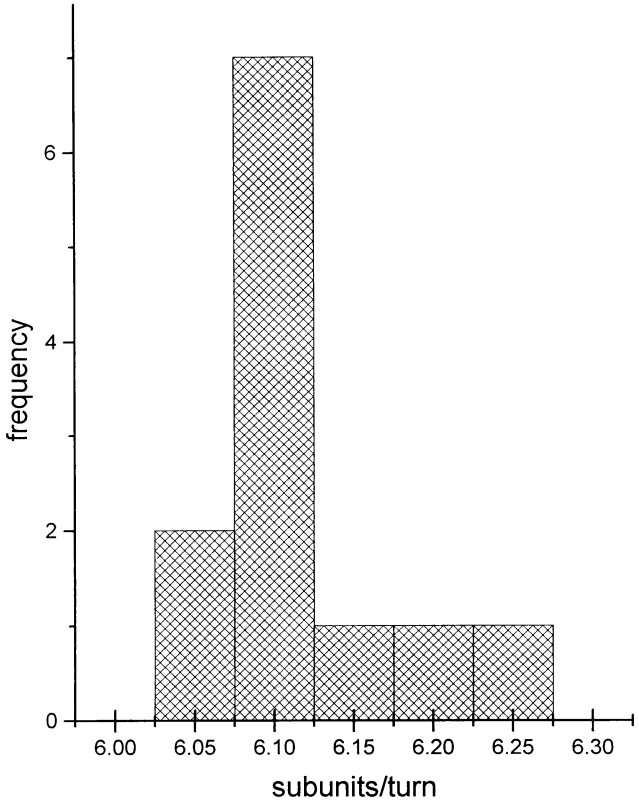

FIGURE 5 The "twist" of the filaments, measured in terms of the number of subunits per turn of the approximately 90-Å pitch, right-handed one-start helix. The values are shown for 12 filaments that were used for three-dimensional reconstruction. The number of subunits per turn was equal to $6.0 + Z_6/(Z_7 - Z_6)$, where Z_6 and Z_7 are the axial positions of the "sixth" and "seventh" layer lines, respectively. The mean is 6.11 ± 0.05 (SD).

associated with each layer line. Based upon the similarity of the protein to the *E. coli* RecA protein, we have used the same indexing, shown in Fig. 6. This assumption appears to be validated by subsequent results, including the three-dimensional reconstruction. Based upon this assumption, however, the histogram in Fig. 5 shows that the twist of the 12 filaments that were subsequently used for three-dimensional reconstruction has a mean of 6.11 Å ± 0.01 Å (SEM). In contrast, the *E. coli* RecA protein has a mean twist of about 6.17 units per turn when examined under similar conditions (Yu and Egelman, 1990). The measured twist and pitch for the *T. aquaticus* RecA protein filaments indicate that the average rise per subunit within the filament is about 14.5 Å. This is consistent with a stoichiometry of 3 base pairs per RecA subunit.

Layer lines were extracted from 12 of the best filaments, as judged by the symmetry shown in the Fourier transforms. The averaged layer line amplitudes and phases are shown in Fig. 6. A measure of the signal-to-noise ratio, as well as of the intrinsic polarity of the filament, is the comparison of average phase residuals between filaments aligned against an average using the "better" polarity versus the alignment using the "worse" polarity. The average "up-down" phase residuals for merging the layer lines from these 12 filaments were 38.0° for the "better" polarity and 66.6° for the "worse" polarity. These averaged layer lines were then used to generate a three-dimensional reconstruction by Fourier-Bessel inversion (DeRosier and Klug, 1968). The reconstruction is shown as a rendered surface (Fig. 7) and in cross sections (Fig. 8). The surface shown in Fig. 7 is very similar to that generated from the *E. coli* RecA protein (Yu and Egelman, 1990, 1993a): there is a very deep helical groove, with subunits arranged so that there is a relatively smooth side of the groove, and with pendulous lobes protruding into the groove on the opposite side. Analysis of filaments formed on ssDNA has shown that the smooth side is the 5′ end of the filament (Stasiak et al., 1988).

Bundles of inactive filaments

In the absence of both DNA and nucleotide cofactor, the *T. aquaticus* RecA protein can be seen in the EM to form several different aggregates, including bundles of filaments and rings (Fig. 1 *c*). There did not appear to be any correlation between the temperature at which the protein was incubated and the prevalence of these bundles and rings. Several different bundle forms were observed, but we chose to concentrate on the particular form shown in Fig. 1 *c*. The

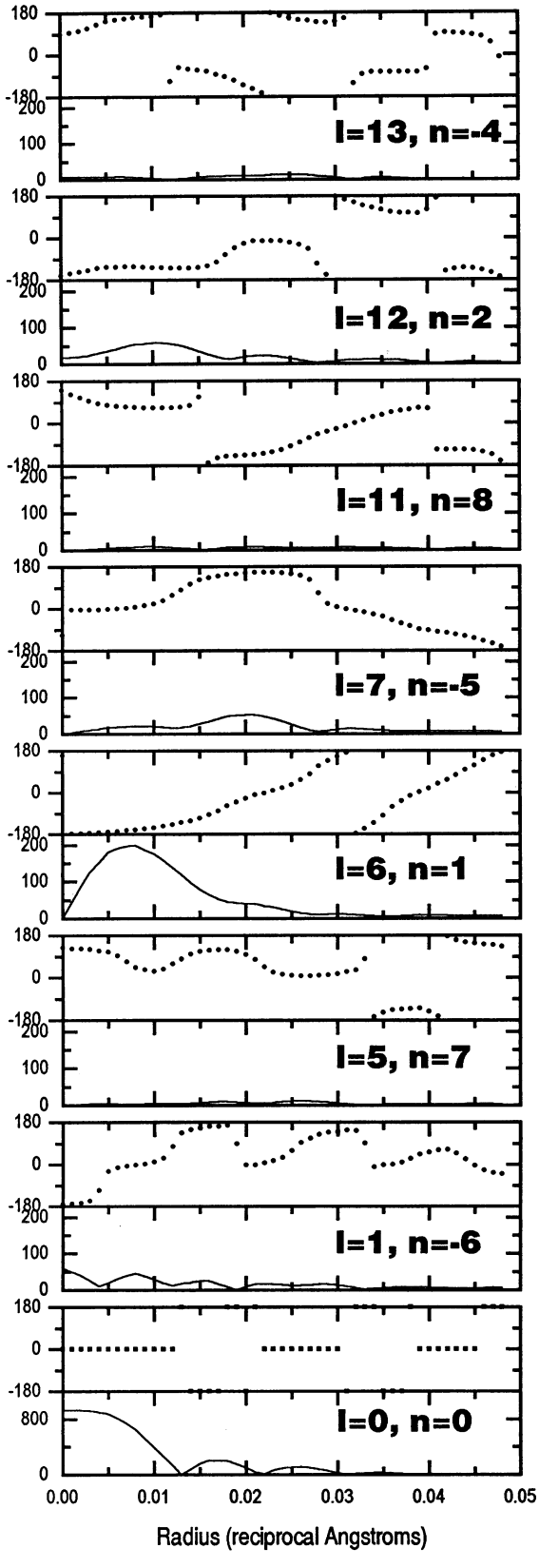

FIGURE 6 The average layer line amplitudes and phases obtained from 12 *T. aquaticus* RecA-dsDNA-ATP-γ-S filaments. Because the "near" and "far" side layer lines from each filament were first averaged together, the final average shown actually contains 24 independent data sets. All layer line amplitudes are on the same scale, with the exception of the equator.

Fourier transform of this type of bundle is shown in Fig. 9. The strong meridional reflection, labeled as $l = 9$, is at a spacing of about 1/50 Å in the transform. Examination of the phases indicated that all layer lines contained only even Bessel orders, and $l = 9$ and $l = 18$ were clearly meridional with $n = 0$. A Fourier-Bessel transform of only the equator revealed that the structure is annular in cross section (data not shown), with density contained between radii of about 110–220 Å. This made indexing the other layer lines quite simple, as the only scheme that yielded helical modulation between these bounds is shown in Fig. 10. The "hand" of this scheme is not determined by this constraint, as an enantiomorphic indexing would also be consistent. However, the n, l plot shown gives rise to component right-handed filaments in the reconstruction that are supercoiling in a left-handed sense, whereas the enantiomorphic scheme gives rise to component left-handed filaments that are supercoiling in a right-handed sense. Because all RecA (Egelman and Stasiak, 1986), Rad51 (Ogawa et al., 1993), and UvsX filaments (Griffith and Formosa, 1985) have been observed to be right-handed, the indexing scheme shown appears to be reasonable. Furthermore, all supercoiling of *E. coli* RecA filaments in bundles has been observed to be left-handed (Egelman and Stasiak, 1988), arising from the approximately 6.17 subunits per turn along the right-handed one-start helix of the filament.

Layer lines were extracted from images of 36 bundles, and the averaged layer line amplitudes and phases are shown in Fig. 11. The average "up-down" phase residuals for merging the layer lines from these 36 bundles were 52.7° for the "better" polarity and 76.5° for the "worse" polarity. These layer lines have been used to generate the reconstruction in Fig. 12 *b* (rendered surface) and Fig. 13, *c* and *d* (cross sections). Although the supercoiling of filaments in this bundle generates long-pitch four-start helices (with a pitch of about 1800 Å) that give rise to the near-equatorial layer lines ($l = 1$–4 in Fig. 10), the bundle in Figs. 12 and 13 clearly has 8 filaments. Thus, the asymmetric unit in this helical bundle contains two filaments, with a nonequivalence of these two filaments. If these filaments were structurally identical, and in identical environments, then the lowest layer lines would arise from eight-start helices. We will look further at the asymmetry between the filaments in the bundle.

Although the component filaments in this helical bundle are clearly right-handed helices themselves, it is evident that there are significant deformations of these individual filaments. For example, the arrows in Fig. 12 *b* indicate large contacts between filaments that appear to exist because of significant deformations of the filaments. Similarly, the cross section in Fig. 13 *d* appears to show very clear helical density belonging to the component filaments. However, if one looks at the cross section 25 Å away (one-half of the 50 Å rise per asymmetric unit) in Fig. 13 *c*, there is much less density present. If the component filaments were undeformed helices, then the filament density in Fig. 13 *c* should

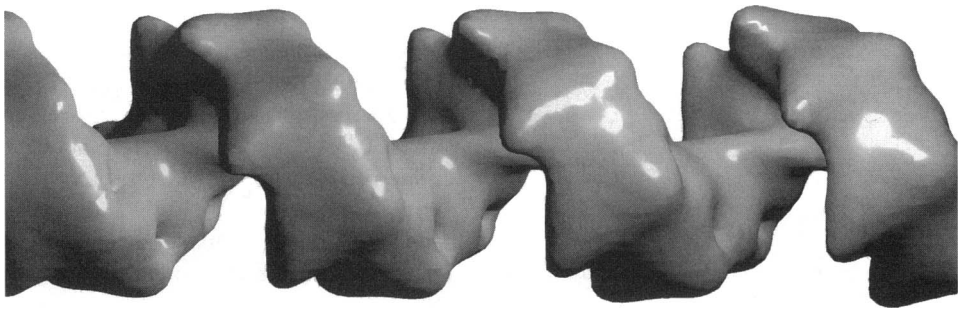

FIGURE 7 The surface of a three-dimensional reconstruction of the *T. aquaticus* RecA protein on linear dsDNA in the presence of ATP-γ-S. The reconstruction has been generated from the averaged layer lines of Fig. 6. The pitch of the right-handed one-start helix is 90 Å, and there are approximately 6.1 subunits per turn of this helix. If we assume that the structural polarity is the same as that of the *E. coli* RecA protein, the 3′ end of a filament formed on ssDNA would be on the left. The surface shown corresponds to a threshold chosen so that 100% of the expected molecular volume is enclosed.

appear similar to the filament density in Fig. 13 *d*, but with a rotation of the mass.

Because the helical reconstruction of the bundle did not involve any assumption about the helical symmetry of the component filaments, or even the assumption that the asymmetric unit contained filaments, we must consider whether limited resolution or other reconstruction artefacts could introduce some of the apparent deformations of the component filaments that are seen. We have therefore generated a model bundle (Figs. 12 *a* and 13, *a* and *b*) to explore this further. The model bundle was generated using eight component helical filaments, arranged with two filaments in the asymmetric unit. Within each filament there were six subunits per repeat. The model filaments had spherical subunits with a radius of 22 Å, placed in a helix with a radius of 35 Å. Each of the two filaments had a freedom of radial position and a rotational freedom. Models were simply compared against the actual average reconstruction to de-

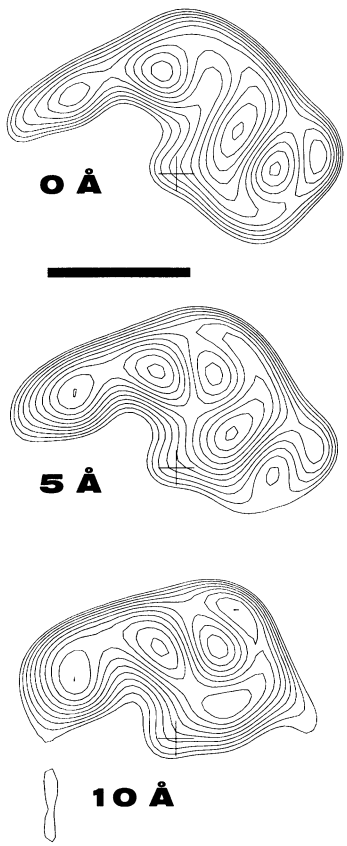

FIGURE 8 Three cross sections from the reconstruction of Fig. 7. Because the rise per subunit in the filament is approximately 15 Å, the next section at $z = 20$ Å would be nearly identical to the section at $z = 0$ Å, but rotated by 360°/6.1, or about 59°. The cross in the center of each section marks the filament axis. The outer contour is the same as that of the surface shown in Fig. 7, and the range between this density level and the highest density level in the electron density map has been divided into 10 equal steps. The scale bar is equal to 40 Å.

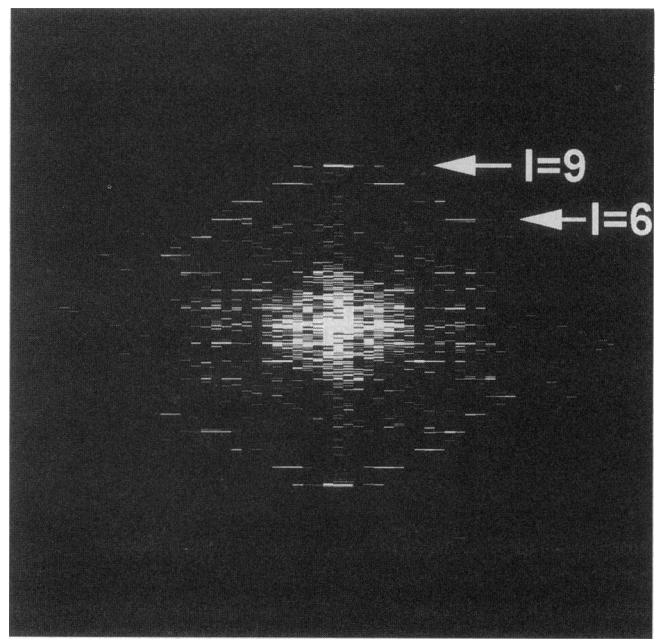

FIGURE 9 A computed Fourier transform of a bundle formed by the *T. aquaticus* RecA protein in the absence of both DNA and nucleotide cofactor (as in Fig. 1 *c*). The layer line indexing for this transform is shown in Fig. 10. The first meridional layer line, $l = 9$, is at a spacing of 1/50 Å. The higher order layer lines, such as $l = 18$ at 1/25 Å, are weaker and more difficult to see.

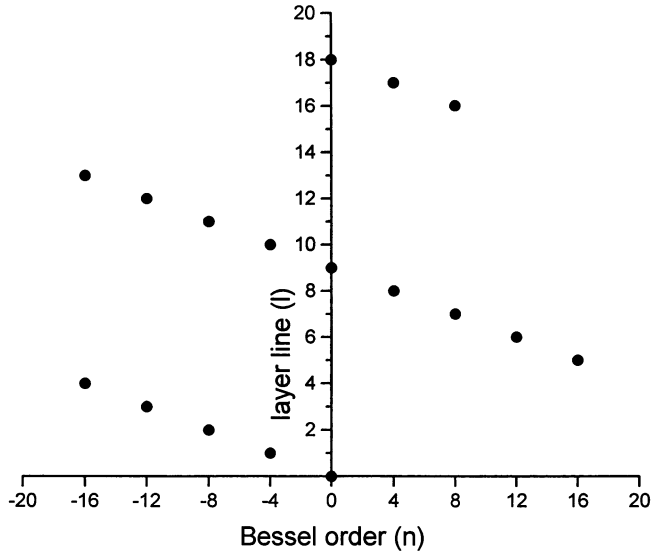

FIGURE 10 An n, l plot showing the indexing that has been used for the bundles (transform in Fig. 9). The first meridional layer line ($l = 9, n = 0$) is at a spacing of 1/50 Å.

termine the best values for these parameters. The best fit model had one filament at a radius of 169 Å, with the second filament at a radius of 138 Å. The supercoil pitch of the observed bundles, 1800 Å, was imposed upon these model filaments, and the symmetry of the bundle was used to generate the other six filaments. The model was projected onto two dimensions and reconstructed with the same layer lines used in Fig. 11. It can be seen in both Figs. 12 *a* and 13, *a* and *b*, that when the model is reconstructed in a fashion similar to that of the actual bundles, the deviations from helical symmetry of the component filaments are small. Thus, the comparison in Fig. 12 between the model and the actual reconstruction highlights the extent to which the filaments are deformed (indicated by the arrows in Fig. 12 *b*). Similarly, the comparison between the cross sections of Fig. 13 *a* (model) and 13 *c* (actual bundle) indicate the extent to which mass is shifted axially in the actual reconstruction. In fact, the strong meridional layer line intensity on $l = 9$ in Fig. 9 arises from this axial deformation of mass, because if the component filaments in the bundle had helical symmetry, the lowest resolution meridional intensity should appear at the reciprocal of the axial rise per subunit, or at about ⅛ Å.

An assumption has been made in generating the model bundle that there are six subunits per repeat in each filament. We have shown above that for the *T. aquaticus* protein there are about 6.1 subunits per turn in the active RecA–dsDNA–ATP-γ-S filament. For the *E. coli* RecA protein, both the active (with ATP) and inactive (without ATP) forms of the filament have about six subunits per turn (Yu and Egelman, 1992). The modulation of the density that can be seen in the reconstruction of the bundle (Fig. 13, *c* and *d*) is also consistent with six subunits per repeat, even though this assumption has not been used in generating the reconstruction. If there are indeed six subunits per filament per repeat, then the supercoil pitch can be used as a measure of the number of subunits per turn in "free" filaments. In one supercoil turn (1800 Å) there are about 36 asymmetric units (50 Å). There are therefore about 216 RecA subunits in one filament in one supercoil (6 × 36). In this distance, each filament contains 36 − 1 turns, because each right-handed filament has made one left-handed supercoil turn. The "twist" of the component filaments is thus about 6.17 subunits/turn (or 216 subunits/35 turns).

Although the axial rise of the asymmetric unit in the bundles is about 50 Å, a consequence of the non-zero supercoil radius and the finite supercoil pitch is that the helical turn within the component filaments will be greater than this. In one supercoil turn, 36 repeats within a filament (at a radius within the bundle of r_{fil}) will have a path length equal to p, where

$$p = (1800^2 + (2\pi r_{fil})^2)^{1/2}$$

Thus, the filaments within the model bundle at 169 Å radius will have a repeat of 58 Å along the filament path, whereas the filaments at a radius of 138 Å will have a repeat of 55 Å along the filament path. Although these distances are greater than 50 Å, they still represent a considerable compression from the active filament with a mean pitch of about 88 Å.

Hexameric rings

In the presence of DNA and ATP-γ-S (Fig. 1, *a* and *b*), or in the absence of both DNA and nucleotide cofactor (Fig. 1 *c*), the *T. aquaticus* RecA protein forms ring-like structures. We have used single-particle averaging to generate an alignment of 1000 images of these rings (Fig. 14 *a*). The average displays a very strong sixfold symmetry, and we have imposed this sixfold symmetry in the average shown in Fig. 14 *b*. The average is about 130 Å in diameter, with a hole in the center that is about 30 Å in diameter.

DISCUSSION

We have shown that the active filament formed by the *T. aquaticus* RecA protein on DNA in the presence of ATP-γ-S is remarkably similar to the filament formed by the *E. coli* RecA protein. This similarity extends to nearly identical parameters for the stretching and untwisting of the DNA, caused by the imposition of the RecA helical symmetry on the DNA. However, whereas the *E. coli* RecA filament is formed in vitro by incubations of the protein with DNA and nucleotide cofactor at 37°C, we have used incubations at 65°C for the *T. aquaticus* protein. Although no technique currently exists for examining hydrated samples at these temperatures in the electron microscope, it is reasonable to assume that the filaments formed at these higher temperatures are not greatly altered as they are prepared for electron microscopy.

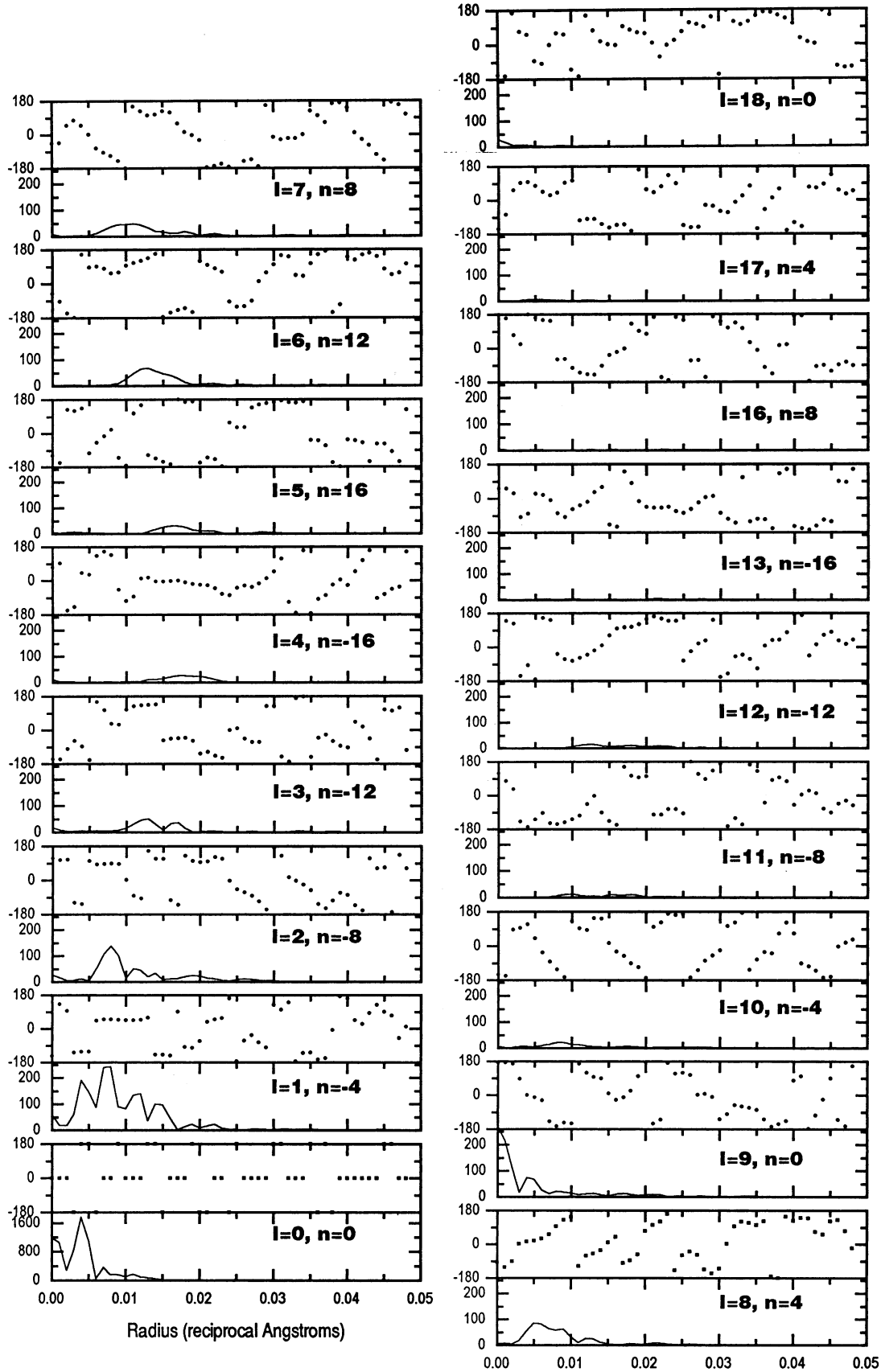

FIGURE 11 The averaged layer line amplitudes and phases from 36 bundles. Because the near and far side layer lines from each bundle have first been averaged together, the final average shown actually contains 72 independent data sets. All layer line amplitudes are shown on the same scale, with the exception of the equator.

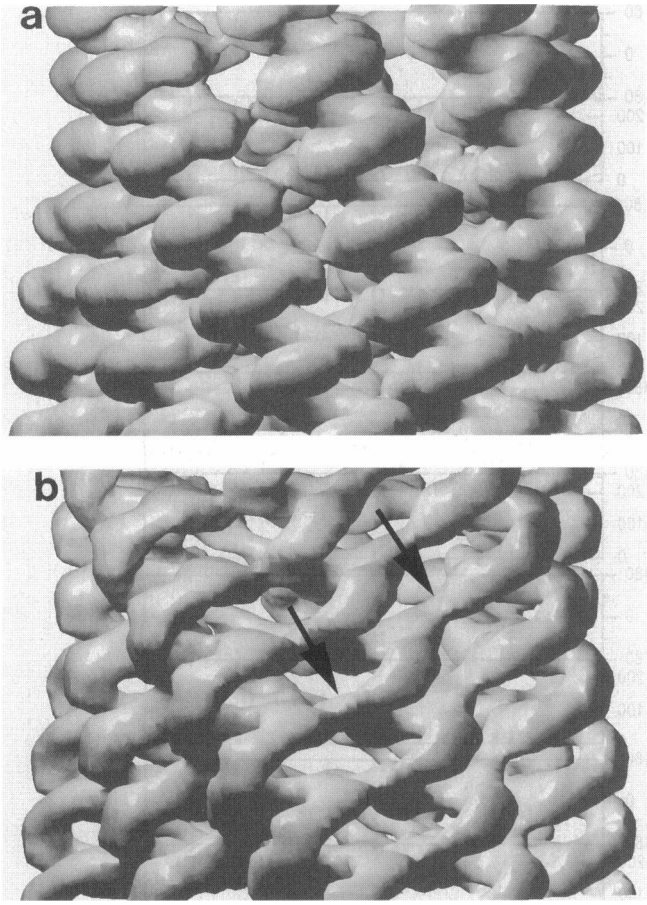

FIGURE 12 The surface of a reconstruction of a model bundle (*a*) and an actual bundle (*b*) of the *T. aquaticus* RecA protein. The model bundle has been generated using component filaments that have perfect helical symmetry. The reconstruction of the bundle makes no assumption about the helical nature of the asymmetric unit (a 50-Å-thick section from each of two adjacent filaments), but the relatively undeformed helical turns of the model filaments can be seen quite clearly. In contrast, the reconstruction at the same resolution of the real bundle (*b*), from the layer lines of Fig. 11, reveals significant distortions of the component filaments. The lower arrow on the left indicates where one side of a helical turn has been pushed down, making a bridge with an adjacent filament on the left, and the upper arrow on the right indicates where the other side of this same helical turn has been pushed up, making a bridge with an adjacent filament on the right.

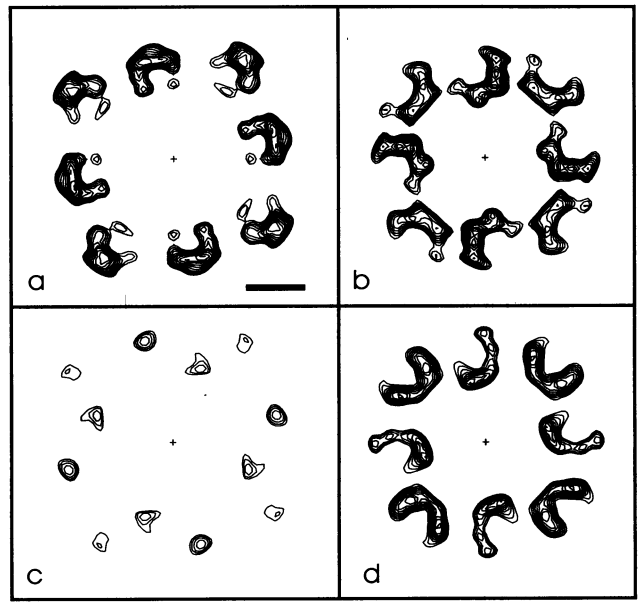

FIGURE 13 Cross sections from the reconstructions of a model bundle (*a, b*) and from the actual bundle (*c, d*). The sections *a* and *c* are from an axial level $z = 0$ Å, and the sections *b* and *d* are from half an asymmetric unit higher, $z = 25$ Å. The scale bar in *a* is 100 Å, and the cross in each section marks the position of the helical axis. It can be seen that at one axial position (*b* and *d*) there is a quite good match between the model and the actual bundle, whereas half an asymmetric unit away (*a* and *c*) the match is quite poor. The failure to see the cross sections of helical component filaments in (*c*) arises from the deformations indicated by the arrows in Fig. 12.

One argument in support of this assumption is that filaments formed at 37°C have a shorter pitch than those formed at 65°C, in addition to the fact that the total extent of polymerization is smaller at the lower temperature. At this temperature, the *T. aquaticus* RecA protein has limited ability to bind to ssDNA as measured by a filter-binding assay (Angov and Camerini-Otero, 1994). There are two possible interpretations for the shorter pitch seen at 37°C. In the absence of nucleotide cofactor, the *E. coli* RecA protein forms an inactive filament that has a significantly smaller pitch (60–75 Å) than that formed in the presence of nucleotide cofactor (90–95 Å) (Yu and Egelman, 1992). Thus, it is possible that the 82 Å mean pitch observed at 37°C corresponds to some intermediate state between the active and inactive conformations. Alternatively, the shorter pitch could result from incomplete saturation of the DNA, resulting in a less than stoichiometric complex. We have no data to distinguish between these possibilities. However, both possibilities indicate that fully saturated, extended filaments do not form at 37°C. We have previously shown for the *E. coli* protein that unstained, fully hydrated RecA–DNA–ATP-γ-S filaments that were cryo-preserved and imaged with cryo-electron microscopy were very similar to negatively stained filaments imaged by conventional electron microscopy (Yu and Egelman, 1992). Thus, it is unlikely that the *T. aquaticus* RecA protein imposes different struc-

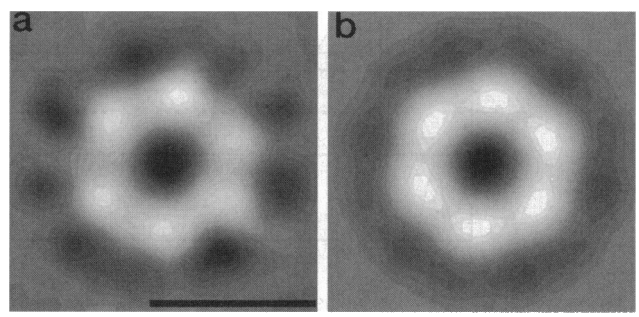

FIGURE 14 A reference-free alignment (Penczek et al., 1992) has been used to average together 1000 images of the rings formed by the *T. aquaticus* RecA protein (*a*). A sixfold symmetry has been imposed upon the average in *a* to produce the image in *b*. The space bar in *a* is 100 Å.

tural parameters on the DNA at 65°C, but that these parameters make a transition to the identical parameters observed for the *E. coli* RecA protein when specimens are prepared for electron microscopy, while the *T. aquaticus* RecA protein incubated at 37° fails to form such structures.

What is the significance of the conservation of these structural parameters? It has been shown that the yeast (Ogawa et al., 1993) and human (Benson et al., 1994) Rad51 proteins and the T4 UvsX protein (Yu and Egelman, 1993a) induce similar extensions and untwisting of the DNA bound within the "presynaptic" complex. Under conditions where the DNA is near the transition for thermal melting, it is very interesting that the *T. aquaticus* RecA protein induces an almost identical structure. This suggests that the configuration of the DNA within this complex may be an important element that has led to the conservation of the structural parameters of the helical protein scaffolding from eubacteria to humans. Although we do not yet have a detailed model for how homologous recognition works, despite new insights (Camerini-Otero and Hsieh, 1993; Rao et al., 1995), it is likely that the particular structure of DNA maintained by these filaments may be an important element in the mechanism.

Two other forms of aggregation of the *T. aquaticus* RecA protein have been examined in this paper, bundles of compressed filaments and rings. In the absence of ATP, the *E. coli* RecA protein forms compressed filaments, with a pitch of 60–75 Å, that can form on ssDNA or can exist as self-polymers (Williams and Spengler, 1986; Yu and Egelman, 1992). It has been suggested that bundles of these self-polymers are a cellular storage form for the inactive RecA protein (Story et al., 1992). One basis for this suggestion was the observation that mutations that appear at the interface between filaments in the RecA crystal can lead to constitutive induction of RecA's coprotease activity, presumably by disrupting such bundles in the cell and allowing RecA to inappropriately polymerize and become active (Story et al., 1992). Additional mutations have since been found that occur at this inter-filament contact region, and their phenotypes are consistent with this suggestion (Liu et al., 1993). This suggestion is also consistent with the fact that overexpressed RecA within *E. coli* can be found in such compressed filaments within inclusion bodies (Ruigrok et al., 1993).

An electron microscopic study (Yu and Egelman, 1992) has suggested that the inactive *E. coli* RecA filaments formed in the absence of nucleotide cofactor are in a conformation similar to that of the 6_1 filaments that exist within the RecA crystal. Thus, the filament-filament contacts within the RecA crystal might be a good model for the contacts that might occur in a cell between inactive filaments. Are the filament-filament contacts within the *T. aquaticus* RecA bundle described here the same? There are actually two classes of contacts within our bundles. That is, there are two different filament environments, with half of the filaments in one environment and half in the other. One set of inter-filament contacts appears to involve only a translation between two filaments, with no rotation, and this would be the symmetry relation found in the $P6_1$ RecA crystals (Story et al., 1992). However, we do not have the resolution to judge whether the actual contacts, and not just the symmetry, are the same. The second inter-filament relation must be completely different from that found in the RecA crystal, because it involves a large rotation between two adjacent filaments. We do not know at this point whether either or both sets of contacts are involved in the putative storage form in *T. aquaticus*, but the existence of two different sets of contacts will make the interpretation of phenotypic effects of mutations in surface residues (Liu et al., 1993) more complicated.

Rings of the *E. coli* RecA protein have been previously observed by electron microscopy (Brenner et al., 1988; Heuser and Griffith, 1989). Although these have been described as being mainly heptamers (Heuser and Griffith, 1989), they have been thought to be the same self-associated aggregation state that has been characterized as hexameric by ultracentrifugation analysis (Brenner et al., 1990) and by light scattering (Benight et al., 1991). It has been unknown whether these rings were actual intermediates in the formation of the active helical polymer on DNA, or a totally different state. Because there are approximately six subunits per turn of the RecA helix, in both the active and inactive filaments, it is suggestive that the *T. aquaticus* rings contain six subunits. However, in one turn of the helix there is a rise along the axis of either about 89 Å for the active *T. aquaticus* RecA filament or about 55–60 Å for the inactive *T. aquaticus* RecA, as seen within the bundles in this paper. Although we have not generated a three-dimensional reconstruction of the rings, it is a reasonable assumption that they are a planar structure. If they were one turn of the helix, there is no reason why they would exist as a stable species and not continue to polymerize. If they were slightly more compressed than one turn, so that there was a steric hindrance to adding additional subunits, it is expected that a large discontinuity would be evident in the images. However, an alignment of 1000 images yielded a structure with a very good sixfold symmetry, with no visible discontinuity. It is, of course, possible that given the limited resolution and poor signal-to-noise ratio that such a discontinuity would not be seen, but the failure to see such a discontinuity suggests that it cannot be large if it is present at all.

Therefore, the rings are most likely planar or very close to planar. This would suggest that the subunit-subunit contacts within the rings must be significantly different from the contacts within the filament: although both would have similar rotations between adjacent subunits, there would be very different axial shifts. It is hard to imagine how the filament contacts might be deformed in such a manner, suggesting that entirely different contacts are made within the rings. Although the active form of a number of ATP-binding helicases is a hexameric ring (Mastrangelo et al., 1989; Stasiak et al., 1994; Egelman et al., 1995), it seems unlikely that the RecA protein functions in this state in any of its activities. Clearly, more work will be needed to

understand the functional significance of this aggregation state.

This work was supported by National Institutes of Health grant GM35269 to EHE.

REFERENCES

Angov, E., and R. D. Camerini-Otero. 1994. The *recA* gene from the thermophile *Thermus aquaticus* YT-1: cloning, expression and characterization. *J. Bacteriol.* 176:1405–1412.

Benight, A. S., D. H. Wilson, D. M. Budzynski, and R. F. Goldstein. 1991. Dynamic light scattering investigations of RecA self-assembly and interactions with single strand DNA. *Biochimie.* 73:143–155.

Benson, F. E., A. Stasiak, and S. C. West. 1994. Purification and characterization of the human Rad51 protein, an analogue of *E. coli* RecA. *EMBO J.* 13:5764–5771.

Brenner, S. L., A. Zlotnick, and J. D. Griffith. 1988. RecA protein self-assembly. Multiple discrete aggregation states. *J. Mol. Biol.* 204: 959–972.

Brenner, S. L., A. Zlotnick, and W. F. I. Stafford. 1990. RecA protein self-assembly. II. Analytical ultracentrifugation studies of the entropy-driven self-association of RecA. *J. Mol. Biol.* 216:949–964.

Camerini-Otero, R. D., and P. Hsieh. 1993. Parallel DNA triplexes, homologous recombination, and other homology-dependent DNA interactions. *Cell.* 73:217–223.

DeRosier, D. J., and A. Klug. 1968. Reconstruction of three-dimensional structures from electron micrographs. *Nature.* 217:130–134.

Egelman, E. H. 1986. An algorithm for straightening images of curved filamentous structures. *Ultramicroscopy.* 19:367–373.

Egelman, E. H. 1993. What do x-ray crystallographic and electron microscopic structural studies of the RecA protein tell us about recombination? *Curr. Opin. Struct. Biol.* 3:189–197.

Egelman, E. H., and A. Stasiak. 1986. The structure of helical RecA-DNA complexes. I. Complexes formed in the presence of ATP-γ-S or ATP. *J. Mol. Biol.* 191:677–697.

Egelman, E. H., and A. Stasiak. 1988. Structure of helical RecA-DNA complexes. II. Local conformational changes visualized in bundles of RecA-ATP-γ-s filaments. *J. Mol. Biol.* 200:329–349.

Egelman, E. H., X. Yu, R. Wild, M. M. Hingorani, and S. S. Patel. 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.

Frank, J., B. Shimkin, and H. Dowse. 1981. SPIDER—a modular software system for electron image processing. *Ultramicroscopy.* 6:343–358.

Griffith, J., and T. Formosa. 1985. The UvsX protein of bacteriophage T4 arranges single-stranded and double-stranded DNA into similar helical nucleoprotein filaments. *J. Biol. Chem.* 260:4484–4491.

Heuser, J., and J. Griffith. 1989. Visualization of RecA protein and its complexes with DNA by quick-freeze/deep-etch electron microscopy. *J. Mol. Biol.* 210:473–484.

Kowalczykowski, S. C., and A. K. Eggleston. 1994. Homologous pairing and DNA strand-exchange proteins. *Annu. Rev. Biochem.* 63:991–1043.

Lee, J. W., and M. M. Cox. 1990a. Inhibition of recA protein promoted ATP hydrolysis. 1. ATP-γ-S and ADP are antagonistic inhibitors. *Biochemistry.* 29:7666–7676.

Lee, J. W., and M. M. Cox. 1990b. Inhibition of recA protein promoted ATP hydrolysis. 2. Longitudinal assembly and disassembly of recA protein filaments mediated by ATP and ADP. *Biochemistry.* 29: 7677–7683.

Liu, S.-K., J. A. Eisen, P. C. Hanawalt, and I. Tessman. 1993. *recA* mutations that reduce the constitutive coprotease activity of the RecA1202(Prt^c) protein: possible involvement of interfilament association in proteolytic and recombination activities. *J. Bacteriol.* 175: 6518–6529.

Mastrangelo, I. A., P. V. C. Hough, J. S. Wall, M. Dobson, F. B. Dean, and J. Hurwitz. 1989. ATP-dependent assembly of double hexamers of SV40 T antigen at the viral origin of DNA replication. *Nature.* 338:658–662.

Ogawa, T., X. Yu, A. Shinohara, and E. H. Egelman. 1993. Similarity of the yeast RAD51 filament to the bacterial RecA filament. *Science.* 259:1896–1899.

Penczek, P., M. Radermacher, and J. Frank. 1992. Three-dimensional reconstruction of single particles embedded in ice. *Ultramicroscopy.* 40:33–53.

Rao, B. J., S. Chiu, L. R. Bazemore, G. Reddy, and C. M. Radding. 1995. How specific is the first recognition step of homologous recombination? *Trends Biochem. Sci.* 20:109–113.

Ruigrok, R. W., B. Bohrmann, E. Hewat, A. Engel, E. Kellenberger, and E. DiCapua. 1993. The inactive form of RecA protein: the "compact" structure. *EMBO J.* 12:9–16.

Saenger, W. 1984. Principles of Nucleic Acid Structure. Springer-Verlag, New York.

Shinohara, A., H. Ogawa, Y. Matsuda, K. Ikeo, N. Ushio, and T. Ogawa. 1993. Cloning of human, mouse and fission yeast recombination genes homologous to RAD51 and recA. *Nature Genet.* 4:239–243.

Shinohara, A., H. Ogawa, and T. Ogawa. 1992. Rad51 protein involved in repair and recombination in *S. cerevisiae* is a RecA-like protein. *Cell.* 69:457–470.

Stasiak, A., and E. DiCapua. 1982. The helicity of DNA in complexes with RecA protein. *Nature.* 229:185–186.

Stasiak, A., E. DiCapua, and T. Koller. 1981. Elongation of duplex DNA by RecA protein. *J. Mol. Biol.* 151:557–564.

Stasiak, A., E. H. Egelman, and P. Howard-Flanders. 1988. Structure of helical RecA-DNA complexes. III. The structural polarity of RecA filaments and functional polarity in the RecA-mediated strand exchange reaction. *J. Mol. Biol.* 202:659–662.

Stasiak, A., I. R. Tsaneva, S. C. West, C. J. B. Benson, X. Yu, and E. H. Egelman. 1994. The *Escherichia coli* RuvB branch migration protein forms double hexameric rings around DNA. *Proc. Natl. Acad. Sci. USA.* 91:7618–7622.

Story, R. M., I. T. Weber, and T. A. Steitz. 1992. The structure of the *E. coli* RecA protein monomer and polymer. *Nature.* 355:318–325.

Voet, D., and J. G. Voet. 1990. Biochemistry. Wiley, New York. 805–806.

Weinstock, G. M., K. McEntee, and I. R. Lehman. 1981. Interaction of the recA protein of *Escherichia coli* with adenosine 5′-O-(3-thiotriphosphate). *J. Biol. Chem.* 256:8850–8855.

West, S. C. 1992. Enzymes and molecular mechanisms of genetic recombination. *Annu. Rev. Biochem.* 61:603–640.

Williams, R. C., and S. J. Spengler. 1986. Fibers of RecA protein and complexes of RecA protein and single stranded φX174 DNA as visualized by negative-stain electron microscopy. *J. Mol. Biol.* 187: 109–118.

Yu, X., and E. H. Egelman. 1990. Image analysis reveals that the *E. coli* RecA protein consists of two domains. *Biophys. J.* 57:555–566.

Yu, X., and E. H. Egelman. 1992. Structural data suggest that the active and inactive forms of the RecA filament are not simply interconvertible. *J. Mol. Biol.* 227:334–346.

Yu, X., and E. H. Egelman. 1993a. The LexA repressor binds within the deep helical groove of the activated RecA filament. *J. Mol. Biol.* 231: 29–40.

Yu, X., and E. H. Egelman. 1993b. 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.