 0014-4754/94/030192-12$1.50 + 0.20

Structure and function of RecA-DNA complexes

A. Stasiak and E. H. Egelman[a]

Laboratoire d'Analyse Ultrastructurale, Bâtiment de Biologie, Université de Lausanne, CH-1015 Lausanne (Switzerland) and [a]Department of Cell Biology and Neuroanatomy, University of Minnesota Medical School, 321 Church St. SE, Minneapolis (Minnesota 55455, USA)

Abstract. While the *E. coli* RecA protein has been the most intensively studied enzyme of homologous recombination, the unusual RecA-DNA filament has stood alone until very recently. It now appears that this protein is part of a universal family that spans all of biology, and the filament that is formed by the protein on DNA is a universal structure. With RecA's role in recombination given new and greatly increased significance, we focus in this review on the energetics of the RecA-mediated strand exchange and the relation between the energetics and recombination spanning heterologous inserts.
Key words. RecA; RecA-DNA complexes; homologous recombination; Rad51 protein; electron microscopy; image analysis.

Introduction

The *RecA* gene of *E. coli* was first identified over 28 years ago[7] and the RecA protein has been the object of intensive genetic, biochemical and structural studies for the past 15 years. These studies have led to a model for RecA function in which this 38kDa molecule polymerizes on either a single- or double-stranded DNA molecule to form a helical nucleoprotein filament. Within this filament, the DNA is greatly stretched and untwisted from its native form: both single and double-stranded DNA are stretched to about a 5.1 Å rise per base or base pair[48], and double-standed DNA is untwisted to about 18.6 base pairs per turn[46]. This picture of the active RecA filament has been the core of the 'RecA paradigm'. However, until quite recently, there has never been any evidence that this remarkable nucleoprotein filament would have any relevance to recombination in organisms other than bacteria. This situation has now dramatically changed.

Electron microscopic observations have now established that the yeast Rad51 protein[35] and the bacteriophage T4 UvsX protein[58] both form nucleoprotein filaments that are nearly identical to those formed by the RecA protein (see fig. 1). Further, the human[44], mouse[44] and chicken homologs[2] of the yeast Rad51 protein have been found, and the sequence comparisons strongly suggest that the filament formed by the yeast protein will be formed by the mammalian and avian analogs (see also Heyer in this issue).

How does this affect our picture of the RecA protein? For one thing, it shows that the formation of similar helical nucleoprotein filaments is the key determinant of biological similarity for this class of molecules. For example, the T4 UvsX protein is so weakly homologous in primary

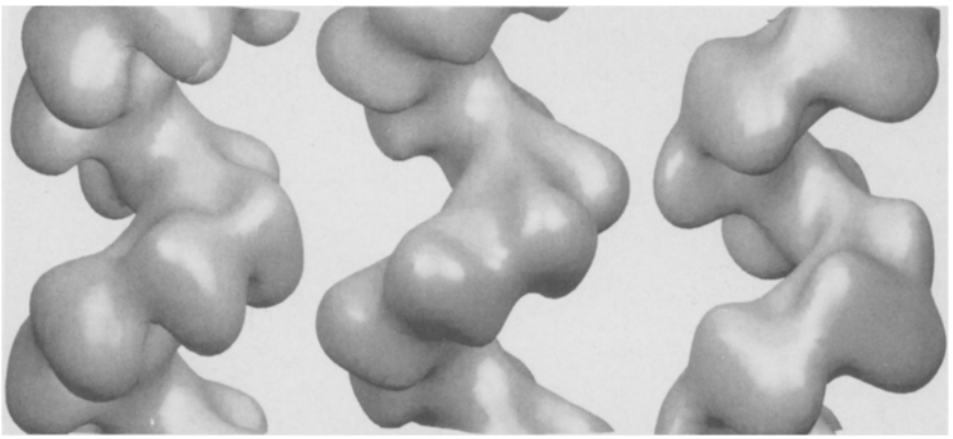


Figure 1. Comparison of helical nucleoprotein filaments formed by *E. coli* RecA (left), *S. cerevisiae* Rad51 (centre) and the bacteriophage T4 UvsX protein (right) on ds DNA. The rendered surfaces are from three-dimensional reconstructions of electron micrographs. The RecA reconstruction is from Yu and Egelman[57], the Rad51 reconstruction is from Ogawa et al.[35], and the UvsX reconstruction is from Yu and Egelman (manuscript in preparation). The Rad51 protein lacks approximately 90 C-terminal residues that are present in RecA and UvsX, and this corresponds to the lack of subunit lobes in the Rad51 reconstruction.

sequence to the RecA protein that two different published alignments between RecA and UvsX are very different[19,41]. The finding that the weakly related sequences form nearly identical three-dimensional structures[58], suggests that structure is conserved more strongly than sequence, but this is not a novel finding. Most importantly, these results suggest that the particular features of the RecA-DNA filament, including the unwinding and stretching of the DNA, may be the main force in the conservation of the structure of these filaments between prokaryotes and humans.

Can we now say that it was inevitable that RecA-like structures were to be found in eukaryotes? We think not. There are many structures and mechanisms that exist in prokaryotes that have no analog in eukaryotes, and vice versa. For example, bacterial motility is extremely interesting, but we know that the prokaryotic flagellum, based upon the single protein flagellin, has no apparent relationship to any eukaryotic systems based upon actin, myosin, tubulin, dynein or other eukaryotic proteins. Both the structures and mechanisms of motility appear to be totally unrelated. The fact that RecA-like filaments are found in eukaryotes, however, suggests that particular features of this complex must be conserved far more strongly than motility across evolution. It also may suggest that unlike motility, where many different mechanisms may exist for mechanochemical transduction, the recognition and exchange of DNA strands between homologous molecules may be quite limited in the number of different paths available. Now that the RecA filament has received its apotheosis, in the form of being elevated from a prokaryotic curiosity to a universal biological structure, more attention will be paid to this filament, since it appears that what we learn about RecA may be directly applicable to repair, recombination and neoplastic transformation in humans. We will not attempt to summarize the biochemistry and genetics of RecA, the focus of several recent reviews[29,37,41,55]. We will concentrate instead on three questions: the relation between the structure of RecA-DNA complexes and the RecA crystal, the function of ATP hydrolysis in RecA reactions, and how RecA exchanges strands between partially heterologous DNA molecules.

The relation of the RecA crystal to the structure of functional RecA-DNA complexes

Most of our knowledge about the structure of the active RecA-DNA-ATP filament has come from electron microscopy. We will not attempt to review what is known about RecA structure, since that is the subject of two recent articles[15,17]. The most important new development has been the solution of a crystal structure for the RecA protein[51]. However, questions about the interpretation of the functional state of RecA in the crystal exist, since it has been shown that the RecA filament in the crystals is much more similar to the inactive RecA filament that forms in the absence of a nucleotide cofactor, with or without single-stranded DNA, than it is to the active ternary complex of RecA, single- or double-stranded DNA, and ATP[56]. Yu and Egelman[56] have also shown that the inactive RecA filament appears to bind single-stranded DNA with a different stoichiometry (5 bases per RecA) than the active filament (3 bases per RecA), which may help explain why these two different filament forms are not interconvertible[30].

Nevertheless, the crystal structure provides the first atomic-level model for RecA structure, and will serve as a starting point for understanding the conformational changes that occur in RecA as a result of the binding of ATP and DNA. An additional limitation of the crystal structure is that the putative DNA-binding sites are not seen, presumably due to the disorder of two loops. It has been suggested that the first loop in the sequence, L1 is the binding site for the second DNA molecule, while the second loop, L2, is the primary binding site for the DNA molecule that becomes ensconced in the presynaptic filament[51]. Supporting evidence of this assignment has come from studies of a co-complex of the RecA-DNA-ATP filament with a non-cleavable mutant LexA repressor[57]. The co-complex shows that the LexA repressor binds deep within the helical groove of the RecA filament, but also shows that one of the two main regions of contact between the LexA repressor and RecA maps quite well onto loop L1 of the RecA crystal structure, the putative secondary DNA binding site. This can explain observations that excess single-stranded DNA can inhibit the RecA-mediated cleavage of the LexA repressor[13,52], since both DNA and LexA appear to be competing for the same or neighboring binding sites, and supports the DNA-binding assignment of Story et al.[51].

Role of ATP hydrolysis in the RecA-mediated recombination reaction

The earliest in vitro studies of RecA showed that an ATPase activity accompanies the ongoing strand reaction[9]. It was demonstrated that the RecA ATPase is induced when RecA forms complexes with DNA[34,40]. Since the RecA-mediated strand exchange reaction is a multistage process, it was important to determine which stage of the reaction requires the energy of ATP hydrolysis. A prerequisite for the strand exchange reaction is the formation of a so-called presynaptic complex with a single-stranded DNA (or double-stranded DNA molecules containing single-stranded gaps). The DNA molecules in the presynaptic complexes are maintained in a configuration ready for homologous recognition. The bases of these DNA molecules are exposed for contacts with protein-free double-stranded DNA molecules approaching the presynaptic filaments[12]. Var-

ious electron microscopic studies demonstrated that the structure of the presynaptic complex was essentially the same when ATP was replaced by the very slowly hydrolyzable analog, ATP-γ-S[16,18,20]. In both cases, the complexes were extended, deeply grooved helical filaments. However, complexes which were formed in the absence of a nucleotide cofactor did not show these features and were in the form of compact filaments[14]. Using morphological criteria it was therefore possible to conclude that while a nucleotide cofactor is needed, ATP hydrolysis is not required for the formation of the extended presynaptic filaments. Once presynaptic filaments are assembled, the next step of the reaction is the search for homology. In 1985 two independent papers showed that presynaptic filaments formed in the presence of ATP-γ-S have the selective ability to bind homologous duplex DNA molecules[20,39]. It was therefore demonstrated that the process of homologous recognition does not require the energy obtained from ATP hydrolysis. The same two papers concluded, though, that the subsequent steps of the strand exchange reaction, i.e., the actual formation and extension of the newly formed heteroduplex regions, do require ATP hydrolysis.

Five years later proper conditions were found for the in vitro strand exchange reaction to be performed in the presence of ATP-γ-S[31,42]. Studying such reactions occurring between full size linear double-stranded and circular single-stranded M13 DNA molecules, Menetski et al.[31] observed joint molecules with heteroduplex regions extending for more than three thousand base pairs. However, despite extensive heteroduplex formation, no complete strand exchange was seen in these reactions. This raised the possibility that the long heteroduplex regions observed in part of the molecules resulted from a spontaneous branch migration between fully homologous DNA strands. To produce a significant proportion of complete products of the strand exchange reaction (nicked circular dsM13 and linear ssM13), one would need to move the branch point by more than 7 kilobases (from one end to the other of linear M13 DNA). A random bidirectional process of spontaneous branch migration should be rather ineffective in pushing branch migration through long stretches[23] and that could be why Menetski et al.[31] did not observe the final products of the reaction.

In 1981 Cox and Lehman[9] proposed that the function of ATP hydrolysis was to drive a directional branch migration, converting short and unstable heteroduplex regions that would initially form into stable products of the strand exchange reaction. Therefore, it was important to determine if the RecA-mediated strand exchange occurring in the absence of ATP hydrolysis can have a preferred polarity in converting substrates into products of the reaction. Rosselli and Stasiak[42] showed that the polarity of the RecA-mediated strand exchange is maintained when ATP is replaced with ATP-γ-S. In the same paper using short, palindrome-free oligonucleotides (to avoid trapping secondary structure of ssDNA within presynaptic filaments formed in the presence of ATP-γ-S) it was also demonstrated that in these reactions almost 100% of the substrate molecules can be converted into products of the strand exchange reaction. This means that the extent of strand exchange reactions in the absence of ATP hydrolysis is not determined by the thermodynamic equilibrium between substrates and products. Therefore, RecA protein in the absence of ATP hydrolysis has the ability to push the reaction in one direction. What kind of energy can be used by RecA to achieve this?

Analysis of the in vitro strand exchange reaction performed in the presence of ATP-γ-S revealed that RecA protein becomes tightly bound to the products of the reaction[42]. Thus, RecA-DNA binding energy can be utilized for the complete conversion of substrates into products of the reaction. Figure 2 illustrates how the DNA binding energy can be used to drive sequential steps of the in vitro recombination reaction. The self assembly of presynaptic complexes is a simple consequence of a high affinity of RecA-nucleoside triphosphate molecules for ssDNA. According to this model, in the presynaptic complexes assembled on ssDNA a second binding site for another DNA molecule becomes available[22,33]. Homologous recognition can be then driven by protein-DNA and DNA-DNA binding energy. When the second DNA binding site is very close to the already bound ssDNA, then residues of the resident ssDNA can interact positively (e.g., hydrogen bonds) and negatively (e.g., steric clash) with the dsDNA molecule occupying the second binding site. Thus, binding of homologous DNA by the second binding site may be much stronger than the binding of unrelated sequences. The actual strand switch from substrate to product configuration can also be driven by the energy of protein-DNA interactions. If after the strand switch the configuration of DNA strands within the RecA-DNA complex is such that it allows a better fit and stronger binding between protein and DNA, then this will cause the strand exchange to proceed till the end of a homologous molecules for simple energetic reasons.

One of the consequences of such a model would be the increasing stability of RecA-DNA complexes during sequential stages of the reaction. Indeed, electron microscopy studies of ATP-γ-S complexes with post-exchange arrangements of DNA strands revealed that these complexes appeared more stable than complexes formed during earlier stages of the reaction[42]. In the absence of ATP hydrolysis the reaction stops at the stage where the product DNA molecules are tightly bound by RecA. Only upon hydrolyzing ATP can RecA dissociate from these stable complexes, releasing the DNA products of the reaction. After renewed binding

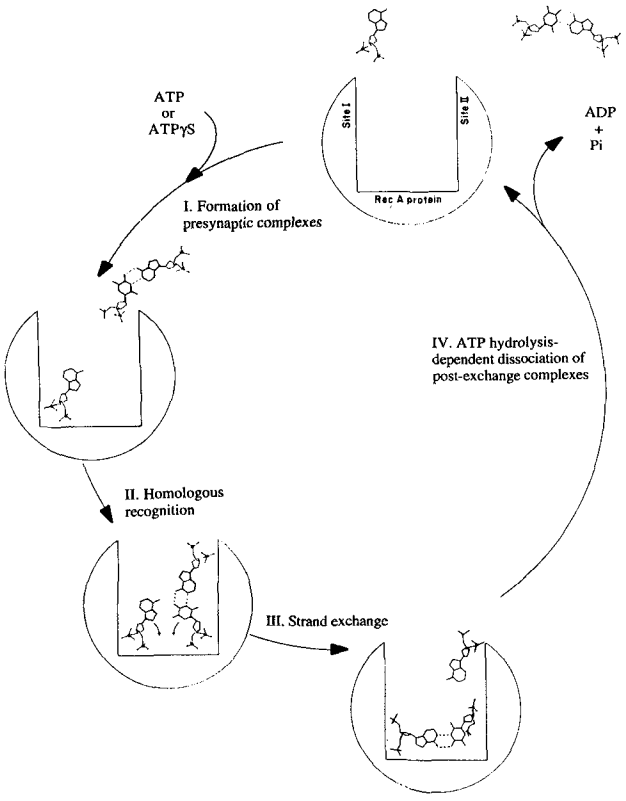


Figure 2. Cycle of the RecA reaction. In this schematic representation of the recombination reaction, DNA molecules and RecA-DNA complexes are seen in cross section. Only a short segment corresponding to a one base pair step is shown. At the beginning of the reaction, ATP or ATP-γ-S play the function of allosteric effectors required for the structural change in the protein that leads to the formation of the active presynaptic filament. In the presence of ATP or ATP-γ-S, the reaction can proceed through: I) formation of presynaptic complexes; II) homologous recognition; and III) DNA strand exchange. These three steps of the recombination reaction are driven by the RecA-DNA binding energy. Therefore, RecA-DNA complexes become more and more stable after each of these steps and their free energy becomes lower (the free energy proceeds from the highest at the top of the cycle to the lowest at the bottom). At the end of the reaction, it is necessary to dissociate the stable postsynaptic RecA-DNA complexes (IV). ATP hydrolysis is necessary for this dissociation, which brings RecA and DNA back to the starting free energy level (top of the drawing), thus closing the cycle of the RecA reaction. ATP-γ-S will maintain the filament in the high-affinity state, and will not produce the last step of the reaction.

of a nucleoside triphosphate the dissociated RecA protomers are capable of forming new presynaptic complexes, ready to promote another round of the recombination reaction[42].

The finding that the energy of ATP hydrolysis is only needed for the very last stage of the strand exchange between completely homologous single- and double-stranded DNA molecules has raised the question why RecA is constantly hydrolyzing ATP during the reaction, instead of having an ATPase activity induced only upon completion of the strand switch process. It is probable that in in vivo reactions the energy of ATP hydrolysis is also needed during the earlier stages of the reaction. The in vitro strand exchange between completely homologous DNA molecules is isoenergetic: for every base pair opened in the substrate molecules a new base pair is formed in the product molecules. The isoenergetic strand exchange can be driven to completion by the relatively small amount of free energy of RecA-DNA binding. However, in the case of strand exchange between partially heterologous DNA molecules, the reaction is not isoenergetic, as not every base pair opened in the substrate molecules is compensated for by the formation of a new base pair in the products. In such cases the energy of RecA-DNA binding might not be sufficient to push the strand exchange through heterologous regions. Since the biological purpose of recombination reactions is to exchange strands between partially heterologous DNA molecules, RecA needs to have the ability to push the strand exchange through heterology[3].

Two recent papers have directly tested whether ATP hydrolysis is required for strand exchange between partially heterologous double- and single-stranded DNA molecules. Rosselli and Stasiak[43] demonstrated that ATP hydrolysis is required to allow RecA to push strand exchange through heterologous inserts as small as 6 base pairs. Kim et al.[26] showed that in reactions performed in the presence of ATP-γ-S the branch migration halts abruptly upon reaching a heterologous region. In another paper, Kim et al.[27] studied the energetic requirements of strand exchange between two double-stranded DNA molecules. They demonstrated that in contrast to reactions between double- and single-stranded DNA, ATP hydrolysis was required for reciprocal strand exchange even in the case of complete homology.

In the reactions performed in the presence of ATP-γ-S the strand exchange between gapped duplex and linear duplex stopped exactly at the end of the gap where reciprocal strand exchange was presumed to start. Apparently, under the conditions used for the reciprocal strand exchange between two homologous duplexes (isoenergetic type of reaction), the RecA-DNA binding energy was insufficient to overcome the activation barrier needed for the simultaneous opening of base pairs in two duplexes. It is probable that the strand exchange reaction in vivo needs to overcome some additional obstacles. For example, different proteins bound to DNA may impede strand exchange. Also, the rotation of DNA molecules which is needed for the strand exchange[21] may be quite difficult in a crowded, viscous cellular interior. Therefore, significant force may be needed to drive the strand exchange reaction in vivo. Perhaps one function of the constitutive ATPase activity of RecA is to provide such a force.

Intermediates formed during strand exchange reaction between homologous double- and single-stranded DNA molecules

There is a growing consensus that the model of Howard-Flanders et al.[22] describes the general features of homologous recognition and strand exchange between homologous single- and double-stranded DNA molecules quite well. According to this model, the RecA-ssDNA complex has the ability to recognize protein-free homologous duplex DNA in a process which involves the incorporation of both strands of substrate duplex DNA into the helical RecA filament. The actual mechanism of homologous recognition is not yet known, although several recent papers propose that additional hydrogen bonds leading to DNA triplex formation can be responsible for initial homologous recognition (reviewed in 45).

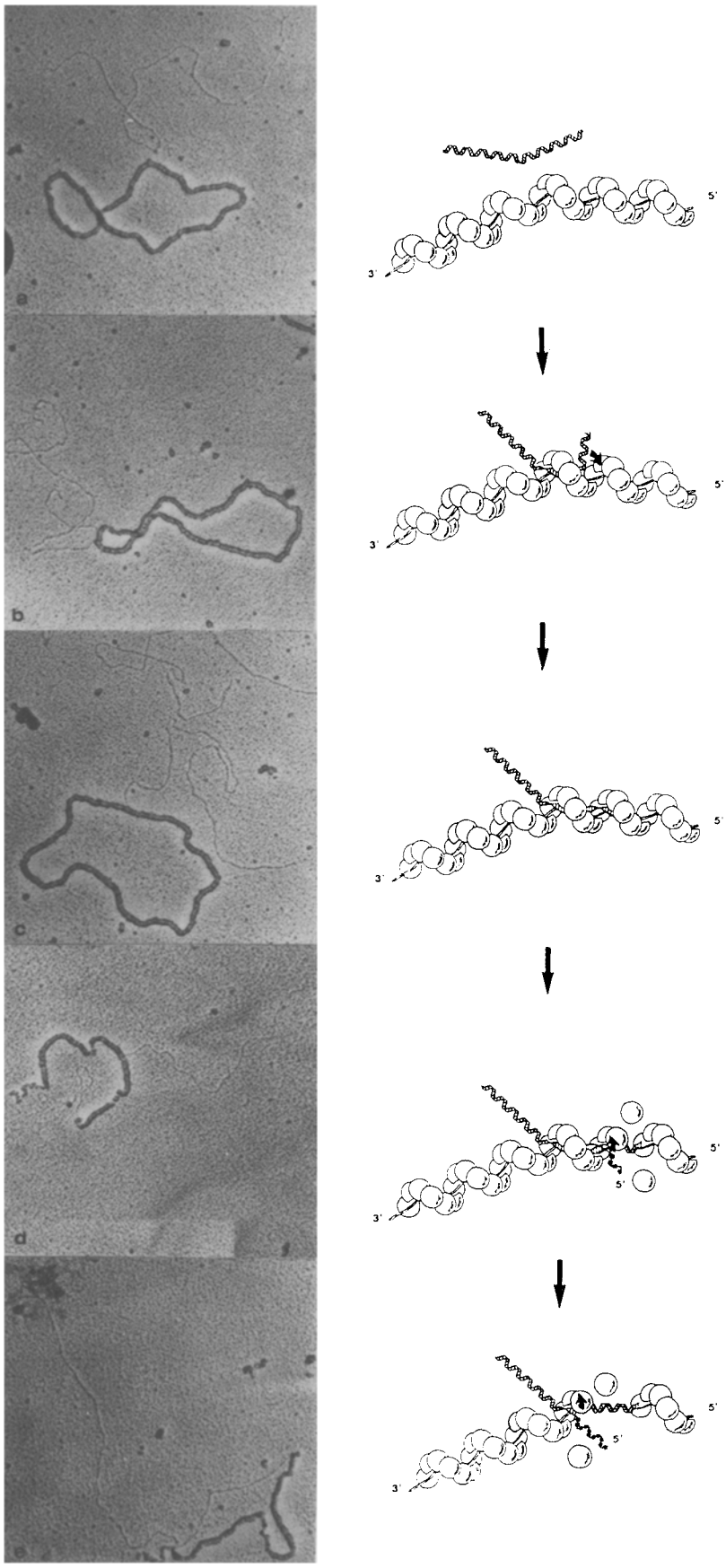
The model of Howard-Flanders et al.[22] was partially based and subsequently confirmed and extended by EM studies from several laboratories[18, 38, 47, 49] which allowed for the reconstruction of the sequence of events during in vitro strand exchange reactions between homologous double-stranded linear and single-stranded circular DNA molecules (fig. 3, a–d). These studies revealed that the prerequisite for the reaction is polymerization of RecA on substrate single-stranded DNA molecules. The complexes formed have a regular helical structure and are called presynaptic filaments[18, 53]. These filaments, in which ssDNA is in a configuration suitable for homologous recognition, play an active role in the process of homologous recognition and strand exchange. At the beginning of the reaction the presynaptic filaments show the ability to bind short regions of protein-free duplex DNA molecules (fig. 3b). This binding seems to be reversible when the sequences in the resident single-strand and in the bound duplex DNA are not in a homologous register. However, when the contacting sequences happen to be in a homologous register then the initial contact site gets extended and the duplex DNA gets progressively enveloped into the helical RecA-DNA filament, whereby homologous register is apparently maintained (fig. 3c). Electron micrographs[47, 50] suggested that this envelopment of the protein-free linear duplex DNA by the RecA-ssDNA complex progresses bidirectionally from the site of the productive initial contact. However, envelopment in a 3′ to 5′ direction (as related to the polarity of the resident ssDNA) seems to be much quicker and this ensures that the proximal end of the linear duplex (the end which contains the 5′ end of the strand to be displaced in the reaction), is quickly brought into coaxial homologous alignment with the single strand resident in the complex.

For the strand exchange reactions between double-stranded linear and single-stranded circular DNA molecules it has been shown that the displaced strand dissociates from joint molecules beginning from its free 5′ end[9, 25, 54]. Only after the proximal portion of the substrate linear duplex has been wrapped around the resident single-strand are there no more topological difficulties in the separation of the displaced strand and the newly formed duplex region. In reactions performed with the addition of SSB protein, SSB was shown to bind to the displaced strand[18, 38, 47] and seemed to pull it out from the synaptic complexes (fig. 3d). In the presence of SSB protein the strand displacement is apparently a faster reaction. Electron micrographs from such reactions show strand displacement progressing in 5′-3′ direction, quickly reaching the point where the linear duplex entered the synaptic complex. In reactions performed in the absence of SSB strand displacement seems to be slow, and this leads to the envelopment of the distal portion of the linear duplex by the synaptic complex before strand displacement reaches the point where the duplex entered the synaptic complex. When this point is reached, the strand exchange seems to proceed further without the coaxial alignment of the protein-free linear duplex with the ssDNA resident in the complex (fig. 3e). SSB protein binds to the displaced strand and RecA seems to help anneal the strand being exchanged with the ssDNA resident in the complex. The reaction can then proceed in this manner till the distal end of the linear duplex is reached[47].

The EM studies presented in figure 3 indicate that there are two distinct mechanisms of strand exchange operating at different stages of the reaction and at different regions of the substrate DNA molecules. Early in the reaction the strand exchange occurs between portions of single- and double-stranded DNA molecules which are accommodated together within the synaptic filaments (as in the model of Howard-Flanders et al.[22]). This type of reaction, occurring within the RecA complex, operates on those portions of the substrate linear duplex which extend from the point of their initial homologous contact until their proximal end (fig. 3c). However, in distal portions of ds linear DNA strand exchange operates by a mechanism which does not seem to place substrate duplex DNA within the RecA-DNA complex (fig. 3e).

How does RecA exchange strands between partially heterologous DNA molecules?

Branch migration between completely homologous DNA molecules is, in principle, a simple reaction; every base pair opened in the substrate molecules leads to a new base pair formed in the product molecules. This type of reaction can also occur spontaneously in deproteinized DNA. However, the strand exchange reaction between partially heterologous single- and double-stranded DNA molecules seems to be much


a
b
c
d
e
3'
5'
3'
5'
3'
5'
3'
5'
5'
3'
5'
5'
3'
5'

more complicated. For example, when the substrate duplex contains a heterologous insertion, this has to undergo complete strand separation without the compensatory formation of base pairs in the product DNA molecules. In addition, heterologous insertions in single- or double-stranded DNA molecules disrupt homologous alignment, hindering the progression of strand exchange. Therefore, after the complete strand separation within the heterologous insert, both of its strands have to be somehow looped out to allow for the reestablishment of homologous alignment after the heterologous insert.

How RecA manages to perform these complicated maneuvers is not yet known. There are only a small number of systematic studies devoted to heterologous strand exchange and their results are not entirely consistent. One of the first models was put forward by Bianchi and Radding[3]. These authors concluded that RecA can span heterologous insertions in ssDNA with a higher efficiency than those in dsDNA. They showed that while a 1308 bp long heterologous insertion in duplex DNA reduced the strand exchange to the background level of the reaction, the same heterologous insertion placed in single stranded DNA was spanned with up to 40% efficiency (as compared with the homologous control). According to the model that they proposed, a heterologous insertion in single-stranded DNA would be simply folded or looped out, thus allowing strand exchange to span the insertion.

In the case of the insertion in double-stranded DNA a more complicated process would be required. A complete base-pair separation of the insert needs to precede a looping of the separated strands in order to reestablish homologous register behind the heterologous insert. The model proposed by Bianchi and Radding[3] did not account for the fact that the ssDNA is a part of a presynaptic complex. This RecA-ssDNA presynaptic filament is 10-fold more rigid than ds DNA[16], suggesting that the folding or looping out of RecA-covered ssDNA in presynaptic complexes would be rather difficult. In addition, a simple looping model of ssDNA did not explain other experimental findings of the same paper, namely that short heterologous insertions in dsDNA were better spanned by RecA than the same length insertion in ssDNA.

In a more recent study, Jwang and Radding[24] placed a 110 base pair insert in various locations within a linear duplex. They concluded that it is required that the heterologous insert be flanked on both sides by substantial homologous regions for the strand exchange to proceed efficiently through heterology. The authors proposed a model explaining how the strands of the heterologous insert can be separated in advance of the actual strand exchange. According to this model, both portions of the linear duplex flanking the heterologous insertion would interact with homologous portions of ssDNA in the synaptic filament. While a portion of the duplex before the heterologous insert would be fixed within the synaptic complex, the portion beyond the heterologous insertion would undergo a rotation around the synaptic filaments. This could bring about a complete unwinding of the heterologous insertions. One mechanism by which this rotation could be achieved was proposed earlier by Cox et al.[11]. According to this proposal, duplex DNA could be moved around RecA filaments by sequential transferring between putative DNA binding sites spaced regularly around the external perimeter of the RecA filament. The model proposed by Jwang and Radding[24] did not elaborate on how insertions in ssDNA are bypassed by RecA, and how separated strands of the heterologous insert in double-stranded DNA are folded or looped out to allow branch migration to pass through the heterology.

In a very recent work by Morel et al.[32] the effect of the location of the heterologous insertions on the efficiency of strand exchange was studied. It was noticed that heterologous insertions of up to 150 bp were efficiently spanned by RecA when these insertions were located at the proximal end of duplex DNA. In the case of medial insertions, RecA's ability to span these decreased as the location of the heterologous insertion was moved from the proximal end toward the distal end of a linear double-stranded substrate. And finally, insertions as small as 22 bp placed on the distal end of linear duplex blocked the progress of strand exchange. Combining these observations with previous EM studies showing that proximal portions of linear duplexes exchange strands upon envelopment by the synaptic complexes while distal portions exchange strands without entering into synaptic complexes, Morel et al.[32] put forward a

Figure 3. Electron micrographs and accompanying models of an in vitro strand exchange reaction between double-stranded linear and single-stranded circular ϕX174 DNA.

a, *b* and *c* Aliquots taken from early time points (up to 10 min) of the reaction mediated by RecA only. Notice that due to glutaraldehyde fixation the helical structure of the presynaptic and synaptic complexes is not visible. Accompanying models present RecA-DNA complexes as deeply grooved helical filaments. The sequence of pictures is intended to illustrate the progress of the reaction. Early in the reaction *a* ssDNA gets covered with RecA while dsDNA remains protein-free. Later *b* presynaptic complex interacts with incoming duplex DNA and upon homologous recognition *c*, the proximal end of the linear duplex gets enveloped by the synaptic complex.

d and *e* Aliquots taken from reaction performed in the presence of RecA and SSB. Notice that the displaced single strand gets covered by SSB protein *d* and this helps to unravel the displaced strand from its coaxial arrangement with the newly formed duplex region. *e* Strand displacement reaches the point where the distal portion of the linear duplex contacts the RecA-ssDNA filament. After this, strand exchange can continue without the envelopment of duplex DNA by the RecA-ssDNA complex. Figure based on Stasiak and Egelman[47].

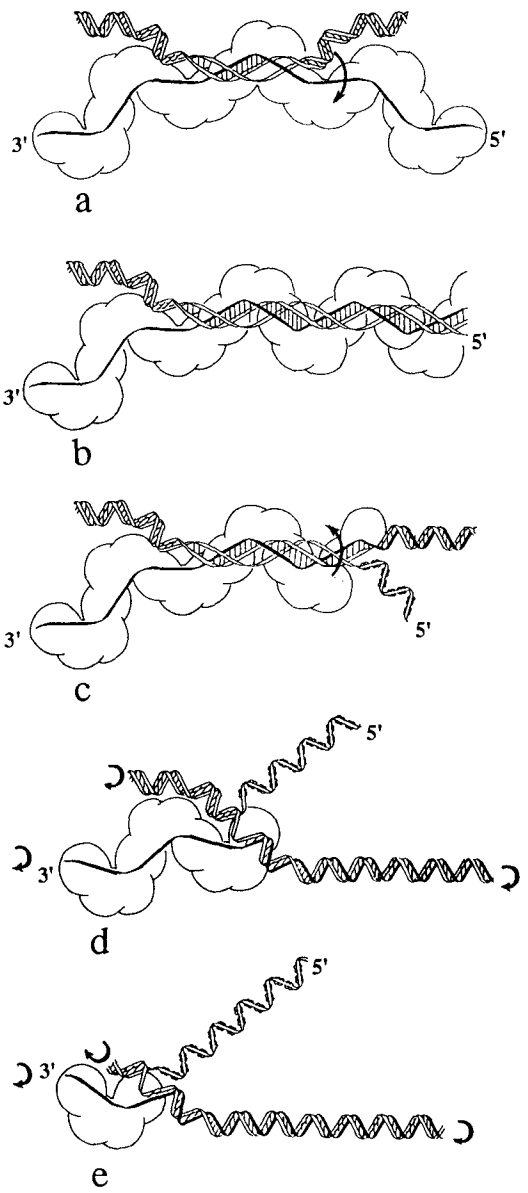


Figure 4. Molecular model of strand exchange reaction between completely homologous DNA molecules. In this schematic representation the substrate duplex DNA is drawn as white, while the substrate single-stranded DNA is drawn as black. RecA protomers in presynaptic or synaptic complexes are depicted only as contours. Coaxial alignment of DNA molecules, their extension and partial unwinding within RecA-DNA complexes are as proposed in Howard-Flanders et al.[22]. SSB proteins on the displaced single strand are drawn as small tetrameric blobs on a scale much smaller than that of RecA and DNA.
a Initiation of strand exchange. Initial pairing of the interacting DNA molecules can occur at any point of mutual homology (it is not essential for the model presented if the initial pairing occurs by homology-specific triplex formation or by direct Watson-Crick base pairing between the resident strand and the locally open region of the homologous duplex). Shortly after homologous recognition strand exchange commences in the region of initial pairing and progresses toward the proximal end of the linear duplex (notice that within synaptic complexes base pairing connects the strands which will form the new duplex molecule).
b Strand exchange in the proximal portion of the linear duplex. Duplex DNA gets progressively enveloped into the synaptic complex in a 3′–5′ direction in relation to the resident single strand. Within the enclosure of the RecA filament the strand exchange reaches the proximal end of the linear duplex.
c Spatial separation of the displaced strand. Starting from its 5′ end the displaced strand can unravel from its coaxial arrangement with the newly formed recombinant duplex. SSB binds the displaced strand while RecA dissociates from the DNA.
d Strand exchange without coaxial alignment of homologous regions. In distal portions of the linear duplex the strand exchange and the actual strand displacement are coupled with each other and proceed toward the 3′ end of the displaced strand. SSB protein binds the displaced strand and RecA helps to anneal strands of the nascent recombinant duplex.

unified model explaining homologous and heterologous strand exchange by essentially the same mechanism.

According to this model, the envelopment of dsDNA by the synaptic RecA-DNA filament leads to strand separation of the duplex DNA. It was proposed that this strand separation does not start from a proximal end of double-stranded DNA (as in some earlier models), but that it starts from the point of initial homologous recognition and then progresses toward the proximal end of the linear duplex. In the case of strand exchange between homologous DNA regions, each open base pair in the incoming duplex allows for the immediate formation of a new base pair in the recombinant duplex being formed. Such strand exchange reactions are thus practically isoenergetic, and can occur in the absence of ATP hydrolysis[31,42]. The proposal that strand separation can initiate internally in the duplex DNA, and that it does not need the energy of ATP hydrolysis when the strand exchange occurs between homologous DNA regions, is supported by the recent studies of the chemical sensitivity of DNA enclosed within synaptic complexes[1]. These studies revealed that in reactions performed in the presence of the slowly hydrolyzable analog of ATP (ATP-γ-S) the incoming duplex within the synaptic complex was in a post-exchange state, i.e., one strand was paired with the resident ssDNA and the second was left unpaired. The conversion to post-exchange state also occurred when the paired region of duplex was flanked by heterologous regions which did not undergo strand separation in the reaction performed in the absence of ATP hydrolysis.

Figure 4a shows how Morel et al.[32] depict strand exchange between completely homologous DNA molecules. It is indicated that in the region between the

initial point of contact and the proximal end of the linear duplex the process of strand switching proceeds in a 3′–5′ direction (as related to the polarity of the resident ssDNA). However, strand switching is not followed by an immediate strand displacement. Only when strand switching reaches the proximal end of linear duplex can the actual strand displacement start and proceed in a 5′–3′ direction. In accord with EM observations (ref. 47 and fig. 3) it is postulated that SSB protein binds to the displaced strand and helps to unravel it from its coaxial arrangement with the other two strands, while RecA protein dissociates from the newly formed duplex region. When strand displacement reaches the point of initial homologous recognition, the reaction proceeds further in a 5′–3′ direction, toward the distal end of the linear duplex DNA. However, in this distal portion of joint molecules strand switching leads to immediate strand displacement: the substrate duplex does not seem to get taken into the synaptic filament.

Figure 5 presents a model for the progress of strand exchange between partially heterologous DNA molecules. In the case shown the heterologous insertion happened to be between the point of initial homologous recognition and the proximal end of the linear duplex. At the beginning, such a reaction can proceed in the same way as depicted in figure 3. However when the progressing strand separation reaches a heterologous insert the process of strand switching is blocked since opening of base pairs in the heterologous insert cannot be followed by immediate base pair formation in the product DNA. The model of Morel et al.[32] postulates that from the point where heterology starts the two strands of the incoming duplex remain separated. The strand to be exchanged in the reaction would be relatively free to move within the filament, while the other strand would be firmly bound by RecA. This proposal is supported by the DNase I protection studies of individual DNA strands in RecA filaments during the process of strand exchange[5,6]. These studies showed that in contrast to the resident strand and the strand to be displaced, which were both well protected against DNase attack, the strand to be exchanged in the reaction was very sensitive to DNase as if it were only weakly bound by RecA.

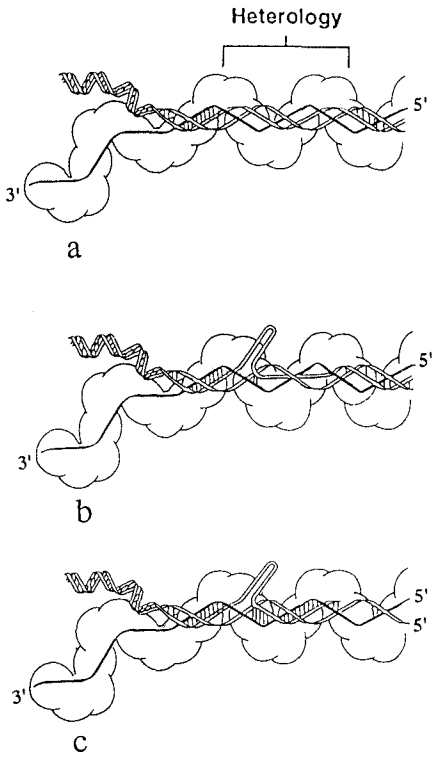


Figure 5. Schematic representation of how strand exchange can resolve medial heterology placed proximally to the point of initial pairing.
a Duplex is enveloped into a complex from the point of initial homologous recognition towards the proximal end. Initially, strands get exchanged, but in the region of heterology base pairing is opened without the possibility of corresponding base pair formation in the new duplex region. Since the ability of RecA to maintain open regions of duplex DNA is limited, the portion of linear duplex somewhat behind the heterologous insert is still in the original base pairing.
b Strand to be exchanged undergoes a lateral slippage within the open domain of duplex DNA. This slippage can be spontaneous (driven by thermal motion) or might be driven by ATP hydrolysis. Depending on the direction of slippage, heterologous insertions or deletions in duplex DNA can be bypassed by this mechanism. Complementarity is found as a result of the slippage, and the first base pairing is formed at the base of the folded loop.
c Base pairing propagates in a 3′–5′ direction in relation to initial single strand. A lateral slippage mechanism helps to move the two strands of initial duplex with respect to each other. When base pairing extends to the proximal end of the linear duplex the spatial separation of product molecules can start, and the reaction can proceed further in a 5′–3′ direction as depicted in figure 4 *c,d,e.*

To span a heterologous insert in a double stranded DNA, the whole length of the heterologous insertion has to undergo strand separation, at least for a short moment. Keeping two complementary strands apart requires, in addition to the substantial energetic input needed for the strand separation (supplied by ATP hydrolysis), a high stability of the synaptic filament. However, the structure of RecA filaments formed in the presence of ATP is rather dynamic. ATP hydrolysis is constantly going on in the synaptic filaments and RecA protomers lose temporarily their affinity for ssDNA after each cycle of ATP hydrolysis (reviewed in 28). At those regions of the synaptic complex where several adjacent RecA protomers lose affinity to ssDNA, the separated strands can snap back. This would limit the extent to which long heterologous regions could be processed by RecA. Under the experimental conditions used by Morel et al.[32] the biggest heterologies successfully spanned by RecA were of the order of 150 bp. Thus, a stretch of synaptic filament which interacts with about 150 bp may represent an upper limit of a complex size in which all RecA promoters are in the high affinity

state. In longer stretches of the synaptic complexes, which contain separated strands of a heterologous insert, some internal reannealing will always occur and this would keep product molecules unresolved, limiting the size of a heterology which could be productively processed by RecA protein. Within the open domain of dsDNA the weakly bound strand could undergo a local stretching (up to a total extension of about 7 Å per bp) in one region combined with compensatory compression or looping out in another region. Such longitudinal movements within synaptic filaments could help to reestablish homologous register between the strand to be exchanged and the resident strand in the synaptic complex.

In the case of a heterologous insertion in dsDNA, stretching of the region behind the heterologous insertion can bring about compression and looping out of the heterologous region. This can allow for spanning this insertion by complementary base pair formation (fig. 5). In the case of a heterologous insertion in ssDNA, stretching progressing from the beginning of heterology would be needed to allow for spanning this type of insertion. The movement of DNA strands in the search for complementarity can be driven by RecA using the energy obtained from ATP hydrolysis. Upon finding complementarity between the resident strand and the strand to be exchanged the process of strand switching can progress without further difficulties, since every base pair separated in the substrate duplex can be followed by new base pair formation in the product molecule. When such a strand switch reaction reaches the proximal end of the linear duplex the actual strand displacement can begin and the reaction could change direction and progress further as depicted in figure 4 c,d,e.

Clearly, the strand separation and lateral slippage of DNA strands are energetically costly and that is why ATP hydrolysis is necessary for the strand exchange between partially heterologous DNA molecules[26,43]. Morel et al.[32] proposed that the efficient strand separation and the lateral strand slippage aided by ATP hydrolysis can take place only within the recombinational scaffold of RecA filaments. Proximal ends are always enveloped by RecA synaptic filaments, and thus can be efficiently opened and manipulated by RecA. Medial insertions can be enveloped into synaptic complexes when the initial homologous recognition occurs in the portion of DNA between the distal end and the heterologous insertion. A heterologous insertion which happened to be located between the site of initial recognition and the distal end will lead to RecA's attempt to exchange strands without enveloping the heterologous duplex by the synaptic complex. Therefore, the probability of envelopment of medial insertions into synaptic complexes and their successful opening would decrease with their distance from the proximal end of the linear substrate. In distal portions of linear duplex, which are not taken into synaptic complexes, the strands of heterologous inserts have difficulty being separated. The fact that 22 bp long heterologous inserts at the distal end block the progressing strand exchange suggests that only short duplex regions can be opened there[32].

It is probable that the spontaneous end breathing of duplex DNA can be slightly enhanced by RecA or SSB and thus utilized to complete the strand exchange and to detach product DNA molecules from each other. In this respect it is important to note that RecA was shown to have a weak strand separation activity when binding to short duplex regions[4]. The largest separated fragment had a size of about 25 bp[4], which is similar in size to the largest distal heterology which could be opened by RecA.

In the process of homologous recombination in vivo other proteins can help or even replace RecA in pushing branch points of recombination intermediates. Although purified RecA, without the participation of other proteins, is able to promote branch migration, this branch migration is rather slow (on the order of 10 bp per second, ref. 10) and inefficient (with respect to the amount of ATP consumed per base pair exchanged in typical in vitro reactions). Functional analysis of various gene products involved in the process of recombination showed that the RuvB protein in association with the RuvA protein is very efficient in promoting branch migration, and that it can also promote branch migration through heterology[36]. The RecG protein was also shown to have a branch migrating activity (reviewed in Müller and West, in this issue). In contrast to the RecA protein, which acts stoichiometrically in the process of strand exchange, RuvB protein acts catalytically. Therefore, small contaminations of RuvB protein in RecA preparations (both proteins have almost the same molecular weight) could explain why in some earlier studies RecA was shown to be very efficient in promoting branch migration through heterology[3]. This highlights the fact that to comprehend the mechanism of homologous recombination one needs to study and understand both the simpler reactions performed by individual proteins as well as the multistage reactions catalyzed by a set of collaborating proteins (see Müller and West, Kowalczykowski in this issue).

Acknowledgments. This work was sponsored in part by Swiss National Science Foundation grant 31-32313.91 (to J. Dubochet and A. Stasiak).

1 Adzuma, K., Stable synapsis of homologous DNA molecules mediated by the *Escherichia coli* RecA protein involves local exchange of DNA strands. Genes Dev. *6* (1992) 1679–1694.

2 Bezzubova, O., Shinohara, A., Mueller, R. G., Ogawa, H., and Buerstedde, J. M., A chicken RAD51 homologue is expressed at high levels in lymphoid and reproductive organs. Nucl. Acids Res. *21* (1993) 1577–1580.

3 Bianchi, M. E., and Radding, C. M., Insertions, deletions and mismatches in heteroduplex DNA made by RecA protein. Cell *35* (1983) 511–520.

4 Bianchi, M. E., Riboli, B., and Magni, G., *E. coli* recA protein posseses a strand separating activity on short duplex DNA. EMBO J. *4* (1985) 3025–3030.

5 Chow, S. A., Honigberg, S. M., Bainton, R. J., and Radding, C. M., Patterns of nuclease protection during strand exchange, J. biol. Chem. *261* (1986) 6961–6971.

6 Chow, S. A., Honigberg, S. M., and Radding, C. M., DNase protection by RecA protein during strand exchange: assymetric protection of the Holliday structure. J. biol. Chem. *263* (1988) 3335–3347.

7 Clark, A. J., and Margulies. A. D., Isolation and characterization of recombination deficient mutants of *E. coli* K12. Proc. natl Acad. Sci. USA *53* (1965) 451–459.

8 Cox, M. M., and Lehman, I. R., RecA protein of *Escherichia coli* promotes branch migration, a kinetically distinct phase of DNA strand exchange. Proc. natl. Acad. Sci. USA *78* (1981) 3433–3437.

9 Cox, M. M., and Lehman, I. R., Directionality and polarity in RecA protein promoted branch migration. Proc. natl Acad. Sci. USA *78* (1981) 6018–6022.

10 Cox, M. M., Soltis, D. A., Liveneh, Z., and Lehman, I.R., On the role of single-stranded DNA binding protein in recA protein-promoted DNA strand exchange. J. biol. Chem. *258* (1983) 2577–2585.

11 Cox, M. M., Pugh, B. F., Schutte, B. C., Lindlsey, J. E., Lee, J., and Morrical, S. W., On the mechanism of recA protein-promoted DNA branch migration, in: DNA Replication and Recombination, pp. 597–607. Eds R. McMacken and T. J. Kelly, Alan R. Liss, Inc., New York 1987.

12 DiCapua, E., and Müller, B., The accessibility of DNA to dimethylsulfate in complexes with RecA protein. EMBO J. *6* (1987) 2493–2498.

13 DiCapua, E., Cuillel, M., Hewat, E., Schnarr, M., Timmins, P. A., and Ruigrok, R. W. H., Activation of RecA Protein. The Open Helix Model for LexA Cleavage. J. molec. Biol. *226* (1992) 707–719.

14 Dunn, K., Chrysogelos, S., and Griffith, J., Electron microscopic visualization of recA-DNA filaments: evidence for a cyclic extension of duplex DNA. Cell *28* (1982) 757–765.

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

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

17 Egelman, E. H., and Stasiak, A., Electron microscopy of RecA-DNA complexes: Two different states, their functional significance and relation to the solved crystal structure. Micron *3* (1993) 309–324.

18 Flory, J., Tsang, S. S., and Muniyappa, K., Isolation and visualization of active presynaptic filaments of RecA protein and single-stranded DNA. Proc. natl Acad. Sci. USA *81* (1984) 7026–7030.

19 Fujisawa, H., Yonesaki, T., and Minagawa, T., Sequence of the T4 recombination gene, uvsX, and its comparison with that of the RecA gene of *Escherichia coli*. Nucl. Acids Res. *13* (1985) 7473–7481.

20 Honigberg, S. M., Gonda, D. K., Flory, J., and Radding, C. M., The pairing activity of stable nucleoprotein filaments made from recA protein, single-stranded DNA, and adenosine 5'-(γ-thio)triphosphate. J. biol. Chem. *260* (1985) 11845–11851.

21 Honigberg, S. M., and Radding, C. M., The mechanics of winding and unwinding helices in recombination: torsional stress associated with strand transfer promoted by RecA protein. Cell *54* (1988) 525–532.

22 Howard-Flanders, P., West, S. C., and Stasiak, A., Role of RecA protein spiral filaments in genetic recombination. Nature *309* (1984) 215–220.

23 Hsieh, P., Camerini-Otero, C. S., and Camerini-Otero, R. D., Pairing of homologous DNA sequences by proteins: evidence for three-stranded DNA. Genes Dev. *4* (1990) 1951–1963.

24 Jwang, K., and Radding, C. M., Torsional stress generated by RecA protein during strand exchange separats strands of a heterologous insert. Proc. natl Acad. Sci. USA *89* (1992) 7596–7600.

25 Kahn, R., Cunningham, R. P., DasGupta, C., and Radding, C. M., Polarity of heteroduplex formation promoted by *Escherichia coli* recA protein. Proc natl Acad. Sci. USA *78* (1981) 4786–4790.

26 Kim, J.-I., Cox, M. M., and Inman, R. B., On the role of ATP hydrolysis in RecA protein-mediated DNA strand exchange. I. Bypasing a short heterologous insert in one DNA substrate. J. biol. Chem. *267* (1992) 16438–16443.

27 Kim, J.-I., Cox, M. M., and Inman, R. B., On the role of ATP hydrolysis in RecA protein-mediated DNA strand exchange. II. Four-strand exchange. J. biol. Chem. *267* (1992) 16444–16449.

28 Kowalczykowski, S. C., Mechanistic aspects of the DNA strand exchange activity of E. coli recA protein. Trends biochem. Sci. *12* (1987) 141–145.

29 Kowalczykowski, S. C., Biochemistry of genetic recombination: energetics and mechanism of DNA strand exchange. A. Rev. Biophys. biophys. Chem. *20* (1991) 539–575.

30 Lee, J. W., and Cox, M. M., Inhibition of recA protein promoted ATP hydrolysis. 1. ATP gamma S and ADP are antagonistic inhibitors. Biochemistry *29* (1990) 7666–7676.

31 Menetski, J. P., Bear, D. G., and Kowalczykowski, S. C., Stable DNA heteroduplex formation catalyzed by the *Escherichia coli* recA protein in the absence of ATP hydrolysis. Proc. natl Acad. Sci. USA *87* (1990) 21–25.

32 Morel, P., Stasiak, A., Ehrlich, S. D., and Cassuto, E., Effect of length and location of heterologous sequences on RecA-mediated strand exchange. J. biol. Chem. (1993) in press.

33 Müller, B., Koller, T., and Stasiak, A., Characterization of the DNA binding activity of stable RecA-DNA complexes; Interaction between the two DNA binding sites within RecA helical filaments. J. molec. Biol. *212* (1990) 97–112.

34 Ogawa, T., Wabiko, H., Tsurimoto, T., Hori, T., Masukata, H., and Ogawa, H., Characteristics of the purified RecA protein and the regulation of its synthesis *in vivo*. Cold Spring Harb. Symp. quant. Biol. *43* (1978) 909–915.

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

36 Parson, C. A., Tsaneva, I., Lloyd, R. G., and West, S. C., Interaction of *Escherichia coli* RuvA and RuvB proteins with synthetic Holliday junctions. Proc. natl Acad. Sci. USA *89* (1992) 5452–5456.

37 Radding, C. M., Helical interactions in homologous pairing and strand exchange driven by RecA protein. J. biol. Chem. *266* (1991) 5355–5358.

38 Register, J. C. III, Christiansen, G., and Griffith, J., Electron microscopic visualization of the recA protein-mediated pairing and branch migration phases of DNA strand exchange. J. biol. Chem. *262* (1987) 12812–12820.

39 Riddles, P. W., and Lehman, I. R., The formation of plectonemic joints by the recA protein of *Escherichia coli*, requirements for ATP hydrolysis. J. biol. Chem. *260* (1985) 170–173.

40 Roberts, J. W., Roberts, C. W., Craig, N. L., and Phizicky, E. M., Activity of *Escherichia coli* recA-gene product. Cold Spring Harb. Symp. quant. Biol. *43* (1978) 917–920.

41 Roca, A. I., and Cox, M. M., The RecA protein: structure and function. Crit. Rev. Biochem. molec. Biol. *25* (1990) 415–456.

42 Rosselli, W., and Stasiak, A., Energetics of RecA-mediated recombination reactions: Without ATP hydrolysis RecA can mediate polar strand exchange but is unable to recycle. J. molec. Biol. *216* (1990) 333–350.

43 Rosselli, W., and Stasiak, A., The ATPase activity of RecA is needed to push the DNA strand exchange through heterologous regions. EMBO J. *10* (1991) 4391–4396.

44 Shinohara, A., Ogawa, H., Matsuda, Y., Ushio, N., Ikeo, K., and Ogawa, T., Recombination genes of human, mouse and fission yeast homologous to *S. cerevisiae* RAD51 and *E. coli* recA. Nature Genet. *4* (1993).

45 Stasiak, A., Three-stranded DNA structure: is this the secret of DNA homologous recognition? Molec. Microbiol. *6* (1992).
46 Stasiak, A., and DiCapua, E., The helicity of DNA in complexes with RecA protein. Nature *229* (1982) 185–186.
47 Stasiak, A., and Egelman, E. H., Visualization of recombination intermediates, in: Genetic Recombination, pp. 265–307. Eds R. Kucherlapati and G. R. Smith. American Society for Microbiology, Washington, D.C. 1988.
48 Stasiak, A., DiCapua, E., and Koller, T., Elongation of duplex DNA by RecA protein. J. molec. Biol. *151* (1981) 557–564.
49 Stasiak, A., Stasiak, A. Z., and Koller, T., Visualization of RecA-DNA complexes involved in consecutive stages of an *in vitro* strand exchange reaction. Cold Spring Harb. Symp. quant. Biol. *49* (1984) 745–749.
50 Stasiak, A. Z., Rosselli, W., and Stasiak, A., RecA-DNA helical filaments in genetic recombination. Biochemie *73* (1991) 199–208.
51 Story, R. M., Weber, I. T., and Steitz, T. A., The structure of the *E. coli* recA protein monomer and polymer. Nature *355* (1992) 318–325.
52 Takahashi, M., and Schnarr, M., Investigation of RecA-polynucleotide interactions from the measurement of LexA repressor cleavage kinetics. Eur. J. Biochem. *183* (1989) 617–622.
53 Tsang, S. S., Muniyappa, K., Azherdian, E., Gonda, D. K., Radding, C. M., Flory, J., and Chase, J. W., Intermediates in homologous pairing promoted by RecA protein. J. molec. Biol. *185* (1985) 295–309.
54 West, S. C., Cassuto, E., and Howard-Flanders, P., Heteroduplex formation by RecA protein: polarity of strand exchanges. Proc. natl Acad. Sci. USA *78* (1981) 6149–6153.
55 West, S. C., Enzymes and Molecular Mechanisms of Genetic Recombination. A. Rev. Biochem. *61* (1992) 603–640.
56 Yu, X., and Egelman, E. H., Structural data suggest that the active and inactive forms of the RecA filament are not simply interconvertible. J. molec. Biol. *227* (1992) 334–346.
57 Yu, X., and Egelman, E. H., The LexA repressor binds within the deep helical groove of the activated RecA filament. J. molec. Biol. *231* (1993) 29–40.
58 Yu, X., and Egelman, E. H., The DNA conformation induced by the bacteriophage T4 UvsX protein appears identical to the conformation induced by the *Escherichia coli* RecA protein. J. molec. Biol. *232* (1993) 1–4.