J. Biochem. **119**, 216-223 (1996)

Homologous Recognition by RecA Protein Using Non-Equivalent Three DNA-Strand-Binding Sites[1]

Hitoshi Kurumizaka[2] and Takehiko Shibata[3]

Laboratory of Cellular & Molecular Biology, The Institute of Physical and Chemical Research (RIKEN), 2-1 Hirosawa, Wako, Saitama 351-01

Received for publication, November 7, 1995

A key step in homologous recombination is the formation of a heteroduplex joint between double-stranded DNA and single-stranded DNA by the homologous pairing and strand-exchange, and this step is also important in recombinational repair of damaged DNA in various organisms. The homologous pairing and the strand-exchange are promoted *in vivo* and *in vitro* by RecA protein of *Escherichia coli* or its homologues of bacteria, virus, and lower and higher eukaryotes. A central question on the mechanism of homologous recombination is how RecA protein (and its homologues) recognizes homologous sequences between single-stranded DNA and double-stranded DNA. Recent studies suggest that RecA protein promotes homologous recognition between these DNA molecules by the formation of a transient and additional pairing of identical sequences *via* non-Watson-Crick interactions to the Watson-Crick-type duplex DNA, and that RecA protein uses three non-equivalent DNA-strand-binding sites in this reaction.

Key words: genetic recombination, homologous pairing, non-Watson-Crick interactions, RecA protein, triplex DNA.

Homologous recombination is a general process in various organisms by which genetic information is re-arranged to create genetic variations. Homologous recombination plays crucial roles in the repair of damaged DNA in various organisms, and is essential to reductive cell-division (meiosis) in eukaryotes. An apparatus for homologous recombination has to find homologous sequences within entire genomic DNA to form inter-molecular duplex joints ("heteroduplex joints") that are general intermediates for recombination (*1*). In 1979, it was discovered that the *Escherichia coli recA* gene product, RecA protein, recognizes and pairs homologous sequences between single-stranded DNA and double-stranded DNA to form stable heteroduplex joints, through a Mg^{2+}- and ATP-dependent reaction *in vitro* (*2, 3*). This reaction consists of two successive steps, "homologous pairing" and "strand-exchange." The homologous pairing is the formation of small heteroduplex joints, and the strand-exchange is unidirectional elongation of the heteroduplex joints (*4, 5*). Recently, eukaryotic homologues of RecA protein were found from yeast to human as proteins of the Rad51-family (*6-9*), and these proteins were shown to promote the homologous pairing and the strand-exchange (*10*). Thus, RecA protein-promoted homologous pairing and the strand-exchange are generally important for homologous recombination.

Homologous pairing promoted by RecA protein is divided into two phases: the presynaptic phase and synapsis (Fig. 1; *11*). In the presynaptic phase, RecA protein first binds to single-stranded DNA with ATP as a cofactor (the primary binding; Fig. 1; *12, 13*), to form helical nucleoprotein filament called "presynaptic filament" (*14-16*). By the presynaptic filament-formation, hydrogen bonds maintaining a folded structure of single-stranded DNA are disrupted (*17, 18*), and the single-stranded DNA is extended up to 1.5-folds longer than the B form DNA (*15*). Then, in the synapsis, the presynaptic filament binds to double-stranded DNA forming a three-component complex without homologous alignment (the secondary binding; Fig. 1; *12, 19, 20*). Homologous sequences of single-stranded DNA and double-stranded DNA are aligned in the three-component complex, and the heteroduplex joints maintained by the intermolecular Watson-Crick base-pairs are formed without net interwinding of the two strands (paranemic joint; Fig. 1; *21, 22*). When a terminus of a DNA-strand involved in the complex is located within a homologous sequence or a topoisomerase is present in the reaction mixture, RecA protein promotes interwinding of the two strands of the heteroduplex joints to form a plectonemic joint (*21, 23, 24*), followed by the strand-exchange (*4, 5*).

How are homologous sequences recognized by RecA protein in the three-component complex? This is a central question to be solved for understanding of the molecular mechanism of homologous recombination. In this review, we focus on DNA-DNA interactions in the three-component complex and on the DNA-binding sites of RecA protein involved in the recognition. Other aspects and comprehensive view of RecA protein were described in other recent reviews (*25-28*).

[1] This study was supported partly by a grant for "Biodesign Research Program" from RIKEN, and partly by a grant from the Ministry of Education, Science and Culture of Japan.

[2] Present address: Laboratory of Molecular Embryology, National Institute of Child Health and Human Development, NIH, Bethesda, Maryland, USA.

[3] To whom correspondence should be addressed. Fax: +81-48-462-4671, E-mail: tshibata@postman.riken.go.jp

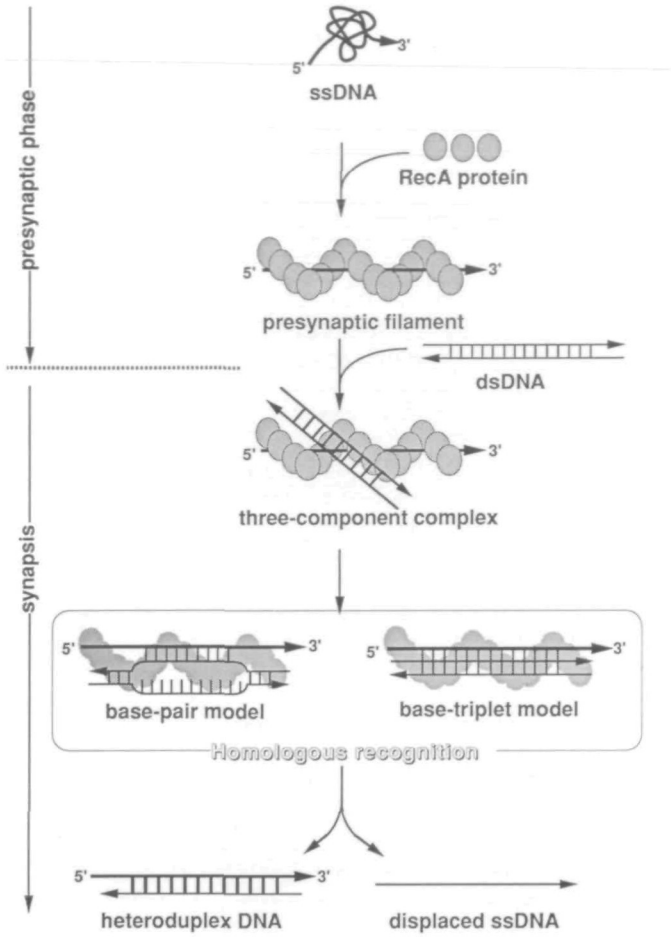

Fig 1 **Steps in the homologous pairing by RecA protein.** The mechanism of the recognition of homologous sequences between double-stranded DNA and single-stranded DNA has been explained by either of two models: *Base-pair model* and *Base-triplet model*

RecA protein

RecA protein is an acidic protein of 38 kDa consisting of 352-amino acid residues (*29, 30*), and has activities to promote DNA-dependent ATP-hydrolysis (*31, 32*) and unwinding of double-stranded DNA (*19, 33, 34*) in addition to the homologous pairing and the strand-exchange *in vitro*. RecA protein promotes proteolytic cleavage of a group of proteins such as lambda repressor (*35-37*) by an ATP- and single-stranded DNA-dependent fashion, but the proteolytic reaction is not dealt with in this review. Unwinding of double-stranded DNA by RecA protein requires ATP or its analogue, ATPγS, and associates with the polymerization of RecA protein along the DNA-strands. The filament formed along single-stranded DNA and that formed along double-stranded DNA are indistinguishable by electron microscopy, since both have the same striated profile and dimensions. Either single-stranded or double-stranded DNA is located in an inner part of the filament (*38*). Even without ATP and DNA, RecA protein forms a filamentous polymer with the similar apparent structure (*39, 40*). Two kinds of RecA filaments have been recognized by electron microscopic studies; an extended filament (95-100 Å of pitch) that represents the active state induced by the binding of ATP (or ATPγS) and DNA, and a compressed filament (70-75 Å of pitch) that represent an inactive state without ATP (*15, 41-43*). Three-dimensional structures of free and ADP-bound RecA proteins have been solved by X-ray crystallography (*44, 45*). In the crystal, RecA protein forms a head-to-tail polymer that has helical periodicity of six monomers per turn with the pitch of 82.7 Å.

Although amino acid-sequences are not very well conserved, virus analogue, UvsX protein of *coli* phage T4 (*46*) and proteins belonging to the Rad51-family (including those of yeast and human; *47, 48*) also formed filaments along DNA with the same morphology as RecA protein of *E. coli*. These findings suggest a general role of protein-single-stranded DNA filament in homology search between the single-stranded DNA and double-stranded DNA for homologous recombination in living organisms.

Mechanistic models for the homologous recognition

To explain molecular mechanisms of homologous recognition in the three-component complex of single-stranded DNA, double-stranded DNA, and RecA protein, two models have been proposed; *i.e.*, (i) *Base-pair model* (Fig. 1) in which parental duplex DNA is partially melted prior to pairing, and then homologous sequences are searched for by the intermolecular Watson-Crick interactions between complementary base-sequences; and (ii) *Base-triplet model* (Fig. 1) in which homologous sequences are searched for by trials to form a triplex joint without the disruption of the parental Watson-Crick pairing, followed by base-pair-switch and strand separation (from the Watson-Crick pairing in parental duplex DNA to paring in heteroduplex joints) to form heteroduplex joints.

The base-pair model

The base-pair model is a straightforward explanation for the fidelity of homologous recognition (Fig. 1). This model had been supported by the following observations: (i) RecA protein catalyzes ATP-dependent renaturation of complementary strands (Fig. 2; *49*). (ii) The double-stranded DNA in the three-component complex formed with ATPγS was shown by a topological technique to be unwound by RecA protein without homologous alignment (*19*), and (iii) electron microscopic observations of RecA-double-stranded DNA complex showed that the helical pitch of double-stranded DNA in the complex is extended 1.5 times longer than the B-form DNA, supporting the unwinding of the double-helix by RecA protein (*33, 41*). Unwinding is an obligatory consequence of melting of base-pairs in the double-helix, and thus, these observations suggested the melting of double-strands by RecA protein. (iv) Consistent with this suggestion, although it can only melt double-strands at most 20 bps, RecA protein was shown to have a helicase activity (*50*).

However, contradictory results have been accumulated later. By a chemical probing method using dimethylsulfate that methylates DNA bases at functional groups involved in the Watson-Crick interaction (guanine N7, adenine N3, and cytosine N3), Di Capua and Muller have shown that the Watson-Crick base-pairs of the double-stranded DNA in the complex are not separated by the binding of RecA protein (*51*). The same conclusion was obtained from neutron-scattering and ultraviolet linear dichroism observations of RecA-double-stranded DNA complex (*52*).

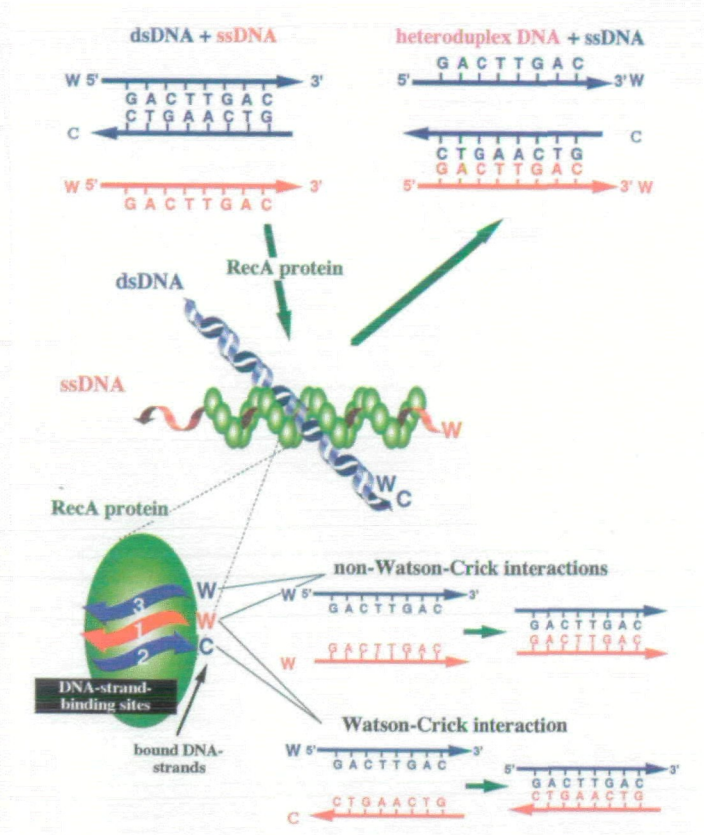

Fig. 2. **Non-equivalent three DNA-binding sites on RecA protein functional to the homologous pairing.** RecA protein has three non-equivalent DNA-strand-binding sites as indicated. RecA protein forms the presynaptic filament by use of site 1. Then, strands of double-stranded DNA bind to site 2 and site 3. Site 2 accepts DNA-strand only in anti-parallel to the DNA-strand bound to site 1. First, identical sequences are searched for between the strand in site 3 and that in site 1 by non-Watson-Crick interactions. Once single-stranded DNA and double-stranded DNA are aligned at a candidate sequence for homologous one, then Watson-Crick base-pairs of the parental duplex DNA are switched to pair the strand in site 1 and the strand in site 2 to try to form a heteroduplex joint. Double-stranded DNA bound to the presynaptic filament could be bent to have the maximum contact with single-stranded DNA. dsDNA, double-stranded DNA; ssDNA, single-stranded DNA; W, Watson-strand; C, Crick-strand

Furthermore, under the conditions for the homologous pairing including ATP, Ohtani *et al.* indicated that most of the unwinding of double-stranded DNA is initiated by the homologous pairing (*34, 53*). Makino *et al.* found that an anti-RecA monoclonal IgG, ARM193, severely inhibits and even reverses the unwinding of double-stranded DNA by RecA protein, but does not the RecA protein-promoted homologous pairing (*54, 55*). These results imply that the recognition of homology in the homologous pairing and the unwinding of the double-helix by RecA protein are independent of each other.

The base-triplet model

Base-triplet model in which the homologous sequences are searched for by additional hydrogen bonding without disrupting the original Watson-Crick base-pairs is more favorable to explain quick search for homologous sequences (Fig. 1). This model was supported by evidences for the formation of triplex DNA by RecA as described below.

When circular single-stranded DNA and linear double-stranded DNA are substrates, following the homologous pairing, RecA protein promotes strand-exchange in which a strand of the parental duplex DNA is replaced by the partner single-stranded DNA. The final product of this reaction is nicked circular heteroduplex DNA and linear single-stranded DNA (*4, 5*). The strand-exchange promoted by RecA protein proceeds only in a direction of 5′ (proximal) to 3′ (distal) as to the parental single-stranded DNA. The presence of sequence heterology at the proximal end of double-stranded DNA allows the homologous pairing, but blocks the strand-exchange by preventing the separation of the strand of the parental duplex DNA to be replaced. Even after RecA protein is removed, the joint molecules formed under these conditions can survive the incubation at a temperature at which the Watson-Crick base-pairs are melted (up to 80°C; *56, 57*). The abnormal stability of the joint molecules suggests the presence of a novel joint structure other than the normal Watson-Crick pairing. A study involving the probing by P1 nuclease that specifically cleaves single-strands showed that all three strands in the joint molecules were protected from the attack by P1 nuclease, suggesting the formation of homology dependent triplex DNA by RecA protein (*58*). Studies on the structure of the deproteinized joint molecules by chemical or P1 nuclease probing methods supported the presence of triplex structures, but revealed that the triplex DNA was not intermediates for the search for homology. Diethylpyrocarbonate (DEPC) modifies unpaired purine residues, and $KMnO_4$ or OsO_4 modifies pyrimidine residues in single-stranded DNA. The results from the chemical probing-experiment using these reagents showed that the stable Watson-Crick base-pairs were formed between single-stranded DNA and its complementary strand from the double-stranded DNA (*58, 59*). This conclusion implies that the triplex DNA detected is not an intermediate for the homologous recognition but a product.

Jain *et al.* studied triplex DNA inside RecA-filament by chemical cross-linking experiments before the removal of the protein, followed by the electron-microscopic observations (*60*). The study indicated that the three-stranded DNA cross-linked by a psoralen derivative was detected only in the RecA-filament and only when heteroduplex joints were formed with ATP (not with an unhydrolyzable analogue, ATPγS) and the strand-exchange was inhibited by placing heterologous sequence at the proximal end. On the other hand, when the psoralen derivative reacted after the removal of RecA protein, only heteroduplex DNA was cross-linked but triplex DNA was not. They suggested that the structure of the triplex DNA in the RecA-filament is different from that detected after the removal of RecA protein, and since the triplex DNA detected by them was ATP hydrolysis-dependent and the homologous pairing did not require ATP-hydrolysis, they doubted that the triplex structure was an intermediate for the homology search (*60*).

Therefore, at present triplex structures responsible to the homology search between single-stranded DNA and double-stranded DNA have not been directly detected.

RecA protein promotes non-Watson-Crick interactions

If triplex DNA is an intermediate for homologous recog-

nition, non-Watson-Crick interactions between single-stranded DNA and double-stranded DNA have to be included. As a non-Watson-Crick-interaction, the Hoogsteen base-pairing was first considered. The Hoogsteen base-pairing was found in a canonical triplex DNA formed non-enzymatically (*61-63*). Substitutions of 7-diazaguanine for all guanines in double-stranded DNA that prevent Hoogsteen base-pairing, did not affect the RecA protein-promoted homologous pairing (*64*). This result suggests that the Hoogsteen base-pairing is not involved in homologous recognition by RecA protein. Similar results were obtained by experiments using methylated DNA at N7 of guanine bases (*65*). Furthermore, in the identified triplex DNA consisting of Hoogsteen base-pairing, a pair of strands connected by paring of identical bases is antiparallel, but in genetic recombination, a pair of DNA-strands bearing an identical sequence must be aligned in parallel, since the cutting and rejoining in homologous recombination occur between identical sequences of the strands with the same polarity (see Ref. *66*).

Rao and Radding found that the presynaptic filament of RecA protein specifically recognized a 33-mer oligo-DNA that has the identical sequence to the single-strand DNA within the presynaptic filament and form a homologous complex with the 33-mer (identical recognition; Fig. 2; *67*). This finding suggests that RecA protein can pair identical strands *via* non-Watson-Crick interactions. Based on this finding, they proposed that the non-Watson-Crick base-pairing between identical strands was a part of base-triplets for the homologous recognition. A support of this idea would be provided from a study involved in mutant RecA proteins that are defective in the homologous pairing yet proficient in the annealing reaction of complementary strands. We found that a chimeric RecA protein of *E. coli* and *Pseudomonas aeruginosa*, the RecAc38 protein, constructed by Ogawa *et al.* (*68*) exhibited such properties *in vitro*. The RecAc38 protein causes partial deficiencies in homologous recombination and the repair of UV damaged DNA *in vivo* (*68*). The RecAc38 protein has nine amino acids substitutions in the region of 58th to 109th amino acid residues compared with the wild-type *E. coli* RecA protein. The purified RecAc38 protein is proficient in the ATP-stimulated annealing of complementary DNA-strands, the single-stranded DNA-dependent hydrolysis of ATP, the formation of the presynaptic filament along single-stranded DNA with ATP, and the ATP-dependent extensive unwinding of the double helix (*69*). We found, however, that the RecAc38 protein is defective in both the homologous pairing of single-stranded DNA and double-stranded DNA, and the recognition of identical single-strands (Fig. 1; *69*). These observations include important findings; *i.e.*, (i) the recognition of identical sequences is likely to play a crucial role in the RecA protein-promoted homologous pairing of single-stranded DNA and double-stranded DNA; and (ii) the DNA-binding sites responsible for the recognition of identical strands are at least partly independent of those for the recognition of complementary strands and those for the formation of the filament along single-stranded DNA or double-stranded DNA.

Considering the instability and short life of complexes responsible for homology search, an approach toward the further understanding of the homologous recognition by RecA protein would be the analysis of tertiary structures of DNA-strand-binding sites, which would determine spacial configurations of three DNA-strands during homology search.

DNA-binding sites of RecA protein

To form the three-component complex that is the apparatus for the homologous recognition by RecA protein, the RecA-filament formed along single-stranded DNA has to bind to double-stranded DNA (secondary DNA-binding), and thus, an active unit of RecA protein has a site(s) for the secondary binding of double-stranded DNA besides the site for the primary binding to single-stranded DNA (*12*). Howard-Flanders *et al.* proposed that a spiral RecA protein filament had two equivalent DNA-binding sites within the filament (*70*), and Horii proposed that the RecA-polypeptides with a single DNA-binding site form a symmetric dimmer as an active unit (*71*). Takahashi and co-workers proposed that RecA protein has three DNA-binding sites by a study using a chemically modified DNA substrates (*72*).

Biochemical experiments, however, have shown the nonequivalence of the DNA-binding sites of RecA protein. As described, the RecAc38 protein is proficient in the formation of the presynaptic filament (primary binding), but was defective in the homologous pairing. The deficiency of the RecAc38 protein in the homologous pairing was attributable to the defective formation of the three-component complex (secondary binding; *69*). These indicate that the RecAc38 protein is proficient in the site for primary binding, but is defective in the site for secondary binding of double-stranded DNA.

On the other hand, the annealing of complementary DNA-stands is mediated by two pathways (*73*): (i) annealing between locally concentrated protein-free complementary sequences that occur under the condition for "aggregation of single-stranded DNA" by RecA protein (*20*) and is the same mechanism of annealing by histone H1 or ethanol; and (ii) annealing between saturated RecA protein-single-stranded DNA complex (identical to the presynaptic filament) and protein-free single-stranded DNA, which is the same mechanisms for the homologous pairing of single-stranded DNA and double-stranded DNA. In the annealing by the latter mechanism, three component complex consisting of RecA protein and two single-stranded DNAs is an intermediate (*74, 75*), and since the RecAc38 protein is proficient in the annealing reaction, the presynaptic filament of RecA38 protein should have a proficient DNA-binding site for the secondary binding of single-stranded DNA. In contrast to the annealing, the presynaptic filament of the RecAc38 protein is defective in the recognition of identical strands (*76*). A simple interpretation of these observations is as follows: (i) the presynaptic filament of RecA protein and single-stranded DNA has two kinds of sites for the secondary DNA-binding; one of them (site 2) is responsible for the recognition of the complementarity between two single-strands by the Watson-Crick interaction and the other (site 3) for the recognition of the identity through non-Watson-Crick interactions; and both the sites are required for the secondary binding of double-stranded DNA. (ii) The RecAc38 protein has proficient site 2 and defective site 3 besides the proficient site (site 1) for the primary DNA-binding (three-sites model; Fig. 2; *76*).

Our three site model includes an assumption that site 1 and site 2 are not equivalent and this assumption is

supported by the following observations: we found that a mutant RecA protein having a single substitution of Arg_{243} by Gln (RecAR243Q) is proficient in the presynaptic filament formation, single-stranded DNA-dependent hydrolysis of ATP, and the recognition and complex formation of the identical DNA-strands, but is defective in both the homologous pairing of single-stranded DNA and double-stranded DNA and the annealing reaction of complementary single-strands (H.K. *et al.*, unpublished observation). This observation indicates that RecAR243Q protein has proficient site 1 and site 3 and defective site 2, and that the formation of the stable heteroduplex joints requires also the RecA protein-promoted (not spontaneous) Watson-Crick interaction. An important role of Arg_{243} in the binding of RecA protein to DNA had been reported previously (*77*).

Location of DNA-binding sites on the RecA-polypeptide

DNA-binding sites on the RecA-polypeptide were first located by similarity or homology of the amino acid-sequence to other DNA-binding proteins. An example is "DNA-binding wing" (Arg_{243}-Gln_{257}) suggested by the comparison of the amino acid sequence of RecA protein with the DNA-binding wings of the gene V-encoded single-stranded DNA-binding proteins of filamentous bacteriophages (*77-79*).

When RecA protein forms filament along single-stranded DNA or double-stranded DNA, the DNA was located at the center of the filament, and both DNA molecules were shown to bind to the same binding sites (*38*). Just after the homologous pairing, both double-stranded DNA and single-stranded DNA are included within the same RecA filament (*60, 80, 81*). These observations suggested that RecA filament has two or perhaps three sites for the DNA-binding inside the filament. A crystal structure of RecA protein revealed two loop-regions (Loop-1, the 157th to 164th amino acid residues counted from the N-terminus; and Loop-2, the 195th to 209th amino acid residues) that included clusters of highly conserved amino acid residues and lay near the polymer axis. Story *et al.* suggested that these loops consist of the primary and secondary DNA-binding sites (*44*). Mutations in these loops were shown to cause deficiencies in the formation of the presynaptic filament (*82, 83*), the ATP-dependent renaturation of complementary strands (*83*), the displacement of single-stranded DNA-binding protein (SSB) from single-stranded DNA (*84*), the homologous pairing and strand-exchange (*84, 85*). These observations are consistent with a role of both loops for the DNA-binding. It is very likely that either or both loops are involved in the primary DNA-binding. On the other hand, photochemical crosslinking experiments and a fluorescence study suggested the involvements of other regions of the RecA-polypeptide in the DNA-binding (*77*).

At the first stage of the homologous recognition, double-stranded DNA outside the presynaptic filament (of RecA protein and single-stranded DNA) has to find "gate-way" to contact with the single-stranded DNA at the center of the filament. On the other hand, RecA protein efficiently recognizes homologous sequences between DNA molecules independent of termini of DNA-strands (*21, 22, 53, 86*). Since the absence of terminus of both DNA molecules within a homologous sequence prevents extensive interwinding of the two molecules, coaxial complexes with extensive interwinding of parental DNA molecules are an unlikely candidate for intermediates for homology search. Thus, an internal region of double-stranded DNA outside the presynaptic filament has to contact with the single-stranded DNA inside the filament for each trial to search for homology. We assumed that amino acid residues in the gate-way would play a primary role in the secondary DNA-binding and that the residues locate in regions facing outside or an opening of the filament to help this trial.

Our guide to select amino acid residues to be studied as candidates of the sites for the secondary DNA-binding was the mapped epitopes of anti-RecA protein IgGs, ARM414 (within 233rd to 250th amino acid residues) and ARM191 (the 280th to the 320th amino acid residues) both of which prevent the binding of double-stranded DNA (*87, 88*). Especially, ARM191 prevents homologous pairing by RecA protein and the binding of the protein to double-stranded DNA, but little affects the single-stranded DNA-dependent ATP-hydrolysis by the protein (*89*). Since the binding of a large IgG to small RecA protein is likely to cause an extensive alteration of the tertiary structures of the protein, and may not directly interfere the DNA-binding, we more seriously consider possible locations of the gate-way described above. We looked especially at basic amino acid residues that were (i) located in regions for a possible gate-way and (ii) well conserved among proteins of the RecA-family (see Ref. *25*). In this assumption, site 2 and site 3 are not really "sites" but paths in which DNA-strands have multiple contacts with amino acid residues.

We replaced conserved Arg and Lys by Gln and Asn, respectively, in a such region suggested by modeling derived from the crystal structure of RecA protein (A. Sarai and H.K., an unpublished work), so that the replacements kept hydrophilicity of the region to give the least effect to the tertiary structure of the protein. We found some replacements of conserved basic amino acid-residues in a region starting from Arg_{243} to Lys_{302} caused deficiencies in homologous pairing but did not affect the ATP-dependent formation of the presynaptic filament along single-stranded DNA and single-stranded DNA-dependent hydrolysis of ATP. Based on these observation, we studied a purified polypeptide consisting of amino acid sequence from Val_{238} to Asn_{332} ($RecA_{238-332}$) that included all the residues mentioned above. We found that this polypeptide was able to bind to either double-stranded DNA or single-stranded DNA, and that the effects of the replacements of amino acid residues that caused deficiencies in the nucleation of the binding of RecA protein to double-stranded DNA were well reproduced by the polypeptide (H.K. and H. Aihara, unpublished observation). These results suggest that this region of the RecA-polypeptide has at least a part of the sites for the secondary DNA-binding.

This suggestion is consistent with the results of photochemical cross-linking experiments that implied the region between 233rd and 243rd amino acid residues (W.M. Rehrauer and S.C. Kowalczykowski, personal communication) and that between 257th and 280th amino acid residues (*90*), in the DNA-binding. The chemical cross-linking experiments that show proximity of an amino acid residue to a DNA-strand but are not necessary to indicate functional significance are complementary to the mutational approaches.

Concluding remarks

Biochemical studies have provided insights and suggestions for the understanding of the homologous recognition by RecA protein as follows: (i) RecA protein forms homology dependent triplex DNA, (ii) the presynaptic filament of RecA protein recognizes identical single-strands *via* non-Watson-Crick interactions, and (iii) the non-Watson-Crick interactions are crucial for homologous recognition between single-stranded DNA and double-stranded DNA in the homologous pairing. The lower stringency in the identical recognition and weak pairing between identical strands might be favorable for quick search for a homologous sequence between DNA molecules (*91*). To understand the mechanism of homologous recognition by RecA protein, it is important to find out the tertiary structure of the intermediates and the DNA-binding sites of RecA protein. Recent advances of NMR spectroscopy and X-ray crystallography would enable us to analyze the tertiary structures of the RecA-DNA complex and DNA molecules in the RecA-filament.

REFERENCES

1. Holliday, R. (1990) The history of the DNA heteroduplex. *Bioessays* **12**, 133-142
2. Shibata, T., DasGupta, C., Cunningham, R.P., and Radding, C.M. (1979) Purified *Escherichia coli recA* protein catalyzes homologous pairing of superhelical DNA and single-stranded fragments. *Proc. Natl. Acad. Sci. USA* **76**, 1638-1642
3. McEntee, K., Weinstock, G.M., and Lehman, I.R. (1979) Initiation of general recombination catalyzed *in vitro* by the *recA* protein of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **76**, 2615-2619
4. Cox, M.M. and Lehman, I.R. (1981) *recA* protein of *Escherichia coli* promotes branch migration, a kinetically distinct phase of DNA strand exchange. *Proc. Natl. Acad. Sci. USA* **78**, 3433-3437
5. Kahn, R., Cunningham, R.P., DasGupta, C., and Radding, C.M. (1981) Polarity of heteroduplex formation promoted by *Escherichia coli* recA protein. *Proc. Natl. Acad. Sci. USA* **78**, 4786-4790
6. Shinohara, A., Ogawa, H., and Ogawa, T. (1992) RAD51 protein involved in repair and recombination in *S. cerevisiae* is a recA-like protein. *Cell* **69**, 457-470
7. Bishop, D.K., Park, D., Xu, L.Z., and Kleckner, N. (1992) Dmc1: A meiosis-specific yeast homolog of *E. coli recA* required for recombination, synaptonemal complex formation, and cell cycle progression. *Cell* **69**, 439-456
8. Basile, G., Aker, M., and Mortimer, R.K. (1992) Nucleotide sequence and transcriptional regulation of the yeast recombinational repair gene RAD51. *Mol. Cell. Biol.* **12**, 3235-3246
9. Shinohara, A., Ogawa, H., Matsuda, Y., Ushio, N., Ikeo, K., and Ogawa, T. (1993) Cloning of human, mouse and fission yeast recombination genes homologous to RAD51 and recA. *Nature Genet.* **4**, 239-243
10. Sung, P. (1994) Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. *Science* **265**, 1241-1243
11. Radding, C.M., Flory, J., Wu, A., Kahn, R., DasGupta, C., Gonda, D., Bianchi, M., and Tsang, S.S. (1983) Three phases in homologous pairing: Polymerization of *recA protein* on single-stranded DNA, synapsis, and polar strand exchange. *Cold Spring Harbor Symp. Quant. Biol.* **47**, 821-828
12. Shibata, T., Cunningham, R.P., DasGupta, C., and Radding, C.M. (1979) Homologous pairing in genetic recombination: Complexes of recA protein and DNA. *Proc. Natl. Acad. Sci. USA* **76**, 5100-5104
13. Cox, M.M. and Lehman, I.R. (1982) RecA protein-promoted DNA strand exchange. Stable complexes of recA protein and single-stranded DNA formed in the presence of ATP and single-stranded DNA binding protein. *J. Biol. Chem.* **257**, 8523-8532
14. Howard-Flanders, P., West, S.C., and Stasiak, A. (1984) Role of recA protein spiral filaments in genetic recombination. *Nature* **309**, 215-220
15. Flory, J., Tsang, S.S., and Muniyappa, K. (1984) Isolation and visualization of active presynaptic filaments of recA protein and single-stranded DNA. *Proc. Natl. Acad. Sci. USA* **81**, 7026-7030
16. Tsang, S.S., Muniyappa, K., Azhderian, E., Gonda, D.K., Radding, C.M., Flory, J., and Chase, J.W. (1985) Intermediates in homologous pairing promoted by recA protein: Isolation and characterization of active presynaptic complexes. *J. Mol. Biol.* **185**, 295-309
17. Shibata, T., DasGupta, C., Cunningham, R.P., and Radding, C.M. (1980) Homologous pairing in genetic recombination: Formation of D loops by combined action of recA protein and a helix-destabilizing protein. *Proc. Natl. Acad. Sci. USA* **77**, 2606-2610
18. Kahn, R. and Radding, C.M. (1984) Separation of the presynaptic and synaptic phases of homologous pairing promoted by recA protein. *J. Biol. Chem.* **259**, 7495-7503
19. Cunningham, R.P., Shibata, T., DasGupta, C., and Radding, C.M. (1979) Single strands induce recA protein to unwind duplex DNA for homologous pairing. *Nature* **281**, 191-195, and **282**: 4264
20. Tsang, S.S., Chow, S.A., and Radding, C.M. (1985) Networks of DNA and recA protein are intermediates in homologous pairing. *Biochemistry* **24**, 3226-3232
21. Cunningham, R.P., Wu, A.M., Shibata, T., DasGupta, C., and Radding, C.M. (1981) Homologous pairing and topological linkage of DNA molecules by combined action of *E. coli* recA protein and topoisomerase I. *Cell* **24**, 213-223
22. Bianchi, M., DasGupta, C., and Radding, C.M. (1983) Synapsis and the formation of paranemic joints by *E. coli* recA protein. *Cell* **34**, 931-939
23. Cunningham, R.P., DasGupta, C., Shibata, T., and Radding, C.M. (1980) Homologous pairing in genetic recombination: RecA protein makes joint molecules of gapped circular DNA and closed circular DNA. *Cell* **20**, 223-235
24. DasGupta, C., Shibata, T., Cunningham, R.P., and Radding, C.M. (1980) The topology of homologous pairing promoted by recA protein. *Cell* **22**, 437-446
25. Roca, A.I. and Cox, M.M. (1990) The RecA protein: Structure and functions. *Critic. Rev. Biochem. Mol. Biol.* **25**, 415-456
26. Radding, C.M. (1991) Helical interactions in homologous pairing and strand exchange driven by recA protein. *J. Biol. Chem.* **266**, 5355-5358
27. Kowalczykowski, S.C. and Eggleston, A.K. (1994) Homologous pairing and DNA strand-exchange proteins. *Annu. Rev. Biochem.* **63**, 991-1043
28. Rao, B.J., Chiu, S.K., Bazemore, L.R., Reddy, G., and Radding, C.M. (1995) How specific is the first recognition step of homologous recombination? *Trends Biochem Sci.* **20**, 109-113
29. Sancar, A., Stachelek, C., Konigsberg, W., and Rupp, W.D. (1980) Sequences of the *recA* gene and protein. *Proc. Natl. Acad. Sci. USA* **77**, 2611-2615
30. Horii, T., Ogawa, T., and Ogawa, H. (1980) Organization and the *recA* gene of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **77**, 313-317
31. Ogawa, T., Wabiko, H., Tsurimoto, T., Horii, T., Masukata, H., and Ogawa, H. (1979) Characteristics of purified *recA* protein and the regulation of its synthesis *in vivo*. *Cold Spring Harbor Symp. Quant. Biol.* **43**, 909-915
32. Roberts, J.W., Roberts, C.W., Craig, N.L., and Phizicky, E.M. (1979) Activity of the *Escherichia coli recA*-gene product. *Cold Spring Harbor Symp. Quant. Biol.* **43**, 917-920
33. Stasiak, A. and Di Capua, E. (1982) The helicity of DNA in complexes with recA protein. *Nature* **299**, 185-186
34. Ohtani, T., Shibata, T., Iwabuchi, M., Watabe, H., Iino, T., and Ando, T. (1982) ATP-dependent unwinding of double helix in closed circular DNA by recA protein of *E. coli*. *Nature* **299**, 86-89
35. Roberts, J.W. and Roberts, C.W. (1975) Proteolytic cleavage of bacteriophage lambda repressor in induction. *Proc. Natl. Acad. Sci. USA* **72**, 147-151
36. Shinagawa, H., Mizuuchi, K., and Emmerson, T. (1977) Induc-

tion of prophage lambda by γ-rays, mitomycin C and *tif*; repressor cleavage studies by immunoprecipitation. *Mol. Gen. Genet.* **155**, 87-91
37. Roberts, J.W., Roberts, C.W., and Craig, N.L. (1978) *Escherichia coli recA* gene product inactivates phage λ repressor. *Proc. Natl. Acad. Sci. USA* **75**, 4714-4718
38. Egelman, E.H. and Yu, X. (1989) The location of DNA in recA-DNA helical filaments. *Science* **245**, 404-407
39. Takahashi, M. (1989) Analysis of DNA-recA protein interactions involving the protein self-association reaction. *J. Biol. Chem.* **264**, 288-295
40. Ruigrok, R.W.H. and Dicapua, E. (1991) On the polymerization state of recA in the absence of DNA. *Biochimie* **73**, 191-198
41. Egelman, E.H. and Stasiak, A. (1986) Structure of helical recA-DNA complexes. Complexes formed in the presence of ATP-gamma-S or ATP. *J. Mol. Biol.* **191**, 677-697
42. Egelman, E.H. and Stasiak, A. (1993) Electron microscopy of RecA-DNA complexes—Two different states, their functional significance and relation to the solved crystal structure. *Micron* **24**, 309-324
43. Hewat, E.A., Ruigrok, R.W.H., and DiCapua, E. (1991) Activation of recA protein: The pitch of the helical complex with single-stranded DNA. *EMBO J.* **10**, 2695-2698
44. Story, R.M., Weber, I., and Steitz, T.A. (1992) The structure of the *E. coli recA* protein monomer and polymer. *Nature* **355**, 318-325
45. Story, R.M. and Steitz, T.A. (1992) Structure of the recA protein-ADP complex. *Nature* **355**, 374-376
46. Harris, L.D. and Griffith, J. (1987) Visualization of the homologous pairing of DNA catalyzed by the bacteriophage T4 uvsX protein. *J. Biol. Chem.* **262**, 9285-9292
47. Ogawa, T., Yu, X., Shinohara, A., and Egelman, E.H. (1993) Similarity of the yeast RAD51 filament to the bacterial RecA filament. *Science* **259**, 1896-1899
48. Benson, F.E., Stasiak, A., and West, S.C. (1994) Purification and characterization of the human Rad51 protein, an analogue of *E. coli* RecA. *EMBO J.* **13**, 5764-5771
49. Weinstock, G.M., McEntee, K., and Lehman, I.R. (1979) ATP-dependent renaturation of DNA catalyzed by the recA protein of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **76**, 126-130
50. Bianchi, M., Riboli, B., and Magni, G. (1985) *E. coli* recA protein possesses a strand separating activity on short duplex DNAs. *EMBO J.* **4**, 3025-3030
51. Di Capua, E. and Muller, B. (1987) The accessibility of DNA to dimethylsulfate in complexes with recA protein. *EMBO J.* **6**, 2493-2498
52. Norden, B., Elvingson, C., Eriksson, T., Kubista, M., Sjöberg, B., Takahashi, M., and Mortensen, K. (1990) Structure of a RecA-DNA complex from linear dichroism and small-angle neutron-scattering in flow-oriented solution. *J. Mol. Biol.* **216**, 223-228
53. Iwabuchi, M., Shibata, T., Ohtani, T., Natori, M., and Ando, T. (1983) ATP-dependent unwinding of the double helix and extensive supercoiling by *Escherichia coli* recA protein in the presence of topoisomerase. *J. Biol. Chem.* **258**, 12394-12404
54. Shibata, T., Makino, O., Ikawa, S., Ohtani, T., Iwabuchi, M., Shibata, Y., Maeda, H., and Ando, T. (1984) Roles of processive unwinding in recombination reactions promoted by recA protein of *Escherichia coli*: A study using a monoclonal antibody. *Cold Spring Harbor Symp. Quant. Biol.* **49**, 541-551
55. Makino, O., Ikawa, S., Shibata, Y., Maeda, H., Ando, T., and Shibata, T. (1987) RecA protein-promoted recombination reaction consists of two independent processes, homologous matching and processive unwinding: A study involving an anti-recA protein-monoclonal IgG. *J. Biol. Chem.* **262**, 12237-12246
56. Hsieh, P., Camerini-Otero, C.S., and Camerini-Otero, R.D. (1990) Pairing of homologous DNA sequences by proteins: Evidence for three-stranded DNA. *Genes Dev.* **4**, 1951-1963
57. Rao, B.J., Dutreix, M., and Radding, C.M. (1991) Stable three-stranded DNA made by RecA protein. *Proc. Natl. Acad. Sci. USA* **88**, 2984-2988
58. Chiu, S.K., Rao, B.J., Story, R.M., and Radding, C.M. (1993) Interactions of three strands in joints made by RecA protein. *Biochemistry* **32**, 13146-13155
59. Adzuma, K. (1992) Stable synapsis of homologous DNA molecules mediated by the *Escherichia coli* RecA protein involves local exchange of DNA strands. *Genes Dev.* **6**, 1679-1694
60. Jain, S.K., Cox, M.M., and Inman, R.B. (1995) Occurrence of three-stranded DNA within a RecA protein filament. *J. Biol. Chem.* **270**, 4943-4949
61. Hoogsteen, K. (1959) The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta Cryst.* **12**, 822-823
62. Hanvey, J.C., Shimizu, M., and Wells, R.D. (1988) Intramolecular DNA triplexes in supercoiled plasmids. *Proc. Natl. Acad. Sci. USA* **85**, 6292-6296
63. Kohwi, Y. and Kohwi-Shigematsu, T. (1988) Magnesium ion-dependent triple-helix structure formed by homopurine-homopyrimidine sequences in supercoiled plasmid DNA. *Proc. Natl. Acad. Sci. USA* **85**, 3781-3785
64. Jain, S.K., Inman, R.B., and Cox, M.M. (1992) Putative three-stranded DNA pairing intermediate in recA protein-mediated DNA strand exchange: No role for guanine *N*-7. *J. Biol. Chem.* **267**, 4215-4222
65. Rao, B.J., Chiu, S.K., and Radding, C.M. (1993) Homologous recognition and triplex formation promoted by RecA protein between duplex oligonucleotides and single-stranded DNA. *J. Mol. Biol.* **229**, 328-343
66. Cameriniotero, R.D. and Hsieh, P. (1993) Parallel DNA triplexes, homologous recombination, and other homology-dependent DNA interactions. *Cell* **73**, 217-223
67. Rao, B.J. and Radding, C.M. (1993) Homologous recognition promoted by RecA protein *via* non-Watson-Crick bonds between identical DNA strands. *Proc. Natl. Acad. Sci. USA* **90**, 6646-6651
68. Ogawa, T., Shinohara, A., Ogawa, H., and Tomizawa, J. (1992) Functional structures of the RecA protein found by chimera analysis. *J. Mol. Biol.* **226**, 651-660
69. Kurumizaka, H., Ikawa, S., Ikeya, T., Ogawa, T., and Shibata, T. (1994) A chimeric recA protein exhibits altered double-stranded DNA-binding. *J. Biol. Chem.* **269**, 3068-3075
70. Howard-Flanders, P., West, S.C., Rusche, J.R., and Egelman, E.H. (1984) Molecular mechanisms of general genetic recombination: The DNA binding sites of recA protein. *Cold Spring Harbor Symp. Quant. Biol.* **49**, 571-580
71. Horii, T. (1991) Head to head dimer model: An alternative model for the strand exchange reaction by recA protein of *Escherichia coli*. *Biochimie* **73**, 177-185
72. Wittung, P., Norden, B., Kim, S.K., and Takahashi, M. (1994) Interactions between DNA molecules bound to RecA filament—Effects of base complementarity. *J. Biol. Chem.* **269**, 5799-5803
73. Muller, B. and Stasiak, A. (1991) RecA-mediated annealing of single-stranded DNA and its relation to the mechanism of homologous recombination. *J. Mol. Biol.* **221**, 131-145
74. Bryant, F.R. and Lehman, I.R. (1985) On the mechanism of renaturation of complementary DNA strands by the recA protein of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **82**, 297-301
75. McEntee, K. (1985) Kinetics of DNA renaturation catalyzed by the recA protein of *Escherichia coli*. *Biochemistry* **24**, 4345-4351
76. Kurumizaka, H., Rao, B.J., Ogawa, T., Radding, C.M., and Shibata, T. (1994) A chimeric RecA protein that implicates non-Watson-Crick interactions in homologous pairing. *Nucleic Acids Res.* **22**, 3387-3391
77. Morimatsu, K. and Horii, T. (1995) Analysis of the DNA binding site of *Escherichia coli* RecA protein. *Adv. Biophys.* **31**, 23-48
78. Prasad, B.V.V. and Chiu, W. (1987) Sequence comparison of single-stranded DNA binding proteins and its structural implication. *J. Mol. Biol.* **193**, 579-584
79. de Jong, E.A.M., van Duynhoven, J.P.M., Harmsen, B.J.M., Tesser, G.I., Konings, R.N.H., and Hilberst, C.W. (1989) Two-dimensional ^{1}H nuclear magnetic resonance studies on the gene V-encoded single-stranded DNA-binding protein of the filamentous bacteriophage IKe. II. Characterization of the DNA-binding wing with the aid of spin-labelled oligonucleotides. *J. Mol. Biol.* **206**, 133-152

80. Stasiak, A., Stasiak, A.Z., and Koller, T. (1984) Visualization of recA-DNA complexes involved in consecutive stages of an *in vitro* strand exchange reaction. *Cold Spring Harbor Symp. Quant. Biol.* **49**, 561-570
81. Muller, B., Koller, T., and Stasiak, A. (1990) Characterization of the DNA binding activity of stable recA-DNA complexes. Interaction between the two DNA binding sites within recA helical filaments. *J. Mol. Biol.* **212**, 97-112
82. West, S.C., Cassuto, E., Mursalim, J., and Howard-Flanders, P. (1980) Recognition of duplex DNA containing single-stranded regions by recA protein. *Proc. Natl. Acad. Sci. USA* **77**, 2569-2573
83. Wabiko, H., Ogawa, T., and Ogawa, H. (1983) Defective complex formation between single-stranded DNA and mutant *recA430* or *recA1* protein of *Escherichia coli*. *Eur. J. Biochem.* **137**, 263-267
84. Bryant, F.R. (1988) Construction of a recombinase-deficient mutant recA protein that retains single-stranded DNA-dependent ATPase activity. *J. Biol. Chem.* **263**, 8716-8723
85. Ikawa, S., Kamiya, N., and Shibata, T. (1989) Defective homologous pairing and proficient processive unwinding by the *recA430* mutant protein and intermediates of homologous pairing by recA protein. *J. Biol. Chem.* **264**, 21167-21176
86. Wu, A.M., Bianchi, M., DasGupta, C., and Radding, C.M. (1983) Unwinding associated with synapsis of DNA molecules by recA protein. *Proc. Natl. Acad. Sci. USA* **80**, 1256-1260
87. Ikeda, M., Makino, O., and Shibata, T. (1990) Epitopes and active sites of the recA protein. *J. Biol. Chem.* **265**, 8948-8956
88. Ikeda, M., Hamano, K., and Shibata, T. (1992) Epitope mapping of anti-recA protein IgGs by region specified polymerase chain reaction mutagenesis. *J. Biol. Chem.* **267**, 6291-6296
89. Makino, O., Shibata, Y., Maeda, H., Shibata, T., and Ando, T. (1985) Monoclonal antibodies with specific effects on partial activities of recA protein of *Escherichia coli*. *J. Biol. Chem.* **260**, 15402-15404
90. Morimatsu, K. and Horii, T. (1995) DNA binding surface of RecA protein: Photochemical cross-linking of the first DNA binding site on RecA filament. *Eur. J. Biochem.*, in press
91. Rao, B.J. and Radding, C.M. (1994) Formation of base triplets by non-Watson-Crick bonds mediates homologous recognition in RecA recombination filaments. *Proc. Natl. Acad. Sci. USA* **91**, 6161-6165