Nucleic Acids Research, 1994, Vol. 22, No. 16 3387–3391

# A chimeric Rec-A protein that implicates non-Watson–Crick interactions in homologous pairing

Hitoshi Kurumizaka[1,2], B.J.Rao[3], Tomoko Ogawa[4], Charles M.Radding[3] and Takehiko Shibata[1,2]*
[1]Laboratory of Cellular and Molecular Biology, The Institute of Physical and Chemical Research (RIKEN), Wako-shi, Saitama 351-01, [2]The Graduate School of Science and Engineering, Saitama University, Urawa-shi, Saitama 338, Japan, [3]Department of Genetics and Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, CT 06510, USA and [4]Department of Biology, Faculty of Science, Osaka University, Toyonaka, Osaka 560, Japan

Received May 6, 1994; Revised and Accepted July 14, 1994

## ABSTRACT

**The helical filament formed by RecA protein on single-stranded DNA plays an important role in homologous recombination and pairs with a complementary single strand or homologous duplex DNA. The RecA nucleoprotein filament also recognizes an identical single strand. The chimeric protein, RecAc38, forms a nucleoprotein filament that recognizes a complementary strand but is defective in recognition of duplex DNA, and is associated with phenotypic defects in repair and recombination. As described here, RecAc38 nucleoprotein filament is also defective in recognition of an identical strand, either when the filament has within it a single strand or duplex DNA. A model that postulates three DNA binding sites rationalizes these observations and suggests that the third binding site mediates non-Watson–Crick interactions that are instrumental in recognition of homology in duplex DNA.**

## INTRODUCTION

Rapid progress has been made in recent years in understanding the biochemistry of homologous recombination (1) which affects the inheritance of every individual, but central questions remain about the mechanisms of homologous recombination. Among these is the riddle of how homologous double-stranded DNA molecules recognize one another. In *Escherichia coli* and other bacteria, RecA protein is essential for homologous genetic recombination and *in vitro* it promotes homologous pairing of single-stranded DNA with double-stranded DNA (2, 3; see refs 1, 4–7 for reviews). To do so, RecA protein forms a right-handed nucleoprotein filament on single-stranded DNA (8, 9). This structure pairs with either naked homologous double-stranded DNA or a naked complementary strand *in vitro*. A yeast homologue of RecA protein, RAD51 protein, forms a very similar nucleoprotein filament as that of RecA protein (10), and homologues of RecA protein have been found recently in various eukaryotes including human beings (11–14).

In contrast to the pairing of complementary single strands which is mediated by Watson–Crick interactions, the way in which the RecA nucleoprotein filament recognizes homology in duplex DNA remains unknown. Two classes of model have been considered. According to the first of these, the base-pair model, the nucleoprotein filament of RecA protein might recognize homology in duplex DNA by partially disrupting base pairs to make use of Watson–Crick complementarity in one strand. According to the second class of model, the base-triplet model, the RecA nucleoprotein filament might form a three-stranded intermediate in which the duplex retains its Watson–Crick base pairs but is joined to the filament by additional non-Watson–Crick interactions.

Recent experiments demonstrated that the RecA nucleoprotein filament can also pair with a strand of identical sequence and hence revealed that the filament can readily form non-Watson–Crick bonds (15). In addition, parallel observations made with duplex DNA in the filament suggested that it also recognized homology by non-Watson–Crick interactions (15). Rao and Radding suggested that the non-Watson–Crick interactions that result in recognition of identical sequences may be a subset of the interactions that pair a single strand with duplex DNA. The behavior of a mutant RecA protein that is defective in the recognition of identical sequences might help to test this hypothesis. The chimeric protein RecAc38 forms nucleoprotein filament on single-stranded DNA that recognizes a complementary strand but is defective in the recognition of homologous duplex DNA (16), and is associated with phenotypic defects in repair and recombination (17). Here, we report further observations on RecAc38 protein that provide insight into the relation between the recognition of an identical strand and the recognition of homology in duplex DNA, and into the specific roles of each DNA-binding site of RecA protein.

*To whom correspondence should be addressed at: Laboratory of Cellular and Molecular Biology, The Institute of Physical and Chemical Research (RIKEN), Wako-shi, Saitama 351-01, Japan

## MATERIALS AND METHODS

### DNA and oligonucleotides

Circular single-stranded M13 DNA was prepared as described (18, 19). Single-stranded oligonucleotides (33-mers and heterologous 30-mer) and hairpin oligonucleotide (proximal hairpin) were the same substrates as described in refs (15) and (20), respectively. Their sequences (5′→3′) were as follows:
minus strand: 5′-GGC TTA GAG CTT AAT TGC TGA ATC TGG TGC TGT-3′
plus strand: 5′-ACA GCA CCA GAT TCA GCA ATT AAG CTC TAA GCC-3′
minus strand, reversed backbone: 5′-TGT CGT GGT CTA AGT CGT TAA TTC GAG ATT CGG-3′
plus strand, reversed backbone: 5′-CCG AAT CTC GAA TTA ACG ACT TAG ACC ACG ACA-3′
heterologous-strand: 5′-ACC CAC TCG TGC ACC CAA CTG ATC TTC AGC-3′
hairpin 70-mer: 5′-GGC TTA GAG CTT AAT TGC TGA ATC TGG TGC TGT CC CC ACA GCA CCA GAT TCA GCA ATT AAG CTC TAA GCC-3′

Oligonucleotides were labeled with [$\gamma$-$^{32}$P]-ATP by using T4 polynucleotide kinase (New England Biolabs). The amounts of DNA and oligonucleotides were expressed as moles of nucleotide residues.

### RecA protein and RecAc38 protein

Wild-type RecA protein was prepared as described previously (21, 22). RecAc38 protein is a product of a chimeric *recA* gene of *Escherichia coli* and *Pseudomonas aeruginosa* that was constructed by homologous recombination between *recA* genes of the two bacteria. RecAc38 protein was expressed on plasmid DNA under the control of the lac promoter and was extensively purified as described (16). Compared with wild-type RecA protein of *E.coli*, RecAc38 protein has the following nine amino acid substitutions: $Met_{58}$ is replaced by Lys, $Gln_{78}$ by Ser, $Ala_{82}$ by Glu, $Arg_{85}$ by Lys, $Glu_{86}$ by Gln, $Lys_{88}$ by Ala, $Ile_{93}$ by Val, $Ile_{102}$ by Asp, and $Arg_{105}$ by Gly (17). The measurements of the amounts of wild-type and chimera RecA proteins were described previously (16).

### Reaction conditions

The standard reaction mixture contained 30 mM HEPES (pH 7.2), 15 mM Mg acetate, 2 mM dithiothreitol, 1−6% glycerol (derived from RecA protein preparation), 1.2 mM ATP, 100 μg/ml bovine serum albumin, and an ATP regenerating system, consisting of 8 mM phosphocreatine and 10 U/ml creatine phosphokinase.

### Assay for the pairing of single-stranded DNA with a duplex hairpin oligonucleotide

M13 circular single-stranded DNA (2 μM in nucleotide residues) and RecA protein (4 μM) were incubated at 37°C for 12 min in the standard reaction buffer, and then the pairing reaction was started by addition of $^{32}$P-labeled hairpin double-stranded oligonucleotide 70-mer (40 nM) which had a hairpin at the proximal end. The molar ratio of homologous regions of single-stranded DNA:double-stranded DNA was 1:2. After the reaction at 37°C, the products (20 μl) were treated with 1% SDS, and then diluted with 5 ml of 1.5 M NaCl and 0.15 M sodium citrate (10× SSC). The amount of joint molecules formed in the reaction was determined by trapping $^{32}$P-labeled hairpin double-stranded oligonucleotide on a nitrocellulose filter in the presence of 10× SSC (D-loop assay; 20, 21).

### Assay for the pairing of single-stranded DNA and single strand oligonucleotides with non-Watson−Crick interactions

M13 circular single-stranded DNA (20 μM) and the indicated amount of RecA protein were incubated at 37°C for 12 min in the standard reaction buffer. After the incubation period, heterologous oligonucleotide 30-mer (200 μM) was added and then the reaction was started by the addition of $^{32}$P-labeled oligodeoxyribonucleotide 33-mer (0.2 μM). After the reaction, 5 μl of reactant were diluted with 95 μl of washing buffer (30 mM HEPES, 15 mM Mg acetate, 2 mM dithiothreitol, 1.5 mM ATP) and then filtered by centrifugation through cellulose filter units that retain molecules larger than 30 kDa (Millipore Ultrafree-MC cellulose filters; centrifugal filtration assay; 15). The cellulose filter units were washed with 50 μl of the washing buffer. The $^{32}$P-labeled oligonucleotide which formed a homology-dependent complex with RecA protein and M13 circular single-stranded DNA was trapped o the cellulose filter, but free oligonucleotide was not.

### Assay for strand exchange associated with homologous pairing

[5′-$^{32}$P]Minus strand was annealed to circular plus single-stranded DNA and the annealed product was purified by gel filtration (see Fig. 1B). The RecA nucleoprotein filaments were formed under the standard reaction conditions containing ATP regeneration with the indicated amounts of RecA protein and 10 μM circular single-stranded M13 DNA with the annealed labeled oligonucleotide. Pairing and strand exchange were initiated by adding 1.2 μM single-stranded 33-mer oligonucleotide and 100 μM heterologous single-stranded 30-mer oligonucleotide. After the incubation for the indicated time, 30 μl aliquots were treated with 20 mM EDTA (pH 8.0), 0.5% SDS and 200 μg/μl proteinase K, and incubated at 37°C for an additional 20 min. The samples were analyzed by electrophoresis on a polyacrylamide gel under nondenaturing conditions in a cold room (5% polyacrylamide in spacer gel plus 15% polyacrylamide in the bottom gel, 89 mM Tris−borate, pH 7.8, and 2 mM EDTA as buffer, 48×20×0.1 cm, 600 V, 150 mA, 18 h). The gel was dried and autoradiographed. The counts associated with labeled substrates and products were quantitated in the dried gel with a PhosphorImager (Molecular Dynamics).

## RESULTS

### RecAc38 protein is defective in the recognition of identical strands.

We tested the pairing of M13 single-stranded DNA (the plus strand) and various 33-mer oligonucleotides (Fig. 1). We first examined the ability of the RecAc38 filament formed on single-stranded M13 DNA to pair with a duplex oligonucleotide or a complementary minus strand oligonucleotide, to be sure that the pairing of oligonucleotides followed the pattern observed with larger substrates (16). Using standard conditions of reaction and assay we compared the pairing of a hairpin oligonucleotide with nucleoprotein filaments formed on single-stranded M13 DNA by wild-type and the chimeric RecAc38 protein. There was a three- to four-fold reduction in the yield of joint molecules formed with a duplex oligonucleotide by RecAc38 filaments vs wild-type filaments (Fig. 2).

To examine the ability of RecAc38 filaments to pair with the oligonucleotides illustrated in Fig. 1, we used an assay that can detect weak interactions such as the pairing of identical strands (15). RecAc38 protein formed double-stranded DNA from plus single-stranded DNA and the minus strand oligonucleotide with almost the same efficiency as the wild-type RecA protein (Fig. 3A).

The RecAc38 filament was specifically defective in the recognition of a plus strand (i.e. an identical strand), which was reduced by a factor of two relative either to its recognition of a complementary minus strand, or to the recognition of either strand by a wild-type filament (Fig. 3A).

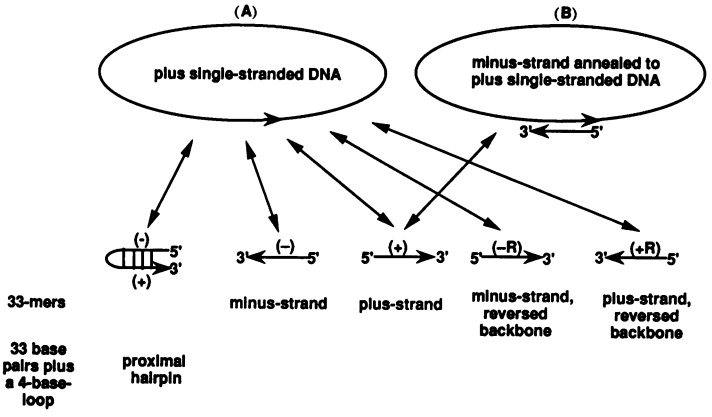


**Figure 1.** DNA substrates. Circular single-stranded M13 DNA, designated as plus single-stranded DNA (**A**) The oligonucleotides used were 33-mers; plus strand (+) and minus strand (−) had the same nucleotide sequence as M13 plus strand and minus strand, respectively. Plus strand with reversed backbone (+R) and minus strand with reversed backbone (−R) were oligonucleotides that contained the same linear arrays of nucleotide residues as in (+) and (−), respectively, but that had the reverse orientation in their sugar−phosphate backbones. In some experiments, plus single-stranded DNA was replaced by plus single-stranded DNA to which the minus strand had been annealed (**B**). Sequences of the oligomers are described in Materials and Methods.

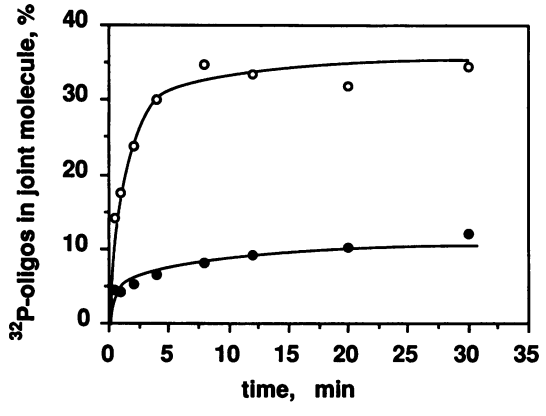


**Figure 2.** Defective homologous pairing by RecAc38 protein of single-stranded DNA with a duplex hairpin oligonucleotide. M13 circular single-stranded DNA (2 μM), $^{32}$P-labeled hairpin oligonucleotide (40 nM) and ATP (1.2 mM) were incubated with 4 μM wild-type RecA protein (○) or RecAc38 protein (●) at 37°C. At the indicated times, 20 μl of the reaction mixture were withdrawn and the joint molecules formed were measured by a D-loop assay.

## Defective recognition of the plus strand precludes strand exchange

Since one of our principle aims was to determine if the recognition of an identical plus strand by the RecA filament is related to the

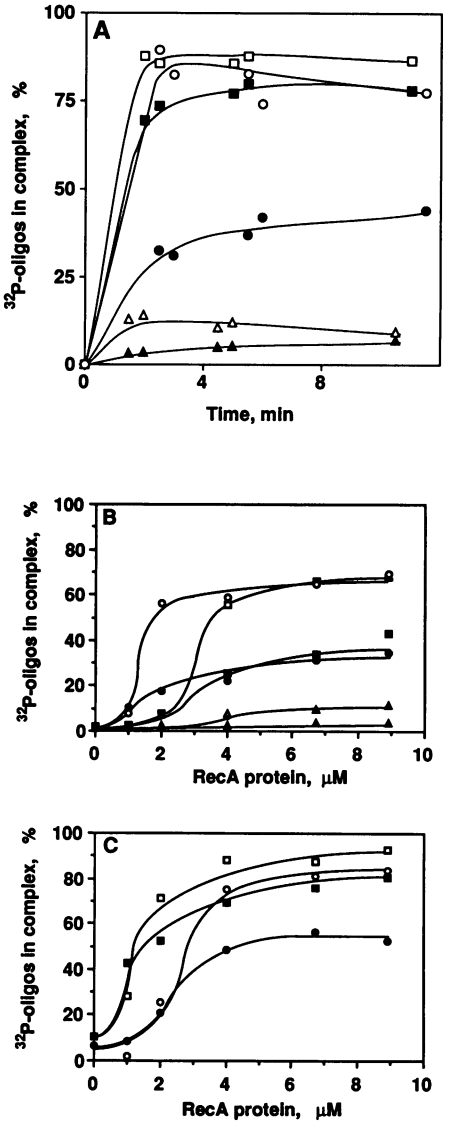


**Figure 3.** RecAc38 protein is defective in various non-Watson−Crick pairings of single strands. RecA protein was incubated with 20 μM M13 plus single-stranded DNA and ATP at 37°C for 12 min in the standard reaction mixture (20 μl) before the pairing reaction was started by addition of the [$^{32}$P]oligonucleotides (0.2 μM) diagramed in Fig. 1. Excess unlabeled heterologous oligonucleotide (200 μM) was added as a nonspecific competitor just before the labeled homologous oligonucleotide. Pairing of oligonucleotides with the plus single-stranded DNA was assayed by a centrifugal filtration assay (See Materials and Methods). Open symbols, wild-type RecA protein; closed symbols, RecAc38 protein. (**A**) Recognition of an identical plus strand *vs* a complementary minus strand. The concentration of RecA protein was 6.7 μM. [$^{32}$P]oligonucleotides: (■, □) complementary minus strand (Watson−Crick base pairing); (●, ○) identical plus strand (non-Watson−Crick base pairing); (▲, △) control labeled heterologous 30-mer. Other non-Watson−Crick pairings. Incubation for the pairing reaction was 3 min. (**B**) [$^{32}$P]oligonucleotides: (■, □) plus strand; (●, ○) plus strand with reversed backbone; (▲, △) control labeled heterologous 30-mer. (**C**) [$^{32}$P]oligonucleotides: (■, □) complementary minus strand control; (●, ○) minus strand with reversed sugar phosphate backbone.

recognition of homologous duplex DNA, we asked whether the RecAc38 filament is also defective in recognition of the plus strand when the filament contains duplex DNA instead of a single strand (See Fig. 1B). To do so, we made use of a previous observation which showed that the wild-type filament containing duplex DNA promoted strand exchange when a plus strand oligonucleotide was added (15). RecAc38 nucleoprotein filaments were formed on M13 plus single-stranded DNA to which a labeled minus strand had been annealed (Fig. 1B). After addition of an unlabeled plus strand, exchange was assessed by the appearance upon gel electrophoresis of a labeled band corresponding to a duplex oligonucleotide (Fig. 4A). At saturating levels of wild-type protein, 42% of label was converted to the duplex oligonucleotide in 30 min, whereas only 14% was converted when RecAc38 protein was used (Fig. 4B). A time-course showed that the rate of strand exchange promoted by RecAc38 protein was a third to one half of that of wild-type (Fig. 4C). As shown by an internal control, the filaments that were formed in this experiment by wild-type protein and RecAc38 protein on M13 plus single-stranded DNA were equally competent in taking up a labeled minus strand (Figs 4A and B).

### The defect of RecAc38 protein extends to other non-Watson–Crick pairings

The wild-type RecA nucleoprotein filament containing a plus single-stranded DNA recognizes not only a complementary minus strand or a plus strand, but it also recognizes either of those sequences with their backbones reversed (15; Fig. 3). The RecAc38 filament by contrast was defective in the recognition of either strand with a reversed backbone. Thus, RecAc38 protein was defective in pairing all three non-Watson–Crick combinations that were presented.

A further effect suggests that both the wild-type and chimera nucleoprotein filaments recognize parallel and antiparallel strands differently. Titrations of wild-type and chimeric RecAc38 protein revealed that a greater threshold concentration of protein, either wild-type or chimera, was required to form complexes between parallel strands than between antiparallel strands (Figs 3B and C).

## DISCUSSION

In a previous study, we found that RecAc38 protein is proficient in the renaturation of complementary single-stranded DNA, but is defective in homologous pairing of single-stranded DNA with

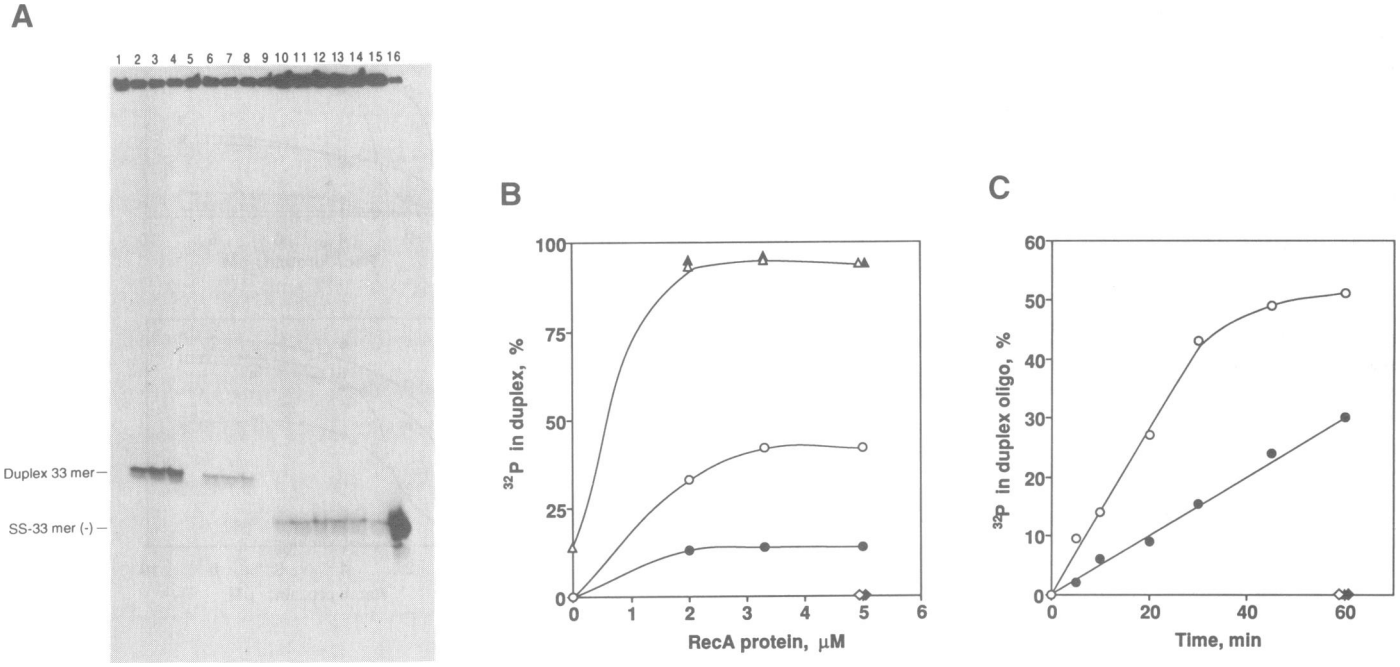


**Figure 4.** Defective recognition of the plus strand precludes strand exchange. (**A, B**) The RecA nucleoprotein filaments were formed under the standard reaction conditions containing ATP regeneration (see legend to Fig. 3) with RecA protein and the partial double-stranded DNA (10 μM). Pairing and strand exchange were initiated by adding 1.2 μM single strand (plus strand in lanes 1–4 and 6–8; minus strand in lanes 5 and 9) and 100 μM heterologous single-stranded 30-mer. After 30 min of incubation, 30 μl aliquots were analyzed by gel electrophoresis. Lanes 2–5, with wild-type RecA protein; lanes 6–9, RecAc38 protein, at 2 μM, 3.3 μM, 5 μM, 5 μM, respectively. Lane 1, without RecA. As controls, simple complementary pairing reactions were performed in parallel with plus single-stranded DNA (10 μM). Pairing was initiated by adding [$^{32}$P]-labeled minus strand (0.05 μM) and unlabeled heterologous 30-mer (100 μM). After 10 min of incubation, 30 μl aliquots were analyzed by gel electrophoresis. Lanes 10–12, with wild-type RecA protein; lanes 13–15, RecAc38 protein at 2 μM, 3.3 μM, 5 μM, respectively. Lane 16, without RecA. The gel was dried and autoradiographed. The counts associated with labeled substrates and products were quantitated in the dried gel with a PhosphorImager (Molecular Dynamics), and plotted in (B). ○, strand exchange by wild-type RecA protein; ●, strand exchange by RecAc38 protein. Controls: △, pairing of complementary single strands by wild-type RecA protein; ▲, pairing of complementary single strands by RecAc38 protein; ◇, wild-type RecA filament containing double-stranded DNA without the addition of single-stranded; ◆, RecAc38 filament containing double-stranded DNA without the addition of single-stranded. (**C**) Time course of strand exchange. As described above, filaments were formed by 7.5 μM RecA protein on 15 μM circular plus single-stranded DNA to which had been annealed [5′-$^{32}$P]minus strand. Pairing and strand exchange were initiated by adding 1.8 μM plus strand and 150 μM heterologous 30-mer. At various times during incubation, 38 μl aliquots were analyzed and the extent of strand exchange was quantitated as described above. ○, wild-type RecA protein; ●, RecAc38 protein. Controls: ▲, no RecA protein; ◇, wild-type RecA and unlabeled minus strand in place of plus strand; ◆, RecAc38 and unlabeled minus strand in place of plus strand.

double-stranded DNA (16). A step in the homologous pairing defective in RecAc38 protein was identified as the binding of double-stranded DNA to RecAc38–nucleoprotein filament. To explain these observations, we postulated three DNA-binding sites on RecA protein (16). The observations made in the present study enable us to specify the role of each DNA-binding site in homologous pairing. Site 1 is the well-known strong binding site for single-stranded DNA that underlies the formation of the nucleoprotein filament. Site 2 is occupied specifically by a strand that is both complementary and antiparallel to the strand in site 1, i.e. the complementary strand of Watson–Crick DNA. Non-Watson–Crick interactions are mediated by site 3, which has less stringent requirements. Site 3 is normally occupied by a parallel plus strand, but it can also be occupied by that strand with a reversed backbone; site 3 can in addition be occupied by a complementary strand with a reversed backbone, whose polarity causes it to be rejected by site 2.

In accord with experimental data, the model holds that recognition of an identical parallel strand and the recognition of oligonucleotides with reversed backbones are mediated by site 3 whether only site 1 is filled, or both sites 1 and 2 are filled. When site 1 is occupied by a plus single strand followed by the binding of duplex DNA, the antiparallel strand of the duplex sits in site 2, its parallel strand sits in site 3, and the identical sequence is scanned via non-Watson–Crick interactions between the strands in site 1 and site 3. Finally, new Watson–Crick base-pairs are formed between strands in site 1 and site 2 to complete homologous pairing. Mutational alteration of site 3 in RecAc38 protein accounts for the defect in recognition of homologous duplex DNA, as well as the defect in recognition of a second plus strand by itself.

Thus, this model rationalizes the observations made on homologous recognition by non-Watson–Crick interactions (15), and the observations made on the properties of RecAc38 protein reported here and previously (16). The unusual specificity of the third site, as revealed by the chimera, suggests that binding there results principally from non-Watson–Crick interactions of homologous bases, and supports the view that such interactions are important in recognition of homology. Since the mutant strain, recAc38, is defective in repair and recombination, these results further suggest that recognition via non-Watson–Crick interactions is important *in vivo*.

## ACKNOWLEDGEMENTS

This study was partly supported by a grant for the Biodesign Research Program from RIKEN and one from the Ministry of Education, Science and Culture, Japan. H.K. is supported by a fellowship from the Japan Society for the Promotion of Science. The contributions of B.J.R. and C.M.R. were supported by grant RO1-GM33504 from the National Institutes of Health, USA.

## REFERENCES

1. West, S. C. (1992) *Annu. Rev. Biochem*, **61**, 603–640.
2. McEntee, K., Weinstock, G. M. and Lehman, I. R. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 2615–2619.
3. Shibata, T., DasGupta, C., Cunningham, R. P. and Radding, C. M. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 1638–1642.
4. Radding, C. M. (1989) *Biochim. Biophys. Acta*, **1008**, 131–145.
5. Roca, A. I. and Cox, M. M. (1990) *Crit. Rev Biochem. Mol. Biol.*, **25**, 415–456.
6. Kowalczykowski, S. C. (1991) *Annu. Rev. Biophys. Biophys. Chem.*, **20**, 539–575.
7. Radding, C. M. (1991) *J. Biol. Chem.*, **266**, 5355–5358.
8. Flory, J., Tsang, S. S., Muniyappa, K., Bianchi, M., Gonda, D., Kahn, R., Azhderian, E., Egner, C., Shaner, S. and Radding, C. M. (1984) *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 513–523.
9. Tsang, S. S., Muniyappa, K., Azhderian, E., Gonda, D. K., Radding, C. M., Flory, J. and Chase, J. W. (1985) *J. Mol. Biol.*, **185**, 295–309.
10. Ogawa, T., Yu, X., Shinohara, A. and Egelman, E. H. (1993) *Science*, **259**, 1896–1899.
11. Bezzubova, O., Shinohara, A., Mueller, R. G., Ogawa, H. and Buerstedde, J. M. (1993) *Nucleic Acids Res*, **21**, 1577–1580.
12. Shinohara, A., Ogawa, H., Matsuda, Y., Ushio, N., Ikeo, K. and Ogawa, T. (1993) *Nature Genet.*, **4**, 239–243.
13. Bishop, D. K., Park, D., Xu, L. Z. and Kleckner, N. (1992) *Cell*, **69**, 439–456.
14. Story, R. M., Bishop, D. K., Kleckner, N. and Steitz, T. A. (1993) *Science*, **259**, 1892–1896.
15. Rao, B. J. and Radding, C. M. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 6646–6651.
16. Kurumizaka, H., Ikawa, S., Ikeya, T., Ogawa, T. and Shibata, T. (1994) *J. Biol. Chem.*, **269**, 3068–3075.
17. Ogawa, T., Shinohara, A., Ogawa, H. and Tomizawa, J. (1992) *J. Mol. Biol.*, **226**, 651–660.
18. Cunningham, R. P., DasGupta, C., Shibata, T. and Radding, C. M. (1980) *Cell*, **20**, 223–235.
19. DasGupta, C., Shibata, T., Cunningham, R. P. and Radding, C. M. (1980) *Cell*, **22**, 437–446.
20. Rao, B. J., Chiu, S. K. and Radding, C. M. (1993) *J Mol Biol*, **229**, 328–343.
21. Shibata, T., Cunningham, R. P. and Radding, C. M. (1981) *J. Biol. Chem.*, **256**, 7557–7564.
22. Shibata, T., Osber, L. and Radding, C. M. (1983) *Methods Enzymol.*, **100**, 197–209.