REFERENCES AND NOTES

1. M. H. Carr, *Nature* **326**, 30 (1987); S. W. Squyres, *Icarus* **79**, 229 (1989).
2. M. H. Carr, *Icarus* **68**, 187 (1986).
3. S. M. Clifford, thesis, University of Massachusetts (1984).
4. B. K. Lucchitta, H. M. Ferguson, C. Summers, *J. Geophys. Res. (Suppl.)* **91**, E166 (1987); T. J. Parker, R. S. Saunders, D. M. Schneeberger, *Icarus* **82**, 111 (1989); V. R. Baker *et al.*, *Nature* **352**, 589 (1991).
5. R. Greeley and B. D. Schneid [*Lunar Planet. Sci.* **XXII**, 489 (1991)] estimated a volume of 384 × 10^6 km^3 for volcanic and plutonic rocks, containing 88.8 m of water. Greeley and Schneid [*Science* **254**, 996 (1991)] later revised their volume measurement upward to 654 × 10^6 km^3. Although they did not revise the water estimate, this estimate would correspond to 150 m of water.
6. M. B. McElroy and Y. L. Yung, *Planet. Space Sci.* **24**, 1107 (1976); E. Anders and T. Owen, *Science* **198**, 453 (1977); M. B. McElroy, M. B. Kong, T. Y. Yung, Y. L. Yung, *J. Geophys. Res.* **82**, 4379 (1977); S. I. Rasool and L. LeSergeant, *Nature* **266**, 822 (1977); T. Owen, J. P. Maillard, C. De Bergh, B. L. Lutz, *Science* **240**, 1767 (1988); Y. L. Yung *et al.*, *Icarus* **76**, 146 (1988).
7. T. A. Scambos and B. M. Jakosky, *J. Geophys. Res.* **95**, 14779 (1990).
8. J. B. Pollack and D. C. Black, *Science* **205**, 56 (1979); M. H. Carr, *Icarus* **87**, 210 (1990); B. M. Jakosky, *ibid.* **94**, 14 (1991).
9. H. Y. McSween, Jr., *Rev. Geophys.* **23**, 391 (1985), and references therein.
10. H. Wanke and G. Dreibus, *Philos. Trans. R. Soc. London Ser. A* **325**, 545 (1988).
11. G. Dreibus and H. Wanke, *Icarus* **71**, 225 (1987); M. H. Carr and H. Wanke, *ibid.* **98**, 61 (1992).
12. C. F. Chyba, *Nature* **330**, 632 (1987).
13. A. E. Fallick *et al.*, *Lunar Planet. Sci.* **XIV**, 183 (1983); J. Yang and S. Epstein, *ibid.* **XVI** (Suppl. A), 25 (1985); J. F. Kerridge, *ibid.* **XIX**, 599 (1988).
14. H. R. Karlsson, R. N. Clayton, E. K. Gibson, Jr., T. K. Mayeda, *Science* **255**, 1409 (1992).
15. T. J. McCoy, G. J. Taylor, K. Keil, *Geochim. Cosmochim. Acta* **56**, 3571 (1992).
16. Z. Qin, F. Lu, A. T. Anderson, Jr., *Am. Mineral.* **77**, 565 (1992).
17. M. C. Johnson, M. J. Rutherford, P. C. Hess, *Geochim. Cosmochim. Acta* **55**, 349 (1991).
18. R. J. Floran *et al.*, *ibid.* **42**, 1213 (1978); A. H. Treiman, *Meteoritics* **20**, 229 (1985).
19. R. P. Harvey and H. Y. McSween, Jr., *Earth Planet. Sci. Lett.* **111**, 467 (1992).
20. Calculated solidification histories ranging from 55 to 75% crystallization of the host mineral pigeonite, corresponding to an initial water content of 1.8 to 1.0%, give plausible parental magma compositions. The mean value of 1.4% is consistent with the observation that the proportions of amphibole in Shergotty inclusions are similar to those of Chassigny.
21. R. P. Harvey, M. Wadhwa, H. Y. McSween, Jr., G. Crozaz, in preparation.
22. A. Jambon and J. L. Zimmerman, *Earth Planet. Sci. Lett.* **101**, 323 (1990).
23. R. P. Harvey and H. Y. McSween, Jr., unpublished data.

24. This work was supported in part by National Aeronautics and Space Administration grant NAG 9-58.

12 November 1992; accepted 15 January 1993

Structural Relationship of Bacterial RecA Proteins to Recombination Proteins from Bacteriophage T4 and Yeast

Randall M. Story, Douglas K. Bishop,* Nancy Kleckner, Thomas A. Steitz†

RecA protein is essential in eubacteria for homologous recombination and promotes the homologous pairing and strand exchange of DNA molecules in vitro. Recombination proteins with weak sequence similarity to bacterial RecA proteins have been identified in bacteriophage T4, yeast, and other higher organisms. Analysis of the primary sequence relationships of DMC1 from *Saccharomyces cerevisiae* and UvsX of T4 relative to the three-dimensional structure of RecA from *Escherichia coli* suggests that both proteins are structural homologs of bacterial RecA proteins. This analysis argues that proteins in this group are members of a single family that diverged from a common ancestor that existed prior to the divergence of prokaryotes and eukaryotes.

RecA protein plays a central role in recombination and DNA repair in prokaryotes. RecA forms a nucleoprotein filament and catalyzes adenosine 5′-triphosphate (ATP)–dependent recognition, pairing, and exchange of strands between two homologous DNA molecules (*1–4*). Bacterial RecA proteins are all closely related (56 to 100% sequence identity) (*2*). Bacteriophage T4 UvsX protein is more distantly related to the bacterial proteins (~20% amino acid identity) (*2, 5*) and has in vitro activities similar but not identical to those of RecA (*6–8*). In the yeast *S. cerevisiae*, three RecA-related proteins have been identified: DMC1, RAD51, and RAD57 (*9–11*). DMC1 has a meiosis-specific function required for meiotic recombination, synaptonemal complex formation, and progress through the meiotic divisions (*9*). The recombination role may be primary, and *dmc1* mutants accumulate recombination intermediates (double-stranded ends with 3′ single-stranded tails) compatible with a defect in strand invasion or exchange (*9*). RAD51 and RAD57 are required both for mitotic repair of DNA double-strand breaks and for full levels of meiotic recombination (*10, 11*). RAD51 and DMC1 are closely related to one another (~45% identity), have less similarity with prokaryotic RecA proteins (~20 to 30% identity), and are even less similar to T4 UvsX (*9, 11*). RAD57 also shows significant sequence similarity to RAD51 and RecA proteins (*12*). Very recently, *RAD51*-like genes have been identified in a number of higher organisms (*13, 14, 15*) and a *DMC1*-like gene has been identified in lily (*16*).

Although the primary sequence relationships among these proteins are suggestive, we have sought further evidence of an overall structural relationship among them by comparing the primary sequences of bacterial RecAs, yeast DMC1, and T4 UvsX with reference to the crystal structure of *E. coli* RecA protein (*17, 18*). One type of evidence for overall structural similarity among DMC1, UvsX, and RecA proteins is provided by mapping the primary sequences of the yeast and T4 proteins onto the tertiary structure of RecA. For DMC1, the primary sequence alignment used for this analysis was that derived previously in the absence of structural information (*9*) with a few minor modifications (Fig. 1). The DMC1-RecA comparison reveals several features that are compatible with an overall similarity in tertiary structure. (i) All of the small insertions and deletions present in the primary sequence alignment occur in surface loops or at the ends of α-helices (Fig. 2A). (ii) Larger insertions and deletions occur at the NH_2- and COOH-termini (Figs. 1 and 2A). DMC1 lacks the COOH-terminal portion of RecA protein and has 32 additional NH_2-terminal residues. In RecA protein, each of these regions represents a structural domain that protrudes away from the body of the protein in both the monomer and polymer (*17*). A proteolytic fragment of RecA protein lacking much of the COOH-terminal domain is proficient in strand exchange (*19*), and this domain may not be essential for homologous pairing and strand exchange but may instead be involved in the regulation of DNA binding and the SOS response in *E. coli* (*17, 20*). (iii) Charged residues occur on the surface of the protein, whereas hydrophobic residues occur mainly in the interior (Fig. 2B).

Similar conclusions pertain to UvsX. A published alignment of UvsX with RecA (*2*) showed the same general features as the

R. M. Story and T. A. Steitz, Departments of Molecular Biophysics and Biochemistry and Chemistry and Howard Hughes Medical Institute, Yale University, New Haven, CT 06511.
D. K. Bishop and N. Kleckner, Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA 02138.

*Present address: Department of Radiation and Cellular Oncology, University of Chicago, Chicago, IL 60637.
†To whom correspondence should be addressed.

DMC1-RecA alignment; however, comparison with the crystal structure allowed us to refine this alignment. The positions of some small insertions and deletions were altered to place them in less disruptive locations, and minor changes were made to keep charged residues on the surface and hydrophobic residues primarily buried in the interior of the protein. The unmodified alignment also supports the major conclusions described, and in particular, conserved residues in classes I and IV are not affected by the changes introduced. The revised alignment is used in the comparisons described below. The most significant difference between RecA and UvsX occurs in the COOH-terminal domain, in which sequence similarity is low and the UvsX sequence can only be accommodated by introducing several sizable insertions relative to the RecA structure (Fig. 1).

A second type of evidence for overall structural similarity among DMC1, UvsX, and RecA proteins is provided by analysis of residues that are invariant or conserved among the three proteins within the central region shared among them (residues 14 to 272 of *E. coli* RecA). These residues fall into several distinct and reasonable categories, moreover, two sets of very highly conserved residues [the adenosine triphosphatase (ATPase) active site and an additional structural motif] occur at widely separated positions in the molecule.

Among 23 bacterial RecA species, 136 residues (53%) in this region are invariant or conserved; when the comparison is expanded to include DMC1 and UvsX, the relevant subset includes 43 of these 136 residues (31%). These 43 residues are widely distributed throughout the primary sequence (Fig. 1), providing additional evidence that the proteins in this group are similar in overall tertiary structure. All but 2 of these 43 residues can be grouped into four classes (Fig. 3A):

In class I, invariant residues cluster in and around the ATP-binding site (Fig. 3B). Ten of the 13 residues that are invariant occur in this region, and Thr^{73} and Gln^{194} are invariant except for conservative substitutions (by Ser and His, respectively) in UvsX. All of these residues are known or proposed to be involved in ATP binding, ATP hydrolysis, or the conformational change hypothesized to be induced by ATP binding, or all of these (Table 1) (*18*).

In class II, residues that are always highly conserved but not invariant occur in the hydrophobic core of the major domain of RecA protein. Of the 30 such highly conserved residues, 25 are large hydrophobic residues (Val, Leu, Ile, Phe, and Met), all but 5 of which are mostly buried, with solvent-accessible surfaces (*21*) of 10 $Å^2$ or less.

In class III, the regions involved in forming the interface between monomers in the RecA polymer (*17*) are less well conserved. However, the remaining five highly conserved large hydrophobic residues not included in class II (Leu^{14}, Phe^{21}, Leu^{99}, Leu^{114}, and Leu^{132}) all lie in the subunit interface (*17*). In addition, the largely hydrophobic character of the surfaces forming the interface is maintained in both DMC1 (Fig. 2B) and UvsX (*22*).

In class IV, four of the remaining six highly conserved and invariant residues cluster together in a structurally unusual region of RecA protein (Fig. 4). The four conserved residues are found in a loop between helix B and β-strand 1 and in regions around this loop. The loop contains the sequence Gly-Ala-Gly-Gly and snakes through a region in close proximity to other regions of the structure. The two glycines (Gly^{54} and Gly^{55}) are invariant, contain unusual Phi and Psi backbone torsion angles, and are in a region with no room for a side chain. Asp^{48} (conserved as Asx) is completely buried and makes a hydrogen bond to the side chain of Thr^{42} (conserved as Thr or Ser). The Gly^{54}-Gly^{55} peptide stacks on the Asp^{48} side chain, and the backbone N-H of Gly^{54} makes a hydrogen bond with the backbone C=O of Asp^{48}. This region of the protein is distant from the ATP-binding site and toward the outside of the polymer (distant from the presumed DNA binding sites).

The above analysis provides good evidence for overall structural similarity and, hence, homology among DMC1, UvsX, and bacterial RecA proteins. Presumably, all three proteins evolved from a single common ancestor existing prior to the divergence of prokaryotic and eukaryotic organisms. The specific characters of residues common to all three proteins is consistent

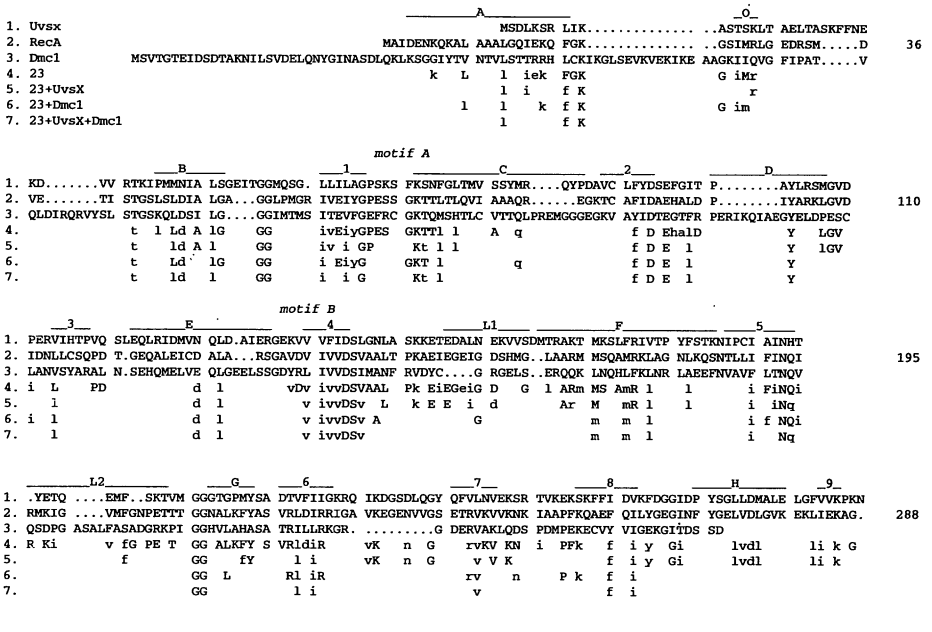

Fig. 1. Alignment of UvsX and DMC1 to *E. coli* RecA and identification of invariant and conserved residues among RecA homologs. Primary sequence alignments of UvsX to RecA and of DMC1 to RecA are shown (lines 1, 2, and 3). The alignments were generated with the three-dimensional structure of RecA protein as a guide to modify previously published alignments (*2, 9*) as described. The locations of RecA secondary structure elements are indicated above the alignment (refer to Fig. 2A). Numbers to the right of the alignment refer to the *E. coli* RecA sequence. The locations of the two purine nucleotide consensus sequences (*34*) are indicated as motif A and motif B. Invariant and conserved amino acids were identified by comparing the *E. coli* RecA sequence with homologous sequences with the alignment shown as well as an alignment of 23 bacterial sequences (*35*). The bacterial sequence alignment is an updated version of an alignment of 16 sequences (*2*) which includes additional sequences from *H. influenzae, L. pneumophila, N. gonorrhoeae, P. cepacia, B. fragilis, B. pertussis,* and *M. flagellatum*. The results of four such comparisons are shown on lines 4 through 7. Invariant residues are indicated in uppercase and residues always conserved in lowercase. Conservative substitutions were defined with the use of the similarity matrix of McLachlan (*36*), and a similarity cutoff of 5. (In addition, His was considered to be a conservative substitution for Gln.) Line 4: invariant and conserved residues among 23 bacterial sequences; line 5: invariant and conserved residues among 23 bacterial sequences and UvsX; line 6: invariant and conserved residues among 23 bacterial sequences and DMC1; and line 7: conserved residues among 23 bacterial sequences, UvsX, and DMC1.

with this view. Clustering of identities around the ATPase active site (class I) is appropriate, because substrate binding and catalysis both require the presence and precise positioning of appropriate functional groups. In contrast, residues in classes II and III are primarily large hydrophobic residues important for tertiary or quaternary structure, and such residues are expected to be conserved in size and polarity rather than in their specific identity. Similarly, class IV residues are conserved but do vary slightly in ways compatible with their apparent structural role. The primary sequence similarities among these proteins are not strong, but are comparable to those observed in other well-studied families of proteins (*23–26*).

Because all three proteins examined play a role in recombination in their respective organisms, it seems likely that they will share many important activities, as UvsX and RecA are already known to do. The structural comparison suggests that all of these proteins almost certainly share common mechanisms for binding and hydrolyzing ATP and coupling ATP hydrolysis to a conformational change (*18*) (Table 1). All three should also share the ability to make filaments: hydrophobic residues at the subunit interface are among those conserved, although the (putative) interfaces appear to have diverged considerably. UvsX is known to make filaments (*8*), and a very recent report documents the formation of RAD51 filaments (*27*). Structural analysis of RecA has not yet provided information about the protein's interaction with DNA. However, two loops that lie toward the center of the RecA polymer (and thus near the DNA) are candidates for relevant regions (L1 and L2, Fig. 2A) (*17*). The region in and around these loops is not highly conserved in UvsX and DMC1 and shows no obvious pattern of positively charged residues (Fig. 1). It is possible that DNA is bound by contacts with the polypeptide backbone that are not reflected in amino acid conservation.

Although the comparison of DMC1, UvsX, and bacterial RecA proteins described above is useful in identifying residues conserved throughout this highly diverged family, such a comparison would not reveal relationships that exist uniquely between any two members of this family. Therefore, we compared DMC1 and UvsX separately to the bacterial RecA proteins in order to identify the conserved and invariant residues (Fig. 1) and located these residues on the structure of *E. coli* RecA protein (Fig. 3, C and D). Of the 156 residues highly conserved or invariant over the entire length of bacterial RecA proteins, 87 (56%) are conserved in UvsX and 64 (41%) in DMC1. In both cases, as above, a number of invariant residues cluster largely around the ATP-binding site, and highly conserved residues occur mainly in the hydrophobic interior of the protein. However, the overall distribution of conserved and invariant residues differs somewhat in the two pairwise comparisons. In the DMC1-RecA comparison, additional invariant residues occur on the face closest to the interior of the RecA filament. In contrast, in the UvsX-RecA comparison, additional residues are more dispersed, with some residues clustered around the disordered region L1. DMC1 and UvsX are more divergent from one another than they are from the bacterial RecA proteins (~10% identity) (Fig. 1).

There are both genetic and biochemical reasons to suggest that the activities of the three proteins will also differ somewhat, presumably because they have each become specialized or adapted for slightly different biological roles. Biochemical comparison of RecA and UvsX illustrates the point that relatively subtle variations in biochemical behavior might have profound influences

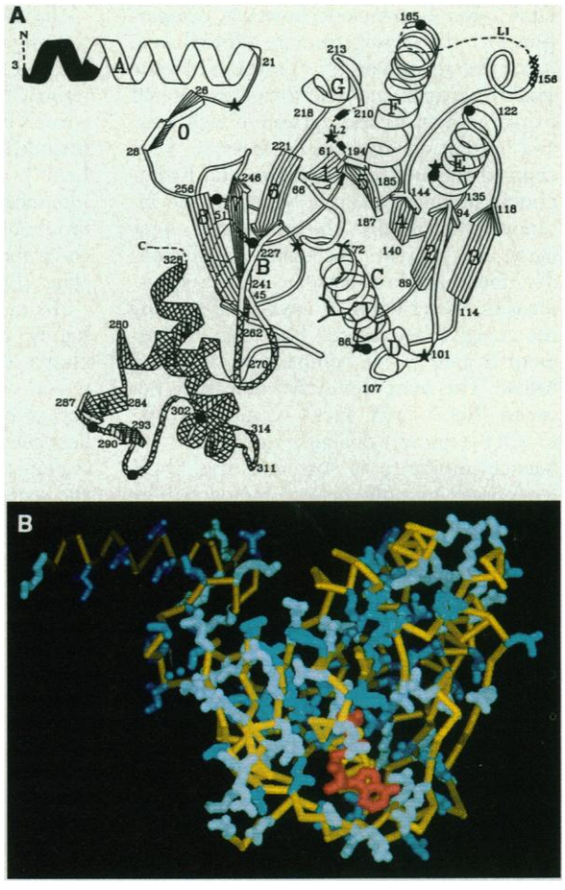

Fig. 2. (**A**) Positions of insertions and deletions in the DMC1 and UvsX sequences (Fig. 1) relative to that of RecA protein. Shown is a ribbon drawing of RecA protein modified from that of reference 17. Regions deleted in UvsX relative to RecA are displayed as shaded, those in DMC1 are cross hatched. Stars indicate the positions of insertions in DMC1 and circles indicate insertions in UvsX. (**B**) Position of large hydrophobic (Val, Ile, Leu, Met, and Phe, in blue) and charged residues (Asp, Glu, Arg, and Lys, in white) of DMC1 built onto the RecA structure with the alignment of Fig. 1. The α-carbon backbone of the model is yellow, and adenosine 5′-diphosphate (ADP) is red. Side chains were built in so that common atoms of RecA and DMC1 were positioned similarly when possible. It should be emphasized, however, that this is a test of the plausibility of the proposed alignment and not intended to be an accurate model of the entire DMC1 structure. It is likely that there will be differences in both main chain and side chain positions between the two structures, particularly in loop regions and far from the active site.

Table 1. Invariant and highly conserved active site residues.

Residue	Reason conserved; function
Gly^{66}	Begins ATP phosphate binding loop; unusual Phi, Psi angles
Lys^{72}	Interacts with β, γ-phosphates of ADP-ATP
Thr^{73}*	Interacts with Mg^{2+}, β-phosphate of ADP-ATP
Asp^{94}	Stabilizes loop containing Glu^{96} by H-bonds to backbone NH of residues 96 and 97
Glu^{96}	Catalytic base for activation of H_2O†
Tyr^{103}	Interacts with adenosine of ADP-ATP
Asp^{144}‡	Interacts with Mg^{2+} (by way of H_2O)
Ser^{145}‡	H-bonds to Glu^{96} carboxylate
Asn^{193}	Involved in ATP-induced conformational change†
Gln^{194}*	Triggering of conformational change by binding γ-phosphate of ATP†
Gly^{211}	Interaction with DNA or conformational change or both†
Gly^{211}	Interaction with DNA or conformational change or both†

*Conservatively substituted in UvsX: Thr^{73} by Ser, Gln^{194} by His. †Proposed function. ‡An unusual *cis* peptide bond is located between residues 144 and 145.

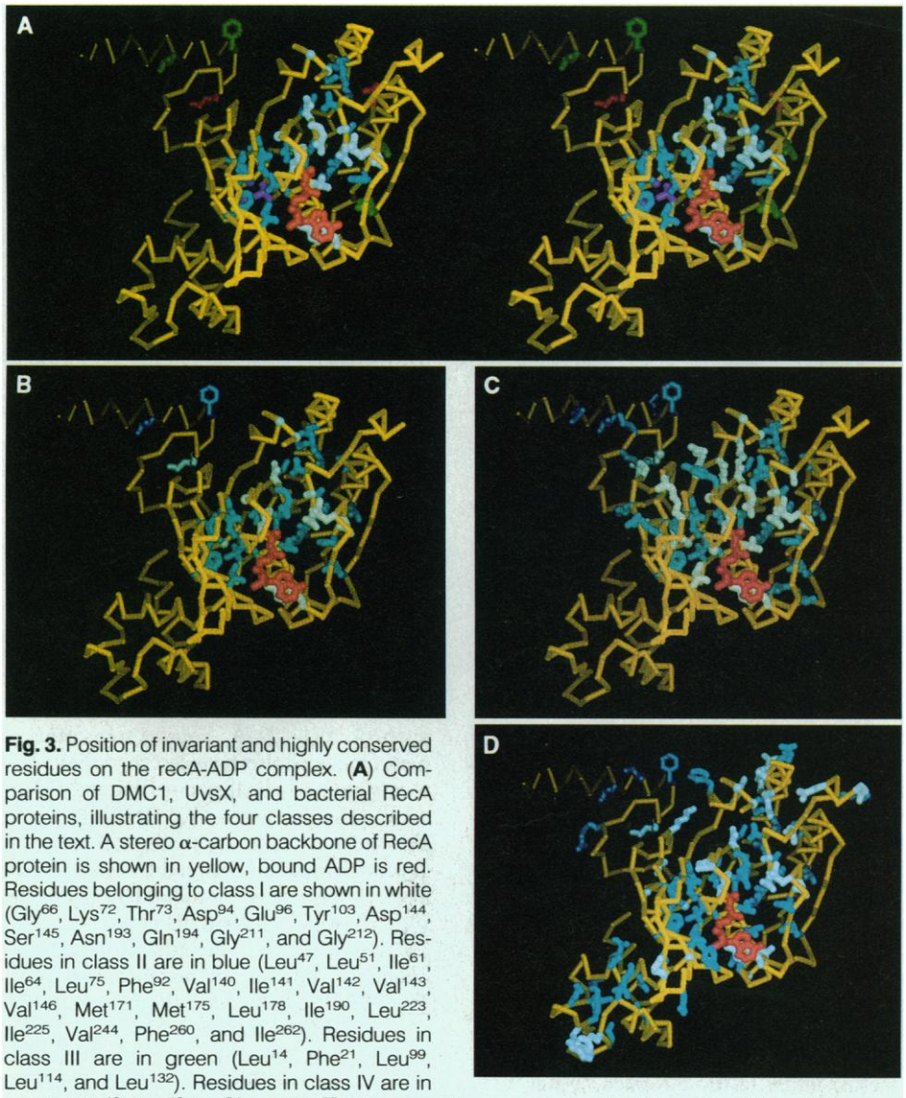

Fig. 3. Position of invariant and highly conserved residues on the recA-ADP complex. (**A**) Comparison of DMC1, UvsX, and bacterial RecA proteins, illustrating the four classes described in the text. A stereo α-carbon backbone of RecA protein is shown in yellow, bound ADP is red. Residues belonging to class I are shown in white (Gly^{66}, Lys^{72}, Thr^{73}, Asp^{94}, Glu^{96}, Tyr^{103}, Asp^{144}, Ser^{145}, Asn^{193}, Gln^{194}, Gly^{211}, and Gly^{212}). Residues in class II are in blue (Leu^{47}, Leu^{51}, Ile^{61}, Ile^{64}, Leu^{75}, Phe^{92}, Val^{140}, Ile^{141}, Val^{142}, Val^{143}, Val^{146}, Met^{171}, Met^{175}, Leu^{178}, Ile^{190}, Leu^{223}, Ile^{225}, Val^{244}, Phe^{260}, and Ile^{262}). Residues in class III are in green (Leu^{14}, Phe^{21}, Leu^{99}, Leu^{114}, and Leu^{132}). Residues in class IV are in purple (Thr^{42}, Asp^{48}, Gly^{54}, and Gly^{55}). The remaining two residues (Lys^{23} and Asp^{130}) are in red and lie on the surface of the protein toward the outside of the filament. (**B**) Comparison of DMC1, UvsX, and bacterial RecA proteins, with residues highly conserved shown in blue and invariant residues in white. (**C**) Comparison of yeast DMC1 and bacterial RecA proteins, with residues highly conserved shown in blue and invariant residues in white. (**D**) Comparison of T4 UvsX and bacterial RecA proteins, with residues highly conserved shown in blue and invariant residues in white.

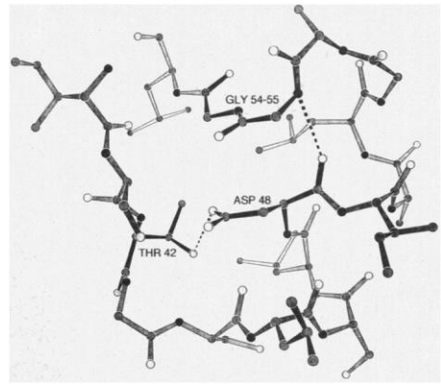

Fig. 4. Position of conserved class IV residues. Residues 40 to 56 of RecA are depicted, with conserved residues Thr^{42}, Asp^{48}, Gly^{54}, and Gly^{55} shown in bold. Hydrogen bonds between these conserved residues are shown as dotted lines.

on how a RecA family member functions in vivo. UvsX closely resembles RecA in its ability to bind and form filaments on double- and single-stranded DNA, to pair a single strand of DNA with an homologous duplex region, and in the modulation of its activities by ATP binding and hydrolysis. However, it differs from RecA in that UvsX protein alone, under typical conditions, can only promote limited invasion of a single strand into an homologous duplex (*6, 7*); additional proteins are required for full strand exchange in vitro (*28, 29*). This difference is biologically appropriate, because the major function of UvsX-promoted strand invasion is to provide a primer for DNA replication rather than to effect extensive strand displacement (*30*). The nucleoprotein filaments formed by UvsX are also less stable than those formed by RecA, and entry and exit of subunits from the filament is proposed to play a more important role in the UvsX reaction than in that promoted by RecA (*7*). In addition to minor alterations in the fundamental activities common to all RecA-like proteins, different members of the family may differ in the extent to which additional proteins are involved in the basic reaction, as exemplified by the intimate relationship between UvsX and a second protein, UvsY (*31*). Moreover, different proteins may have additional unique activities not common to the family as a whole, such as the ability of bacterial RecAs to promote cleavage of certain repressor proteins and UmuD (*32, 33*).

Because the structural relationship between DMC1 and RecA is as close as that between UvsX and RecA, the analogous strong functional similarities and subtle differences seem likely. The same argument applies to RAD51, which is related to RecA by the same structural criteria as are UvsX and DMC1 (*22*), as well as to a third yeast protein, RAD57. Each of these proteins must differ in the nature, location, or timing of its activities, because single mutations in each of the three genes confer unique phenotypes (*9, 10*). Presumably the existence of these several related genes reflects the type of functional specialization that is well known for other types of proteins such as topoisomerases, helicases, and polymerases.

REFERENCES AND NOTES

1. S. C. West, *Annu. Rev. Biochem.* **61**, 603 (1992).
2. A. I. Roca and M. M. Cox, *Crit. Rev. Biochem. Mol. Biol.* **25**, 415 (1990).
3. S. C. Kowalczykowski, *Ann. Rev. Biophys. Biophys. Chem.* **20**, 539 (1991).
4. C. M. Radding, *J. Biol. Chem.* **266**, 5355 (1991).
5. H. Fujisawa, T. Yonesaki, T. Minagawa, *Nucleic Acids Res.* **13**, 7473 (1985).
6. T. Yonesaki and T. Minagawa, *EMBO J.* **4**, 3321 (1985).
7. T. Kodadek, M. L. Wong, B. Alberts, *J. Biol. Chem.* **263**, 9427 (1988).
8. J. Griffith and T. Formosa, *ibid.* **260**, 4484 (1985).
9. D. K. Bishop, D. Park, L. Xu, N. Kleckner, *Cell* **69**, 439 (1992).
10. M. A. Resnick, in *Meiosis*, P. B. Moens, Ed. (Academic, New York, 1987), p. 157.
11. A. Shinohara, H. Ogawa, T. Ogawa, *Cell* **69**, 457 (1992).
12. G. Basile, M. Aker, R. K. Mortimer, *Mol. Cell. Biol.* **12**, 3235 (1992).
13. A. Shinohara, H. Ogawa, Y. Matsuda, N. Ushiro, K. Ikeo, T. Ogawa, personal communication.
14. O. Bezzubova, A. Shinohara, R. G. Mueller, H. Ogawa, J. M. Buerstedde, *Nucleic Acids Res.*, in press.
15. R. Cheng, T. I. Baker, C. E. Cords, R. J. Radloff, *Mutat. Res.*, in press.
16. T. Kobayashi, Y. Hotta, S. Tabata, *Mol. Gen. Genet.*, in press.
17. R. M. Story, I. T. Weber, T. A. Steitz, *Nature* **355**, 567 (1992).

18. R. M. Story and T. A. Steitz, *J. Mol. Biol.* **355**, 374 (1992).
19. R. C. Benedict and S. C. Kowalczykowski, *J. Biol. Chem.* **263**, 5450 (1991).
20. S. Tateishi, T. Horii, T. Ogawa, H. Ogawa, *J. Mol. Biol.* **223**, 115 (1992).
21. B. Lee and F. M. Richards, *ibid.* **55**, 379 (1971).
22. R. M. Story and D. K. Bishop, data not shown.
23. D. Bashford, C. Chothia and A. M. Lesk, *J. Mol. Biol.* **196**, 199 (1987).
24. L. T. J. Delbaere, W. L. B. Hutcheon, M. N. G. James, W. E. Thiessen, *Nature* **257**, 758 (1975).
25. L. A. Kohlstaedt, J. Wang, J. M. Friedman, P. A. Rice, T. A. Steitz, *Science* **256**, 1783 (1992).
26. M. Delarue, O. Poch, N. Tordo, D. Moras, P. Argos, *Protein Eng.* **3**, 461 (1990).
27. T. Ogawa, X. Yu, A. Shinohara, E. H. Egelman, *Science* **259**, 1896 (1993).
28. T. Kodadek and B. M. Alberts, *Nature* **326**, 312 (1987).
29. L. D. Harris and J. D. Griffith, *J. Mol. Biol.* **206**, 19 (1989).
30. T. Formosa and B. M. Alberts, *Cell* **47**, 793 (1986).
31. T. Kodadek, D. C. Gan, K. Stemke-Hale, *J. Biol. Chem.* **264**, 16451 (1989).
32. J. W. Little, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1375 (1984).
33. S. E. Burckhardt, R. Woodgate, R. H. Sheuermann, H. Echols, *ibid.* **85**, 1811 (1988).
34. J. E. Walker, M. Saraste, M. J. Runswick, N. J. Gay, *EMBO J.* **1**, 945 (1982).
35. A. I. Roca and M. M. Cox, personal communication.
36. A. McLachlan, *J. Mol. Biol.* **61**, 409 (1971).
37. We thank A. I. Roca and M. M. Cox for kindly providing an unpublished comparison of bacterial RecA sequences. This work was supported in part by NIH grant GM22778 to T.A.S.

8 January 1993; accepted 5 March 1993

Similarity of the Yeast RAD51 Filament to the Bacterial RecA Filament

Tomoko Ogawa, Xiong Yu, Akira Shinohara, Edward H. Egelman*

The RAD51 protein functions in the processes of DNA repair and in mitotic and meiotic genetic recombination in the yeast *Saccharomyces cerevisiae*. The protein has adenosine triphosphate–dependent DNA binding activities similar to those of the *Escherichia coli* RecA protein, and the two proteins have 30 percent sequence homology. RAD51 polymerized on double-stranded DNA to form a helical filament nearly identical in low-resolution, three-dimensional structure to that formed by RecA. Like RecA, RAD51 also appears to force DNA into a conformation of approximately a 5.1-angstrom rise per base pair and 18.6 base pairs per turn. As in other protein families, its structural conservation appears to be stronger than its sequence conservation. Both the structure of the protein polymer formed by RecA and the DNA conformation induced by RecA appear to be general properties of a class of recombination proteins found in prokaryotes as well as eukaryotes.

The RecA protein of *E. coli* is the most intensively studied enzyme involved in genetic recombination (*1–4*). In the presence of adenosine triphosphate (ATP) or an ATP analog, RecA can polymerize on single- or double-stranded DNA to form a helical nucleoprotein filament with about 6.2 RecA protomers per filament turn and a pitch of approximately 92 to 97 Å (*5*). Within this filament, both double-stranded DNA (dsDNA) and single-stranded DNA (ssDNA) are extensively stretched and dsDNA is unwound. The active form of RecA as determined by adenosine triphosphatase (ATPase) activity, strand-exchange activity, and the cleavage of the LexA and phage repressors exists only in this extended filament. We now show that the RAD51 protein from the yeast *S. cerevisiae*, which functions in DNA repair and recombination (6), induces a similar conformation in DNA and forms helical nucleoprotein filaments similar to those formed by RecA.

When cloned and purified from *E. coli*, RAD51 has ATP-dependent DNA binding activities similar to those of RecA (6). The RAD51 protein contains 400 amino acid residues. Within a 221-residue core region, 30% of the residues are identical and 24% are similar to those of the corresponding region of *E. coli* RecA (6). Thus, 119 out of 400 residues, or 30% of the total sequence of RAD51, are either identical or similar to those of RecA. We have purified RAD51 from *S. cerevisiae* to 95% purity (*7*). In the presence of ATP, this protein forms filaments on dsDNA that are similar to the filaments formed by RecA protein (Fig. 1). After covering circular dsDNA molecules with RAD51, we measured the extension of the DNA induced by RAD51. The DNA was greatly extended from a 3.4 Å rise per base pair found in B-DNA to an axial rise per base pair of 5.1 Å (Fig. 2). This extension of DNA is the same as that induced by RecA protein (*8*).

Fourier transforms of filament images were used to determine the pitch of the RAD51-ATP-DNA filament from both negatively stained and frozen-hydrated filaments (Fig. 3). The mean pitch of 99 Å is greater than that of comparable RecA-ATP-DNA filaments, which have a mean pitch of about 92 Å. Fourier transforms of such filaments also revealed that these filaments have a helical symmetry nearly identical to that of RecA-ATP-DNA filaments. This helical symmetry made it possible to generate three-dimensional reconstructions from single views of individual filaments (9). Negatively stained filaments were used for the reconstruction rather than frozen-hydrated filaments for several reasons. (i) There was no detectable change of pitch as a result of dehydration with the negative staining (Fig. 3). (ii) We found no evidence that the frozen-hydrated filaments were better preserved than the negatively stained. (iii) In the frozen-hydrated filaments, weak contrast and a poor signal-to-

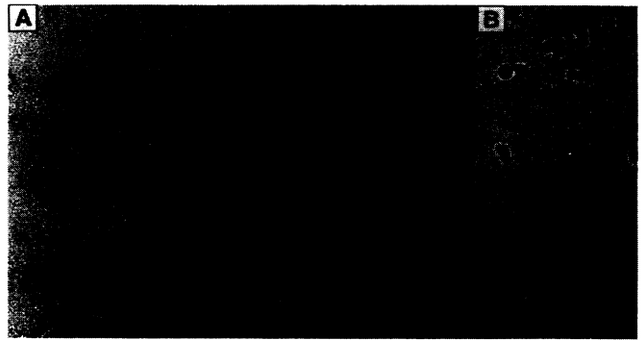

Fig. 1. Electron micrographs of RAD51 protein polymerized on DNA. (**A**) Negatively stained filaments of RAD51 on circular ΦX174 dsDNA observed by conventional transmission electron microscopy. (**B**) Unstained, frozen-hydrated RAD51 filament on linear dsDNA observed by cryo-electron microscopy (*28*). Under the conditions used (*29*), RAD51 protein (unlike RecA protein) does not form filaments on ssDNA with ATP or on dsDNA with the slowly hydrolyzable ATP analog ATP-γ-S (*30*). This may explain the inability of RAD51 to catalyze ssDNA-dsDNA strand-exchange reactions because although RAD51 has a ssDNA-dependent ATPase activity (*7*), the RecA-catalyzed ssDNA-dsDNA exchange reaction proceeds by way of the formation of the active filament on ssDNA, and RAD51 fails to form filaments on ssDNA. The tobacco mosaic viruses seen in (A) are about 200 Å in diameter, and both micrographs are shown at the same magnification.

T. Ogawa and A. Shinohara, Department of Biology, Osaka University, Osaka 560, Japan.
X. Yu and E. H. Egelman, Department of Cell Biology and Neuroanatomy, University of Minnesota Medical School, Minneapolis, MN 55455.

*To whom correspondence should be addressed.