A Molecular Model for RecA-Promoted Strand Exchange via Parallel Triple-Stranded Helices

Guillaume Bertucat, Richard Lavery, and Chantal Prévost
Laboratoire de Biochimie Théorique, CNRS UPR 9080, Institut de Biologie Physico-Chimique, 75005 Paris, France

ABSTRACT A number of studies have concluded that strand exchange between a RecA-complexed DNA single strand and a homologous DNA duplex occurs via a single-strand invasion of the minor groove of the duplex. Using molecular modeling, we have previously demonstrated the possibility of forming a parallel triple helix in which the single strand interacts with the intact duplex in the minor groove, via novel base interactions (Bertucat et al., *J. Biomol. Struct. Dynam.* 16:535–546). This triplex is stabilized by the stretching and unwinding imposed by RecA. In the present study, we show that the bases within this triplex are appropriately placed to undergo strand exchange. Strand exchange is found to be exothermic and to result in a triple helix in which the new single strand occupies the major groove. This structure, which can be equated to so-called R-form DNA, can be further stabilized by compression and rewinding. We are consequently able to propose a detailed, atomic-scale model of RecA-promoted strand exchange. This model, which is supported by a variety of experimental data, suggests that the role of RecA is principally to prepare the single strand for its future interactions, to guide a minor groove attack on duplex DNA, and to stabilize the resulting, stretched triplex, which intrinsically favors strand exchange. We also discuss how this mechanism can incorporate homologous recognition.

INTRODUCTION

RecA protein has been extensively studied over the past 20 years, notably for its ability to promote recognition and strand exchange between homologous single-stranded (ss) and double-stranded (ds) DNA (for reviews, see Camerini-Otero and Hsieh, 1993; Kowalczykowski and Eggleston, 1994; Kowalczykowski et al., 1994; Takahashi and Nordén, 1994; Kurumizaka and Shibata, 1996). Other functions of this protein involve ATP hydrolysis and SOS induction through LexA repressor cleavage (Roberts et al., 1978; Little et al., 1980). All of them participate in homologous genetic recombination and repair mechanisms within prokaryotes. Homologous recombination is now also considered to be an important process in higher organisms, where it plays a role in both the replication and the conservation of the genetic material (Baumann and West, 1998). Analogs of RecA, such as Rad51, are involved in this process and seem to use similar mechanisms, with the exception of reduced ATPase activity.

The active form of RecA is the so-called presynaptic complex, where the protein polymerizes around ssDNA (in the $5' \rightarrow 3'$ direction and using ATP as a cofactor) to form a right-handed helical nucleoprotein filament. Subsequently, during synapsis, this filament incorporates dsDNA, aligning homologous regions of ss and dsDNA, and strand exchange occurs. The displaced strand is ultimately released in the $5' \rightarrow 3'$ direction. Recognition is not fully specific, and strand exchange is hardly affected by a 3% heterology between the ss and ds sequences (DasGupta and Radding, 1982a; Bianchi and Radding, 1983). The reaction, however, is strongly dependent on both the length of homologous regions, with a minimum of 20 consecutive base pairs (Watt et al., 1985), and the degree of negative superhelicity (when circular dsDNA is involved; DasGupta and Radding, 1982b; Wong et al., 1998).

Considerable progress in understanding the mechanism of association and strand exchange has been achieved by the development of methods to investigate exchange kinetics, the nature of the reaction intermediates and factors directing homologous recognition. Notably, assays based on energy transfer between fluorescent dyes attached to DNA strands, combined with stop-flow spectrofluorometry, have allowed the direct observation of strand association and dissociation (Ellouze et al., 1997b; Bazemore et al., 1997a,b; Gumbs and Shaner, 1998). Independently of homology between the ss and dsDNA, association is believed to involve the formation of a coaxial, stretched and unwound ternary (triple stranded) intermediate (Rosselli and Stasiak, 1990; Kiianitsa and Stasiak, 1997; Wong et al., 1998). Indeed, RecA can accommodate up to three strands independently of their sequence (Kubista et al., 1990). Both Ellouze et al. (1997b) and, more recently, Malkov and Camerini-Otero (1998) have established that the dissociation of a complex between RecA/ssDNA and either ss or dsDNA occurs very quickly in the presence of heterologies, thus interrupting the strand exchange process. This important result, however, does not determine whether such heterologies are detected during strand exchange (the "duplex melting" mechanism; see Zhou and Adzuma, 1997) or within the initial ternary complex (the "triple helix" mechanism; see Camerini-Otero and Hsieh, 1993; Rao et al., 1995; Kurumizaka and Shibata, 1996), a question that was raised earlier by Howard-Flanders et al. (1984). Although neither of these mecha-

Received for publication 22 March 1999 and in final form 14 June 1999.
Address reprint requests to Dr. Chantal Prévost, Laboratoire de Biochimie Théorique, CNRS UPR 9080, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France. Tel.: 33-1-43-25-26-09; Fax: 33-1-43-29-56-45; E-mail: chantal.prevost@ibpc.fr.

0006-3495/99/09/1562/15 $2.00

nisms has been clearly established so far, results from fluorescence energy transfer experiments suggest two levels of recognition, one during dsDNA association with the filament and the other during strand exchange (Bazemore et al., 1997a). The observed destabilization of ternary intermediates caused by "base mutations" in the ssDNA has also been interpreted as involving two distinct levels of heterology detection (Malkov and Camerini-Otero, 1998).

Two recent studies have clarified the kinetics of the initial steps of strand exchange by indicating a fast initial association step (rate constant 0.2 s^{-1}) followed by a slower reorganization of the ternary complex (rate constant 0.02 s^{-1}) (Bazemore et al., 1997b; Gumbs and Shaner, 1998). Two or three ternary intermediates in strand exchange are implied by these studies. One of them may be the putative triple helix already isolated and characterized in the presence of RecA, which appears to be stabilized by non-Watson-Crick interactions between the ss and ds components and is remarkably stable upon deproteinization (Hsieh et al., 1990; Muller et al., 1992; Rao et al., 1991). This structure is sometimes considered to be a proof of the triple-helix mechanism, but the parallel triplex in question, if it is really involved in the reaction, would correspond to the product of the strand exchange reaction, involving a heteroduplex dsDNA and the outgoing single strand (Chiu et al., 1993; Rao et al., 1993). A true intermediate triplex stabilized by sequence alignment remains to be identified.

Given the rapidity of the recognition and strand exchange steps, direct experimental identification of such an intermediate is difficult. Modeling has thus proved to be a useful tool for constructing plausible, stretched, parallel triple helices, where the third strand interacts either in the major or minor grooves. The major groove hypothesis was favored in two early modeling studies (Hsieh et al., 1990; Zhurkin et al., 1994; see Kowalczykowski and Eggleston, 1994, for a review). The major groove of B-DNA is undoubtedly more accessible than the minor groove for single-strand association. Moreover, the resulting base triplets more easily explain the fidelity of homologous base association (Zhurkin et al., 1994). However, a number of experimental observations support single-strand invasion of the duplex via its minor groove (Kumar and Muniyappa, 1992; Baliga et al., 1995; Podyminogin et al., 1996; Tuite et al., 1997; Zhou and Adzuma, 1997). A model of the parallel triple helix with the third strand placed in the major groove was proposed by Chiu et al. (1993) as a product of strand exchange, as opposed to the association complex.

We have recently demonstrated that stretching of the 3′ extremities of a duplex DNA by a factor of 1.5 opens the minor groove sufficiently to allow the introduction of a single DNA strand (Bertucat et al., 1998). This study led to a model of parallel triple-helical DNA stabilized by a new type of minor groove interactions (the "minor groove triple helix"; Fig. 1 *a*). This triplex resembles RecA bound to ss or dsDNA in its global characteristics, by being extended by a factor of 1.5 (that is, an average rise of 5.1 Å; Stasiak et al., 1981), unwound to a twist of 20° (Stasiak and DiCapua,

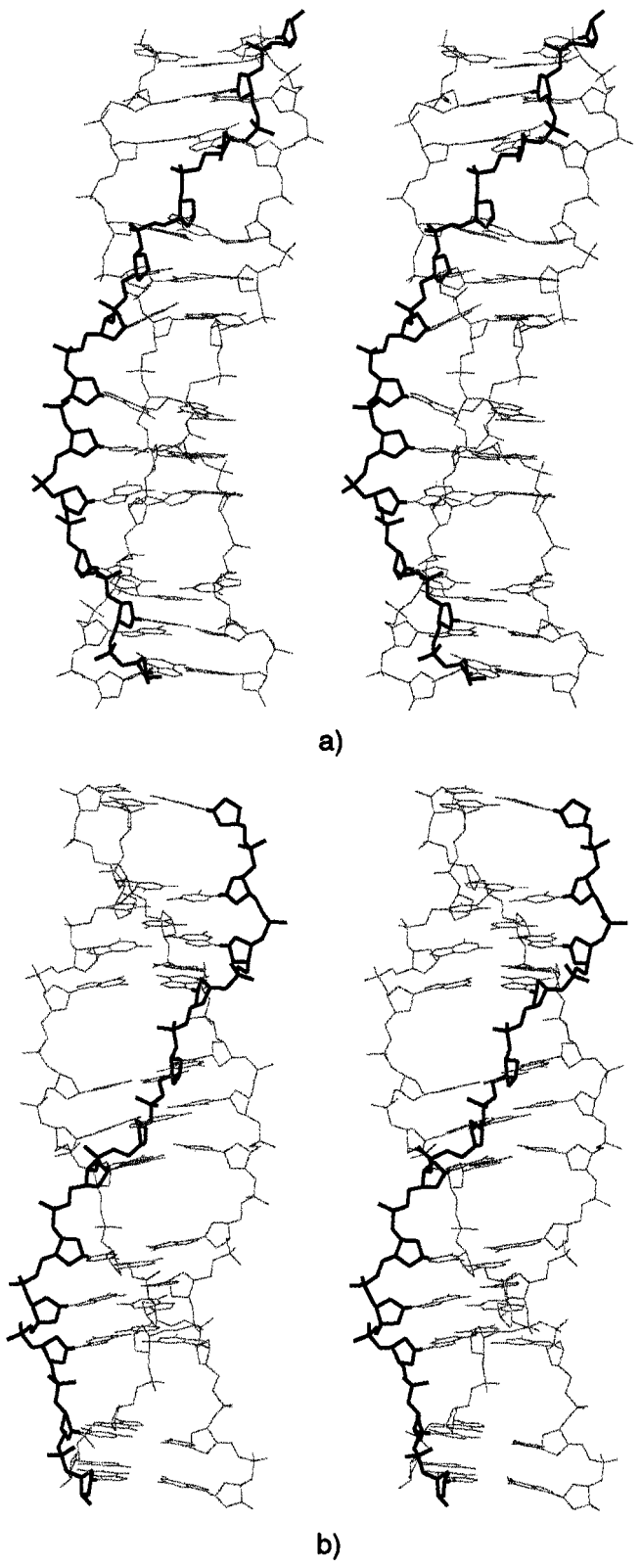

FIGURE 1 Stereoscopic view of the model minor groove (*a*) and major groove (*b*) triple helices. The single-strand DNA (backbone in *bold* and 5′-3′ *pointing upward*) interacts in an orientation parallel to the homologous strand of the duplex, via the minor and major grooves, respectively. The two triple helices are stretched by a factor of 1.5 with respect to canonical B-DNA.

1982), and having base triplets perpendicular to the helix axis (Nordén et al., 1992). Moreover, the structure of the ssDNA interacting in the minor groove is compatible with available NMR data (Nishinaka et al., 1997). Interestingly, our calculations show that the stretching and unwinding deformations are necessary for the stability of the triplex. Its formation, however, is enthalpically unfavorable without the intervention of RecA. This structure thus contrasts with the R-DNA model of Zhurkin et al. (1994), which can exist in a compact, protein-free form. In line with the experimental results, these observations make the minor-groove triple helix a good candidate for the ternary intermediate of association, which is known experimentally to be unstable in the absence of RecA (Reddy et al., 1995). R-DNA would then correspond to the putative triple helix mentioned above, isolated experimentally, supposedly formed after strand exchange and stable upon RecA removal.

Several questions need to be answered to clarify the role of these parallel triplexes in the process of recognition/strand exchange. First, it is important to know whether strand exchange is possible within such structures. We now address this question by modeling base pair switching within our minor-groove triple-helical model, to produce a new triplex, the "major groove triple helix," that we compare with the R-DNA model. The properties of the minor groove and major groove triple helices are tested against available structural information corresponding to the different stages of the strand exchange reaction.

A second question concerns homologous sequence recognition. We make some preliminary tests of this process by introducing heterologies in the invading single strand and analyzing their effects on the patterns of minor-groove interaction that we have established for the four possible homologous, minor-groove triplets (Bertucat et al., 1998). We relate the results to the recent experimental study of Malkov and Camerini-Otero (1998), who measured the sequence-dependent rate of dissociation of the homologously aligned complex.

Our studies employ the internal coordinate Jumna program and the associated Flex force field (Lavery et al., 1986, 1995; Lavery, 1988, 1995), already used for the construction and optimization of the stretched minor-groove triple-helix model (Bertucat et al., 1998). This program is particularly well adapted to our problem because it permits controlled structural deformation and extensive exploration of the conformational space. It has already been used to study the mechanism of extreme stretching deformations in relation to nanomanipulation experiments (Cluzel et al., 1996; Lebrun and Lavery, 1996), and these structures have turned out to be very helpful in deriving our triple-helical model.

METHODS

The Jumna program and the Flex force field have been extensively described in previous publications (Lavery et al., 1986, 1995; Lavery, 1988, 1995). Basically, Jumna uses a mixture of helical and internal coordinates (valence and dihedral angles) to describe nucleic acid flexibility. The helical parameters position each 3′-monophosphate nucleotide with respect to a fixed axis system. Junctions between successive nucleotides are maintained with quadratic restraints on the O5′-C5′ distances. In addition to a reduced number of variables with respect to Cartesian coordinate programs, the choice of physically meaningful variables allows large, concerted conformational moves during minimization, together with an efficient control of the structure and easy introduction of constraints or restraints. Available tools include both adiabatic mapping and combinatorial searches with respect to chosen structural parameters. Particularities of the Flex force field include the presence of a specific term to account for the angular dependence of hydrogen bonding and the possibility of electrostatic energy screening with a sigmoidal dielectric function (Hingerty et al., 1985; Lavery et al., 1995),

$$\epsilon(R) = D - (D - D_0)/2[(RS)^2 + 2RS + 2]\exp(-RS)$$

where R is the distance between two charges. The slope S, the plateau value at long distance D, and the initial value D_0 of the function are adjustable, with default values of 0.16, 80, and 1, respectively.

We have performed the present study using two assumptions already employed for constructing the minor-groove triple helix (Bertucat et al., 1998). First, the base triplets are restrained to be coplanar to avoid any possible interbase triplet interactions. Such interactions easily form during the construction of stretched helices but cannot play a role in recognition or strand exchange, since these processes are independent of the overall sequence. We have checked that the optimized structure of the minor-groove triplex is independent of these restraints. The second assumption, in line with the stoichiometry of RecA/DNA complexes, which show three nucleotides per RecA monomer, is the use of trinucleotide helical symmetry. For this reason we have also limited our preliminary studies to sequences with trinucleotide repeats.

Specific restraints or constraints are needed for triplex construction and manipulation. These include the "plateau" restraints and the trinucleotide symmetry constraints, described previously (Bertucat et al., 1998). The "plateau" restraint maintains the coplanarity of the bases forming a triplet, while allowing the rotations and displacements required for base pair switching. The trinucleotide symmetry constraint implies the equivalence of the variables describing each successive group of three nucleotides. Stretching dsDNA so that the twist decreases and the minor groove opens was previously achieved by restraining the distance between the terminal O3′ atoms of our trinucleotide symmetry unit. We have modified this restraint slightly because the O3′-O3′ distance can be altered by a lateral displacement of the backbones during strand exchange. In the present work, only the component of the O3′-O3′ vector

parallel to the helix axis is restrained. Finally, restraints on the groove width, calibrated with the help of numerical Poisson-Boltzmann electrostatic calculations, are used to avoid groove narrowing due to the lack of explicit solvent molecules (see Bertucat et al., 1998).

Base pair switching is studied by base rotation, using the approach defined by Bernet et al. (1997). This involves a restraint applied to the angle θ between the glycosidic bond (purine: C1′-N9 or pyrimidine: C1′-N1) and the vector joining the two C1′ atoms of a base pair, projected on the plane perpendicular to a local helical axis (Fig. 2). θ has a value of 55° in canonical B-DNA. Modeling base pair switching for a chosen base involves an adiabatic variation of θ from 65° to −10° by steps of 2° while maintaining both "plateau" and stretching restraints.

Protocols for optimization and adiabatic mapping

Optimization of the minor groove and major groove triple helices with trinucleotide repeat sequences is systematically performed under trinucleotide symmetry constraints. The protocol includes minimization followed by combinatorial searches on the nine independent sugar puckers, while maintaining "plateau" restraints. These restraints are released in a final minimization step. In some cases of minor groove triple helices, restraints on the hydrogen bond distances were added to avoid spontaneous switching of the base pairs during these highly perturbing combinatorial searches (see Results). Optimization of the major-groove triple helix is performed under stretching restraints.

Adiabatic mapping of the conformational space of the minor groove and major groove triplexes with respect to both rise and twist was carried out by modifying the projected O3′-O3′ distance in steps of 0.5 Å, from 40 Å (that is, an average rise of 3.4 Å) to 65 Å (an average rise of 5.2 Å) and the mean twist in steps of 1°, from 17° to 38°. The resulting structures were then minimized under trinucleotide symmetry constraints and "plateau" restraints. In the case of the minor-groove triplex, restraints on the hydrogen bond distances were added.

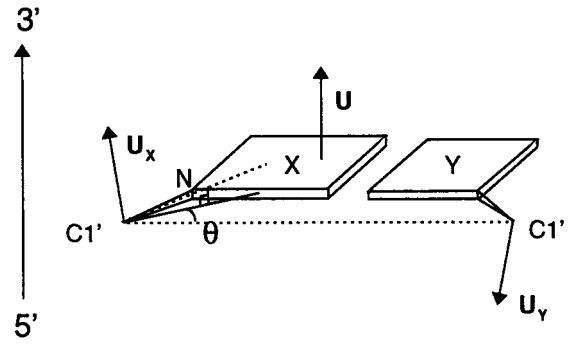

FIGURE 2 Definition of the base pair opening restraint, applied to a base X. θ is the angle between the glycosidic bond C1′-N in base X and the vector joining the C1′ atoms of complementary bases X and Y, projected onto the plane perpendicular to vector **U**. Vector **U** is defined as the mean of the vectors $\mathbf{U}_X$ and $-\mathbf{U}_Y$ joining the C1′ atoms preceding and following X and Y bases in each strand. This projection ensures that base inclination is not interpreted as opening. Base rotation is considered to be positive for right-hand rotation around a vector pointing in the 5′-3′ direction.

To assess the importance of RecA binding for the stability of the triplexes, their energies are calculated relative to the relaxed dsDNA and ssDNA components. The ssDNA is minimized in a helicoidal conformation that does not take into account the possible formation of secondary structures in solution. If such structures exist, they imply that the triplex would require even more stabilization by RecA.

In the studies involving local sequence modification or local base pair switching, our DNA fragment is divided into three sections. Two blocks comprising, respectively, the first three and last four nucleotide triplets have frozen conformations. A kink is then introduced at the center of the structure to allow relative movements of the two blocked segments. Finally, the central base triplet is used for sequence modification trials, and the two adjacent triplets are unconstrained. Each locally modified sequence was structurally optimized by minimization and combinatorial sugar pucker searches under "plateau" restraints.

Relative stability of heterologous minor groove triplexes

"Base mutation" in the single strand of a minor-groove triple helix leads to a new (heterologous) triplex whose energy cannot be directly compared with that of the initial structure. We thus calculate the effect of introducing a single mismatch on the enthalpy of formation of the triplexes by subtracting the energy of the ssDNA extended by a factor of 1.5, unwound to 20°, and having the bases perpendicular to the helical axis (as in the complex with RecA) from that of the optimized triplex. The relative value of this energy difference (ΔE_{diff}) is an indication of the relative stabilities of the two triplexes. Note that the structure of the isolated dsDNA need not be considered, because the duplex part of heterologous and homologous triplexes remains identical.

RESULTS

Sequence effects within the minor-groove parallel triple helices

The triple helix we have studied consists of a single-stranded DNA interacting in the minor groove of a homologous duplex DNA. In what follows, we will term strand 1 and strand 2 the 5′-3′ and 3′-5′ strands of the dsDNA duplex and strand 3 the 5′-3′ ssDNA with a sequence identical to that of strand 1.

In our previous study, we noted that the three base triplets of each symmetry unit were not equivalent. This is due to an irregular distribution of the stretching and unwinding deformation, characterized by a single gap in stacking (Fig. 1 *a*). Before studying the effects of sequence and of mismatches on base pair switching, we need to determine whether the

base sequence affects the minor-groove triple helix conformation. To this end, structures of sequence $d(GCG)_4$, containing purine/purine steps, and $d(CGC)_4$, containing pyrimidine/pyrimidine steps, were used. For $d(GCG)_4$, the positioning of the gap at the GpG step was tested by sliding the sequence with respect to the symmetry constraints: $(GCG)_4 \rightarrow (GGC)_4 \rightarrow (CGG)_4$. Using a common starting structure, we thus placed the gap, respectively, at GpG, CpG, and GpC steps. These calculations confirmed a preference for the GpG site by almost 10 kcal·mol^{-1}. A similar study with the $d(CGC)_4$ sequence showed no such preference, and although a gap always existed, it could be positioned indifferently at CpG, GpC, or CpC steps.

Optimizing other dodecamer sequences, $d(ATA)_4$, $d(GGG)_4$, $d(CCC)_4$, $d(AAA)_4$, $d(TTT)_4$, confirmed the modes of minor-groove interactions presented previously (Bertucat et al., 1998) but also revealed a new pattern of minor-groove interaction for T.AxT, where the ssDNA thymine only binds to adenine, via two weak interactions involving O4(T)-HC2(A) and HN3(T)-N3(A). All of the triplet patterns are reproduced in Fig. 3 (*top*). The triplets have average inclinations below 10° and are associated with mixed sugar puckers (C3′-endo and C2′-endo) and common backbone modifications involving *trans*, *trans* α, γ configurations.

Strand exchange and the major-groove parallel triple helix

Rotation of a strand 2 base toward the minor groove of the duplex DNA leads to an exchange of Watson-Crick pairing from the 2–1 strands to the 2–3 strands. This was carried out in the $d(GCG)_4$ triplex, in the plane of the base triplets and under trinucleotide symmetry constraints. Strand 1 thus locally becomes a single strand interacting with the major groove of the heteroduplex formed by strand 3 and strand 2. Complete strand exchange was achieved by successively switching the base pairing of the three triplets of a symmetry unit, while optimizing the intermediate structures. Note that because of symmetry constraints, each base pair switch actually implies the rotation of four symmetrically equivalent bases within the dodecamer triplex. It is also remarked that both the initial duplex and the resulting heteroduplex were held under a 1.5× stretching restraint.

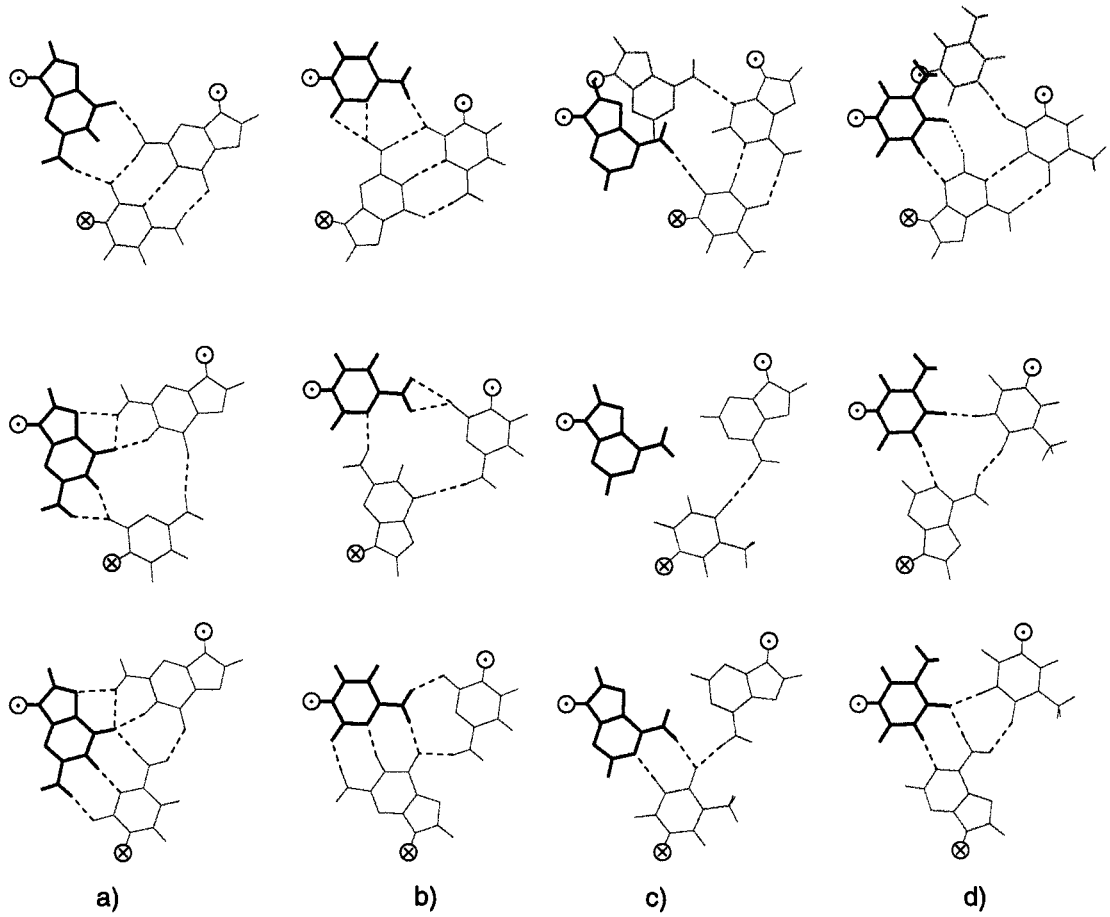

FIGURE 3 Patterns of interactions within the four triplets (*a*) G.CxG, (*b*) C.GxC, (*c*) A.TxA, and (*d*) T.AxT. Hydrogen bond interactions are represented by broken lines (– – –). (*Top*) Patterns of interactions in the minor groove triple helix. For the AT triplets, alternative patterns are represented with thin lines. (*Middle*) Patterns of interaction at an intermediate stage of base pair switching. (*Bottom*) Patterns of interaction in the major groove triple helix.

The helical parameters of the new triple helix are given in Table 1 and compared to those of the minor-groove starting triplex. The major-groove triple helix (Fig. 1 *b*) conserves many features characterizing the starting structure (Fig. 1 *a*). The bases remain perpendicular to the helix axis. The stretching and unwinding deformations generally conserve a gap every three stacked base triplets and local variations of the twist parameter, although these are slightly less pronounced than in the starting triplex. Once again the extension of the backbone is characterized by many *trans/trans* configurations for the (α, γ) backbone dihedrals and by a number of C3′-endo sugars. Following combinatorial sugar pucker searches, the latter represent more than half of all of the sugars, slightly more than in the minor-groove triplexes. It was also found that the $d(GCG)_4$ major-groove triple helix exhibits a preferred gap, although its position is now at the CpG step and not at GpG, as in the minor groove triplex. This position stabilizes the structure by 13 kcal·mol^{-1} per trinucleotide symmetry unit with respect to gaps at GpG or GpC steps and by 7 kcal·mol^{-1} with respect to a more regular structure having small gaps at every step. In the latter structure, the triplets are separated by 4.4–5.5 Å, whereas the twist varies over the range 21°-25°. Remarkably, the loss of stacking interactions between successive base triplets (12 kcal·mol^{-1} per symmetry unit) in this conformation is almost compensated by stabilization of the backbones, which now contain only C3′-endo sugars. This agrees with the results of Zhurkin et al. (1994), who predicted increased base separation for structures presenting a uniform C3′-endo sugar, and our structure almost exactly corresponds to the stretched R-form described by Zhurkin et al. (1994) and modeled under mononucleotide symmetry constraints. This structure loses its regularity upon release of the "plateau" restraints, showing important buckle deformations within the triplets, which help to regain the lost van der Waals stacking interactions.

Global modifications of the triple helix are also observed upon strand exchange. Although the total rise has been restrained to a constant value, the average twist of the model helices increases slightly, from 20° to 23°. This means that the number of RecA subunits per turn in the filament falls from 6.0 to 5.2 and the pitch of the triplex, 95 Å in the case of the minor-groove conformation, now drops to 80 Å. The triplex grooves are naturally redistributed, with widths of 11 Å between strands 1 and 2 and 11.5 Å between 2 and 3, compared to values of 18.2 Å and 7.2 Å in the minor-groove conformation. Overall, the major-groove triplex has better separated backbones and shows a 20% average increase in diameter measured at the C1′ atoms (Table 2). This change is anisotropic, however, because the C1′-C1′ distances between strands 1 and 2 increase much more than the other two C1′-C1′ distances (see Table 2). The evolution of these distances during a single base exchange for a G.CxG triplet is given in Fig. 4 *b*.

We have checked the strand 1 conformation of our major groove triplex against the NMR data of Nishinaka et al. (1997), which pertains to a four-nucleotide single strand of DNA interacting with RecA/ATPγS. The H2′-H68, H2″-H68, and H3′-H68 distances, which correspond to the characteristic NOE signals, are given in Table 3 and are compared with the same distances in strand 3 of the minor-groove triplex. Large internucleotide distances are found at the stacking gap within the major groove triplex (step C2pG3), exceeding those of the corresponding minor-groove triplex gap (step G3pG4). This leads to a large variance in the internucleotide distances, making direct comparison with the NMR data difficult. It nevertheless appears that the major-groove triplex is less compatible with the experimental results than the minor-groove conformation.

Patterns of interaction in the major groove triple helix

The patterns of interaction presented in Fig. 3 (*bottom*) have been obtained by optimizing modified sequences derived from the major-groove model triplex $d(GCG)_4$, under trinucleotide symmetry constraints, or as a result of base pair switching within individual triplets. Each of the four possible triplets successfully underwent the exchange. For the triplets A.TxA and T.AxT, base pair switching was possible only when the base of strand 3 initially interacted with that of strand 2 (see bold lines in Fig. 3, *top*). Note that in these two cases, the rotation of the base of strand 3 had to be restrained to keep it from accompanying the paired strand 2 and blocking the exchange. Interestingly, the patterns resulting from strand exchange (Fig. 3, *bottom*) are almost all identical to those proposed by Zhurkin et al. (1994). They

TABLE 1 Comparison between the helical parameters in the optimized major-groove triple helix and minor-groove triple helix of sequence $d(GCG)_4$

		Xdisp (Å)		Ydisp (Å)		Rise (Å)		Inc (°)		Tip (°)		Twist (°)	
G1xC.G	*G1.CxG*	**−4.1**	*1.6*	**−0.5**	*−1.2*	**3.9**	*3.8*	**−3.3**	*7.1*	**4.1**	*−1.0*	**29.3**	*24.2*
C2xG.C	*C2.GxC*	**−4.3**	*0.5*	**−0.9**	*−2.4*	**7.2**	*4.1*	**−2.7**	*1.0*	**−2.8**	*0.9*	**13.3**	*28.7*
G3xC.G	*G3.CxG*	**−4.4**	*1.9*	**0.4**	*−0.8*	**4.0**	*7.4*	**0.3**	*9.9*	**7.1**	*7.8*	**25.8**	*5.7*

Bold, major-groove triple helix; italics, minor-groove triple helix (from Bertucat et al., 1998). The parameters have been calculated using the program Curves (Lavery and Sklenar, 1988, 1989). For the major-groove triplex, the calculations have been done using a strand order 3-2-1 to allow comparison with the starting triplex. Note that Xdisp and Ydisp refer to the duplex, and the other parameters concern the whole triplex. The rise and twist displayed at each line are between the designated triplet and the following one. The gap is situated at step C2pG3 in the optimized major-groove triplex and at step G3pG1 in the minor-groove triplex.

TABLE 2 Comparison between the intratriplet C1′-C1′ distances in the minor-groove and the major-groove triple helices, for the four possible base triplets

	Minor-groove triplex (Å)				Major-groove triplex (Å)			
	C1′-C1′ distances				C1′-C1′ distances			
Triplet	1–2	2–3	1–3	Diameter	1–2	2–3	1–3	Diameter
G.CxG/GxC.G	10.9	9.3	11.9	*12.7*	14.2	10.9	12.0	*14.7*
C.GxC/CxG.C	10.8	8.3	9.1	*10.1*	13.0	10.8	10.4	*13.5*
A.TxA/AxT.A	11.1 (10.7)	9.2 (9.1)	12.5 (9.2)	*13.0*	14.4	10.9	12.3	*14.9*
T.AxT/TxA.T	10.7 (11.)	7.7 (8.2)	10.3 (8.4)	*11.3*	13.7	10.7	10.8	*13.8*

The distances have been measured in the optimized structures. The distances given for triplets A.TxA and T.AxT in the minor-groove triplex correspond to the patterns that successfully underwent base pairing exchange in our calculations. The distance relative to the alternate patterns (*thin lines* in Fig. 3, *c–d, top*) are given in parentheses.

are conserved during combinatorial searches of sugar puckers and are independent of the surrounding base sequence. As seen in Fig. 3, one or two hydrogen bonds can be conserved during base switching within the triplets. During the intermediate stage (Fig. 3, *middle*), each of the three bases interacts in a triangular pattern with its two neighbors, through non-Watson-Crick interactions. Fig. 3 also indicates that the relative movement of the identical bases in strands 1 and 3 mainly involves a shift, with rotation of less than 15° in all cases except A.TxA (around 35°).

Relative stability of the two triple helices

Strand exchange within $d(GCG)_4$ triplex under constant stretching leads to a stabilization of 34 kcal·mol^{-1} per trinucleotide unit. Comparison of the energy components of the minor groove and major groove helices shows that an important part of this stabilization (~55%) concerns base interactions at the level of each triplet. Triplets G.CxG, C.GxC, A.TxA, and T.AxT respectively gain 5.6, 2.1, 1.4, and 5.4 kcal·mol^{-1} upon switching. The interactions are also improved both within and between the dsDNA and ssDNA components of the triplex.

In contrast, base switching within a single triplet of the minor-groove triple helix does not necessarily stabilize the structure. Although the base interactions within the switched triplet are always improved, this gain is partially compensated for by intrastrand destabilization, which depends on the sequence and the position of the stacking gap. The energy variation during the switch from G.CxG to GxC.G is represented in Fig. 4 *a*. Activation requires 12 kcal·mol^{-1} in this case, leading to the exchanged triplet via a stable intermediate form. The activation energy may be due to steric hindrance or electrostatic repulsion between functional groups. It can probably be lowered by optimizing the pathway for switching. For T.AxT triplets a barrier of 5 kcal·mol^{-1} was obtained. Further investigation of the pathway is in progress.

Under trinucleotide symmetry constraints, the passage of the first triplet of each trinucleotide symmetry unit of $d(GCG)_4$, followed by optimization (see Fig. 4 *c*), leads to a form stabilized by 3 kcal·mol^{-1}. The main stabilization of the major groove triplex (24 kcal·mol^{-1}) occurs upon switching the second triplet of each unit, followed by a final gain of 7 kcal·mol^{-1} after switching the last triplet. These data suggest cooperative effects, which nevertheless need to be confirmed in the absence of symmetry constraints.

Although strand exchange does not dramatically alter the characteristics of the RecA bound triplexes, it does strikingly change their behavior with respect to stretching and winding deformation. We have previously seen that the minor-groove triplex is stable in its extended and underwound state (Bertucat et al., 1998). We confirm this observation here by monitoring the energy variations for the axially projected O3′-O3′ distance and for the twist (Fig. 5 *a*). In contrast, the same deformations acting on the optimized major-groove triple helix (Fig. 5 *b*) indicate clear stabilization upon compression and rewinding. This behavior agrees perfectly with the experimentally observed stability of the triple helices after deproteinization. The evolution of the sugar puckers of the major-groove triplex, presented in Fig. 5 *c* (the same is true for the minor-groove triplex, data not shown), agrees with the results of a Fourier transform infrared study on RecA-complexed triplexes, probably following strand exchange (Dagneaux et al., 1995a). This study detected sugars in a north conformation (C3′-endo) in the extended RecA-complexed form, but not in the relaxed conformation.

Note that Fig. 5 *a* also shows a secondary energy well for the minor-groove helix, which was not found in our previous one-dimensional energy mapping. This conformation has a minimum at 31° for the average twist and 51 Å for the O3′-O3′ distance (corresponding to an average rise of 4.3 Å). This structure is stable upon removal of the "plateau" restraints, but is clearly incompatible with the 1.5× stretching imposed by complexation with RecA (corresponding to an average rise of 5.1 Å). It can be noted from Fig. 5, a and b, that, in contrast to the major-groove triplex, the minor-groove triplex is enthalpically unstable with respect to its ssDNA and dsDNA components (independent of its conformation) and thus requires RecA for its formation.

Effects of heterologies

Confronted with an enormous number of possible base sequence heterologies, we chose to concentrate here on

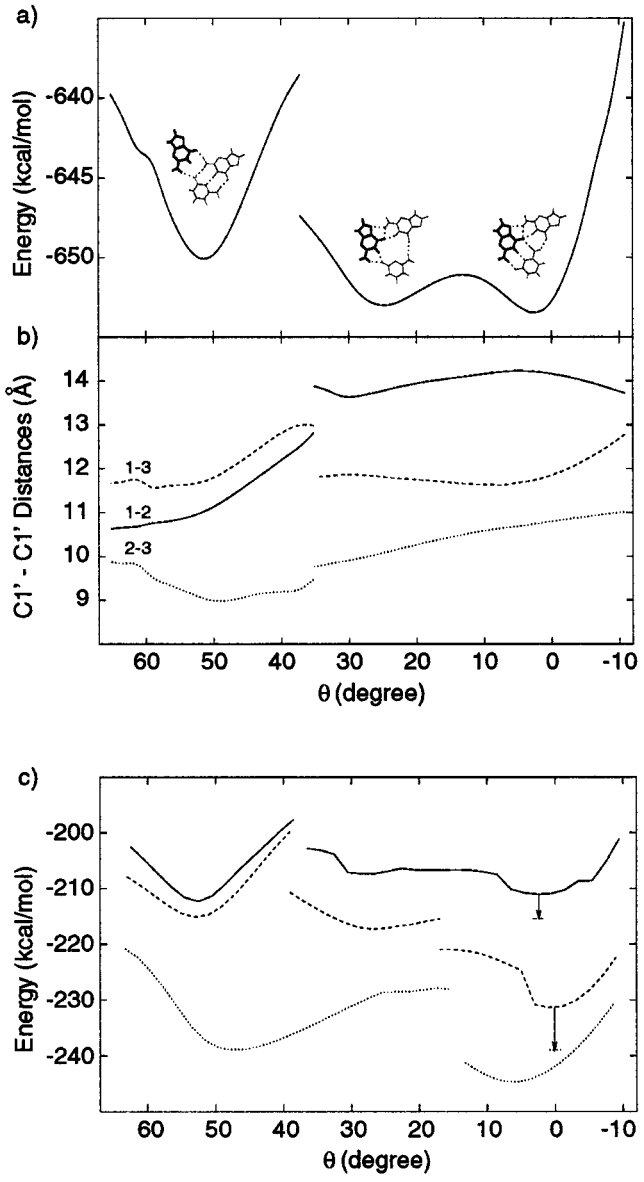

FIGURE 4 (*a*) Evolution of the energy of the $d(CGC)_3$ oligomer during base pair switching of the central G.CxG triplet. The state of intertriplet interactions, as described in Fig. 3 *a*, is represented for each energy well. (*b*) Evolution of the three C1′-C1′ distances of the central G.CxG base triplet during base pair switching within $d(CGC)_4$. The coordinate θ is the angle of rotation of cytosine in strand 3, as defined in Fig. 2. The different curves correspond to C1′ atoms in strands 1 and 2 (——), 1 and 3 (– – –), and 2 and 3 (·····). (*c*) Evolution of the energy of a trinucleotide unit of $d(GCG)_4$ during strand exchange under trinucleotide symmetry constraints. Each of the three triplets of a symmetry unit is switched in turn, followed by optimization. Arrows represent the gain in energy due to optimization. The three curves respectively represent the switching of G_3.CxG (——), C_4.GxC (– – –), and G_5.CxG (·····) and symmetry-related triplets. In *a*–*c*, the breaks in the curves correspond to abrupt conformational changes in the backbones.

X.YxM mismatches in minor-groove triplexes with the sequence d(GCGGCXGCGGCG), where X is G or C, Y is the complementary base, and M is a heterologous base. The underlined sequence corresponds to frozen segments (see Methods section), and a stacking gap is present between the modified triplet and the following G.CxG triplet. When present on the 3′ side, these mismatches destabilize the ternary intermediates of strand exchange in the study of Malkov and Camerini-Otero (1998), in the order A (G.CxA or C.GxA) < G/C (G.CxC or C.GxG) < T (G.CxT or C.GxT). Mismatches at A.T or T.A are more favorable, with the exception of A.TxC. The authors proposed that the destabilized intermediate is then the complex of association, and the effects of mismatches on the 5′ side were attributed to a destabilization of the resulting heteroduplex. We calculated the relative variation in enthalpy of association, ΔE_{diff}, for each "mutation" at the center of the single strand, as detailed in the Methods section. The results for G.C and C.G pairs are shown in Table 4. They indicate destabilization in all cases, but with an order slightly different from that observed experimentally. Preliminary investigation of mismatches at A.T and T.A did not always lead to destabilization, but the variability of the interaction patterns for these bases requires a more complete study. It will also be necessary to look into the effect of mismatches on the base exchange pathway. This work is in progress.

DISCUSSION

We have previously proposed a plausible conformation for a triple helix in which the invading single strand interacts in the minor groove of a homologous duplex under the conditions of stretching and unwinding imposed by the RecA filament. We show here that such a triplex can easily transform into a more conventional triplex (which can be assimilated with the so-called R-DNA form), with the single strand in the major groove of the duplex. This is achieved by rotation of the bases of the complementary strand of the duplex, leading to a switch in base pairing (see similar suggestions made by Nishinaka et al., 1998). At constant stretching, this transformation is energetically very favorable, because of the stronger non-Watson-Crick interactions within the resulting major-groove triplex, together with a stabilization of both the ssDNA and dsDNA elements of this triplex. These results clearly favor the hypothesis of an initial association of the ss and dsDNA via a triple helix and provide new elements to support the experimental indications of a single-strand invasion of the duplex by its minor groove. This hypothesis is further supported by the different properties of the two triplexes with respect to stretching and winding. As expected (see Introduction), the major-groove triple helix is stabilized when compressed and rewound, in agreement with the experimentally isolated ternary product of strand exchange. Moreover, a recent study by Wong et al. (1998) has shown an unwinding requirement for circular dsDNA by at least six negative superturns for its homologous association with the filament, but not following synapsis. All of these elements support the identification of the minor-groove and major-groove triplexes with, respectively, the complex of association and the product of strand

TABLE 3 Comparison of the interproton distances H2′-H6/8, H2″-H6/8, and H3′-H6/8 for the ssDNA of the minor-groove and major-groove triplexes

	G1pC2		C2pG3		G3pG4		TpA
	Strand 3	**Strand 1**	Strand 3	**Strand 1**	Strand 3	**Strand 1**	NMR
H2′-H6/8	3.0 (0.4)	**2.6 (0.5)**	4.1 (0.2)	**5.8 (0.9)**	3.9 (1.1)	**3.8 (0.9)**	*4.5*
H2″-H6/8	3.6 (0.3)	**3.2 (0.7)**	3.6 (0.1)	**6.8 (1.2)**	4.1 (1.4)	**3.2 (0.7)**	*3.1*
H3″-H6/8	4.6 (0.4)	**4.1 (0.7)**	4.0 (0.3)	**5.3 (0.8)**	5.3 (1.0)	**4.8 (0.5)**	*4.3*

Strand 3 was used in the case of the minor-groove triplex (Bertucat et al., 1998); strand 1 was used in the case of the major-groove triplex (bold). The values (in Å) in the first six columns are average values for selected sets of structures following combinatorial searches on the nine independent sugars. The values in parentheses show the corresponding variance. These values are compared with the model proposed by Nishinaka et al. (1997) on the basis of NMR data (italics). The gap is situated at steps G3pG4 in the minor-groove triplex, C2pG3 in the major-groove triplex.

exchange. These two triplexes could thus characterize, at the atomic level, two of the ternary intermediates revealed by the kinetic studies of Ellouze et al. (1997b) and Gumbs and Shaner (1998), which exchange via a first-order reaction.

Driving force of strand exchange

Our present results suggest that RecA takes advantage of the intrinsic properties of extended triple helices rather than actively promoting strand exchange. In our model, the role of the protein appears to consist mainly of directing the approach of the duplex via the minor groove and trapping it in an extended form (see Léger et al., 1998). This complex is stabilized by its interaction with the already stretched ssDNA, provided that the degree of homology is sufficient. Further stabilization is then achieved via strand exchange and can be improved if some compression and winding accompany the exchange. We are not aware of any concrete data supporting this possibility. Little is known about the pitch of the nucleoproteic complex resulting from strand exchange, and it is not clear whether the 19.5 base triplets per turn, recently measured by Kiianitsa et al. (1997), corresponding to an average twist of 19°, concern the complex of association or the product of strand exchange.

The stabilization of the triple helix following the passage from the minor-groove to the major-groove form could account for the short lifetime of the association intermediate and its quick transformation into the strand exchange product, provided that the activation energy is sufficiently low. Although we have not yet accurately determined the value of this energy for all triplets, our calculations indicate that it may be on the order of 5 kcal·mol^{-1}. In addition, the results obtained under trinucleotide symmetry constraints indicate a possible cooperativity between neighboring triplets during strand exchange due to the disappearance of backbone strain as a contiguous stretch of the major-groove triplex is formed. It is probable that this is an important driving force for strand exchange, and calculations of its propagation along the triplex are under way.

An NMR-based modeling study of a RecA nucleofilament (Nishinaka et al., 1998) has proposed that sugar repuckering from C3′-endo to C2′-endo could play an active role in strand exchange. This hypothesis arose from the sugar-dependent pitches obtained when duplex DNA was modeled with NMR constraints. The pitches, respectively 95 Å (C3′-endo) and 64 Å (C2′-endo), correspond to those measured in active and inactive RecA filaments, the average rise value remaining at 5.1 Å in both cases. Our results do not confer such an important role on sugar conformation because, for identical rise values, both the minor-groove and major-groove triplexes present a significant proportion of north sugars, and this proportion can even be higher in certain major-groove triplexes. The sugar puckers do not show any notable variation during base pair switching (data not shown). Fig. 5 *c* shows that within a given triplex structure, a decrease in the helical pitch at constant rise is indeed accompanied by a decrease in the proportion of C3′-endo sugars. We therefore believe that the effect observed by Nishinaka et al. may simply be related to variations in pitch and not to strand exchange-induced conformational changes. Repuckering the sugars from south to north has previously been observed within the R-DNA form by Zhurkin et al. (1994) by increasing both the pitch and the average rise and was experimentally confirmed by a Fourier transform infrared study (Dagneaux et al., 1995b). Note that in the study of Nishinaka et al. (1998), as well as in that of Zhurkin et al. (1994), the rigidity introduced in the structures by assuming a uniform distribution of the stretching deformation induces a homogeneity in the sugar puckers that is not imposed in our modeling.

Related theoretical studies

The proposed model of association via the minor groove is supported by previous studies of DNA structural deformations performed in our laboratory. First, Lebrun and Lavery (1996) showed that upon further extension, the 3′-3′ stretched dsDNA that we used as a starting point to construct the minor groove triplex opens its base pairs toward the minor groove. A study of base pair opening as a function of twist also showed that opening toward the minor groove was enthalpically favored when the double helix was locally unwound to 16° per base pair step, over a 3-bp stretch (Bernet et al., 1997). An exchange mechanism implying base pair opening toward a third strand situated in the minor groove is thus favored both by unwinding and by stretching. In parallel with these results, a model of sequence recognition via the juxtaposition of two double-stranded helices has

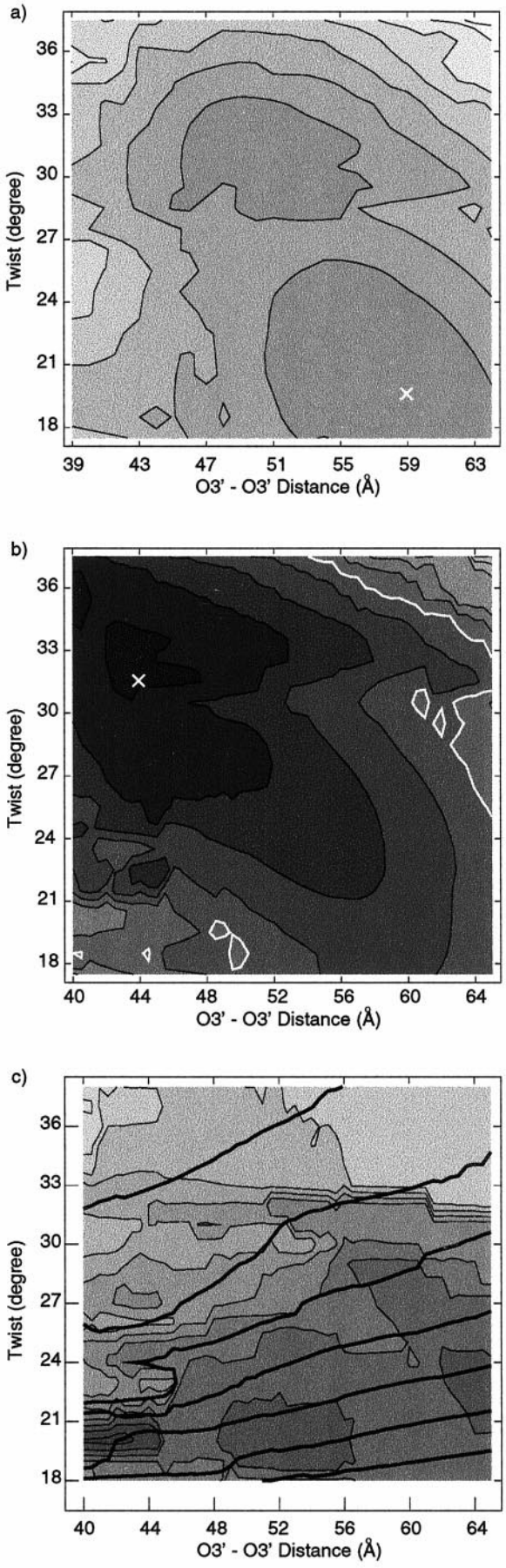

TABLE 4 Effect of a mutation on the enthalpy of formation of minor-groove triplexes with the sequence d(GCGGCXGCGGCG), where X is G or C

	G	C	A	T
G.CxM	0	4	6	10
C.GxM	6	0	2	6

The values, given in kcal · mol^{-1}, correspond to the enthalpy of formation of triplexes presenting all possible mutations in the ssDNA (base M in strand 3), relative to the most stable case (underlined). The enthalpies are calculated as described in the Methods section.

been proposed by McGavin as early as 1971 (McGavin, 1971, 1973) and subsequently developed by Wilson to provide a mechanism for the nick-free formation of stable quadruplexes (Wilson, 1979). This mechanism consists of recognition via the minor-groove interaction of two duplexes that have been stretched and unwound to avoid steric hindrance. This is followed by simultaneous base pair opening of the two double helices toward the minor grooves to form heterologous base pairs, which now interact via their major-groove faces (Fig. 6 *a*). The resulting quadruplex can then rewind and recover a canonical rise value. The structure has been constructed by Lebrun and Lavery (1995). It proved remarkably stable and showed sequence specificity, in support of the McGavin model (see McGavin, 1971). It is remarked that two Watson-Crick base pairs interacting via their minor groove faces have also been observed crystallographically by the association of two DNA loops belonging to the sequence d(GCATGCT) (Leonard et al., 1995; Brown and Hunter, 1997).

In the triple helix that we propose as a recognition model between ssDNA and dsDNA, the bases adopt an interesting pattern with respect to the McGavin and Wilson model. As shown in Fig. 3 (*top*) and schematically in Fig. 6 *b*, the ssDNA base is already positioned roughly perpendicularly to the dsDNA base pair, so that it does not need to rotate significantly for strand exchange to occur. In fact, only the base of the complementary strand needs to rotate. We have effectively obtained a strand exchange reaction by simply

FIGURE 5 Adiabatic mapping of the relative energy of each symmetry unit with respect to the projected O3′-O3′ distance (*horizontal axis*) and the average twist (*vertical axis*), respectively, in the (*a*) minor groove triple helix and (*b*) major groove triple helix. The energy is calculated with respect to relaxed ssDNA and dsDNA components. Consecutive gray levels are separated by 10 kcal·mol^{-1}, ranging from −50 (*black*) to 100 (*white*). The zero isoenergy curve is indicated by a white line, and the absolute minimum is shown in each map by a white cross. The isoenergy curves range from 20 to 80 kcal·mol^{-1} for the minor groove triplex, from −40 to 50 kcal·mol^{-1} for the major groove triplex. (*c*) Coupling between the average sugar phase and the pitch of the structures generated during adiabatic mapping. The gray levels, ranging from 200° (*black*) to 0° (*white*) by steps of 10°, concern the variations of the sugar phase averaged over the nine independent sugars of the trinucleotide symmetry unit. This average value varies between 70° and 171°. The variations in pitch *P*, from 32 Å to 109 Å, are represented by bold lines superimposed on the map. The separation between two lines is 10 Å. The pitch increases from the upper left to the lower right corner.

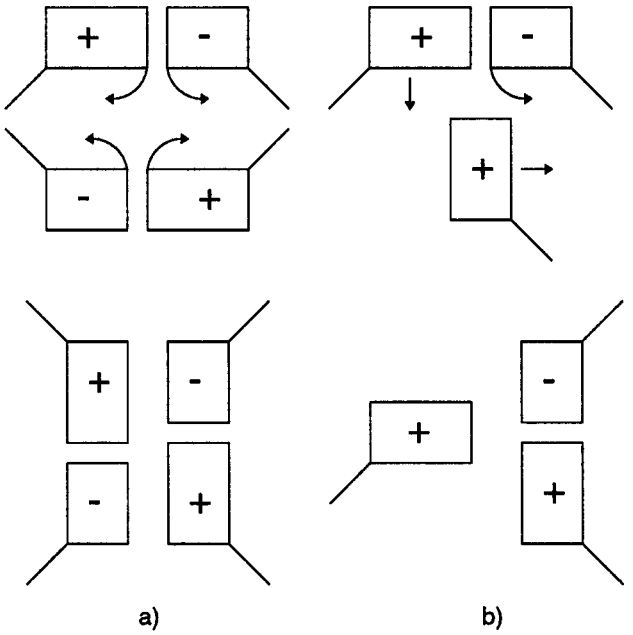

FIGURE 6 (*a*) Schematic representation of the model proposed by Wilson (1979) for the formation of two heteroduplexes, via base pair switching in two juxtaposed duplexes interacting through their minor grooves. (*b*) Simplified representation of the model we propose for RecA-promoted strand exchange. Rectangles represent the nucleic bases, and the line in the corner represents the glycosidic bond. In the top view, the ssDNA interacts in the minor groove of the duplex. In the bottom view, after strand exchange, the new ssDNA interacts with the major groove of the duplex. The + and − signs respectively designate 5′-3′ and 3′-5′ strand directions.

rotating this base toward the minor groove by 50° to 60°. In this process, the relative movement of the identical bases in strands 1 and 3 is principally translational.

Structure of the RecA filament during strand exchange

Our model of strand exchange only concerns the nucleic acid partner of the biologically relevant complex. It is interesting, however, to anticipate how the different properties of the minor-groove and major-groove triplexes may influence the structure of the RecA filament. Three states have been postulated for the filament, based on its degree of affinity for DNA (Kowalczykowski, 1991). Allosteric changes between these states are supposedly controlled by the nucleotide cofactors and produce the different geometric properties of the filament as determined by electron microscopy (Stasiak and DiCapua, 1982; Flory et al., 1984; Heuser and Griffith, 1989; Yu and Egelman, 1990, 1992), small-angle neutron scattering (SANS) (DiCapua et al., 1990; Ellouze et al., 1995), and x-ray crystallography (Story et al., 1992). Thus the "high-affinity DNA binding" or "active" state, with ATP as a cofactor, shows a pitch of 95–100 Å for a diameter of 100 Å. The "low-affinity DNA binding" state, where ADP is the cofactor, has a pitch of 83 Å and a 120-Å diameter. Both states have about six RecA monomers per turn, which implies an average rise of 4.6 Å for DNA complexed to RecA/ADP. The last state, with no cofactor, does not bind dsDNA. The mechanical properties of the various states of the filament also differ, as underlined by the distinct relationships established between the pitch and the number of RecA monomers per turn (Yu and Egelman, 1992).

In our model triple helices, the diameters defined by the C1′ atoms of the triplets increase by 20% when going from a minor groove triplex to the major groove form. This is the same increase as observed in the filament when going from a RecA/ATP to a RecA/ADP form. Added to the clear fit between the minor-groove triple helix and the active form of the RecA filament and to the propensity of the major-groove triplex to reduce its pitch, this suggests that the product of strand exchange may have a geometry adapted to the RecA/ADP form of the filament. Whether the strand exchange process by itself is capable of inducing a conformational change in the filament, similar to that induced by ATP hydrolysis, is worth considering. There are several indications of strong interactions between RecA and the phosphodiester groups of the DNA strands that could support this hypothesis (see Pugh and Cox, 1988; Mazin and Kowalczykowski, 1996, 1998). The reorganization of the three strands upon strand exchange could then result in a reorientation of the RecA monomers, similar to that observed when going from the active form of the filament to the inactive one or to the RecA/ADP form (Yu and Egelman, 1992). The energy necessary for this conformational change in the filament may arise from the energy release following strand exchange.

However, the stable triplexes resulting from strand exchange have only been observed when the cofactor is hydrolyzable ATP (Chiu et al., 1993; Rao et al., 1993; Jain et al., 1995). Jain et al. (1995) have even specified that the hydrolysis of ATP is a requirement for the appearance of these triple helices. Although strand exchange effectively occurs in the presence of ATPγS or ADP-AlF_4^- and gives rise to a ternary complex where the outgoing strand remains in the filament, several lines of evidence indicate that this strand does not interact strongly with the heteroduplex (Adzuma, 1992; Kowalczykowski and Krupp, 1995; Podyminogin et al., 1996; Zhou and Adzuma, 1997). Notably, it fails to cross-link to the duplex (Jain et al., 1995) and can easily be displaced by SSB (single strand binding protein) or by a nonhomologous single strand that presents a higher affinity for RecA (Mazin and Kowalczykowski, 1998). In view of our atomic model of strand exchange, these results may signify a low compatibility between the triplex resulting from strand exchange and the active form of the RecA filament, blocked by the absence of ATP hydrolysis. Ensuing distortions in the major groove triplex may lead to the observed separation of the outgoing strand from the heteroduplex. After strand exchange, an overall inclination of the bases, by at least 20° to 30° with respect to the helix axis, would be necessary to conserve the original backbone diameter, and thus it is interesting to note that such an incli-

nation has been observed in a complex between RecA/ATPγS and a classical antiparallel triplex (Kim et al., 1995). A duplex DNA obtained by RecA/ATPγS promoted renaturation and probably occupying the same position in the filament as the heteroduplex (see below) also exhibited inclinations of, respectively, 20° and 30° for the two strands (Nordén et al., 1998).

The three-stranded product of strand exchange, imbedded in a RecA filament in its active form, before or in the absence of ATP hydrolysis, may constitute the third ternary intermediate detected by Gumbs and Shaner (1998). It may also present a particular interest when strand exchange is promoted by Rad51, known for its low ATPase activity, and which does not seem to induce a conformational change in the filament of the same amplitude as that observed with RecA (Ellouze et al., 1997a).

Even if the allosteric change of the RecA filament is not driven by strand exchange, the relative displacement of the three backbones is likely to promote local modifications that may favor ATP hydrolysis, for example by bringing together the amino acid residues involved in the reaction. Likewise, the allosteric change induced in the RecA filament by ATP hydrolysis may facilitate base pair switching by positioning the three DNA backbones in the conformation they occupy in the major-groove triple helix. This could explain how ATP hydrolysis allows small heterologous stretches to be bypassed, in accord with experimental observations (DasGupta and Radding, 1982a; Rosselli and Stasiak, 1991). In this sense, it seems that RecA locally retains the role of motor molecule analogous to that of actin or myosin, where ATP hydrolysis is responsible for large structural changes (Amos and Cross, 1997; see Cox, 1994).

The coupling between ATP hydrolysis and the unidirectional release of ssDNA (Bedale and Cox, 1996), necessarily initiated at a free homologous 5′ end, remains to be interpreted within our model. If the major-groove triplex really fits the RecA/ADP form of the filament as proposed here, the interactions between the outgoing single strand and the heteroduplex should be very stable, and the release of the single strand should be disfavored. This is in fact the case when there is no free 5′ end. The biological role of a major groove triple helix that would appear only when a heterologous 5′ end blocks the strand exchange is also puzzling. A possible role could be the recycling of homologous stretches imbedded in heterologous DNAs, in the course of the dynamic turnover observed in several studies and conditioned by ATP hydrolysis (Burnett et al., 1994; Reddy et al., 1995).

Location of the three strands within the RecA filament

Coupled with recent findings concerning the role and specificity of the different binding sites in RecA (Mazin and Kowalczykowski, 1996, 1998), our structural model suggests some modifications to the scheme for RecA-mediated strand exchange (Fig. 7), initially proposed by Howard-Flanders et al. (1984). Favoring a major-groove approach, these authors anticipated the coaxial arrangement of the three interacting strands within RecA filaments, together with the necessity of at least two independent binding sites. In Fig. 7, the position of each strand with respect to these sites is specified, taking into account available experimental information. (Identical positions were proposed by Zhurkin et al. (1994; their figure 8) in the context of their major groove approach model.) In addition, the respective positions of each of the three strands and their movements during strand exchange reflect the nature of our model.

The incoming ssDNA first binds, slowly but very tightly, to what is usually called site 1 (Fig. 7 *a*; RecA filaments can accommodate three independent single strands, and site 1 is the first to be occupied; Takahashi et al., 1991). We can assume that the position of this strand within site 1 is very

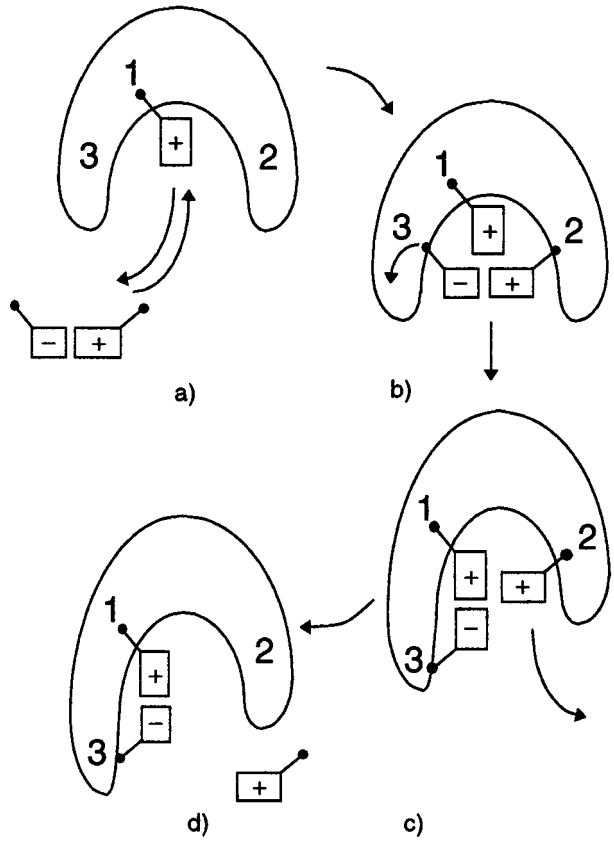

FIGURE 7 Model for the RecA-promoted strand exchange. A section of the filament perpendicular to its axis is represented at various stages of the strand exchange reaction. The bases are represented in the same way as in Fig. 6. RecA is only illustrated by a schematic cross section, which may contain more than one protein monomer. Numbers indicate the different sites for DNA complexation (see text). (*a*) Uptake of the dsDNA by the ssDNA/RecA filament, via its minor groove side. Depending on the degree of homology, the dsDNA will go back into solution or (*b*) remain associated with the ssDNA, forming a minor groove triple helix. This triplex readily undergoes strand exchange, producing (*c*) the major groove triplex. (*d*) The outgoing single strand, situated at site 2 of the protein filament, is released upon ATP hydrolysis or displaced by SSB.

well defined and that it probably involves interactions with the DNA backbone. Next, dsDNA binds less tightly to sites 2 and 3 (Fig. 7 *b*). Its 5′-3′ strand, destined to become the leaving strand after strand exchange, lies at site 2, as argued by Mazin and Kowalczykowski (1996, 1998). Strand exchange occurs by a rotation at site 3, where the complementary strand is only weakly bound, whereas the outgoing strand principally undergoes a shift with respect to incoming ssDNA. This shift does not significantly alter the distance between the backbones at sites 1 and 2, but modifies the angle formed by the vector joining the C1′ atoms and the glycosidic bonds (Fig. 7 *c*). It may be remarked that the three strands need not necessarily interact with the same RecA monomer at a given level.

In the framework of this model, the interpretation of experimental results involving only two DNA strands requires consideration of the binding sites involved and the possible state of the filament. It is reasonable to assume, within our scheme, that the two antiparallel strands of a double-stranded DNA will occupy either sites 2 and 3, as in the initial triplex (Fig. 7 *b*), or sites 1 and 3, as in the triplex after strand exchange (Fig. 7 *c*). In both cases, the duplex interacts with RecA through its minor groove, as experimentally established (DiCapua and Muller, 1987; Kumar and Muniyappa, 1992). In the second case, we would expect the duplex to be distorted in the absence of ATP hydrolysis, as argued above. In a study in which they monitored the different sites of binding, Wittung et al. (1994) have observed that, in the presence of ATP, renaturation occurred at both sites 2 and 3 and at sites 1 and 3, the latter combination being the most efficient. It was also found possible between sites 1 and 2, which would imply a shift of the second strand from site 2 to site 3 for pairing to occur. This corresponds to the situation in which base inclinations of 20° and 30° were measured for two paired strands in the presence of ATPγS (Nordén et al., 1998). Under the same conditions, the formation of regular Watson-Crick pairing could not be observed in the study of Wittung et al. (1994), whether the strands occupied sites 1 and 2 or sites 1 and 3. It would be interesting to determine whether this would still be the case for strands in positions 2 and 3. The shift from position 1,2 to position 1,3 that we postulate for pairing two complementary strands may also account for the puzzling results of a linear dichroism experiment (Wittung et al., 1996), obtained in the presence of RecA/ATPγS. A loss of signal was observed, indicating orientation disorder, whereas association of an identical or a heterologous ssDNA reinforced the signal.

Homology recognition within the minor-groove triplex

A last fundamental point remains to be discussed, namely RecA-controlled homologous recognition. Without any doubt, recognition within a triple-helical structure is more efficient via interactions in the major groove, which are both stronger and more specific than those in the minor groove. However, experimental observations rule out a simple mechanism for recognition. Strong discriminating interactions would impede any rapid search of long DNA stretches, currently believed to occur via collisions followed by quick release of the duplex DNA in the case of insufficient homology (Ellouze et al., 1997b; Malkov and Camerini-Otero, 1998; Adzuma, 1998). The results of our preliminary investigations show that the minor-groove triple helix is commonly although not always destabilized by the introduction of a heterologous base into the ssDNA already complexed to RecA. Given the low stability of the resulting triplex, this destabilization may be sufficient in many cases for a return to the reactants. However, it is probable that strand exchange also participates in detecting heterology (Bazemore et al., 1997a), via a dynamic process that produces different results at the 3′ and 5′ extremities according to the experiments of Malkov and Camerini-Otero (1998). We are continuing our modeling of the effects of heterology on both triplex stability and the strand exchange reaction to attempt to clarify these proofreading mechanisms.

CONCLUSIONS

By deliberately considering only the nucleic acid component of the RecA filament coupled with the global conditions of stretching imposed by RecA, we have been able to propose a detailed atomic model for homologous duplex DNA association and strand exchange. By suggesting coupling between strand exchange and the state of the filament, this model appears to rationalize various experimental results and throws new light on a possible role for ATP hydrolysis. Moreover, it supports the proposition of Kowalczykowski and Krupp (1995) that RecA binding alone may be sufficient to induce strand exchange. Last, our preliminary results imply that homologous recognition can occur partly during the formation of the initial minor-groove triplex, although further proofreading, via perturbation of the exchange pathway or by destabilization of the major-groove triplex following local strand exchange, may also be important.

REFERENCES

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.

Adzuma, K. 1998. No sliding during homology search by RecA protein. *J. Biol. Chem.* 273:31565–31573.

Amos, L. A., and R. A. Cross. 1997. Structure and dynamics of molecular motors. *Curr. Opin. Struct. Biol.* 7:239–246.

Baliga, R., J. W. Singleton, and P. B. Dervan. 1995. RecA.oligonucleotide filaments bind in the minor groove of double-stranded DNA. *Proc. Natl. Acad. Sci. USA.* 92:10393–10397.

Baumann, P., and S. C. West. 1998. Role of the human RAD51 protein in homologous recombination and double-stranded-break repair. *Trends Biochem. Sci.* 23:247–251.

Bazemore, L. R., E. Folta-Stogniew, M. Takahashi, and C. M. Radding. 1997a. RecA tests homology at both pairing and strand exchange. *Proc. Natl. Acad. Sci. USA.* 94:11863–11868.

Bazemore, L. R., M. Takahashi, and C. M. Radding. 1997b. Kinetic analysis of pairing and strand exchange catalyzed by RecA. Detection by fluorescence energy transfer. *J. Biol. Chem.* 272:14672–14682.

Bedale, W. A., and M. Cox. 1996. Evidence for the coupling of ATP hydrolysis to the final (extension) phase of RecA protein-mediated DNA strand exchange. *J. Biol. Chem.* 271:5725–5732.

Bernet, J., K. Zakrzewska, and R. Lavery. 1997. Modelling base pair opening: the role of helical twist. *J. Mol. Struct. Theochem.* 398–399: 473–482.

Bertucat, G., R. Lavery, and C. Prévost. 1998. A model for parallel triple helix formation by RecA: single-strand association with a homologous duplex via the minor groove. *J. Biomol. Struct. Dyn.* 16:535–546.

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

Brown, T., and W. N. Hunter. 1997. Non-Watson-Crick base associations in DNA and RNA revealed by single crystal x-ray diffraction methods: mismatches, modified bases and non-duplex DNA. *Biopolymers.* 44: 91–103.

Burnett, B., B. J. Rao, B. Jwang, G. Reddy, and C. M. Radding. 1994. Resolution of the three-stranded recombination intermediate made by RecA protein. An essential role of ATP hydrolysis. *J. Mol. Biol.* 238: 540–554.

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

Chiu, S. K., B. J. Rao, R. M. Story, and C. M. Radding. 1993. Interactions of three strands in joints made by RecA protein. *Biochemistry.* 32: 13146–13155.

Cluzel, P., A. Lebrun, C. Heller, R. Lavery, J.-L. Viovy, D. Chatenay, and F. Caron. 1996. DNA: an extensible molecule. *Science.* 271:792–795.

Cox, M. M. 1994. Why does RecA protein hydrolyse ATP? *Trends Biochem. Sci.* 19:217–222.

Dagneaux, C., H. Porumb, J. Liquier, M. Takahashi, and E. Taillandier. 1995a. Conformations of three-stranded DNA structures formed in presence and in absence of the RecA protein. *J. Biomol. Struct. Dyn.* 13:465–470.

Dagneaux, C., A. K. Shchyolkina, J. Liquier, V. L. Florentiev, and E. Taillandier. 1995b. A triple helix obtained by specific recognition of all 4 bases in duplex DNA can adopt a collapsed or an extended form. *C. R. Acad. Sci. III.* 318:559–562.

DasGupta, C., and C. M. Radding. 1982a. Polar branch migration promoted by RecA protein: effect of mismatched base pairs. *Proc. Natl. Acad. Sci. USA.* 79:762–766

DasGupta, C., and C. M. Radding. 1982b. Lower fidelity of RecA protein catalysed homologous pairing with a superhelical substrate. *Nature.* 295:71–73.

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

DiCapua, E., M. Schnarr, R. W. Ruigrok, P. Lindner, and P. A. Timmins. 1990. Complexes of RecA protein in solution. A study by small angle neutron scattering. *J. Mol. Biol.* 214:557–570.

Ellouze, C., M. Takahashi, P. Wittung, K. Mortensen, M. Schnarr, and B. Nordén. 1995. Evidence for elongation of the helical pitch of the RecA filament upon ATP and ADP binding using small-angle neutron scattering. *Eur. J. Biochem.* 233:579–583.

Ellouze, C., H.-K. Kim, K. Maeshima, E. Tuite, K. Morimatsu, T. Horii, K. Mortensen, B. Nordén, and M. Takahashi. 1997a. Nucleotide cofactor-dependent structural change of *Xenopus laevis* Rad51 protein filament detected by small-angle neutron scattering measurements in solution. *Biochemistry.* 36:13524–13529.

Ellouze, C., B. Nordén, and M. Takahashi. 1997b. Dissociation of non-complementary second DNA from RecA filament without ATP hydrolysis: mechanism of search for homologous DNA. *J. Biochem. (Tokyo).* 121:1070–1075.

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

Gumbs, O. H., and S. L. Shaner. 1998. Three mechanistic steps detected by FRET after presynaptic filament formation in homologous recombination. ATP hydrolysis required for release of oligonucleotide heteroduplex product from RecA. *Biochemistry.* 37:11692–11706.

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

Hingerty, B. E., R. H. Richtie, T. L. Ferrell, and J. E. Turner. 1985. Dielectric effects in biopolymers: the theory of ionic saturation revisited. *Biopolymers.* 24:427–439.

Howard-Flanders, P., S. C. West, and A. Stasiak. 1984. Role of RecA protein spiral filaments in genetic recombination. *Nature.* 309:215–219.

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

Jain, S. K., M. M. Cox, and R. B. Inman. 1995. Occurrence of three-stranded DNA within a RecA protein filament. *J. Biol. Chem.* 270: 4943–4949.

Kiianitsa, K., and A. Stasiak. 1997. Helical repeat of DNA in the region of homologous pairing. *Proc. Natl. Acad. Sci. USA.* 94:7837–7840.

Kim, S. K., M. Takahashi, and B. Nordén. 1995. Binding of RecA to anti-parallel poly(dA)·2poly(dT) triple helix DNA. *Biochem. Biophys. Acta.* 1264:129–133.

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

Kowalczykowski, S. C., D. A. Dixon, A. K. Eggleston, S. D. Lauder, and W. M. Rehrauer. 1994. Biochemistry of homologous recombination in *Escherichia coli. Microbiol. Rev.* 58:401–465.

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

Kowalczykowski, S. C., and R. A. Krupp. 1995. DNA-strand exchange promoted by RecA protein in the absence of ATP: implications for the mechanism of energy transduction in protein-promoted nucleic acid transactions. *Proc. Natl. Acad. Sci. USA.* 92:3478–3482.

Kubista, M., M. Takahashi, and B. Nordén. 1990. Stoichiometry, base orientation, and nuclease accessibility of RecA complexes seen by polarized light in flow-oriented solution. Implications for the mechanism of genetic recombination. *J. Biol. Chem.* 265:18891–18897.

Kumar, K. A., and K. Muniyappa. 1992. Use of structure-directed DNA ligands to probe the binding of RecA protein to narrow and wide grooves of DNA and on its ability to promote homologous pairing. *J. Biol. Chem.* 267:24824–24832.

Kurumizaka, H., and T. Shibata. 1996. Homologous recognition by RecA protein using non-equivalent three DNA-strand-binding sites. *J. Biochem.* 119:216–223.

Lavery, R. 1988. Junctions and bends in nucleic acids: a new theoretical modeling approach. *In* Structure and Expression, Vol 3: DNA Bending and Curvature. W. K. Olson, R. H. Sarma, M. H. Sarma, and M. Sundaralingam, editors. Adenine Press, Schenectady, NY. 191–211.

Lavery, R. 1995. Modelling the DNA double helix: techniques and results. *In* Non-Linear Excitation in Biomolecules. M. Peyrard, editor. Springer-Verlag, Editions de Physique, Berlin and New York. 57–82.

Lavery, R., and H. Sklenar. 1988. The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* 6:63–91.

Lavery, R., and H. Sklenar. 1989. Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dyn.* 6:655–667.

Lavery, R., H. Sklenar, K. Zakrzewska, and B. Pullman. 1986. The flexibility of the nucleic acids. II. The calculation of internal energy and applications to mononucleotide repeat DNA. *J. Biomol. Struct. Dyn.* 3:989–1014.

Lavery, R., K. Zakrzewska, and H. Sklenar. 1995. JUMNA (junction minimisation of nucleic acids). *Comput. Phys. Commun.* 91:135–158.

Lebrun, A., and R. Lavery. 1995. Modelling a strand exchange tetraplex conformation. *J. Biomol. Struct. Dyn.* 13:459–464.

Lebrun, A., and R. Lavery. 1996. Modelling extreme stretching of DNA. *Nucleic Acids Res.* 24:2260–2267.

Léger, J. F., J. Robert, L. Bourdieu, D. Chatenay, and J. F. Marko. 1998. RecA binding to a single double-stranded DNA molecule: a possible role of DNA conformational fluctuations. *Proc. Natl. Acad. Sci. USA.* 95: 12295–12299.

Leonard, G. A., S. Zhang, M. R. Peterson, S. J. Harrop, J. R. Helliwell, W. B. Cruse, B. Langlois d'Estaintot, O. Kennard, T. Brown, and W. N. Hunter. 1995. Self-association of a DNA loop creates a quadruplex: crystal structure of d(GCATGCT) at 1.8 Å resolution. *Structure.* 3:335–340.

Little, J. W., S. H. Edmiston, L. Z. Pacelli, and D. W. Mount. 1980. Cleavage of the *Escherichia coli* lexA protein by the recA protease. *Proc. Natl. Acad. Sci. USA.* 77:3225–3229.

Malkov, A. M., and R. D. Camerini-Otero. 1998. Dissociation kinetics of RecA protein-three-stranded DNA complexes reveals a low fidelity of RecA-assisted recognition of homology. *J. Mol. Biol.* 278:317–330.

Mazin, A. V., and S. C. Kowalczykowski. 1996. The specificity of the secondary DNA binding site of RecA protein defines its role in DNA strand exchange. *Proc. Natl. Acad. Sci. USA.* 93:10673–10678.

Mazin, A. V., and S. C. Kowalczykowski. 1998. The function of the secondary DNA-binding site of RecA protein during DNA strand exchange. *EMBO J.* 17:1161–1168.

McGavin, S. 1971. Models of specifically paired like (homologous) nucleic acid structures. *J. Mol. Biol.* 55:293–298.

McGavin, S. 1973. Chromosome pairing. *Nature.* 242:330.

Muller, B., I. Burdett, and S. C. West. 1992. Unusual stability of recombination intermediates made by *Escherichia coli* RecA protein. *EMBO J.* 11:2685–2693.

Nishinaka, T., Y. Ito, S. Yokoyama, and T. Shibata. 1997. An extended DNA structure through deoxyribose-base stacking induced by RecA protein. *Proc. Natl. Acad. Sci. USA.* 94:6623–6628.

Nishinaka, T., A. Shinohara, Y. Ito, S. Yokoyama, and T. Shibata. 1998. Base pair switching by interconversion of sugar puckers in DNA extended by proteins of RecA-family: a model for homology search in homologous genetic recombination. *Proc. Natl. Acad. Sci. USA.* 95: 11071–11076.

Nordén, B., C. Elvingson, M. Kubista, P. Sjöberg, H. Ryberg, M. Ryberg, K. Mortensen, and M. Takahashi. 1992. Structure of RecA-DNA complexes studied by combination of linear dichroism and small-angle neutron scattering measurements on flow-oriented samples. *J. Mol. Biol.* 226:1175–1191.

Nordén, B., P. Wittung-Stafshede, C. Ellouze, H. K. Kim, K. Mortensen, and M. Takahashi. 1998. Base orientation of second DNA in RecA.DNA filaments. Analysis by combination of linear dichroism and small angle neutron scattering in flow-oriented solution. *J. Biol. Chem.* 273: 15682–15686.

Podyminogin, M. A., R. B. Meyer, Jr., and H. B. Gamper. 1996. RecA-catalyzed, sequence-specific alkylation of DNA by cross-linking oligonucleotides. Effects of length and non-homologous base substitutions. *Biochemistry.* 35:7267–7274.

Pugh, B. F., and M. Cox. 1988. High salt activation of RecA protein ATPase in the absence of DNA. *J. Biol. Chem.* 263:76–83.

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

Rao, B. J., S. K. Chiu, and C. M. Radding. 1993. Homologous recognition and triplex formation promoted by RecA protein between duplex oligonucleotides and single-stranded DNA. *J. Mol. Biol.* 229:328–343.

Rao, B. J., M. Dutreix, and C. M. Radding. 1991. Stable three-stranded DNA made by RecA protein. *Proc. Natl. Acad. Sci. USA.* 88: 2984–2988.

Reddy, G., B. Burnett, and C. M. Radding. 1995. Uptake and processing of duplex DNA by RecA nucleoprotein filaments: insights provided by a mixed population of dynamic and static intermediates. *Biochemistry.* 34:10194–10204.

Roberts, J. W., C. W. Roberts, and N. L. Craig. 1978. *Escherichia coli* recA gene product inactivates phage lambda repressor. *Proc. Natl. Acad. Sci. USA.* 75:4714–4718.

Rosselli, W., and A. Stasiak. 1990. Energetics of RecA-mediated recombination reactions. Without ATP hydrolysis RecA can mediate polar strand exchange but is unable to recycle. *J. Mol. Biol.* 216:335–352.

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

Stasiak, A., and E. DiCapua. 1982. The helicity of DNA in complexes with RecA protein. *Nature.* 299:185–186.

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

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

Takahashi, M., M. Kubista, and B. Nordén. 1991. Co-ordination of multiple DNA molecules in RecA fiber evidenced by linear dichroism spectroscopy. *Biochimie (Paris).* 73:219–226.

Takahashi, M., and B. Nordén. 1994. Structure of RecA-DNA complex and mechanism of DNA strand exchange reaction in homologous recombination. *Adv. Biophys.* 30:1–35.

Tuite, E., U. Sehlstedt, P. Hagmar, B. Nordén, and M. Takahashi. 1997. Effects of minor and major groove-binding drugs and intercalators on the DNA association of minor groove-binding proteins RecA and deoxyribonuclease I detected by flow linear dichroism. *Eur. J. Biochem.* 243:482–492.

Watt, V. M., C. J. Ingles, M. S. Urdea, and W. J. Rutter. 1985. Homology requirements for recombination in *Escherichia coli. Proc. Natl. Acad. Sci. USA.* 82:4768–4772.

Wilson, J. H. 1979. Nick-free formation of reciprocal heteroduplexes: a simple solution to the topological problem. *Proc. Natl. Acad. Sci. USA.* 76:3641–3645.

Wittung, P., L. R. Bazemore, M. Takahashi, B. Nordén, and C. Radding. 1996. Second-site RecA-DNA interactions: lack of identical recognition. *Biochemistry.* 35:15349–15355.

Wittung, P., B. Nordén, and M. Takahashi. 1994. Spectroscopic observation of renaturation between polynucleotides interacting with RecA in the presence of ATP hydrolysis. *Eur. J. Biochem.* 224:39–45.

Wong, B. C., S.-K. Chiu, and S. A. Chow. 1998. The role of negative superhelicity and length of homology in the formation of paranemic joints promoted by RecA protein. *J. Biol. Chem.* 273:12120–12127.

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

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

Zhou, X., and K. Adzuma. 1997. DNA strand exchange mediated by the *Escherichia coli* RecA protein initiates in the minor groove of double-stranded DNA. *Biochemistry.* 36:4650–4661.

Zhurkin, V. B., G. Raghunathan, N. B. Ulyanov, R. D. Camerini-Otero, and R. L. Jernigan. 1994. A parallel DNA triplex as a model for the intermediate in homologous recombination. *J. Mol. Biol.* 239:181–200.