THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 265, No. 28, Issue of October 5, pp. 16898-16912, 1990
Printed in U.S.A.
© 1990 by The American Society for Biochemistry and Molecular Biology, Inc.

Triple-helical DNA Pairing Intermediates Formed by recA Protein*

(Received for publication, February 14, 1990)

Scott W. Umlauf, Michael M. Cox‡, and Ross B. Inman

From the Department of Biochemistry, Program in Cellular and Molecular Biology, and Institute for Molecular Virology, University of Wisconsin, Madison, Wisconsin 53706

RecA protein aligns homologous single- and double-stranded DNA molecules in three-stranded joints that can extend over thousands of base pairs. When cross-linked by 4′-amino-4,5′,8-trimethyl-psoralen the joint structure observed is nonuniform and divided into multiple substructures each a few hundred base pairs long. Two paired substructures are observed; at least one, and possibly both, are right-handed triple helices. Sites of homologous contact are interspersed with regions where the DNA molecules are arranged side-by-side without contact. These substructures alternate in all combinations. The length and frequency of joints is much greater when one of the DNA substrates is linear, and interwinding is unrestricted, than when there are topological restrictions between the pairing partners. The results are consistent with the idea that recA protein facilitates the formation of a right-handed triple-helical DNA pairing intermediate during strand exchange. The results further suggest that recA filaments do not promote the formation of structures that provide efficient topological compensation for right-handed interwinding of two paired DNA molecules.

The recA protein of *Escherichia coli* promotes some of the central steps in homologous genetic recombination including the prerequisite alignment of homologous DNA sequences. RecA-mediated DNA pairing has been characterized *in vitro* with a wide variety of DNA substrates (for review, see Cox and Lehman, 1987; Radding, 1988). The first complex in strand exchange reactions in which homologous pairing has occurred is termed a paranemic joint, ("not truly interwound") (Bianchi *et al.*, 1983; Crick, 1957). Paranemic joints were originally observed and defined by Radding and colleagues (DasGupta and Radding, 1982; Wu *et al.*, 1983) using DNA substrates in which a net exchange of strands is precluded topologically. These joints are unstable in the absence of recA protein. If homologous contact occurs at an end or break in one of the DNA molecules, the alignment leads to a stable plectonemic, or interwound joint in which an exchange of strands has been initiated (Cassuto *et al.*, 1980; Shibata *et al.*, 1980). Such a joint is extended in recA-mediated DNA strand exchange leading to stable heteroduplex products (Cox and Lehman, 1981). An understanding of the structure of paranemic joints is of fundamental importance to the study of all aspects of recA protein-mediated strand exchange. The structure may also provide a model for homologous DNA interactions in meiotic and mitotic eukaryotic cells.

Many of the known properties of paranemic joints are governed by the structure of the recA filament. RecA protein binds cooperatively to single-stranded DNA, or duplex DNA that contains a single-stranded gap, to produce a polar, right-handed helical nucleoprotein filament. The DNA in this filament is underwound and extended to 18 bases per turn with an axial base separation of 5.1 Å (Egelman and Stasiak, 1986). The resulting nucleoprotein filament is capable of binding a second duplex DNA molecule (for review see Cox and Lehman, 1987; Radding, 1988). If this second molecule is at least partially homologous to the first, it is also underwound by the recA filament to permit alignment (Schutte and Cox, 1988; Wu *et al.*, 1983). Indirect measurements indicate that homologous alignment can extend over thousands of base pairs (Schutte and Cox, 1987; 1988). Measurement of the lengths of paranemic joints in reactions where the recA filaments have been fixed chemically has revealed contiguous homologous contacts ranging between 300 and 1200 bp (Register *et al.*, 1987). Three-stranded joints are formed much more readily than four-stranded joints even though recA protein promotes a four-strand exchange reaction (Lindsley and Cox, 1990). In three-stranded joints we refer to the single-stranded substrate on which the filament first forms as the initiating strand.

The structure of paranemic joints remains one of the central open questions in the recA system. The most intriguing possibility for DNA-DNA pairing in these joints is the formation of a DNA triple helix with base-specific cross-strand hydrogen bonds in the major groove of the duplex molecule. This idea was first advanced by Howard-Flanders and colleagues (1984), and is now widely accepted as the most likely possibility. DNA structures that may contain three strands have been observed by Stasiak *et al.* (1984), Register *et al.* (1987), and Griffith *et al.* (1990). In each case long regions that may contain three DNA strands are observed in gaps in recA filaments during strand exchange reactions between two completely homologous DNA substrates. Since junctions between single- and double-stranded molecules, which would confirm the presence of all three strands, are never observed within the gaps, the presence of all three strands is inferred on the basis of DNA length measurements and assumptions about recA protein binding patterns during strand exchange. Alternative explanations are possible in each case. For example, if recA protein remains bound to heteroduplex DNA (as indicated by several different solution measurements (Chow *et al.*, 1986; Pugh and Cox, 1987; Schutte and Cox, 1987)), then the circular triplexes of Register *et al.* (1987) and Griffith *et al.* (1990) instead could be the duplex products of strand exchange. Nevertheless, the reports suggest that recA protein facilitates the formation of a structure in which all three strands coincide over thousands of base pairs, and these observations have had a catalytic effect on efforts to examine

* This work was supported by National Institutes of Health Grants GM32335 (to. M. M. C.) and GM14711 (to. R. B. I.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed.

Downloaded from www.jbc.org at University of Wisconsin-Madison, on January 26, 2012

paranemic joints. The observed structures were fixed with glutaraldehyde, and their structure or stability could not be specified. In particular, there is as yet no evidence that the three strands are arranged in a triple helix.

Stimulated in part by the reports cited above, a variety of models has been advanced for the structure of three-stranded joints. These are presented schematically in Fig. 1. *Panels A–C* depict uniform structures and *panels D–F* depict composite structures in which two different segments with structures A, B, or C alternate. In *panel A* the DNA molecules are arranged "side-by-side"; homologous contacts in such a structure could occur only periodically, once per turn of the helical duplex (Cox *et al.*, 1987). In *panels B* and *C* are right- and left-handed triple helices, respectively, with hydrogen bonding patterns unspecified. Model B is functionally equivalent to a structure proposed by Howard-Flanders and colleagues (Howard-Flanders *et al.*, 1984). In *panel D* is a combination of the right-handed triple helix with regions where the DNA molecules are side-by-side and not involved in homologous interactions. This is not a trivial extension of model B because the side-by-side segments may have functional significance in DNA strand exchange (Cox, 1990). *Panel E* depicts the right-handed triple helix in combination with regions where the two DNA molecules are wrapped loosely in a left-handed fashion. This is similar to a model first proposed by Wilson (1979). In *panel F* is the model proposed by Fishel *et al.* (1988) in which right- and left-handed triple helices alternate.

If a joint is to be truly paranemic (no net or global interwinding of the two DNA molecules), local helical interwinding must be balanced globally by an equivalent number of helical turns of the opposite handedness. This topological problem presumably is encountered by DNA molecules undergoing recombination *in vivo*. An important codicil to the general problem of paranemic joint structure is the question of whether the recA system solves the topological problem internally by facilitating the formation of a topologically balanced strucutre. Models A, E, and F were largely inspired by this consideration. In model A, no local interwinding occurs, therefore no global compensation is necessary. Models E and F provide other possible solutions to this topological problem that might be facilitated by recA protein itself. Model E accomplishes this by means of compensating left-hand toroidal wrapping of one DNA molecule about the other outside of the regions of homologous contact. The toroidal wrapping could be incidental or facilitated in some manner by the recA filament. In model F, right-handed interwinding is balanced by left-handed turns within the recA filament so that no net interwinding occurs. Models B–D, however, involve net interwinding. If one of these models is correct, the topological problem introduced by forming one of these structures in cellular DNA must be solved by means external to the recA system (*e.g.* topoisomerases).

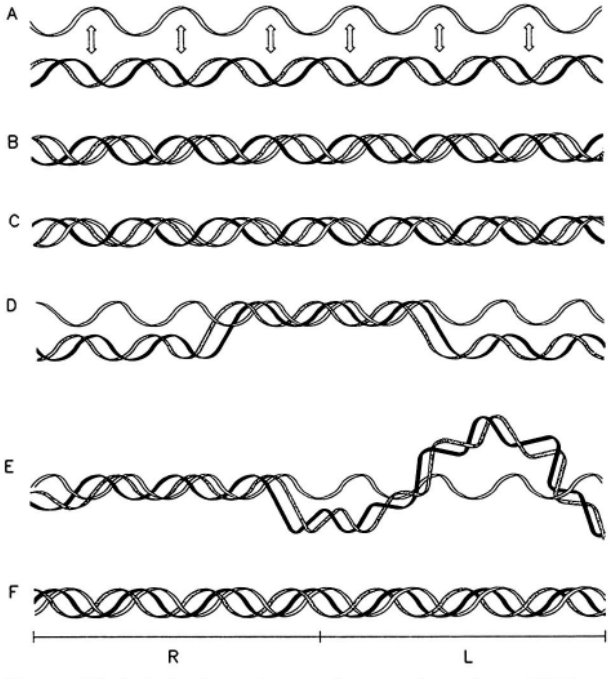


FIG. 1. **Models for homologous interaction of two DNA molecules (three strands) in a joint formed by recA protein.** *A*, a "side-by-side" homologous interaction. Local interwinding is less than one helical turn, global interwinding is zero. *B*, a right-handed triple helix. Local and global interwinding are nonzero. *C*, a left-handed triple helix. Local and global interwinding are nonzero. *D*, alternation of right-handed triple helix with side-by-side regions that are not in contact. Local and global interwinding are nonzero. *E*, alternation of right-handed triple helix with DNA regions not in homologous contact but which provide topological compensation in the form of interdomainal left-handed wraps. The beginnings of a left-handed counter-wrap are shown at the right of the figure. Local interwinding is nonzero, global interwinding is zero. *F*, alternating left- and right-handed triple helices. Local interwinding is nonzero, global interwinding is zero.

In the standard *in vitro* reaction, the initiating strand on which the recA filament forms is circular. If recA protein promotes the formation of a joint that is truly paranemic, *i.e.* the topological problem is solved by the recA filament in a facilitated process resembling models A, E, or F, we expect to see no difference in size or structure between joints formed with linear duplex DNA with heterologous ends and those formed with circular duplex DNA (Fig. 2, *A* and *B*). However, if homologous joints take a form similar to models B, C, or D of Fig. 1, the joints formed with these two duplex substrates may have very different structures. For example, joints formed with the linear duplex, where net interwinding is possible (Fig. 2*C*), may be much longer than those formed with a

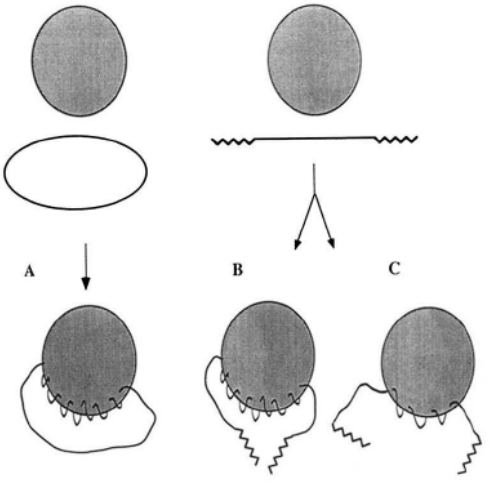


FIG. 2. **Topology of homologous alignment.** The *gray circle* in each drawing represents a single-stranded DNA molecule contained within a recA filament. The plane defined by the circle is shaded to ease identification of the handedness of nodes. The additional lines represent double-stranded DNA. Jagged lines depict heterologous regions of DNA. *A*, pairing restricted by topology. The joints must be globally paranemic in that all right-handed helical turns must be compensated by left-handed turns. *B*, *C*, pairing restricted by heterologous terminal sequences. *B*, a paranemic joint in which right-handed helical wraps are compensated by left-handed wraps. *C*, right-handed helical interwinding of the DNA molecules (a plectonemic joint).

Downloaded from www.jbc.org at University of Wisconsin-Madison, on January 26, 2012

circular duplex that precludes net interwinding (Fig. 2A).

It should be noted that a triple-helical recombination intermediate must differ structurally from the stable triple-helical DNA structures reported over the past few years (Hanvey *et al.*, 1988; Htun and Dahlberg, 1988; Moser and Dervan, 1987). In particular, triple helices observed to date involve antiparallel like strands and occur only where there is a strong pyrimidine or purine strand bias. In a three-stranded joint formed by the action of recA protein, the like strands must be parallel and the sequence must be irrelevant, other than the requirement for homology between partners.

All of the proposed structures for three-stranded homologous joints, with the exception of model A, are triple-helical in the sense that all three strands will be wound around a common axis. Homology-specific hydrogen bonding within such a helix could take many forms. The initiating strand could form hydrogen bonds to one or both of the strands in the duplex DNA.

In this study we examined the DNA structure in homologous three-stranded joints formed between DNA molecules in which strand exchange is blocked either by topology or by heterologous ends. We have used topological and chemical fixation to trap this structure in the presence of recA protein, and then removed the protein and examined the resulting DNA structures in the electron microscope. The results extend the work of Stasiak *et al.* (1984) and Register *et al.* (1987), and they provide evidence that model D of Fig. 1 depicts most accurately the overall structure of joints formed between circular single-stranded and linear duplex DNA. In addition, in a series of experiments carried out under identical conditions, joints are shown to be longer when interwinding of the DNA molecules is not restricted by topology, indicating that recA filaments do not facilitate the formation of any structure that provides topological compensation for a right-handed triple helix.

EXPERIMENTAL PROCEDURES

Enzymes, Chemicals, and DNA—E. coli recA protein was purified as described previously (Cox *et al.*, 1981). The concentration of recA protein in stock solutions was determined by the absorbance at 280 nm, using an extinction coefficient of $\epsilon_{280} = 0.59\ A_{280}\ mg^{-1}$ ml (Craig and Roberts, 1981). *E. coli* SSB protein was purified by the method of Lohman *et al.* (1986) and was stored frozen at −70 °C in a buffer containing 20 mM Tris, 40% cation (pH 8.4), 0.15 M NaCl, 1 mM EDTA, 1 mM β-mercaptoethanol, and 50% glycerol. The concentration of SSB protein was determined by the absorbance at 280 nm using an extinction coefficient of $\epsilon_{280} = 1.5\ A_{280}\ mg^{-1}$ ml (Lohman and Overman, 1985).

Single-stranded DNA (ssDNA)[1] derived from bacteriophage M13 mp8 was prepared as described (Messing, 1983). Circular double-stranded DNA (dsDNA) from bacteriophage M13Gori1 (Kaguni and Ray, 1979) as well as plasmid DNA was prepared as described (Garger *et al.*, 1983). Stock concentrations of ssDNA and dsDNA were determined by absorbance at 260 nm, with 36 and 50 $\mu g\ ml^{-1}\ A_{260}^{-1}$ used as conversion factors, respectively. DNA concentrations are expressed as total nucleotides. The construction of pMMC3 is described elsewhere (Cox, 1983). Nicked circular duplex DNA was prepared as described (DasGupta *et al.*, 1981). In brief, M13Gori1 form I DNA was treated with *Xho*I restriction enzyme in the presence of ethidium bromide. This produced ≈70% form II and ≈30% form III DNA in the preparations used in this study.

Restriction enzymes were purchased from New England Biolabs and Promega Biotech. Topoisomerase I was purchased from Bethesda Research Laboratories. 4′-Amino-4,5′,8-trimethyl-psoralen (AMT) was purchased from Calbiochem. Glutaraldehyde was purchased from Ladd Research Industries (Burlington, VT). All other biochemicals and enzymes were purchased from Sigma.

[1] The abbreviations used are: ss, single-stranded; ds, double-stranded; AMT, 4′-amino-4,5′,8-trimethyl-psoralen; Ac, acetyl; bp, base pairs.

Reaction Conditions—Reactions were carried out at 37 °C in standard reaction buffer consisting of 25 mM Tris-OAc (80% cation, pH 7.5), 10 mM MgOAc, 3 mM potassium glutamate, 3 mM ATP, and 1 mM dithiothreitol. All reaction volumes were 50 μl. Reactions were made up containing all components except SSB protein, ATP, and dsDNA. This mixture was preincubated at 37 °C for 2 min. At this point SSB and ATP were added and the reaction incubated for 1 min. Then dsDNA was added and the reaction incubated for another 2 min. Reactions were then subjected to fixation or halted by other means as described below. Reaction mixtures for experiment 1 contained 2.25 μM ssDNA, 1.13 μM recA protein, 0.15 μM SSB protein, and 6.0 μM dsDNA and a creatine kinase/phosphocreatine ATP-regenerating system as described (Bianchi *et al.*, 1983). Reaction mixtures for experiments 2–11 contained 1.5 μM ssDNA, 0.75 μM recA protein, 0.1 μM SSB protein, and 3.0 μM dsDNA, and a pyruvate kinase/phosphoenol pyruvate ATP-regenerating system as described elsewhere (Schutte and Cox, 1987).

For the experiments 4–10, recA and SSB proteins were removed after cross-linking by treatment with 1% sodium dodecyl sulfate (final w/v) and 1 mg/ml proteinase K for 30 min at 37 °C.

Cross-linking Conditions—Samples crosslinked with glutaraldehyde were treated with glutaraldehyde at 0.5% for 2 min at room temperature then transferred to ice. Samples cross-linked with topoisomerase I were treated with 10 units (BRL) of enzyme for 2 min at 37 °C. For samples cross-linked with AMT, the psoralen derivative was added to a final concentration of 30 μg/ml, the samples were incubated for 3 min at room temperature followed by a 5-min exposure to long-wave UV light at 0 °C. The UV light was generated with two fluorescent black light tubes. Samples were placed 8 cm below the light source. Intensity of the light ranged between 2 and 4 milliwatts/cm^2 at 365 nm (as measured with a Blak-ray J-221 long wave UV meter). A glass microscope slide was placed above the sample to filter light below 320 nm. Samples were checked for AMT cross-linking density by heating to 70 °C in the presence of 5% formaldehyde. All fixed or cross-linked samples were dialyzed into 20 mM NaCl, 5 mM EDTA overnight at 4 °C before Kleinschmiddt spreading.

Electron Microscopy—Unfixed samples (Fig. 4) were prepared for electron microscopy by separating filaments from unbound protein using gel filtration chromatography, pipetting onto an activated carbon film and positive staining with UAc. In particular, reactions were run through a spun column of Sepharose 4B preequilibrated in 25 mM Tris-OAc (80% cation, pH 7.5), 10 mM MgOAc, 3 mM potassium glutamate, 3 mM ATP, 1 mM dithiothreitol, and 20% glycerol. Columns were prepared in 0.5-ml Eppendorf centrifuge tubes and spun at 4000 rpm for 15 s in an Eppendorf centrifuge. Eluted fractions were placed on a carbon film activated with 1 μg/ml polylysine (Williams, 1977). Samples were washed on the surface of a solution of 20% glycerol, positively stained with two drops of 0.5% UAc in 20% glycerol, then washed again with 20% glycerol. Samples were blotted in an upright position with a piece of filter paper, then, after evaporation of the glycerol, shadowed with platinum at a high angle (17°) for maximum detection of double-stranded DNA near recA filaments.

Samples in which the DNA was cross-linked were prepared for electron microscopy by a modification of the cytochrome *c* spreading technique in which DNA (and in some cases covalently fixed recA protein) is spread onto a small water drop (Inman and Schnos, 1970). The spreading solution contained approximately 0.4 mg of DNA/ml, 0.01% cytochrome *c*, 7.3% formaldehyde, 14.3% formamide, 16 mM Na_2CO_3, 27 mM NaCl, and 2.5 mM EDTA. The solution was adjusted to pH 8.0 ± 1 (25 °C) with HCl before the addition of DNA and cytochrome *c*.

Computer-assisted length measurements were made from electron micrographs by tracing and recording on a semiautomatic digitizer, as described previously (Littlewood and Inman, 1982). All measured lengths were calculated in base pairs relative to substrate lengths of 7229 and 8623 bp for M13mp8 and M13Gori1, respectively. In cases where three strands were visible such tracings were done independently for each strand. A small percentage of contacts were separated by lengths which were unequal on two or more strands. These contacts were discarded as nonhomologous artifacts of two-dimensional compression due to the spreading process.

RESULTS

Experimental Design

The goal of this study was to define several features of the DNA structure formed when recA protein aligns two homol-

Downloaded from www.jbc.org at University of Wisconsin-Madison, on January 26, 2012

ogous DNA molecules; in particular the number of strands connected by homology-specific hydrogen bonds, and the topology of DNA strands. To permit pairing but preclude net strand exchange, substrates we used to study paranemic joints included circular single-stranded DNA molecules in combination with either linear double-stranded DNA molecules with heterologous sequences at the ends, or circular double-stranded DNA molecules nicked within a heterologous region. These are similar to the substrates used originally to define these joints (DasGupta and Radding, 1982; Wu *et al.*, 1983). Paranemic joints are formed within a recA filament and thus the DNA structure is obscured from view in the electron microscope by protein. We first attempted, unsuccessfully, to detect a stable three-stranded structure after pairing and removal of recA protein. Experiments with a wide variety of substrates and several different reaction conditions indicated that joints were too unstable to detect once recA protein was removed. Therefore in order to preserve some features of the structure of these joints, the DNA strands within a joint were fixed with either topoisomerase I or 4′-amino-4,5′,8-trimethyl-psoralen in the presence of long-wave UV light prior to the removal of recA protein. The resulting structures were then observed by electron microscopy after removal of recA protein. Multiple approaches were used to minimize the potential for artifacts.

Reaction conditions were chosen to reduce the occurrence of large DNA aggregates which are difficult to interpret. The data are based generally on definable joints involving only one double-stranded molecule and one single-stranded molecule. In a few cases, one single strand was aligned with two different duplexes, forming two distinct three-stranded joints. Similarly, a single duplex DNA molecule was sometimes found containing two distinct joints involving separate circular single strands. Since joints involving three DNA molecules were usually interpretable, they were included in the data sets. These joints did not exceed 10% of the definable joints in any experiment presented here.

Reaction conditions were varied as little as possible between experiments. All fixation conditions were optimized so that both the fixative concentration and time of fixation are as low as possible without loss of cross-linking efficiency. Except in the case of the control experiments with unfixed recA filaments, the electron microscopy was carried out on samples spread on a protein monolayer and shadowed with platinum. These spreading conditions allowed us to visualize and discriminate readily between double- and single-stranded regions of DNA. In this system, regions which contain three strands are identified unambiguously by the presence of junctions between single-stranded molecules and homologous duplexes.

Presented in Fig. 3 is a schematic drawing of the DNA molecules used in this study. As shown, M13mp8 (Messing, 1983) single-stranded circles and M13Gori1 (Kaguni and Ray, 1979) double-stranded DNA were used for all of the experiments involving homologous alignment. These share 6407 bp of homology and thus are listed as homologous DNA pairs in the text and figures. In all combinations, however, the duplex substrates contained heterology at the ends or near-nicks, precluding net strand exchange. For experiments 1, 2, 4, 6, 8, and 11, double-stranded DNA was linearized at the *Xho*I restriction site (see "Experimental Procedures") resulting in heterologous ends of 1003 and 1213 bp (Fig. 3*A*). For experiment 9, the double-stranded DNA was nicked specifically at the *Xho*I restriction site (Fig. 3*B*). In the control experiments listed in Table I involving completely heterologous DNA (experiments 3, 5, 7, and 10), pMMC3 (9480 bp) linearized at the *Nco*I site was used (Cox, 1983).

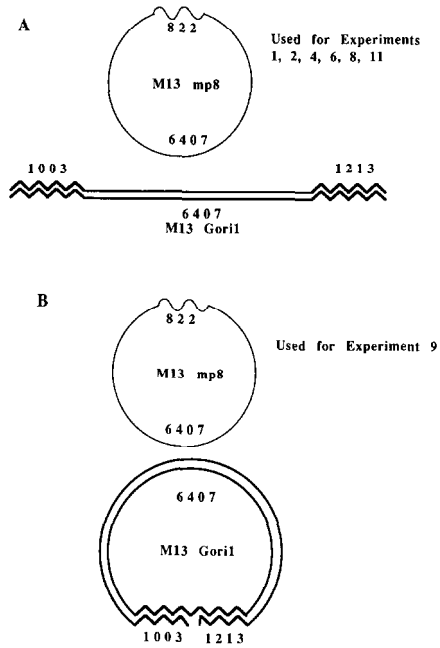


FIG. 3. **Substrate DNA molecules.** Maps of the DNA substrates used for the experiments in this paper (see "Experimental Procedures"). Single-stranded M13mp8 was always circular. Homologous regions are indicated by *smooth lines*, and heterologous regions are indicated by *wavy lines*. Lengths of substrates are indicated by base pairs. M13Gori1 duplex DNA was linearized (*A*) or nicked (*B*) at the *Xho*I site. Lengths of homologous and heterologous DNA segments are shown. Experiments in which substrates were used are also indicated.

The observed efficiency of joint formation under each of the experimental conditions used in this study is presented in Table I. These data reflect only the propensity of molecules to be trapped or cross-linked under a certain set of conditions, neither the length nor the structure of joints is indicated. Each experiment was repeated multiple times. Differences between experiments reflect intrinsic efficiencies of the particular cross-linking reagent rather than variability between trials. For example, the results of three independent trials of experiment 8 varied by less than 5%. When the recA filaments were preserved without fixation, 68% of double-stranded molecules were observed in definable joints (experiment 1). Fixation of protein and DNA with glutaraldehyde also resulted in the retention of a large percentage of homologous joints (experiment 2). When the duplex substrate was entirely heterologous, contacts were not observed (experiment 3), indicating that the observed joints in experiment 2 represented molecules in homologous alignment. When recA protein was removed in the absence of cross-linking (experiments 4 and 5), only 15% of homologous duplex molecules and 5% of heterologous duplex molecules remained in some form of contact with the single-stranded circles. In both cases these contacts were invariably very short. These control experiments confirmed that the substrates chosen do not permit strand exchange and that any paired structures formed were highly unstable in the absence of recA protein.

When fixing agents were added before removing recA protein (experiments 6–10), a very high percentage of the added DNA was trapped in the form of homologous joints. AMT trapped 71% of the DNA substrates shown in Fig. 3*A* in definable joints (experiment 8), indicating that the homologous pairing reaction itself was efficient under the conditions employed. The results also indicate that AMT cross-linked heterologous pairs of molecules at a frequency of 51% (experiment 10). This does not reflect excessive reactivity or con-

Downloaded from www.jbc.org at University of Wisconsin-Madison, on January 26, 2012

TABLE I

Frequencies of alignment and cross-linking

The frequencies of pairing and cross-linking are expressed as a percentage of total double-stranded DNA molecules. Definable joints are those which contain two or three total molecules: either one or two double-stranded and one or two single-stranded DNA molecules. Aggregates contain more total molecules than definable joints.

Substrates and method of fixation	Frequencies of joint formation			No. counted	Figure
	Free	Definable joints	Aggregates		
		%			
1. Homologous form III RecA present during pairing No fixation	13	68	19	100	4
2. Homologous form III RecA present during pairing Glutaraldehyde	0	79	21	100	5
3. Heterologous form III RecA present during pairing Glutaraldehyde	92	8	0	100	
4. Homologous form III RecA present during pairing No fixation/RecA removed	85	15	0	100	
5. Heterologous form III RecA present during pairing No fixation/RecA removed	95	5	0	153	
6. Homologous form III RecA present during pairing Topoisomerase I/RecA removed	49	45	6	114	6
7. Heterologous form III RecA present during pairing Topoisomerase I/RecA removed	86	14	0	100	
8. Homologous form III RecA present during pairing AMT/RecA removed	9	70	20	119	7
9. Homologous form II RecA present during pairing AMT/RecA removed	37	39	24	157	9
10. Heterologous form III RecA present during pairing AMT/RecA removed	41	51	8	140	
11. Homologous form III No RecA AMT	99	1	0	174	

centrations of AMT since cross-linking was dependent on recA protein (experiment 11). Thus it would appear that AMT trapped some contacts made during the "search for homology" carried out by a recA-single-stranded DNA nucleoprotein filament. Although this would seem to represent an unworkably high background, the heterologous and homologous contacts were easily distinguishable: cross-links formed between heterologous DNA molecules were limited to one or two contacts per pair and were always too short to measure (*i.e.* ≤40 bp). Thus these contacts were fundamentally different than the extended AMT-linked homologous contacts described below.

In order to discuss the complex structures observed in this study, it is necessary to define a set of descriptive terms. These were chosen for utility, and assumptions implied by their use are described. In all of the experiments in this study, including those in which recA protein was still present at the observation stage, there were often multiple homologous contacts between two aligned molecules separated by DNA which was apparently not involved in homologous contact. Paired segments with uniform cross-linking patterns are generally referred to as "contiguous regions of homologous contact." The phrase "total aligned DNA" includes the length of all DNA from the point of the first homologous contact to the last, including regions of unpaired DNA.

It is also useful to identify the strands involved within a particular joint, where possible. The initiating, circular, single-stranded DNA molecule is referred to as strand S+. The identical strand in the duplex DNA is labeled D+, and the complementary strand of the duplex DNA is labeled D−. Regions where all three strands are cross-linked are referred to as triple-stranded or TS. Regions where the original strands of the duplex are cross-linked are referred to as D+D−. Double-stranded regions where S+ and D− are cross-linked are referred to as S+D−. It should be noted that strand identification relies on the assumption that when one strand of the duplex is cross-linked extensively to the single-stranded circle, that it is the D− or complementary strand. Double-stranded regions which could be either D+D− or S+D− are simply referred to as ambiguous and designated A. Regions where all three strands are single-stranded are labeled only on strand D−. Some homologous contacts were too short to be characterized reliably. These are referred to as point contacts and given the designation P. When they were present in a joint with other contacts, point contacts were identified as homologous if they were separated from another definable contact by equal distances on two or more strands.

Protein-coated Joints

To provide a structural point of reference, joints were examined prior to removal of recA protein, both with and without glutaraldehyde fixation. Experiment numbers refer to Table I.

Downloaded from www.jbc.org at University of Wisconsin-Madison, on January 26, 2012

Experiment 1, No Fixation—To minimize disruption of the recA filament during spreading, native samples were prepared for microscopy using a different protocol than that used for the other experiments in this study (see "Experimental Procedures"). The critical factor in this protocol seems to be the exposure of the samples to 20% glycerol prior to and during drying. Examples of homologous joint molecules prepared without fixation are shown in Fig. 4, *A* and *B*. In Fig. 4*C* is shown an example of a recA-single-stranded DNA nucleoprotein filament prepared without a second interacting DNA molecule. This protocol was sufficiently gentle that the repeating helical structure observed previously within recA filaments (Dunn *et al.*, 1982; Egelman and Stasiak, 1986; Flory *et al.*, 1984; Williams and Spengler, 1986) was preserved in some regions of the recA filaments, even though no ATPγS was present.

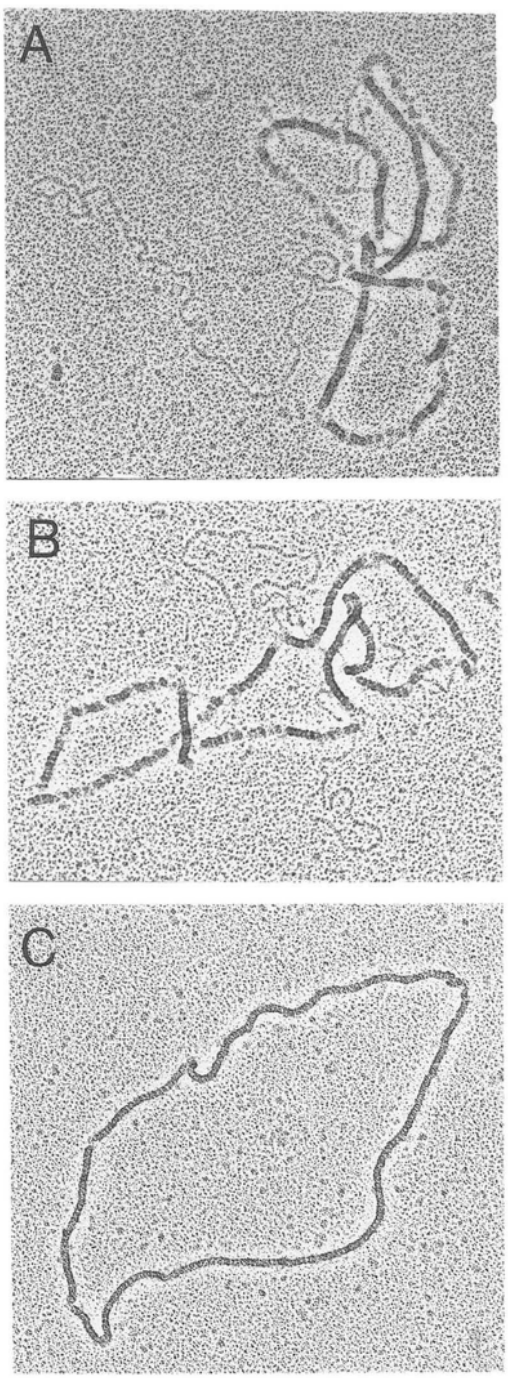


FIG. 4. **Experiment 1: homologous joints visualized without fixation.** *A* and *B*, joint molecules formed between circular single-stranded M13mp8, and *C*, linear duplex M13Gori1, and recA filaments on M13mp8 in the absence of a duplex DNA molecule were prepared without fixation in the presence of 3 mM ATP (see "Experimental Procedures").

With this protocol, joints were observed with both homologous and completely heterologous duplex DNA molecules. In the samples containing heterologous duplex DNA a large fraction of the filaments were present in aggregates that precluded efforts at quantitation (data not shown). This is consistent with the coaggregation of DNA molecules observed by Gonda and Radding (1986) under similar conditions. Samples with homologous duplex DNA, however, exhibited very little aggregation. In these samples 68% of the duplex DNA molecules were present in definable joints. Most homologous joints were quite complex, with the duplex DNA disappearing into and emerging from the filament several times. The filaments in these joints were fragmented slightly and somewhat supertwisted by comparison to filaments formed with heterologous DNA or in the absence of duplex DNA, suggesting that the formation of homologous joints induces torsional stress. The complexity of the homologous joints and the presence of protein generally precluded length measurements.

Experiments 2 and 3. Glutaraldehyde Fixation—To eliminate heterologous contacts as much as possible in the presence of recA filaments, the approach of Register *et al.* (1987) was used in which joints are trapped by GA and samples spread under conditions which tend to denature unfixed protein (Fig. 5). There were slight differences in technique from the published protocol (see "Experimental Procedures"); notably no attempts were made to separate joint molecules from free protein; glutaraldehyde alone was used; and both the concentration of cross-linking agent and time of cross-linking were reduced substantially. The fixation conditions chosen did not preserve the recA filament completely but tended to fix regions of DNA-DNA interactions. A total of 29 double-stranded molecules involved in joints were traced and compared to an equivalent number of unpaired molecules added between fixation and spreading. The distribution of total lengths of the molecules containing joints was not substantially different from that of free molecules, indicating that the correct path for tracing was chosen and that under these spreading conditions the cross-linked regions were not extended relative to unbound DNA. As shown in Table I, 79% of the duplex DNA molecules in this experiment were found in definable joints and an additional 21% in aggregates. In the heterologous control, only 8% of duplex molecules were in joints and these were all point contacts.

Two typical joints are shown in Fig. 5, *A* and *B*. As observed previously (Register *et al.*, 1987), the double-stranded DNA simply disappeared into the recA filament in the homologous region and then reemerged; the structure of the recA filament in joined regions was indistinguishable from that of unjoined regions in these samples. As seen in Fig. 5*C*, regions of contiguous contact were fairly long, averaging 370 bp and stretching up to 1700 bp in length. However, 77% of joined molecules contained multiple regions of contact separated by loops of DNA not in contact with the filament. The length of total aligned DNA averaged 2160 bp and extended up to 5360 of the possible 6407 bp, or 84% of the available homology. We cannot rule out the possibility that, prior to the preparation for microscopy, the unpaired DNA loops were in contact with the single-stranded DNA in the filament, or the filament itself, and were excluded because of insufficient cross-linking. However, these loops were observed frequently in all three trials of this experiment.

Downloaded from www.jbc.org at University of Wisconsin-Madison, on January 26, 2012

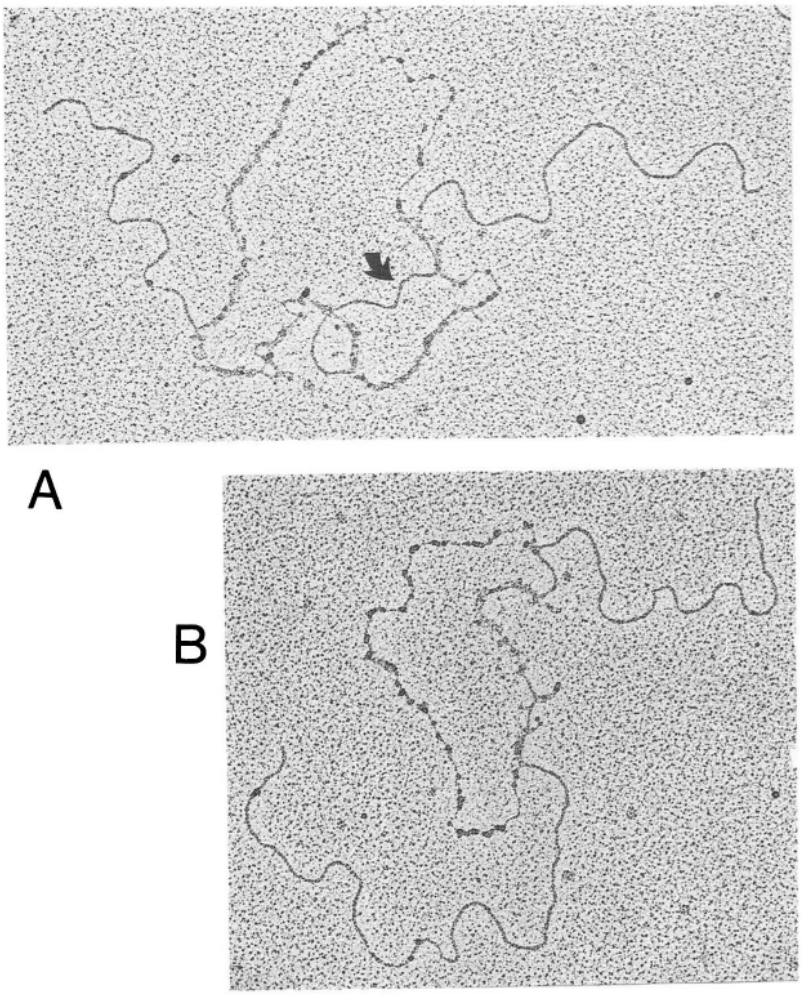


FIG. 5. **Experiment 2: joints fixed with glutaraldehyde.** *A* and *B*, examples of glutaraldehyde-fixed joints formed between circular single-stranded M13mp8 and linear duplex M13Gori1. A loop of unpaired DNA between paired regions is indicated by the *arrow* in *A*. *C*, length of joined regions. The length of contiguous regions of homologous contact and total aligned DNA, expressed in base pairs, are shown in histogram form. The *y* axis indicates the number of samples within units of 100 bp.

Joints Observed after DNA Fixation Followed by Removal of RecA Protein

In the experiments presented in this section, DNA-specific cross-linking reagents were used to preserve some of the features of the DNA structure of paranemic joints after removal of recA protein.

Experiments 6 and 7. Topoisomerase I Trapping—These experiments relied on the ability of topoisomerase I to nick a DNA strand, allow another strand to pass through, and then reclose. The resulting structure resembles a hemicatenane, but does not represent a true topological link because the S+ strand is trapped by noncovalent base-pairing in the heterologous ends. The protocol used here resembles that used by Radding and colleagues (Bianchi *et al.*, 1983; Cunningham *et al.*, 1981) to demonstrate recA-mediated homologous alignment away from ends. Here, efforts focused on the fine structure of the aligned regions. Observed pairing efficiency is given in Table I.

An example of joints fixed in this way is shown in Fig. 6*A*, along with an interpretive drawing. Structures extending between the first and last homologous contact were generally not uniform. Regions were observed in which all three strands coincided (TS). Although the TS regions are short, their assignment is rendered unambiguous by the junctions of a single- and double-stranded DNA observed at both ends of the joint. Also seen were point contacts (P), and regions where the identity of strands in double-stranded DNA was ambiguous (A). These structures are labeled on the interpretive diagram in Fig. 6*A*. This example contains two single-stranded circles linked with a single double-stranded DNA molecule. The circle labeled S+1 makes only a point contact with the double-stranded molecule, and S+2 has more extensive contacts. Other regions where the S+ and D− strands coincided and regions where all three strands were single-stranded were observed in the experiment but are not found in this particular example. The lengths of the different types of contact observed are shown in histogram form in Fig. 6*B*. Length measurements were corrected for the fact that single-stranded DNA is 2.2% longer than double-stranded DNA under these spreading conditions. Triple-stranded regions were twice as common as unambiguous S+D− regions but were roughly one third as long. The S+D− regions may be more common than indicated because there are many regions where the identity of the duplex DNA is ambiguous. The triple-stranded regions averaged 50 bp with a maximum of 110 bp. The unambiguous S+D− region lengths averaged 200 bp with a maximum of 370. As with the glutaraldehyde cross-linked joints, loops of ambiguous (A) or D+D− duplex DNA separated multiple homologous contacts. The total lengths over which DNA molecules were aligned in single joints averaged 1050 bp with a maximum of 4370 bp.

AMT Cross-linking Experiments—Several psoralen derivatives were tried before AMT was chosen as the most efficient cross-linking agent for paranemic joints. Like all psoralens, AMT is an intercalating molecule which reacts with pyrimidine bases (predominately thymines) in the presence of UV light to produce cross-strand linkages (for review, see Cimino *et al.*, 1985). The protocol used (see "Experimental Procedures") reflects the minimum concentration of AMT and the

Downloaded from www.jbc.org at University of Wisconsin-Madison, on January 26, 2012

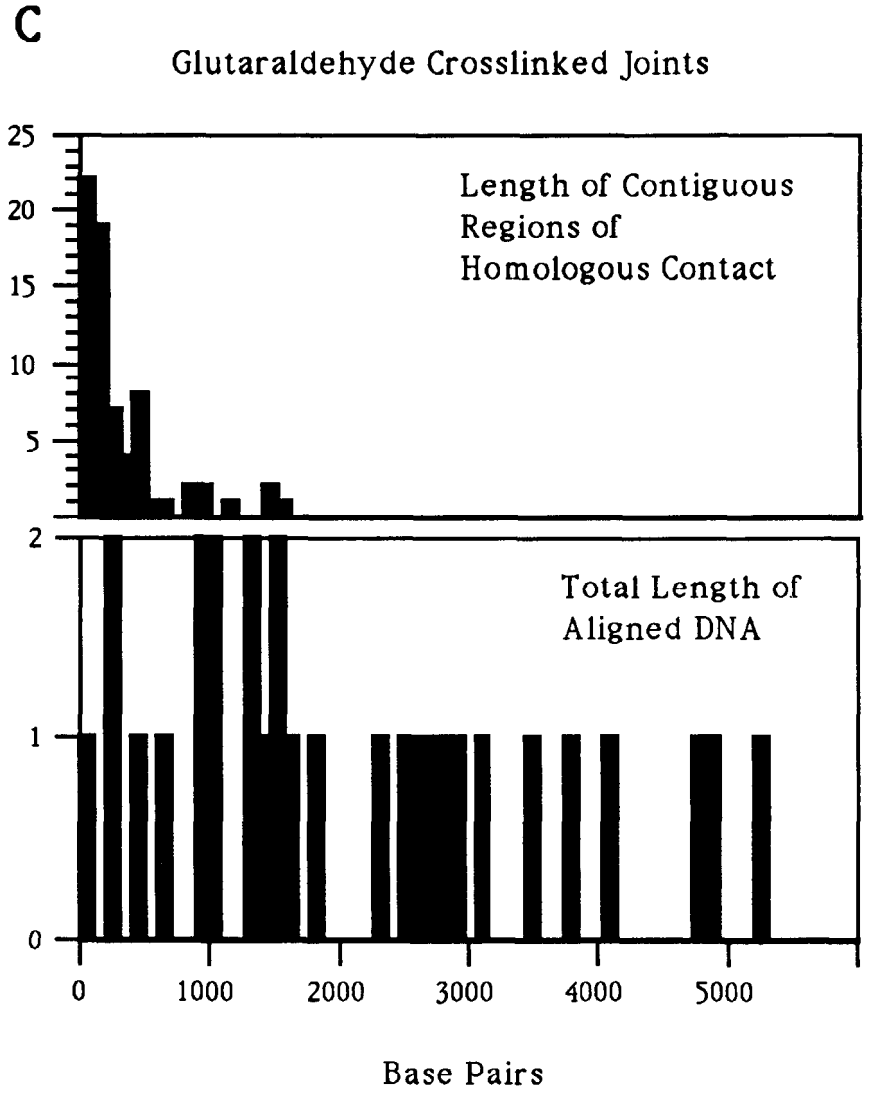


FIG. 5—*continued*

shortest incubation times required to cross-link unbound duplex DNA to a high extent. When unbound duplex DNA was subjected to these conditions and completely denatured to measure the cross-linking density, an average of at least 77 cross-links per molecule was found. This corresponds to at least one cross-link in every 112 bp. The efficiency of AMT cross-linking was slightly higher on recA-bound duplex DNA than on unbound duplex (data not shown). This set of experiments relies on the assumption that AMT cross-links only DNA strands. Studies of the reactivity of psoralen derivatives with proteins has revealed that these reagents do not attach covalently to protein, however they do produce oxygen radicals which may inactivate enzymes (Granger *et al.*, 1982; Veronese *et al.*, 1982). To test this inactivation of recA protein by AMT treatment, we checked the ability of recA filaments to hydrolyze ATP after the AMT cross-linking protocol described under "Experimental Procedures." These results indicated that ≈75% of ATPase activity remained (data not shown). Attempts were also made to observe recA filaments directly after AMT fixation (before deproteinization). However, in both AMT and samples incubated on ice for 5 min without AMT, recA filaments could not be preserved as in Fig. 4 (data not shown). Although it is not clear that the techniques used to visualize samples without fixation (Fig. 4) preserves recA filaments under all conditions, this may indicate some disruption of the filaments as a result of the reduction in temperature.

Experiment 8, Joints Formed between Single-stranded Circles and Linear Duplex DNA—Examples of joints formed between single-stranded circles and linear duplex DNA after cross-linking with AMT and removal of recA protein are shown in Fig. 7, *A–C*. These examples were chosen to illustrate certain common structural features, but are representative of the data pool. Pairing efficiencies are noted in Table I. When homologous DNA molecules were paired, over 90% of the DNA was present in either definable joints or aggregates. These joints were much longer than joints formed in the heterologous controls, and the individual strands could be identified and traced unambiguously in about half of the structures. Regions in which all three strands coincided could again be assigned unambiguously since they were all bounded by junctions between single- and double-stranded DNA.

The lengths, positions, and (where possible) strand identity of the various structural segments were determined for 24 joint molecules selected at random and photographed from a single grid. The grid chosen was qualitatively identical to grids prepared in eight separate trials. As seen in the molecules in Fig. 7, *A–C*, and the histograms in Fig. 7*D*, regions were again observed in which all three strands coincided (TS) and others where the S+D− strands were cross-linked. Both types of contact were longer than observed with topoisomerase I. This is probably due to the higher efficiency of AMT cross-linking. Contiguous regions of triple-stranded contact averaged 120 bp and ranged up to a maximum of 510 bp. Regions

Downloaded from www.jbc.org at University of Wisconsin-Madison, on January 26, 2012

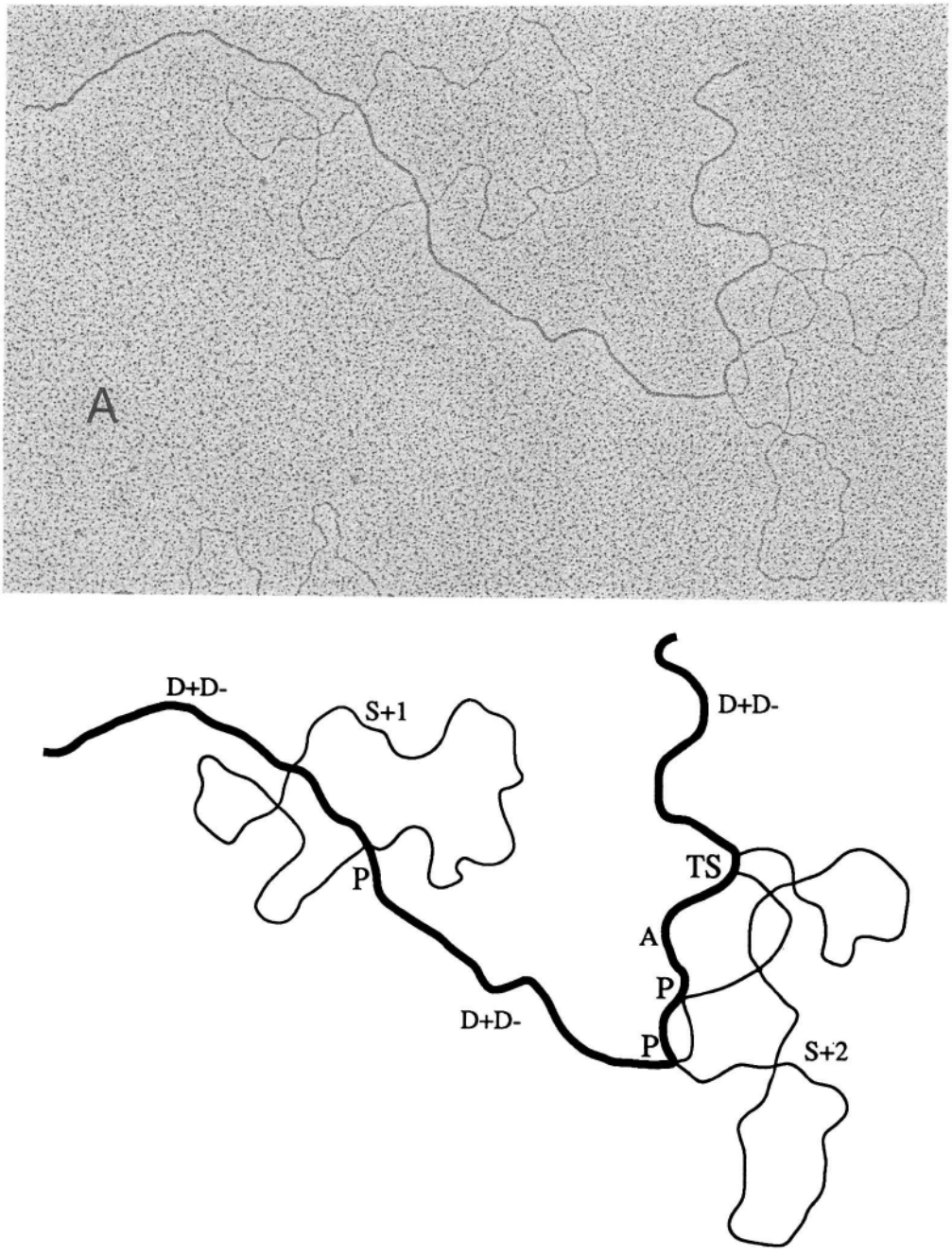


FIG. 6. **Experiment 6: regions of contact trapped by topoisomerase I.** *A*, examples of aligned joints formed between two circular single-stranded M13mp8 molecules and linear duplex M13Gori1, trapped by topoisomerase I. In the interpretative drawings, triple-stranded regions are indicated by TS, regions where the original duplex is unperturbed and indicated by D+D−, ambiguous double-stranded DNA is indicated by *A* and point contacts are indicated with a *P*. *B*, lengths of observed structures. The length of contiguous regions of homologous contact either TS or S+D− and total aligned DNA, is indicated in histogram form. The *y* axis represents the number of samples of each type within units of 200 bp.

of S+D− cross-linking averaged 310 bp with a maximum of 1250 bp. If samples were denatured by heating in the presence of formaldehyde prior to spreading, joints retained both the TS and S+D− structures in proportions and lengths similar to those observed in native joints, indicating that both forms of contact were indeed covalently cross-linked by AMT (data not shown). Both the TS and S+D− structures were common, often occurring several times in a single joint separated by loops of D+D− or ambiguous double-stranded DNA. Short TS lengths were found in 71% of the joint molecules. The total length of DNA aligned in joints averaged 2460 bp with a maximum of 5880 bp. These numbers correspond closely to the numbers obtained for recA protein-bound joints cross-linked with glutaraldehyde. In some cases the D+ and D− strands were both single-stranded near contacts with S+. The average length of unlinked D− strand in these regions was 150 bases, and the maximum observed length was 720 bases.

A representative sample of the data collected from 12 of the 24 joints is presented in a graphic form in Fig. 8. It is evident that there was a great diversity in the order and position of the structural elements within the joints. There was also diversity, again, in the total lengths over which the DNA molecules were aligned. In all cases, cross-linked contacts were restricted to regions of the duplex DNA that

Downloaded from www.jbc.org at University of Wisconsin-Madison, on January 26, 2012

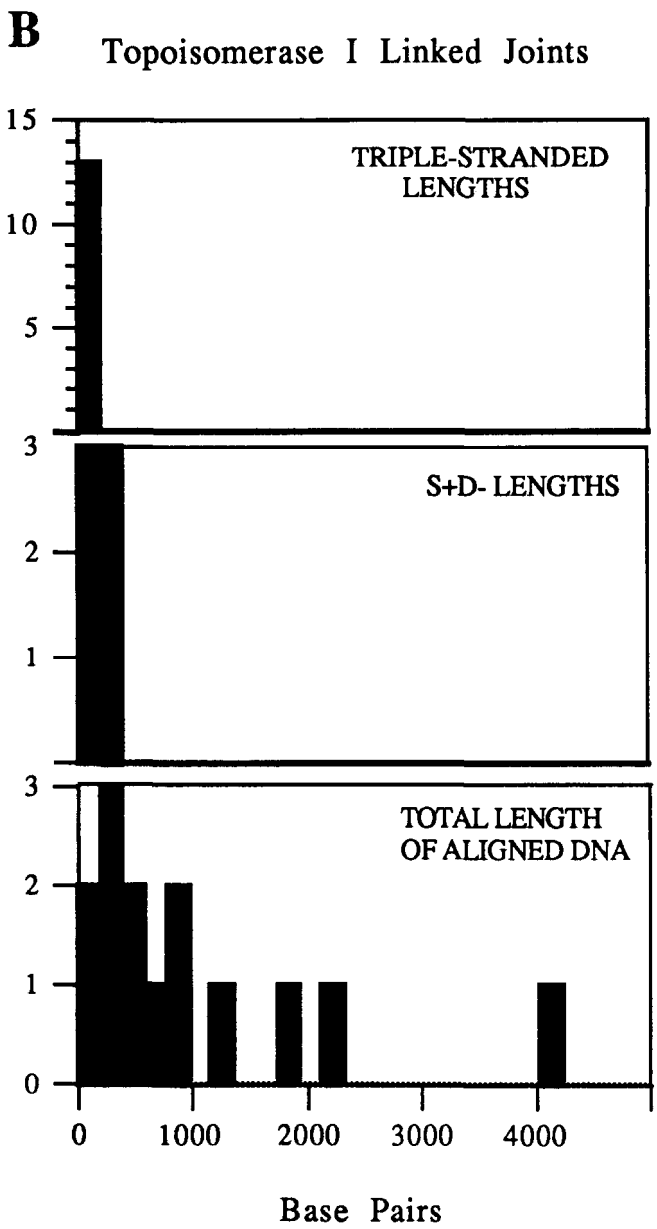


FIG. 6—*continued*

contained sequences homologous to the single-stranded DNA.

As can be seen in Fig. 7*A*, regions of S+D− contact can alternate frequently with D+D− regions. Although this example is particularly simple, in general, the S+D− segments observed in this experiment did not contain a large number of point contacts or nodes between the displaced D+ strand and the S+D− region. The possible implications of this result will be considered in more detail under "Discussion."

In all of the previous experiments, regions of double-stranded DNA (presumably D+D−) were observed separating regions of homologous contact suggesting that this DNA was not involved in homologous interactions. This conclusion is supported by the joint presented in Fig. 7*C*. In this micrograph, two single-stranded circles are aligned with one double-stranded molecule. The circle above, labeled S+1, is aligned over a long region containing both TS and S+D− contacts. The second circle, labeled S+2, has a point contact with the double-stranded molecule in a region of cross-linked double-stranded DNA that is probably D+D−. This indicates that a D+D− region internal to the joint was accessible to interactions with other DNA molecules at the time of cross-linking. This configuration was observed in 2 of 24 randomly chosen joint molecules, and additional examples were located in a less quantitative scan of these grids.

Experiment 9: Joints Formed between Circular Single Strands and Nicked Circular Duplex DNA—In order to test the influence of topological restrictions on the structure of these joints, cross-linked paranemic joints formed between double- and single-stranded circles were examined. Examples of these molecules are shown in Fig. 9. All types of structure described above were again seen but the distribution was skewed very heavily in favor of point contacts at the expense of TS regions and S+D− regions. Out of 63 joints scored, 48 contained only point contacts, 10 contained short TS regions, and 5 contained identifiable S+D− regions. Both the length of contiguous contacts and the length of total aligned DNA was greatly reduced for paranemic joints formed with these topological constraints. In the small number of measurable joints, no contiguous TS or S+D− region of contact was longer than 70 bp. The longest stretch of total alignment was 880 bp with the average only 680 bp. The examples shown in Fig. 9 represent some of the longest joints examined in this experiment. The occurrence of strand cross-overs at nonhomologous points was also much higher in these samples than in those formed with linear duplex DNA. These nonhomologous contacts are consistent with the need for topological nodes to compensate for regions of alignment of a single handedness. This result suggests that the alternating structural segments observed in the joints formed with linear duplex DNA molecules (described above) do not represent an alternation of structures which provide topological compensation for each other.

DISCUSSION

This work provides evidence for the formation of an interwound three-stranded DNA structure as part of the mechanism of homologous DNA pairing promoted by recA protein. A triple helix was proposed as a strand exchange intermediate by Howard-Flanders *et al.* (1984) and DNA structures formed during recA-mediated reactions that appear to contain three strands have been observed experimentally (Register *et al.*, 1987; Stasiak *et al.*, 1984). This study extends these observations by attempting to define some of the properties of paranemic joints. Using the covalent DNA cross-linking agent AMT we were able to trap and visualize the length and cross-linked structure of three-stranded joints after removal of recA protein. The most fundamental observation from this work comes from the comparison of pairing reactions carried out with a circular single-stranded DNA and either a linear or circular duplex substrate. Linear duplex substrates form joints of 2460 bp average total cross-linked length, but circular substrates form joints with an average cross-linked length of only 190 bp. The simplest explanation of this difference is that a net helical interwinding must occur to form long joints. The observation that joints formed with linear duplexes are longer than those formed with circular duplexes is consistent with data reported previously (Honigberg *et al.*, 1985), although the differences observed in the present study are particularly dramatic.

Building on this evidence of helical interwinding, the cross-linking data indicates that all three strands are positioned very near one another over at least short segments of this structure. This is consistent with the formation of a hydrogen-bonded triple helix. The most direct demonstration of this are regions in which all three strands coincide and are cross-linked for stretches that average 120 bp and can be up to 510 bp in length. Any helix formed in this system will almost

Downloaded from www.jbc.org at University of Wisconsin-Madison, on January 26, 2012

FIG. 7. **Experiment 8: joints linked by AMT cross-linking.** *A–C*, examples of joint molecules formed between circular single-stranded M13mp8 and linear duplex M13Gori1 and cross-linked by AMT. Labels in the interpretive drawings are explained in the text and legend to Fig. 6. *D*, lengths of observed structures. The length of individual substructures: TS, S+D−, and D−; and total aligned DNA is indicated in histogram form. The *y* axis represents the number of samples of each type within units of 100 bp.

Downloaded from www.jbc.org at University of Wisconsin-Madison, on January 26, 2012

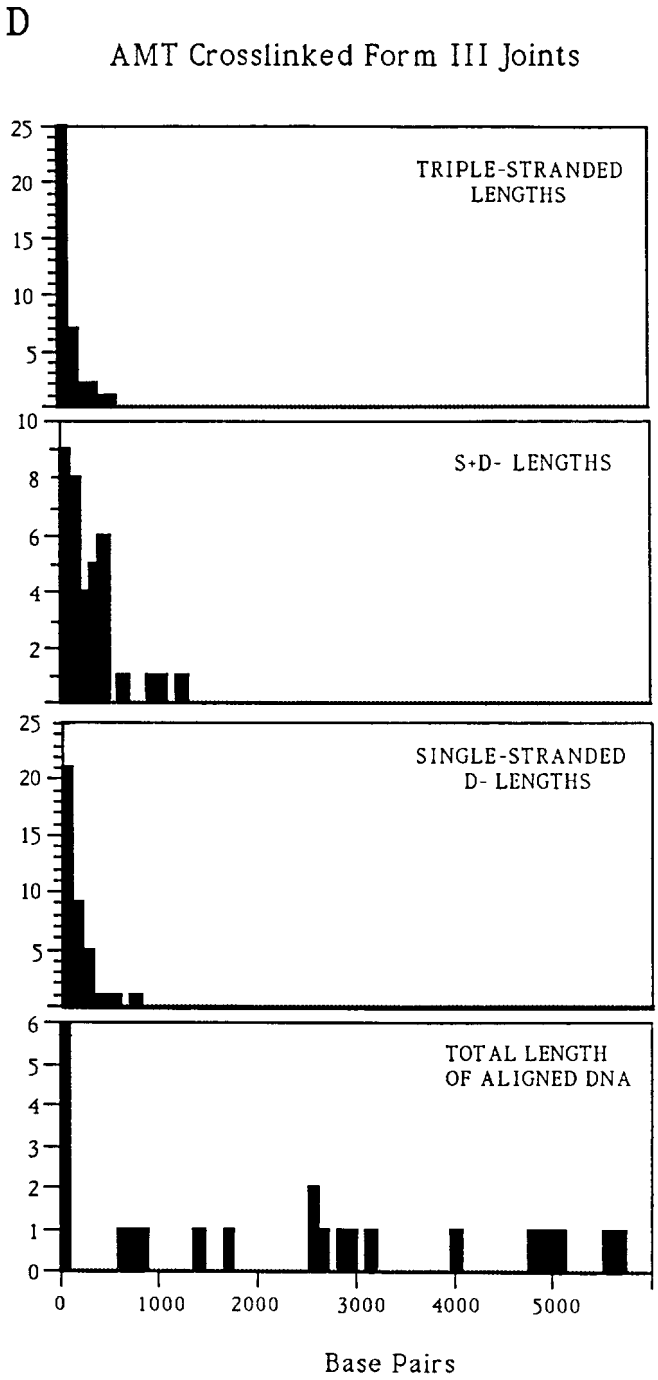


FIG. 7—*continued*

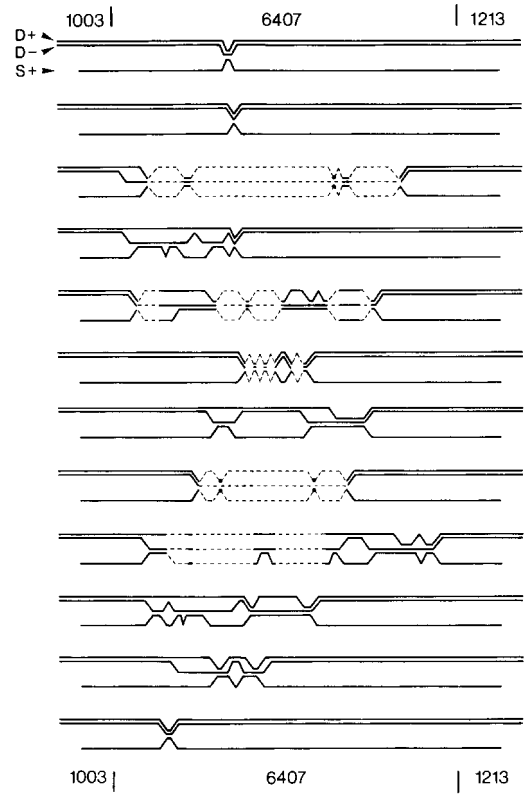


FIG. 8. **Schematic diagram of length and type of aligned structures formed between circular single-stranded M13mp8 and linear duplex M13Gori1.** The single-stranded substrate is shown as a linear molecule for the sake of simplicity, and it is positioned so that homologous sequences in the two DNA molecules are aligned horizontally. These molecules contain 6407 bp of homology, and the duplex substrates have 1003 and 1213 bp of heterology at their ends. Cross-linking is denoted by the close juxtaposition of two or all three lines. Ambiguous regions (where two of three strands are cross-linked but strand identities cannot be determined) are indicated by *dashed lines*. Angles of diagonal lines are constant and chosen arbitrarily. The horizontal distances correspond to lengths of the substructures indicated.

certainly be right-handed because recA protein always binds DNA in a right-handed helical filament, even when the DNA is in Z form prior to binding (Kim *et al.*, 1989). These regions are unambiguously three-stranded because they are defined by junctions between single- and double-stranded DNA molecules. This work thus lends support to the idea that recA protein promotes the formation of right-handed, triple-helical, homologous joints.

The same data generally argue against the idea that helical winding is balanced in any way by a compensating structure formed within the joint. Where topological restrictions exist, joints size is simply limited. This observation argues against those models for recA-stabilized paranemic joints in which no helical winding is introduced (model A in Fig. 1), or in which helical winding is balanced within the nucleoprotein filament either by alternating right- and left-handed helical pairing domains (model F), or by left-handed toroidal winding of the two DNA molecules outside of right-handed helical pairing domains (model E). This result suggests that, in the cell, any topological problems brought about by interwinding two DNA molecules are solved by means external to the recA system. One caveat here is that some disruption of the filaments may occur during cross-linking, as discussed further below. Interwinding of the DNA molecules in the experiment with linear duplex may selectively stabilize this DNA-DNA interaction and artificially enhance the difference in the lengths of cross-linked joints in the linear *versus* circular duplex comparison. This possibility, however, does not affect the conclusion that the long joints formed with linear duplex involve interwinding of the two DNA molecules.

The cross-linking patterns observed here are complex and can be interpreted in two ways: either a single-paired structure is formed and slight variations in cross-linking efficiency or structural changes in the recA filament during cross-linking bring about the multiple structures; or the variations represent multiple classes of substructures that exist within long joints. With caveats outlined below, we believe the data overall are more consistent with the latter intepretation. As noted under "Results," there are four interesting patterns seen within the overall structure of the cross-linked joints. The first pattern is the region where all three strands coincide. In these regions there are sufficient cross-links between all three

Downloaded from www.jbc.org at University of Wisconsin-Madison, on January 26, 2012

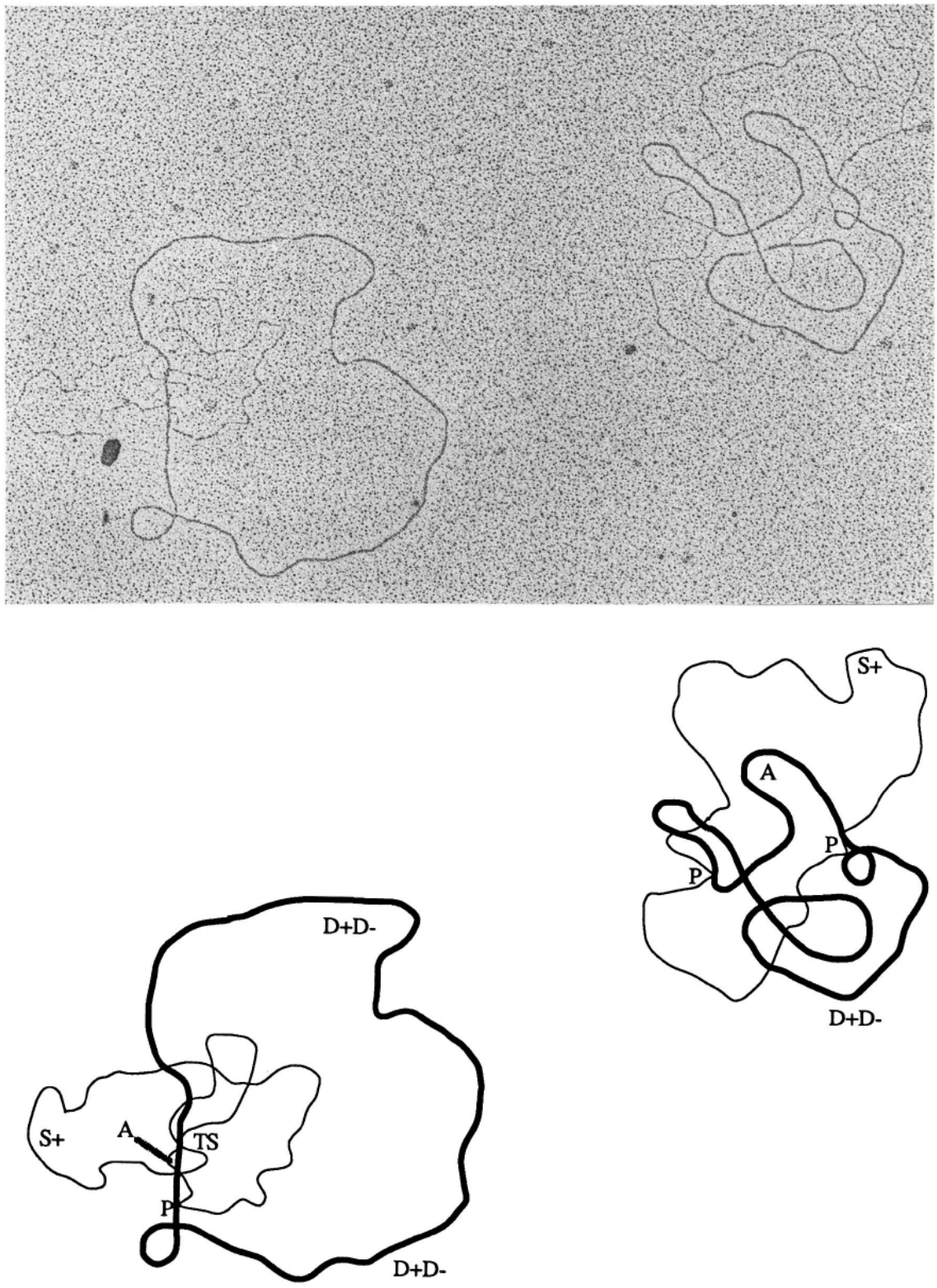


FIG. 9. **Experiment 9: examples of aligned molecules formed between circular single-stranded M13mp8 and circular duplex M13Gori1 and cross-linked by AMT.** Labels in the interpretive drawings are explained in the text and legend to Fig. 6.

strands so that the regions are not disrupted even under denaturing conditions. The second pattern is that in which the S+ strand of the single-stranded circle is cross-linked to one of the strands of the duplex while the other strand is looped out. This structure is also most likely a manifestation of triple helix, either the same structure as the TS regions or a distinct structure in which a local strand switch has occurred that pairs the D− strand with the S+ strand so that these are cross-linked and the D+ strand is not. The third pattern is the presence of regions where none of the strands are cross-linked and any interwinding that may be present is eliminated after recA protein removal, possibly as a result of torsional stress. The fourth pattern is the D+D− cross-linked regions. These are interpreted as loops where the D+D− duplex is not

Downloaded from www.jbc.org at University of Wisconsin-Madison, on January 26, 2012

in contact with S+ and these segments separate points of homologous contact. We interpret these loops as an indication that homologous contacts promoted by recA protein can be discontinuous.

The interpretation of these patterns is subject to a number of caveats. The most important of these concerns the cross-linking protocol itself. AMT has not been observed to cross-link protein (Wollenzien *et al.*, 1979), but the low temperature used and the O_2 radicals generated during cross-linking appear to produce some inactivation and/or disruption of the recA filaments. This could obviously affect the observed cross-linking patterns in unpredictable ways. We therefore cannot eliminate the possibility that the nonuniformity of these patterns is produced in whole or in part as an uncharacterized artifact of the method itself. Other results in this study, however, tend to support the conclusion that the joint structure is variable. The patterns of strand contact observed with AMT and topoisomerase I were qualitatively identical, although the latter joints were shorter. As mentioned under "Results," the lengths of joints obtained in the glutaraldehyde and AMT cross-linking experiments were nearly identical. All of the experiments in this study, including those in which there is no cross-linking, indicate the presence of loops of the incoming duplex DNA. These loops are internal to many joints and are not in homologous contact with the single-stranded circle. The assignment of the D+D− structures as side-by-side regions lacking homologous contacts is based not only on the correlation with the duplex loops seen within joints fixed with glutaraldehyde, but also on the occasional observation of a second molecule of single-stranded DNA interacting with D+D− segments internal to a joint (Fig. 7*C*). The D+D− regions that we have observed may simply represent interdomainal unbound loops of duplex DNA that separate the triple helices. Alternatively, they may be bound to putative secondary DNA binding sites on the exterior of the recA nucleoprotein filament (Cox, 1990). The simplest overall interpretation is that these joints represent discontinuous triple helices, best illustrated by model D in Fig. 1.

This proposed interpretation of the cross-linking patterns is illustrated schematically in Fig. 10. In this figure two independent regions of triple helix are formed, separated by a side-by-side region without homologous contact. Each of the triple helices are shown to contain two distinct segments that give rise to the TS and S+D− cross-linking patterns. The side-by-side region yields a D+D− cross-linking pattern. Once recA protein is removed, the triple helices must unravel as shown in order to generate the observed separation of the D− strand in the S+D− regions.

This interpretation of the observed structures is predicated on several additional assumptions. Our controls confirm that DNA is cross-linked with an efficiency more than sufficient to explain the observed structures, and uniform structures exceeding 100 bp generally contained multiple cross-links. This efficiency may not be great enough to trap all triple helical structures, however, with the possible result that structure lengths may be underestimated. Another problem is that AMT cross-links strands only when suitable pyrimidine bases are in a rather precise spatial arrangement. Strands in close proximity will not always be cross-linked and our interpretation of the structures generally takes this caveat into account. These properties of AMT have the advantage of placing a useful constraint on attempts to build structural models for the triple-helix evidently formed here. Our attempts to trace DNA strands rest on the stated assumption that when S+ is cross-linked to one strand of the duplex, it is the complementary D− strand. This is based on the simple fact that strands

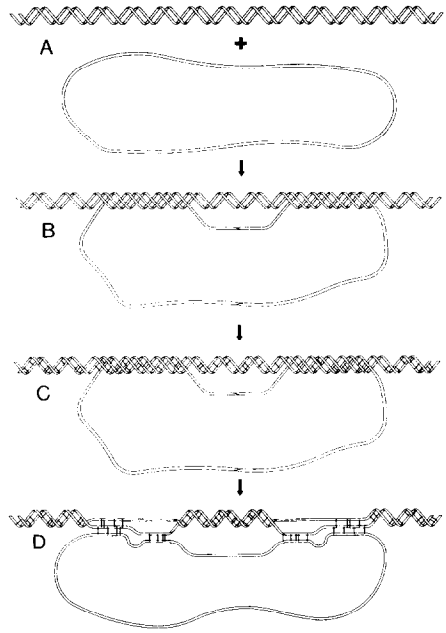


FIG. 10. **Schematic representation of AMT cross-linked structures.** DNA molecules are shown without recA protein for clarity. *A*, substrates prior to alignment. *B*, substrates have formed two regions of homologous alignment separated by a side-by-side region. *C*, AMT covalently cross-links double- and triple-stranded regions. *D*, recA protein is removed and strands unravel to the extent permitted by cross-links.

paired in the Watson-Crick conformation are cross-linked efficiently (Cimino *et al.*, 1985), but other possible cross-linking patterns are speculative. Even if some or all of the S+D− structures actually represent S+D+ cross-links, however, the general argument that these represent a form of triple-helix is still valid. The homologous joints observed here are highly unstable; they fall apart rapidly and completely when recA protein is removed unless they are fixed. Therefore the structures observed in Figs. 6, 7, and 9 presumably represent the remnants of joints that have unraveled to the extent that DNA fixation will permit. All three strands coincide on the grid only where all three are fixed in some way. Unraveling of joints after DNA fixation and removal of recA protein could also result in an underestimation of the amount of triple-helical structures. It is clearly important to find additional probes that will permit further analysis of these unstable and possibly dynamic structures.

These considerations are especially important in the case of the S+D− structures. If our assignment of the D− strand is correct, we can interpret the joint segments in which the S+ and D− strands are cross-linked by AMT to be a form of triple helix in which the D+ strand is in a conformation that makes it refractory to AMT cross-linking. This may be the same structure as the TS regions, or a second distinct form of the triple helix. This interpretation is based on the topological argument that any interwinding of D− about S+ must involve an equivalent degree of interwinding of D+ about S+. This is because D− and D+ are joined inseparably by base pairing in the long heterologous ends in the linear duplex substrates. The S+D− structure is potentially very interesting in that it suggests that local strand separation of D+ and D− occurs within these joints accompanied by a strand switch so that D− and S+ are paired, possibly in a Watson-Crick mode, and cross-linked. From the preceding arguments, it follows that the only way D+ could be displaced without cross-over nodes (as we have observed by electron microscopy), is if the

Downloaded from www.jbc.org at University of Wisconsin-Madison, on January 26, 2012

cross-linked S+D− structure observed on the grid has lost most or all of its helical pitch as a result of unraveling of the triple helix that occurred after recA protein was removed. This suggests that a helical structure exists under significant torsional stress that is maintained only in the recA filament. If the S+D− structures indeed represent a local strand switch, this suggests a mechanism for the partial strand exchange reactions observed by Menetski *et al.* (1990), *i.e.* the fundamental structure in which DNA molecules are paired by recA protein is one in which strand rearrangement is facilitated coincident with homologous pairing. Both results suggest that the strand switch *per se* is not coupled to ATP hydrolysis, but instead is a direct byproduct of the formation of this triple-helical paired intermediate.

The complexity of the observed cross-linking patterns and the cautionary notes inserted above point to a general need to apply additional methods to this problem. Two sensitive but somewhat indirect solution methods have been used previously to estimate the length of paranemic joints. Changes in ATP hydrolysis (Schutte and Cox, 1987) indicate that all available homology (6407 bp here) is sampled when a circular single-stranded DNA molecule is paired with a linear duplex. Linear and circular duplex substrates were compared side-by-side in this study; the results tend to support the idea that the joints formed with linear duplex substrates are longer. Underwinding of the duplex (Schutte and Cox, 1988) indicates that about 30% of the homology (1900 bp) on average is aligned when the duplex substrate is circular. Both sets of results also suggest that the AMT cross-linking protocol used here underestimates joint size and refinements may provide useful additional information.

Our conclusions can be combined in a comprehensive model for recA protein-promoted homologous alignment of DNA, that, in most respects, follows established lines of thought in this system. The fundamental form of homologous contact between a single-stranded DNA molecule within a recA filament and a second, double-stranded DNA molecule is a right-handed triple-helix formed in the helical groove of the nucleoprotein filament. Formation of this structure induces global interwinding whenever topology is not restrictive. The homologous interaction is mediated by an as yet undetermined pattern of hydrogen bonding probably involving the bases of all three DNA strands in the major groove of the duplex. More than one hydrogen bonding pattern may exist. This triple helix has 18 base triplets per turn and an axial base separation of 5.1 Å. Like strands (S+ and D+) are parallel, and the structure is nonspecific with respect to DNA sequence. This structure is probably quite dynamic,[2] with paired regions migrating in the manner observed by Register *et al.* (1987). The triple helix may migrate, perhaps unidirectionally, in a wave coupled to ATP hydrolysis. Within this triple helix, the strands can evolve into another form of contact in which the D+ and D− strands are separated, and D− is paired with S+. In the case where such a strand switch occurs at a homologous end, this structure represents the initiation of a strand exchange reaction. Where such an end is not available, this process might still result in strand exchange *in vivo* with the aid of topoisomerases, ultimately leading to a genetic recombination event. This scenario can be readily accommodated in the context of current models for recA-mediated DNA strand exchange (Cox, 1990; Howard-Flanders *et al.*, 1984).

[2] S. Umlauf, unpublished observations.

Acknowledgments—We wish to thank Maria Schnös and David Inman for technical assistance.

REFERENCES

Bianchi, M., DasGupta, C., and Radding, C. M. (1983) *Cell* **34,** 931–939
Cassuto, E., West, S. C., Mursalim, J., Conlon, S., and Howard-Flanders, P., F. P. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77,** 3962–3966
Chow, S. A., Honigberg, S. M., Bainton, R. J., and Radding, C. M. (1986) *J. Biol. Chem.* **261,** 6961–6971
Cimino, G. D., Gamper, H. B., Isaacs, S. T., and Hearst, J. E. (1985) *Annu. Rev. Biochem.* **54,** 1151–1193
Cox, M. M. (1983) *Proc. Acad. Natl. Sci. U. S. A.* **80,** 4223–4227
Cox, M. M. (1990) in *The Biology of Nonspecific DNA-Protein Interactions* (Revzin, A., ed) CRC Press, Boca Raton, FL
Cox, M. M., and Lehman, I. R. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 3433–3437
Cox, M. M., and Lehman, I. R. (1987) *Annu. Rev. Biochem.* **56,** 229–262
Cox, M. M., McEntee, K., and Lehman, I. R. (1981) *J. Biol. Chem.* **256,** 4676–4678
Cox, M. M., Pugh, B. F., Schutte, B. C., Lindsley, J. E., Lee, J. W., and Morrical, S. W. (1987) in *DNA Replication and Recombination* (Kelly, T., and McMacken, R., eds) pp. 597–607, Alan R. Liss, New York
Craig, N. L., and Roberts, J. W. (1981) *J. Biol. Chem.* **256,** 8039–8044
Crick, F. H. C. (1957) in *The Chemical Basis of Heredity* (Elroy, W. D., and Glass, B., eds) pp. 532–539, The Johns Hopkins Press, Baltimore, MD
Cunningham, R. P., Wu, A. M., Shibata, T., DasGupta, C., and Radding, C. M. (1981) *Cell* **24,** 213–223
DasGupta, C., and Radding, C. M. (1982) *Nature* **295,** 71–73
DasGupta, C., Wu, A. M., Kahn, R., Cunningham, R. P., and Radding, C. M. (1981) *Cell* **25,** 507–516
Dunn, K., Chrysogelos, S., and Griffith, J. (1982) *Cell* **28,** 757–765
Egelman, E. H., and Stasiak, A. (1986) *J. Mol. Biol.* **191,** 677–697
Fishel, R. A., Detmer, K., and Rich, A. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 36–40
Flory, J., Tsang, S. S., and Muniyappa, K. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81,** 7026–7030
Garger, S. J., Griffith, O. M., and Grill, L. K. (1983) *Biochem. Biophys. Res. Commun.* **117,** 835–842
Gonda, D. K., and Radding, C. M. (1986) *J. Biol. Chem.* **261,** 13087–13096
Granger, M., Toulme, F., and Helene, C. (1982) *Photochem. Photobiol.* **36,** 175–180
Griffith, J., Bortner, C., Christiansen, G., Register, J., and Thresher, R. (1990), in *Molecular Mechanisms in DNA Replication and Recombination* (Lehman, I. R., and Richardson, C. C., eds) pp. 105–114, Alan R. Liss, New York
Hanvey, J. C., Shimizu, M., and Wells, R. D. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 6292–6296
Honigberg, S. M., Gonda, D. K., Flory, J., and Radding, C. M. (1985) *J. Biol. Chem.* **260,** 11845–11851
Howard-Flanders, P., West, S. C., and Stasiak, A. (1984) *Nature* **309,** 215–220
Htun, H., and Dahlberg, J. E. (1988) *Science* **241,** 1791–1796
Inman, R. B., and Schnös, M. (1970) *J. Mol. Biol.* **49,** 93–98
Kaguni, J., and Ray, D. S. (1979) *J. Mol. Biol.* **135,** 863–878
Kim, J.-I., Heuser, J., and Cox, M. M. (1989) *J. Biol. Chem.* **264,** 21848–21856
Lindsley, J. E., and Cox, M. M. (1990) *J. Biol. Chem.* **265,** 10164–10171
Littlewood, R. K., and Inman, R. B. (1982) *Nucleic Acids Res.* **10,** 1691–1706
Lohman, T. M., and Overman, L. B. (1985) *J. Biol. Chem.* **260,** 3594–3603
Lohman, T. M., Green, J. M., and Beyer, R. S. (1986) *Biochemistry* **25,** 21–25
Menetski, J. P., Bear, D. G., and Kowalczykowski, S. C. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 21–25
Messing, J. (1983) *Methods Enzymol.* **101,** 20–78
Moser, H. E., and Dervan, P. B. (1987) *Science* **238,** 645–650
Pugh, B. F., and Cox, M. M. (1987) *J. Biol. Chem.* **262,** 1337–1343
Radding, C. M. (1988) in *Genetic Recombination* (Kucherlapati, R., and Smith, G., eds) pp. 193–229, American Society for Microbiology, Washington, D. C.
Register, J. C. III, Christiansen, G., and Griffith, J. (1987) *J. Biol. Chem.* **262,** 12812–12820
Schutte, B. C., and Cox, M. M. (1987) *Biochemistry* **26,** 5616–5625
Schutte, B. C., and Cox, M. M. (1988) *Biochemistry* **27,** 7886–7894
Shibata, T., DasGupta, C., Cunningham, R. P., and Radding, C. M. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77,** 2606–2610
Stasiak, A., Stasiak, A. Z., and Koller, T. (1984) *Cold Spring Harbor Symp. Quant. Biol.* **49,** 561–570
Veronese, F. M., Schiavon, O., Bevilacqua, R., Bordin, F., and Rodighiero, G. (1982) *Photochem. Photobiol.* **36,** 25–30
Williams, R. C. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 2311–2315
Williams, R. C., and Spengler, S. J. (1986) *J. Mol. Biol.* **187,** 109–118
Wilson, J. H. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76,** 3641–3645
Wollenzien, P. L., Hearst, J. E., Squires, C., and Squires, C. (1979) *J. Mol. Biol.* **135,** 285–291
Wu, A. M., Bianchi, M., DasGupta, C., and Radding, C. M. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80,** 1256–1260

Downloaded from www.jbc.org at University of Wisconsin-Madison, on January 26, 2012