Pairing of homologous DNA sequences by proteins: evidence for three-stranded DNA

Peggy Hsieh, Carol S. Camerini-Otero, and R. Daniel Camerini-Otero

Genetics and Biochemistry Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892 USA

We show that recombinases form joint molecules over very short regions of homology. When these molecules are deproteinized the three strands are in a structure that is surprisingly resistant to dissociation by branch migration, even at elevated temperatures. The joint molecules dissociate at temperatures comparable to those required to melt DNA duplexes of the same length and sequence. We also show that nonenzymatically formed structures of the same length and sequence, which have a free third strand ready to branch migrate, dissociate at much lower temperatures. These results provide compelling evidence that the three DNA strands in the region of pairing are hydrogen bonded to each other. Our observations suggest that such a novel three-stranded DNA molecule, or a structure very similar to it, may be the intermediate in general recombination that is used in the recognition of sequence homology. We discuss some of the structural features implicit in this molecule containing any base sequence and compare them with those manifest in true DNA triple helices containing special sequence motifs.

[Key Words: Homologous recombination; recA; gene targeting; DNA triple helix]

Received June 29, 1990; revised version accepted August 23, 1990.

General genetic recombination involves the exchange of genetic information between sequences residing on two homologous chromosomes. A key step in models of homologous recombination is DNA strand exchange which, in general terms, involves the pairing of two DNA duplexes followed by the exchange of strands to yield Holliday structures containing heteroduplex DNA. Several purified proteins that carry out strand exchange in vitro have been shown to be essential for homologous recombination. These are the *Escherichia coli* recA protein, the uvsX protein from T4, and the rec1 protein from the fungus *Ustilago maydis* (for review, see Cox and Lehman 1987).

The most extensively studied homologous recombination protein, recA, is required for induction of the SOS response, DNA repair, and genetic recombination in *E. coli* (for review, see Radding 1982; Cox and Lehman 1987). In a typical in vitro strand exchange assay, recA can pair a linear duplex DNA and a homologous single-stranded circular DNA. Heteroduplex DNA, thousands of base pairs in length, can be formed with displacement of one of the original duplex DNA strands and formation of nicked, circular [form II] DNA.

Previously, we have described the partial purification of strand exchange activities, termed recombinase fractions, from the human B lymphoblastoid cell line RPMI 1788 [Hsieh et al. 1986], HeLa cells [Hsieh and Camerini-Otero 1989], and *Drosophila melanogaster* embryos [Eisen and Camerini-Otero 1988]. DNA strand exchange or pairing activities from mitotic and meiotic cells of *Saccharomyces cerevisiae* [Kolodner et al. 1987; Sugino et al. 1988; Halbrook and McEntee 1989], *Drosophila melanogaster* embryos [McCarthy et al. 1988], and human cells [Fishel et al. 1988] have been reported by others.

The central problem in the mechanism of homologous recombination is how two homologous sequences are paired during the early steps of strand exchange. Unlike proteins that mediate site-specific recombination such as the Int protein from phage λ (for review, see Landy 1989), proteins involved in general recombination do not bind DNA in a site-specific manner. Instead, these proteins must be able to recognize and pair all DNA sequences on the basis of shared homology. Although it is well established that a complete strand exchange reaction by recA protein yields heteroduplex DNA and a displaced single strand, it is unclear when the strand separation step occurs. One possibility is that strand separation or melting of the duplex DNA occurs very early and facilitates the pairing or annealing of a homologous single-stranded DNA giving rise to heteroduplex DNA.

Arguing strongly against this possibility is our observation that the HeLa and *Drosophila* recombinase fractions, as well as *E. coli* recA protein, recognize and pair homologous sequences without significant melting of duplex DNA prior to strand exchange [Hsieh and Ca-
merini-Otero 1989). We suggested, as have others, that the homology search process takes place prior to strand separation within a three-stranded DNA intermediate [Howard-Flanders et al. 1984; Stasiak et al. 1984; Register et al. 1987].

Shown in Figure 1 are two possible structures involved in early steps of pairing by recombinase proteins. The first structure is formed by an exchange event in which the third strand replaces one of the duplex strands with no net change in the number of noncovalent interactions—in this case, hydrogen bonds. In the lower structure, a three-stranded DNA is formed in which we postulate a net increase in the number of noncovalent interactions. Here, we ask directly whether recombinase proteins can form such a three-stranded DNA. In the past, attempts to describe such a structure have been confounded by two factors: (1) The interactions between the two DNAs have extended over large stretches of DNA homology; and (2) recombinase protein has been associated with the intermediates and/or products formed. In this paper we have studied a protein-free structure involving a very short region of homology. The advantages of this approach are the following. First, by limiting the reaction to the nucleation event, an intermediate might be trapped. Subsequent steps in the strand-exchange reaction such as strand separation might require longer regions of homology. Second, by removing the protein, we could ascribe the structure and its stability to only one macromolecular component, the nucleic acid. Finally, a striking difference between a three-stranded intermediate and one consisting of a heteroduplex DNA with a displaced strand is that only the former would be resistant to dissociation by branch migration, i.e., a protein-independent exchange of hydrogen bonds [see Fig. 1]. A critical test for three-stranded DNA, then, would be the demonstration that structures formed by recombinase protein are very stable. Since branch migration is a random walk process, the time required to reach a destination (e.g., the end of the duplex DNA) is proportional to the square of the distance traveled [Thompson et al. 1976]. Thus, it is difficult to deduce structural information from a study of the stability of joint molecules in which pairing extends over several thousand base pairs. In contrast, the stability of joint molecules involving a region <60 bp on the end of a 3-kb linear duplex should unequivocally allow one to distinguish between the two possible structures shown in Figure 1. Our results provide compelling evidence that recombinase proteins form three-stranded DNA.

**Results**

*Joint molecules with short regions of homology*

To test for the formation of three-stranded DNA, we first monitored the ability of recombinase proteins to form stable joint molecules with short regions of homology. Recombinase proteins were incubated with a linear duplex DNA and a homologous circular single-stranded DNA to form a joint molecule product in which the two substrates are noncovalently joined. Stable joint molecule formation initiates from the ends of the linear duplex. The circular single-stranded DNA substrate is the viral [plus] strand of M13mp18 phage. The linear duplex substrate is a plasmid, pGEM4 [Fig. 2A]. pGEM4 and M13mp18 share two significant regions of homology, a 59-bp region from the E. coli lacI gene and a homologous 57-bp region that constitutes the polylinker cloning sites of both pGEM4 and M13mp18. These two homologous regions are separated in pGEM4 by 37 bp of nonhomologous sequence. When pGEM4 is linearized in the polylinker region, the polarity of strand exchange by the human [HeLa cell] and Drosophila recombinase fractions dictates that only homologous sequences at the right end of the pGEM4 linear duplex, as shown in Figure 2A, are utilized [Hsieh et al. 1986; Eisen and Camerini-Otero 1988]. The right end of pGEM4 was drawn exposes the 3' end of the plus strand of the M13mp18 polylinker sequence.

HeLa recombinase fraction formed joint molecules between M13mp18 single-stranded DNA and linearized pGEM4 double-stranded DNA [Fig. 2B]. The samples were deproteinized prior to electrophoresis. Surprisingly, as few as 13 bases of shared homology were sufficient for the formation of stable joint molecules (lane 4). Homologous sequences 5 bases long on the right end of the linear duplex and 52 bases long on the left end of the duplex were not utilized (lane 5). This result confirms the directionality of strand exchange by the human recombinase fraction. Control experiments using M13mp19 single-stranded DNA containing the negative strand of the polylinker region resulted in the appearance of stable joint molecules when pGEM4 was linearized with EcoRI but not HindIII (data not shown). Deletion of all but 5 bp of homology in assays using pGEM4 digested with both HindIII and EcoRI (lane 6) also
Figure 2. Recombinases form joint molecules using short homologies. [A] A partial map of the pGEM4 linear duplex substrate used in joint molecule assays is shown. Boxes indicate regions homologous to M13mp18. Numbers in the polylinker region at the right of the duplex indicate the lengths of homology available for pairing with M13mp18 single-stranded DNA when pGEM4 is linearized at each of these sites. [B–D] Joint molecule assays using pGEM4 linear duplex and M13mp18 viral strand DNA and HeLa recombinase fraction [B], Drosophila recombinase fraction [C], or E. coli recA and SSB proteins [D] were carried out as described in Materials and methods and were deproteinized by the addition of SDS. Lengths of homology available for pairing are indicated. [Lanes C] Control DNAs incubated alone; [lanes 1] HindIII-linearized pGEM4; [lanes 2] SalI-linearized duplex; [lanes 3] BamHI-linearized duplex; [lanes 4] KpnI-linearized duplex; [lanes 5] EcoRI-linearized duplex; [lanes 6] duplexes digested with both HindIII and EcoRI [B, D], or ligation control, duplex alone [C].

yielded no product indicating that the second region of shared homology located at an internal site 37 bp from the left end of the duplex is not utilized by the HeLa recombinase fraction in forming joint molecules. The joint molecule assay, using very short sequence homologies, was surprisingly efficient with 16–26% of the duplex converted to joint molecules. Under similar assay conditions using completely homologous DNAs, human recombinase fraction converts from one-third to one-half of the linear duplex to joint molecules (Hsieh et al. 1986; Hsieh and Camerini-Otero 1989).

Like the human recombinase fraction, the Drosophila recombinase fraction also formed joint molecules with as few as 13 bp of homology [Fig. 2C, lane 4], did not utilize 5 bp of homology, and exhibited the expected directionality of joint molecule formation [lane 5]. Approximately 20% of the linear duplex was converted to joint molecules in lanes 1–3. Under similar assay conditions using completely homologous DNAs, Drosophila recombinase fraction converts two-thirds of the duplex to joint molecules (Eisen and Camerini-Otero 1988).

Joint molecule assays using pGEM4 linear duplex, M13mp18 single-stranded DNA, recA protein, and E. coli SSB protein are shown in Figure 2D. Assay conditions were those that have been shown by others to promote extensive strand exchange by recA over thousands of base pairs (for review, see Radding 1982). We observed that 56 or 38 bp of homology [lanes 1 and 2] was sufficient for the formation of joint molecules by recA; 28% and 19% of the linear duplexes were converted to joint molecules, respectively. However, no joint molecules were observed when the homology was limited to 26 bp [lane 3]. Under identical reaction conditions using completely homologous substrates, we observe that recA converts essentially all of the duplex to higher-order structures [data not shown]. In this assay, the directionality of recA was identical to that of the eukaryotic recombinase fractions, that is, only homology on the right end of the linear duplex was utilized by recA protein [cf. lanes 1 and 5]. As we observed for the human recombinase fraction, joint molecules were formed efficiently with EcoRI-linearized pGEM4 but not with HindIII-linearized duplex when recA and SSB proteins were incubated with M13mp19 single-stranded DNA [data not
shown). At present, we do not understand why the polarity of strand pairing by recA with short regions of homology appears different from that observed by others with substrates sharing extensive regions of homology (for review, see Radding 1982). It has been observed, however, that the directionality of strand exchange by recA, unlike that of the eukaryotic recombinases, is substrate-dependent (Konforti and Davis 1990).

Electron microscopy of joint molecules formed by recA confirmed that the pairing of a linear duplex to one single-stranded circular DNA occurred at the end of the duplex (data not shown). Although no displaced strand was seen, the resolution was insufficient for determining the configuration of the third strand in the region of pairing. In addition, the formation of joint molecules by all three recombinases required the simultaneous presence of both DNAs and recombinase protein and was not due to exonucleolytic processing of the duplex and subsequent annealing of the single-stranded DNA (see Fig. 6). When M13mp18 single-stranded DNA and HindIII-linearized pGEM4 DNA (56 bp of homology) were separately incubated with recombinase, followed by coinubcation of the treated substrates at 65°C for 10 min in the presence of 1% SDS and 10% [wt/vol] polyethylene glycol to promote nonenzymatic annealing, no joint molecules were formed (data not shown).

A novel joint molecule assay

An explanation for the stability of these joint molecules is that the circular single strand is nonspecifically held in place or “pinched” by the duplex at the junction of the region of homology and nonhomology. To rule this and other forms of trapping out, we have designed a novel assay in which the single strand has two ends, is very short, and is completely homologous to the duplex. Furthermore, we demonstrate that recognition of short homologies is not limited to the polylinker sequence of M13mp18. A novel assay for the formation of joint molecules between a linear duplex DNA and a 5'-32P-labeled oligonucleotide identical to the 3' end of one strand of the linear duplex is shown in Figure 3A. The formation of stable joint molecules was monitored by the appearance of label migrating at the position of linear duplex DNA on agarose gels.

Stable joint molecules were formed by the human recombinase fraction between pdel 9 linear duplex DNA and an oligonucleotide 33 bases or 20 bases in length (Fig. 3B, lanes 1 and 3). No joint molecules were formed with an oligonucleotide 13 bases long (lane 5) or with a nonhomologous oligonucleotide (data not shown). Approximately 50% of the linear duplex in lane 1 was converted to joint molecules. Control experiments in lanes 2, 4, and 6 demonstrate that joint molecules were not formed by degradation of the duplex by potential exonucleases in the recombinase fraction preparations, thereby exposing the duplex strand that could anneal to the oligonucleotide; when the two DNA substrates were incubated separately with recombinase fraction and then coincubated for 10 min at 65°C in the presence of SDS to stop activity, virtually no joint molecules were observed. No joint molecules were formed when an oligonucleotide corresponding to the negative strand of pdel 9 was used, reflecting the 3' to 5' directionality of the human recombinase fraction, that is, pairing of the oligonucleotide initiates at the 3' end of the noncomplementary strand of the duplex (data not shown).

Thermal stability of joint molecules

If the deproteinized joint molecules with short regions of homology discussed above had a displaced strand, branch migration would lead to their rapid dissociation. DNA branch migration is a random walk process where, depending on the geometries of the DNAs examined, the time required to step through a base pair is between 12 and 200 μsec (Thompson et al. 1976; Radding et al. 1977; Green and Tibbetts 1981). Thus, the time required to reach the end of the duplex is approximately equal to the distance (56 bp) squared, multiplied by the step time (0.2 msec) divided by 2 or ~300 msec (Feller 1957). Surprisingly, we observed that these deproteinized joint molecules were stable for hours (>104 sec) at room temperature.

The stability of joint molecules at room temperature reflects the formation of additional interactions between

![Figure 3. A novel joint molecule assay.](#)

**Figure 3.** A novel joint molecule assay. (A) Scheme for joint molecule formation between a linear duplex DNA and a homologous 32P-labeled oligonucleotide. (B) Joint molecule assays were performed at 37°C for 10 min with HeLa recombinase fraction, 100 ng of HindIII-linearized pdel 9 duplex, and 5 ng of a homologous 32P-labeled oligonucleotide 33 bases long (lane 1), 20 bases long (lane 3), or 13 bases long (lane 5). Samples were deproteinized with 1% SDS. Control experiments in lanes 2, 4, and 6 demonstrate that joint molecules are not formed by potential contaminating exonucleases. Duplexes and oligonucleotides (33-mer, lane 2; 20-mer, lane 4; 13-mer, lane 6) were incubated separately with recombinase at 37°C, brought to 1% SDS, and combined and annealed at 65°C without quenching on ice prior to electrophoresis at room temperature. Under these annealing conditions, DNA duplexes 33 or 20 bp long can be formed and are stable (data not shown).
the third strand and the duplex. To ascertain the strength of these interactions and to rule out covalent interactions, we examined the thermal stability of joint molecules and compared their thermal stability with the stabilities of partially duplex molecules and molecules that can undergo nonenzymatic branch migration (see Fig. 4A). The short duplex regions of the partially duplex molecule, I, corresponded exactly in sequence and length to the presumed paired regions of joint molecules formed by recombination between pGEM4 linear duplexes digested with HindIII (56 bp), SalI (38 bp), BamHI (26 bp), and KpnI (13 bp) and M13mp18 single-stranded DNA shown in Figure 2. The branch migration structure, II, corresponded exactly in sequence, length, and direction of potential branch migration to the joint molecules formed by recombinases (see Materials and methods). Nonenzymatic branch migration would result in the displacement of the 32P-labeled oligonucleotide. In this assay, displacement of the annealed fragment was via homology-dependent branch migration, as no displacement of a 32P-labeled fragment was observed when the second oligonucleotide contained a 10-base annealing site attached to a sequence not homologous to the duplex region (data not shown).

Representative data from stability experiments for a homologous region of 38 bp corresponding to joint molecules formed between SalI-linearized pGEM4 and M13mp18 DNAs are shown in Figure 4B. While the duplexes (58% G-C) were stable after 10 min at 65°C, melting at 75°C, branch migration structures dissociated after 10 min at 45°C. Interestingly, joint molecules formed by human recombinase fraction were stable after 10 min at 65°C and dissociated at 75°C. This represents a 30°C difference between the relative stability of joint molecule products formed by recombinase versus branch migration structures of identical length, sequence, and direction of branch migration. The faint band migrating slower than joint molecules in the ethidium bromide-stained gel is a dimer of pGEM4 resulting from a contaminating ligase activity.

Similar thermal stability data for duplexes, branch migration structures, and joint molecules formed by the three recombinases over four different lengths of homology are summarized in Figure 5. Joint molecules were exceedingly stable, dissociating at temperatures 20–30°C higher than those of branch migration structures. In all cases, observed melting temperatures of the duplexes were in close agreement with calculated \( T_m \)s.
bases. We attempted to reconstruct joint molecule structures nonenzymatically to assess their stability relative to joint molecules formed by recombinases. Heat-denatured linear pGEM4 and M13mp18 single-stranded DNA were annealed at neutral pH or at pH 5.5 (Voloshin et al. 1988) in the presence of 10% (wt/vol) polyethylene glycol and then electrophoresed on agarose gels. No joint molecules were observed. Similarly, annealing heat-denatured linear pGEM4 with 32P-labeled oligonucleotides failed to yield stable joint molecules (data not shown).

*Stable joint molecules have three intact strands*

The surprising stability of these joint molecules could be explained if the 3' end of the plus strand of the linear duplex had been exonucleolytically degraded, leaving only heteroduplex DNA in the joint molecule (see Fig. 1, top). In the following experiment, we show that stable joint molecules have an intact third strand (the plus strand of the polylinker region).

Joint molecule assays using HeLa recombinase fraction or recA and SSB proteins were carried out between M13mp18 single-stranded DNA and a fragment of pGEM4 32P-labeled at the extreme 3' end of the plus strand of the linear duplex (Fig. 6A). Restriction enzyme digests of the labeled duplexes confirmed that >99% of the 32P label was restricted to the 3' terminal 10 bases of the plus strand of the pGEM4 fragment (see Materials and methods). Thus, formation of 32P-labeled joint molecules would indicate that the third strand is intact because heteroduplexes of 9 bp or less are very unstable and impossible to recover under these conditions (cf. Fig. 5). Because 15–20% of the labeled duplex was converted to joint molecules in these assays, virtually all of the 32P-labeled joint molecules must have intact third strands. The thermal stability of 32P-labeled joint molecules formed by recA (Fig. 6B) indicated that molecules having three intact strands dissociated at relatively high temperatures (cf. Fig. 5). The rapidly migrating lower band represents the nonenzymatic loss of 32P label from the 32P-labeled duplex that occurred at higher temperatures in the absence of recombinase protein (data not shown).

Quantitation of the relative stabilities of 32P-labeled joint molecules formed by recA and HeLa recombinase fraction and the corresponding 56-base branch migration structure is shown in Figure 6C. The 32P-labeled joint molecules formed by human recombinase fraction, like those formed by recA protein, were exceedingly stable. These results, using a truncated pGEM4 linear duplex, also demonstrate that the stability of joint molecules derives from pairing restricted solely to the region of homology in the polylinker at the right end of the linear duplex.

*Stable joint molecules are devoid of protein*

A possible explanation for the stability of joint molecules is that they remain associated with recombinase protein in the presence of 1% SDS. In Figure 7, we demonstrate that the stable DNA joint molecules formed by recA protein are devoid of any associated recA molecules after treatment with several deproteinizing agents. Joint molecules having 56 bp of homology were deproteinized with proteinase K in the presence of 1% SDS and extracted with phenol/chloroform as described in Materials and methods. The level of residual recA pro-
protein was assessed by Western blotting using a polyclonal antibody against recA protein (Fig. 7A). Identically treated joint molecules were also analyzed for thermal stability and dissociated between 75 and 85°C (Fig. 7B). On the basis of our limit of detection of recA protein by Western blotting (0.1 ng or 2.5 fmole; Fig. 7A, lane 5) and the recovery of deproteinized joint molecules (30 fmole, Fig. 7B, lane 3), we estimate that >90% of the joint molecules was free of recA protein (Fig. 7A, lane 6). No detectable recA polypeptides were observed by silver staining [data not shown]. In the blot, no recA was observed migrating at the position of recA polypeptides or single-stranded or double-stranded DNA.

We also observed that even if active recombinases were present, they did not interfere with branch migration and subsequent dissociation of the DNA molecule. The 32P-labeled partial duplex structures shown in Figure 4A were incubated at 37°C with a second oligonucleotide containing a 10-base annealing site in the presence of recA and SSB proteins or HeLa recombinase fraction under joint molecule assay conditions. Within 10 min, branch migration structures were formed followed by displacement of the 32P-labeled oligonucleotide (P. Hsieh and R.D. Camerini-Otero, unpubl.).

Discussion

The experiments presented here provide compelling evidence that an intermediate in strand exchange is a novel three-stranded DNA in which additional noncovalent interactions are formed between the third strand and the duplex. This conclusion is based on the stability of joint molecules formed by recombinase proteins in which the region of DNA homology available for pairing is extremely small, <60 bp.

Formation of joint molecules with short regions of homology

We observed that HeLa and Drosophila recombinase fractions formed stable joint molecules between a linear duplex and a circular single-stranded DNA, sharing as few as 13 bp of homology. In this assay, purified E. coli recA protein formed joint molecules with 38 bp of homology. Previously, Gonda and Radding (1983) observed that 151 bp of homology, but not 30 bp, was sufficient for recA to pair a linear duplex and circular single-stranded DNA. Our findings establish that the length of homology recognized by recombinases is significantly shorter than thought previously. We also observed that the human recombinase fraction can form stable joint molecules with short homologies using different DNA sequences and single-stranded DNA substrates.

Our data provide insight into the mechanism of homologous pairing and suggest that recombinases can nucleate pairing by rapidly sampling very discrete regions of DNA. In fact, it appears that the initial pairing event mediated by recA involves less than one turn of the DNA helix (P. Hsieh, C.S. Camerini-Otero, and R.D. Camerini-Otero, unpubl.).
Figure 7. Joint molecules are stable in the absence of recA protein. (A) Joint molecules formed by recA were deproteinized as described in Materials and methods and analyzed for residual recA protein on a 12% SDS-polyacrylamide gel followed by Western blotting with anti-recA antibody and 125I-labeled secondary antibody. (Lanes 1–5) 5, 1, 0.5, 0.2, and 0.1 ng, respectively, of purified E. coli recA protein; (lane 6) total deproteinized joint molecules from five assays; (lane 7) 0.2 ng purified recA protein. (B) Thermal stability assays of the deproteinized joint molecules. (Lane 1) DNA substrates alone; (lane 2) a reacted joint molecule assay with recA stopped with SDS and EDTA and loaded on the agarose gel directly; (lanes 3–7) thermal stability assays of deproteinized joint molecules carried out as described in Fig. 4. (ds and ss) Mobilities of double-stranded and single-stranded DNA, respectively.

A large body of genetic evidence from both prokaryotic and eukaryotic systems shows that during recombination, the cell recognizes substrates that involve regions of homology as short as those we have used in this study. These data substantiate the biological relevance of the structures we are studying. Recombination involving <50 bp of homology has been observed in prokaryotes [Albertini et al. 1982, Singer et al. 1982, Watt et al. 1985; Shen and Huang 1986; Whorsikey et al. 1987]. Recombination involving <30 bp of homology has been observed in S. cerevisiae [Ahn et al. 1988], in D. melanogaster embryos [Pont et al. 1988], and in cultured mammalian cells [Anderson et al. 1984; Rubnitz and Subramani 1984, Ayares et al. 1986; Rommerskirch et al. 1988].

DNA pairing occurs within three-stranded DNA
Howard-Flanders and colleagues proposed that the nuclease site for the pairing of homologous duplex DNAs by recA protein was a region of three-stranded DNA [West et al. 1981; Howard-Flanders et al. 1984]. Some suggestive evidence for this model has come from studies of recA protein. When joint molecule intermediates containing a linear duplex, a completely homologous single-stranded circular DNA, and recA were viewed by electron microscopy, some regions devoid of recA had no displaced single-stranded DNA. The absence of displaced strands, even in long stretches devoid of recA, was interpreted as being consistent with a three-stranded structure [Stasiak et al. 1984; Register et al. 1987]. It is difficult, however, to exclude whether the observed structure was stabilized by the recA protein in adjacent regions. Recently, Rao et al. [1990] reported that the potentially displaced third strand in joint molecules is not in an accessible conformation. For example, the third strand is not a cofactor for the single-stranded DNA-dependent ATPase activity of recA and is not digested by PI nuclease.

In this paper we demonstrate that a three-stranded structure can be formed by recombinase proteins by determining that joint molecules are quite stable even when pairing interactions are limited to very small regions of DNA. We have ruled out potential artifacts that might account for this stability such as [1] trapping of the single strands, [2] formation of joint molecules by contaminating exonucleases, and [3] stabilization of joint molecules by residual protein. The unusual thermal stability of the joint molecules argues strongly against the dissociation of joint molecules by the exchange of hydrogen bonds [branch migration]. Because we observed previously that the recombinase proteins do not mediate homology recognition and pairing of DNA sequences by melting the duplex DNA [Hsieh and Camerini-Otero 1989], our conclusion is that recombinase proteins form three-stranded joint molecules in which the third strand is interacting with an intact duplex DNA. The exact relationship between the three-stranded DNA structure associated with the recA protein filament and the deproteinized three-stranded structure described here remains to be determined. Nevertheless, we feel that many of the key structural features can be studied with our model system.

Previously, we [Hsieh et al. 1986; Eisen and Camerini-Otero 1988] and others [see Cox and Lehman 1987] observed that joint molecules contained a displaced single strand when viewed by electron microscopy. We do not know whether three-stranded structures are stable in the 50–60% formamide used in preparing our samples for electron microscopy. Another possibility is that strand exchange [and strand displacement] occurs when the paired regions are longer, as is typical in electron microscopy studies, than those considered here.

A new three-stranded DNA structure
We propose a possible structure for this three-stranded DNA with the aim of arriving at a model with certain
predictive power. We examine how these predictions are borne out by known properties of recombinase proteins and discuss how we might test the model in the future. The overriding constraint of such a structure is that it allows for the dependence of its formation on sequence homology between the third strand and one of the duplex strands.

Broadly speaking, we can consider two major classes of structures, those in which the third strand is hydrogen bonded to a duplex with intact Watson–Crick pairing and those in which the third strand is intercalated between the bases of such a duplex. Although we cannot rule out the intercalation model, it is less attractive for the following reasons. First, to our knowledge there is no experimental evidence for a triple-stranded structure based on intercalation. Second, the requirement for stringency in homologous pairing dictates that intercalation should occur at every base pair; this would probably require an axial rise per base pair of double-stranded DNA and per base of single-stranded DNA of ~6.8 Å. We note that such a structure is inconsistent with the extent of lengthening of these DNAs by recA, resulting in an axial rise of 5.1 Å [see below]. Furthermore, depending on the extent of unwinding of the duplex, such a structure would probably require distortions of the phosphate backbone approaching the stereochemical maximum of 7.5 Å. Finally, intercalation at every base pair would introduce a huge distortion of the duplex; this geometrical distortion has been postulated to account for the fact that most intercalating agents interact with only every other base pair of duplex DNA (the nearest-neighbor exclusion principle; see Saenger 1984).

A more attractive structure is one in which the third strand is hydrogen bonded to the duplex in a fashion similar to that observed in true DNA triple helices. Triple helices were first described three decades ago (Felsenfeld et al. 1957) and have recently been observed to form by intermolecular pairing of a duplex and an oligonucleotide (Moser and Dervan 1987) and by intramolecular rearrangement of supercoiled plasmid DNA (Vosloshin et al. 1988). Properties of these model triple helices include (1) the requirement for strands containing only purines or pyrimidines, (2) positioning of the third strand in the major groove of the duplex, (3) an antiparallel orientation of the third strand with respect to the identical strand of the duplex, (4) pairing of the third strand [purine or pyrimidine] with the purine strand of the duplex, (5) two hydrogen bonds formed between the third strand and the purine, and (6) protonation of cytosine residues in the third strand (for review, see Wells et al. 1988). Although the protein-independent formation of DNA triple helices containing oligopurine–oligopyrimidine sequences is well established and relatively facile, it is not known whether under certain special conditions or in the presence of proteins, triple helices can form between a duplex and a homologous single-stranded DNA of any sequence.

Although we have no direct evidence for the formation of a structure resembling a triple helix by recombinases, our hypothetical structure would probably share some properties with model triple helices, but differ in some critical respects [see Fig. 8A]. Because most triplexers containing oligopurine–oligopyrimidine sequences melt at temperatures at or below those of corresponding duplexes (cf. Krakauer and Sturtevant 1968) and these triplexers have two hydrogen bonds between the third strand and the duplex, we propose that there are at least two hydrogen bonds between the bases of the third strand and the duplex in our hypothetical structure. In addition, as there are too few potential hydrogen bond donors and acceptors on pyrimidines and in the minor groove, the third strand must pair with purines in the major groove of the duplex DNA. The above are all properties of model triple helices.

In sharp contrast with the situation in previously described triple helices, in our hypothetical structure, because the single strand contains both purines and pyrimidines and must pair by homology to the duplex, the third strand must be parallel to its identical strand in the duplex (Fig. 8A). In addition, the third strand must alternate pairing with the two duplex strands [dashed lines in Fig. 8A], depending on whether the purine is located on the Watson or Crick strand of the duplex. Wherever such a strand switch occurs, there will be a reduction in the normal stacking interactions observed between adjacent bases (Cantor and Schimmel 1980), that is, the successive bases in the third strand are no longer related by a uniform angular rotation as they are in model triple helices.

Despite these constraints, a number of possible base triplets could be envisioned (Saenger 1984). Preliminary model building indicates that the triplets shown in Figure 8B can be accommodated in a triple helix while maintaining acceptable phosphodiester backbone lengths (P. Hsieh and R.D. Camerini-Otero, unpubl.). With respect to the issue of the specificity of pairing, it appears that backbone constraints may play a role in disallowing certain mismatches.

Some predictions of the model

As indicated above, the third strand would have to switch pairing between the parallel and antiparallel strands. This would entail considerable unstacking of the bases in the single strand and distortion of the phosphodiester backbone at every strand switch. Both of the above would contribute to a large activation energy in the formation of these three-stranded DNAs. We predict that the proteins involved in homologous recombination would facilitate the formation of these three-stranded DNAs by lowering this activation energy. This is borne out by some of the properties of recA. When recA forms filaments on single-stranded DNA, the axial rise per base increases to ~5.1 Å, indicating that the bases are unstacked (Flory et al. 1984). This nucleoprotein filament is the active species that initiates pairing in the
strand exchange reaction and has a high affinity for double-stranded DNA (Cox and Lehman 1987). recA elongates duplex DNA to about the same extent and, just as importantly, unwinds the duplex by ~12–15° per base pair [for review, see Cox and Lehman 1987]. We imagine that the nucleoprotein filament formed between recA and single-stranded DNA also elongates and unwinds duplex DNA. This unwinding would minimize the magnitude of the excursions involved in strand switches. In contrast with the central role of recombinase proteins in site-specific recombination, the search for the target sequence in homologous recombination would not be mediated solely by the protein component of the protein–DNA complex. We propose, instead, that homologous recombinases promote conformational changes in the DNA molecules that allow the DNAs themselves to mediate the recognition process. In essence, these recombinase proteins could promote the annealing of a single strand to duplex DNA.

The ability to localize pairing interactions to short, defined regions of DNA and to isolate stable structures free of any recombinase protein will facilitate future studies required to test critically this novel DNA structure. For example, formation of joint molecules might be inhibited by the presence of blocking groups in the major but not the minor groove of the duplex, and one should be able to identify specific groups both in the duplex and the single strand that are critical for hydrogen bonding. In addition, one might be able to assess the energetic costs of strand switching by determining the effect of sequence context on both joint molecule formation by recombinase proteins and homologous recombination in cells. Finally, it is interesting to speculate that this unusual three-stranded DNA structure might be recruited by the cell in other processes where intimate DNA–DNA contacts are required.
Materials and methods

Recombinase proteins

Partially purified HeLa recombinase fraction was prepared from nuclear extracts as described elsewhere [Hsieh and Camerini-Otero 1989]. Partially purified Drosophila recombinase, fraction IV, was prepared from 24-hr embryos as described previously [Eisen and Camerini-Otero 1988]. The human recombinase preparation was free of any 3'→5' exonuclease activities, as determined by trichloroacetic acid (TCA)-precipitable counts. No ^32P cpm were released from ^32P-3'-end-labeled duplex DNAs after incubation with human recombinase fraction using joint molecule assay conditions (data not shown). As described previously, the Drosophila recombinase fraction removed <10% of ^32P cpm in a similar assay [Eisen and Camerini-Otero 1988]. Although not relevant to the joint molecule assays described here, 5'→3' exonuclease activity in recombinase fraction preparations was exceedingly low, <5% ^32P cpm and 10–15% ^32P cpm were released as TCA-soluble counts by the Drosophila and human recombinase fractions, respectively [Eisen and Camerini-Otero 1988, and data not shown]. Purified E. coli recA and SSB proteins were generously provided by Dr. Stephen C. Kowalczykowski, Northwestern University Medical School (Kowalczykowski and Krupp 1987). Both E. coli proteins migrate as a single homogeneous polypeptide by SDS-PAGE and contain no detectable exonuclease activities as judged by release of TCA-soluble counts [S. Kowalczykowski, pers. comm., and data not shown].

DNA substrates

pGEM4 plasmid DNA was obtained from Promega Biotec. M13mp18 viral DNA and restriction enzymes were obtained from New England Biolabs. M13mp19 viral DNA and pdel 9 plasmid DNA, a derivative of pBR322 deleted from nucleotides 1745 to 2505 (Brenner et al. 1985), were prepared according to standard procedures. Oligonucleotides were synthesized and purified as described elsewhere [Hsieh and Camerini-Otero 1989]. Oligonucleotides homologous to the plus strand of pdel 9 spanned pBR322 positions 17–29, 10–29, and 4359–29 (Sutcliffe 1978) for the 13-mer, 20-mer, and 33-mer, respectively. Oligonucleotides homologous to the polylinker region of pGEM4 were derived from the negative strand of M13mp19 and mapped to M13mp19 positions listed below (Yannisch-Perron et al. 1985). Oligonucleotides used for partial duplexes spanned positions 6278–6290, 6265–6290, 6253–6290, and 6235–6290 for 13, 26, 38, and 56 bp, respectively, those used for branch migration structures spanned 6278–6300, 6265–6300, 6253–6300, and 6235–6300 for 13, 26, 38, and 56 bp, respectively.

Oligonucleotides used were labeled with [γ-^32P]ATP (New England Nuclear) and T4 polynucleotide kinase (Pharmacia) and desalted by passage over G25 spin columns (Boehringer Mannheim) or by dialysis. Preparation of ^32P-labeled partial duplexes from labeled oligonucleotides and M13mp19 viral strand DNA was as described elsewhere [Hsieh and Camerini-Otero 1989]. The branch migration structure was formed by incubating the partial duplex molecule with an unlabeled oligonucleotide identical in sequence to the ^32P-labeled annealed oligonucleotide, but containing an additional 10-base annealing site immediately 5' to the ^32P-labeled fragment. DNA concentrations are expressed as moles of nucleotides or by weight.

The length of homology shared between a single-stranded and double-stranded substrate is defined as the maximum number of bases that could be paired between the single-stranded DNA and the complementary strand of the linear duplex substrate in a joint molecule. Two significant regions of homology shared between pGEM4 and M13mp18 were found by use of the University of Wisconsin BESTFIT program. They are in opposite orientations with respect to each other in pGEM4. Map positions for the 59-bp region from the lacI gene are M13mp18 6001–6059 [Yannisch-Perron et al. 1985] and pGEM4 104–162. The 57-bp polylinker region maps to M13mp18 positions 6231–6287, and in pGEM4, to positions 10–66.

pGEM4 linear duplex with a unique 3' ^32P label

pGEM4 DNA was linearized with HindIII and labeled by filling in 3' ends with [α-^32P]dATP (New England Nuclear) and cold deoxynucleotides using the Klenow fragment of DNA polymerase [Pharmacia]. The reaction was stopped by heating at 65°C for 10 min and was brought to 100 mM NaCl and digested with BglII. The 1635-bp BglI–HindIII fragment containing the polylinker region was gel purified and dialyzed extensively against 10 mM Tris-HCl (pH 8.0) and 1 mM EDTA. When this ^32P-labeled fragment was digested with PstI and analyzed by densitometry after electrophoresis on an 8% denaturing polyacrylamide gel or a 1% agarose gel at 99% of the label was removed from the 1.6-kb fragment (data not shown). The ^32P label thus resided within 10 bases of the 3' end of the plus strand of the duplex substrate.

Joint molecule assays

Joint molecule assays (20 μl) were carried out for 10 min at 37°C as described [Hsieh and Camerini-Otero 1989] by use of 150 pmol (50 ng) of M13mp18 viral DNA and 150 pmol (50 ng) of linear pGEM4 DNA or 300 pmol (100 ng) of linear pdel 9 and 5 ng ^32P-labeled oligonucleotide with 100 ng HeLa protein or 40 ng Drosophila protein. Assays containing 7 μg recA protein and 700 ng SSB protein were carried out for 10 min at 37°C with preincubation as described [Hsieh and Camerini-Otero 1989]. Assays were brought to 1% SDS and 10 μM EDTA prior to electrophoresis at room temperature on 0.7% agarose gels in TAE buffer [40 mM Tris-acetate (pH 8), 1 mM EDTA] containing ethidium bromide for 12–16 hr at 0.6 V/cm. Quantitation was accomplished by densitometry. We thank Dr. Kunio Nagsa, Frederick Cancer Research Facility (Frederick, MD) for electron microscopy using the modified Kleinschmidt technique.

Thermal stability assays

Partial duplexes (150 pmol) labeled with ^32P, consisting of a ^32P-labeled oligonucleotide annealed to M13mp19 viral strand DNA, were incubated for 10 min at the indicated temperatures or on ice in strand exchange buffer containing 1% SDS and 10 μM EDTA. TAE buffer containing 0.1% bromophenol blue and 50% glycerol was added and the mixtures were electrophoresed for 12–16 hr at 0.6 V/cm on 0.7% agarose gels in TAE buffer containing ethidium bromide followed by autoradiography. For stability assays of branch migration structures, 150 pmol of the ^32P-labeled partial duplexes and 4 ng of a second oligonucleotide containing a 10-base annealing site to M13mp19 were incubated for 10 min at the indicated temperature or on ice in strand exchange buffer containing 1% SDS, 10 μM EDTA, and 10% (wt/vol) polyethylene glycol. The reactions were electrophoresed on agarose gels followed by autoradiography as described above. Joint molecules were formed by recombinases. The reactions were stopped in 1% SDS and 10 μM EDTA, incubated for an additional 10 min at the indicated temperatures or on ice, and electrophoresed on agarose gels.
**Hsieh et al.**

**Deproteinizing joint molecules**

Reacted joint molecule assays containing recA and SSB proteins were brought to 1% SDS, 10 mM EDTA, and 20 µg/ml proteinase K (Boehringer Mannheim) and incubated for an additional 20 min at 37°C. The assays were extracted sequentially with phenol, phenol–chloroform, and chloroform. The samples were extracted four times with anhydrous ethyl ether saturated with water just before use. Residual ether was removed under nitrogen. The deproteinized joint molecules were analyzed by SDS-PAGE (Laemmli 1970) on 12% polyacrylamide gels (Novex) followed by silver staining (Hochstrasser et al. 1988) or Western blotting. Electrophoresis and transfer to nitrocellulose using a Schleicher & Schuell Miniblot apparatus were performed according to manufacturer's instructions. Following transfer, the nitrocellulose filters were blocked in 10% milk in phosphate-buffered saline (PBS) for 1 hr. The filters were incubated for 1 hr with 20 µl of a 1:500 dilution of rabbit antisera raised to purified recA protein (Hazleton Biotechnology) in 5% milk, 0.1% [vol/vol] Tween 20, and PBS. After extensive washing, the filters were incubated for 1 hr with 20 µl of 5% milk and 0.1% Tween 20 in PBS containing 15 µCi of [3H]-labeled donkey anti-rabbit IgG (3000 Ci/mmol, Amersham) followed by extensive washing. Autoradiography was done for 3 days using intensifying screens.

**Acknowledgments**

We thank David Davies, Gary Felsenfeld, Michael Lichten, Kiyoji Mizuuchi, and Howard Nash of the National Institutes of Health for comments on the manuscript. We thank Andrew Eisen for Drosophila recombinase fraction, Stephen Kowalczykowski for E. coli recA and SSB proteins, and Ryoiti Kiyama for recA antibodies. We are grateful to members of the Genetics and Biochemistry Branch for helpful comments and to Linda Taylor for her assistance.

The publication costs of this article were defrayed in part by advertisement in accordance with 18 USC section 1734 solely to indicate this fact.

**References**


Three-stranded DNA


Pairing of homologous DNA sequences by proteins: evidence for three-stranded DNA.

P Hsieh, C S Camerini-Otero and R D Camerini-Otero

Genes Dev. 1990, 4:
Access the most recent version at doi:10.1101/gad.4.11.1951

This article cites 44 articles, 23 of which can be accessed free at:
http://genesdev.cshlp.org/content/4/11/1951.full.html#ref-list-1

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.