RecA Realigns Suboptimally Paired Frames of DNA Repeats through a Process That Requires ATP Hydrolysis

Subhojit Sen, G. Karthikeyan,§ and Basuthkar J. Rao*

Department of Biological Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai 400005, India

Received April 3, 2000; Revised Manuscript Received May 17, 2000

ABSTRACT: Microsatellite repeats such as mono-, di-, and trinucleotides are highly abundant and viable targets for homologous recombination in the genome. However, if recombination ensues in such repetitive regions, they are intrinsically prone to frame misalignments during pairing and might eventually give rise to genetic instabilities. Suboptimally paired frames lead to an abrogation of branch migration at the junctions of mixed sequences and repeats, due to a heterologous register. If so, can recombination machinery rectify such misalignments in order to avoid subsequent arrest in branch migration? We analyzed *Escherichia coli* RecA, the universal prototype of a recombinase, for its pairing abilities across repeats. We used a complementary pairing assay to test whether RecA can mediate realignments of stochastically paired suboptimal frames to a maximally aligned register. Here, we demonstrate that RecA−single stranded DNA filament indeed facilitates such a realignment, probably by sliding the paired strands across mono- and di- as well as trinucleotide repeats. These realignments apparently have no net directional bias. Such a putative "motor" function of RecA seems to be ATP hydrolysis-dependent.

Escherichia coli RecA forms a classical paradigm for understanding the mechanisms of DNA homology search, sequence recognition and subsequent exchange of strands in homologous recombination (*1*−*3*). It is a DNA-dependent ATPase that cooperatively polymerizes on ssDNA[1] and generates a helical nucleoprotein filament where the bases are highly unstacked (*4*−*6*). Such an extended RecA−ssDNA helix is a hallmark of the presynaptic invasive strand that aligns with its homologue during recombination. The search mechanism that precedes homologous alignment remains unclear, wherein sliding as well as three-dimensional processive steps are invoked (*7*−*12*). Both steps of the RecA reaction, namely, pairing and strand exchange, are intimately involved in testing homology (*13*). Although it had been demonstrated that initial pairing is promiscuous (*14*−*17*), the Watson−Crick complementarity principle of base pairing may be the confirmatory check which the system employs to accomplish productive sequence alignments between the recombining strands.

When recombination ensues across repetitive DNA, Watson−Crick pairing runs into frame misalignment problems. This problem becomes acute if the repeat units are much smaller such as those in microsatellites. Genetic and biochemical approaches have shown that such repeat regions indeed show high instabilities (*18*−*21*). Mechanistic studies suggest that base mismatches/loops/ssDNA flaps arising due to replicational slippages lead to such instabilities when remain uncorrected. These genetic instabilities are exaggerated when the genes that govern such repair functions are mutated (*22*−*26*).

This study is prompted by a search for special properties of RecA pairing that might mitigate the problem of frame misalignments in microsatellite repeats. When DNA strands anneal within a repeat region, they do so stochastically, wherein the chemistry of pairing plays no role in specifying the optimal frame of alignment, thus giving rise to a heterogeneous population of annealed products (Figure 1). How different this population is when the same is mediated by RecA protein is the focus of this paper. We had designed a targeted ligation assay (TL assay) to assess the hierarchical role of base mismatches in RecA-mediated pairing (*17*). Using the same assay here, we analyzed whether RecA pairing between simple DNA repeats shows a dynamic realignment following an initial stochastic mode of pairing. Indeed, we find that RecA mediates such realignments of suboptimally paired frames in a slow and time-dependent manner. Thermally annealed strands showed no such dynamic changes. We find no description of such an activity of RecA in the literature to date and discuss the implications of this novel "motor" activity.

MATERIALS AND METHODS

Materials. RecA protein was purified as described (*27*). T4 polynucleotide kinase, T4 DNA ligase, *E. coli* DNA ligase, and DNase I were purchased from Amersham

* To whom correspondence should be addressed: Department of Biological Sciences, Tata Institute of Fundamental Research, Colaba, Bombay 400005, India. Phone: 91 22 2152971, ext 2606; fax 91 22 2152110/2152181; e-mail bjrao@tifr.res.in.

§ Present address: Department of Health and Human Services, NIEHS, National Institutes of Health, P.O. Box 12233, Research Triangle Park, NC 27709.

[1] Abbrevations: ATP, adenosine triphosphate; ATPγS, adenosine 5′-*O*-(3-thiotriphosphate); BSA, bovine serum albumin; dsDNA, double-stranded DNA; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; NAD, nicotinamide adenine dinucleotide; SDS, sodium dodecyl sulfate; SSB, single-stranded DNA binding protein; ssDNA, single-stranded DNA; TL assay, targeted ligation assay.

10.1021/bi000753y CCC: $19.00
Published on Web 07/25/2000

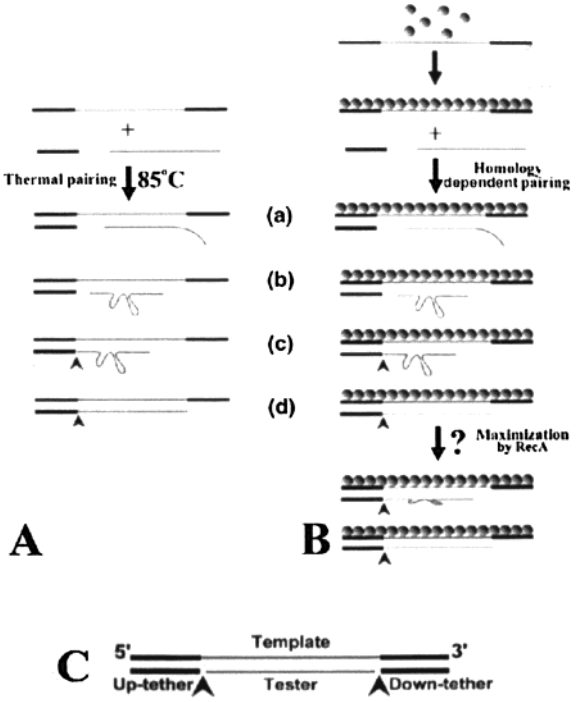

FIGURE 1: Cartoon depiction of the plausible intermediates that ensue during TL assay: Repeat DNA (gray lines) is flanked by mixed sequences (black lines) on either side. (A) Denaturation of strands at 85 °C followed by slow cooling mediates thermal annealing. (B) RecA monomers (shaded circles) polymerize on template strand (presynapsis) to generate RecA−ssDNA filament that mediates pairing. (C) The nomenclature of the substrate system is depicted where the up-tether, down-tether, and tester annealing to the template leads to the formation of ligatable nicks. The arrowhead indicates a ligatable nick. Stochastic pairing across DNA-repeats generates a population of annealed duplexes (a−d), some of which exist in a ligatable alignment (c and d) while the rest (a and b) need to be realigned by a process of "maximization".

Pharmacia Biotech. Adenosine triphosphate, phosphocreatine, creatine phosphokinase, dithiothreitol, and nuclease-free BSA were purchased from Sigma. ATPγS was purchased from Boehringer Mannheim.

DNA Substrates. Single-stranded oligonucleotides were purchased from either Keck Biotechnology Resource Laboratory (Yale, University, New Haven, CT) or from DNA Technology (Denmark). All oligonucleotides were purified on denaturing polyacrylamide gels and desalted further as described (*17*). The purity of an oligonucleotide was judged by ^{32}P labeling at the 5′-end by T4 polynucleotide kinase, followed by analysis on a 10% denaturing polyacrylamide gel which revealed that they were more than 90% pure. The oligonucleotides used are listed in Table 1.

End Labeling of Oligonucleotides. Oligonucleotides were end-labeled by T4 polynucleotide kinase and [γ-^{32}P]ATP as described (*17*).

Targeted Ligation Assay Reactions. Presynapsis was done by incubating template (8 μM) with RecA (2.8 μM) at 37 °C for 10 min in a reaction buffer containing 33 mM Tris-HCl (pH 7.5), 1.2 mM magnesium acetate, 2 mM DTT, 1.2 mM ATP, 8 mM phosphocreatine, creatine phosphokinase (10 units/mL) and BSA (nuclease-free) (100 μg/mL) followed by pairing (synapsis) with tether (2.7 μM) and tester (3.4 μM for 30-mer and 2.8 μM for A_{25}). In various experiments, either tester or tether was ^{32}P-labeled at 5′-end, the details of which are given in the legends. For thermal reactions, all the oligonucleotides used in the reaction were denatured together at 85 °C for 2 min, followed by cooling to room temperature (~25 °C) over a period of 45 min, followed by reconstitution to reaction buffer conditions. TL assay for both thermal and RecA reactions was initiated at room temperature by adding T4 DNA ligase (200 units/mL) and magnesium acetate (10 mM). Reaction aliquots were withdrawn at various time points and quenched by adding EDTA (20 mM), SDS (0.5%), and formamide (50%). The samples were analyzed on a 10% denaturing polyacrylamide gel at 50 °C. For analyses on a native gel, the TL assay samples were deproteinized by proteinase K (100 μg/mL) in 0.5% SDS and 20 mM EDTA at 37 °C for 20 min, followed by the addition of 10% glycerol and analyses on a 10% native polyacrylamide gel at 4 °C. Following electrophoresis, the gels were dried and the radioactivity was quantified by ImageQuant software on a PhosphorImager (Molecular Dynamics). TL assay efficiencies were expressed as a percentage of the ligated product in each lane and plotted as a function of time.

RESULTS

Experimental Setup. Complementary pairing plays a major role when RecA-coated presynaptic strand homologously aligns with a negatively supercoiled duplex, leading to D-loops (*28*−*30*). In the present study, we have used a complementary pairing reaction to assess whether RecA protein has the ability to realign suboptimal pairings that may happen between DNA repeats (Figure 1). A 70-mer "template" strand (T_{30} repeat flanked on either side by a 20-mer mixed sequence) is coated by RecA in an ATP-regeneration system (T-template, Figure 1B). It is known from other studies as well as demonstrated here that a strand of such a length does yield stable RecA−ssDNA filaments (*31*, *32*). Two strands, one that is complementary to the flanking mixed sequence (tether) and the other to the middle repeat (tester) in the 70-mer template, are paired with the RecA−ssDNA filament (Figure 1C). Only a "maximized" pairing frame across the repeats generates a product that is ligatable with the tether (frames c and d, Figure 1). Any other frame that is stable to form (described as suboptimal frames a and b, Figure 1) is unligatable. We measured the specific ligation product on a denaturing gel using either a labeled tether or tester. We did a comparative analysis of ligatable frames in RecA versus thermally annealed reactions as a function of time. More importantly, we tested whether RecA reaction showed a time-dependent accumulation of forms c and d over that of a control (thermal) reaction, which perhaps meant a process of maximization of alignment by RecA (Figure 1).

RecA Promotes Maximization of Alignment between Repeat Sequences. We compared targeted ligation of A_{30}-tester on T-template in RecA versus thermal pairing reactions. RecA reactions showed a time-dependent accumulation of TL assay products whereas thermal reactions failed to show the same [Figure 2 parts A (i) and C]. To assess the annealing efficiencies of these reactions in the absence of ligase, we analyzed the products on a native polyacrylamide gel (Figure 2B). Annealing went to completion in RecA as well as in thermal reactions. These results taken together suggest that although annealing was complete, a significant fraction of the thermally annealed products remained "frozen" in unligatable frames whereas RecA realigns the same

Table 1.

oligonucleotide	sequence (5′ to 3′)
T-template	acgcacatactaggctgtat**ttttttttttttttttttttttttttttttt**ttcagtacagtcatgacagt
GT-template	acgcacatactaggctgtat**gtgtgtgtgtgtgtgtgtgtgtgtgtgtgt**ttcagtacagtcatgacagt
CAG-template	acgcacatactaggctgtat**cagcagcagcagcagcagcagcagcagcag**ttcagtacagtcatgacagt
A_{30}-tester	aaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
A_{25}-tester	aaaaaaaaaaaaaaaaaaaaaaaaa
$(AC)_{15}$-tester	acacacacacacacacacacacacacac
$(CTG)_{10}$-tester	ctgctgctgctgctgctgctgctgctg
up-tether	atacagcctagtatgtgcgttag
down-tether	gatactgtcatgactgtactgaa
mixed-sequence template	ttgataagaggtcatttttgcggat**ggcttagagcttaattgctgaatctggtgctgt**agctcaacatgttttaaatatgcaa
mixed-sequence down-tether	ttgcatatttaaaacatgttgagct
mixed-sequence (−5) down-tether	tttagttgcatatttaaaacatgtt
mixed-sequence tester	acagcaccagattcagcaattaagctctaagcc

to maximized and ligatable forms. Thus RecA maximizes suboptimal frames more efficiently than what is allowed by the thermodynamics of naked DNA annealing. However, the observed kinetics of realignment is rather slow (Figure 2C), which perhaps reflects on the nature of the experimental system being studied, where realignment is mediated by a single protein, while the same is likely to be assisted by a myriad of other helper molecules in vivo (see Discussion).

It is known that intermolecular ligations across blunt ends or cohesive ends in plasmid duplexes are stimulated by RecA (*33*). Is the stimulation observed in the RecA reactions of TL assay [Figure 2 parts A (i) and C] due to similar effects of RecA promoted crowding of DNA ends? It may not be so because TL assay scores an intramolecular nick that is less likely to be influenced by the crowding effects of DNA ends or be limited by diffusion. The following controls have been done to rule out such effects of stimulation in TL assay and thus confirm that RecA does not ligate suboptimally aligned frames spuriously (frames a and b, Figure 1). Essentially, these are repeats of the controls described earlier (see Figure 3 in ref *17*). We compared TL assay efficiencies using a template that has a mixed sequence rather than a repeat in the middle (mixed-sequence template, Materials and Methods). A positive control was done with a 33-mer tester (mixed-sequence tester) that pairs immediately adjacent to the 25-mer tether (mixed-sequence down-tether), forming a ligatable nick. A negative control included a 20-mer tether falling short of five nucleotides [mixed-sequence (−5) down-tether] at the ligating end. Hence, the former represents a ligation control of a nick while the latter depicts a gap ligation that might represent suboptimal frames discussed above (frames a and b, Figure 1). For various time points, TL assay efficiencies obtained for RecA versus thermal annealing have been compared [Figure 2A (ii)]. Nick ligation products formed with equal efficiencies in RecA versus thermal reactions at all time points monitored (Figure 2C). Although ligation efficiencies at very early time points (<10 min) in RecA vs thermal reactions might be different, effects of these early differences, if any, were abolished by the first time point analyzed (i.e., 10 min) [Figure 2 parts A (ii) and C]. Under similar conditions, we performed a gap ligation control in the presence of excess tester to be able to detect even minor levels of spuriously ligatable nicks, if any. Despite the presence of an 8-fold molar excess of tester, neither RecA nor thermal annealing could generate spuriously ligatable nicks across the five nucleotide gap [Figure 2A (iii)]. A parallel control nick ligation performed in thermal and RecA reactions showed 11% and 12% efficiencies that amounted to a completion of the reaction in molar equivalents of the tester used. These two controls, taken together, reveal that the assay scores authentic nicks, thereby indicating that a RecA-dependent accumulation of TL assay products in repetitive sequences stem from a novel "maximization" function of the protein. This aspect has been studied further.

RecA Promotes Maximization of Frame Alignment by a Nonreiterative Mode That Might Involve "Sliding". Is the apparent maximization of repeat tester, exhibited by RecA, a result of reiterative or nonreiterative (sliding, for example) search following a stochastic mode of annealing? We tested this mechanism by a competition experiment based on the assumption that a reiterative search involves multiple rounds of dissociation of the stochastically annealed tester, until it encounters a "maximized" frame. The transient intermediates that ensue during such a reiterative search can be competed out by an exogenous tester.

Two RecA reactions were initiated in parallel (as described in Figure 2A), one with labeled tester and the other with unlabeled tester, and were allowed to anneal for 15 min, following which severalfold molar excess of competitor DNA was added. After 15 min of competition all reactions were monitored by TL assay. It must be noted that, in ATP-regeneration reactions, RecA coating on short oligonucleotide DNA molecules is dynamic and hence can lead to transfer of RecA from the template to the competitor ssDNA, which is in molar excess. Such a transfer might abolish the realignment properties being manifested by the template filament. Hence, to determine a safe concentration range of competitor ssDNA that does not affect the original RecA reaction, we tested realignment in the presence of an increasing molar excess of heterologous tester (mixed-sequence tester, Table 1). The reaction initiated with labeled A_{30}-tester was challenged by heterologous competitor for 15 min, followed by the addition of ligase, and monitored by TL assay [Figure 3A (i)]. The results revealed that the realignment kinetics remained unaffected up to 5-fold molar excess of heterologous competitor. However, at 10- and 25-fold excess competitor, the realignment was stalled as shown by a lack of time-dependent increase in TL assay signals (Figure 3B). Thus, this experiment revealed that when competitor DNA concentration approaches 10-fold molar excess, RecA can be titrated away from the template filament, thereby rendering the filament inactive in strand realignment. Therefore, the experiment using homologous competitor was performed only at concentrations below 10-fold molar excess.

To test the reiterative or nonreiterative nature of the realignment intermediates, we carried out a competition experi-

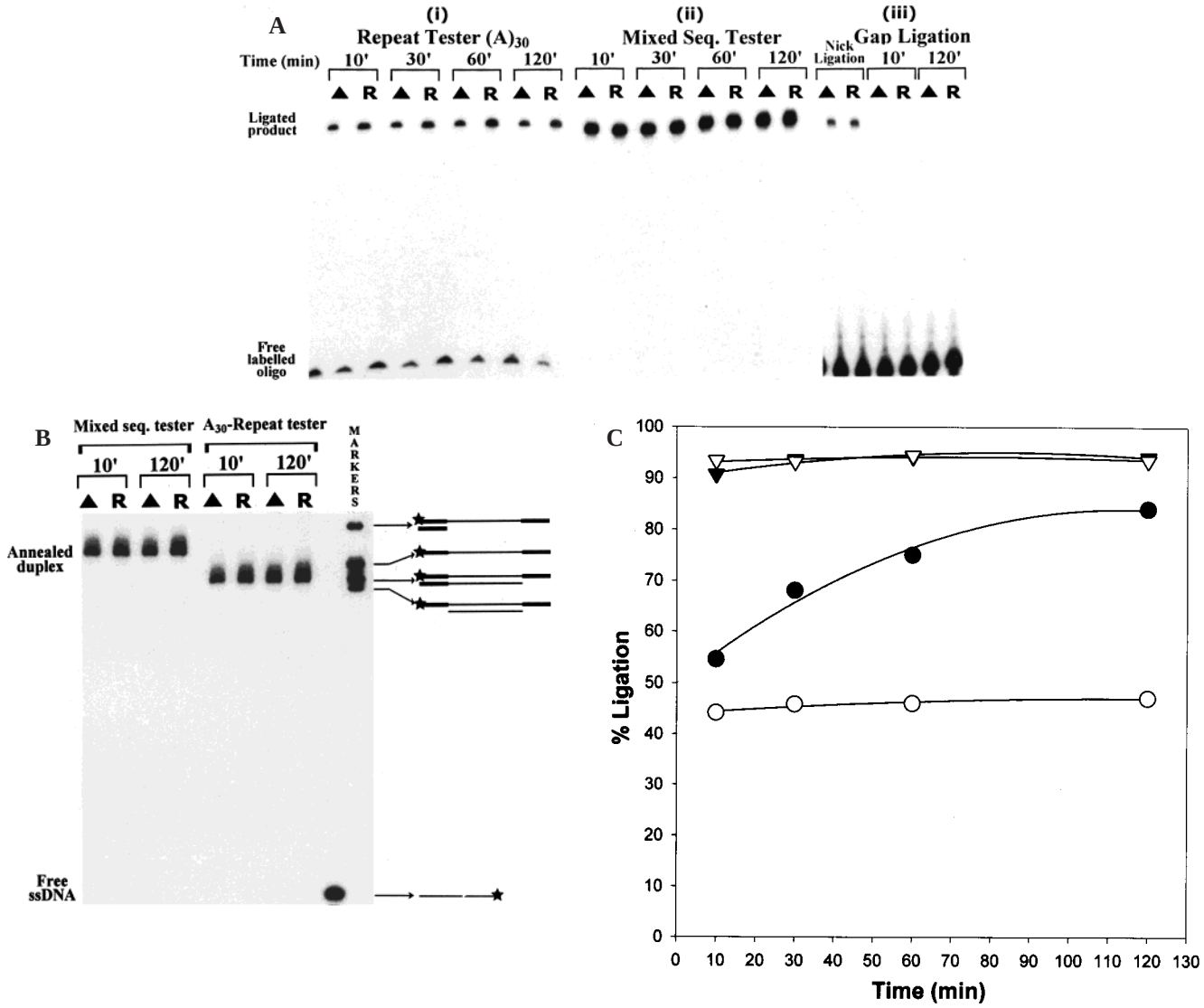

FIGURE 2: (A) Time courses of strand realignment by TL assay: Assay conditions and the substrates are described under Materials and Methods. Experiments described in panel i were with T-template, A_{30}-tester, and labeled up-tether. Panel ii shows nick ligation controls with all mixed-sequence substrates, namely, template, down-tether, and labeled tester. Panel iii has a nick ligation of the same shown in panel ii at 120 min but with 5-fold molar excess of labeled tester. Gap-ligation controls in panel iii were done with the same template and molar excess of labeled tester. However, the tether [mixed-sequence (−5) down-tether] was shifted by five nucleotides from the ligating end of tester, creating a gap. Aliquots of thermal (▲) and RecA (R) reactions were analyzed by denaturing polyacrylamide gel electrophoresis. TL assay products and labeled free oligomer are indicated. (B) Experiments were repeated at conditions identical to those described in part A, panels i and ii, with labeled testers but without ligase. Select samples were analyzed by native 10% polyacrylamide gel electrophoresis. Mobility positions of the free A_{30}-tester, T-template, and various annealed products of these with the corresponding up-tether are indicated as cartoon next to the marker lanes. The star indicates ^{32}P-label at the 5′-end. The difference in the mobility of annealed products with mixed-sequence versus repeat tester is due to different lengths of template strands involved (see Materials and Methods). (C) Quantitation of strand realignment by TL assay. TL assay efficiencies of A_{30}-tester (○, ●) and mixed sequence tester (▽, ▼) (described in part A, panels i and ii) were expressed as a percentage of the ligated product in each lane and plotted as a function of time. RecA, solid symbols; thermal, open symbols.

ment where the reaction was initiated with an unlabeled A_{30}-tester followed by a challenge with molar excess of labeled homologous competitor [Figure 3A (ii)]. In this design, the emergence of a radioactive ligation signal from a "cold" background renders the assay extremely sensitive to detecting a reiterative search mode. As mentioned above, the competition was initiated after 15 min of RecA-mediated annealing, and TL assay was performed after 15 min of competition. Quantitative analysis revealed that labeled competitors did not "chase" the unlabeled tester from the initial pairing (Figure 3B). Even at the highest concentration of competitor used (8×), where RecA filament remains stable, the competition was very marginal and remained independent of the dose of added competitor. One must note that, under these conditions, a parallel positive control initiated with labeled tester indicated that almost the entire population of testers was engaged in realignment (Figure 3B). Therefore, a lack of competition observed here strongly suggested that RecA-mediated maximization involves a nonreiterative internal realignment of frames, perhaps by a mechanism analogous to sliding. A reciprocal competition experiment where the pairing was initiated by a labeled tester, followed by a challenge with increasing concentrations of unlabeled homologous tester, also showed a lack of competition (data not

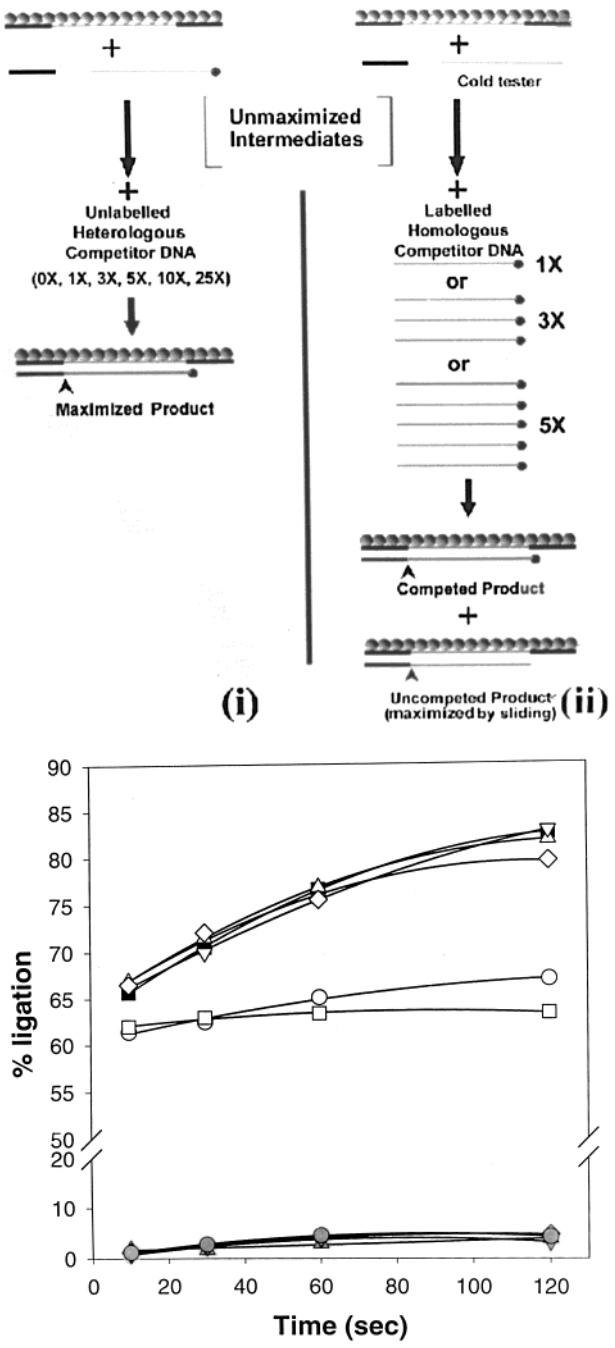

FIGURE 3: (A) Competitor "chase" assay to test whether strand realignment involves physical dissociation of repeat tester strands: (i) RecA-mediated realignment reaction using labeled A_{30}-tester and up-tether followed by challenge with molar excess of heterologous competitor (mixed-sequence tester, Table 1). (ii) RecA-mediated pairing was done for 15 min with cold A_{30}-tester without ligase, followed by the addition of varying doses of labeled A_{30}-tester as competitor. TL assay was initiated by the addition of ligase after 15 min of competition. Circles on template strand indicate RecA monomers, the darker circles on tester strands indicate the ^{32}P label, and the arrowheads indicate ligatable nicks. (B) Competitor chase analysis by TL assay reveals no physical dissociation of A_{30}-tester during strand realignment: TL assay was performed with labeled A_{30}-tester in the presence of 0×(■), 1×(△), 3×(▽), 5×(◇), 10×(○), and 25×(□) heterologous competitor (top graphs). Parallel reactions were initiated with cold A_{30}-tester followed by competition with increasing concentrations of labeled homologous tester [1×(▲), 3×(▼), 5×(◆), and 8×(●)].

shown). These experiments corroborate each other, thereby suggesting a nonreiterative nature of repeat realignment.

Although the lack of competition observed suggests a nonreiterative mode of search as the most plausible and simplest mechanism for RecA-mediated realignment, other more complex scenarios cannot be ruled out. For example, if RecA−DNA complexes that are formed prior to the addition of competitor involve extensive network of aggregates, competitor strands added will not be able to access the target sites during competition, thereby negating the effect of competitor. It must be noted that although RecA−DNA networks are a serious concern in such competition experiments, they are formed only when DNA substrates several hundred base pairs long are used (*34*). The network of RecA−DNA interactions is either of unstable nature or absent when the system involves short DNA substrates (*8*). Therefore, in our experimental system it is very unlikely that the lack of competition observed is simply due to a preclusion of the competitor from such network aggregates.

Repeat Tester Realigns Even When It Is Tethered; Realignment Has No Directional Bias. As discussed above, if realignment of frames between the template and the repeat tester involves sliding of the strands, does there exist a directional bias to it? This was analyzed by use of a tester that was shorter by five nucleotides (A_{25}-tester) compared to its full-length frame of 30 nucleotides. During realignment, the A_{25}-tester has a choice of aligning next to either the left tether or the right tether, depending upon its intrinsic directional bias. We believe that if sliding has such a bias, it might be reflected in the kinetics and/or the final steady-state levels of TL assay signals, as monitored by ligations with left versus right tether (up-tether versus down-tether, Figure 1C).

In this design, the A_{25}-tester is lodged between two tethers, simultaneously on either side. Here we hypothesize that, after aligning with one of the tethers and thus yielding a single ligation product, the tester may subsequently slide over the repeat stretch toward the second tether, thereby generating a double-ligation product, involving a "loop-out" of the template. This supposition is based on the fact that, in pairing across mixed sequences, an equivalent gap precludes ligation, perhaps due to an inability to slide the paired strands [Figure 2A (iii); see Figure 3B in ref *17*]. On the basis of the sizes, one can separately monitor single versus double ligation products in the same sample. Moreover, by placing the label appropriately, one can independently analyze left versus right single ligations that eventually generate a common product of double ligation. We did all these analyses by comparing TL assay time courses in RecA versus thermal reactions. In the first set, the left tether was labeled and the tester was kinased with cold ATP. In the second set, the right tether was labeled while both the left tether and tester was kinased with cold ATP. By this design, we could sequentially study left ligation followed by right ligation in the first set and the reverse of the same, in the second. We also wanted to assess whether the first ligation that generates an anchored tester affects realignment of the repeat at the other end. The results showed that, specifically in RecA reactions, the anchored first ligation product realigned to the other end, perhaps by sliding, thereby revealing a time course of the double ligation product (Figure 4A,B). Whether the ligation was monitored from left to right (first set, Figure 4A,B) or from right to left (second set, Figure 4A,C), the kinetics as well as the final steady-state levels attained were similar.

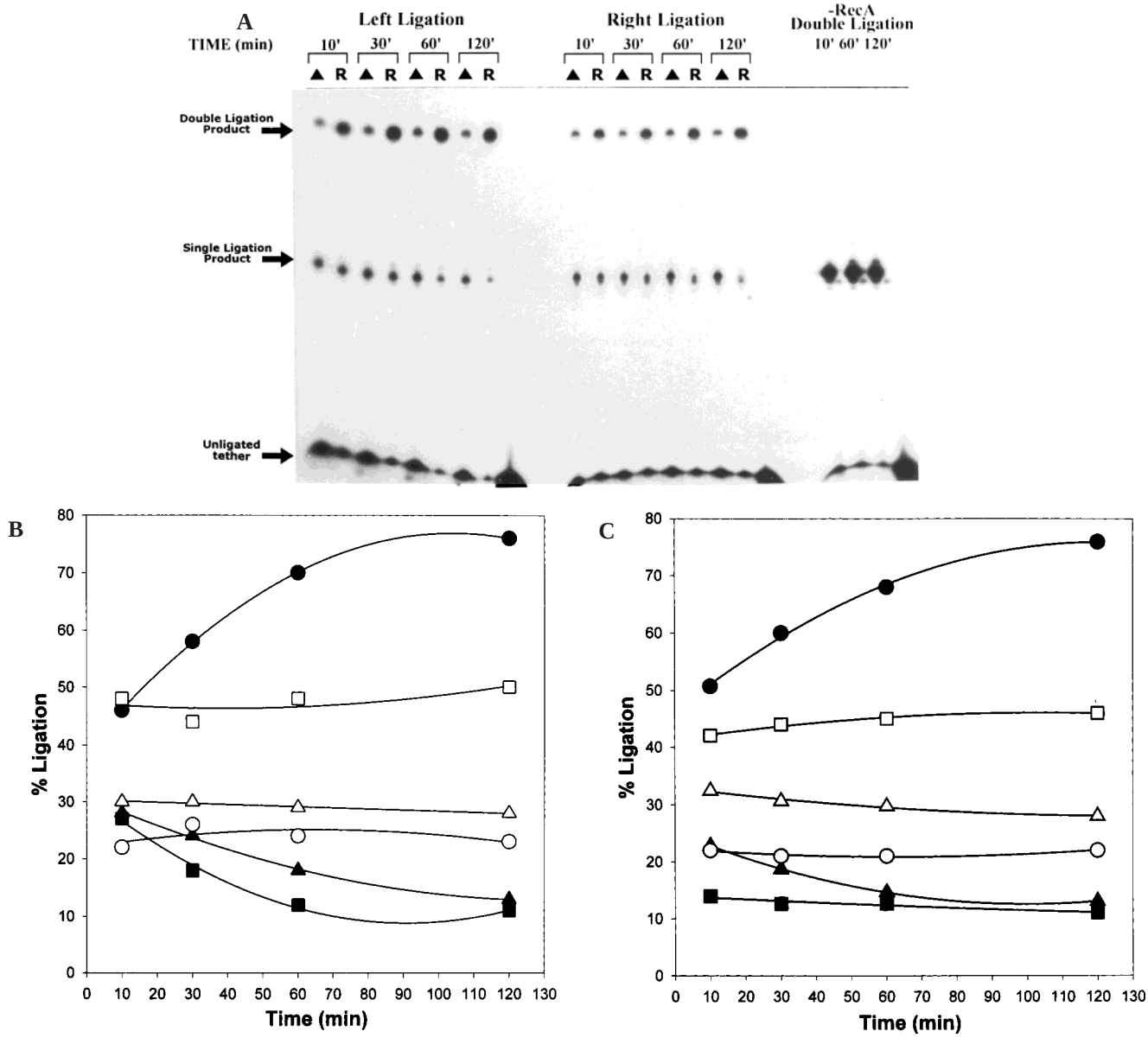

FIGURE 4: (A) Analysis of left versus right ligation events: a test of directionality in strand realignment. TL assays contained up-tether as well as down-tether (2.7 μM each) in each reaction. All other assay conditions were as described (Materials and Methods). Left ligation was monitored by using labeled up-tether, cold-kinased A_{25}-tester, and unkinased down-tether. Right ligation involved cold-kinased up-tether, cold-kinased A_{25}-tester, and the labeled down-tether. Last three lanes pertain to a control set identical to that of RecA reactions (left ligation), that lacked RecA but contained equivalent buffer. Single and double ligation products and free tether are indicated. RecA, R; thermal, ▲. (B, C) Quantitative analyses of left (B) versus right (C) ligations. Single and double ligation products described in part A were quantitated and expressed as a function of time. RecA, solid symbols; and thermal, open symbols. Unligated labeled tether (□, ■); single ligation product (△, ▲); double ligation product (○, ●).

This observation suggested that RecA realignment does not possess a directional bias. No such time-dependent changes were observed in the thermal samples, implying that following stochastic alignments a significant fraction of the paired strands remain "frozen" in suboptimal frames of alignment (Figure 4). A parallel minus-RecA control where the DNA substrates were "passively" annealed (without denaturation and slow cooling) revealed only a single ligation product, thereby indicating the frozen nature of the annealed substrates. These experiments taken together strongly suggest that RecA realigns the repeat tester at the free end despite the other end being covalently tethered into a stable frame. Since gaps cannot be sealed by ligase [Figure 2A (iii)], it becomes obligatory that the tester strand, running short by five nucleotides, realigns by looping out the template strand elsewhere in the repeat region to form a ligatable nick. In fact, the generation of a loop itself probably stems from the ability of the tester to slide over the repeat region toward the ligating end in a drive to maximize alignment of the paired strands.

Distribution of RecA Protein on Template Strand during Tester Realignments: DNase I Protection Assay. The experiments described below were performed to obtain an insight about the distribution of RecA on the putative intermediates that are likely to form during repeat realignments. Several partial duplexes were generated that mimic such intermediates that may ensue during realignment. The T-template labeled at the 5′ end was coated with RecA. Four separate RecA-mediated annealing reactions were initiated, each differing in the combinations of tester and tethers used.

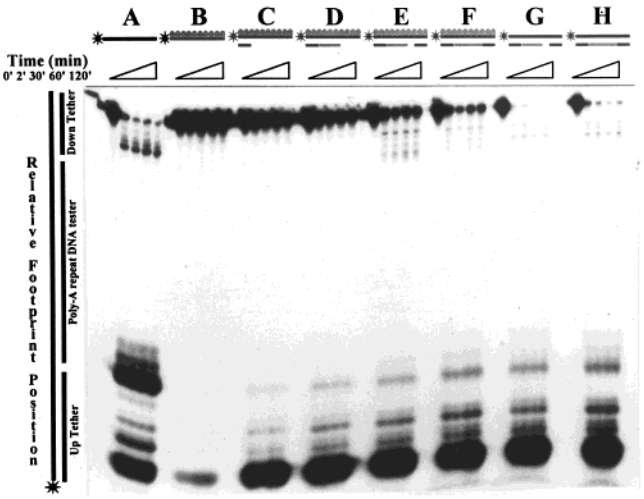

FIGURE 5: Analyses of RecA localization in intermediates of strand realignment: DNase I protection assay of RecA−DNA complexes. Labeled T-template was paired with tester and/or tethers as per TL assay conditions but without ligase. Cartoon (C−H) depicts various annealed products (C−F, RecA annealed; G and H, thermally annealed) where the top strand is the template and the bottom strands are tether (black bars) and/or tester (gray bars). Reactions B−F employ RecA-coated templates. A, G, and H are controls for corresponding RecA reactions B, E, and F, respectively. Reactions E and G employ A_{25}-tester, whereas F and H have A_{30}-tester. Aliquots (5 μL) were withdrawn from each reaction at various time points and subjected to DNase I (0.03125 unit; 2.5 units/μg) digestion for 2 min. Reactions were quenched with 5 μL of 95% formamide-containing indicator dyes and analyzed on a 10% denaturing polyacrylamide gel. The 0 min time point in each set is without DNase I digestion, thus representing the original input of the labeled template strand. A cartoon depiction on the left side indicates the relative location of tether and tester annealing vis-à-vis the zone of DNase I sensitivity on the template strand.

Appropriate thermal annealing reactions were set up as controls. At different time points of annealing, aliquots were withdrawn and subjected to DNase I probing followed by analyses on a denaturing gel (Figure 5). Template strand coated with RecA showed almost complete protection against DNase I, whereas naked template strand was fully sensitive under similar conditions (compare set B with A). A similar level of sensitivity was shown by other naked annealed products as well (G and H). RecA-coated single-stranded template became gradually sensitive to DNase I, specifically in those zones where annealing ensued (C and D). This indicates that RecA can sense the status of homologous pairing, following which it dissociates from the template strand. The same trend was seen when the template was almost completely duplexed on either side by tether (E and F). However, RecA annealing in these samples did not render the template strand as much sensitive as their corresponding naked DNA controls (compare E and F with G and H respectively). These RecA complexes did show a measurable level of full-length protection of labeled template strand (E and F). Such full-length protection was evident at all time points of annealing, suggesting that RecA was still associated with these complexes. Moreover, suboptimal alignments, as exemplified by an annealed duplex that contained a small gap of five nucleotides (set E), showed better full-length protection than the corresponding complete duplex (set F). In addition, a difference in DNase I sensitivity was observed specifically in the template region corresponding to the down-tether of set E and not of set F. The higher steady-state levels of RecA in set E might result in an increased sensitivity to DNase I at the borders of RecA-coated microgaps, in the same way as stable footprints in classical DNase I footprinting assays are usually abutted by flanking regions of enhanced sensitivity. This again reflects on the higher levels of RecA localization in the repeat realignment region. In fact, recent biophysical analyses using the same experimental system suggests that RecA is able to sense small single-stranded DNA gaps and persist onto such suboptimally aligned duplexes, until the realignment is complete (Sen et al., submitted for publication). Thus, the DNase I protection results point out to a property of RecA that senses the unmaximized substrates containing gaps that are as small as even five nucleotides (set E). The repeat region, whether single-stranded or double stranded, was intrinsically more resistant to DNase I (A, G, and H) and hence it was not possible to map RecA distribution in this zone, where the realignment was in progress. Nevertheless, full-length protection of template in sets E and F suggests that a sufficient level of RecA was still lingering on these complexes to facilitate a slow time course of tester realignment.

Realignment Kinetics of Mono-, Di-, and Trinucleotide Repeats Is Identical. If realignment involves sliding, can RecA overcome the energy barriers associated with the sliding of di- and trinucleotide repeats? As depicted in Figure 6, during sliding, every second frame in dinucleotide repeats and every two out of three frames in trinucleotide repeats face an energy barrier due to base mispairing. To obtain an insight of whether RecA can overcome such barriers, we compared realignment kinetics across mono-, di-, and trinucleotide repeats. Three template strands having different repeat regions were used. The repeat, either A_{30} or $(CA)_{15}$ or $(CAG)_{10}$, was flanked on either side by mixed-sequence tethers (Table 1). TL assays were performed with labeled up-tether and unlabeled tester. RecA and thermal annealing reactions were compared in parallel. As observed earlier, thermal annealing did not show a time course for any of the repeats (Figure 6). However, the quantitative levels of TL assay signals were different for each repeat, where the trinucleotide repeat showed the highest signal, followed by di- and mononucleotide repeats. This apparent gradation may be due to the short nature of repeat duplexes used here. Theoretically, if one takes all possible pairing frames into consideration in a 30-mer repeat, mononucleotides, dinucleotides and trinucleotides could potentially misalign in 29, 14, and 9 frames, respectively. If one were to consider a fraction of these frames that lead to stable duplexes, the frequency of misaligned products would gradually decrease from mono- to di- to trinucleotides, thereby reciprocally leading to higher TL assay signals. In contrast, RecA reactions showed clear time courses that were nearly identical for all the three different kinds of repeats, indicating that RecA mediates pairing and realignments across repetitive DNA irrespective of the length of the repeat units (tested here). These results demonstrate the ability of RecA in overcoming the energy barriers associated with unfavorable frames that ensue during sliding.

ATP-Hydrolysis is Required for RecA-Mediated Realignment. RecA filament is active only in the presence of ATP (*1*−*3*). In fact, RecA is one of the most actively ATP-hydrolyzing enzymes when bound to DNA (*35*). We intended to study whether hydrolysis of ATP was required for strand

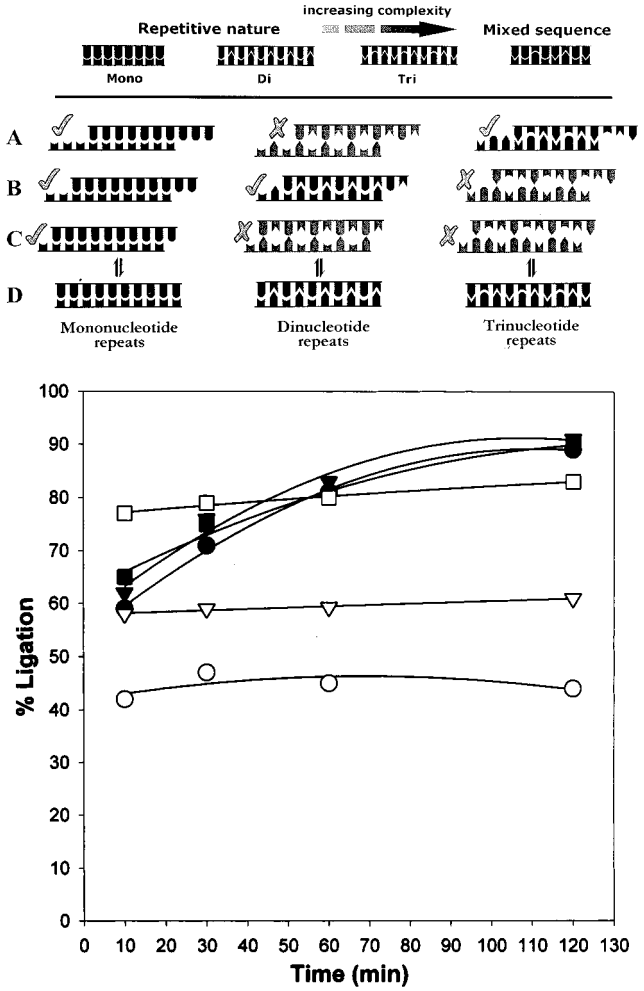

FIGURE 6: (Top) Frames of realignment versus energy barriers associated with sliding of the repeats. Top portion depicts increasing sequence complexity of repeats (mono- < di- < trinucleotide < mixed sequence). In the bottom portion A−D depict sequential frames of alignment during sliding wherein frame (D) represents the maximized frame of alignment. The check marks indicate energetically favorable frames while the cross-marks indicate unfavorable ones. (Bottom) RecA-mediated realignment kinetics of mono-, di-, and trinucleotide repeat testers is similar. Three types of testers, namely, A_{30} (●, ○), $(AC)_{15}$ (▼, ▽), and $(CTG)_{10}$ (■, □), were paired with T-template, GT-template, and CAG-template, respectively. TL assay was done in the presence of a common labeled up-tether. RecA, solid symbols; thermal, open symbols.

realignment activity of RecA that has been described here. We addressed this issue by adopting a protocol wherein an ongoing RecA reaction in ATP is "poisoned" by the addition of a poorly hydrolyzable ATP analogue, namely, ATPγS, followed by the biochemical analyses of RecA (*36*, *37*). Such a protocol was designed earlier to study the dynamics of "RecA treadmilling" and strand-exchange vis-à-vis ATP hydrolysis (*35*, *36*). Since TL assay involves T4 DNA ligase, which in turn requires ATP, the effect of ATPγS on a standard ligation control in this protocol was studied. NAD dependent *E. coli* DNA ligase that does not require ATP was used as a control. For this purpose, we chose the ligation of a nick wherein the mixed-sequence tether and tester strands were juxtaposed to each other on a RecA-coated template strand [Figure 2A (ii)]. Following RecA annealing (in ATP), we "spiked" the reactions with varying levels of ATPγS and analyzed them by TL assay with either T4 DNA ligase or *E. coli* DNA ligase. TL assays involving either of

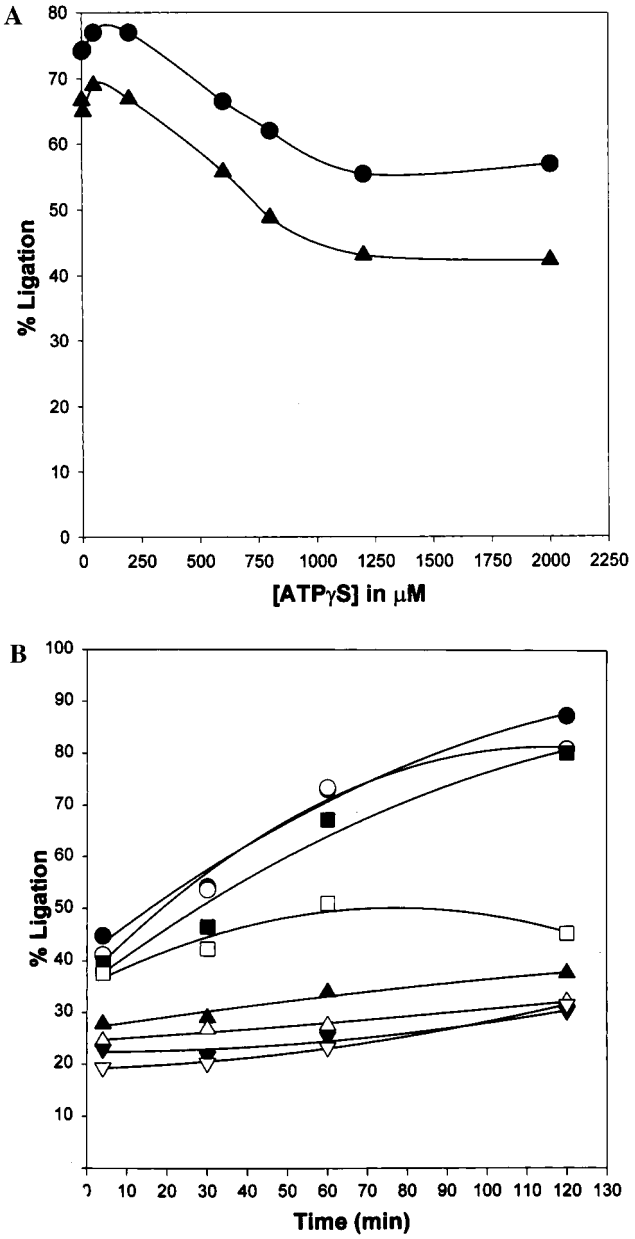

FIGURE 7: (A) Effect of ATPγS on nick ligation. The nick ligation control of RecA pairing across mixed sequences was assayed [as described in Figure 2A (ii)] after the addition of increasing concentrations of ATPγS, following presynapsis in a standard reaction containing 1.2 mM ATP (see Materials and Methods). TL assay was scored at 30 min and plotted as a function of ATPγS concentration [T4 DNA ligase (●) and *E. coli* DNA ligase (▲)]. *E. coli* DNA ligase reaction was performed in the presence of 26 μM NAD at otherwise identical conditions described for TL assay. (B) Effect of ATPγS on realignment kinetics of A_{30}-tester in the TL assay. RecA-mediated realignment of A_{30}-tester was monitored by TL assay with T4 DNA ligase [see Figure 2A (i)]. As described in part A, ATPγS was added following presynapsis. Parallel reactions were done with increasing concentrations of ATPγS. TL assay efficiencies were plotted as a function of time. ATPγS concentrations used were 0 μM (●), 5 μM (○), 50 μM (■), 200 μM (□), 600 μM (▲), 800 μM (△), 1200 μM (▼), and 2000 μM (▽).

the ligases were similarly affected (Figure 7A). There was no effect on TL assay signals until about 200 μM ATPγS, above which the signals dropped gradually. This drop tapered off at 30−40% of the total by about 1.5 mM ATPγS.

We repeated the realignment reaction using A_{30}-tester and labeled up-tether with both T4 DNA ligase and *E. coli* DNA

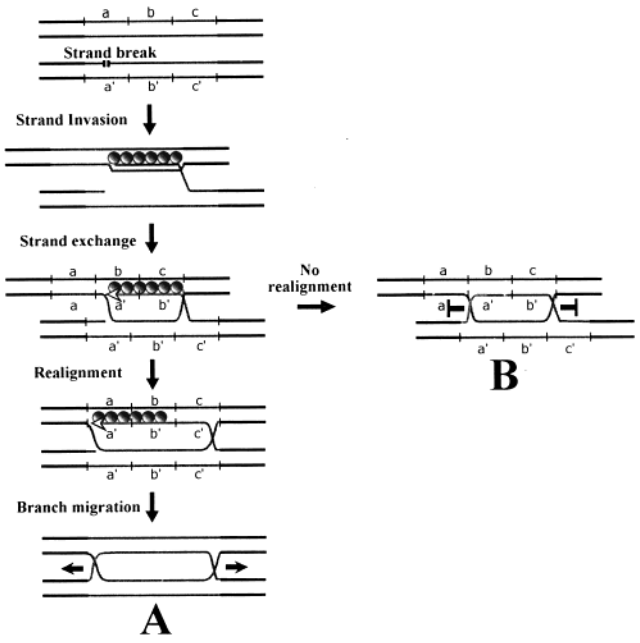

FIGURE 8: Strand realignment relieves a block in the migration of Holliday junctions: a model. Two recombining duplexes have homologous repetitive (gray lines representing repeat units a, b, c) as well as mixed sequences (black lines). A single nick in segment a initiates invasion on to the other duplex at a homologous segment b. Strand exchange without realignment leads to a block in the branch migration (B). RecA-mediated movement of invading end vis-à-vis the rest of the Holliday junction realigns the two duplexes to an optimal frame that can freely branch-migrate (A).

ligase (see Figure 2A). A range of ATPγS concentrations was tested in RecA−ATP-mediated realignment reactions. T4 DNA ligase patterns were similar to those of *E. coli* DNA ligase (data shown only for T4 DNA ligase, Figure 7B). Repeat realignments were found to be more sensitive to the inhibitory effects of ATPγS than the nickligation controls. For example, at a concentration of ATPγS (i.e., 200 μM) where the nickligation was unaffected, repeat realignment almost came to a halt (compare effects in Figure 7 panel A versus panel B). At higher concentrations of ATPγS, all repeat realignments virtually stopped whereas the nick ligations were only partially affected. These results taken together indicate that realignments catalyzed by RecA are sensitive to ATPγS, thereby suggesting an involvement of ATP hydrolysis in maximization.

DISCUSSION

If the very end of a RecA-coated invasive single strand were to contain an entire stretch of microsatellite repeats, homology search is unlikely to pair this with the recipient duplex into an "optimal frame". An optimal frame is defined as one that aligns repeats such that the adjoining mixed sequences automatically fall into a homologous register. Any alignment that misses this optimal frame leads to an abrogation of branch migration at the junctions of the mixed sequences and repeats, due to a heterologous register (Figure 8). Misalignment errors are unlikely to be rectified during homology search steps that are believed to be fast and diffusion controlled (*7*−*9*, *11*). Hence we surmised that such errors might be rectified at the next step in the pathway, namely, strand exchange. Since strand exchange in the repeat region proceeds normally until it is arrested at a mixed-sequence junction, a release of the block is possible only if the repeat strand is realigned in the heteroduplex (Figure 8). Such a process brings back the recombining duplexes into homologous register and thus branch migration ensues freely. We tested this hypothesis in this paper.

The TL assay was primarily designed to quantitatively monitor RecA-mediated pairing across different kinds of mismatches (*17*). When pairing ensues across repeat regions, TL assay specifically scores optimally aligned frames (i.e., ligatable, c and d, Figure 1) among a pool of several other stochastically paired suboptimal frames (a and b, Figure 1). In appropriate controls, we demonstrated that the assay was equally efficient in scoring nick ligations of RecA as well as thermally annealed duplexes, involving mixed sequences [Figure 2 parts A (ii) and C]. Besides, the assay was also accurate in avoiding spurious ligations across gaps and heterologous pairings [Figure 2A (iii); see also Figure 3 in ref *17*]. Using such an assay, we tested the "repeat realignment" hypothesis described above. Here, we demonstrate that a RecA filament that is paired with a repeat strand does indeed realign suboptimally paired frames in a slow and time-dependent manner (Figures 2, 4, and 6). In addition, time-resolved fluorescence experiments with a labeled repeat DNA probe showed similar kinetics of strand realignment specifically in RecA pairing reactions (Sen et al., submitted for publication). On the other hand, fluorescence changes associated with thermal pairings were rather abrupt and did not show any such kinetics. Moreover, RecA-annealed duplexes containing gaps of ssDNA as short as five nucleotides, impart better protection against DNase I than fully duplexed controls (Figure 5). This result indicates that suboptimally paired frames retain higher levels of RecA than fully maximized products, thereby suggesting an active role of RecA filament in repeat realignments.

Although RecA strand invasions preferentially happen by 3′-ended presynaptic filaments, it can also be mediated by 5′ ends (*38*−*40*). Hence, it seems likely that RecA should rectify misalignments resulting from both 3′- as well as 5′-ended invasions, failing which branch migration might be affected (Figure 8). Our in vitro observations, which show no directional bias in realignment, fit well with such a requirement (Figure 4).

Strand realignments in a heteroduplex tract will be energetically expensive if they involve reiterative rounds of association and dissociation. An alternative mechanism that involves no measurable dissociation of strands invokes a "sliding" motion between the repeats. Such a sliding motion might involve either a continual movement of the strands or a "caterpillar-like" crawling of a microloop in one strand vis-à-vis the other strand via a process analogous to branch migration (*21*, *41*). The competitor-chase experiment did indeed point to such a mechanistic possibility (Figure 3). In fact, sliding may be facilitated by continual RecA treadmilling, followed by repeated extension/relaxation of the helix along the duplex, that results in a caterpillarlike motion of one strand vis-à-vis the other. One must note that although various models have been invoked that involve either one -or three-dimensional modes for the initial steps of homology search (*7*−*9*, *12*), the realignment process studied here refers to a distinct and hitherto undescribed post-homology-search event leading to maximization of homology.

How does RecA overcome the energy barriers associated with the sliding motion across di- and trinucleotide repeats? During sliding across such repeats, every other frame in dinucleotides and every two out of three frames in trinucleotides would face an energy barrier (Figure 6). Identical kinetics observed for realignments across mono-, di-, and trinucleotide repeats suggests that RecA has an efficient mechanism of overcoming these energy barriers (Figure 6). Since RecA houses 3−4 nucleotides (at site I) and their complementary bases (at site II) (*1*−*3*) in a single monomer (perhaps in the same binding domain), it is possible to invoke a cross-talk between the bases across the paired stacks (*42*, *43*). We hypothesize that unstacked bases from two different strands that share a single monomer may mediate a "hydrogen bonding cross-talk" between a base at site I with any other base at site II. Such a cross-talk may constantly reassure "homology register" and thereby mediate a dynamic "slippage" between repeat sequences, to completion. However, as RecA encounters increasingly complex repeats tending toward mixed sequences, such slippages get abrogated because of high energy barriers that flank the original "register" (Figure 6). This failure to mediate slippage probably signals the end of homologous alignment in mixed sequences and leads to an eventual dissociation of RecA from the duplex (Figure 5).

RecA−ssDNA filament can bring about homologous pairing and base switching even in the absence of ATP hydrolysis [i.e., when a poorly hydrolyzable ATP analogue (ATPγS) is present] (*44*,*45*). Various models propose that treadmilling of the protein and/or the concerted rotation of DNA helices within the RecA filament are steps that are likely to consume the energy of ATP hydrolysis (*2*). Although strand realignment activity described here requires ATP hydrolysis, the relationship between the above models and strand realignment needs further investigation. Recent studies by Nishinaka et al. (*46*) have suggested a model for RecA−dsDNA filament where ATP hydrolytic cycles are coupled to interconversion of sugar puckers between N and S type that cause bases to rotate. When such rotation of bases ensues in a duplex containing repetitive tracts, it might result in a constant realignment of their pairing frames, mediating a slippage between the complementary strands. Substantial affinity differences existing between the two putative DNA binding sites on a single monomer (*1*−*3*) may impart a relative fluidity between the paired strands and mediate sliding through a process that requires ATP hydrolysis. These slippages promoted by various RecA monomers ought to synchronize by some means to finally manifest the strand realignment observed in TL assay. Perhaps it is at this stage that the process requires energy input and encounters rate-limiting steps, the former leading to cessation of realignment in ATPγS and the latter to the slowness of the process.

Very few reports in the literature have addressed how RecA brings about recombination in microsatellite regions. In a recent study by Dutreix (*47*) it was demonstrated that strand exchange slows down markedly as it encounters microsatellite sequences. As expected, this effect was directly proportional to the total length of the repeat. We believe that these results are indicative of the intrinsically slow steps of repeat realignments that precede a productive strand exchange. Our results directly point out to such a mechanistic possibility. However, such slow kinetics is not compatible with the expected rates in vivo. The in vitro model described here depends entirely on a single protein, RecA, whereas in vivo the same is likely to be assisted by several helper proteins such as helicases, RecO, RecR, and SSB (*48*−*50*). Such a concerted action might result in better rates in vivo.

ACKNOWLEDGMENT

We thank Professor K. V. R. Chary of the Department of Chemical Sciences Tata Institute of Fundamental Research (TIFR), for critically reading the manuscript. We thank the Department of Atomic Energy (TIFR) for financial support. We thank Instrumentation In-charge at the Cancer Research Institute (CRI, Mumbai) for the usage of the Phosphor Imager. We would also like to thank Mr. Jatin and Mr. Vijay from the photography section of TIFR for their help.

REFERENCES

1. Radding, C. M. (1991) *J. Biol. Chem. 266*, 5355−5358.
2. Roca, A. I., and Cox, M. M. (1997) *Prog. Nucleic Acid Res. Mol. Biol. 56*, 129−223.
3. Kowalczykowski, S. C., and Eggleston, A. K. (1994) *Annu. Rev. Biochem. 63*, 991−1043.
4. Flory, J., and Radding, C. M. (1982) *Cell 28*, 747−756.
5. Egelman, E. H., and Stasiak, A. (1986) *J. Mol. Biol. 191*, 677−697.
6. Nishinaka, T., Ito, Y., Yokoyama, S., and Shibata, T. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 6623−6628.
7. Gonda, D. K., and Radding, C. M. (1983) *Cell 34*, 647−654.
8. Gonda, D. K., and Radding, C. M. (1986) *J. Biol. Chem. 261*, 13087−13096.
9. Julin, D. A., Riddles, P. W., and Lehman, I. R. (1986) *J. Biol. Chem. 261*, 1025−1030.
10. Yancey-Wrona, J. E., and Camerini-Otero, R. D. (1995) *Curr. Biol. 5*, 1149−1158.
11. Bazemore, L. R., Takahashi, M., and Radding, C. M. (1997) *J. Biol. Chem. 272*, 14672−14682.
12. Adzuma, K. (1998) *J. Biol. Chem. 273*, 31565−31573.
13. Bazemore, L. R., Folta-Stogniew, E., Takahashi, M., and Radding, C. M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 11863−11868.
14. DasGupta, C., and Radding, C. M. (1982) *Nature 295*, 71−73.
15. Malkov, V. A., Sastry, L., and Camerini-Otero, R. D. (1997) *J. Mol. Biol. 271*, 168−177.
16. Malkov, V. A., and Camerini-Otero, R. D. (1998) *J. Mol. Biol. 278*, 317−330.
17. Karthikeyan, G., Wagle, M. D., and Rao, B. J. (1998) *FEBS Lett. 425*, 45−51.
18. Farber, R. A., Petes, T. D., Dominska, M., Hudgens, S. S., and Liskay, R. M. (1994) *Hum. Mol. Genet. 3*, 253−256.
19. Pearson, C. E., and Sinden, R. R. (1998) *Curr. Opin. Struct. Biol. 8*, 321−330.
20. Djian, P. (1998) *Cell 94*, 155−160.
21. Karthikeyan, G., Chary, K. V., and Rao, B. J. (1999) *Nucleic Acids Res. 27*, 3851−3858.
22. Gordenin, D. A., Kunkel, T. A., and Resnick, M. A. (1997) *Nat. Genet. 16*, 116−118.
23. Freudenreich, C. H., Kantrow, S. M., and Zakian, V. A. (1998) *Science 279*, 853−856.
24. Morel, P., Reverdy, C., Michel, B., Ehrlich, S. D., and Cassuto, E. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 10003−10008.
25. Strand, M., Prolla, T. A., Liskay, R. M., and Petes, T. D. (1993) *Nature 365*, 274−276.
26. Strand, M., Earley, M. C., Crouse, G. F., and Petes, T. D. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 10418−10421.
27. Shibata, T., Cunningham, R. P., and Radding, C. M. (1981) *J. Biol. Chem. 256*, 7557−7564.
28. Beattie, K. L., Wiegand, R. C., and Radding, C. M. (1977) *J. Mol. Biol. 116*, 783−803.

29. Hsieh, P., Camerini-Otero, C. S., and Camerini-Otero, R. D. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6492−6496.
30. Adzuma, K. (1992) *Genes Dev. 6*, 1679−1694.
31. Brenner, S. L., Mitchell, R. S., Morrical, S. W., Neuendorf, S. K., Schutte, B. C., and Cox, M. M. (1987) *J. Biol. Chem. 262*, 4011−4016.
32. Bianco, P. R., and Weinstock, G. M. (1996) *Nucleic Acids Res. 24*, 4933−4939.
33. Register, J. C., III and Griffith, J. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 624−628.
34. Tsang, S. S., Chow, S. A., and Radding, C. M. (1985) *Biochemistry 24*, 3226−3232.
35. Cox, M. M. (1994) *Trends Biochem. Sci. 19*, 217−222.
36. Lindsley, J. E., and Cox, M. M. (1989) *J. Mol. Biol. 205*, 695−711.
37. Lindsley, J. E., and Cox, M. M. (1990) *J. Biol. Chem. 265*, 9043−9054.
38. DasGupta, C., Wu, A. M., Kahn, R., Cunningham, R. P., and Radding, C. M. (1981) *Cell 25*, 507−516.
39. Konforti, B. B., and Davis, R. W. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 690−694.
40. Dutreix, M., Rao, B. J., and Radding, C. M. (1991) *J. Mol. Biol. 219*, 645−654.
41. Hsieh, P., and Panyutin, I. G. (1995) *Nucleic Acids Mol. Biol. 9*, 42−65.
42. Aymami, J., Coll, M., Frederick, C. A., Wang, A. H., and Rich, A. (1989) *Nucleic Acids Res. 17*, 3229−3245.
43. Coll, M., Frederick, C. A., Wang, A. H., and Rich, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8385−8389.
44. Menetski, J. P., Bear, D. G., and Kowalczykowski, S. C. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 21−25.
45. Rosselli, W., and Stasiak, A. (1990) *J. Mol. Biol. 216*, 335−352.
46. Nishinaka, T., Shinohara, A., Ito, Y., Yokoyama, S., and Shibata, T. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 11071−11076.
47. Dutreix, M. (1997) *J. Mol. Biol. 273*, 105−113.
48. Harmon, F. G., and Kowalczykowski, S. C. (1998) *Genes Dev. 12*, 1134−1144.
49. Shan, Q., Bork, J. M., Webb, B. L., Inman, R. B., and Cox, M. M. (1997) *J. Mol. Biol. 265*, 519−540.
50. Webb, B. L., Cox, M. M., and Inman, R. B. (1997) *Cell 91*, 347−356.

BI000753Y