Non-Watson-Crick base pairs modulate homologous alignments in RecA pairing reactions

G. Karthikeyan, Mahendra D. Wagle, Basuthkar J. Rao*

Molecular Biology Unit, Tata Institute of Fundamental Research, Colaba, Bombay-400 005, India

Received 31 December 1997; revised version received 16 February 1998

Abstract Complementary pairing by RecA was examined in vitro to investigate how homology is deciphered from non-homology. Somewhere in a window of 40–50% sequence complementarity, RecA pairing begins to manifest the specificity of homology. Quantitation reveals a hierarchy among non-Watson-Crick mispairs: RecA reaction treats six out of 12 possible mispairs as *good* ones and three each of the remaining ones as *moderate* and *bad* pairs. The mispairs seem to function as independent pairing units free of sequence context effects. The overall strength of pairing is simply the sum of the constituent units. RecA mediated gradation of mispairs, free of sequence context effects, might offer a general thumb-rule for predicting the pairing strength of any alignment that carries multiple mispairs.

Key words: Homology search; Mismatch; Recombination; RecA

1. Introduction

Early steps of recombination involve a genome-wide search for homology by the initiator strand of DNA. It is believed that in vivo, RecA-like proteins coat the initiator single strands to catalyze such a search [1]. *Escherichia coli* RecA is one of the best studied pairing proteins and is likely to represent a universal paradigm for DNA pairing rules [2]. In this paper, we study how RecA protein starts to discriminate homology from non-homology and the role non-canonical base pairs play in such a step. We chose complementary pairing for this study since complementarity in base pairing is fundamental to the specificity of recognition between two strands of nucleic acid [3]. It is very likely that complementary recognition plays a major role when the RecA-coated initiator strand synapses with negatively supercoiled duplexes leading to D-loops [4–6]. In vivo, chromosomal DNA retains a high level of negative superhelicity at a steady state which makes D-loop mediated complementary pairing significant [7]. In addition, models of homologous pairing have been described that involve annealing of single strands between two different parental molecules (single strand annealing model) [8,9]. Studies on the annealing reaction promoted by *E. coli* RecA, a universal prototype of a pairing enzyme, might also offer fundamental insights into the basic rules of pairing applicable to the SSA model. Here we address the role of RecA catalyzed promiscuous base pairing that might guide the early steps of homology search. The experimental system was designed to quantitatively monitor targeted pairing in its native conditions so that even highly dynamic reaction intermediates housing multiple mispairs could be effectively captured. Such an analysis makes an important beginning to assessing the quantitative effects of all different mispairs in the dynamic pairing states of complementary recognition.

RecA-ss-DNA complex is unusual in DNA conformation with its almost completely unstacked string of DNA bases [10]. A strong prediction from such a structure is that constraints imposed on chemistry of base pairing due to base stacking in naked DNA helix might significantly change. This might open up new rules of base pairing that are more RecA-specific. The results described here reveal that RecA-based DNA recognition in complementary pairing undoes sequence context effects and manifests a high level of mispair tolerance. Quantitation of these targeted pairings reveals an intrinsic hierarchy prevalent amongst the non-canonical pairs. These mispairs essentially function as separable individual modules of pairing units and the overall strength of a pairing reaction could simply be the sum of the constituent units.

2. Materials and methods

2.1. Materials

RecA protein was purified as described [11]. T4-polynucleotide kinase and T4-DNA ligase were from Amersham Life Science. Adenosine triphosphate, phosphocreatine, creatine phosphokinase, dithiothreitol and nuclease-free BSA were from Sigma.

2.2. DNA substrates

Single-stranded oligonucleotides were synthesized by an Applied Biosystems DNA synthesizer at the Keck Biotechnology Resource Laboratory at Yale. All oligonucleotides were further purified by electrophoresis in 10% polyacrylamide gels containing 6 M urea. Care was taken to maintain the temperature of the gel between 65°C and 70°C to prevent the formation of secondary structures. The band corresponding to the full length oligomer was located by shadowing the gel with UV light, following which the oligomer was eluted by diffusion from thin slices of gel overnight in 2–3 ml of 20 mM Tris-HCl (pH 7.5) containing 1 mM EDTA. The sample of oligonucleotide was subsequently desalted by passing through Sep-pak C18 cartridge [12]. The purity of all oligomers was judged by ^{32}P labeling of a small portion by T4-polynucleotide kinase, followed by analysis on a 12% polyacrylamide sequencing gel.

2.3. End-labeling of oligonucleotides

The standard protocol was used to phosphorylate 5′ ends with either cold ATP (1 mM) (on tester oligomers) or γ-^{32}P-ATP (10 μCi) (on tether oligomers) in 5 μl reactions containing 100 μM (total nucleotide concentration) of oligomer. Subsequently the sample was diluted to 50 μl and heated at 70°C for 10 min to heat-inactivate T4-polynucleotide kinase. Standard reaction conditions and assays are as described in the figure legends.

*Corresponding author. Fax: (91) (22) 215 2110/81.
E-mail: BJRAO@TIFRVAX.TIFR.RES.IN

PII S0014-5793(98)00195-1

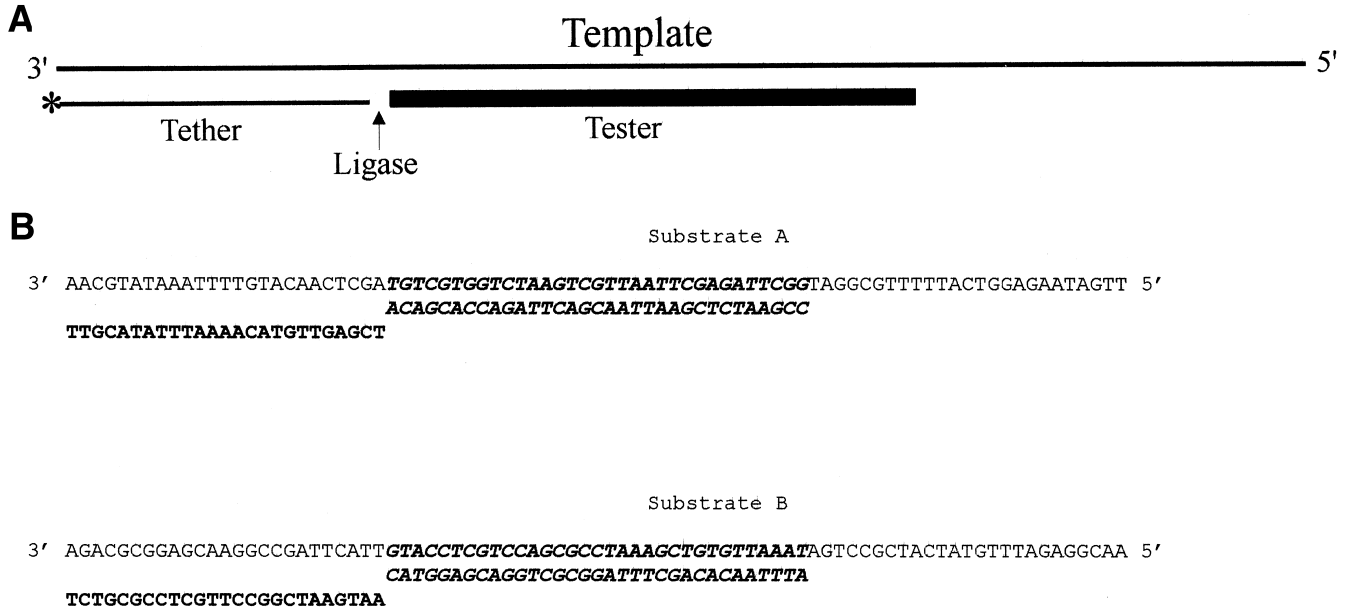

Fig. 1. A: Schematic representation of targeted pairing assay. Asterisk at the 5′ end denotes ^{32}P label. Tester (25-mer) and tether (33-mer) oligonucleotides lie immediately adjacent to each other and are ligatable on pairing with template (83-mer) oligonucleotide. B: Two substrates used in the experiments. The template in both cases is the negative strand sequence of M13 duplex genome (locations 182–265 for substrate A; 1082–1165 for substrate B) and represents approximately equal distribution of purines and pyrimidines in the region complementary to the tester. Tether and tester strands are complementary and antiparallel to the template strand.

3. Results

3.1. Recognition system and targeted ligation assay

To elucidate the effects of specific sets of non-Watson-Crick base pairs, targeted recognition was studied which was monitored as a ligatable alignment between a tester and a reference tether immediately upstream of it (Fig. 1A). Two different substrates were studied to elucidate sequence context effects, if any (Fig. 1B). In each case, RecA was coated on an 83-mer which is long enough to promote RecA binding in the presence of ATP [13,14]. The tether is a 25-mer, fully complementary to one end of the 83-mer. The tester is a 33-mer which carried base substitutions that reduced Watson-Crick complementarity between pairing substrates. In all the assays, unless otherwise indicated, the tether was labeled at the 5′ end with ^{32}P and the tester was kinased with ATP (non-radioactive). In this assay, any other tester alignment that is out of frame with respect to the tether will not be scored as the ligatable ends will not juxtapose with each other. The assay focuses directly on the strengths of precisely aligned pairs with one caveat. The caveat has to do with the effects of sequence bulges within a pairing alignment and the consequent mispairs. However, such effects seem unlikely in the sequences analyzed here (see Section 4). All oligomers used in these experiments were full length and phosphorylatable at the 5′ end to similar efficiencies (data not shown). RecA and thermal pairings were compared under identical conditions of reaction. The ligation product was scored as a labeled 58-mer on a denaturing polyacrylamide gel and quantitated by phosphorImager analysis. When fully homologous (complementary) to the 83-mer, essentially stoichiometric levels of tester were ligatable (Figs. 2 and 3). The time course of ligation revealed that within about 5–10 min, ligase captured almost all paired molecules. The specificity of ligase capture was demonstrated when a tester of random sequence yielded essentially no product even after 2 h of ligation. Additional negative controls revealed no ligation when the tether was shifted upstream with respect to a homologous tester or when template strand was omitted (bars A and B, Fig. 3).

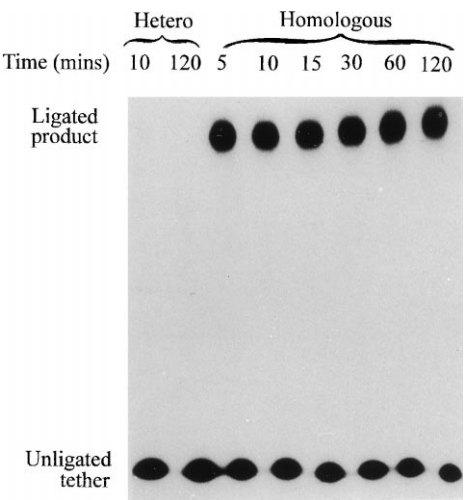

Fig. 2. Targeted pairing assay. Substrate A (see Fig. 1B) was used in homologous reactions. In heterologous control, tester was from substrate B. Template (6 μM) was incubated with RecA (3 μM) in a reaction containing 33 mM Tris-HCl (pH 7.5), 1.2 mM magnesium acetate, 2 mM DTT, 1.2 mM ATP, 8 mM creatine phosphate, creatine phosphokinase (10 U/ml) and BSA (100 μg/ml) for 15 min at 37°C followed by pairing with tester (3 μM) (phosphorylated at 5′ with cold ATP, see Section 2) and 5′ ^{32}P-labeled tether (2.5 μM) and incubated further for 15 min. Ligation was done by adding magnesium acetate (10 mM) and T4 DNA ligase (150 U/ml) followed by incubation at room temperature. Aliquots were withdrawn at various time points as indicated, made to 20 mM EDTA, 0.2% SDS and denatured by adding an equal volume of formamide sequencing buffer. The samples were then analyzed on a 10% polyacrylamide denaturing sequencing gel.

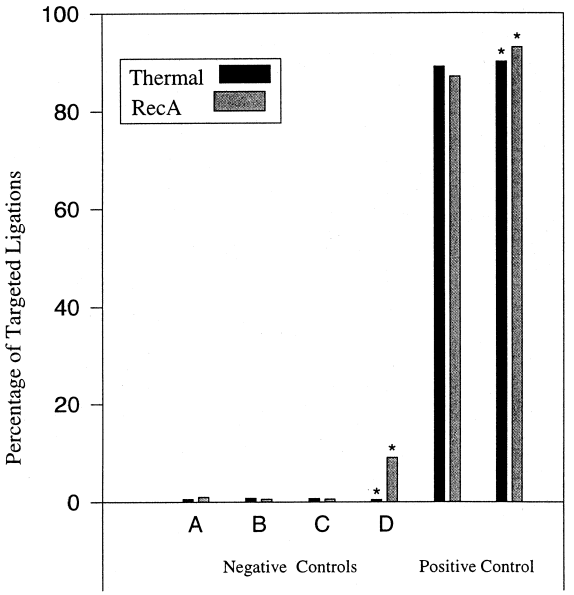

Fig. 3. PhosphorImager (Molecular Dynamics) quantitation by Image Quant software of negative and positive controls in the targeted pairing assay. RecA pairing reactions were as described in the legend to Fig. 2, except that template strand was 10 μM. Thermal annealing was done by incubating the template (10 μM), tether (2.5 μM) and tester (3 μM) strands in 33 mM Tris-HCl (pH 7.5), 1.2 mM magnesium acetate at 90°C for 3 min followed by slow cooling to room temperature. Subsequently, thermal annealing samples were made to the same buffer composition as that of RecA reactions by adding DTT, ATP, creatine phosphate, creatine phosphokinase and BSA. All ligations were done for 10 min and processed as described in the legend to Fig. 2. Percentage targeted ligation was expressed as a ratio of PhosphorImager counts associated with ligated product to the total input radioactivity (ligated product+unligated tether). First bars in all pairs refer to thermal annealing and the second ones to RecA pairings. The first set in positive controls is from substrate B and the second (with the asterisk) from substrates A. Negative controls were: (A) everything except template in substrate A; (B) tether shifted 5 nucleotides upstream to leave a gap between tester and tether in substrate A; (C) template and tether from substrate B with tester from substrate A; (D) template and tether from substrate A and tester from substrate B.

3.2. Base substitutions

In a transient alignment between two non-homologous strands of DNA, stochastically 25% of base pairs would be of the Watson-Crick type distributed randomly amongst 75% of non-Watson-Crick base pair combinations. Such random alignments are too transient to be captured by the ligase assay as shown in the above controls. An alignment that resulted in 40% Watson-Crick pairs was unligatable in one case and marginally so in the other for RecA reactions (bars C and D, Fig. 3). To gain insight into how 'homology' is deciphered in a plethora of 'non-homology', a minimum threshold of Watson-Crick pairs that markedly discriminates between the two needs to be examined. Since we had observed a basal level of ligation with 40% Watson-Crick base pairs (bar D, Fig. 3), we decided to further increase the percentage of Watson-Crick contribution in a rational manner. Since both substrate systems (Fig. 1B) had testers with an equal preponderance of purines and pyrimidines, by substituting either all purines or all pyrimidines to other bases we achieved a 50% Watson-Crick contribution. Four sets of rational substitutions were done which included transitions and transversions (Table 1). The resultant mispairs are: GT+AC, GA+AG, TG+CA and TC+CT where the first base of the mispair is from the 83-mer and the second from the tester. To our surprise, even with 50% mispairs, most of the ligation reactions went nearly to completion in either substrate, the strong exceptions being the reactions containing CT+TC mispairs (Fig. 4A). In RecA and thermal annealings, GT+AC mispairs were almost as proficient as wild type in both substrates. Important differences were seen in TG+CA and GA+AG mispairs due to sequence context effects which were undone by RecA (see Section 4). Since our main interest was to study the rules of RecA-mediated pairings and since non-RecA (thermal) control annealings seem be heavily subject to sequence context effects, we did not study thermal controls further in this paper. More extensive quantitative evaluations of RecA versus thermal annealings across several sequence backgrounds are under way and will be discussed elsewhere due to space constraints here.

All the pairing reactions described in this set were scored by ligase for 2 h. In order to check whether this was a cumulative response, we examined a shorter time of ligation. Since a wild type reaction goes almost to completion within 5–10 min (see Fig. 2), we chose 10 min as the shortest duration for optimal ligation in a set that was less prone to sequence context effects (substrate B). Of the three pairings that worked with almost full efficiency in a 2 h assay, the one carrying TG+CA mispairs showed reduced ligation in a 10 min score (Fig. 4B). Thus by ligating for a shorter period, it was possible to further discriminate between pairing efficiencies (see Section 4).

3.3. Analysis of individual mispairs

We wanted to assess the relative efficiencies of the 12 differ-

Table 1
Rational double changes in tester to get 50% non-Watson-Crick mispairs

Sr. No.	Change	Mispair
1	C→T	GT
	T→C	AC
2	C→A	GA
	T→G	AG
3	A→G	TG
	G→A	CA
4	A→C	TC
	G→T	CT

The second column represents the changes made in the wild type testers (shown in Fig. 1B) so as to get the non-Watson-Crick mispairs represented in the third column. The specified changes are made in the entire length of the tester. The first base in the mispair is from the template and the second from the tester.

Table 2
Single changes in the tester of substrate B to get 25% non-Watson-Crick mispairs

Sr. No.	Change	Mispair
1	A→G	TG
2	A→C	TC
3	A→T	TT
4	G→A	CA
5	G→T	CT
6	G→C	CC
7	C→T	GT
8	C→A	GA
9	C→G	GG
10	T→C	AC
11	T→G	AG
12	T→A	AA

All details are as in Table 1.

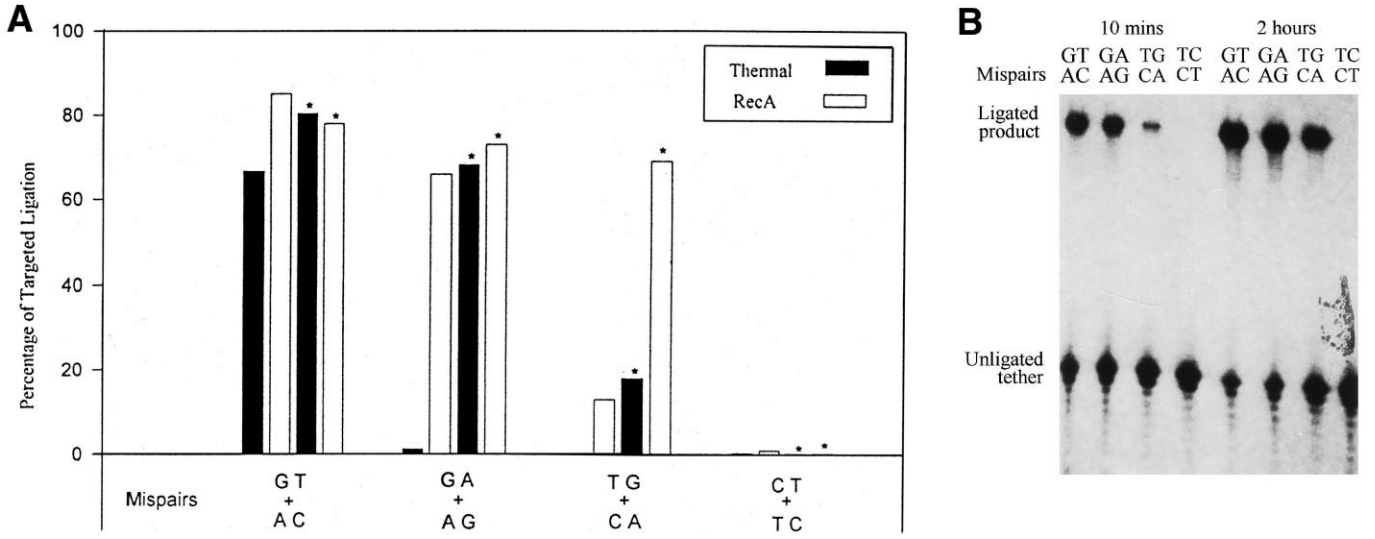

Fig. 4. A: PhosphorImager analysis of targeted pairings in double changes. RecA and thermal annealings were as described in Fig. 3. Ligations were done for 2 h which were followed by gel analysis and quantitation. Base changes and notations of mispairs were as described in Table 1. The data bars with asterisks are from reactions on mismatches derived from substrate A (see Fig. 1B) whereas those without asterisk are from substrate B. B: Comparison of short versus long times of ligation with RecA pairings of reactions containing double changes in substrate A. Details are as described in Fig. 3.

ent non-Watson-Crick base pairs in RecA pairings. Tester oligomers, based on substrate B, were made in which a specific base was substituted at all locations (Table 2). We did this in substrate B as it has equal preponderance of all four bases. Such substitutions led to pairings in which about 25% of base pairs were converted to the non-Watson-Crick type. All pairings were done in RecA reaction conditions and followed by ligation for 10 min. PhosphorImager quantitation of targeted ligation revealed that all mispairs could be grouped under three categories: *good, moderate* and *bad*. *Good* mispairs (TT, TG, GT, AA, AG and AC) are almost as efficient as wild type whereas the *bad* ones (TC, CA, and CC) show a relative strength that was less than 50% and *moderate* ones (CT, GA and GG) ranked in between (Fig. 5). We surmised

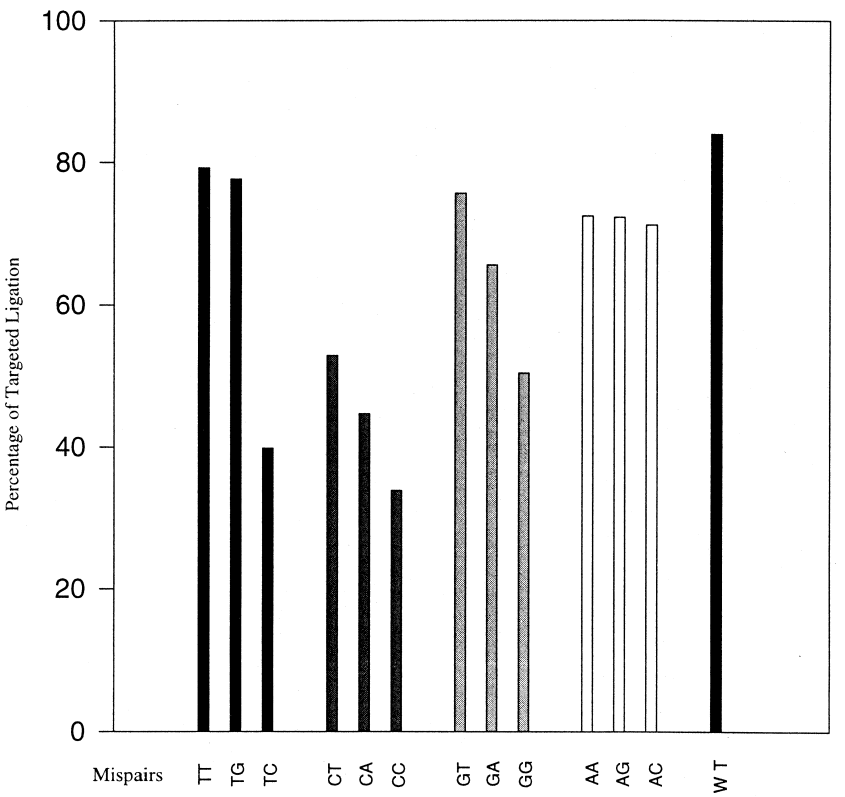

Fig. 5. PhosphorImager analysis of targeted pairings in single changes. All are RecA pairings (as described in Fig. 3) of reactions containing single changes in substrate B. Ligations were done for 10 min. Base changes and notations of mispairs were as described in Table 2.

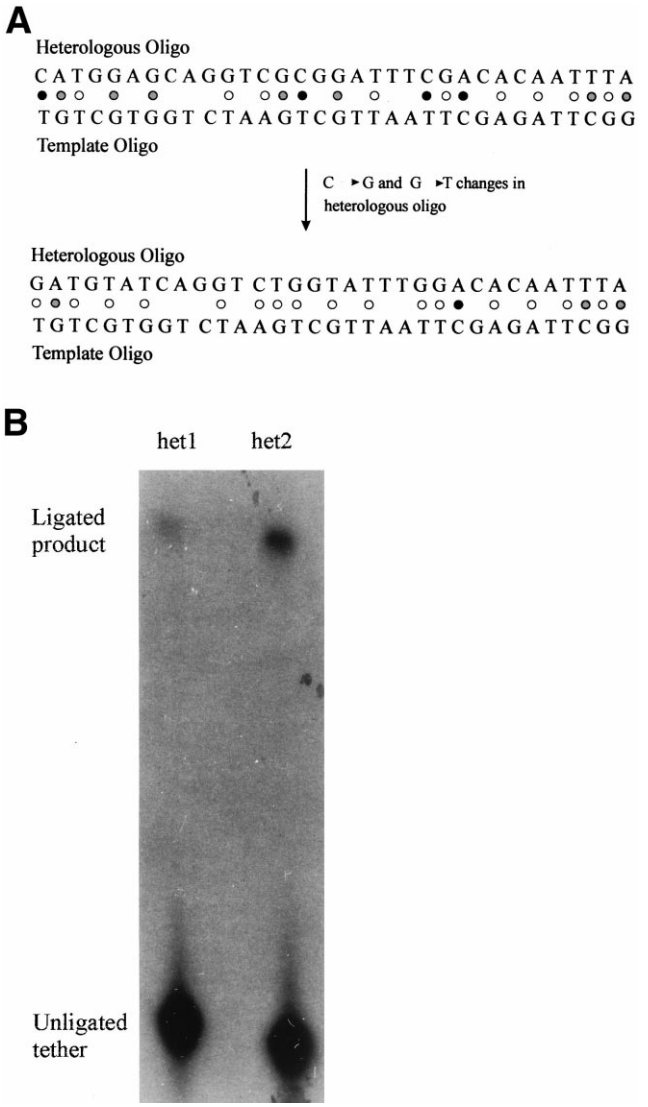

Fig. 6. A: A test on base replacements. The heterologous alignment described in Fig. 3 (second bar in D, a negative control) was tested for effect of base replacements on pairing. Top panel shows the distribution of three grades of mispairs (9 *good* (white circles), 7 *moderate* (gray circles), 4 *bad* (black circles)) before the test replacements were introduced. Following C to G and G to T replacements, as shown in bottom panel, the distribution of mispairs changes (16 *good* (white circles), 3 *moderate* (gray circles), 1 *bad* (black circles)). B: Targeted pairing assay for alignments described in A. RecA pairings were ligated for 2 h as described earlier. Het 1 is the reaction from original heterologous tester and Het 2 is from the modified tester.

that such a hierarchy might be independent of sequence context effects and may offer a rational basis for predicting pairing stability in homologous alignment (see Section 4). This was found to be true on dissecting the composition of the two non-homologous controls and grouping them under these categories (bars C and D in Fig. 3). Of the two controls, one paired better (9%) than the other (0.5%) even though both had the same number of Watson-Crick base pairs (13 out of 33) and *good* mispairs (9 out of 33) (Fig. 5). However, the control that worked better had a higher number of *moderate* mispairs than the other (7 vs. 3) and reciprocally had fewer number of *bad* mispairs (4 vs. 8). Thus the relative pairing strengths of these two controls could be rationalized by the hierarchical strengths of resident mispairs.

We checked the general predictive nature of the proposed hierarchy by a direct test. We asked whether a given pairing alignment could be improved by merely replacing a set of *moderate/bad* mispairs with the *good* ones without altering the number or location of Watson-Crick pairs. We chose the non-homologous control reaction that originally showed a basal signal and contained four *bad* and seven *moderate* mispairs (bar D in Fig. 3). Substitutions were made which replaced three *bad* and four *moderate* mispairs by *good* ones (Fig. 6A). A direct comparison revealed that the new tester paired measurably better than the original one as scored by the ligase assay (Fig. 6B). PhosphorImager quantitation revealed an improvement of new tester (Het2) by 3–4-fold over the old one (Het1) in various repeats. This result strongly reinforced not only the relative hierarchy of various mispairs but also their predictive value (see Section 4).

4. Discussion

Experiments described here critically examine the rules of complementary recognition promoted by RecA for the first time. The results indicate that specific pairing ensues between two sequences even when half of the base pairs are of the non-canonical type. These pairings seem specific enough and reveal an intrinsic hierarchy (Fig. 4A,B). The hierarchy was further clarified when substitutions were made in each base separately which converted one fourth of the base pairs into non-canonical ones (Table 2). The hierarchy seen was not due to differential levels of secondary structures in tester oligos. We analyzed all the tester variants containing single changes for intramolecular as well as intermolecular complementarity by the 'Oligo' program. The maximal number of complementary base pairs observed ranged from 0 to -5 for intramolecular and from 4 to 12 for intermolecular pairings and showed no correlation whatsoever with the pairing efficiencies of these oligos with the template. Moreover, these numbers of intermolecular base pairs were far lower than those expected between the tester and template in targeted alignment. This was so also when the tester sequences were paired with the template even with the allowance of sequence bulges in either strands. All such frameshifted alignments tested by the program yielded only 4–10 complementary base pairs which is far less than that (about 24 out of a total of 33 base pairs) expected in a precise alignment that is ligatable. Therefore, we believe that in the sequences tested, the unknown effects of other pairings that contain bulges are rather low and we could draw conclusions on mispairings in the ligatable frame with a high degree of confidence.

The results showed that in single changes, specific pairings revealed three strengths of non-canonical pairings: the *good, moderate* and *bad* (Figs. 5 and 6). Such a gradation allowed a rational explanation of the hierarchy seen in reactions containing two types of mispairs. For example, the overall hierarchy amongst double mispair sets of GT (*good*)+AC (*good*) > GA (*moderate*)+AG (*good*) > TG (*good*)+CA (*bad*) > CT (*moderate*)+TC (*bad*) stems from the combined effects of individual mispairs. It is to be noted that TG+CA showed the above hierarchy more clearly when a shorter time of ligation was used (Fig. 4B). An interesting aspect of TG+CA pairing is that *bad* mispairs tend to dominate over

the *moderate* or *good* ones. The differential signal observed in controls (Fig. 3A, second bars in C and D) could be explained using the same rationale. The control that worked better had higher number of *moderate* mispairs than the other (7 vs. 3) and reciprocally had fewer number of *bad* mispairs (4 vs. 8).

Based on these observations we hypothesized that RecA pairing across a sequence could be relatively free of sequence context effects and base pairs function as independent modules of pairing. Such a freedom may be ascribed specifically to RecA reaction due to RecA protein's ability to stretch and unstack the DNA bases [10]. This interesting notion was further supported by additional controls. Firstly, RecA pairing could undo sequence context related pairing failures seen in thermal annealings of GA+AG and TG+CA mispairs (Fig. 4A). In each of these cases, RecA pairings were markedly higher than that of thermal. Secondly, by merely replacing *bad*/*moderate* mispairs with *good* ones, pairing efficiencies were enhanced by 3–4-fold (Fig. 6A,B). We believe that the gradation of mispairs, free of sequence context effects, might offer a thumb-rule for predicting the pairing strength of any RecA-mediated alignment that carries multiple mispairs. However, as a word of caution, it is important to note that in situations where sequence context effects become more complex due to either the aspects of DNA secondary structure polymorphism (Z DNA, triplex DNA, etc.) or rigidity of DNA helix, the rules of RecA-mediated mispairing may deviate.

In the experiments described here, alignments between two apparently unrelated sequences (non-homologous) that show about 40% Watson-Crick pairs fail to show significant pairing. In the set where double mispairs were studied, 50% Watson-Crick pairs start exhibiting pairing efficiencies that are specific and comparable to a wild type control. Therefore, within the window of 40–50% sequence similarity, RecA pairing begins to discriminate homology from non-homology. For example, the five mispairs that function poorly in single mismatches (TC, CT, CA, CC and GG) (Fig. 5), in spite of 75% Watson-Crick contribution, underscore their negative dominance in homology search. Homology search involves transient interactions between DNA base pairs before sufficiently stable homologous alignments are achieved. In fact, it is hypothesized that such a search is mediated through weak interactions that are rapid and not rate limiting [15–18]. The experiments described here are the first set of RecA reactions that have to do with high levels of mispairs which border between non-homology and partial homology and hence have a bearing on the early steps of homology search.

Ligation efficiency of various pairing reactions was not a trivial outcome of the nature of terminal base pairs at the ligating end. In GT+AC and GA+AG double mispair reactions, both of which are as proficient as wild type, one substrate set ended with a non-Watson-Crick base pair at the ligating end and the other with a Watson-Crick base pair. In the set with TG+CA mispairs, the pairing that was proficient ended with TG mispair whereas the inefficient one ended with GC, again showing a lack of correlation between Watson-Crick end and ligation proficiency. In fact, the alignment that ends with GC at the ligating end showed no ligation at all in the CT+TC set (Fig. 4A). Among single mispair reactions, the poorest set of CT (*moderate*), CA (*bad*), CC (*bad*) have their first mispair placed farthest away from the ligating end. Conversely, the mispairs GT (*good*), GA (*moderate*) and GG (*moderate*) fare much better even though they are at the ligatable end. Earlier studies have demonstrated the laxity tolerated by T4-DNA ligase at the site of ligation [19]. Phosphodiester bonds that abut on either an abasic site or base mispairs are efficiently ligated by T4-DNA ligase in a substrate system similar to ours (Fig. 1A) [19,20]. Therefore we believe that different ligation efficiencies observed in our experiments truly reflect the intrinsic differences in the strengths of pairing reactions.

Numerous studies have been done in yeast to address the effects of sequence heterogeneity on mitotic recombination rates and the involvement of mismatch repair machinery in regulating homologous interactions [21–26]. These studies used substrates with less than 85% homology and the magnitude of the effects on recombination rates was quite variable depending upon whether plasmid or chromosomal events were monitored. Indeed, it is known that plasmid versus chromosomal recombination events have different genetic requirements which explain the differential effects seen [27]. Also, chromosomal recombination in mammalian cells is much more sensitive to sequence divergence than is extrachromosomal recombination [28]. However, a general point to note both in prokaryotes and in eukaryotes is that reduced homology between DNA substrates decreases recombination frequencies, although to varying extents [28,29]. Defects in yeast mismatch repair can elevate intrachromosomal recombination between 91% homologous inverted repeats by as much as 100-fold while they have only a modest effect on recombination between 77% homologous ones [25]. In this system, the anti-recombination role of the mismatch repair system is so strong that even a single mismatch between two otherwise identical sequences inhibits recombination fully [26]. So, most of these studies reflect a combined biological outcome of the acceptance of divergent sequences by recombination machinery and the rejection of the same by mismatch repair machinery in vivo [30–32].

Homologous recombination in vivo is best exemplified in a conjugational cross between *Salmonella typhimurium* (Hfr) and *E. coli* (F^-) in which the recipient is $muts^-$. Extensive recombination is revealed in such a cross between the two genomes that show an overall 18% sequence divergence, although the degree of divergence varies greatly at different regions of the genome with the highly expressed genes showing much less divergence [33]. However, the upper limit of sequence divergence that $muts^-$ *E. coli* recombination machinery tolerates in vivo is not known. A high degree of mispairing seen in the present study is consistent with the promiscuity of RecA pairing that were reported earlier [18,34]. RecA paired G4 and ϕX-174 phage DNA sequences only when one of the substrates was supercoiled and not when relaxed [34]. Such a reaction between sequences that have about 30% sequence divergence is believed to form D-loop complexes that are sustained essentially by complementary pairings and not by stable three stranded pairings [4–6]. The results described in this paper on a complementary recognition paradigm such a D-loop three stranded synapse. However, in a stable three stranded pairing, when duplex substrate is in a topologically relaxed state, pairing is much more intolerant to mismatches and is detectable only when there is as little as 3% sequence divergence [35]. Such a pairing might most likely involve base triplet recognitions that are more stringent than complementary interactions. Analysis of three stranded inter-

mediates formed by RecA pairings in vitro has revealed some aspects of such triplex structures [36–38].

Results demonstrating the ability of RecA proteins to show low fidelity pairing demand an intervention by the mismatch repair machinery during homology search. What is the relationship between the base pair hierarchy seen in RecA pairing reactions here vis-à-vis that of mismatches repaired postreplicationally by the Mut system in *E. coli*? Perhaps the relationship is rather remote because the former has to do with the early steps of homology search whereas the latter has to do with methyl-directed repair of heteroduplexes generated by polymerase errors during DNA replication, the two sets of events possibly are unconnected in vivo. In vivo studies in *E. coli* on transfected heteroduplexes have revealed that transition mismatches are repaired better than transversions on either strand of unmethylated DNA and CC, TC, GA mismatches are highly refractory to repair [39–41]. Moreover, this repair is subject to sequence context effects [42]. In vitro repair results are largely consistent with in vivo heteroduplex transfection studies [43]. Estimation of the apparent affinity of MutS for each of the mispairs revealed only a very rough correlation between MutS binding and efficiency of correction [43,44]. For instance, the rates of repair of AC and CT are similar even though the affinity of MutS for AC is significantly higher than that for CT. The finding that there is no simple correlation even between binding and repair in methyl-directed postreplicational mismatches by MutS underscores a related challenge posed by the present results. Thus it is even more difficult to draw a correlation between the efficiencies (binding or correction) by the mismatch repair system and its ability to disallow mispairs in the context of RecA-mediated pairing. We are separately addressing this issue in vitro to understand how the MutS machinery handles different DNA mispairs in the context of the RecA filament (manuscript in preparation). This study was motivated, in part, also by the findings that MutS/MutL proteins efficiently undo RecA-mediated three stranded branches in vitro [45]. We believe that the facet of the ability of MutS/MutL proteins to efficiently undo heteroduplex joints within the RecA filament heralds a novel enzymatic reaction of the MutS machinery that is fully independent of methyl-directed postreplicational events. Future studies should unravel more details of connections between rules of RecA mispairing and those of MutS unpairing.

Acknowledgements: We would like to thank Dr. S.K. Mahajan of BARC for the usage of the phosphorImager and Mr. Narayan for his assistance with the software. We would also like to thank Mr. Jatin from the photography section of TIFR for his unstinting efforts in getting high quality photographs to our satisfaction.

References

[1] Kowalczykowski, S.C., Dixon, D.A., Eggleston, A.K., Lauder, S.D. and Rehrauer, W.M. (1994) Microbiol. Rev. 58, 401–465.
[2] Radding, C.M. (1993) Curr. Biol. 3, 358–360.
[3] Saenger, W. (1984) Principles of Nucleic Acid Structure, Springer, New York.
[4] Beattie, K.L., Wiegand, R.C. and Radding, C.M. (1977) J. Mol. Biol. 116, 783–803.
[5] Hsieh, P., Camerini-Otero, C.S. and Camerini-Otero, D. (1992) Proc. Natl. Acad. Sci. USA 89, 6492–6496.
[6] Adzuma, K. (1992) Genes Dev. 6, 1679–1694.
[7] Sinden, R.R., Carlson, J.O. and Pettijohn, D.E. (1980) Cell 21, 773–783.
[8] Carrol, D., Wright, S.H., Wolff, R.K., Grzesiuk, E. and Maryon, E.B. (1986) Mol. Cell. Biol. 6, 2053–2061.
[9] Seidman, M.M. (1987) Mol. Cell. Biol. 7, 3561–3565.
[10] Egelman, E.H. (1993) Curr. Opin. Struct. Biol. 3, 189.
[11] Shibata, T., Cunningham, R.P. and Radding, C.M. (1981) J. Biol. Chem. 256, 7557–7564.
[12] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) in: Molecular Cloning: A Laboratory Manual, 2nd edn., p. 11.39, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[13] 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.
[14] Bianco, P. and Weinstock, G.M. (1996) Nucleic Acids Res. 24, 4933–4939.
[15] von Hippel, P.H. and Berg, O.G. (1989) J. Biol. Chem. 264, 675–678.
[16] Kleckner, N. and Weiner, B.M. (1993) Cold Spring Harbor Symp. Quant. Biol. LVIII, 553–565.
[17] Yancey-Wrona, J.E. and Camerini-Otero, D. (1995) Curr. Biol. 5, 1149–1158.
[18] Malkov, V.A., Sastry, L. and Camerini-Otero, R.D. (1997) J. Mol. Biol. 271, 168–177.
[19] Wu, D.Y. and Wallace, B.R. (1989) Gene 76, 245–254.
[20] Goffin, C., Bailly, V. and Verly, W.G. (1987) Nucleic Acids Res. 15, 8755–8771.
[21] Harris, S., Rudnicki, K.S. and Haber, J.E. (1993) Genetics 135, 5–16.
[22] Mezard, C., Pompon, D. and Nicolas, A. (1992) Cell 70, 659–670.
[23] Priebe, S.D., Westmoreland, J., Nilsson-Tillgren, T. and Resnick, M.A. (1994) Mol. Cell. Biol. 14, 4802–4814.
[24] Selva, E.M., New, L., Crouse, G.F. and Lahue, R.S. (1995) Genetics 139, 1175–1188.
[25] Datta, A., Adjiri, A., New, L., Crouse, G.F. and Robertson, S.J. (1996) Mol. Cell. Biol. 16, 1085–1093.
[26] Datta, A., Hendrix, M., Lipsitch, M. and Robertson, S.J. (1997) Proc. Natl. Acad. Sci. USA 94, 9757–9762.
[27] Sugawara, N., Ivanov, E.L., Fishman-Lobel, J., Ray, B.L., Wu, X. and Haber, J.E. (1995) Nature 373, 84–86.
[28] Waldman, A.S. and Liskay, R.M. (1987) Proc. Natl. Acad. Sci. USA 84, 5340–5344.
[29] Shen, P. and Huang, H.V. (1989) Mol. Gen. Genet. 218, 358–360.
[30] Rayssiguier, C., Thaler, D.S. and Radman, M. (1989) Nature 342, 396–401.
[31] Matic, I., Rayssiguier, C. and Radman, M. (1995) Cell 80, 507–515.
[32] Vulic, M., Dionisio, F., Taddei, F. and Radman, M. (1997) Proc. Natl. Acad. Sci. USA 94, 9763–9767.
[33] Sharp, P.M. (1991) J. Mol. Evol. 33, 23–33.
[34] Dasgupta, C. and Radding, C.M. (1982) Nature 295, 71–73.
[35] Bianchi, M.E. and Radding, C.M. (1983) Cell 35, 511–520.
[36] Chiu, S.K., Rao, B.J. and Radding, C.M. (1993) Biochemistry 32, 13146–13155.
[37] Rao, B.J., Chiu, S.K. and Radding, C.M. (1993) J. Mol. Biol. 229, 328–343.
[38] Kim, M.G., Zhurkin, V.B., Jernigan, R.L. and Camerini-Otero, R.D. (1995) J. Mol. Biol. 247, 874–889.
[39] Kramer, B., Kramer, W. and Fritz, H.-J. (1984) Cell 38, 879–887.
[40] Dohet, C., Wagner, R. and Radman, M. (1985) Proc. Natl. Acad. Sci. USA 82, 503–505.
[41] Jones, M., Wagner, R. and Radman, M. (1987) Genetics 115, 605–610.
[42] Fazakerley, G.V., Quignard, E., Woisard, A., Guschbauer, W., van der Marel, G.A., van Boom, J.H., Jones, M. and Radman, M. (1986) EMBO J. 5, 3697–3703.
[43] Su, S.-S., Lahue, R.S., Au, K.G. and Modrich, P. (1988) J. Biol. Chem. 263, 6829–6835.
[44] Jiricny, J., Su, S.-S., Wood, S.G. and Modrich, P. (1988) Nucleic Acids Res. 16, 7843–7853.
[45] Worth Jr., L., Clark, S., Radman, M. and Modrich, P. (1994) Proc. Natl. Acad. Sci. USA 91, 3238–3241.