

Research article

Open Access

Phage annealing proteins promote oligonucleotide-directed mutagenesis in *Escherichia coli* and mouse ES cells

Youming Zhang^{1,2}, Joep PP Muylers^{1,2}, Jeanette Rientjes^{1,2} and A Francis Stewart*^{1,3}

Address: ¹Gene Expression Program, EMBL, Meyerhofstr. 1, 69117, Heidelberg, ²Gene Bridges GmbH, Pfotenhauerstr. 108, 01307 Dresden and ³BIOTEC, Technische Universitaet Dresden, c/o MPI-CBG, Pfotenhauerstr. 108, 01307 Dresden

Email: Youming Zhang - youming.zhang@genebridges.com; Joep PP Muylers - joepmuyjers@hotmail.com; Jeanette Rientjes - jeanette.rientjes@genebridges.com; A Francis Stewart* - stewart@mpi-cbg.de

* Corresponding author

Published: 16 January 2003

Received: 16 September 2002

BMC Molecular Biology 2003, 4:1

Accepted: 16 January 2003

This article is available from: <http://www.biomedcentral.com/1471-2199/4/1>

© 2003 Zhang et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: The phage protein pairs, RecE/RecT from *Rac* or Red α /Red β from λ , initiate efficient double strand break repair (DSBR) in *Escherichia coli* that has proven very useful for DNA engineering. These phage pairs initiate DSBR either by annealing or by another mechanism that is not defined.

Results: Here we report that these proteins also mediate single strand oligonucleotide repair (ssOR) at high efficiencies. The ssOR activity, unlike DSBR, does not require a phage exonuclease (RecE or Red α) but only requires a phage annealing protein (RecT or Red β). Notably, the P22 phage annealing protein Erf, which does not mediate the same DSBR reactions, also delivers ssOR activity. By altering aspects of the oligonucleotides, we document length and design parameters that affect ssOR efficiency to show a simple relationship to homologies either side of the repair site. Notably, ssOR shows strand bias. Oligonucleotides that can prime lagging strand replication deliver more ssOR than their leading complements. This suggests a model in which the annealing proteins hybridize the oligonucleotides to single stranded regions near the replication fork. We also show that ssOR is a highly efficient way to engineer BACs and can be detected in a eukaryotic cell upon expression of a phage annealing protein.

Conclusion: Phage annealing proteins can initiate the recombination of single stranded oligonucleotides into endogenous targets in *Escherichia coli* at very high efficiencies. This expands the repertoire of useful DNA engineering strategies, shows promise for applications in eukaryotic cells, and has implications for the unanswered questions regarding DSBR mediated by RecE/RecT and Red α /Red β .

Background

In a search for a simple way to use homologous recombination in *E. coli* for DNA engineering, we found that the phage protein pairs of either RecE/RecT, from the *Rac* prophage, or Red α /Red β from λ , phage, mediate very useful double stranded break repair (DSBR) in the absence of

RecBC [1]. Concomitantly, Murphy [2] reported similar findings from studies with λ Red recombination. Notably, we found that this DSBR requires only very short homology regions for efficient recombination and the ratio of homologous to non-homologous recombination is very high [1]. Together with other advantages, these

remarkable properties opened a new logic for engineering DNA using recombination *in vivo* that circumvents many of the limitations inherent with the *in vitro* use of restriction enzymes, ligases and PCR or the *in vivo* use of the *recA/recBC* pathway. A variety of implications have been developed, by us [3–6] and others ([7,8,8–15]; for reviews see [16–18]). Initially we termed this DSBR mechanism "ET recombination" after RecE/RecT, but the alternative terms "Red" after the λ phage proteins, and 'recombineering' are also in use (Copeland et al, 2001). In the interests of clarity, we will refer to homologous recombination in *E. coli*, mediated by either RecE/RecT or Red α /Red β in the absence of RecBCD activity, as Red/ET recombination.

The Red/ET recombination mechanism is the simplest example of DSBR yet described with distinct similarity to DSBR in yeast [19]. Like in yeast, it can be employed for both efficient gene replacement and gap repair using short homology regions [6,20]. Red/ET recombination involves co-operation between a 5'-3' exonuclease (RecE or Red α , see [21] for Red α crystal structure) with a single strand DNA binding and annealing protein (RecT or Red β , see [22,23] for analyses of these proteins on DNA by electron microscopy). In a comparative mechanistic study, we showed that DSBR initiated by either phage protein pair is functionally and operationally equivalent [24]. This study also revealed that co-operation between the 5'-3' exonuclease and its cognate annealing protein is required, due to a specific protein/protein interaction between the two components. Furthermore, we presented evidence showing that DSBR by Red/ET encompasses two mechanisms, one of which concurs with the accepted annealing mechanism [25,26] and another, which remains undefined, could be due to a strand invasion activity. On this point, *in vitro* binding studies have shown that both RecT and Red β can mediate strand displacement [27,28].

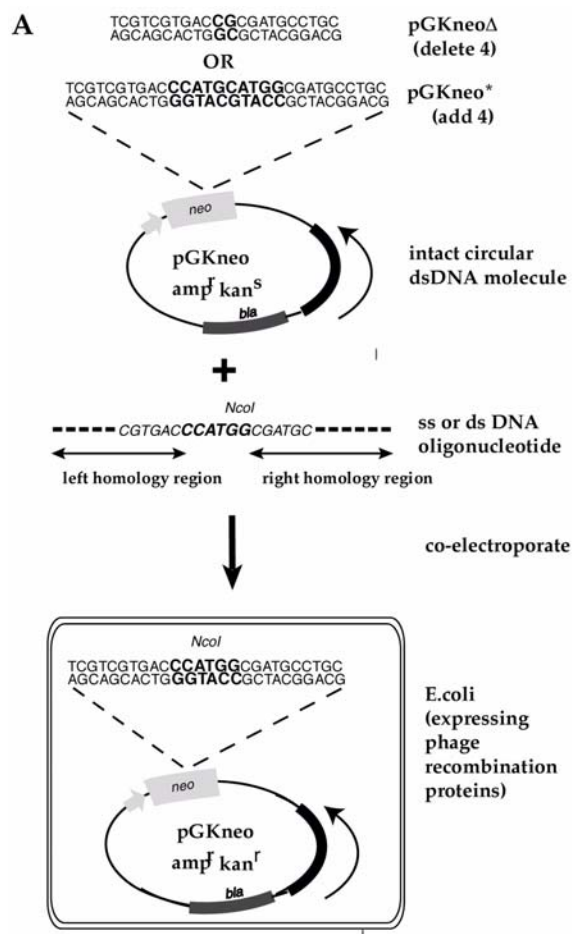
Here we report an unexpected further activity with single stranded DNA that requires only the single stranded binding proteins, RecT or Red β . The same activity has been recently reported by Ellis *et al* [29]. We characterize the activity, here referred to as ssOR for 'single stranded oligo repair', and thereby establish several criteria that distinguish it from Red/ET DSBR. In agreement with the chromosomal studies of Ellis *et al*, we show that ssOR displays strand bias on pUC plasmid origin targets. Repair by oligos that hybridize with the lagging strand template (lagging) is consistently higher than repair by oligos that hybridize with the leading strand template (leading). This suggests the involvement of the replication fork. Interestingly, ssOR also occurs in eukaryotic chromatin when a phage single strand binding protein is expressed.

Results

Identification of a novel single strand repair activity

Previously we applied a double-strand break repair (DSBR) assay to analyze aspects of homologous recombination mediated by the phage protein pairs, RecE/RecT and Red α /Red β [24]. With the intention to explore the DSBR mechanism further, we developed a DNA repair assay using a mutated kanamycin resistance gene (*neo*) present in pGKneo (Fig. 1A). pGKneo was cut with Nco1 and either filled in to create a 4 nucleotide insertion (pGKneo*), or resected to create a four nucleotide deletion (pGKneo Δ). Oligonucleotides (oligos) containing the wild type sequence across the Nco1 site will, if incorporated correctly into pGKneo* or pGKneo Δ , restore kanamycin resistance and thus be easily detected. Most of the following experiments were performed with both pGKneo templates. Since virtually identical results were obtained for both, except that results with pGKneo* consistently yielded slightly more correctly repaired products, only results obtained with pGKneo Δ are presented. In the first experiment, a 46 nt double stranded oligo, containing 21 bps of homologous sequences on both sides of the intact Nco1 site, was used to evaluate the frequency of repair in the presence or absence of either RecE/RecT or Red α /Red β protein pairs. Repair was dependent on expression of phage proteins and both insertion and deletion substrates were suitable (Fig. 1B, and data not shown). Since Red/ET DSBR requires linearised, double stranded DNA and does not work with single stranded DNA [1], we expected that the corresponding single stranded oligos (ss oligos) would serve as good negative controls in this experiment. Surprisingly, each ss oligo delivered notable levels of phage protein-dependent repair (Fig. 1B). This unexpected result suggested that the phage proteins mediate a repair reaction different from their activities in DSBR and prompted further analysis.

In addition to the requirement for double stranded DNA, Red/ET DSBR requires the co-expression of an orthologous phage protein pair, either RecE/RecT or Red α /Red β [24]. Since the ss oligo repair (ssOR) activity was unexpected, we examined the requirement for expression of an orthologous pair. As also shown in Fig. 1B, ssOR was mediated by either phage annealing protein alone without the need for its orthologous exonuclease partner. This presents a second piece of evidence that the ssOR activity is distinct from Red/ET DSBR. In contrast to the results of Ellis *et al* [29], ssOR by either annealing protein alone was less efficient than that obtained when the cognate exonuclease partner was co-expressed (Fig. 1B). Since it is unlikely that the exonuclease activity of RecE or Red α is beneficial, this result may reflect an aspect of the specific protein/protein interaction between these protein pairs [24,30]. Alternatively, it is possible that the difference between our results and those of Ellis *et al* on this point



B

Strand	Vector	RecET	Redβ _α	RecE	RecT	Red _α	Redβ
double	0	435	929	0	137	0	128
lagging	0	116	151	0	83	0	90
leading	0	30	34	0	12	0	15

Figure 1
ssOR requires only one strand and expression of a phage annealing protein. **A.** Outline of the ssOR assay in *E. coli*. The Tn5 gene for kanamycin resistance (*neo*) in pGKneo was cut at its NcoI site and either end filled or treated with exonuclease to create pGKneo* or pGKneoΔ respectively. pGKneo plasmids also encode for ampicillin resistance (*bla*). The direction of leading strand synthesis initiated from the pUC-derived origin is indicated. A pGKneo plasmid was mixed with either single or double stranded oligonucleotides encoding the unmutated sequence across the NcoI site and co-electroporated into an *E. coli* host containing or not various phage proteins. Relative oligo repair efficiencies were scored by counting kanamycin resistant colonies adjusted for transformation efficiency by scoring ampicillin resistant colonies for each data point. **B.** Two 48 mer oligos encompassing 22 nts of homology either side of the inner 4 nts of the NcoI site were synthesized and used either alone (lagging, leading) or after annealing together (double). They were co-electroporated with pGKneoΔ into *E. coli* hosts that contained the expressed phage proteins indicated; Vector; the pBAD24 expression plasmid without any cloned phage protein. "Lagging" denotes the oligo that hybridizes to the lagging strand template, hence could prime Okazaki fragment synthesis. Numbers denote the relative ssOR efficiency.

Table 1: Evaluation of host context and dependence on phage protein expression in ssOR.

Strain + plasmid	Genotype + expressed proteins	Normalised ssOR
MM294	wild-type	0
+ pBADET	+ RecE, RecT	50
+ pBADET γ	+ RecE, RecT, Red γ	46
+ pBAD $\alpha\beta$	+ Red α , Red β	55
+ pBAD $\alpha\beta\gamma$	+ Red α , Red β , Red γ	52
JC8679	<i>SbcA</i> , <i>recBC</i>	15
JC5519	<i>RecBC</i>	0
+ pBADrecET	+RecE, RecT	49
+ pBADred $\alpha\beta$	+ Red α , Red β	52
+ pBADerf	+ Erf	35
JC9366	<i>recA</i>	0
+ pBADrecET	+RecE, RecT	45
+ pBADrecET γ	+RecE, RecT, Red γ	50
+ pBADred $\alpha\beta$	+ Red α , Red β	54
+ pBADred $\alpha\beta\gamma$	+ Red α , Red β , Red γ	51
+pBADerf	+ Erf	32
JM103	<i>sbcBC</i>	0.1
JC9387	<i>sbcBC</i> , <i>rac</i> , <i>recBC</i>	0.3
JC15329	<i>sbcBC</i> , <i>recA</i> , <i>recAB</i> , <i>rac</i>	0
JC8111	<i>sbcBC</i> , <i>rac</i> , <i>recBC</i> , <i>recF</i>	0.1

Various *E. coli* hosts were transformed with pBAD24 expression plasmids or not as indicated before co-electroporation with a 48 mer oligonucleotide and pGKneo Δ to test for ssOR. The strains used were MM294, JC8679 [44], JC5519 [45], JC9366 [45], JM103 [46], JC 9387 [45], JC15329 [47] and JC8111 [48]. Relevant aspects of genotype and expressed proteins, after arabinose induction, are listed in the central column. Numbers refer to the ratio of kanamycin to ampicillin resistant colonies and were generated from the same master mixture of oligonucleotide and pGKneo Δ .

reflect an operational aspect of experimental design, such as protein expression levels or stabilities, either in our experiments or those of Ellis *et al.* Notably, the lagging oligo (that is, the oligo that would prime lagging strand synthesis) delivered better efficiency than the leading for all configurations (Fig. 1B). Also notable is the observation that the double stranded (ds) oligo delivered better efficiency than either ss oligo in all cases. In the cases where either phage protein pair was co-expressed, this appears to be simply the additive outcome of both DSBR and ssOR activities. In the cases in which only one annealing protein was expressed, this may reflect a simple addition of lagging and leading ssOR. Alternatively, it could reflect an operational parameter such as stabilization of the oligos against decay, or a beneficial contribution of a presumptive ds oligo unwinding step. To focus on the ssOR activity, further experiments were performed with ss oligos.

To evaluate the requirements for the ssOR activity in more detail, a series of differently configured *E. coli* hosts were examined (Table 1). No significant activity could be found in the absence of a phage annealing protein, including in the wild type *E. coli* strain, MM294, the *recBC* strain JC5519, or *recA* strain JC9366. Furthermore, no significant activity was found in several *sbcBC* strains. Importantly, ssOR in the presence or absence of RecBC activity was the same. This conclusion is established by compar-

ing strains with or without RecBC (MM294, JC8679, JC5519, JC9366) and with or without expression of the RecBC inhibitor, Red γ [31]. The independence of ssOR from RecBC presents a third operational difference to Red/ET DSBR, since phage DSBR is diminished by the presence of RecBC. The results in Table 1 also show that ssOR on plasmids is independent from RecA, which is neither required nor an impediment.

Interestingly, the P22 phage annealing protein, Erf, also mediates efficient ssOR (Table 1). The P22 recombination system, which includes Erf, Arf, Abc1 and Abc2, shares some mechanistic similarities with that of the RecE/RecT and Red α /Red β phage protein pairs [32,33] but does not appear to include a 5'-3' exonuclease activity. Consistent with this absence, we have been unable to find notable DSBR activity with any combination of these four P22 proteins (unpublished results). The observation that Erf alone, like RecT and Red β , delivers ssOR is a fourth piece of evidence that this activity is distinct from Red/ET DSBR.

Characterisation of the ss oligo repair reaction

The contribution of ss DNA to ssOR was evaluated in four ways.

First, the relationship between length of homology and repair efficiencies was determined by use of a series of

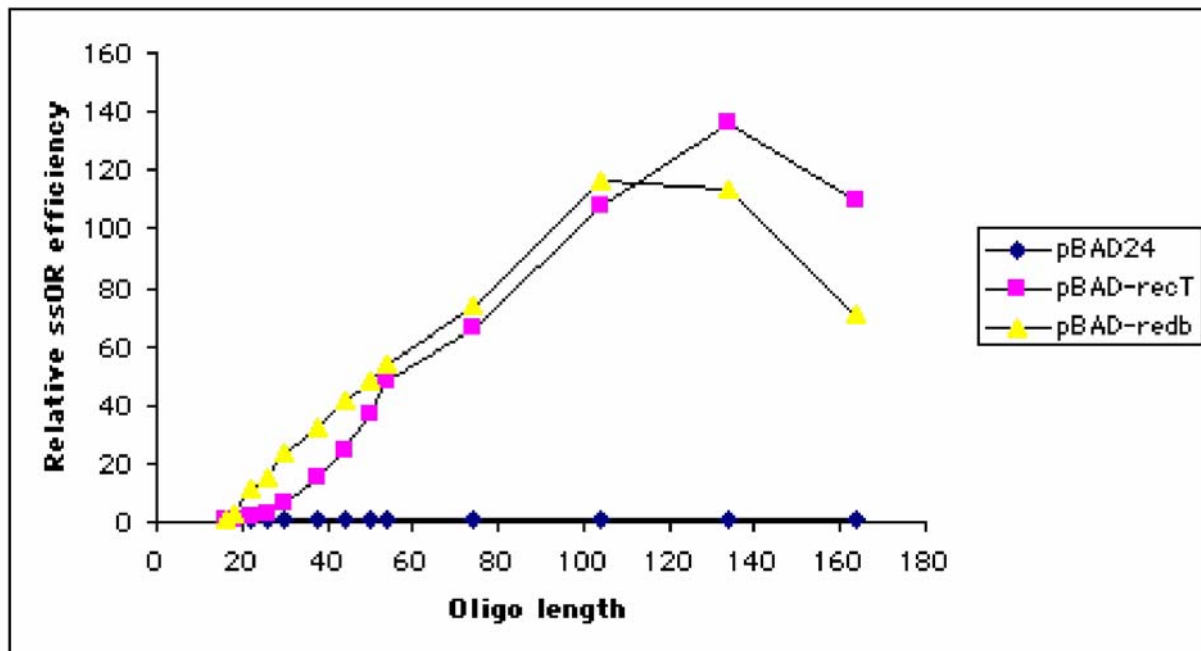


Figure 2

Relationship between oligo length and ssOR efficiency. The graph shows relative ssOR efficiencies, scored as kanamycin resistance colonies normalized to ampicillin resistant colonies for oligos of the indicated lengths centered around the NcoI site in the *neo* gene. All oligos were lagging and were co-electroporated with pGKneoΔ into single batches of competent cells carrying either pBAD24 vector (diamonds), pBAD24-recT (squares) or pBAD24-redβ (triangles) that had been cultured with arabinose to induce expression from the pBAD promoter. Results are from one experiment and a repetition using 130 and 160 mers synthesized a second time gave the same profile.

lagging oligos from 16 – 160 nucleotides, each with the repair site in the center (Fig. 2). A little ssOR mediated by Redβ was observed with the 22 mer ss oligo (i.e. 9 nt homology arms either side of the inner 4 bps of the NcoI site) and efficiencies increased to an apparent maximum near 120 nts. RecT displayed a very similar profile except that a slightly longer oligo (30 nt) was required before repair was detected. These results concur with, and extend, the chromosomal analysis of Ellis *et al* [29] who examined oligo length dependence with Redβ in the range 20 – 70. Notably, the efficiencies of both Redβ and RecT ss oligo repair were reduced at the longest length examined. This may reflect an implicit limit of ssOR, or an operational parameter such as oligonucleotide quality. Interestingly, the

relationship between total homology included in the oligos and ssOR efficiencies is very similar to the relationship observed for single homology arm length and Red/ET DSB repair efficiencies [1,24], including the shortest lengths required (around 27) and the apparent maximums observed (around 120 nts).

Second, the relative contributions of homologies on the 5' and 3' sides of the repair site were evaluated. No repair was observed when the repair site was located at either the 5' or 3' ends, regardless of the length of the single-sided homology region, up to 100 nt (Fig. 3A and not shown). Hence homology regions on both sides of the repair site are required. Next, a series of six 50 mers were synthe-

sized, (3 pairs for lagging and leading) differing in the relative position of the repair site within the oligo (Fig. 3B). For both lagging and leading oligos, the centrally positioned repair site (23/4/23) was most efficient. Again the lagging oligo delivered better efficiency than the leading. For both lagging and leading oligos, more homology 5' to the repair site (35/4/11) worked better than oligos with more 3' homology (11/4/35). Notably, the 11/4/35 and 35/4/11 oligos functioned with RecT whereas the 11/4/11 26 mer oligo did not (Fig. 2). This indicates that the stability of hybrid formation on one side of the repair enhances the efficacy of the hybrid formed on the other side.

In a third variation, three new substrate plasmids were created that had larger deletions around the Nco1 site, being deletions of 15, 33 and 60 nucleotides (Fig. 3B). Oligos were synthesized that retained 22 nts homology on either side of the deletions and included the extra 4, 15, 33 or 60 nts required for repair of the *neo* gene for kanamycin resistance. Repair efficiencies decreased with the increasing size of the deletion. Hence ssOR is less efficient when more of the oligo itself needs to be inserted. In contrast, Ellis *et al* [29] obtained a very similar efficiency of ssOR in the converse situation, namely when more of the template was deleted. They showed equivalent efficiencies when repairing a point mutation or deleting 3.3 kb with the same oligo. Together these data imply that oligo/template hybrid formation is differentially sensitive. Looping out of the oligo from the template appears to be more deleterious than looping out of the template from the oligo.

In a fourth experiment, additional point mutations were introduced in the homology region near to the repair site. The point mutations were chosen to be silent with respect to the kanamycin resistance protein reading frame, and also to be easily identifiable by alteration of the restriction digestion pattern. The restriction digestion pattern therefore will display complete incorporation during ssOR, or further repair, of the oligo. In all configurations of this experiment, the point mutation reduced ssOR efficiencies (Fig. 3C and data not shown) and was removed in all but 2 cases out of 168 examined. Both of these two cases were found with the 44/14/29 lagging oligo repairing pGKneo* (2/28 cases examined; data not shown).

Taken together, the data of Fig. 2 and Table 3 imply that the efficiencies of ssOR simply reflects the stability of the hybrid formed between the entire oligo and the template on both the 5' and 3' sides of the repair site, up to an optimum size of approximately 120 nts. Whereas ssOR may not be impaired by intervening, non-homologous sequence in the template, it appears to be impaired by non-homologous sequence in the oligo.

Phage proteins mediate ss oligo repair in mouse ES cells

Since the above characterization of ssOR simply highlighted the importance of stable hybrid formation, we reasoned that it may also work in any host, including eukaryotes. On the other hand, a failure to work in a non-prokaryotic host may indicate a specific requirement for other prokaryotic factors. Hence an experiment in mouse ES cells was performed (Table 2). The defective *neo** gene was cloned into pcDNA to create pcDNA/PGK-*neo**. Then *redβ* was cloned in to create pcDNA-*redβ*/PGK-*neo**. Both vectors were stably integrated into ES cells by selection for hygromycin resistance and 12 independent colonies each were taken, electroporated with a repairing oligo, followed by selection for G418 resistance. No G418R colonies were observed with any pcDNA/PGK-*neo** clone, whereas half of the pcDNA-*redβ*/PGK-*neo** clones gave rise to G418 resistant colonies. Hence Redβ can function in a eukaryotic context to mediate ssOR and no factor specific to *E. coli* is required. Similar results in ES cells were obtained with RecT (data not shown).

Evaluation of absolute efficiencies and application to BAC engineering

To evaluate the utility of ssOR in DNA engineering, two experiments were performed with BACs in *E. coli*. In the first, the *neo* gene was interrupted, at its Nco1 site, by inserting a gene encoding for zeocin resistance, and this cassette was integrated, by Red/ET DSB, into a BAC containing 143 kb of mouse DNA including most of the Mll gene (Fig. 4). Ss oligo repair on the endogenous BAC will delete the zeocin gene and restore kanamycin resistance. Unlike the co-transfer plasmid assay used above, where the calculation of absolute efficiencies is complicated by other parameters (such as the high-copy number of the plasmid target and variable co-transformation efficiencies), this assay presents a simple way to evaluate absolute ssOR efficiencies. The total number of chloramphenicol resistant colonies scores the total number of endogenous BAC targets in the experiment and kanamycin resistant colonies represents the number of ssOR events. Five oligos with increasing lengths of homology either side of the Nco1 site were evaluated. The most efficient was a 106 mer, that is with 55 nt homology either side of the zeocin gene to be deleted. The longer oligos did not give better results. A remarkable efficiency of 3% was observed with the 106 mer. At these efficiencies, it is clearly straightforward to identify correct ssOR products by physical methods such as colony PCR or hybridization.

In the second BAC experiment, the merits of concerted usage of antibiotic selection and counterselection were examined. A cassette containing the *rpsL* and *neo* genes was integrated into the Mll BAC using Red/ET DSB by selection for kanamycin resistance (Fig. 5). Two oligos with 25 nts homology arms were used for ssOR. In the centre of

A		RecT	Redβ
23	23	46	51
35	11	35	33
11	35	19	22
0	20	0	0
0	100	0	0
20	0	0	0
100	0	0	0
leading			
23	23	22	24
35	11	18	19
11	35	9	11
B			
pGKneoΔ + ssDNA oligo			
pGKneoΔ +		40	41
pGKneoΔ15 +		36	34
pGKneoΔ33 +		18	21
pGKneoΔ60 +		7	10
C			
		61	101
		29	42
		20	28

Figure 3

Exploration of ssOR by alterations in oligo design. The contribution of the oligo to ssOR was explored by altering the oligos in three different ways. **A.** The position of the repair site was shifted from a central position in 50 mers (23/23) to more 3' (35/11) or 5' (11/35), as well as at the 5' (0/20; 0/100) or 3' (20/0; 100/0) terminae, for both lagging and leading oligos. Results are from a single experiment performed with competent cell batches made in parallel after induction of RecT or Redβ expression. **B.** The ability of ssOR to repair larger deletions in the template was evaluated with three further derivatives of pGKneoΔ that lacked 15, 33 or 60 bps around the NcoI site of the kanamycin resistance gene. Oligos that retained 22 nucleotides of homology either side of the the 4, 15, 33 or 60 nucleotides required to restore kanamycin resistance were co-electroporated with pGKneoΔ. Results are from a single experiment performed with lagging oligos and competent cell batches made in parallel after induction of RecT or Redβ expression. **C.** The impact of a second mismatch in the oligo was evaluated using 92 mers with the NcoI site off center leaving a short side of 29 nts from either the 5' or 3' end. The longer side of 59 nts was interrupted either by a point mutation 14 or 29 nts from the NcoI site. Results shown are from a single experiment using a lagging oligo. Qualitatively similar results were obtained with the corresponding leading strand oligos (data not shown). The two, rare, events in which the point mutation was also incorporated were found with the 44/14/29 oligo illustrated.

Table 2: ssOR in ES cells.

Clones	1	2	3	4	5	6	7	8	9	10	11	12
pcDNA/PGK-neo*	0	0	0	0	0	0	0	0	0	0	0	0
pcDNA-redb/PGK-neo*	0	4	20	0	7	0	0	0	11	0	6	5

ES cells were electroporated with pcDNA derivatives containing the filled-in *neo* gene (*neo**) under a PGK promoter without or with *redβ* cloned under the CMV promoter. Stable colonies were isolated by hygromycin resistance, conveyed by the pcDNA vector, and 12 randomly picked clones were taken for each of the 2 vectors. These colonies were expanded separately and electroporated with a 70 nt repairing oligo for the *neo** gene, followed by plating and selection for G418 resistance. The numbers shown are the numbers of colonies that arose for each of the 2 × 12 clones.

these oligos, either an additional 6 nts encoding an XhoI site, or an additional 34 nts encoding the minimal FRT site, were included. Candidate colonies containing the correctly modified BACs were identified by selection for streptomycin resistance (i.e. elimination of the *rpsL* gene in HS996 host background). In both cases 20/22 colonies examined by restriction digestion analysis of BAC DNAs showed the correct ssOR event without any apparent, unintended, secondary recombinations elsewhere in the BAC (data not shown). Hence this selection/countersélection strategy is qualitatively very efficient. Furthermore this strategy does not rely on the absolute efficiencies required for physical screening methods, such as colony PCR, but permits a practical way to use suboptimal ssOR oligos. In this experiment, the oligos were suboptimal since they included both significantly shorter regions of homology than optimal (Fig. 2, Fig. 4) and centrally located non-homologous sequences that impair absolute efficiency (Fig. 3B).

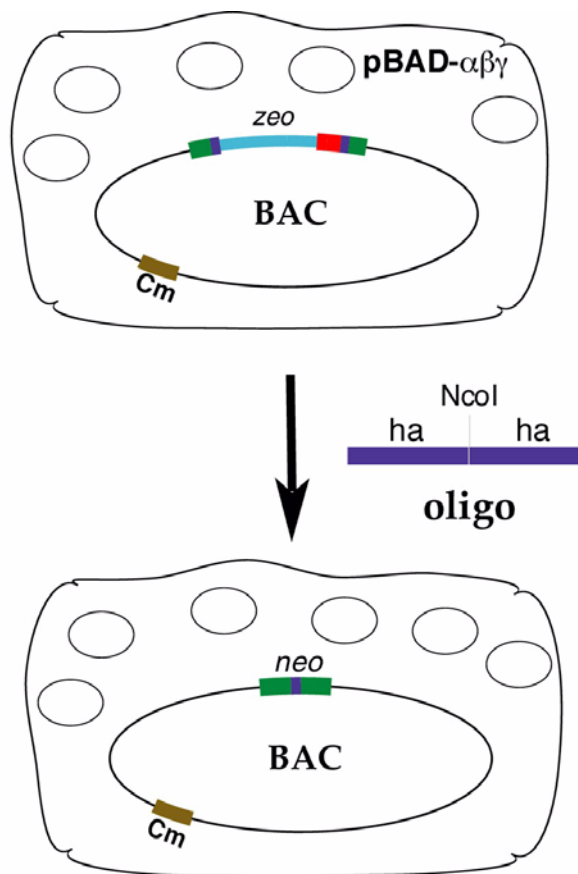
Discussion

We developed a DNA repair assay to explore further the Red/ET DSBR activity, however unexpectedly observed an activity with single stranded DNA, termed here ssOR (single stranded oligo repair). The two activities differ with respect to required protein expression. Red/ET DSBR requires the co-expression of a phage 5'-3' exonuclease and annealing protein pair [24] whereas ssOR requires only the expression of either phage annealing protein, RecT or Redβ. Consistent with this difference, the P22 phage annealing protein, Erf, which does not appear to have a phage 5'-3' exonuclease partner and does not mediate DSBR, also mediates ssOR.

The simplest explanation for the ssOR reaction is that the oligos serve as primers for DNA replication after the phage annealing proteins establish a hybrid molecule spanning the repair site. A notable characteristic of ssOR is strand bias that concurs with the orientation of replication origins. In our plasmid assay, the oligos that would prime lagging strand replication delivered greater efficiencies than those that would prime the leading strand. This characteristic was also noted by Ellis *et al* [29], who carefully

examined the phenomenon at various sites around the *E. coli* chromosome. They were able to show that ssOR strand bias was independent of the direction of local transcription and simply concordant with the direction of replication. Before their study was published, we had drawn a similar conclusion by integration of the *neo** and *neoΔ* repair substrates into the *lacZ* gene on the *E. coli* chromosome. Since these results do not extend those of Ellis *et al*, they have not been included in this paper. The strand bias strongly implicates the asymmetric exposure of single stranded regions at the replication fork [34] as the sites to which the phage annealing proteins establish stable oligo hybrids. If this simple physical explanation, referred to here as the BARF model (Bias Annealing at the Replication Fork; Fig. 6), is sufficient, we reasoned that the phage annealing proteins may be able to mediate ssOR near any replication fork. Testing of this idea in a eukaryotic context provided by mouse ES cells proved encouraging. More work is required to determine whether conditions can be found for useful application of oligo repair mediated by the phage proteins in eukaryotic mutagenesis. Also an interesting possibility to explore involves the use of ssOR strand bias assays using the phage annealing proteins to map replication origins in eukaryotic cells.

The limitations and utility of ssOR were examined by use of a variety of differently configured oligos. Several conclusions emerged. First, the efficiency of ssOR requires hybrid formation either side of the repair site and efficiencies increased up to an apparent optimum of 55–65 symmetrically on each side. Shifting the repair site off center reduced repair efficiency, however shifting the repair site towards the 5' end of the oligo was worse than shifting to the 3' end. This indicates that the stability of the 5' end of the oligo hybrid is more important than the stability of the 3' oligo hybrid and suggests that the 5' end of the oligo hybrid serves to block destabilizing influences, such as 5'-3' helicases or polymerases, and the 3' end of the oligo hybrid serves mainly as a primer for DNA replication. Second, additional mismatches in the oligo decreased repair efficiencies. Furthermore, when these oligos did effect repair, the additional mismatch was not included in most cases. Hence the oligo itself is subject to further editing.



oligo length	ha	total CmR colonies	total KanR colonies	% of deletion
46	20	2.2×10^8	1.7×10^4	0.0077
76	35	2.4×10^8	3.6×10^5	0.15
106	50	2.4×10^8	7.5×10^6	3.13
136	65	2.3×10^8	2.3×10^6	1.00
166	80	2.2×10^8	8.6×10^5	0.39

Figure 4

Efficient deletion in a BAC using ssOR. The illustration at the top shows *E. coli* HS996 harbouring the pBAD-γβα plasmid and a 143 kb BAC containing mouse genomic MII DNA. This BAC had previously been modified by Red/ET recombination to carry a cassette containing the *neo* gene interrupted at its *NcoI* site by the zeocin resistance gene as depicted. A series of oligos, listed at the bottom, symmetrically centred on the *NcoI* site were electroporated into the *E. coli* host and plated to score ssOR frequencies. In this case, ssOR will delete the approximately 800 bp zeocin gene to restore kanamycin resistance. Total colony numbers were assessed by counting chloramphenicol resistance conveyed by the BAC vector.

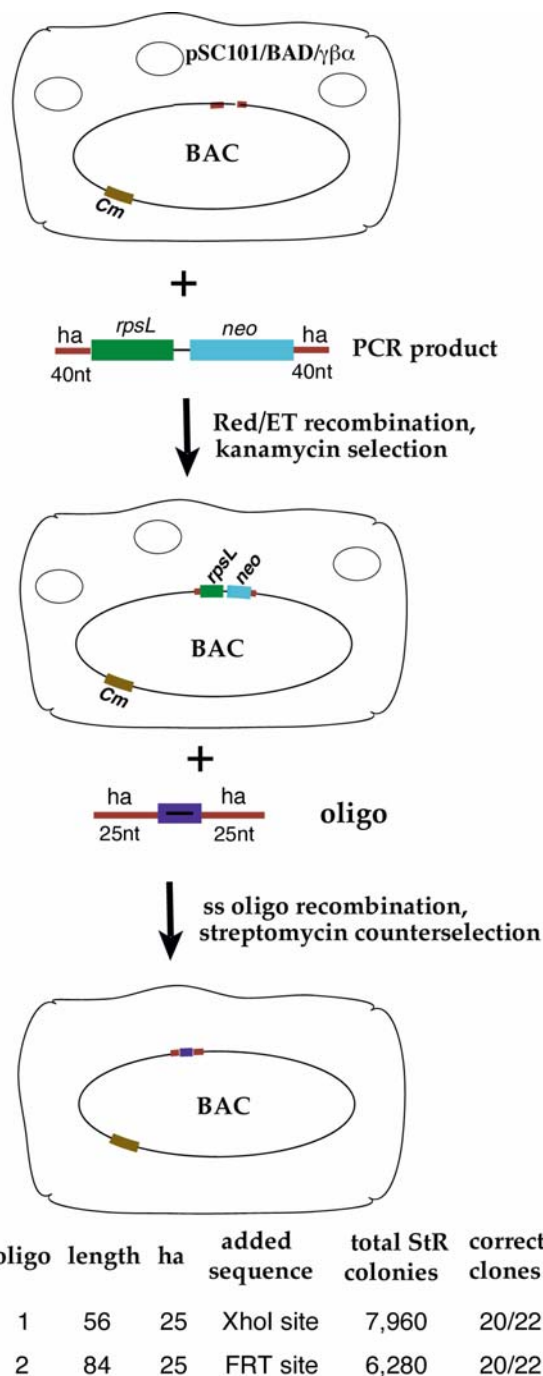


Figure 5
A counterselection strategy for engineering with ssOR. The illustration at the top shows *E. coli* HS996 harbouring the low copy, temperature sensitive, expression plasmid, pSC101/BAD/γβα and the 143 kb mouse genomic MII BAC. In the first step, a PCR product of the *neo* and *rpsL* genes (50) flanked by two 50 nt homology arms sited 908 bps apart in the BAC was introduced into the BAC by selection for kanamycin resistance. In the second step, either of the two oligos listed below were used for ssOR with selection for loss of *rpsL* and restoration of streptomycin resistance. Both oligos had 25 nts of homology to the BAC immediately flanking the *rpsL-neo* cassette and included either 6 nts of an XhoI site or 34 nts of an FRT. In addition, the temperature sensitive pSC101 plasmid was eliminated by culture at 37°C. BAC DNAs prepared from 22 streptomycin resistant colonies for each experiment were analyzed by restriction enzyme digestions and in both cases, 20 were correct.

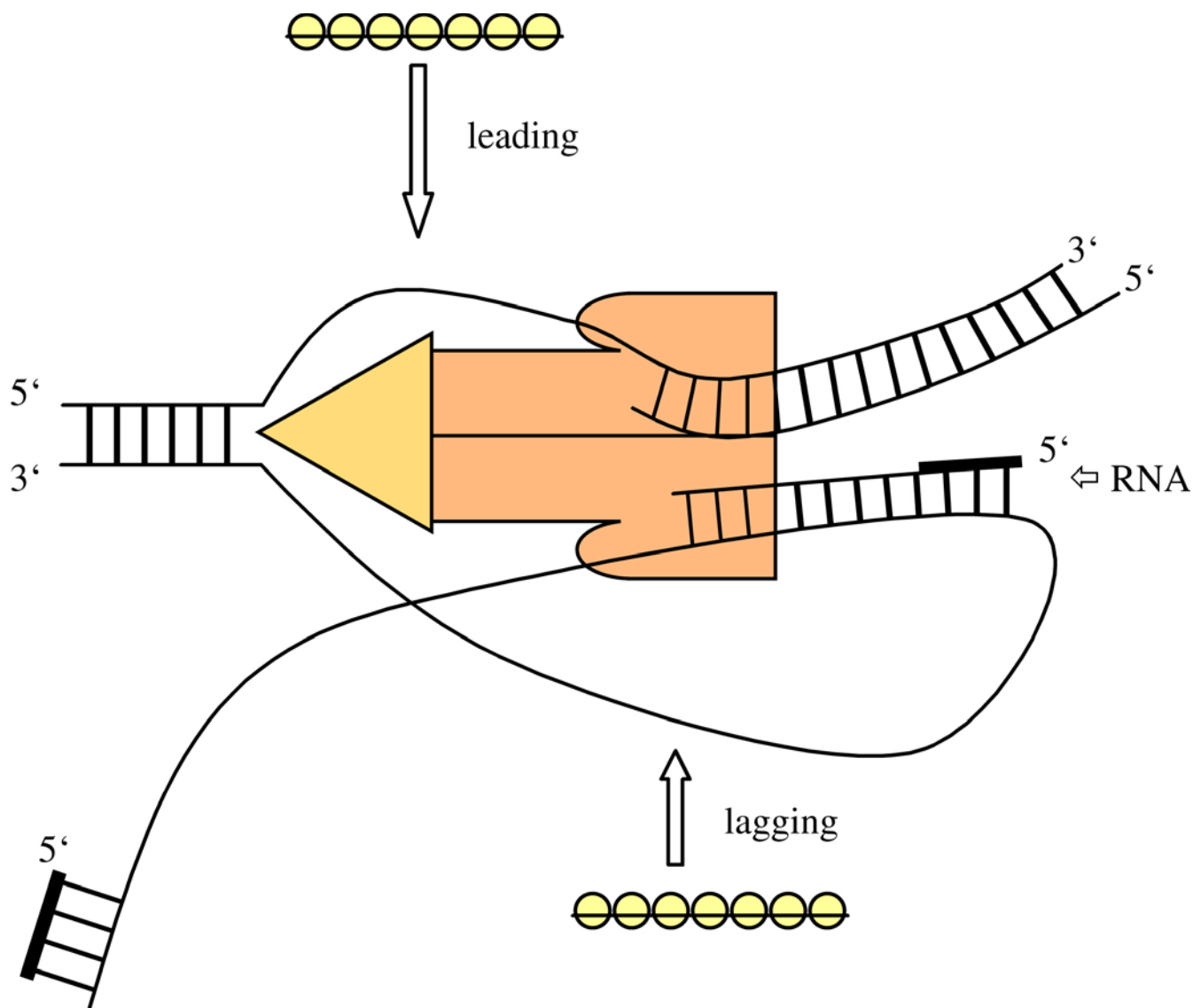


Figure 6
The BARF (Bias Annealing at the Replication Fork) model. A model for the replication fork, adapted from [34] is illustrated showing a complex between the helicase DnaB (yellow triangle) and DNA PolIII (blue shape) proceeding to the left, with leading strand product emerging at the above right. On either side of DnaB, single stranded DNA is depicted as being exposed before entry into DNA PolIII for 5'-3' synthesis in a similar way for both strands. For this to be possible, the single stranded lagging strand template must make a loop. The complex of phage annealing protein and ss oligo is depicted above and below with the exposed ss template regions indicated by arrows.

Third, the relationship between oligo length and repair efficiencies was measured in two different assays. Both assays revealed optimums of repair efficiency lying between 105 and 130 nts. Oligo repair was not detected with oligos less than 22 nts. Intriguingly, these upper and lower homology length limits concord with the homology length requirements for Red/ET DSBR, which detected a little DSBR first at 27 bps homology arms and peaked at

around 120 bps [1,24]. The similar relationship between nucleotide length and efficiency in both DSBR and ssOR suggests that the two mechanisms share a common aspect. This common aspect could be due to the implicit binding properties of the annealing proteins for single stranded DNA. The *in vitro* properties of Redβ and RecT lend some support to this speculation. For Redβ, stable single strand DNA binding *in vitro* requires at least 28 nts

[35]. Also for Red β , electron microscopic observations led to a model in which homomultimeric wheels promote annealing of bound ss DNA concomitant with a transition to a helical Red β -ds DNA filament of approximately 100 bps per turn [24]. Whether that observation relates to the apparent optima for both DSBR and ssOR mediated by the phage annealing proteins remains to be determined. An interesting analogy to cite here is the case of the poly A tail binding protein (PABP2), whose homomultimerization binding characteristics correlate with the natural lengths of poly A tails [36].

As discussed above, the BARF model is the simplest way to account for ssOR and its distinctive characteristic of strand bias. This model posits that the strand bias arises from the difference of exposed single strandedness on the lagging and leading strand templates. A simple annealing mechanism at the replication fork is attractive since most evidence shows that these phage single strand binding proteins act by annealing [26]. Whether they can also act to promote strand invasion is not clear, although evidence indicating strand displacement *in vitro* has been presented [27,28]. When analyzing Red/ET DSBR, we presented functional evidence that two mechanisms were encompassed, one by classical annealing and another by a different mechanism. However our evidence did not prove a strand invasion mechanism [24]. Given the possibility that ssOR represents a subsection of the DSBR activity, we speculate that Red/ET DSBR is based on annealing to single stranded regions at the replication fork. If so, then this would provide another link between homologous recombination and DNA replication [37–40].

Although appealingly simple, the BARF annealing at the replication fork model may not be correct and needs to be challenged. One reason for doubt can be found in the evidence that there is little, or no, single stranded exposure on the leading strand. Kim *et al.*, [41] showed that DnaB, the helicase at the replication fork, makes a specific protein-protein contact to, and is driven by, DNA Pol III in the direction of leading strand synthesis. Hence there may be no exposed ss DNA on the leading strand and a much larger strand bias than the 2 – 10 fold observed would be expected. Also, the model implies that the phage proteins can efficiently displace SSB loading onto the lagging strand template [34]. Possibly the BARF model is correct but ssOR as observed here encompasses a second mechanism that may, or may not, show a strand bias.

Interestingly, oligo repair with a significant strand bias has been previously observed in *S. cerevisiae* [42]. The major annealing protein in eukaryotes is RAD 52 and Passy *et al* [23] proposed a functional similarity between RAD52 and Red β . However Yamamoto *et al* [43] showed that oligo repair in *S. cerevisiae*. does not require Rad52. It will be in-

teresting to determine what annealing activity in *S. cerevisiae* is responsible.

Conclusions

Phage annealing proteins can initiate the recombination of single stranded oligonucleotides into endogenous targets in *Escherichia coli* at very high efficiencies. This expands the repertoire of useful DNA engineering strategies, shows promise for applications in eukaryotic cells, and has implications for the unanswered questions regarding DSBR mediated by RecE/RecT and Red α /Red β .

Methods

All ssOR methods were essentially the same as previously described for Red/ET Recombination in *E. coli* [24] except that cells were electroporated with 50 pmoles of oligonucleotide. For normalizing the transformation efficiency in the plasmid assays, aliquots were plated on ampicillin (50 μ g/ml) as well as kanamycin (50 μ g/ml) plates. Efficiencies of ssOR were calculated as the ratio of kanamycin to ampicillin resistant colonies. All results shown are from single experiments performed in triplicate, each performed with a single batch of competent cells, or when comparing different phage proteins, single batches made in parallel. The experiments of Figures 1 and 2 were independently repeated and gave similar quantitative, and identical qualitative, results. The experiments of Figures 1,2 and 3 were performed in DH10B strain and HS996 strain was obtained from Research Genetics.

Oligonucleotides

Most oligo sequences can be deduced from the descriptions in the text and the sequence of the longest lagging oligo used, namely the 160 mer of Fig. 2, which, with central Nco1 site in bold, is –
 5'GCATCAGGGGCTCGCGCCAGCCGAACGTTCGCCA
 CGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCG
 TCGTGACCCATGGCGATGCCTGCTTGCCGAATATCAT-
 GGTGAAAATGGCCGCTTTTCTGGATTCATCGA 3'

The silent point mutant oligos used in the experiment of Fig. 3C are underlined and bold, being C to T to destroy an Sph1 site and G to A to create a BglII site.

The oligos used to replace rpsL-neo cassette in the BAC were:

XhoI site: 5' GTCGTAACCTGTAGAACGGAGTAACCTC-
 GAGTGAATTGTAATACGACTCACTATAG 3'

FRT site: 5'
 GTCGTAACCTGTAGAACGGAGTAACGAAGTTCCTATTC
 TCTAGAAAGTATAGGAACCTTCTGAATTGTAATACGACT-
 CACTATAG 3'

Plasmids

pGKneo was linearised with NcoI. To generate pGKneo* or pGKneo Δ , the 5' overhangs of the NcoI site were filled in using Klenow and nucleotides, or removed using Mung Bean nuclease according to supplier's instructions (New England Biolabs), followed by ligation to generate an intact circular plasmid. To generate pGKneo Δ 15, pGKneo Δ 33 and pGKneo Δ 60, NcoI digested pGKneo was treated with BAL-31 nuclease according to manufacturer's instructions (New England Biolabs).

pSC101-BAD- $\gamma\beta\alpha$ (tet) was generated by PCR of the lambda red operon and cloning into pBAD24 (Invitrogen) between NcoI and HindIII sites to form pBAD- $\gamma\beta\alpha$. The pSC101 origin and repA gene were copied by PCR using pMAK-705 (49) plasmid as a template. The tetracyclin gene flanked by homology arms to the pSC101 ori + repA cassette was generated by a PCR reaction from pBR322, and the two PCR fragments were recombined in YZ2000 cells (6) to generate pSC101-tet. A pSC101-repA-tet cassette flanked with homology arms to the red operon of pBAD- $\gamma\beta\alpha$ was generated by PCR reaction using pSC101-tet as a template. pSC101-BAD-gba (tet) was generated by replacing the ColE1 ori plus amp in pBAD-gba with pSC101-repA-tet cassette mediated by Red/ET recombination (6).

pBAD_{Erf} was generated by PCR using P22 phage DNA as a template and insertion into pBAD24 followed by sequencing.

The PGKneo* cassette from pGKneo* was inserted into the Bst 1107I site of the pcDNA31/Hygro mammalian expression vector (Stratagene) to form pcDNA/PGK-neo*. The red β gene was inserted under the CMV promoter between NheI and HindIII sites in pcDNA-PGKneo* to form pcDNA-red β /PGK-neo*.

ES cells

After linearization, 20 μ g of pcDNA/PGK-neo* or pcDNA-red β /PGK-neo* were electroporated into ES cells as described [3], followed two days later by selection with hygromycin (200 μ g/ml). Twelve independent clones were picked from each transfected line. For repairing the neo*, 1 nmol of repairing oligo carrying 28 nt of homology either side of the NcoI site was used for electroporation and selection with G418 (80 μ g/ml) using the same conditions.

Authors' contributions

Y.Z. performed the experiments of Figs. 1, 2, 4, 5 and Table 2.

J.P.P.M. performed the experiments of Figs 1, 2, 3 and Table 1.

J.R. provided technical assistance.

A.F.S. provided supervision and wrote the paper.

Acknowledgements

This work was partly funded by a grant from the VW Foundation, Program on Conditional Mutagenesis. It was initiated at EMBL, Heidelberg, and continued at Gene Bridges GmbH, and the Technische Universitaet Dresden after A.F.S. moved from EMBL in the middle of 2001.

References

- Zhang Y, Buchholz F, Muyrers JPP and Stewart AF **A new logic for DNA engineering using recombination in *E. coli***. *Nature Genetics* 1998, **20**:123-128
- Murphy KC **Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli***. *J Bacteriol* 1998, **180**:2063-71
- Angrand PO, Daigle N, van der Hoeven F, Schoeler HR and Stewart AF **Simplified generation of targeting constructs using ET recombination**. *Nucleic Acids Res* 1999, **27**:e16
- Muyrers JPP, Zhang Y, Testa G and Stewart AF **Rapid modification of Bacterial Artificial Chromosomes by ET-recombination**. *Nucleic Acids Res* 1999, **27**:1555-1557
- Muyrers JPP, Zhang Y, Benes V, Testa G, Ansoerge W and Stewart AF **Point mutation of bacterial artificial chromosomes by ET recombination**. *EMBO Reports* 2000, **1**:239-243
- Zhang Y, Muyrers JPP, Testa G and Stewart AF **DNA cloning by homologous recombination in *Escherichia coli***. *Nature Biotechnology* 2000, **18**:1314-1317
- Narayanan K, Zhang Y, Stewart AF, Williamson R and Ioannou PA **Efficient and precise engineering of a 200 kb b-globin human/bacterial artificial chromosome in DH10B using an inducible homologous recombination system**. *Gene Therapy* 1999, **6**:432-441
- Datsenko KA and Wanner BL **One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products**. *Proc Natl Acad Sci USA* 2000, **97**:6640-6645
- Hill F, Benes V, Thomasova D, Stewart AF, Kafatos FC and Ansoerge W **BAC trimming; minimizing clone overlaps**. *Genomics* 2000, **64**:111-113
- Murphy KC, Campellone KG and Poteete AR **PCR-mediated gene replacement in *Escherichia coli***. *Gene* 2000, **246**:321-330
- Yu D, Ellis HM, Lee EC, Jenkins NA, Copeland NG and Court DL **An efficient recombination system for chromosome engineering in *Escherichia coli***. *Proc Natl Acad Sci USA* 2000, **97**:5978-83
- Lee EC, Yu D, Martinez de Velasco J, Tessarollo L, Swing DA, Court DL, Jenkins NA and Copeland NG **A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA**. *Genomics* 2001, **73**:56-65
- Swaminathan S, Ellis HM, Waters LS, Yu D, Lee EC, Court DL and Sharan SK **Rapid engineering of bacterial artificial chromosomes using oligonucleotides**. *Genesis* 2001, **29**:14-21
- Casanova E, Lemberger S, Fehsenfeld S, Greiner E, Stewart AF and Schutz G **Rapid localisation of a gene within BACs and PACs**. *Biotechniques* 2002, **32**:240-2
- Zhang P, Li MZ and Elledge SJ **Towards genetic genome projects: genomic library screening and gene-targeting vector construction in a single step**. *Nat Genet* 2002, **30**:31-39
- Muyrers JPP, Zhang Y and Stewart AF **Recombinogenic Engineering: new options for manipulating DNA**. *Trends in Biochemical Sciences* 2001, **26**:325-331
- Poteete AR **What makes the bacteriophage lambda Red system useful for genetic engineering: molecular mechanism and biological function**. *FEMS Microbiol Lett* 2001, **201**:9-14
- Copeland NG, Jenkins NA and Court DL **Recombineering: a powerful new tool for mouse functional genomics**. *Nat Rev Genet* 2001, **2**:769-779
- Paques F and Haber JE **Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae***. *Microbiol Mol Biol Rev* 1999, **63**:349-404

20. Shashikant CS, Carr JL, Bhargava J, Bentley KL and Ruddle FH **Recombinogenic targeting: a new approach to genomic analysis – a review.** *Gene* 1998, **223**:9-20
21. Kovall R and Matthews BW **Toroidal structure of lambda-exonuclease.** *Science* 1997, **277**:1824-7
22. Thresher RJ, Makhov AM, Hall SD, Kolodner R and Griffith JD **Electron microscopic visualization of RecT protein and its complexes with DNA.** *J Mol Biol* 1995, **254**:364-71
23. Passy SI, Yu X, Li Z, Radding CM and Egelman EH **Rings and filaments of beta protein from bacteriophage lambda suggest a superfamily of recombination proteins.** *Proc Natl Acad Sci USA* 1999, **96**:4279-84
24. Muylers JPP, Zhang Y, Buchholz F and Stewart AF **RecE/RecT and Red α /Red β initiate double stranded break repair by specifically interacting with their respective partners.** *Genes and Development* 2000, **14**:1971-1982
25. Stahl MM, Thomason L, Poteete AR, Tarkowski T, Kuzminov A and Stahl FW **Annealing vs. invasion in phage lambda recombination.** *Genetics* 1997, **147**:961-77
26. Kuzminov A **Recombinational repair of DNA damage in Escherichia coli and bacteriophage lambda.** *Microbiol Mol Biol Rev* 1999, **63**:751-813
27. Noiroit P and Kolodner RD **DNA strand invasion promoted by Escherichia coli RecT protein.** *J Biol Chem* 1998, **273**:12274-80
28. Li Z, Karakousis G, Chiu SK, Reddy G and Radding CM **The beta protein of phage lambda promotes strand exchange.** *J Mol Biol* 1998, **276**:733-44
29. Ellis HM, Yu D, DiTizio T and Court DL **High efficiency mutagenesis, repair and engineering of chromosomal DNA using single-stranded oligonucleotides.** *Proc Natl Acad Sci USA* 2001, **98**:6742-6
30. Muniyappa K and Radding CM **The homologous recombination system of phage lambda. Pairing activities of beta protein.** *J Biol Chem* 1986, **261**:7472-8
31. Murphy KC **Lambda Gam protein inhibits the helicase and chi-stimulated recombination activities of Escherichia coli RecBCD enzyme.** *J Bacteriol* 1991, **173**:5808-21
32. Poteete AR and Fenton AC **Efficient double-strand break-stimulated recombination promoted by the general recombination systems of phages lambda and P22.** *Genetics* 1993, **134**:1013-21
33. Murphy KC **Bacteriophage P22 Abc2 protein binds to RecC increases the 5' strand nicking activity of RecBCD and together with lambda bet, promotes Chi-independent recombination.** *J Mol Biol* 2000, **296**:385-401
34. Neidhardt FC, Curtiss III R, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M and Umberger HE **Escherichia coli and Salmonella.** Eds. *ASM Press, Washington, DC* 1:749-63
35. Mythili E, Kumar KA and Muniyappa K **Characterization of the DNA-binding domain of beta protein, a component of phage lambda red-pathway, by UV catalysed cross-linking.** *Gene* 1996, **182**:81-7
36. Keller RW, Kuhn U, Aragon M, Bornikova L, Wahle E and Bear DG **The nuclear poly (A) binding protein, PABP2, forms an oligomeric particle covering the length of the poly (A) tail.** *J Mol Biol* 2000, **297**:569-83
37. Kogoma T **Stable DNA replication: interplay between DNA replication, homologous recombination and transcription.** *Microbiol Mol Biol Rev* 1997, **61**:212-238
38. Kuzminov A and Stahl F **Double-strand end repair via the RecBC pathway in Escherichia coli primes DNA replication.** *Genes Dev* 1999, **13**:345-56
39. Cox MM, Goodman MF, Kreuzer KN, Sherratt DJ, Sandler SJ and Marians KJ **The importance of repairing stalled replication forks.** *Nature* 2000, **404**:37-41
40. Marians KJ **Replication and recombination intersect.** *Curr Opin Genet Dev* 2000, **10**:151-6
41. Kim S, Dallmann HG, McHenry CS and Marians KJ **Coupling of a replicative polymerase and helicase: a tau-DnaB interaction mediates rapid replication fork movement.** *Cell* 1996, **84**:643-50
42. Yamamoto T, Moerschell RP, Wakem LP, Komar-Panicucci S and Sherman F **Strand-specificity in the transformation of yeast with synthetic oligonucleotides.** *Genetics* 1992, **131**:811-9
43. Yamamoto T, Moerschell RP, Wakem LP, Ferguson D and Sherman F **Parameters affecting the frequencies of transformation and co-transformation with synthetic oligonucleotides in yeast.** *Yeast* 1992, **8**:935-48
44. Gillen JR, Willis DK and Clark AJ **Genetic analysis of the RecE pathway of genetic recombination in Escherichia coli K-12.** *J Bacteriol* 1981, **145**:521-32
45. Willetts NS and Clark AJ **Characterization of some multiply recombination-deficient strains of Escherichia coli.** *J Bacteriol* 1969, **100**:231-9
46. Hanahan D **Studies on transformation of Escherichia coli with plasmids.** *J Mol Biol* 1983, **166**:557-80
47. Cohen A and Clark AJ **Synthesis of linear plasmid multimers in Escherichia coli K-12.** *J Bacteriol* 1986, **167**:327-35
48. Lovett ST and Clark AJ **Genetic analysis of the recJ gene of Escherichia coli K-12.** *J Bacteriol* 1984, **157**:190-6
49. Hashimoto-Gotoh T and Sekiguchi M **Mutations of temperature sensitivity in R plasmid pSC101.** *J Bacteriol* 1977, **131**:405-412
50. Russell CB and Dahlquist FW **Exchange of chromosomal and plasmid alleles in Escherichia coli by selection for loss of a dominant antibiotic sensitivity marker.** *J Bacteriol* 1989, **171**:2614-8

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

