Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 4768–4772, July 1985
Genetics

Homology requirements for recombination in *Escherichia coli*

(length dependence/insulin gene/synthetic oligonucleotides/supF miniplasmid)

VALERIE M. WATT*†, C. JAMES INGLES*‡, MICKEY S. URDEA§, AND WILLIAM J. RUTTER*

*Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143; †Department of Physiology, University of Toronto, Toronto, Ontario, Canada M5S 1A8; ‡Banting and Best Department of Medical Research, University of Toronto, Toronto, Ontario, Canada M5G 1L6; and §Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608

Contributed by William J. Rutter, March 25, 1985

ABSTRACT **The DNA sequence homology required for recombination in *Escherichia coli* has been determined by measuring the recombination frequency between insulin DNA in a miniplasmid πVX and a homologous sequence in a bacteriophage λ vector. A minimum of ≈20 base pairs in a completely homologous segment is required for significant recombination. There is an exponential increase in the frequency of recombination when the length of homologous DNA is increased from 20 base pairs to 74 base pairs and an apparently linear increase with longer DNA segments. Mismatches within a homologous segment can dramatically decrease the frequency of recombination. Thus, the process of recombination is sensitive to the length of precisely base-paired segments between recombining homologues.**

Synapsis of DNA molecules during recombination in *Escherichia coli* involves the homologous alignment of single-stranded DNA with duplex DNA to produce a nascent heteroduplex structure (for review, see ref. 1). In contrast to the presynaptic reaction and to strand exchange, synapsis is a rapid process (2, 3). The first stage in synapsis, conjunction (4), involves the invasion of duplex DNA by single-stranded DNA and is independent of sequence homology; subsequently, during the second stage of synapsis, homologous alignment of complementary sequences occurs either by facilitated diffusion or by translocation of the DNA molecules in a ternary complex with the RecA protein (4). The initial structure between paired complementary sequences is thought to be a heteroduplex molecule in which the DNA strands form hydrogen bonds but are not interwound in a stable helix (5). A transient nick then allows interwinding of the two strands; this converts the unstable heteroduplex into a stable joint molecule. Subsequently, strand exchange occurs forming long heteroduplex joints between the DNA molecules.

The degree of sequence homology required for alignment of complementary strands of DNA during synapsis has not been defined. It has been suggested that >30 base pairs (bp) are required for the formation of the recombination intermediate, because DNA molecules with 151 bp of homologous sequence could be paired, but those with only 30 bp of homology could not be paired *in vitro* by RecA (4). Although not dependent on RecA protein, 50 bp of homology appeared to be required for the primary pathway of recombination between the *rII* cistrons of bacteriophage T4 (6). Recombination did occur below this threshold, albeit at lower efficiency and perhaps involving a second mechanism (6). To systematically assess the dependence of recombination on short lengths of homology, we have chemically synthesized oligonucleotides containing up to 53 bases of human insulin DNA. We have studied the effects both of length and of sequence mismatch on the frequency of recombination between the plasmid πVX, containing an easily assayed suppressor and the synthetic DNA, and bacteriophage λ carrying the insulin gene. Recombination can then be detected by measuring phage growth in an appropriate indicator cell. This technique of measuring *in vivo* recombination was developed by Seed (7) as a facile method for screening genomic libraries in bacteriophage λ. We have carried out an analysis of recombination with this system in order to characterize the process of homologous recombination in *E. coli* and also to assess the use of recombinational screening to isolate specific genomic DNA sequences.

METHODS

Bacterial Strains and Bacteriophage. *E. coli* strains W3110r$^-$m$^+$, W3110r$^-$m$^+$(P3), W3110r$^-$m$^+$(P3)πVX, NK5486(*lacZ am*), MC1061(P3), and MC1061(P3)πAN7 were provided by Seed (7). The λ Charon 4A bacteriophage (*red*$^-$ *gam*$^-$) containing the human insulin gene was λhI1 described by Bell *et al.* (8). The λ Charon 4A bacteriophage carrying the rat insulin I gene (λrI1) was isolated from the rat Sargent library (9) by M. Crerar and, apart from the absence of the 6.2-kilobase-pair (kb) *Eco*RI fragment, is similar to the phage rI1 isolated from this same library by Lomedico *et al.* (10). We confirmed by DNA sequencing that this phage contained the authentic rat insulin I gene.

Synthesis of Oligonucleotides. The different oligonucleotides were synthesized for these studies by the phosphoramidite method (11) as described by Urdea *et al.* (12). Oligonucleotides were synthesized beginning with the four bases, ACGT, the *Pst* I cohesive end that was used in subcloning (below), followed by bases complementary to human insulin. To obtain sequences of various lengths complementary to human insulin, aliquots of the silica solid support were removed at different steps of nucleotide addition between 20 bases and 53 bases (Fig. 1*A*).

Miniplasmid Probe Constructions. To introduce a blunt cloning site into the πVX miniplasmid (7), a *Hpa* I site was constructed between the *Bgl* II and *Hin*dIII sites of the polylinker. This plasmid is termed πVb. Equimolar amounts (75 ng/μl) of two synthetic single strands of DNA encoding a *Hpa* I site and either a *Bgl* II or *Hin*dIII cohesive end were annealed by heating to 95°C for 2 min and cooling to room temperature. They were then ligated in 25-fold molar excess with πVX DNA previously cut with *Bgl* II and *Hin*dIII and used to transform MC1061(P3) cells. The chemically synthesized insulin DNA fragments were cloned between this *Hpa* I site and the *Pst* I site of πVb. The synthetic single-stranded DNAs encoding part of the insulin gene were annealed at 14°C at a molar ratio of 1:5 with a 12-bp oligonucleotide complementary to the insulin DNA immediately 5′ to the *Pst* I cohesive end by heating at 95°C for 2 min and cooling to

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: bp, base pair(s); kb, kilobase pair(s).

room temperature; they were then ligated in 25-fold molar excess to πVb DNA cut with *Hpa* I and *Pst* I, filled in with Klenow DNA polymerase, and religated. Insulin DNA sequences longer than 53 bp were generated by appropriate restriction enzyme digestion of the insulin gene DNA (Fig. 1 *D* and *E*). A 313-bp *Pst* I fragment encoding the B chain of human insulin was inserted into the *Pst* I site of πVX. The other large fragments were generated by cutting the 313-bp *Pst* I fragment with *Alu* I for the 74-bp and 239-bp pieces or with *Hae* III for the 56-bp and 143-bp fragments and subcloning these fragments into πVb cut with *Pst* I and *Hpa* I. A 224-bp *Pst* I fragment encoding the B chain of anglerfish insulin (13) was also subcloned into πVX. πAN7 probes were constructed by inserting *Eco*RI fragments of πVb miniplasmids that contained the polylinker and the synthetic insulin DNA into the *Eco*RI site of πAN7.

Characterization of πVb Insulin Constructions. Plasmid DNAs were prepared (14) from saturated cultures in L-broth containing ampicillin and tetracycline. DNA was digested with the appropriate enzymes and the sizes of vector and insert fragments were characterized by polyacrylamide electrophoresis. To confirm the sequence of the synthetic insulin DNA constructions, *Eco*RI fragments of πVb probes containing the insulin fragments were subcloned into M13 mp8 replicative form DNA. Both strands of each insulin probe were sequenced; single-stranded DNA templates, dideoxy-sequencing reactions, and sequencing gels were prepared essentially as described (15).

Phage–Plasmid Recombination. Aliquots of 10^6 phage were used to infect 0.25 ml of overnight cultures of MC1061(P3) cells or W3110(P3) cells containing the πVb-insulin miniplasmids and plated on single 100-mm plates as described by Seed (7). Phage harvested from these plates were plated on K803 cells (*SupE, SupF*) to determine the total number of phage and on the *sup*^{0} host, NK5486, to determine the number of recombinant phage. The observed frequencies of recombination cannot be attributed to different growth characteristics of recombinant phage, because recombinants containing fragments of 20, 53, 143, and 313 bp each exhibited a similar burst size and rate of replication. Furthermore, since none of the insulin sequences used in the πVb constructions contain a *chi* sequence, this sequence could not have contributed to observed differences in recombination frequency. The phage that grew on the *sup*^{0} host, NK5486, were shown to represent bonafide recombinants. All recombinants between λhI1 and either the 313- or the 20-bp probe (>100 of each were tested) made blue plaques on the *sup*^{0} *lacZ am* host in the presence of isopropyl β-D-thiogalactoside and 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside, thus indicating the presence of the *supF* gene. None of the phage in these screens was the result of reversion of the amber mutations in A or B λ phage functions. Where indicated, individual plaques formed by bacteriophage containing the *supF* gene were isolated, plate stocks were prepared, and DNA was made from these phage (16).

RESULTS

Construction of πVb Miniplasmids Containing Short Fragments of Synthetic Insulin DNA. To study the dependence of recombination frequency on the length of homologous DNA, we chemically synthesized oligonucleotides of various lengths containing from 16 to 53 bases (Fig. 1*A*). The sequences were homologous to human insulin DNA so that we could use the human insulin gene, already present in a λ bacteriophage vector (8), in this study. The largest synthetic fragment, 53 bases, represented nucleotides 500–553 in the sequence of human DNA reported by Bell *et al.* (8) and encoded the first 17 amino acids of insulin. The shorter DNAs were synthetic intermediates of the 53-base fragment and began at its 5′ end.

We cloned the synthetic oligonucleotides into πVb, a plasmid derived from the miniplasmid πVX, which was specifically designed for detecting *in vivo* recombination (7). Some of the probes when subcloned into πVb were heterogenous both in length and in sequence. For example, only 25% of the clones generated from the aliquot removed after addition of the 53rd base contained full-length 53-bp inserts. Of four 53-bp fragments sequenced, two contained the correct 53-bp sequence; one contained a single nucleotide deletion at position 25 (πhI-53:Δ25), and one had a cytosine to thymine transition at 17 bp (πhI-53:C→T^{17}). This heterogeneity was likely generated during the oligonucleotide synthesis, because the sequences of each of these different

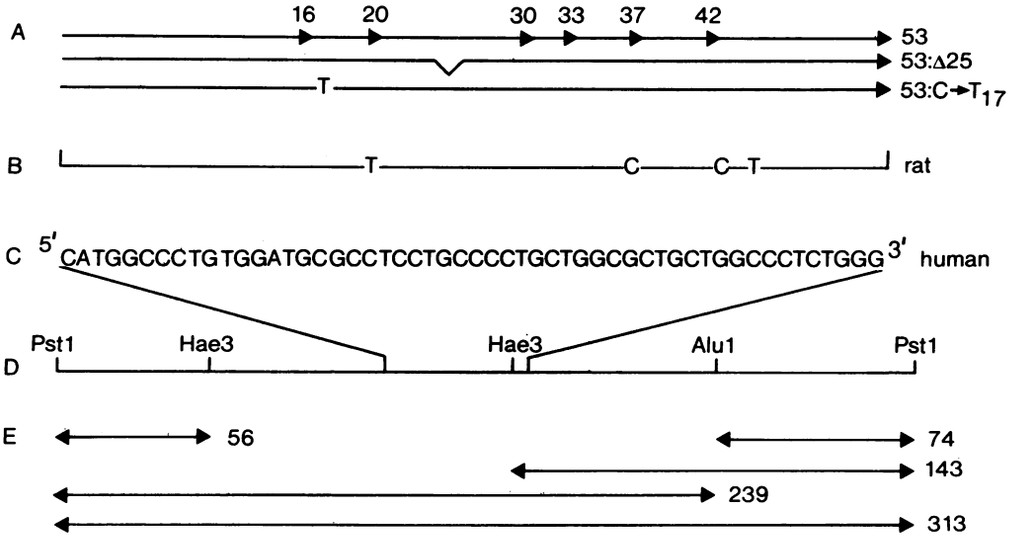

FIG. 1. Insulin DNA. (*A*) Synthetic DNA fragments inserted into the πVb miniplasmid. Numbers above the sequence indicate length of the insulin DNAs. Alterations from the human insulin sequence are shown. (*B*) Rat insulin DNA in λrI1 with mismatches to human insulin DNA. (*C*) Sequence of human insulin DNA fragment showing the complement of the synthetic fragment. (*D*) Map of human insulin genomic DNA showing the site of the synthetic DNA and restriction enzyme sites used for the generation of fragments inserted into the miniplasmid. (*E*) Position and length of restriction fragments inserted into the miniplasmid.

probes were identical in at least two independent subclones in M13. These variant sequences have proven useful in assessing the effects of specific mismatches on recombination frequency.

Dependence of Recombination Frequency on Length of Homology. The πVb probes containing the synthetic DNAs, as well as longer DNA fragments generated by restriction enzyme digestion of the human insulin gene (Fig. 1), together with a λ Charon 4A vector containing the human insulin gene were used to determine the dependence of recombination frequency on length of homology. The πVb miniplasmids, in addition to the insulin DNA, carry the selectable *E. coli* tyrosine tRNA amber-suppressor gene, *supF*. Recombination *in vivo* between πVb insulin probes and bacteriophage that carry the insulin gene yields bacteriophage carrying the *supF* gene. These recombinant bacteriophage can be readily distinguished from other phage by plating on an appropriate indicator strain of *E. coli* lacking the *supF* gene, such as NK5486. This simple plaque assay was used to measure the frequency of recombination.

The frequency of recombination between the πVb plasmids and the λ bacteriophage increased with the length of homology in a biphasic manner with a breakpoint at ≈74 bp (Fig. 2). The recombination frequency with probes from 74 to 313 bp exhibited an approximately linear dependence on length; a 3-fold increase in frequency with the 313-bp probe was observed when compared to the 74-bp probe. Below 74 bp there was a dramatic decline in the recombination frequency with decreasing length of homology. There was a reduction by a factor of ≈10 in recombination at each step as the homologous DNA sequence was reduced from 74 to 53, to 30, and to 20 bp (Fig. 2). The recombination frequency of a 16-bp probe with λhI1 was 1/100th that of a 20-bp probe and was not significantly above the background recombination

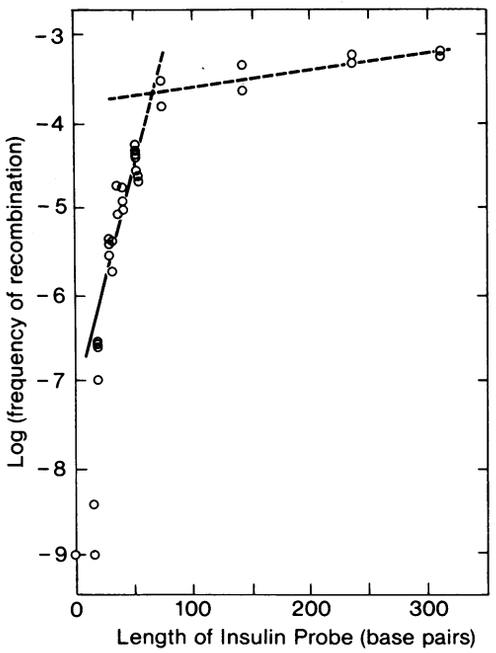

FIG. 2. Frequency of recombination as a function of length of homologous DNA. The frequencies of recombination between insulin DNA fragments in the miniplasmid and the insulin gene in λhI1 have been plotted against the length of the insulin DNA in the probe. Each experimental point represents a single determination of recombination frequency from separate recombination experiments. By straight line regression, for lengths from 20 to 74 bp, $r = 0.93$ by using semilogarithmic analysis; for lengths from 74 to 313 bp, $r = 0.84$ by using semilogarithmic analysis, and $r = 0.89$ by using linear analysis.

between πVb and λhI1 (Fig. 2). The recombination with probes ≥20 bp appeared to be homologous because the DNA of recombinant phage arising from four independent recombination experiments between λhI1 and the 313-bp probe or the 20-bp probe were shown to have new restriction enzyme sites consistent with the integration of the πVb insulin probes at the homologous site in the large *Eco*RI fragment (12.5 kb) containing the human insulin gene (data not shown).

Similar experiments were carried out with synthetic insulin DNA fragments (20, 30, 42, and 53 bp) inserted into the related miniplasmid, πAN7 (7), which replicates to a higher copy number. The frequency of recombination between λhI1 and each sequence in πAN7 was increased by ≈6-fold over the frequency observed with the same sequence in πVX. πAN7 was present at ≈5-fold higher copy number than πVb as assessed by DNA extraction and electrophoresis on agarose gels. Thus, recombination frequency apparently varies linearly with intracellular DNA concentration.

Effects of Mismatch on Recombination Frequency. Selected mismatches within an homologous stretch of DNA were generated either by altering the sequence of the synthetic DNA inserted into the πVb plasmids or by using the rat insulin gene instead of the human gene in the λ Charon 4A vector (Fig. 1 *A* and *B*). For these experiments, we first showed that there was no variation in *in vivo* packaging and plating efficiency or burst size between the human and rat insulin genes in λhI1 and λrI1 by using a miniplasmid probe that contained 20 bp of perfect homology to both the rat and human genes. The frequencies of recovery of recombinants were similar ($2 \times 10^{-7} \pm 0.5$, $n = 4$; $1 \times 10^{-7} \pm 0.3$, $n = 4$; $P \leq 0.1$). Thus, any differences should be due to the recombination frequency of the probe.

As shown in Table 1, mismatches in the region of DNA homology reduced the frequency of recombination. For example, a single mismatch in 53 bp reduced the frequency of recombination by a factor of 4 (compare πhI:53 X human with πhI-53:Δ25 X human and πhI-53:C→T[17] X human, $P \leq 0.005$; Table 1). Additional mismatches between plasmid and phage DNA sequences along this 53 bp were generated when these 53-bp fragments encoding the human insulin gene in πVb were used in similar experiments with the phage containing the partially homologous rat insulin DNA. These mismatches reduced the frequency of recombination to levels below those of plasmid probes whose lengths approximated

Table 1. Effects of single base nonhomologies on recombination frequency

Recombination event* (πhI- X gene)		Stretches of homology,† bp	Frequency of recombination
16	X rat	17	10^{-9}
53:C→T[17]	X rat	17, 16	5×10^{-7}
30	X human	30	4×10^{-6}
74	X rat	18	10^{-7}
143	X rat	18, 11	10^{-7}
20	X rat	20	10^{-7}
30	X rat	20, 10	10^{-7}
53:Δ25	X rat	20, 11	10^{-7}
37	X rat	20, 16	10^{-6}
53	X rat	20, 16	10^{-6}
37	X human	37	10^{-5}
53:C→T[17]	X human	36, 16	10^{-5}
53:Δ25	X human	28, 24	10^{-5}
53	X human	53	4×10^{-5}

*Recombination event describes the length and alterations of the human insulin DNA in the miniplasmid πVb, and the insulin gene in the bacteriophage; rat (in λrI1) or human (in λhI1).
†Stretches of homology >9 bp long.

the total length of the remaining two longest homologous fragments. As shown in Table 1, recombination between a homologous domain of 37 bp (πhI-37 X human) was 10-fold greater than recombination between sequences that contained regions of homology of 20 and 16 bp interrupted by a single mismatch (πhI-53 X rat). A similar relationship was observed for recombination between a homologous domain of 30 bp (πhI-30 X human) compared to recombination between stretches of DNA containing homologous lengths of 17 bp plus 16 bp (πhI-53:C→T^{17} X rat) or 20 bp plus 11 bp (πhI-53:Δ25 X rat).

By comparing the differences in recombination frequency between miniplasmid probes and phage where mismatches had been generated, the ability of additional domains of homology to affect the frequency of recombination could be determined. For example, the frequency of recombination of a 17-bp probe was increased 500-fold by the presence of an additional 16-bp region of homologous DNA (compare πhI-16 X rat with πhI-53:C→T^{17} X rat; Table 1). This increase in frequency was much greater than the summation of the recombination frequency with the homologous 17- and 16-bp fragments. Furthermore, the length of this additional domain of homologous DNA is critical. For example, the addition of either 10- or 11-bp fragments of homology did not affect the recombination frequency of 18- or 20-bp fragments (Table 1). Fragments that are 16 bp long, however, did substantially increase the recombination frequencies of the 17- and the 20-bp probes. In addition, a 24-bp fragment increased the recombination frequency of a 28-bp fragment (πhI-53:Δ25 X human) when compared to the recombination frequency of a 30-bp fragment (πhI-30 X human). It appears, however, that the degree of influence of the second domain varies inversely with the length of the first. Whereas the 16-bp domain increased the recombination frequency of the 17-bp probe by 500-fold and that of a 20-bp probe by 10-fold, the addition of a 16-bp domain of homology did not affect the recombination frequency of a 36-bp fragment. The recombination frequency with the human insulin gene of πhI-53:C→T^{17} was the same as that of πhI-37 (Table 1).

Over short distances, the distance between domains does not appear to influence the critical length of the second domain. As shown in Table 1, the lack of effect of 10- and 11-bp regions of homology on a 20-bp fragment occurred whether the regions were separated by one mismatch (πhI-30 X rat) or by 6 bp, including two mismatches (πhI-53:Δ25 X rat). Conversely, a 16-bp fragment greatly increased recombination frequencies when separated by 4 bp from a 17-bp fragment (πhI-53:C→T^{17} X rat) or by 1 bp from a 20-bp fragment (πhI-37 X rat and πhI-53 X rat).

Use of Recombination to Isolate Partially Homologous Genes from Genomic Libraries. Recombinants between the human insulin probe (πhI-313) and the rat insulin gene (λrI1) were isolated at a frequency of 10^{-6}, 1/600th the frequency of the recombinants between πhI-313 and the homologous human gene. This recombination frequency is consistent with the degree of homology between the rat and human insulin DNA sequences: there are only three stretches of homology >11 bp: 19 bp, 18 bp, and 16 bp. This 313-bp human DNA probe in the πVX miniplasmid recombined with similar efficiency to a synthetic 37-bp DNA sequence containing two regions of homologous DNA 20 and 16 bp long. The many additional small regions of homology within the 313-bp sequence of the rat and human genes contributed little to recombination frequency.

Recombinants between the anglerfish insulin probe and λhI1 were isolated at a frequency of 3×10^{-9}. This is also consistent with the limited homology between the anglerfish and human insulin DNA sequences. For example, over the 90-bp stretch of DNA coding for their respective B-chain peptides, homology is 72%; however, no more than 10 consecutive base pairs are identical. As would be predicted from this low frequency of recombination, we did not recover any recombinants between the anglerfish insulin probe and phage containing the human insulin gene in 10 different screens of a human λ Charon 4A genomic library (17). Using the 313-bp human insulin probe, however, the phage λhI1 containing the insulin gene was readily obtained from the same genomic library. Recombinants between the anglerfish insulin probe and phage that did not contain the human insulin gene were, however, isolated in these experiments. Since we observed recombination between the miniplasmid without an insert and λhI1 at a frequency of $\approx 10^{-9}$, the anglerfish–human recombinants may result from this background of illegitimate recombination. We have also observed that not all phage containing homologous sequences are isolated in a library screen. For example, we did not recover λhi2 (8) in screens of the human genomic library with the 313-bp human insulin probe. This could have been due to an underrepresentation of this particular isolate upon amplification of the library or to the constraints of DNA length on packaging after recombination.

DISCUSSION

To determine the requirements of recombination in *E. coli* for sequence homology, we have used an assay that measures the frequency of recombination between sequences cloned into a miniplasmid, πVX, and bacteriophage λ. This system was originally developed by Seed (7) as a facile method for screening genomic libraries for isolates carrying homologous genomic DNAs. Our studies provide additional information about the parameters required for using recombination to isolate bacteriophage containing DNA homologous to sequences in the plasmid. At least 74 bp of homologous sequence is required to maximize recombination frequency and to minimize the constraint on phage DNA length imposed by packaging. Seed (7) also suggested that cross-species probes might be used to isolate families of related genes. Our results using synthetic DNAs and insulin DNA from three different species suggest that application of miniplasmid screening for the isolation of only partially homologous genes is limited because the effect of mismatches on recombination is so marked. A correlation of recombination frequency with the presence of stretches of perfect homology rather than percentage of homology over longer stretches of DNA explains the recombination between different globin genes measured by Seed (7). These parameters, which govern recombination between plasmid and bacteriophage, are also likely to be applicable to the method of recombinationally screening cosmid libraries described by Poustka *et al.* (18).

We were able to discern several features of the dependence of recombination in *E. coli* on the length of sequence homology. We found that the minimum length of homology necessary to effect homologous recombination is only slightly longer than the length required for the stable interaction of DNA–DNA duplexes (19) and for the occurrence of a genetically unique oligonucleotide in the *E. coli* chromosome (20). Significant recombination could be detected only when homologous regions of the DNA molecules were at least 20 bp long; 17 bp was not enough. A significant increase in the recombination frequency by a second domain required ≈16 bp of homology; 11 bp was inadequate. For lengths >15 nucleotides, the incremental increase in Tm (the temperature of the midpoint of the thermal denaturation transition) for an oligonucleotide length increase of one nucleotide is very small *in vitro* (19); it would seem reasonable to assume that somewhat similar conditions exist *in vivo*. Thus, it appears unlikely that the large increment in recombination frequency from 16 to 20 bp can be attributed solely to DNA duplex stability. Perhaps the additional homology is required for the

formation of a stable recombination complex comprised of protein(s) as well as DNA. The length of homologous DNA we observed to be necessary for efficient recombination was also longer than the length (12 bp) of a sequence that has a predicted frequency of <1 of reoccurring by chance elsewhere in the *E. coli* genome (20). The marked reduction in recombination frequency from 20 to 16 bp of homology observed in our experiments might therefore suggest that the enzyme system involved in recombination has evolved to prevent nonhomologous recombination. On a statistical basis, a system requiring 20 bp of homology has 1/1000th the chance of recombining heterologous molecules than a system requiring only 15 bp of homology (ref. 21; see also ref. 20). The importance of a minimum recognition length for homologous pairing in guarding against the mistake of nonhomologous recombination has previously been recognized (20).

Although a stretch of only 20 bp of homologous DNA appears to be the minimum required for recombination between plasmid and phage in *E. coli*, our results suggest that efficient recombination requires ≈75 bp of homology. The frequency of recombination increases exponentially from 20 to 74 bp; at >74 bp, the frequency of recombination exhibits an approximately linear dependence on length. Thus, between 74 and 313 bp, the frequency of recombination appears to be directly proportional to the number of sites available for recombination. These results are similar to those of Singer *et al.* (6) who observed a linear dependence on length of homology between 50 and 200 bp for recombination in a T4 bacteriophage system. The recombination frequency between deletion mutants of bacteriophage T4 extrapolated to zero recombination at ≈50 bp of homology and the recombination frequency between bacteriophage T4 and plasmids containing bacteriophage T4 DNA was significantly above background when the lengths of homologous DNA were 76 but not 53 bp (6). These results agree with our observation of a sharp reduction in recombination frequency below 74 bp of homology.

Although all the components of the recombination complex are not known, current models of recombination (22) would suggest the involvement of RecA protein, perhaps in association with a helix-destabilizing protein and topoisomerase. A dependence on RecA has previously been reported for this *in vivo* recombination between miniplasmid and bacteriophage λ (7). We have also observed a requirement for RecA in the present system. The frequency of recombination between the 313-bp probe in πVX and the human insulin gene in λSep6A was reduced ≈70% in a deficient RecA strain (unpublished data). *In vitro* analysis of the RecA-dependent homologous DNA pairing (4) showed that 151 bp of homology between phage M13 and G4 DNA was adequate for homologous pairing but 30 bp was not. In our experiments, DNAs with a 30-bp homologous segment recombined 1/100th as well as DNAs with 150 bp of homology. Thus, synapsis between 30-bp homologous fragments may have been too infrequent to be detected in the *in vitro* experiments.

The exponential decrease in recombination frequency from 74 to 20 bp could reflect reduced binding of any of the components of the recombination complex, but involvement of a second less efficient pathway of recombination is not ruled out. Two pathways of recombination that involve the RecA protein, the RecBC pathway and the RecF pathway, have been described (23). It is possible that one of these pathways is utilized for recombination between DNA molecules containing only short regions of homology, while the other pathway more efficiently catalyzes recombination between DNA molecules with longer stretches of homology. By making use of appropriately mutated bacterial strains, the role of each of these pathways in plasmid–phage recombination could be assessed.

We are particularly indebted to Brian Seed for helpful discussions and for the bacterial strains used in recombinational screening. We thank Orgad Laub for insulin plasmid DNA and Michael Crerar for the phage carrying the rat insulin gene. This work was supported by a grant to V.M.W. from the J. P. Bickell Foundation, by Grant MT4649 from the Medical Research Council of Canada to C.J.I., and by National Institutes of Health Grants AM21344 and GM28520 to W.J.R.

1. Radding, C. M. (1982) *Annu. Rev. Genet.* **16,** 405–437.
2. Wu, A. M., Kahn, R., DasGupta, C. & Radding, C. M. (1982) *Cell* **30,** 37–44.
3. Kahn, R. & Radding, C. M. (1984) *J. Biol. Chem.* **259,** 7495–7503.
4. Gonda, D. K. & Radding, C. M. (1983) *Cell* **34,** 647–654.
5. Bianchi, M., DasGupta, C. & Radding, C. M. (1983) *Cell* **34,** 931–939.
6. Singer, B. S., Gold, L., Gauss, P. & Doherty, D. H. (1982) *Cell* **31,** 25–33.
7. Seed, B. (1983) *Nucleic Acids Res.* **11,** 2427–2445.
8. Bell, G. I., Pictet, R. L., Rutter, W. J., Cordell, B., Tischer, E. & Goodman, H. M. (1980) *Nature (London)* **284,** 26–32.
9. Sargent, T. D., Wu, J., Sala-Trepat, J. M., Wallace, R. B., Reyes, A. A. & Bonner, J. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 3256–3260.
10. Lomedico, P., Rosenthal, N., Efstratiadis, A., Gilbert, W., Kolodner, R. & Tizard, R. (1979) *Cell* **18,** 545–558.
11. Beaucage, S. L. & Caruthers, M. H. (1981) *Tetrahedron Lett.* **22,** 1859–1862.
12. Urdea, M. S., Merryweather, J. P., Mullenbach, G. T., Coit, D., Heberlein, U., Valenzuela, P. & Barr, P. J. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 7461–7465.
13. Hobart, P. M., Shen, L. P., Crawford, R., Pictet, R. L. & Rutter, W. J. (1980) *Science* **210,** 1360–1363.
14. Ish-Horowicz, D. & Burke, J. F. (1981) *Nucleic Acids Res.* **9,** 2989–2998.
15. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
16. Yamamoto, K. R., Alberts, B. M., Benzinger, R., Lawthorne, L. & Trieber, G. (1970) *Virology* **40,** 734–744.
17. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15,** 1157–1174.
18. Poustka, A., Rackwitz, H.-R., Frischauf, A. M., Hohn, B. & Lehrach, H. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 4129–4133.
19. Smith, M. (1983) in *Methods of RNA and DNA Sequencing*, ed. Weissman, S. M. (Praeger, New York), pp. 23–68.
20. Thomas, C. A. (1966) *Prog. Nucleic Acid Res. Mol. Biol.* **5,** 315–337.
21. von Hipple, P. H. (1979) in *Biological Regulation and Development*, ed. Goldberger, R. E. (Plenum, New York), Vol. 1, pp. 279–347.
22. Dressler, D. & Potter, H. (1982) *Annu. Rev. Biochem.* **51,** 727–761.
23. Clark, A. J. (1980) in *Mechanistic Studies of DNA Replication and Genetic Recombination*, ICN-UCLA Symposia on Molecular and Cellular Biology, ed. Alberts, B. (Academic, New York), Vol. 19, pp. 891–899.