Focused Genetic Recombination of Bacteriophage T4 Initiated by Double-Strand Breaks

Victor Shcherbakov,*,† Igor Granovsky,† Lidiya Plugina, Tamara Shcherbakova,*, Svetlana Sizova, Konstantin Pyatkov,† Michael Shlyapnikov† and Olga Shubina*

*Institute of Problems of Chemical Physics, Russian Academy of Sciences, Chernogolovka, Moscow Region, 142432 Russia and †Skryabin Institute of Biochemistry and Physiology of Microorganisms, RAS Pushchino, Moscow Region, 142292 Russia

Manuscript received June 21, 2002
Accepted for publication July 2, 2002

ABSTRACT

A model system for studying double-strand-break (DSB)-induced genetic recombination in vivo based on the etsI segΔ strain of bacteriophage T4 was developed. The etsI, a 66-bp DNA fragment of phage T2L containing the cleavage site for the T4 SegC site-specific endonuclease, was inserted into the proximal part of the T4 rIIB gene. Under segC+ conditions, the etsI behaves as a recombination hotspot. Crosses of the etsI against rII markers located to the left and to the right of etsI gave similar results, thus demonstrating the equal and symmetrical initiation of recombination by either part of the broken chromosome. Frequency/distance relationships were studied in a series of two- and three-factor crosses with other rII and rIIA mutants (all segC+) separated from etsI by 12–2100 bp. The observed relationships were readily interpretable in terms of the modified splice/patch coupling model. The advantages of this localized or focused recombination over that distributed along the chromosome, as a model for studying the recombination-replication pathway in T4 in vivo, are discussed.

DOUBLE-STRAND breaks (DSBs) in DNA and their repair are perceived now as a crossroads among a great variety of cell metabolic pathways, including maintenance of chromosomal stability, regulation of the cell cycle, immunoglobulin gene rearrangement, intron homing, recombination-dependent replication, repair of damaged replication forks, initiation of mitotic and meiotic recombination, and genetic recombination in prokaryotes (for reviews see Thaler and Stahl 1988; Shinohara and Ogawa 1995; Belfort and Roberts 1997; Osman and Subramani 1998; Haber 1998, 1999, 2000; Kuzminov 1999; Paques and Haber 1999; Flores-Rozas and Kolodner 2000; Kreuzer 2000; Cox 2001; Smith 2001; van Gent et al. 2001). T4 phage, a classical object of molecular genetics, presents one of the most suitable model systems to study interconnections of general or homologous genetic recombination with other fundamental genetic processes. In the T4 life cycle, DNA replication and recombination are integrated in one complex pathway; the ends of phage chromosomes initiate recombination, a step that is identical to the corresponding step in DSB repair; T4 has group I introns that can be transferred to intronless targets via a mechanism similar to DSB-promoted gene conversion (Belfort and Roberts 1997; Mosig 1998; Kreuzer 2000). The exceptionally well-developed biochemistry and genetics of T4 enables the most flexible and sophisticated experimental approaches to be used.

Recombination analysis based on crosses between T4 rII mutants is a powerful tool for studying genetic processes in vivo. Present-day availability of DNA sequences for recombining markers and knowledge of the functional properties of the numerous enzymes and other proteins involved in DNA metabolism, reproduction of many elementary genetic processes, and DNA transactions in vitro make interpretation of the recombination data much more explicit than was possible in the “golden” times of classic molecular genetics. However, our experience in recombination analysis tells us that a lot of work still remains to be done before recombination data can be unambiguously interpreted in terms of physicochemistry and molecular biology. The obstacles that hamper the progress are the real complexity of the phenomenon: intertwining of recombination with other genetic processes, the multiplicity of the interconnected recombination pathways (Mosig 1998), definitely non-uniform distribution of the recombination events along the chromosome on a fine scale, and influence of the markers themselves on this distribution. In this work, we developed a model system to study genetic recombination initiated by a DSB at a unique site within the rII region in a hope that this focused recombination will be devoid of at least some of the above obstacles.

What advantages are expected from this model system? First of all, we think that the focused recombin-
We inserted a DNA fragment from the intergenic 5–6 region of T2L, \( ets1 \), containing the recognition site for the T4 \( \text{segC} \) endonuclease, into the \( rII \) gene of T4 \( \text{seg} \). The crosses of the phage T4 \( \text{seg} \) \( ets1 \) against the T4 \( rII \) mutants bearing the \( \text{segC}^+ \) allele were used to study genetic recombination initiated by DSBs within the \( ets1 \) sequence, \( i.e. \), within the \( rII \) region.

**MATERIALS AND METHODS**

**Oligonucleotides**: Oligonucleotide primers \( \text{scs}-up \), \( 5'-\text{CCT CTCGTCGTAAGAAAAATCTC-3'} \) (complementary to T4 nucleotides 80,105–80,129), and \( \text{scs}-lo \), \( 5'-\text{CGGAAGACGGCTTAGTGATG-3'} \) (complementary to 80,868–80,890), were designed to anneal sequences flanking the \( \text{segC} \) gene of phage T4 (GenBank accession no. GI: 11079640). Oligonucleotide primers \( \text{rII}-up \), \( 5'-\text{CGCGTTAGTCGAGAAACGTTG-3'} \) (complementary to 622–647), and \( \text{rII}-lo \), \( 5'-\text{ATGCGTTCGTCG-3'} \) (complementary to 167,755–167,782), were designed to anneal to the \( rII \) region of T4. Oligonucleotide primers \( \text{scs}-up \), \( 5'-\text{GCGTACGCTTAGTGATG-3'} \) (complementary to 989–966), and \( \text{scs}-lo \), \( 5'-\text{CGGAATCTGAACTCAATCAATTCATGACAGCG-3'} \) (complementary to 80,868–80,890), were designed to anneal to the 25–56 intergenic region (GI: 18463948). Oligonucleotide primers \( \text{laczZb}-up \), \( 5'-\text{GAGATTGAAATGTTCTGCTGTCGC-3'} \) (complementary to 364,497–364,519), and \( \text{laczZb}-lo \), \( 5'-\text{CGAGGTACCGCCTGATCGACACG-3'} \) (complementary to 364,461–364,483), were designed to anneal to the \( \text{laczZ} \) ORF of Escherichia coli K-12 strain MG1655 (GI: 16127994). Nucleotides that are mismatched with the original sequences are italicized.

**Plasmids**: Plasmid pUSC\( \Delta \) (I. E. Granovsky, F. A. Kadyrov and V. M. Kryukov, unpublished data) was constructed by insertion of the T4 genome region 79,522–81,310 bp, which includes the \( \text{segC} \) gene (80,182–80,602 bp), into plasmid vector pUC18 (Vieira et al. 1982) and subsequently deleting 100–293 bp from the \( \text{segC} \) gene sequence.

**Bacteriophages**: The map, structure, and a source of \( rII \) strains used in this study are presented in Figure 1 and Table 1. The origin of most of the phages was described earlier (Shcherbakov et al. 1982, 1995; Shcherbakov and Plugina 1991). The mutants \( \text{amn116} \) and \( \text{amh17} \) (amber mutations in topoisomerase genes 39 and 52, respectively) were kindly supplied by F. Stahl.

**Bacteria**: Most E. coli strains were described earlier (Shcherbakov et al. 1982). E. coli B, on which \( rII \) mutants form large sharp-edged plaques (\( \tau \) phenotype), was used to discriminate \( rII \) and wild-type plaques of T4. E. coli BB and amber-suppressor E. coli CR\( 63 \) were used as hosts for preparing phage stocks, for phage titration, and for measuring total phage yields. E. coli CR\( 63 \) \( \lambda \) was used to derive \( \text{segC} \) mutants of T4. Genomic DNA of E. coli K-12 strain MG1655 (Heath et al. 1992) was used as a template for amplification of the \( \text{laczZ} \) ORF fragment by polymerase chain reaction (PCR).

**Media, bacterial culturing, and procedures for standard crosses**: These were essentially the same as described earlier (Shcherbakov et al. 1982).

**Standard cross procedure**: A sample of E. coli BB overnight culture was diluted 100-fold into Luria (L) broth and aerated at 37\( ^\circ \). At a cell concentration 1 \( \times 10^8 \) ml, the suspension was cooled to 0\( ^\circ \), pelleted by centrifugation, and resuspended in...
cooled L broth to a density $4 \times 10^8$ cells/ml. To 0.5 ml of the cooled BB suspension, a mixture of phage parents in a volume of 0.5 ml (the multiplicity of each parent being 5) was added. The infected cells were incubated for 10 min at 33°C, diluted 1000-fold into prewarmed L broth, incubated at the same temperature for another 80 min, and cell lysis was completed by adding 0.3 ml chloroform.

**Determination of recombinant frequency and plating efficiency:** Recombinant frequencies were calculated by dividing the titer determined on a λ-lysogenic host by the total lysate titer. The resulting frequencies were corrected for plating efficiency that was determined as follows. Phage strains with a genotype of the expected recombinant, e.g., wild type (an equal mixture of segC and segΔ variants), were used to infect the standard culture of E. coli BB with a total multiplicity of 5. The infected cells were incubated and processed in the same way as those in the standard crosses. The lysates in a proper dilution were plated on the E. coli strain that was used for determining the total progeny titer in crosses, e.g., BB, and on the relevant λ-lysogenic host, e.g., 594(λ). The ratio of the titer on BB to that on 594(λ) was then used as a plating efficiency quotient to correct the titers of the recombinants observed on the given λ-lysogen. This rather complex procedure to measure plating efficiency was used because of our occasional observations that the phages in freshly made lysates may differ in the plating efficiency from those in permanent phage stocks. The phages segC and segΔ may also differ somewhat in the plating efficiency despite the segC gene being nonessential for phage growth.

**Correction for chance coincidence of two independent recombination events:** For three-factor crosses amNI6 ets1 × i, where i designates a variable central marker, we corrected the measured frequencies of the wild-type (double-exchange) recombinants $R_3f$ for a chance coincidence of statistically inde-

### TABLE 1

<table>
<thead>
<tr>
<th>Mutation i</th>
<th>Position* (bp)</th>
<th>Change of sequence</th>
<th>Type of mutation</th>
<th>Sourcea</th>
</tr>
</thead>
<tbody>
<tr>
<td>a6</td>
<td>~200</td>
<td>Point</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N35</td>
<td>~350</td>
<td>Point</td>
<td></td>
<td></td>
</tr>
<tr>
<td>amH72</td>
<td>466</td>
<td>C to T</td>
<td>Amber</td>
<td>Y.V.</td>
</tr>
<tr>
<td>a2</td>
<td>1093</td>
<td>G to T</td>
<td>Opal</td>
<td></td>
</tr>
<tr>
<td>a3</td>
<td>1140</td>
<td>G to T</td>
<td>Ochre</td>
<td></td>
</tr>
<tr>
<td>N23</td>
<td>1546, 1545</td>
<td>Del C at 1546</td>
<td>fs−</td>
<td>N.M.</td>
</tr>
<tr>
<td>UV200</td>
<td>1871</td>
<td>T to A</td>
<td>Ochre</td>
<td>S.C.</td>
</tr>
<tr>
<td>amHB129</td>
<td>1951</td>
<td>C to T</td>
<td>Amber</td>
<td>B.S.</td>
</tr>
<tr>
<td>amHB84</td>
<td>2000</td>
<td>G to A</td>
<td>Amber</td>
<td>B.S.</td>
</tr>
<tr>
<td>106</td>
<td>2001</td>
<td>G to A</td>
<td>Opal</td>
<td>B.S.</td>
</tr>
<tr>
<td>C6</td>
<td>2048, 2043</td>
<td>del A</td>
<td>fs−</td>
<td>S.C.</td>
</tr>
<tr>
<td>N21</td>
<td>2143</td>
<td>C to T</td>
<td>Ochre</td>
<td>S.C.</td>
</tr>
<tr>
<td>490</td>
<td>2168</td>
<td>T to C</td>
<td>Base substitution</td>
<td>B.S.</td>
</tr>
<tr>
<td>X504</td>
<td>2195</td>
<td>G to A</td>
<td>Ochre</td>
<td>B.S.</td>
</tr>
<tr>
<td>UV375</td>
<td>2220</td>
<td>G to T</td>
<td>Ochre</td>
<td>J.D.</td>
</tr>
<tr>
<td>X511</td>
<td>2229</td>
<td>C to A</td>
<td>Ochre</td>
<td>S.C.</td>
</tr>
<tr>
<td>UV357</td>
<td>2248</td>
<td>C to T</td>
<td>Ochre</td>
<td>J.D.</td>
</tr>
<tr>
<td>375</td>
<td>2265</td>
<td>C to T</td>
<td>Ochre</td>
<td>S.C.</td>
</tr>
<tr>
<td>N24</td>
<td>2286</td>
<td>C to T</td>
<td>Ochre</td>
<td>B.S.</td>
</tr>
<tr>
<td>insZ</td>
<td>2298</td>
<td>Insertion of 66 bp</td>
<td>Insertion</td>
<td></td>
</tr>
<tr>
<td>ets1</td>
<td>2298</td>
<td>Insertion of 66 bp</td>
<td>Insertion</td>
<td></td>
</tr>
<tr>
<td>FC47</td>
<td>2317, 2318</td>
<td>TT to CTG</td>
<td>fs+</td>
<td>S.C.</td>
</tr>
<tr>
<td>oc31</td>
<td>2380</td>
<td>A to T</td>
<td>Ochre</td>
<td></td>
</tr>
<tr>
<td>N17</td>
<td>2559</td>
<td>C to T</td>
<td>Ochre</td>
<td>B.S.</td>
</tr>
<tr>
<td>N12</td>
<td>2751</td>
<td>C to T</td>
<td>Ochre</td>
<td>S.C.</td>
</tr>
<tr>
<td>b25</td>
<td>2910, 2907</td>
<td>del A</td>
<td>fs−</td>
<td></td>
</tr>
<tr>
<td>b26</td>
<td>~3060</td>
<td>Point</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Positions of the mutations are shown as distances (bp) from the beginning of the rIIA gene. The markers of insertion and deletion type may appear to occupy different apparent positions depending on the relative position (left or right) of another marker in a cross.

*The mutants rII of phage T4 were gifts from Yu. Vinetsky (Y.V.), N. Matvienko (N.M.), J. Drake (J.D.), S. Champe (S.C.), and B. Singer (B.S.); the rest are from our laboratory. For some mutants the sequence could be deduced from the genetic localization, origin, phenotype, and known sequence in corresponding wild-type sites (Pribnow et al. 1981); some were sequenced in Gold’s lab (Shinedling et al. 1987); amH72, a2, a3, N23, C6, oc31, N12, and b25 were sequenced during this work.
and so the phage with

\[ \Delta_{R} = \frac{2}{F} [2R_3 - (R_1 + R_2 - R_3) - R_1R_2], \]

where \( R_1, R_2, \) and \( R_3 \) are the measured wild-type recombinant frequencies in crosses \( \text{am} \text{N116} \times i, i \times \text{ets} \), and \( \text{am} \text{N116} \text{ets} \times i \), respectively; \( F \) is the correction factor (Lennon et al. 1953) equal to 0.9 for crosses with a multiplicity of infection 5:5; and \( \Delta_R \) is an allowance for \( R_3 \) value for chance coincidence of two statistically independent recombination events. \( R_2 \) and \( R_3 \) values are presented in Table 3 as \( R_3 \) and \( R_3 \), respectively; \( R_1 \) values were found not to differ significantly for 16 different \( i \) markers, and so the mean value 5.75% was used throughout.

Sequencing of \( \text{rlf} \) mutants: The relevant regions in \( \text{rlf} \) and \( \text{rlf} \) genes were amplified by PCR and the PCR products were purified using a QIAGEN (Valencia, CA) purification kit. The nucleotide sequences for \( \text{rlf} \) and \( \text{rlf} \) mutants were then determined by the dideoxy-sequencing method using a Cy-Reader DNA sequencing kit (MBI Fermentas). Determined by the dideoxy-sequencing method using a Cy-Reader DNA sequencing kit (MBI Fermentas). Other methods for molecular cloning and gene engineering were as described by Sambrook et al. (1989).

Constructing \( T4 \) segC mutants: \( E. coli \) [MJ109], bearing plasmid pUSC\( \Delta \), was infected with T4D, and the phage progeny was plated on an \( E. coli \) [MJ109] lawn. The plaques were tested with a \( ^3P \)-DNA probe for the deleted segment of segC, and those with negative response were selected. The presence of segC was further confirmed by PCR analysis, in which the oligonucleotide primers sCsc-up and sCsc-lo complementary to the sequences flanking segC were used.

The selected strain \( T4 \) segC was then used to construct a set of \( \text{rlf} \) strains carrying segC. For this, \( \text{rlf} \) mutants were crossed with \( T4 \) segC at a multiplicity of infection of 1 and 9 for \( \text{rlf} \) and segC, respectively. The lysates were plated on \( E. coli \) B, and selected plaques were backcrossed to the original \( \text{rlf} \) mutant and analyzed by PCR with the primers sCsc-up and sCsc-lo to identify segC. The mutants \( \text{am} \text{N116} \) (gene 39) and \( \text{H17} \) (gene 52), bearing segC, were obtained similarly by selecting the segC strains that did not grow on \( E. coli \) B at 25°C.

Constructing \( T4 \) strains bearing the cleavage site for SegC exonuclease: To clone the \( \text{rlf} \) region, a \( T4 \) DNA fragment 1768 bp long was amplified by PCR using the primers rII-up and rII-lo and inserted into the \( \Phi d\text{X174} \) sites of the plasmid vector pUC18 resulting in the plasmid pUrinB18.

A T2L DNA fragment 66 bp long from the intergenic \( 5-6 \) region [named endonuclease target sequence (ets)], containing the cleavage site for SegC exonuclease (I. Granovsky, F. Kadyrov and V. Kryukov, unpublished data), was amplified by PCR, using primers sccs-up and sccs-lo, and placed in the plasmid pUrinB18 at a HindII restriction site. This HindII restriction site is located 2298 bp from the beginning of the gene \( \text{rlfA} \), or 108 bp from the beginning of the gene \( \text{rlfB} \). The recombinant plasmid pUrinHts1, containing the ets sequence in one of two possible orientations, was selected and sequenced. The ets1 contains two in-frame termination codons, TGA and TAA; hence, a phage with this insertion was expected to have an \( \text{rlfB} \) mutant phenotype.

To construct phage strains with ets1, the plasmid pUrinHts1 was crossed with \( T4 \) segC and the progeny were plated on \( E. coli \) B. An \( r \)-type plaque was tested on \( E. coli \) B(\( \lambda \)) for \( \text{rlf} \) phenotype. The selected T4 strain was checked with PCR using the primers sccs-up and sccs-lo to confirm the presence of ets1. Analogously, the \( T4D \) insZ strain, containing an inert piece of \( E. coli \) gene lacZ 66 bp long in the same HindII site, was then crossed with \( E. coli \) gene lacZ. The ets1 sequence also contains in frame two termination codons, TGA and TAA, and so the phage with \( \text{insZ} \) has \( \text{rlfB} \) phenotype.

RESULTS

DSB-stimulated recombination depends on both the ets1 sequence and segC function: To check recombination frequency of the segC/ets1 system, the phage T4 \( \text{ets} \) and T4 \( \text{insZ} \) were crossed in different segC backgrounds against the double \( \text{rlf} \) mutant amH72 b25 and the single mutants amH72 and b25. The \( \text{ets}1 \) is the SegC exonuclease cut site inserted in the \( \text{rlfB} \) gene and \( \text{insZ} \) is a neutral insertion in the same position and of the same length as \( \text{ets}1 \) (see Figure 1). The lysates were plated on \( E. coli \) strains BB and 594(\( \lambda \)) to determine total progeny and wild-type recombinants, respectively. The results (Table 2) clearly show that in all the cases where DSBs were not expected to be produced because of the absence of either the SegC activity or its cleavage site, or both, the ordinary low recombinant frequencies were observed, the insertions \( \text{ets}1 \) and \( \text{insZ} \) giving similar results. (For an unknown reason, the crosses \( H72 \) b25 \( \times \) \( \text{insZ} \) and \( b25 \) \( \times \) \( \text{insZ} \) gave a twofold increase in recombinant frequencies in the absence of SegC exonuclease.) If, however, DSBs could be produced, the recombinant frequencies were drastically enhanced: >10-fold in the two-factor crosses and ~70-fold in the three-factor cross. The more pronounced DSB effect for the double-exchange recombinants may be attributed to the fact that, in the absence of SegC exonuclease, \( \text{ets}1 \) must behave as an ordinary 66-bp-long insertion. Being a central marker, such insertions display very low apparent recombinant frequencies (see, e.g., cross \( H72 \) b25 \( \times \) \( \text{insZ} \)) most probably originating from the chance coincidence of two single exchanges. Since \( \text{ets}1 \), under segC conditions, greatly enhances the frequencies of the single exchanges to the left and to the right of \( \text{ets}1 \) (see crosses \( H72 \times \text{ets}1 \) and \( \text{ets}1 \times b25 \)), a probability for their chance coincidence must enhance as a product. Indeed, after correction of the observed double-exchange frequencies for chance coincidence of two independent recombination events (Table 2), the DSB-dependent enhancing effect becomes ~500-fold. Besides, as is argued in the discussion, the high frequency of multiple-exchange recombinants is expected for the DSB-promoted process.

To check the efficiency of cleaving \( \text{ets}1 \), the wild-type \( T4 \) was crossed on \( E. coli \) BB with \( \text{ets}1 \) segC or \( \text{insZ} \) segC at a multiplicity of five particles of each parent per cell. The lysates were plated on \( E. coli \) B, on which \( r \)-type and wild-type plaques differ. It was found that \( r \) plaques account for <3% of the total progeny in the cross \( \text{ets}1 \times \text{ut} \) (2.52 ± 0.31% for four independent determinations). Several of these \( r \) plaques were picked up at random and proved genetically and functionally to be \( \text{ets}1 \). They gave no recombination with the original \( \text{ets}1 \) and showed high recombination with the other \( \text{rlf} \) mutants (H72 and b25). Thus, under conditions of standard cross, ~5% of the \( \text{ets}1 \) alleles survive. Since \( \text{ets}1 \) has an \( \text{rlf} \) phenotype, its incomplete cleavage does not hamper
TABLE 2

Frequencies of rII* recombinants in crosses of rII mutants in different segC backgrounds

<table>
<thead>
<tr>
<th>Cross</th>
<th>Frequency of recombinants (R ± SD)* (%)</th>
<th>Ratio (R + Δ) (%)</th>
<th>Ratio (R + Δ)/segC− (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>segC+</td>
<td>segC−</td>
<td>segC+/segC−</td>
</tr>
<tr>
<td>H72 b25 × etsl</td>
<td>14.1 ± 0.2</td>
<td>0.20 ± 0.04</td>
<td>70.5</td>
</tr>
<tr>
<td>H72 b25 × insZ</td>
<td>0.14 ± 0.01</td>
<td>0.24 ± 0.05</td>
<td>0.58</td>
</tr>
<tr>
<td>b25 × etsl</td>
<td>21.8 ± 1.8</td>
<td>1.60 ± 0.14</td>
<td>13.6</td>
</tr>
<tr>
<td>b25 × insZ</td>
<td>0.79 ± 0.07</td>
<td>1.61 ± 0.11</td>
<td>0.49</td>
</tr>
<tr>
<td>H72 × etsl</td>
<td>43.0 ± 3.3</td>
<td>3.90 ± 0.38</td>
<td>11.0</td>
</tr>
<tr>
<td>H72 × insZ</td>
<td>3.77 ± 0.22</td>
<td>4.02 ± 0.31</td>
<td>0.94</td>
</tr>
</tbody>
</table>

segC+ means that one of the parents bears the wild-type segC allele; segC− means that both parents are segCΔ; R + Δ is R corrected for chance coincidence of two independent recombinational events, as explained in Materials and Methods.

*aThe mean of four determinations ± standard deviation.

recombination analysis. In crosses insZ × wt, the particles with an r phenotype comprised about one-half of the total progeny (0.47 ± 0.06 for three independent determinations), proving that insZ is an inert insertion.

Distance dependence of DSB-initiated recombination: In Figure 2, the frequency/distance relationships obtained in three series of crosses, FC47 × i, ets1 × i, and amN116 ets1 × i, are presented (see Table 3). The markers i (rIIb and rIIa point mutations located to the left of ets1, all segC+) were the same in all three series; FC47 is a rIIb frameshift mutation located 19 bp to the right of ets1. The series FC47 × i (Figure 2A) illustrates the distance dependence of the ordinary recombination (in the absence of ets1) between rII mutants in the same region of the chromosome where we monitor DSB-promoted recombination. The two-phase relationship, expected for recombination via patches and splices (Stahl 1979; Toompuu and Shcherbakov 1980), is clearly seen on this plot, the first phase being a little distorted by a recombination anomaly at the intergenic divide, which is partly related to the contribution of the terminal heterozygotes to the apparent frequencies in intergenic (rIIb × rIIa) crosses due to complementation (Shcherbakov et al. 1994). The abscissa of the inflection, which supposedly corresponds to the median length of the patches (Stahl 1979; Toompuu and Shcherbakov 1980), is equal to ~700 bp.

In the series ets1 × i (Figure 2B), all the i markers were segC+, whereas ets1 was segCΔ. Four distinct phases of the frequency/distance relationship could be readily discerned. The first one, 0 to ~90 bp, demonstrates a very fast increase in the frequency as the distance increases. Then the frequencies level off at the value slightly <25% and virtually do not increase further until ~350 bp (the second phase, “intermediate plateau”). In the third phase, the frequencies increase again,
reaching the value \( \approx 45\% \) at the distance \( \approx 800 \) bp. A further increase of the distance up to 2100 bp leads only to a limited, if any, increase in the frequency (the fourth phase, “final plateau”).

The frequency/distance relationship for the double-exchange recombinants was studied in three-factor crosses \( amN116 \) \( ets1 \times i (R_{ef}) \) (Figure 2C). The side marker \( amN116 \) is an amber mutation in the gene 39 located \( >4000 \) bp to the left of \( ets1 \). In these crosses, wild-type recombinants can arise only via double exchanges. This enables one to measure separately only patch-related recombinants. The side marker \( amN116 \) is located far enough from the \( i \) markers not to occur frequently in the same heteroduplex with \( i \). The first three phases of the relationship are similar to those in two-factor crosses and they occupy the same positions: 0–90 bp, 90–350 bp (intermediate plateau), and 350–750 bp, respectively. But then a decline follows, the frequencies in crosses with the most distant central markers \( a3, a2, N35, \) and \( a6 \) falling to values lower than those obtained at a shorter distance. The differences between the frequency in the cross with \( N23 \) and those with \( N35 \) and \( a6 \) (4.22 \( \pm 1.00 \) and 4.13 \( \pm 1.02 \), respectively) are significant at \( P < 0.01 \) (by Student’s \( t \)-criterion).

**Leftward and rightward recombination:** To compare recombination to the left and to the right of \( ets1 \), we performed a symmetrical series of crosses \( amN116 \) \( ets1 \times i (R_{ef}) \) and \( i \) and \( ets1 \) \( amH17 \times i \) (Table 4), the markers \( i \) being mutations located to the left or to the right of \( ets1 \). The marker \( amH17 \) is an amber mutation in gene 52 located \( >4000 \) bp to the right of \( ets1 \). The couples of left and right \( i \) markers were chosen at approximately the same distances from \( ets1 \). Lysates from the crosses were plated on CR63 to measure total progeny titer. To measure recombinant titer, the lysates were exchanged recombinants was studied in three-factor crosses \( amN116 \) \( ets1 \times i (R_{ef}) \) and \( amH17 \) (Figure 3C, D). The couples of left and right \( i \) markers were chosen at approximately the same distances from \( ets1 \). Lysates from the crosses were plated on CR63 to measure total progeny titer. To measure recombinant titer, the lysates were plated on CR63(\( \lambda_0 \)) and on 594(\( \lambda \)). On CR63(\( \lambda_0 \)), amber-type markers \( amN116 \) and \( amH17 \) are suppressed, hence the crosses are two-factor. The left and the right series gave largely the same frequency/distance relationships both in two-factor (Figure 3A) and three-factor (Figure 3B) crosses. So, the left and the right half-molecules must have an equal ability to initiate recombination. Comparison of these series of crosses with those presented in Figure 2 and Table 3 shows also that the conditions when one of the parents is a topoisomerase mutant (gene 39, in which \( amN116 \) is located) do not influence recombination.

**DISCUSSION**

**Basic recombination model:** At least two types of pathway were suggested for DSB repair: double-strand-break repair (DSBR; Resnick 1976; Szostak et al. 1983; Sun et al. 1991) and synthesis-dependent strand annealing

### TABLE 3

<table>
<thead>
<tr>
<th>Distance (bp)</th>
<th>Frequency of recombinants (%)</th>
<th>( R_{ef} \pm SD )</th>
<th>( R_{ij} \pm SD )</th>
<th>( R_{ij} + \Delta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N24 )</td>
<td>31</td>
<td>0.15 ( \pm 0.005 )</td>
<td>2.24 ( \pm 0.18 )</td>
<td>0.53 ( \pm 0.02 )</td>
</tr>
<tr>
<td>( 375 )</td>
<td>52</td>
<td>0.21 ( \pm 0.004 )</td>
<td>3.24 ( \pm 0.22 )</td>
<td>0.92 ( \pm 0.14 )</td>
</tr>
<tr>
<td>( UV375 )</td>
<td>69</td>
<td>0.30 ( \pm 0.006 )</td>
<td>16.7 ( \pm 0.8 )</td>
<td>4.14 ( \pm 0.39 )</td>
</tr>
<tr>
<td>( X511 )</td>
<td>88</td>
<td>0.38 ( \pm 0.021 )</td>
<td>18.0 ( \pm 1.6 )</td>
<td>4.75 ( \pm 0.35 )</td>
</tr>
<tr>
<td>( UV375 )</td>
<td>97</td>
<td>0.37 ( \pm 0.029 )</td>
<td>19.3 ( \pm 0.6 )</td>
<td>5.09 ( \pm 0.45 )</td>
</tr>
<tr>
<td>( X504 )</td>
<td>122</td>
<td>0.47 ( \pm 0.042 )</td>
<td>22.4 ( \pm 0.8 )</td>
<td>6.22 ( \pm 0.63 )</td>
</tr>
<tr>
<td>( 490 )</td>
<td>141</td>
<td>1.08 ( \pm 0.06 )</td>
<td>23.8 ( \pm 2.4 )</td>
<td>6.15 ( \pm 0.58 )</td>
</tr>
<tr>
<td>( N21 )</td>
<td>174</td>
<td>1.31 ( \pm 0.06 )</td>
<td>24.5 ( \pm 1.0 )</td>
<td>6.15 ( \pm 0.43 )</td>
</tr>
<tr>
<td>( C6 )</td>
<td>274</td>
<td>1.53 ( \pm 0.07 )</td>
<td>25.0 ( \pm 1.6 )</td>
<td>6.20 ( \pm 0.60 )</td>
</tr>
<tr>
<td>( 160 )</td>
<td>316</td>
<td>1.45 ( \pm 0.08 )</td>
<td>24.1 ( \pm 1.5 )</td>
<td>12.6</td>
</tr>
<tr>
<td>( amN16 )</td>
<td>317</td>
<td>1.50 ( \pm 0.06 )</td>
<td>24.1 ( \pm 1.5 )</td>
<td>18.1</td>
</tr>
<tr>
<td>( amH17 )</td>
<td>446</td>
<td>1.77 ( \pm 0.05 )</td>
<td>27.7 ( \pm 2.3 )</td>
<td>7.52 ( \pm 0.65 )</td>
</tr>
<tr>
<td>( UV200 )</td>
<td>446</td>
<td>2.67 ( \pm 0.12 )</td>
<td>38.0 ( \pm 3.8 )</td>
<td>9.02 ( \pm 0.90 )</td>
</tr>
<tr>
<td>( N23 )</td>
<td>722</td>
<td>3.04 ( \pm 0.24 )</td>
<td>44.1 ( \pm 3.7 )</td>
<td>8.00 ( \pm 0.88 )</td>
</tr>
<tr>
<td>( a3 )</td>
<td>1168</td>
<td>3.01 ( \pm 0.29 )</td>
<td>44.9 ( \pm 2.6 )</td>
<td>7.66 ( \pm 0.74 )</td>
</tr>
<tr>
<td>( a2 )</td>
<td>1224</td>
<td>3.85 ( \pm 0.12 )</td>
<td>42.8 ( \pm 4.3 )</td>
<td>16.5</td>
</tr>
<tr>
<td>( amH17 )</td>
<td>1851</td>
<td>3.80 ( \pm 0.20 )</td>
<td>46.0 ( \pm 4.7 )</td>
<td>4.80 ( \pm 0.43 )</td>
</tr>
<tr>
<td>( N35 )</td>
<td>1967</td>
<td>3.92 ( \pm 0.09 )</td>
<td>46.5 ( \pm 2.4 )</td>
<td>4.89 ( \pm 0.49 )</td>
</tr>
</tbody>
</table>

* The mean \( \pm \) standard deviation; superscripts show the number of repeated crosses.

1 The mean of seven determinations \( \pm \) standard deviation.

2 The mean of four determinations \( \pm \) standard deviation.

\( \Delta \) corrected for chance coincidence of two independent recombination events, as explained in materials and methods.

**DISCUSSION**

**Basic recombination model:** At least two types of pathway were suggested for DSB repair: double-strand-break repair (DSBR; Resnick 1976; Szostak et al. 1983; Sun et al. 1991) and synthesis-dependent strand annealing...
TABLE 4
Recombinant frequencies in crosses amN116 ets1 × i and amH17 ets1 × i

<table>
<thead>
<tr>
<th>Marker i</th>
<th>Distance ets1-i (bp)</th>
<th>Position of i relative to ets1</th>
<th>Frequency of recombinants (%) (R ± SD)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>375</td>
<td>33</td>
<td>Left</td>
<td>3.54 ± 0.35 0.78 ± 0.09</td>
</tr>
<tr>
<td>FC47</td>
<td>19</td>
<td>Right</td>
<td>4.54 ± 0.72 0.85 ± 0.14</td>
</tr>
<tr>
<td>XS04</td>
<td>103</td>
<td>Left</td>
<td>23.0 ± 2.7 5.48 ± 0.62</td>
</tr>
<tr>
<td>oc31</td>
<td>82</td>
<td>Right</td>
<td>21.6 ± 2.7 7.05 ± 0.93</td>
</tr>
<tr>
<td>106</td>
<td>297</td>
<td>Left</td>
<td>25.5 ± 2.1 7.02 ± 0.82</td>
</tr>
<tr>
<td>N17</td>
<td>261</td>
<td>Right</td>
<td>25.4 ± 2.5 7.51 ± 0.88</td>
</tr>
<tr>
<td>UV200</td>
<td>427</td>
<td>Left</td>
<td>30.8 ± 2.4 8.06 ± 0.29</td>
</tr>
<tr>
<td>N12</td>
<td>453</td>
<td>Right</td>
<td>36.3 ± 0.40 8.66 ± 0.81</td>
</tr>
<tr>
<td>N23</td>
<td>753</td>
<td>Left</td>
<td>39.3 ± 3.1 10.3 ± 1.19</td>
</tr>
<tr>
<td>b26</td>
<td>750</td>
<td>Right</td>
<td>40.5 ± 2.8 10.9 ± 1.2</td>
</tr>
</tbody>
</table>

On CR63(λb) On 594(λ)

*The mean of three determinations ± standard deviation.

E. coli CR63(λb) suppresses amber mutations amN116 (gene 39) and amH17 (gene 52) and so rII+ recombinants form plaques on this strain regardless of their 39 or 52 genotype (i.e., a sum of single- and double-exchange recombinants is registered), whereas only wholly wild-type, double-exchange, recombinants (rII+ 39+/52+/rII- 39−/52−) grow on 594(λ).

(SDSA; for review see Paques and Haber 1999; Haber 2000). Both DSBR and most variants of SDSA postulate limited DNA synthesis and involvement of both DNA ends of the broken chromosome in a single repair event. This repair paradigm, while looking justifiable in the case of eukaryotic cells where the integrity of the chromosomes is crucial and their number is strictly regulated, looks unnecessarily constrained in the context of the T4 life cycle with its multiple chromosomes, the initiation of recombination by DNA ends, and the recombinational initiation of DNA replication. The concerted behavior of the two ends of broken DNA presumes some sophisticated (and evidently unnecessary) mechanism for keeping the DNA ends close to each other during the search for homologous sequences in the pool of the unbroken chromosomes. Indeed, it was shown by George and Kreuzer (1996) in their study of plasmid recombination in vivo during T4 infection that DSBs stimulate a normal, semiconservative DNA replication from each broken end, without coordinating the processing of the two ends. The model of extensive chromosome replication presented by George and Kreuzer is in fact a variant of the normal recombination-replication pathway supposedly operating during T4 infection (Mostig 1983, 1998).

We are also inclined to guide our analysis of DSB-initiated recombination in terms of the normal T4 recombination-replication pathway, more exactly in terms of a splice/patch coupling (SPC) model, which was presented by Shcherbakov et al. (1992) to explain the results of the single-burst analysis of recombinant progeny in T4 multifactor crosses. Double and single ex-
molecule 1, bearing the SegC cleavage site (ets1 molecule, thin lines), is cleaved in the middle of the ets1 insertion by T4 SegC endonuclease, producing ends with two-nucleotide 3’ tails (I. Granovsky, F. Kadyrov and V. Kryukov, unpublished data). It is supposed that 5’ strands are further resected to produce the extended recombinogenic 3’ single-stranded ends. Exonucleases involved might be gp46/47 or RNase H (George and Kreuzer 1996; Huang et al. 1999). The broken molecules 2 interact with the unbroken molecules bearing the i marker (i molecules, thick lines) via single-strand exchange to produce the structure 3 with a displaced loop and a protruding single-stranded “whisker.” In this special case, the 3’ terminus of the broken chromosome is heterologous and thus obligatorily nonpaired, although the whisker is expected in true homologous interactions as well. Special reasons for postulating the intermediate with protruding 3’ whisker were discussed earlier (Shcherbakov et al. 1992, 1995). This step, the search for homology, is suggested to be promoted by UvsX protein in collaboration with accessory UvsY protein and helix destabilizing protein gp32 (Griffith and Formosa 1985; Yonesaki and Minogawa 1985, 1989; Yonesaki et al. 1985; Formosa and Alberts 1986a,b; Hinton and Nossal 1986; Harris and Griffith 1987).

The subsequent single-stranded and double-stranded branch migration (step c) may be facilitated by the gene 41 protein, a DNA helicase bound to the displaced DNA strand by the gene 59 protein (Salinas and Kodadek 1995) and/or UvsW protein (Karles-Kinch et al. 1997).

Intermediate 3 is shared by many modern models of recombination. It can be processed by different pathways. We favor here the transition of structure 3 to 4 via double-stranded branch migration, resulting in formation of a true Holliday junction. This step converts the process to the major recombination-replication pathway (Shcherbakov et al. 1992, 1995). The protruding single-stranded 3’ end (whisker) is removed by 3’→5’ exonuclease of T4 DNA polymerase (43Exo), a step inevitably followed by initiation of DNA replication directed (as shown here) to the right. The loading of a whole replication machine is supposed. It is known that 43Exo, after removing the nonpaired strand, excises complementary nucleotides (Hersfield and Nossal 1972; Roth et al. 1982; Shcherbakov et al. 1995). This must lead to an irreversible loss of the corresponding sequence at the end of the broken DNA. The Holliday junction, molecule 5, is resolved by endonuclease VII in two equal ways (Müller et al. 1990), by cutting the “crossed” strands or by alternative cutting of “non-crossed” strands, producing different replication forks, 6 and 6’, respectively. This step may precede the whisker removal (step d). The replication (step f or f’) results in the splice-and-patch couples of recombinant molecules, 7 and 8 or 7’ and 8’.

The position corresponding to the site where the fragment ets1 has been inserted (ets+ allele) is marked...
with the vertical bar in all the structures. The position of the variable marker $i$ and its $i^*$ allele is not shown: It may happen at any place on the thick ($i$) or thin ($i^*$) line of the scheme, respectively. In three-factor crosses, the ets1 molecule bears an additional marker $j$. If marker $i$ occurs within the hybrid regions (HRs), the primary recombinant products will be heteroduplex heterozygotes (HHs). We do not distinguish parental and new synthesized strands to make the genetic consequences of the transformations more apparent.

For the crosses ets1 × $i$, four different distance-dependent outputs are expected as shown in Figure 4: (1) If the site $i$ is within the part removed by 43Exo [zone 1 (z1)] no recombinant will be produced; (2) if the site $i$ is within the part resected by a 5′ exonuclease but not removed by 43Exo (z2), one HH and one parental progeny chromosome ($i$ chromosome) will be produced; (3) if the site $i$ happens within the region of double-stranded branch migration (z3), two HHs will be produced; and (4) if the site $i$ is more distal than the distal border of the HR (z4), both strands of the splice-type molecule 7 or 8′ will be recombinant (thin lines), while the patch-type product 7′ or 8 will be parental ($i$ type, thick lines). In the three-factor cross of a type $j$ ets1 × $i$ with the central marker $i$, the wild-type recombinants can arise only if $i$ is within a patch, provided the HR does not extend to the side marker $j$.

According to the model in Figure 4, two populations of the HRs are expected: Longer HRs result from the combined contribution of the single-stranded branch migration and the double-stranded branch migration, whereas shorter HRs are produced only by the double-stranded branch migration. One long HR and one short HR arise simultaneously in each individual recombination event.

To make use of the model presented in Figure 4 for calculating the expected recombinant frequencies in the crosses, one must take into account some statistical complications. The primary patches and splices shown in Figure 4 may be replicated once or more to produce DNA molecules without HRs. Since only one-half of the vertical bars mark the position corresponding to the left and to the right of the DSB. Since segC is a late T4 gene (Sharma et al. 1992), dozens or probably hundreds of the phage chromosomes must be present in each infected cell at the moment of initiation of DSBs by SegC endonuclease, one-half of which contain ets1 and eventually get broken. We assume that the left replication forks (Figure 4, molecules 6 and 6′) have equal probability to arise earlier or later than the right ones. In Figure 5A, the left replication fork has replicated the right splices and patches, whereas the opposite is shown in Figure 5B. Note that each replication fork replicates only one of the two primary products, either patch or splice, but not both; therefore the other, splice or patch, remains unreplicated. Thus, each individual coupled event produces three progeny chromosomes. These events can arise from 16 possible combinations of four alternatives: two broken ends (left and right) to initiate the event, two possible ways of Holliday junction resolution, two templates for replication (splice or patch), and two variants of the timing (either left replication fork replicates right patch or splice or right replication fork replicates left patch or splice). This gives 16 possible events and therefore 48 possible types of progeny chromosome. Although only 32 variants have unique genetic structures, all 48 products should be considered as statistically independent variants arising with equal probability, since we have no reason to confine the process to only certain outputs.

In Figure 5C, the recombination consequences of collisions of the left and right replication forks, leading to only limited (repair) DNA synthesis, are presented (all eight possible products are shown). This special case imitates the concerted action of the two parts of the broken DNA with limited DNA replication, resembling in this respect other DSBR models, e.g., the model of Szostak et al. (1983) as modified by Sun et al. (1991), except our model presumes formation of full replication forks with synthesis of both leading and lagging strands of DNA, a situation that is probably realistic also for DSB repair in yeast (Holmes and Haber 1999). The vertical bars mark the position corresponding to the site of ets1 insertion ($ets^+$).

The probability for the marker $i$ to occur within the zones 1, 2, 3, or 4 depends on the distance $i$ets1. So, the schemes in Figure 5 enable one to deduce quantitative expectations for the recombinant frequencies depending on the distance of marker $i$ from ets1. Ignoring inevitable statistical nonrandomness in formation of the different recombinant products, taking the probability for the fork collision (Figure 5C) to be low, and assuming the 48 recombinant products to arise with equal probability, the expected recombinant frequencies in two-factor crosses are zero within zone 1, 21% within zone 2, and 42% within zones 3 and 4. These values can be calculated from Figure 5 as follows. There are three heteroduplex and one wholly recombinant (thin lines) regions within zone 2 among the 12 chromosomes
Figure 5.—Final products of the coupled recombination-replication processes initiated to the left and to the right of a DSB, according to the SPC model (Figure 4). (A) Recombination-replication process to the right of the DSB (a mirror image of that shown in Figure 4) was completed earlier than the corresponding process to the left of the DSB. Subsequently, the left replication fork (structure 6 in Figure 4) has replicated either the right patch (steps $a \rightarrow b$) or the right splice (steps $a' \rightarrow b'$), resulting, respectively, either in the pair of daughter molecules and one unreplicated molecule with the splice or in another pair of daughter molecules and one unreplicated molecule with the patch, the left arms of the chromosomes retaining their primary unreplicated structures. (B) The process to the left of the DSB (shown in Figure 4) was completed earlier than the corresponding process to the right of the DSB. Subsequently, the right replication fork (mirror image of structure 6 in Figure 4) has replicated either the left patch (steps $a \rightarrow b$) or the left splice (steps $a' \rightarrow b'$), resulting, respectively, either in the pair of daughter molecules and one unreplicated molecule with the splice or in another pair of daughter molecules and one unreplicated molecule with the patch, the right arms of the chromosomes retaining their primary unreplicated structures. (C) Collisions of left and right replication forks: Two left and two right variants of the forks (depending on the direction of Holliday junction resolution) give four possible combinations, resulting in eight possible recombinant products. Thick and thin lines designate DNA strands of the $i$ parent and $i'$ parent, respectively. The vertical bars mark the position of the ets1 insertion ($ets^+$ allele). The vertical dotted lines separate different zones in the final DNA products. To calculate the expected recombinant frequency in cross ets1 $i$ $\times \ i$, take into account either the left or the right arm of the molecules. The thick and thin lines should be regarded as parental and recombinant strands, respectively (allele ets1 is virtually always lost during the process, so the strands of progeny chromosomes drawn here by thin lines present $rII^+$ alleles); in three-factor crosses with $i$ as a central marker, only patches and their homoduplex derivatives contribute to the double-exchange recombinant frequency. All the homoduplex and one-half of the heteroduplex recombinants produce plaques on the $\lambda$-lysogenic host (Hertel 1965; Doermann and Parma 1967; Stahl 1979).
presented in Figure 5, A and B. Since only one of two HHs produces a plaque on a λ-lysogenic host (Hertel 1965; Doerrmann and Parma 1967; Stahl 1979), we expect the recombinant frequency to be 2.5/12 = 0.21. There are six heteroduplex and two wholly recombinant (thin lines) regions within zone 3 among the 12 chromosomes and so the expected recombinant frequency is 5/12 = 0.42. There are five wholly recombinant regions among the 12 chromosomes within zone 4 and so we expect a recombinant frequency of 5/12 = 0.42. Calculations involving all 48 possible progeny chromosomes would give the same expectations.

In three-factor crosses only patches produce wild-type recombinants, and so the expected frequencies are 10.5 and 21% within zones 2 and 3, respectively, while no double-exchange recombinants are expected if the middle marker is within zone 4 as a result of a single recombinant event. If, however, the concerted action of the two parts of the broken DNA molecule (Figure 5C) is frequent, the expected values are somewhat higher: For two-factor crosses ets1 × i, we would expect 25% within zone 2 and 50% within zones 3 and 4.

Before embarking on the quantitative comparison of the experimental data to the model predictions, let us make one general note. The schemes in Figures 4 and 5 were drawn assuming that all the events initiated by DSBs do not differ from each other by such physical parameters as the size of the resected and removed tracts and the length of HRs. In reality, these parameters are most probably stochastically distributed, and zones 1–4 likely differ in individual events and partially overlap.

Phase one: The first phase in Figure 2B (0–90 bp) demonstrates a very abrupt increase of the recombinant frequency with distance. The interpretation in terms of the model in Figures 4 and 5 looks rather easy: The closer the i⁺ allele is to the tip of the single-stranded end (i.e., to the position of the DSB), the more probable is its removal by 43Exo (i.e., its occurrence within zone 1) and the less is the probability for recombinant formation. Consequently, the first phase in the frequency/distance relationship (Figure 2B) reflects, according to our model, the probability for the i⁺ allele to be removed by 43Exo. Note that the recombinant frequencies at the shortest distances (the crosses with N24 and 375, Table 3) are especially small, and a huge jump of the frequency both in two- and three-factor crosses is observed when the distance increases only from 33 to 50 bp, a reminiscence (and a possible explanation) of the well-known 50-bp limit in the minimal homology required for recombination (see, e.g., Singer et al. 1982). Shcherbakov et al. (1995) concluded that 43Exo, after removing the nonpaired 3’ end (whisker), excises on the average ~25 complementary nucleotides (see also Hershfield and Nossal 1972; Roth et al. 1982). Since, in the crosses with X504 (103 bp from ets1), the recombinant frequency reaches already the first plateau value, we infer that 43Exo rarely removes >100 complementary nucleotides from the 3’ end of the invading DNA strand. The involvement of 43Exo in intron homing, demonstrated by Huang et al. (1999), is in general congruence with the postulated role for this enzyme in a DSB-stimulated process.

Intermediate plateau: Within the range of distances of 90–350 bp, recombinant frequencies are virtually independent of the distance giving the intermediate plateau in the two-factor crosses ets1 × i at the value slightly <25%. This frequency is close to the 21% predicted for zone 2. Zone 2 in our model corresponds to the region of single-stranded branch migration, resulting in only one recombinant strand of four progeny DNA strands. If so, the median length of the 5’ strand, resected during step a (Figure 4), appears to be ~350 nucleotides. The transitional phase that follows after the intermediate plateau may reflect a length distribution of the resected tract.

Final plateau: The final plateau in the recombinant frequencies (~45%), observed in two-factor crosses ets1 × i (Figure 2B), presumably corresponds to zones 3 and 4. Transition from the single-stranded branch migration to the double-stranded branch migration, as shown in Figure 4, explains well this phase in the frequency/distance relationship. This transition determines the formation of the four-stranded DNA complex, resolution of which gives two recombinant and two parental DNA strands in the primary recombinant products both within the HR zone 3 and out-of-HR zone 4 (instead of one recombinant strand and three parental overlap. strands within zone 2). The observed frequencies within the final plateau do not differ significantly from the expected 42%.

Double-exchange recombinants: One of the predictions of the SPC model is the equal production of patches and splices. Correspondingly, we expect one-half as many recombinants in three-factor crosses with the central marker i as in two-factor crosses within zones 2 and 3 (Figures 4 and 5). However, while running in parallel with the frequencies in two-factor crosses (Table 3 and Figure 2C), the frequencies in three-factor crosses are significantly lower than those predicted by the model: The mean frequency within the intermediate plateau is 6.2% instead of the expected 10.5%. One reason for this underestimation is purely statistical. We deal here with very high frequencies of recombinants, including very high frequencies of the double-exchange recombinants that arise from the patches. This implies a substantial contribution of multiple independent events to the observed recombinant frequencies. Indeed, the correction of the frequencies for the chance coincidence of two independent events (Table 3) gives frequencies that are in better congruence with the model predictions. George and Kreuzer (1996), using a physical assay and a plasmid substrate, also observed
that DSB repair in bacteriophage T4 involves conversions and exchanges of flanking DNA ~50% of the time.

As is seen in Figure 2C, the frequency of the double-exchange recombinants reaches its maximum value at a distance of ~750 bp, the crosses at longer distances giving progressively lower frequencies. This pattern is quite conceivable if the double-exchange recombinants originate primarily from patch-type HRs and if the median length of the patches is ~750 bp. The high apparent recombinant frequencies, ~5%, in the crosses with N35 and a6, separated from ets1 by ~2000 bp, may mean that the HRs of such length are produced with a low but appreciable frequency. The contribution of the ordinary (non-ets1-SegC-dependent) general recombination may account for only a small part of this value, not >1% (T. S. SHCHERBAKOVA, unpublished data).

Relation to general recombination: A good quantitative agreement between the experimental data and the predictions of the SPC model was observed here. The measured frequencies in two-factor crosses would be in even better congruence with the predicted values (21 and 42% for the intermediate and the final plateau, respectively) if the contribution of general recombination (as determined, e.g., in the crosses FC47 × i, Table 3) is taken into account. Other events, such as collisions of replication forks (Figure 5C), may also positively contribute to the observed recombinant frequencies.

All the phenomenology observed here is satisfactorily explained in terms of the model, which earlier was suggested and substantiated for general genetic recombination in T4 bacteriophage (Shcherbakov et al. 1992). Specific recombination parameters, such as the length of HRs in recombination intermediates, are well within the range of the estimates in general recombination (for review, see Stahl 1979; Most C. 1994). The data in this article (Figure 2A) may also be used for an estimation of HR length in ordinary recombination. In terms of the models of recombination via HRs (Stahl 1979; Toompuu and Shcherbakov 1980), the abscissa of the inflection between phases 1 and 2 (~700 bp) corresponds to the median length of the patches. This value is close to what we inferred from the three-factor crosses amN116 ets1 × i for recombination promoted by a DSB (Figure 2C). It is tempting to conclude that the recombination pathway operating in the DSB-initiated process is the same and probably the major pathway that operates in general genetic recombination. As we hoped from the very beginning, the recombination data obtained with this model system turned out to be much easier for interpretation in molecular terms than those related to recombination with stochastically distributed events.

About gene conversion: The SPC model is related to the pathways that are known presently as break-induced replication (BIR). BIR is supposed to account for the very long tracts (up to the end of the chromosome arm) of mitotic gene conversions in yeast, whereas a common view of meiotic gene conversion is that it involves short conversion tracts, on average 1–2 kb (for review, see Paques and Haber 1999; Smith 2001). A variant of the SPC model, which we regard here as most plausible for T4, assumes nonconcerted action of the two ends of a broken chromosome, each end priming unilateral replication of the unbroken chromosome (i.e., BIR), although the processes initiated to the left and to the right of a DSB are statistically coupled. It is easy, however, to imagine that the stronger coupling resulting in obligatory collision of the left and the right replication forks may operate in the systems in which uncontrolled reduplication of a chromosome is forbidden and only limited (repair) DNA synthesis is allowed. In fact, formation of two Holliday junctions on both flanks of a DSB before initiation of DNA synthesis would be enough to prevent chromosome reduplication. All possible recombinant products predicted by the SPC model with the collision of replication forks are shown in Figure 5C. Zone 4 corresponds here to the regions flanking zones 1–3, where the conversion events take place. They have crossover (recombinant) or noncrossover (parental) configurations. The SPC pathway leads to formation of a splice/patch pair as a result of a single recombination event regardless of the manner of Holliday junction resolution. Thus, we do not need to postulate isomerization of Holliday junctions to explain the occurrence of splices and patches in recombination intermediates. It is especially important for eukaryotic chromosomes where the isomerization may be hampered by the axial or lateral elements of the synaptonemal complex. It is important to distinguish the patches and splices arising in a one-sided (left or right) process (Figure 4) and noncrossover and crossover flanking configuration of the conversion region resulting from joint (left and right) DSB processing (Figure 5). The latter depends on the way the Holliday junctions are resolved. The configuration will be crossover or noncrossover if left and right Holliday junctions are resolved in the same or in opposite sense, respectively.

It is tempting to speculate that BIR-like pathways and those resulting in the short-tract gene conversions are variants of a common mechanism, differing, e.g., in the ability to form two Holliday junctions before starting DNA synthesis. A meaningful example of such a relationship may be the data by Malkova et al. (1996), who used the MATa/MATa system to study the repair of HO endonuclease-induced DSB in Saccharomyces cerevisiae. They observed short-tract gene conversions in wild-type (RAD51 RAD52) strains, but long-tract gene conversions, named BIR, in rad51Δ diploids. Relevantly, the pattern of recombination observed in crosses ets1 × i in the absence of UvsX protein (which, like Rad51 protein, is a T4 analog of RecA protein) agrees with the supposition that in a uvsX background the DSB initiated process occurs without transition to the double-stranded branch migration stage (V. P. Shcherbakov, T. S.
From mechanistic and enzymatic points of view, the SPC model seems to be quite feasible for recombinational DSB repair in eukaryotes. All the enzymes that supposedly act in this pathway, including polymerase-associated proofreading exonucleases, are known in eukaryotes (Paques and Haber 1999; Shinohara and Ogawa 1999).

The authors thank Tanya Kolycheva for technical assistance in experiments. This work was supported by Russian Foundation for Basic Research grant 01-04-48794.

LITERATURE CITED


Harris, F. D., and J. Griffith, 1987 Visualization of the homologous pairing of DNA catalyzed by the bacteriophage T4 UvsX protein. J. Biol. Chem. 262: 9285–9292.


Shinohara, A., and T. Ogawa, 1995 Homologous recombination


Communicating editor: G. R. Smith