Exclusion of long heterologous insertions and deletions from the pairing synapsis in pneumococcal transformation

Franck Pasta and Michel A. Sicard

We have studied the mode of recombination of six insertions during genetic transformation of Streptococcus pneumoniae. The six heterologous insertions are located at the same site in the ami locus of the pneumococcal chromosome; insertion sizes range from 4 to 1374 bp. With respect to single-point markers we found that the number of transformants in one-point crosses is reduced, while the number of wild-type transformants in two-point crosses is drastically increased, what we call hyper-recombination. The magnitude of the shift is correlated with the size of the insert. This effect could result either from a special repair pathway of multibase heteroduplexes or from the exclusion of multibase heterologous insertions out of the pairing synapsis. To test these hypotheses we have used insertions in two kinds of three-point crosses. The repair model predicts that the excess of wild-type transformants remains in one set of crosses but is suppressed in the second set. The results we obtained are reversed, ruling out the hypothesis of a repair process, but in agreement with predictions based on the exclusion model. Moreover, we have re-examined the situation of deletions, our previous results suggesting that deletions were likely to be converted at the heteroduplex step. Genetic evidence we obtained in this work no longer supports this hypothesis. Thus, long heterologous insertions are partly excluded at the pairing step.

**Keywords:** Streptococcus pneumoniae, recombination, insertions, deletions, heteroduplexes

**INTRODUCTION**

During pneumococcal transformation, one DNA strand enters the cell and recombines with the chromosome thereby inducing the formation of a donor/recipient heteroduplex joint (for a review, see Lacks, 1988). Once replicated, such heterozygotes generate one transformant and one recipient genotype. However, mismatches may undergo repair prior to replication. The origin of this hypothesis was the discovery that several mutations are poorly transformable in wild-type Streptococcus pneumoniae strains (Ephrussi-Taylor & Gray, 1966). Subsequently, mutants were found, called hex, which integrate all markers with the same high efficiency (Lacks, 1970; Tiraby & Sicard, 1973a). Further studies showed that donor/recipient mismatches may trigger a complete excision of the donor DNA from the heteroduplex structure, followed by a resynthesis of the recipient strand (Méjean & Claverys, 1984). This mismatch repair system, which is absent in hex mutants, has been called Hex. The mutator phenotype of hex strains strongly suggests that it also acts on mismatches spontaneously arising in DNA (Tiraby & Sicard, 1973b; Tiraby & Fox, 1973). Heteroduplexes A/G and C/T corresponding to transition mutations are efficiently corrected by Hex. Those resulting from transversions are partly repaired, except C/C and A/G (for a review, see Claverys & Lacks, 1986). Mispairs corresponding to small heterologous insertions, i.e from 1 to 4 bases, are also processed by Hex (Gasc et al., 1987). Due to the multiplicity of mispairs processed, Hex is considered as a generalized mismatch repair system. Heteroduplexes corresponding to heterologous insertions longer than 4 bases escape the action of Hex (Gasc et al., 1987). Moreover several studies showed that long deletions do not behave as point mutations during transformation. Ghei & Lacks (1967) found that the kinetics of integration were much slower for deletions than for point mutations. Also, the transformation efficiency of deletions is reduced (Lacks, 1966; Lefèvre et al., 1989). Moreover,
in two-point crosses, when donor deletions are used to transform a strain containing a linked point mutation, wild-type recombinants occur in very large excess relative to the physical distance separating the two mutations (Lefèvre et al., 1989). This hyper-recombination could have two origins. First, the heteroduplex forms a multibase loop which might be corrected to the wild-type. Second, recombination might frequently stop at the region of heterology, excluding it from the pairing synopsis. Genetic evidence that we previously obtained for long deletions argues for the first possibility (Lefèvre et al., 1989).

In order to define more precisely how long heterologous insertions recombine, we have carried out a study using a set of chromosomal insertions of different sizes, located at the same site on the pneumococcal chromosome. We must point out that the formation of donor/recipient heteroduplexes generated by either donor deletions or by donor insertions should differ in more aspects than just the symmetry of donor/recipient heteroduplexes. During DNA entry, natural breaks will occur within an insertion, generating a loss of flanking homology, while deletions escape entry nicks just as point mutations. In addition, pairing of donor fragments containing insertions certainly requires the single-stranded heterologous insertion for folding, while recombination of deleted fragments should involve the bypass of the double-stranded chromosomal region. Both processes are likely to be different, and it was not directly possible to extrapolate previous results obtained with deletions to the behaviour of insertions in genetic crosses. This work was therefore carried out in order to clarify this.

**METHODS**

**Strains, media and transformation procedures.** Streptococcus pneumoniae strains used and constructed in this study derive from strain R36A of Avery (Avery et al., 1944). They have a R801 genetic background (Lefèvre et al., 1979) i.e. wild-type except for a mutation in the hexB gene. Thus they transform all point mutations with the same high efficiency (Tiraby & Sicard, 1973a). The str-4I marker, conferring resistance to streptomycin, was present in strain R119 (Tiraby & Fox, 1973). polA mutations were characterized in S. pneumoniae (Martinez et al., 1986) and transferred into the R801 background as described by Pasta & Sicard (1994). Culture media, preparation of chromosomal DNA and pneumococcal transformation methods were as described by Claverys et al. (1980). Routinely for a transformation involving 1 ml diluted competent cells, we used about 0.2 µg chromosome and 0.1 µg plasmid, i.e. saturating amounts of DNA. Plasmid- and M13-derived recombinant vectors were amplified in the Escherichia coli strain JM101 [A lac proAB trpF' trpD6 pro AB lac' F' lacZAM15] (Messing et al., 1981). Double-stranded and single-stranded DNAs were prepared as described by Maniatis et al. (1982).

**Pneumococcal crosses involving the ami locus.** The ami locus is an operon involved in oligopeptide transport (Alloing et al., 1990). ami strains are resistant to 0.02 mM methotrexate, but their growth is inhibited on synthetic medium containing excess isoleucine (Sicard, 1964). Conversely, ami- bacteria are sensitive to methotrexate, but they grow on synthetic medium containing excess isoleucine. One-point crosses can therefore be carried out in both directions: transformation of an ami+ recipient with ami DNA or vice versa. The transformation efficiency of the donor site is then measured as the number of ami os ami+ transformant colonies divided by the number of transformants for the reference point mutation str-4I, also carried on the donor DNA and conferring resistance to streptomycin. The possible screening of ami+ bacteria allows one to perform two-point crosses. In such crosses, a strain mutated within the ami locus is transformed with another ami+ mutation, together with the reference marker str-4I. ami+ bacteria, generated by recombination between the two sites involved in the cross, are selected on synthetic medium. Their number is divided by the number of Str+ transformants and multiplied by 100 to give the recombination index. In hex backgrounds, the recombination index obtained with point ami+ mutations depends mainly on the distance separating the two sites. On average, 27 bp lead to a recombination index of 1% (Claverys et al., 1979).

**Insertion of heterologous insertions at the BamHI site of the ami locus.** The EcoRI A fragment of the ami locus is approximately 1800 bp in size, with a BamHI site at approximately 1100 bp from the EcoRI right end (Fig. 1). The recombinant plasmid pR6 carries almost all of the wild-type ami locus on a 5800 bp insert (Méjean et al., 1981). We have subcloned the wild-type allele of the A fragment from pR6 into the EcoRI site of a M13mp11 replicative form (RF), from which we previously deleted the Smal–HindII fragment, containing the only BamHI site in M13mp11 (Messing & Vieira, 1982). The resulting recombinant vector, RV15, carries the wild-type A fragment with the unique BamHI site. RV15 was linearized with BamHI to insert heterologous insertions. A small insertion of 4 bp was introduced by filling in the recessed BamHI ends with "Klenow fragment" (Maniatis et al., 1982) and ligating with T4 DNA ligase. Among clones resistant to BamHI digestion, two were sequenced (Sanger et al., 1977); both carry the 4 bp insertion at the BamHI site. Longer insertions correspond to Sau3A fragments of the pBR322 plasmid (Rodriguez et al., 1977). A Sau3A digest of pBR322 generates 22 fragments from 8 to 1374 bp. The two longer ones, 1374 and 665 bp, well separated on an agarose gel, were directly purified and ligated to RV15, linearized with BamHI and 5'-dephosphorylated (Maniatis et al., 1982). Smaller Sau3A fragments from pBR322 are located within a smear. To define more precisely the Sau3A sequences to be cloned, we first generated prefragments that separate well on a gel by hydrolysing pBR322 with enzymes that cut rarely. Two such prefragments were purified. One is Avel–HindII, of 368 bp. It contains three Sau3A sites generating four fragments, including 18 bp and 46 bp Sau3A sequences that can be ligated to BamHI termini. The second prefragment is BamHI–NruI, of 598 bp. When hydrolysed with Sau3A it generates 92 and 359 bp fragments that can be cloned into BamHI. Sau3A digests of prefragments were phenol-extracted, ethanol-precipitated and ligated to RV15, hydrolysed with BamHI and dephosphorylated. All ligation mixtures were used to transform separately the E. coli strain JM101. RF and viral DNAs were prepared from individual plaques. Recombinant RV15-pBR322 vectors were checked by sequencing and comparing the inserted fragment to the expected sequence of pBR322 (from the EMBL database). We have isolated clones containing the expected sequences of 1374, 665, 359 and 52 bp. Three clones, generated by ligating the Sau3A digest of the Avel–HindII prefragment into RV15, were sequenced. None carried the 18 bp or the 46 bp sequence, but they carried the complete 64 bp sequence instead. We kept six recombinant clones, each carrying a defined insertion in the BamHI site of the ami locus (of 4, 64, 92, 359, 665 and 1374 bp).
Isolation of pneumococcal strains carrying the ami insertions. RF DNA corresponding to a recombinant clone for a given size of insert was used to transform strain R801. Methotrexate-resistant transformants were selected on plates. Such transformants correspond mainly to the desired allelic replacement of the recipient ami locus by the donor ami-inserted DNA. However, 'insertion-duplication' events integrating the full vector into the chromosome (Vasseghi et al., 1981) also lead to methotrexate resistance, by disrupting the ami locus. The first class of mutations almost never reverts to ami, while the second class should easily revert by homologous recombination excising the plasmid. For each size of insertion, three ami transformant colonies were picked, subcultured three times in rich medium and plated on synthetic medium to score for ami revertants. Out of 18 ami transformants, 2 gave ami colonies. For each of the six ami-insertion constructs, one nonreverting mutant was kept for further analysis. One set of similar mutants carrying the streptomycin-resistance gene was also constructed by transformation.

Construction of double ami mutant strains. Five double ami mutants have been constructed for the purpose of this study: ami-42 ami-6; ami-ins-1374 ami-6; ami-ins-665 ami-6; ami-ins-359 ami-6; ami-42 ami-109 (see Fig. 1 for mutations). Double mutants were constructed by transformation of a defined mutant with a second mutation. To increase transformation frequency at a genetic site, we used cloned ami mutations as donor DNAs. The point mutant ami-42 and insertion mutants ami-ins-1374, ami-ins-665 and ami-ins-359 were transformed with the recombinant plasmid pR32, which carries the ami-6 point mutation on the recipient chromosome. The recombination index depends on those with donor insertions, it appears that deletions allow the measurement of the transformability of a set of donor deletions of increasing size, flanked by the same donor sequences. We still observe a decrease in the efficiency of transformation, inversely related to the size.

RESULTS

Transformation efficiencies of insertions

Chromosomal DNA from the six strains containing a heterologous insertion within the ami locus was used to transform R801, which is a hex strain. This study is therefore free of any interference from Hex, the generalized mismatch repair system of S. pneumoniae.

In agreement with previous findings (Gasc et al., 1987), the smallest insertion of 4 bp transforms as efficiently as a point mutation, since its transformation efficiency is approximately 1 (Table 1). With respect to this reference value, we observe that longer insertions have a reduced transformation efficiency. This is particularly noteworthy for the longest insertion, which displays an efficiency of 0.55 (Table 1). Insertions therefore exhibit a decreased efficiency of transformation, inversely related to their size.

To understand more completely the efficiencies of transformation, we performed the reciprocal crosses. The six strains containing an insertion were used as recipients. Donor DNA was ami and Str'. We must point out that this situation, i.e. wild-type locus on the donor and insertion on the recipient, is physically identical to crosses with deletions on the donor. Thus, these reciprocal crosses allow the measurement of the transformability of a set of donor deletions of increasing size, flanked by the same DNA sequences. We still observe a decrease in the efficiency, ranging from 1 for the smallest insertions down to 0.36 for the longest ones. Comparing these efficiencies to those with donor insertions, it appears that deletions transform slightly less efficiently than donor insertions.

Recombination of insertions in two-point crosses

In the two-point crosses between linked ami mutations, one was located on the donor DNA, the other on the recipient chromosome. The recombination index depends on breaks and recombination events occurring between the two ami sites and is therefore mostly proportional to the physical distance separating them. On average, for point mutations in a hex background, a recombination index of 1% corresponds to a distance of 27 bp between
Table 1. Effect of insertion length on transformation efficiency

Efficiency is measured as the ratio of either \(ami\) to \(Str'\) transformants (donor insertions, wild-type recipient; central column), or \(ami^+\) to \(Str'\) transformants (wild-type donor, recipient insertions; right column). Each value is the mean efficiency measured from three independent transformations with two platings, each time screening for \(ami\) or \(ami^+\) and \(Str'\) colonies.

<table>
<thead>
<tr>
<th>Insertion length (bp)</th>
<th>Transformation efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(ami)-ins donor (\times) (ami^+) recipient</td>
</tr>
<tr>
<td>1374</td>
<td>0.55</td>
</tr>
<tr>
<td>665</td>
<td>0.63</td>
</tr>
<tr>
<td>359</td>
<td>0.64</td>
</tr>
<tr>
<td>92</td>
<td>0.78</td>
</tr>
<tr>
<td>64</td>
<td>0.80</td>
</tr>
<tr>
<td>4</td>
<td>1.03</td>
</tr>
</tbody>
</table>

Table 2. Effect of insertion length on recombination index in two-point crosses involving donor insertions

Chromosomal DNAs carrying both an \(ami\) insertion and the \(Str'\) marker were used to transform either the \(ami-22\) or the \(ami-42\) mutant. \(ami^+\) recombinants and \(Str'\) transformants were counted from three platings and the mean values are indicated below.

<table>
<thead>
<tr>
<th>Insertion length (bp)</th>
<th>Mean recombination index observed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(ami-22) recipient</td>
</tr>
<tr>
<td>1374</td>
<td>22</td>
</tr>
<tr>
<td>665</td>
<td>23.4</td>
</tr>
<tr>
<td>359</td>
<td>16</td>
</tr>
<tr>
<td>92</td>
<td>7.3</td>
</tr>
<tr>
<td>64</td>
<td>8.8</td>
</tr>
<tr>
<td>4</td>
<td>4.8</td>
</tr>
</tbody>
</table>

The mutations (Claverys et al., 1979). One set of crosses was carried out with insertions located on the donor DNA. The recipient mutation was either \(ami-42\) or \(ami-22\). These two point mutations are located, respectively, on the left and on the right of the insertion site (\(BamHI\)) and very close to it (Fig. 1). Crosses involving the 4 bp insertion gave the recombination indices expected for a point mutation, based on the locations of \(ami-22\) and \(ami-42\) relative to the insertion site, i.e. 4.8 and 1.2%, respectively (Table 2). When using longer insertions, experimental indices of \(ami^+\) recombinants exceed these reference values, i.e. insertions trigger hyper-recombination. This behaviour amplifies with the length of the heterologous insertion and appears to reach a plateau for the two longest insertions (Table 2). We also observed that hyper-recombination was higher when crosses were performed with \(ami-22\) than with \(ami-42\). For instance, the 665 bp insert generates about 19% \(ami^+\) transformants above the 4.8% expected with \(ami-22\), but 15% above the 1.2% expected with \(ami-42\). This differential hyper-recombination between the proximal and the distal recipient marker remains for each insertion (Table 2) and was previously observed in crosses involving long donor deletions (Lefèvre et al., 1989). Whatever the mechanism responsible for hyper-recombination, the close vicinity of a long heterologous insertion apparently affects recombination at a genetic site. Similar crosses were also performed for each donor insertion in an R801 \(polA\) background with either \(ami-22\) or \(ami-42\). The lack of DNA polymerase I had no effect on hyper-recombination (data not shown).

A second set of crosses was performed, using insertions in the recipient. Donor DNA was \(ami-22\) \(str-41\) or \(ami-42\) \(str-41\) (Table 3). Crosses involving the 4 bp insertion generate numbers of \(ami^+\) colonies very close to those obtained above, still in agreement with the distances between genetic sites. For longer insertions, we no longer observe hyper-recombination. The numbers of \(ami^+\) transformants are even below the reference value obtained with the 4 bp insertion, decreasing slightly as the length of the heterologous insertion increases. This decrease mostly disappears when dividing the recombination index for each insert by the corresponding efficiency obtained in crosses.
Exclusion of heterologous insertions in *S. pneumoniae*

**Table 3.** Effect of insertion length on recombination index in two-point crosses involving recipient insertions

Chromosomal DNA carrying the point mutations *ami-22* or *ami-42*, together with the Str' marker, was used to transform *ami*-insertion recipient strains. *ami*" recombinants and Str' transformants were counted on three platings and the mean values are indicated below.

<table>
<thead>
<tr>
<th>Insertion length (bp)</th>
<th>Mean recombination index observed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>ami-22 donor</em></td>
</tr>
<tr>
<td>1374</td>
<td>2</td>
</tr>
<tr>
<td>665</td>
<td>2</td>
</tr>
<tr>
<td>359</td>
<td>2.6</td>
</tr>
<tr>
<td>92</td>
<td>3.5</td>
</tr>
<tr>
<td>64</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>5.8</td>
</tr>
</tbody>
</table>

**Fig. 2.** Models accounting for the reduced transformation efficiency and hyper-recombination of insertions. (a) Pairing is completed successfully, generating a single-stranded loop which may be repaired to wild-type. (b, c) Exclusion or one-sided pairings: recombination stops because heterology is a strong barrier and/or because homology on one side is absent or very short. (d) Both flanking homologies are too short; there is no recombination at all. Situations (a), (b), (c) and (d) account for the reduced efficiency of insertions. Situations (a) and (b) also explain hyper-recombination in the two-point cross symbolized in this figure. Black line, donor strand; grey lines, recipient strands; dashed line, heterologous *ami*-insertion.

between donor *ami*" and recipient *ami*-insertions (calculation not shown). Thus, the reduced efficiencies and the slight hyporecombination observed when using recipient insertions are likely to have the same cause.

**Testing for a loop repair or one-sided pairings**

The results described above show that, with respect to point mutations, long donor insertions have a reduced transformation efficiency in one-point crosses and induce hyper-recombination in two-point crosses. We propose two models to account for this behaviour (Fig. 2). The same models may also account for hyporecombination due to recipient insertions.

First, recombination of a fragment carrying an insertion should generate a single-stranded heteroduplex loop. Such structures might undergo correction biased to the wild-type sequence and be closely localized around the loop to prevent co-conversion, as shown in Fig. 2(a). In addition, this correction would not be dependent on DNA polymerase I, since *polA* mutants allow hyper-recombination to occur. The second model, the exclusion model, is based on the fact that a long insert is a barrier which might inhibit recombination of donor DNA. Moreover, donor fragments broken within the insertion cannot fully recombine. Such one-sided pairings, shown in Fig. 2(b, c), will exclude the long insertion, broken or not, from the pairing synapsis, thus reducing their efficiency. These events will also produce hyper-recombination (Fig. 2b).

In order to distinguish between the two possibilities, we have carried out two kinds of three-point crosses. In a first set of crosses (three-point cross of type I, Fig. 3), the double mutant strain *ami-42 ami-6* was used as recipient. *ami-42* and *ami-6* flank the insertion site BamHI (Fig. 1). *ami*-insertions were located on the donor DNA. In this situation a loop repair event should generate a wild-type transformant when the central insertion and both flanking
Fig. 3. Three-point cross of type 1. This class of cross was used to test the loop repair model in a positive way. ami-insertions were used to transform the ami-42 ami-6 double mutant. A localized loop-repair event (a) should mostly generate a wild-type transformant, as in the two-point cross of Fig. 2(a). One-sided pairings (b), which exclude the heterologous insertion and flanking DNA, should lead to a mutant genotype, while they generate an ami+ genotype in the two-point crosses of Fig. 2(b). Line symbols are as in Fig. 2.

Table 4. Effect of insertion length on recombination indices in three-point crosses of type I

The double mutant ami-42 ami-6 was transformed with the ami-insertions, as shown in Fig. 3.

<table>
<thead>
<tr>
<th>Insertion length (bp)</th>
<th>Observed*</th>
<th>Predicted by a loop repair model†</th>
<th>Predicted by double recombinations‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>1374</td>
<td>0.68</td>
<td>10.5</td>
<td>0.4</td>
</tr>
<tr>
<td>665</td>
<td>0.62</td>
<td>12.5</td>
<td>0.4</td>
</tr>
<tr>
<td>359</td>
<td>0.53</td>
<td>6.7</td>
<td>0.4</td>
</tr>
<tr>
<td>92</td>
<td>0.37</td>
<td>2.7</td>
<td>0.4</td>
</tr>
<tr>
<td>64</td>
<td>0.19</td>
<td>3.7</td>
<td>0.4</td>
</tr>
<tr>
<td>4</td>
<td>0.38</td>
<td>–</td>
<td>0.4</td>
</tr>
</tbody>
</table>

* Observed indices were calculated as described in Methods.
† These are the expected indices if repair is the cause of hyper-recombination. With respect to the situation of Fig. 2(a), the number of loop repairs leading to a wild-type strand should be reduced by about 23%, i.e. the frequency of events integrating ami-42* and ami-ins, without integrating ami-6*. Values observed in two-point crosses (Table 2, ami-42 recipient) were multiplied by 77% to give the predicted indices. This calculation is not relevant for the control insertion of 4 bp which does not induce hyper-recombination.
‡ Values correspond to independent recombination events: integration of ami-42* without ami-ins (about 1.8%) and integration of ami-6* without ami-ins (about 23%).

Table 4 shows the effect of insertion length on recombination indices in three-point crosses of type I. The double mutant ami-42 ami-6 was transformed with the ami-insertions, as shown in Fig. 3. The results of these crosses do not support a loop repair model. On the other hand, they are compatible with a one-sided pairing model, although they do not directly prove it.

To obtain evidence directly supporting the latter possibility, i.e. the exclusion model, also called a one-sided pairing model, we have carried out a second kind of cross (three-point cross of type II, Fig. 4). Three double ami mutant strains were constructed, containing one of the larger ami-insertions and the ami-6 mutation. The streptomycin-resistance marker was also present in these strains. Chromosomal DNA from these mutants was used to transform the ami-42 recipient strain. The following is assumed: with respect to crosses between the ami-42 recipient and the ami-insertion donors, the exceedingly large number of ami* bacteria should be mostly unchanged in such three-point crosses provided that one-sided pairings are responsible for hyper-recombination. Indeed, according to this model the ami-6 mutation which flanks the heterologous insertion should be excluded with it, or unlinked by breaks occurring within the insertion (Fig. 4a). In addition, one should note that in this kind of cross, loop repair events cannot generate a wild-type
Exclusion of heterologous insertions in S. pneumoniae

**Fig. 4.** Three-point cross of type II. This class of cross was used to test the exclusion model in a positive way. (a) The insertion being excluded (the linked point mutation ami-6) should also be excluded from the pairing or should be unlinked by entry breaks within the insertion. This should lead to a wild-type recombinant as in the two-point cross of Fig. 2(b). (b) Loop corrections should not frequently lead to a wild-type transformant, with respect to the two-point cross of Fig. 2(a). Line symbols are as in Fig. 2.

---

**Table 5.** Effect of insertion length on recombination indices in three-point crosses of type II

The ami-42 mutant was transformed with the doubly mutated DNA (ami-ins-1374 ami-6, ami-ins-665 ami-6 or ami-ins-359 ami-6) as shown in Fig. 4.

<table>
<thead>
<tr>
<th>Insertion length (bp)</th>
<th>Mean recombination index values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed*</td>
</tr>
<tr>
<td>1374</td>
<td>16</td>
</tr>
<tr>
<td>665</td>
<td>14</td>
</tr>
<tr>
<td>359</td>
<td>7.6</td>
</tr>
</tbody>
</table>

* Observed indices are calculated as described in Methods.
† Expected values if exclusion is the cause of hyper-recombination. The prediction is as follows. As ami-6 is excluded with the heterologous insertion (Fig. 4a), the situation is identical to that of Fig. 2(b) and should lead to the same indices. Thus values predicted from the exclusion model are those from Table 2 (ami-42 recipient).
‡ Predicted ami+ indices, based on a loop correction model. The prediction is as follows. The situation becomes identical to Fig. 2(a) when ami-ins integrates without ami-6. As the disjunction frequency between ami-6 and the insertion site is about 23%, ami+ numbers should be reduced to 23% of those observed in two-point crosses (Table 2, ami-42 recipient), thus giving the numbers indicated.

---

The results with donor insertions indicate that reduced efficiency and hyper-recombination were also characteristic of donor deletions. In particular, results obtained with the 395-base deletion ami-109 (Lefèvre et al., 1989) indicated about 20% hyper-recombination. However, this excess of ami+ recombinants was attributed to a loop correction pathway, not to an exclusion of the donor deletion. To determine whether completely different mechanisms for donor insertion and donor deletion account for similar results, i.e. reduced efficiency and hyper-recombination, we have reinvestigated the behaviour of a donor deletion. For this purpose we carried out the same three-point crosses as described above, but involving ami-109 as a long heterologous insertion.

First, the double mutant strain ami-42 ami-6 was transformed with donor DNA carrying ami-109 (Fig. 5). This deletion is located approximately in the middle of the two sites (Fig. 1). A conversion of the deletion to the wild-type sequence suggests that hyper-recombination remains (Fig. 5a). Actually, the situation is identical to a cross between the ami-109 deletion donor and the ami-6 recipient.

---

**Using the ami-109 deletion in three-point crosses**

The results with donor insertions indicate that reduced efficiency and hyper-recombination are not consistent with a loop correction model. Reduced transforming efficiencies and hyper-recombination were also characteristic of donor deletions. In particular, results obtained with the 395-base deletion ami-109 (Lefèvre et al., 1989) indicated about 20% hyper-recombination. However, this excess of ami+ recombinants was attributed to a loop correction pathway, not to an exclusion of the donor deletion. To determine whether completely different mechanisms for donor insertion and donor deletion account for similar results, i.e. reduced efficiency and hyper-recombination, we have reinvestigated the behaviour of a donor deletion. For this purpose we carried out the same three-point crosses as described above, but involving ami-109 as a long heterologous insertion.

First, the double mutant strain ami-42 ami-6 was transformed with donor DNA carrying ami-109 (Fig. 5). This deletion is located approximately in the middle of the two sites (Fig. 1). A conversion of the deletion to the wild-type sequence suggests that hyper-recombination remains (Fig. 5a). Actually, the situation is identical to a cross between the ami-109 deletion donor and the ami-6 recipient,
provided that the ami-109 deleted site co-recombines with the ami-42\textsuperscript{+} site. According to the positions of ami-42 and ami-109, disjunction between the two sites should be about 10\%. Consequently, according to a loop correction model, the ami\textsuperscript{−} index generated in the present three-point cross should be reduced by 10\% as compared with crosses between the ami-109 donor and the ami-6 recipient. Since this latter cross produces around 27\% ami\textsuperscript{+} colonies, the expected value based on a correction model should be approximately 24\% ami\textsuperscript{+} colonies. On the other hand, if one-sided pairings are responsible for the excess of ami\textsuperscript{+} transformants, this excess should disappear in this cross since one lateral mutation remains on the chromosome (Fig. 5b). We have obtained 0.71\% ami\textsuperscript{+} transformants, i.e. close to the expected number based on independent recombinations integrating both ami-42\textsuperscript{+} and ami-6\textsuperscript{+} sites without integrating the central deletion ami-109. Such events may be estimated at around 1\%, since the distance between ami-109 and ami-6 or ami-42 is about 10\%.

In addition, we have constructed a double mutant ami-109 ami-42. Its DNA was used to transform the ami-6 mutant strain (Fig. 6). Complete pairing and conversion of the deletion to the wild-type (Fig. 6a) should lead to a drastically reduced number of ami\textsuperscript{+} colonies relative to the cross between the ami-109 donor and the ami-6 recipient. Since recombination between the ami-109 site and the ami-42 site is approximately 10\%, the ami\textsuperscript{−} index in the present three-point cross should fall between 2 and 3\%. Conversely, if pairing stops at the heterologous insertion, the ami-42 mutation is excluded (Fig. 6b), and the number of ami\textsuperscript{−} colonies should be mostly the same as in a cross between the ami-109 DNA and the ami-6 recipient. We have observed an ami\textsuperscript{−} frequency identical to the one obtained in the cross between the ami-109 donor and the ami-6 recipient, i.e. between 25\% and 30\%. These results firmly support the prediction based on exclusion events to account for recombination of donor deletions. The conversion model, previously proposed (Lefèvre et al., 1989), is not confirmed by the present study; we have no satisfactory hypothesis to explain previous observations.

**DISCUSSION**

Several insertions and one deletion were used to elucidate the recombination behaviour of long heterologous insertions. By a combination of two- and three-point crosses, we have found that both donor insertions and donor deletions are not repaired at the heteroduplex level. Reduced transforming efficiency and hyporecombination, induced by long deletions and insertions, are best explained by pairing inhibition excluding the donor heterologous insertion from the heteroduplex joint. Such exclusion may also account for the low efficiency and hyporecombination observed in crosses involving recipient insertions (Table 1, right column, Table 3). In that situation, recombination is identical to transformation involving donor deletions, i.e. the wild-type site should be frequently excluded from the pairing synopsis. In two-point crosses, while exclusion of donor insertions directly generates a wild-type transformant, exclusion of donor wild-type sites does not. Moreover, such exclusion events are 'lost' for recombining the wild-type site into the chromosome, thus explaining the slight hyporecombina-
tion observed in two-point crosses involving recipient insertions (Table 3).

Thus, effective pairing of homologous sequences would be initiated by the invading single-strand donor DNA, and the pairing process would be partly blocked at the level of long heterologous insertions. Inhibition is only partial, since the transformation efficiency of long heterologous insertions is reduced approximately to one-third or one-half of the efficiency of point mutations. This is consistent with in vitro studies showing that although RecA of *E. coli* is able to catalyse recombination of long heterologous insertions, the process is less efficient than for homologous DNA (Bianchi & Radding, 1983). In addition, regarding pneumococcal transformation, donor insertions must undergo nicks during entry. As the mean size for entering fragments is about 6000 bases (Morrison & Guild, 1972), insertions of a few hundred bases should be broken with frequencies close to 10% and more if longer. Such broken insertions cannot be efficient in one-point crosses whatever the situation, i.e. no recombination at all or recombination on one side only (Fig. 2). Natural breaks must explain, in the case of donor insertions, a significant part of the observed reduction of efficiency and hyper-recombination.

In this study we also found that insertions transform as well as or more efficiently than deletions of corresponding size. This observation was unexpected, because deletions have no physical size and should behave as single point mutations with regard to entry, while insertions undergo breaks. However, this observation is consistent with previous in vitro findings of Bianchi & Radding (1983): heteroduplexes are spanned more efficiently by RecA when the heterologous insertion is on the single-stranded partner rather than on the duplex DNA. Thus, we propose that the unwinding of a heterologous chromosomal region is the limiting step for the transformation of a donor deletion. It should be a more limiting step than single-stranded DNA folding is for the transformation of a donor insertion. Indeed, in spite of entry breaks destroying them, donor insertions appear to be more transformable markers than donor deletions, at least for the heterologous insertion sizes studied in this work. Lacks (1966) found that the transformation efficiency of the multisite mutation *malE9*, probably a very long deletion, is 0.4. Conversely, when transforming a *malE9* strain with *mat* DNA, i.e. in a situation identical to a donor insertion, the transformation efficiency of the wild-type locus decreased to 0.009. As the *malE9* deletion is very large, i.e. close to the mean size of an entering fragment, almost all *mat* entering fragments should be broken within the sequence deleted on the chromosome, thus explaining the very low efficiency. In our study, insertions are small compared to the mean length of entering DNAs. Interestingly, when carried on donor DNA, the *malE9* deletion had a transformation efficiency of 0.4 (Lacks, 1966). This value is very close to those observed in the present study when transforming strains containing the 665 or the 1374 bp insertions with *ami* DNA. This observation could indicate that, although a multibase chromosomal heterologous insertion partly inhibits recombination of the donor single-stranded fragment, it can be bypassed in vivo, and this bypass could occur efficiently whatever the heterologous insertion size. This initial decrease of recombination for DNAs carrying heterologous insertions beyond a few hundred bases, followed by a plateau for longer heterologous insertions, is reminiscent of the findings of Bi & Liu (1994). They characterized two kinds of recombination in *E. coli*: recA-independent recombination, sensitive to the heterologous insertion length between two repeats, and recA-dependent recombination, sensitive to homology sizes.

The crosses carried out in this study suggest that no pathway specialized for correcting multibase loop heteroduplexes exists in *S. pneumoniae*. Nevertheless, such structures must be formed when recombination of a long donor heterologous insertion is completed successfully; this is likely to occur as long heterologous insertions are transformable markers. Thus, *S. pneumoniae* appears not to correct big looped-out heteroduplexes, while most point mismatches are repaired; regarding the respective bulks, this may appear paradoxical. However, it is likely that this situation is satisfactory for bacteria. Similarly one should recall, for example, that C/C mismatches appear to be very destabilizing for the DNA structure (Gasc et al., 1989). Nevertheless C/C escapes correction by the Hex pathway (Claverys et al., 1983) and is poorly repaired by the *E. coli* Mut system (Kramer et al., 1984). Indeed, this mismatch should not be detrimental to bacteria: transversions, especially G → C, do not arise frequently because of the proofreading activity of DNA polymerase III (Wu et al., 1990). Conversely, A/C and G/T, which are efficiently repaired by the Hex and Mut systems, result from transitions, which are not efficiently repaired by the proofreading activity of DNA polymerase III (Wu et al., 1990). In the same perspective, we can imagine that looped-out heteroduplexes do not arise frequently in bacterial chromosomes and that specific counteracting systems have not evolved in bacteria. In addition, considering natural transformation, pneumococci could benefit by retrieving through recombination, long heterologous or heterospecific DNA fragments. Such horizontal transfers of blocks of foreign DNA have been proposed to explain the emergence of highly penicillin-resistant strains of *S. pneumoniae* among hospital isolates (Dowson et al., 1989). Indeed, we can imagine that like heterologous insertions in this study, highly divergent regions could, sometimes, lead to the formation of heteroduplex bubbles and be integrated into a foreign chromosome, provided there is enough homology on flanking sides.

In *E. coli*, no repair system specializing in long multibase loop heteroduplexes was found (Carraway & Marinus, 1993). On the other hand, based on tetrad analysis, it was concluded that two long deletions undergo conversion in yeast (Fogel et al., 1981). Correction mechanisms acting on single-strand bubble heteroduplexes have been demonstrated in mammals (Ayares et al., 1987; Weiss & Wilson, 1987). The large size and high sequence redundancy of eukaryotic genomes may frequently induce polymerase slippage or interfere with meiotic recombinations. Subsequent looped-out structures could arise in high prop-
ortions and be toxic for cells unless they are repaired by specific pathways.

ACKNOWLEDGEMENTS

We are indebted to Caroline Monod for linguistic assistance. This work was supported, in part, by the Université Paul Sabatier.

REFERENCES


Received 7 August 1995; revised 3 October 1995; accepted 13 October 1995.