Rearrangement of the Bacterial Chromosome: Forbidden Inversions

ANCA SEGALL, MICHAEL J. MAHAN,* JOHN R. ROTH

The order of genes in the chromosome of enteric bacteria has been evolutionarily conserved despite the existence of mechanisms for rearrangement. Homologous chromosomal sequences in the same orientation recombine to form deletions or duplications. When homologous sequences in inverse orientation recombine, one expects to form an inversion of the intervening chromosomal segment. This expectation was tested by placing pairs of homologous sequences in inverse order at various points in the chromosome. Sequences at many pairs of sites (permissive) do recombine to generate the expected inversion, while the same sequences placed at other pairs of sites (nonpermissive) do not form an inversion. For the one nonpermissive interval tested, the missing inversion type can be constructed by an alternative transductional method; strains with this inversion are viable. Thus mechanistic limitations must prevent sequences at particular sites from undergoing the recombination event required to form an inversion.

TWO ENTERIC BACTERIA, *Salmonella typhimurium* AND *Escherichia coli*, are thought to have descended from a common ancestor that lived 150 million years ago; since then, these organisms have probably shared little genetic exchange (*1*). Over this time span, DNA sequences have diverged significantly; essential coding sequences typically show 10 to 20 percent divergence (*2*), whereas nonessential sequences have diverged more widely (*3*). Recombination between the two organisms is typically reduced by a factor of more than 10^5 (*4*). The chromosomes of these bacteria have been extensively mapped genetically (*5*). A comparison of the maps reveals that, during 150 million years, gene order has diverged very little. It is not known what selective forces or mechanistic restrictions act to conserve gene order. While even slight growth disadvantages might explain map conservation, experimental results suggest that mechanistic limitations exist. We have studied intrachromosomal recombination events that form duplications and inversions in hopes of identifying factors that limit rearrangement.

Tandem duplications are frequent mutations that arise by recombination between repeated chromosomal sequences in the same orientation (*6*). The occurrence of duplications seems to be limited only by detrimental effects of some large duplications and by lethal effects of duplications that include the terminus of replication (*7*). Tandem duplications are reversible in that recombination between the two copies causes loss of one copy and returns the chromosome to its original haploid structure. Therefore, while duplications are common and may have short-term selective value, their reversibility reduces the probability that they will lead to stably heritable changes in chromosome structure.

Inversions, which arise by recombination between repeated sequences in opposite orientation, are more likely to lead to permanent rearrangement of the genome, since reversal of an inversion is no more likely than the occurrence of a second, different inversion by recombination between a second pair of repeated sequences. Very few examples of inversion mutations in bacteria have been reported (*8–10*). Konrad devised a method for deliberate construction of inversions; none were found, suggesting a barrier to formation or survival of inversion mutants (*11*). In contrast, Hill and co-workers have observed frequent inversions of a different chromosomal interval (*9*). To study chromosomal inversions, we previously developed a system for comparing the frequency of inversion and duplication mutations having a breakpoint in a single chromosomal region (*12, 13*). Inversions revealed by this system are 100 times less frequent than duplications.

To study inversion formation in more detail, and investigate potential barriers to inversion, we constructed strains with pairs of homologous sequences placed in inverse orientation at various points in the chromosome of *S. typhimurium*. Recombination between such sequences is expected to invert the intervening chromosomal segment (Fig. 1). We have modified methods of Konrad (*11*) in order to develop two systems for selecting recombinants; both systems allow survivors to arise either by inversion or by alternative recombination events. These systems can detect inversion formation at many sites in the chromosome.

The selection systems. In selection I, portions of the *his* operon serve as recombining sequences. Part of the *his* operon (the promoter, *hisG*, *hisD*, and part of the *hisC* gene) is placed at a site far from the normal *his* locus; these *his* sequences are in inverse order vis-à-vis the normal *his* operon. This distant, inverse sequence is flanked by direct repeats of transposon Tn*10*. Methods for constructing these strains have been described (*14, 15*). The normal *his* locus has a deletion mutation that removes the promoter, *hisG*, and part of the *hisD* gene, but leaves the rest of the operon (part of the *hisD* gene and the *hisC-hisE* genes) intact. The resulting strain is phenotypically His^- because the genes at the *his* locus lack their normal promoter. The two parts of the operon share about 1 kilobase (kb) of sequence homology (part of the *hisD* and part of the *hisC* genes). If these shared homologies recombine reciprocally, a normal his^+ operon is reconstructed. This exchange inverts the chromosomal segment between the *his* locus and the site of the distant *his* sequences.

An alternative set of events can generate a His^+ clone without

The authors are in the Department of Biology, University of Utah, Salt Lake City, UT 84112.

*To whom correspondence should be sent.

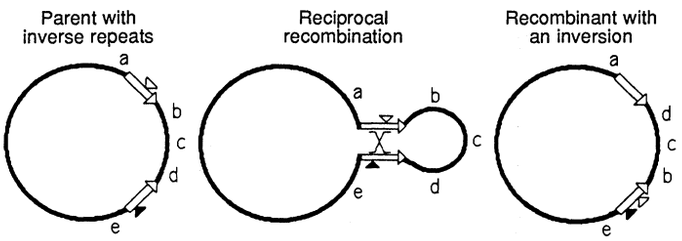

Fig. 1. Inversion formation by intrachromosomal recombination. Sequences positioned in the circular bacterial chromosome in inverse orientation can recombine to invert the intervening chromosomal segment. This event can be positively selected if the sequences involved are mutant alleles of a gene whose function can be selected. In the example above, the homologies are copies of the *lacZ* gene (open arrows), each with an inserted transposon (triangles), as described in the legend to Fig. 3. Nonreciprocal and double exchanges between such sequences can generate a Lac$^+$ recombinant without an inversion.

causing an inversion. The direct repeats of the Tn*10* element can recombine to excise *his* material from the distant site, generating a small circular chromosome fragment. This circle can recombine at the normal *his* locus, inserting itself and generating a complete *his*$^+$ operon without inversion formation (*15*).

In selection II, *lac* operon sequences provide the separated chromosomal homologies. The *lac* operons used are included in derivatives of phage Mu (Mud prophages) originally constructed by Casadaban and co-workers (*16*) and modified by Hughes and Roth (*17*). These prophages include a *lac* operon and a selectable drug resistance determinant; the prophages are able to transpose to a variety of sites in the chromosome. We have constructed strains that have two such prophages, inserted in inverse order at widely separated chromosomal sites. The *lacZ* gene of one prophage is defective because of insertion of a Tn*10* element; the second prophage has a Tn*5* insertion in the *lacZ* gene (for some experiments) or is inactive due to the absence of a functional promoter. Such strains are phenotypically Lac$^-$ because neither of the *lacZ* alleles is functional. Recombination between the two defective *lac* alleles can generate a normal *lacZ*$^+$ gene; such recombinants can be selected since they gain the ability to use lactose as a carbon source. A single reciprocal recombination event between the mutant sites of two *lacZ* genes generates a *lac*$^+$ allele and forms an inversion of the chromosomal segment between the prophage insertion sites (Fig. 1).

Alternative events can generate a *lac*$^+$ recombinant without inversion formation. The major class of noninversion recombinants has lost one of the *lacZ* insertion mutations, possibly by gene conversion. The second noninversion recombinant type is formed by two reciprocal exchanges that generate one *lac*$^+$ allele and one allele with both of the parental defects. Both alternative types arise by recombination between the two *lac* regions (*18*).

In applying these two selection systems, we constructed a series of strains with these inverse order homologies flanking various chromosomal intervals. Recombinants (His$^+$ or Lac$^+$ clones) were tested for an inversion of the intervening chromosomal segment. Methods for detecting inversions are described in the legends to Figs. 2 and 3. Using these methods, we scored the frequency of inversion-bearing strains and alternative types among the selected recombinants. An inversion construction system similar to those described here has been independently developed by Francois *et al.* (*19*).

Chromosome segments that are nonpermissive for inversion. The results of these experiments are striking in that the intervals tested fall into two distinct classes. Some intervals invert frequently, while others have never been seen to invert. Using selection I (*his*), we scored 20 or more His$^+$ recombinants for each strain (Fig. 2). For strains whose interval is classified as permissive, 65 to 90 percent of the His$^+$ recombinants carry an inversion. For strains classified as nonpermissive, none of the His$^+$ recombinants carry an inversion; only the alternative recombinant type is found. Permissive and nonpermissive intervals show a comparable overall frequency of His$^+$ recombinants.

The results of selection II (*lac*) are just as clear-cut (Fig. 3). Several hundred Lac$^+$ recombinants were scored for each strain. For strains with permissive intervals, 25 to 80 percent of the Lac$^+$ recombinants carry an inversion. Intervals classified as nonpermissive yield no inversions, only the alternative Lac$^+$ recombinant types. Two nonpermissive intervals were tested more rigorously; for both, no inversions were found among more than 2000 survivors of the selection. The frequency of total Lac$^+$ recombinants was typically 1×10^{-4}; two exceptional intervals are *ara-thr* (3×10^{-3}) and *his-trp* (2×10^{-7}). There was no correlation between frequency of Lac$^+$ clones and the presence of inversion recombinants.

Chromosomal position, not adjacent sequence, dictates permissivity. Permissive and nonpermissive intervals were identified by both selection systems. In each selection system, one particular pair

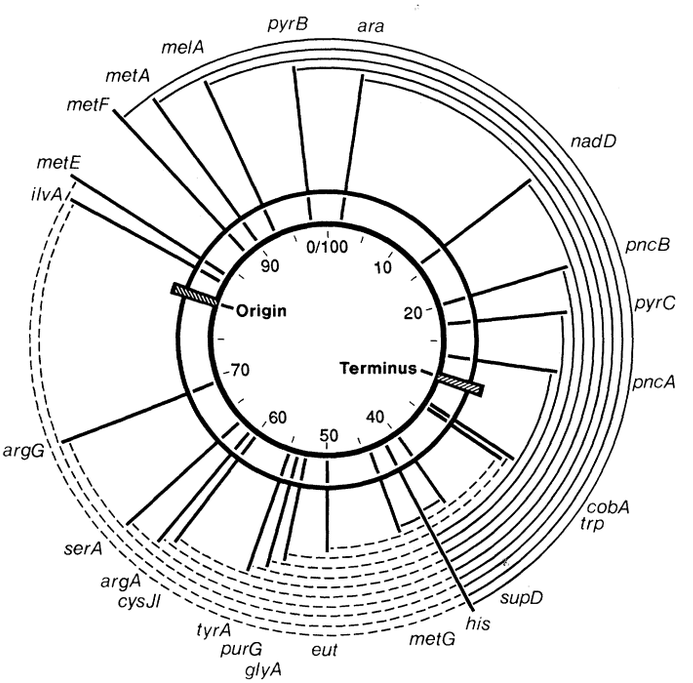

Fig. 2. Inversion generated by recombination between *his* sequences. Solid lines indicate chromosomal segments that invert when one selects for recombination between sequences present at the *his* locus and homologous sequences placed at a distant site in inverse orientation. Dashed lines indicate chromosomal segments that fail to invert under these selective conditions (Selection I). His$^+$ recombinants carrying an inversion are identified by transductional crosses requiring inheritance of a large *his* deletion associated with a drug resistance determinant. This deletion–drug resistance marker can only be inherited by recipient strains whose *his* region has both of the normal flanking sequences. Strains with an inversion have normal flanking sequence on only one side of the *his* region; on the other side they have sequence derived from a distant point in the chromosome. Recipients with such a rearrangement inherit the deletion–drug resistance marker at a 6- to 33-fold reduction in frequency compared to the parental strains; the transductants recovered appear to arise from recipient cells in which the inversion has "reflipped" restoring the normal chromosome structure. Recipients without an inversion are transduced at a frequency indistinguishable from that of the parents. The above test has been confirmed with the use of inversions whose structure has been verified by conjugation crosses (*15*). Strains inferred to carry inversions also show linkage disruption at the other join point of the inversion; disruption at this outlying site was scored as described in the legend to Fig. 3.

of homologous sequences was used; permissive and nonpermissive strains differ only by the chromosomal position of these sequences. Furthermore, several chromosomal intervals have been tested with both selection systems; the two systems give the same result. For example, the *his-trp* interval is nonpermissive for both systems and the *his-ara* interval is permissive for both systems. Clearly, the sequences involved directly in recombination do not determine ability to form inversions.

Sequences immediately adjacent to the recombining sequences do not dictate invertability. This is apparent in the distribution of intervals seen in Fig. 2; permissive and nonpermissive intervals are not randomly distributed but show a regional distribution. Furthermore, in selection system I, the *his* material at the outlying site is flanked by copies of Tn*10* (10 kb); for both permissive and nonpermissive intervals these 10-kb sequences separate the recombining sequences from adjacent chromosomal sequences that might dictate permissivity. This point was tested further when *his* sequences were placed at three sites within the *trp* operon and at one site near the closely linked *cobA* gene; *his* sequences at each site were tested individually for recombination with sequences at the *his* locus. All positions in this general region proved to be nonpermissive, suggesting that sequences immediately adjacent to the recombining sites do not dictate permissivity.

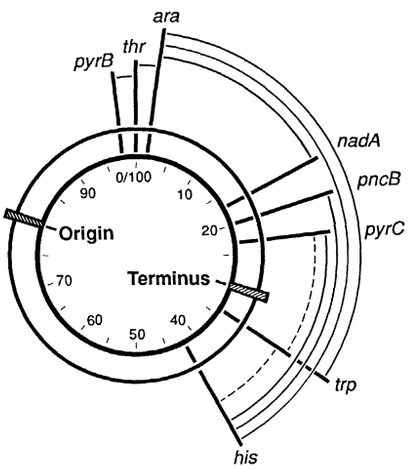

Fig. 3. Inversions generated by recombination between *lac* sequences present in Mu*d* (*lac*) prophages. Solid lines indicate chromosomal segments that invert when one selects for recombination between *lac* sequences placed, in inverse order, at the ends of the segment. Dashed lines indicate chromosomal segments that fail to show inversion under the same selective conditions. The *lac* sequences used are part of Mu*d-lac* prophages inserted within the indicated genes in inverse orientation (selection system II). Lac^+ recombinants inferred to carry an inversion are identified by their disruption of genetic linkage at both of the sites at which the Mud prophages are located. This is scored by transducing Lac^+ strains to prototrophy for nutritional requirements generated by insertion of Mud prophages in the parental strain. In the parental strain and in any Lac^+ recombinant without an inversion, the auxotrophic requirements can be repaired individually at high frequency by wild-type transduced fragments. In strains with an inversion, each of the auxotrophic Mud insertions is flanked by sequences that are normally widely separated in the chromosome. Therefore, it is impossible to repair either of the auxotrophies by recombination with a single wild-type transduced fragment. Strains inferred to carry an inversion show an 11- to 33-fold reduction in the number of transductants repaired for either of the two requirements. The validity of this method was demonstrated previously for several inversion intervals (*12, 13, 15*). The few transductants recovered appear to arise from recipient cells in which the inverted region has "reflipped."

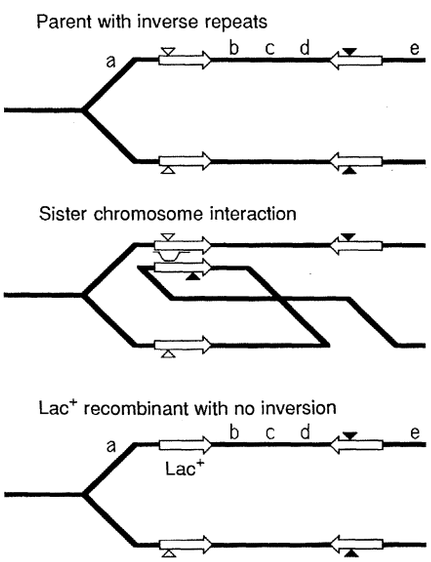

Fig. 4. Sister chromosome exchanges that yield Lac^+ recombinants without inversions. Inversely oriented sequences present in sister chromosomes can interact nonreciprocally or by two reciprocal exchanges to generate a Lac^+ recombinant that does not carry an inversion. An example of a nonreciprocal exchange is presented above.

A search for rules governing invertability. The distribution of permissive and nonpermissive intervals (Figs. 2 and 3) suggests the existence of a pattern that might be described by rules. We have been unable to generate simple rules that account for all the data, but the following generalizations seem relevant. (i) The four short intervals (less than 2 minutes on the genetic map) are permissive. (ii) Nonpermissive intervals are found within permissive intervals. (iii) End points of nonpermissive intervals are frequently permissive when tested for recombination with sequences at other chromosomal sites.

Inspection of the data from system I (Fig. 2) suggests that all permissive intervals must include the origin or terminus of replication. However, we do not believe this is a general rule since it does not fit with the behavior of many intervals without an end point in *his*. For example, *ara-nadA* is permissive without including the origin, and the *trp-pyrC* interval includes the terminus but is nonpermissive (Fig. 3).

We do not yet understand what forces "forbid" inversions and generate the pattern seen in Fig. 2. In approaching such an understanding, we have considered two explanations. Nonpermissive intervals might be dictated by either functional or mechanistic considerations.

Functional barriers to inversion detection. Inversions of a nonpermissive interval may form but have lethal consequences that prevent detection of the inversion-bearing recombinants. Several functional consequences of inversions can be imagined.

1) Since the bacterial chromosome is replicated bidirectionally from a fixed origin, genes near the origin replicate early and are present in higher copy number (on average, over the cell cycle) than genes located near the terminus (*20*). Inversions might be lethal or detrimental because they change the location of many genes vis-à-vis the origin of replication, thereby altering by a slight amount the expression of many gene products.

2) If the structure of the nucleoid (the folded bacterial chromosome) is determined by sequences at particular chromosomal sites, inversions might disturb the position of these sequences and have lethal effects on the folded structure.

3) Some chromosomal sequences may impede replication when placed in one orientation but not the other. An inversion might be detrimental because it placed such sequences in the obstructive orientation.

4) Inversion could disturb the symmetrical positioning of the origin and terminus, requiring one of the growing forks to copy more than half of the chromosome.

If some inversion mutations can be lethal, we would expect that many inversions that are detected might show sublethal growth defects. Of the inversions that we have studied, only one causes a marked reduction in growth rate. The exceptional detrimental inversion lies between the *ara* and *trp* loci (Fig. 3). Inversions of this interval occur (selection II) at a frequency comparable to that seen for other invertable intervals, but strains with the *ara-trp* inversion grow slowly. These slow-growing inversion strains revert frequently

to faster growth by secondary recombination events that reverse the inverted segment, thus restoring a normal chromosomal structure. Some, but not all, of these revertants have lost the Lac^+ phenotype. This is expected since reversal is caused by an exchange between a wild-type and a mutant *lac* region; depending on the position of the exchange, the product is either Lac^+ or Lac^-. Louarn (*10*) and Hill (*21*) and their co-workers have described intervals of the *E. coli* chromosome whose inversion causes a reduction in growth rate. The existence of detrimental inversions demonstrates that inversions can have functional consequences. However, the low frequency of detrimental inversions suggests to us that nonviability is not the major determinant of permissive and nonpermissive intervals.

Mechanism problems underlie some nonpermissive intervals. Failure to find a particular inversion would be explained if the cell had no mechanism capable of forming that rearrangement. For one nonpermissive interval, the failure to invert is clearly caused by a mechanism problem, rather than by lethal consequences of the finished rearrangement. The *his-trp* interval is nonpermissive by both of the original selection methods; no inversions of this segment (8 percent of the chromosome) have been detected among more than 2000 recombinants. If this inversion had lethal consequences, strains carrying it would be inviable regardless of how they were constructed. We have constructed this inversion by genetic methods (outlined below) distinct from the original selection systems (*18*) Strains bearing this inversion grow well and show no strong tendency to revert to the normal chromosome order. Since the *his-trp* inversion is not lethal, the failure to detect inversions by the original selections must reflect a mechanistic problem in forming the inversion under the original conditions. We believe that such mechanistic problems explain many nonpermissive intervals.

The problem that originally prevented formation of the *his-trp* inversion was circumvented by transducing two linear DNA fragments into a normal recipient cell; each transduced fragment included, in effect, sequences identical to one join point of the desired final inversion. Simultaneous inheritance of these fragments causes inversion of the intervening segment of the recipient chromosome (*12, 18*). The join point of the constructed inversion is just counterclockwise of the *his* operon; the other is within the *trp* operon. Both are within 5 kb of the sites (within *his* and *trp*) that were nonpermissive for direct formation of inversion in selection systems I and II. We believe that this transduction method succeeds because it involves exchanges between extrachromosomal fragments and the bacterial chromosome; no direct exchanges are required between two sequences present in the same circular chromosome.

Sequences in the same chromosome can interact directly or by sister chromosome exchanges. Repeated chromosomal sequences can interact in two general ways. The simplest is direct recombination between two sequences located within the same circular chromosome, as shown (Fig. 1) for sequences in inverse orientation. Alternatively, recombination can occur between copies of the same sequences present in different sister chromosomes after replication (Fig. 4). The noninversion recombinant types found in selection II (double recombinants and apparent gene conversion products) can arise by either of the two sorts of recombinational interactions, but sister chromosome interactions are sufficient to account for all noninversion recombinants.

Inversions can only be generated by recombination between two sequences present in the same circular chromosome (Fig. 1). In principle, inversions can also be formed by sister chromosome recombination; however, this requires two independent sister strand exchanges between two separated pairs of inverse repeats. From work on duplications, we know that a single sister strand exchange in regions of this size occurs at a frequency of 10^{-4}; on the basis of this, inversions requiring two coincident events would be expected to arise at a frequency of 10^{-8}. We could not detect such rare events by either of the selection systems that we used (detection would require testing more than 10^4 recombinants). Therefore, within the scope of our experiments, we believe inversions can arise only by a single exchange between sequences in the same chromosome.

A proposal. Since inversions require an exchange between two sites in the same circular chromosome, inversions would not be seen if sequences at some sites were excluded from such recombination events. We propose that inverse order sequences at nonpermissive sites can interact only by sister strand exchange; therefore they can generate the alternative recombinant types seen, but they cannot complete the event required for inversion. Sequences at permissive sites can engage in both intrachromosomal and sister strand exchange, allowing formation of both inversions and the alternative recombinant types.

If this proposal is correct, direct order sequences at nonpermissive sites should be able to recombine to form duplications, since this rearrangement can form by a single sister chromosome exchange. Thus, sites that are not permissive for inversion should be permissive for duplications. We tested this using selection system II; *lac* operons were placed in direct order at the ends of the *trp-his* interval, which is nonpermissive for inversions. Duplications represented about 10 percent of the total Lac^+ recombinants. This demonstrates that exchanges of flanking material can occur between direct order sequences at nonpermissive sites. We believe that these duplications arise by sister chromosome exchanges. While this is consistent with our proposal, we cannot eliminate the possibility that some of these duplications arose by intrachromosomal recombination between direct repeats, generating a circle that reinserted in a sister chromosome.

Although we do not know how recombination between particular sites might be prevented, we offer several possibilities. (i) A rigid three-dimensional structure of the chromosome might limit contact between specific sites. (ii) Differential supercoiling of domains of the chromosome might leave some pairs of sites recombinationally incompatible. (iii) The movement of replication forks may be such that, in growing cells, nonpermissive intervals frequently include a single replication fork. Inversion of a region including a replication fork may lead to an aberrant chromosome structure that cannot be corrected; lethal effects of these intermediate structures could prevent recovery of inversions of these regions; in this case, the finished inversion might not be functionally lethal once established, but the process of intrachromosomal recombination would generate structures with lethal consequences.

REFERENCES AND NOTES

1. H. Ochman and A. C. Wilson, Escherichia coli *and* Salmonella typhimurium: *Cellular and Molecular Biology*, F. C. Neidhart *et al.* Eds. (American Society for Microbiology, Washington, DC, 1987), p. 1649.
2. B. P. Nichols and C. Yanofsky, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 5244 (1979); I. Crawford, B. P. Nichols, C. Yanofsky, *J. Mol. Biol.* **142**, 489 (1980); M. S. Carlomagno, L. Chiariotti, P. Alifano, A. G. Nappo, C. B. Bruni, *ibid.*, in press.
3. D. J. Brenner, G. R. Fanning, K. E. Johnson, R. V. Citarella, S. Falkow, *J. Bacteriol.* **98**, 637 (1969).
4. R. B. Middleton, *Genetics* **69**, 303 (1971); C. Yanofsky, S. Li, V. Horn, J. Rowe, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 286 (1977).
5. B. J. Bachman, *Microbiol. Rev.* **47**, 180 (1983); K. E. Sanderson and J. R. Roth, *ibid.*, p. 410.
6. R. P. Anderson and J. R. Roth, *Annu. Rev. Microbiol.* **31**, 473 (1977).
7. D. Hillyard and J. R. Roth, unpublished results.
8. K. E. Sanderson and C. A. Hall, *Genetics* **64**, 215 (1979); F. Casse, M. C. Pascal, M. Chippaux, *Mol. Gen. Genet.* **124**, 213 (1973); D. J. Savic, S. P. Romac, S. D. Ehrlich, *J. Bacteriol.* **155**, 943 (1983); X. M. Xia and M. Enomoto, *Mol. Gen. Genet.* **205**, 376 (1986); M. Riley and S. Krawiec, Escherichia coli *and* Salmonella typhimurium: *Cellular and Molecular Biology*, F. C. Neidhart *et al.* Eds. (American Society for Microbiology, Washington, DC, 1987), p. 967.
9. C. W. Hill and B. W. Harnish, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7069 (1981).
10. J.-M. Louarn *et al.*, *Mol. Gen. Genet.* **201**, 467 (1985).
11. E. B. Konrad, thesis, Harvard University, Cambridge, MA (1969); *J. Bacteriol.* **130**, 167 (1977).

12. M. B. Schmid and J. R. Roth, *Genetics* **105**, 517 (1983).
13. ———, *ibid.*, p. 539.
14. ———, *ibid.*, **94**, 15 (1980).
15. M. J. Mahan and J. R. Roth, *ibid.* **120**, 23 (1988).
16. M. J. Casadaban and S. N. Cohen, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4530 (1984); B. A. Castilho, P. Olfson, M. J. Casadaban, *J. Bacteriol.* **158**, 488 (1984).
17. K. T. Hughes and J. R. Roth, *J. Bacteriol.* **159**, 130 (1984).
18. A. Segall and J. R. Roth, in preparation.
19. V. Francois, J. Louarn, J. Patte, J.-M. Louarn, *Gene* **56**, 99 (1987).
20. M. B. Schmid and J. R. Roth, *J. Bacteriol.* **169**, 2872 (1987).
21. C. W. Hill and J. A. Gray, *Genetics* **119**, 771 (1988).
22. Supported by USPHS GM 27068 (J.R.R.) and by NIH predoctoral training grant T32-GM 07464-11 (M.J.M. and A.M.S.).

29 March 1988; accepted 10 August 1988