Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 3398–3402, June 1982
Biochemistry

P1 site-specific recombination: Nucleotide sequence of the recombining sites

(dyad symmetry/*lox* site/BAL-31 deletion/plasmid recombination)

RONALD H. HOESS, MARCIA ZIESE, AND NAT STERNBERG

Cancer Biology Program, NCI-Frederick Cancer Research Facility, Frederick, Maryland 21701

Communicated by Philip Leder, February 8, 1982

ABSTRACT **Site-specific recombination between molecules of bacteriophage P1 DNA occurs at sites called *loxP* and requires the action of a protein that is the product of the P1 *cre* gene. Although recombination between two *loxP* sites is very efficient, recombination between *loxP* and a unique site in the bacterial chromosome (*loxB*) is inefficient and generates two hybrid *lox* sites called *loxR* and *loxL*. We present here the nucleotide sequences of all four *lox* sites. Analysis of these sequences indicates that (*i*) a region of extensive homology is not present at the *loxP* × *loxB* crossover point, in contrast to the 15-base pair common-core sequence in the bacteriophage λ *att* sites, and (*ii*) the sites contain a region of dyad symmetry with 8- to 13-base pair inverted repeats separated by an 8- to 9-base pair sequence. The *loxP* × *loxB* crossover point falls in the sequence that separates the inverted repeats, and deletions that remove either the left or the right inverted repeat of *loxP* inactivate the site. These two observations are consistent with the conclusion that the region of dyad symmetry is important in *lox* recombination. We have shown further that the *loxP* × *loxP* crossover point occurs in a 63-base pair sequence containing the *loxP* × *loxB* crossover point, suggesting that, despite the great difference in efficiencies of the two reactions, the crossover points may occur at the same place in both. Explanations for the different recombination properties of the various *lox* sites are discussed.**

Bacteriophage P1 is a temperate virus that has both lytic and lysogenic phases in its life cycle. In contrast to other temperate phages, such as λ, P2, and P22, phage P1 rarely integrates into the chromosome of its host in its lysogenic mode but maintains itself as an autonomous unit-copy plasmid (1). Recently, the existence of a site-specific recombination system in P1 has been demonstrated (2), and a number of functions have been ascribed to it. Among these are the cyclization of phage P1 DNA injected into a *recA*$^-$ host and the rare integration of P1 into the bacterial chromosome. Perhaps the most important role for this recombination, however, is to ensure proper segregation of P1 plasmid molecules to daughter cells at cell division (3). We have postulated that dimer plasmid molecules formed by homologous recombination between monomers interfere with orderly segregation of the products of plasmid replication to daughter cells. P1 has overcome this problem by being able to resolve dimeric molecules rapidly into monomers by means of its site-specific recombination system.

The site-specific recombination system of P1 consists of two elements, a site on the DNA called *lox* (for locus of crossing-over) and a phage-encoded protein, the product of the P1 *cre* gene (4). The site on P1 DNA is termed *loxP*, and the site on the bacterial chromosome into which the P1 plasmid integrates with low efficiency is called *loxB*. On integration of P1 DNA by recombination between *loxP* and *loxB* sites, two hybrid sites that we have designated *loxL* and *loxR* are generated. Because P1 DNA can integrate into *loxB* in either orientation, two sets

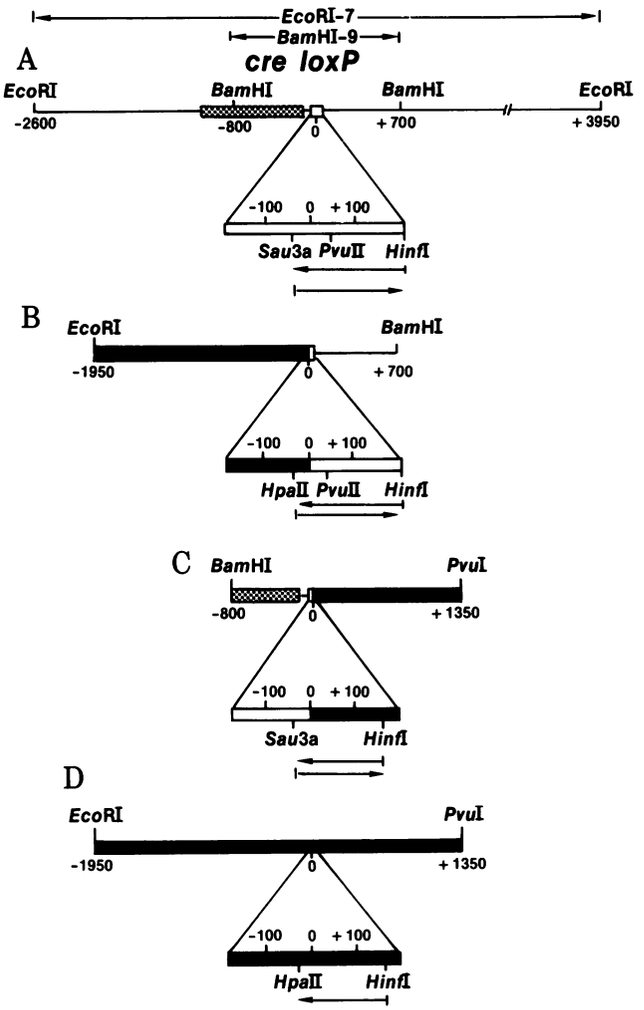

FIG. 1. Primary restriction fragments containing the *lox* sites and DNA sequence analysis strategy. (*A*) *loxP*. (*B*) *loxR*. (*C*) *loxL*. (*D*) *loxB*. Primary restriction fragments containing the *lox* sites were isolated from their respective hybrid λ phages (5) or pBR322 hybrid plasmids. ——, P1 DNA; ■, bacterial DNA; ▨, approximate location of the gene that mediates recombination at *lox* sites, *cre* (4). Below each primary fragment, an expanded restriction map of the region containing the *lox* site is given, with relevant restriction sites noted. →, direction and extent of DNA sequence obtained. Numbers indicate distance in base pairs from an arbitrary point located at the center of the *loxP* site.

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

Abbreviation: bp, base pair(s).

of *loxL* and *loxR* sites can be produced. In this paper, we confine our discussion to a set of *loxL* and *loxR* sites resulting from P1 integration in one orientation. All of these *lox* sites have been isolated on λ phages, and their ability to carry out *cre*-mediated recombination in λ phage crosses has been measured (5). These and other genetic experiments allow us to make a number of predictions regarding the DNA sequence of the *lox* sites. First, all *lox* sites cannot be structurally the same because each site has its own characteristic recombination efficiency. For instance, *loxP* and *loxL* sites are very efficient in *cre*-mediated recombination whereas *loxR* and *loxB* are very inefficient for this type of recombination (5). Second, the *loxB* site is probably palindromic since DNA containing *loxP* can integrate into *loxB* in either orientation with almost equal frequencies. We show in this report that the P1 sequences that define a functional *loxP* site consist of a 63-base pair (bp) region that also contains the *loxP* × *loxB* crossover point, suggesting that the crossover point may occur at the same place in both *loxP* × *loxP* and *loxP* × *loxB* reactions.

MATERIALS AND METHODS

Bacterial and Phage Strains. λ Hybrid phages carrying the *Eco*RI 7 restriction fragment of P1 and *loxL* or *loxR* sites have been described (5). The isolation of a λ *loxB* phage is described below. The isogeneic pair of strains N215 (*recA*$^+$) and N205 (*recA*$^-$) (4) was used to assay phage titers, isolate *loxB*, and measure *loxP* activity.

Chemicals and Enzymes. Restriction endonucleases, T4 DNA ligase, and BAL-31 were purchased from Bethesda Research Laboratories. Calf intestinal phosphatase and polynucleotide kinase were obtained from Boehringer Mannheim. [γ-^{32}P]ATP (≈3,000 Ci/mmol; 1 Ci = 3.7 × 10^{10} becquerels) was from New England Nuclear.

Isolation of the Chromosomal *loxB* Site. To isolate that segment of the *Escherichia coli* chromosome containing *loxB*, we took advantage of the fact that λ P1 *loxR* phage DNA cannot integrate into the bacterial chromosome either by λ integrative recombination (*attP* and *int* are deleted) or by P1 *lox* recombination (*cre* is deleted). Thus, when strain N215 is infected by λ P1 *loxR*, the lysogens that arise are the result of integration of λ DNA into the bacterial chromosome via recombination between homologous sequences in λ P1 *loxR* and the bacterial chromosome (5) (Fig. 1). Because the bacterial sequences in λ P1 *loxR* DNA are adjacent to *loxB*, integration places λ sequences adjacent to *loxB*. When the λ prophage present in these lysogens is induced, ≈ 1 in 10^4 phages is produced by an illegitimate excision of prophage DNA from the chromosome, replacing some λ genes with bacterial DNA sequences adjacent to the integration site. Phages arising by such a recombination event were selected for loss of the λ *gam* gene [the Spi$^-$ phenotype (6)]. The DNA of 15 Spi$^-$ phages was analyzed by restriction enzymes, and three phages were found to contain an *Eco*RI/*Pvu* I fragment (3.3 kilobases) that should contain *loxB* (Fig. 1). This fragment was then recloned between the single *Eco*RI and *Pvu* I sites of plasmid pBR322 (7) to facilitate sequence analysis.

DNA Sequence Determination. DNA was analyzed according to Maxam and Gilbert (8) with previously described modifications (9).

***In Vivo* Test for *loxP* Site Plasmids.** The test for the presence of a functional *loxP* site on a plasmid was similar to that previously used for screening for functional λ attachment sites on plasmids (10). Strain N205 (*recA*$^-$) containing the *loxP* plasmid is infected with λ-P1:7 (*cre*$^+$*loxP*$^+$). The lysate generated from this infection is then used to transduce the plasmid drug-resistance marker, in this case ampicillin resistance, to sensitive recipient cells. Since both host and λ recombination systems are inactive, the ability of the phage to transduce the plasmid is dependent on the P1 site-specific recombination to generate phage–plasmid hybrids. When the λ lacks the P1 *Eco*RI fragment 7 or pBR322 lacks the 1,500-bp *Bam*HI fragment containing *loxP* (*Bam*HI fragment 9), the plasmid is transduced infrequently.

Plasmid Construction. A series of plasmids containing *loxP* sites shortened with BAL-31 (11) was constructed as outlined in *Results*.

RESULTS

DNA Sequences of *loxP*, *loxL*, *loxR*, and *loxB* Sites. We have previously reported the isolation of λ phages containing *loxP*, *loxL*, and *loxR* sites (5). Subsequently, a λ phage containing the *loxB* site was isolated and a fragment containing that site was recloned in plasmid pBR322. DNA was prepared from the plasmid containing *loxB* and from phages containing *loxP*, *loxL*, and *loxR*. The *loxR* and *loxL* phages used here were derived from two separate *loxP* × *loxB* recombination events, both of which resulted in integration of *loxP* into *loxB* in the same orientation. A restriction map of the regions surrounding each of

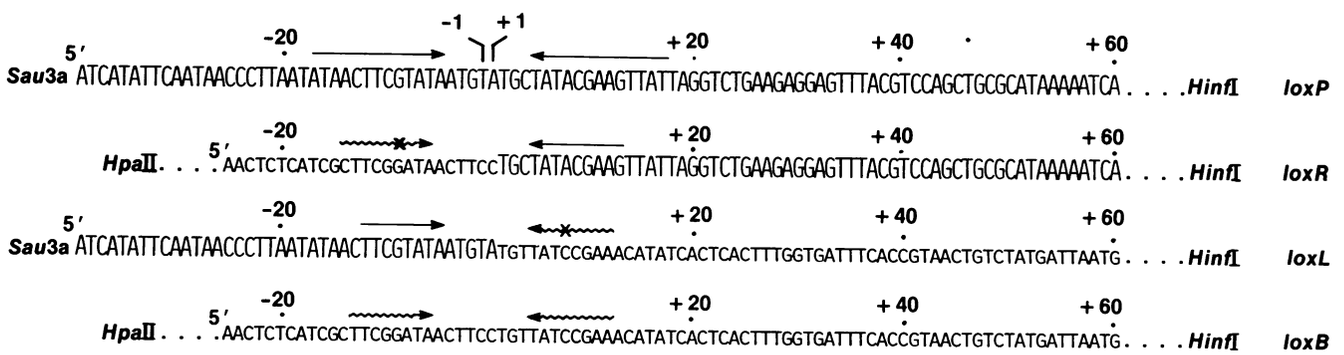

FIG. 2. Nucleotide sequences of the *loxP*, *loxR*, *loxL*, and *loxB* sites. The nucleotide sequence of one strand from each of the *lox* sites immediately surrounding the point of crossing-over is shown. Large type represents phage DNA sequences and small type represents sequences of bacterial origin. The axis of dyad symmetry between the largest palindromic sequences in *loxP* (→) is arbitrarily designated as the starting point for numbering bases, with the base immediately to the right +1 and the base immediately to the left −1. ⇝, Bacterial sequences capable of forming palindromic structures with the 13-bp inverted repeats in the phage DNA; x, mismatch with the complementary base in the phage inverted repeat. These sequences are drawn so that the ends of the palindrome (indicated by the head of the arrows) start at the same base in all four sites. The sequences shown here depict the crossover point as being between bases +1 and +2. Since there is a 2-bp homology between *loxP* and *loxB* at +2 and +3 (T-G), the crossover point could just as well be between +2 and +3 or immediately to the right of +3.

the sites was developed (5) (Fig. 1). Since *loxL* and *loxR* are hybrid sites generated by *loxP* × *loxB* crossover, comparison of their restriction patterns with those of *loxP* and *loxB* sites reveals the approximate location of the crossover point. The appropriate restriction fragments were isolated and the sequence analysis strategy outlined in Fig. 1 was followed.

The sequence data are shown in Fig. 2. As predicted by the restriction maps, the *loxL* and *loxR* sequences are hybrids between *loxP* and *loxB* sequences. From these data, it is not possible to say which internucleotide bond is broken during recombination since all four sites contain two bases of homology (5′ T-G 3′) at the point where *loxP* integrates into the *loxB* site. Unlike the λ attachment sites, there is no extensive common-core sequence shared by *loxP*, *loxB*, and the recombinant sites, *loxL* and *loxR* (12). The sequences shown in Fig. 2 also show that integration of *loxP* into *loxB* does not result in either addition or deletion of bases. A feature common to all of the sites is a region of perfect or nearly perfect dyad symmetry containing inverted repeat sequences of 8–13 bp.

Localization of the Crossover Region in *loxP* × *loxP* Recombination. Probably the most important role for P1 site-specific recombination is the resolution of dimeric P1 molecules produced either during or following DNA replication. This reaction requires recombination between two *loxP* sites. Since *loxP* × *loxP* recombination is 10^3–10^4 times more efficient than *loxP* × *loxB* recombination (5), the question arises whether the crossover region in *loxP* × *loxP* recombination is the same as that in *loxP* × *loxB* recombination.

Since the sequences involved in *loxP* × *loxP* recombination are completely homologous, it is not possible to define the point of crossover by the method used for *loxP* × *loxB*. However, assuming that the point of crossover is within the essential *loxP* sequences, its approximate location can be defined by determining the portion and size of the minimum functional *loxP* site. To address this question, we first cloned the 1,500-bp *Bam*HI fragment 9 containing *loxP* into pBR322, generating pRH30. Then, we constructed a series of deletions within this fragment (Fig. 3).

The first deletion plasmid, pRH31 removes all of the *Bam*HI fragment 9 DNA to the right of the *Pvu* II site [the *Pvu*II site is 44 bp from the axis of dyad symmetry (Fig. 4)]. This plasmid still retains a functional *lox* site that is capable of recombining at high efficiency with a *loxP* site carried on a λ phage. Likewise a BAL-31-generated deletion starting at the *Pvu* II site of pRH31 that removes sequences up to 6 bp from the right inverted repeat sequence of *loxP* (pRH31-56) leaves *loxP* functional (Fig. 4). A deletion generated in plasmid pRH31 to the left of the *loxP* inverted repeat starting at a *Bcl* I site and removing all of the *Bam*HI fragment 9 to the left of this site also retains a functional *loxP* site. The plasmid produced by this deletion event is called pRH38. In contrast, *loxP* can be inactivated by a deletion that removes most of the 84 bp between the *Bcl* I and the *Pvu* II sites (pRH31-39) or by deletions that remove sequences that include either the left (pRH31-56-2) or the right 13-bp inverted repeat (pRH31-57) (Fig. 4).

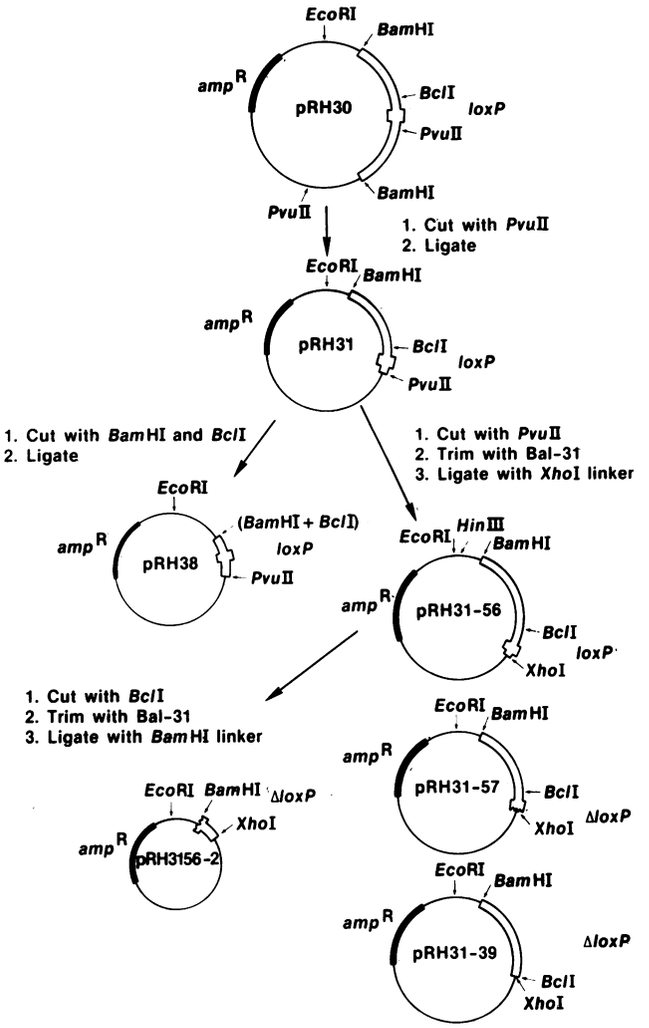

FIG. 3. Construction of plasmids containing shortened *loxP* sites. Plasmid pRH30 containing the 1.5-kilobase *Bam*HI fragment 9 was digested with *Pvu* II and religated under conditions that favored loss of the sequences between the two *Pvu* II sites. The resulting plasmid, pRH31, was restricted with *Pvu* II, trimmed with BAL-31, ligated with *Xho* I linkers, and reclosed. This gave plasmids pRH31-56, pRH31-57, and pRH31-39. Also, plasmid pRH31 was digested with *Bam*HI/*Bcl* I and ligated so as to delete the intervening sequences between these sites, giving pRH38. Plasmid pRH31-56-2 was derived by restricting pRH31-56 with *Bcl* I, trimming with BAL-31, and ligating with *Bam*HI linker. Prior to closure of the plasmid by ligation, the DNA was digested with *Bam*HI to remove all *Bam*HI fragment 9 sequences to the left of the *Bcl* I site. DNA linkers were obtained from Collaborative Research, Waltham, MA.

DISCUSSION

The existence of a site-specific recombination system in bacteriophage P1 has previously been demonstrated by genetic analysis (2, 4, 5). We report here the nucleotide sequence of four *lox* sites, *loxP*, *loxL*, *loxR*, and *loxB*, that are used as substrates for this recombinational process.

The most striking feature of the *loxP* site is a perfect 13-bp inverted repeat whose axis of dyad symmetry falls within an 8-bp region separating the repeats (Fig. 2). It is tempting to speculate that the inverted repeats are functionally important because they flank the *loxP* × *loxB* crossover point and because BAL-31-generated deletions that remove either the left or the right repeat sequence inactivate *loxP*. However, the location of the crossover point is only suggestive and conclusions based on the deletions are not definitive because the deletions remove not only the repeat sequence but also the crossover point. More conclusive are newly isolated BAL-31-generated deletions that remove only part of one of the inverted repeats and that inactivate *loxP* (data not shown). We suggest two functions for a region of dyad symmetry in recombination. (*i*) A rearrangement of base pairing within this region could produce cruciform structures in which the 8-bp region between the 13-bp inverted repeats exists as a single-stranded loop. That loop could be the substrate for cutting by the *cre* recombinase. A model for re-

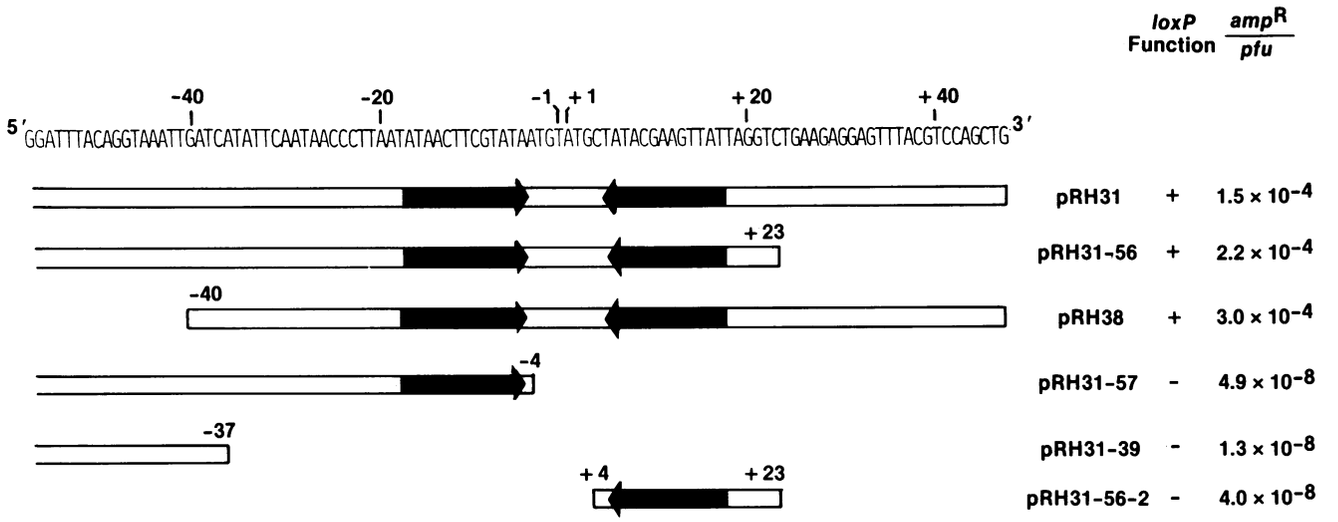

FIG. 4. Deletion analysis of *loxP* site. □, Extent of the *loxP* site preserved in the plasmid; ➡, inverted repeat. Deletion endpoints, which have been verified by DNA sequence analysis, are given as numbers corresponding to the given base in the *loxP* sequence. *loxP* function as determined by the ability of an incoming λ-P1:7 phage to transduce the plasmid is expressed as ampicillin-resistant (amp^R) transductants per plaque-forming unit (pfu). The yield of λ phages obtained in preparing the various transducing lysates was 10–25 pfu per infected cell. Control experiments with pBR322 give a frequency of 1.2×10^{-8} amp^R colonies per pfu.

combination that invokes cruciform structures as intermediates has been proposed previously (13). Support for the existence of such structures comes from recent results showing that regions of dyad symmetry similar to that at *loxP* can produce single-stranded regions that are sensitive to nuclease S1 (14, 15). (*ii*) The region of dyad symmetry may simply serve as a binding site for a recombinase protein that must bind and act symmetrically.

An inverted repeat sequence, albeit smaller (10 bp) than that at *loxP*, is also found in *loxB*. In this case, it also surrounds the crossover point. The sequences comprising this repeat share considerable homology (8 bp for each repeat) with the *loxP* site inverted repeats. The symmetry exhibited by *loxB* could account for the previous observation that *loxP* can integrate into *loxB* in either orientation (5). Because of the homology between the inverted repeats of *loxB* and *loxP*, the hybrid sites *loxR* and *loxL* also retain a portion of the inverted repeat structure. In both *loxL* and *loxR* sites, a complementary set of inverted repeats of 8 or 9 bp, respectively, with a single mismatch can be found (Fig. 2). In spite of the similarity in sequence between *loxP*, *loxB*, and the two hybrid sites, we know that *loxP* and *loxL* work efficiently in *cre*-mediated recombination whereas *loxB* and *loxR* do not (5). A possible explanation for this difference is that *loxB* and *loxR* sites have 9 bp between their inverted repeats rather than the 8 bp found in *loxP* and *loxL* sites. Alternatively, additional sequences other than the inverted repeat structure are necessary for efficient recombination at a *lox* site and these sequences reside in both *loxP* and *loxL*.

Since the *loxP* × *loxB* reaction is very inefficient, we were concerned that the crossover point as defined by this reaction might not be the same as that used in recombination between two *loxP* sites. By first cloning *loxP* onto a plasmid and then generating a number of *in vitro* deletions to either side of the inverted repeats, we have demonstrated that (*i*) all of the *cre* recognition site(s) must be located in a 63-bp sequence (+23 to −40; Fig. 4) that flanks the region of dyad symmetry and (*ii*) if the *loxP* × *loxP* crossover point occurs in P1 sequences, it could be the same or very close to the one seen in *loxP* × *loxB* recombination. Perhaps more importantly, we have the first indication that a functional *lox* site may be quite small, perhaps fewer than 50 bp. The functional site of pRH38 is 84 bp long. Based on the ability of pRH31-56 to recombine efficiently, we assume that an additional 21 bp can be deleted from the region to the right of the inverted repeats without loss of *loxP* function. However, we do not know whether, in fact, two small sites, such as the one in pRH38, can recombine efficiently with each other. There is the possibility that, as in the case of the phage λ *att* site, one of the sites in each reaction must be a large or dominant site whereas the other can be a small or passive site (16, 17). Since our *loxP* assay tests for function by infecting with a λ-P1:7 phage, which certainly contains enough P1 DNA to encompass a theoretically large *loxP* site, we cannot say for certain whether the site on pRH38 truly represents a fully intact *loxP* site or is simply a passive but efficient *lox* site.

We have implied here that the crossover point in *loxP* × *loxP* and in *loxP* × *loxB* recombination is unique and is limited to one or possibly a few adjacent internucleotide bonds. Although our analysis of the products of two different *loxP* × *loxB* recombination events (the *loxL* and *loxR* set described here) indicates that this may be true for *loxP* × *loxB* recombination, there is no compelling reason to believe that it is also true for *loxP* × *loxP* recombination. In fact, for λ *att* recombination, the crossover point is a region of 7 base pairs (18) and, for *loxP* × *loxP* recombination, at least some recombination events can involve crossovers located as much as 1 kilobase from the *loxP* sequence shown in Fig. 2 (19).

In summary, we have described results that begin to specify which DNA sequences are needed for *lox* recombination and that implicate a region of dyad symmetry as being important for recombination. More extensive analysis of mutations that alter *loxP* function and of recombination *in vitro* is needed to further elucidate the mechanism by which two *loxP* sites are cut and ligated.

We thank K. Abremski, S. Austin, M. Yarmolinsky, and H. Nash for helpful discussion during the course of this work. We also thank C. Benyajati and S. Austin for critical appraisal of this manuscript. This research was sponsored by the National Cancer Institute under Contract N01-CO-75380 with Litton Bionetics, Inc.

1. Ikeda, H. & Tomizawa, J. I. (1968) *Cold Spring Harbor Symp. Quant. Biol.* **33**, 791–798.
2. Sternberg, N. (1978) *Cold Spring Harbor Symp. Quant. Biol.* **43**, 1143–1146.

3. Austin, S., Ziese, M. & Sternberg, N. (1981) *Cell* **25,** 729–736.
4. Sternberg, N. & Hamilton, D. (1981) *J. Mol. Biol.* **150,** 467–486.
5. Sternberg, N., Hamilton, D. & Hoess, R. (1981) *J. Mol. Biol.* **150,** 487–507.
6. Zissler, J., Signer, E. & Schaefer, F. (1971) in *The Bacteriophage Lambda* eds., Bukhari, A. I., Shapiro, J. & Adhya, S. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 455–468.
7. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L. & Boyer, H. W. (1977) *Gene* **2,** 95–113.
8. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
9. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
10. Mizuuchi, K. & Mizuuchi, M. (1978) *Cold Spring Harbor Symp. Quant. Biol.* **43,** 1111–1114.
11. Legerski, R. J., Hodnett, J. L. & Gray, H. B., Jr. (1978) *Nucleic Acids Res.* **5,** 1445–1464.
12. Landy, A. & Ross, W. (1977) *Science* **197,** 1147–1160.
13. Sobell, H. M. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 2483–2487.
14. Lilley, D. M. J. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 6468–6472.
15. Panayotatos, N. & Wells, R. D. (1981) *Nature (London)* **289,** 466–470.
16. Mizuuchi, M. & Mizuuchi, K. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 3220–3224.
17. Hsu, P.-L., Ross, W. & Landy, A. (1980) *Nature (London)* **285,** 85–91.
18. Mizuuchi, K., Weisberg, R., Enquist, L., Mizuuchi, M., Buraczynska, M., Foeller, C., Hsu, P.-L., Ross, W. & Landy, A. (1980) *Cold Spring Harbor Symp. Quant. Biol.* **45,** 429–437.
19. Sternberg, N. (1981) *J. Mol. Biol.* **150,** 603–608.