Mutations That Affect TnS Insertion into pBR322: Importance of Local DNA Supercoiling

JENNIFER K. LODGE[†] AND DOUGLAS E. BERG^{*}

Departments of Molecular Microbiology and of Genetics, Box 8230, Washington University Medical School, St. Louis, Missouri 63110

Received ¹⁸ May 1990/Accepted ⁸ August 1990

The major hot spot of transposon TnS insertion in plasmid pBR322 (hot spot I) is in the promoter for the tetracycline resistance gene (tet). We made a series of $pBR322$ derivatives with mutations in and around this promoter and assayed their effects on insertion of TnS into hot spot I. Those mutations which reduced transcription from the tet promoter also reduced the frequency of insertion into hot spot I. Transcription and translation of tet are thought to cause the formation of paired domains of negative and positive supercoiling in pBR322. An amber codon in tet, ³⁴⁵ base pairs from hot spot I, decreases the negative supercoiling of the DNA segment containing hot spot I because it terminates translation of tet prematurely. We report here that this amber mutation also reduces insertion into hot spot I. These results suggest that the ability of Tn5 to insert into its major hot spot in pBR322 depends directly on negative supercoiling of the target DNA.

The bacterial transposon Tn5 inserts into many sites within a single gene but shows a preference for certain sites. Five hot spots of Tn5 insertion were identified in the tet gene of plasmid pBR322 (2, 16). Of 150 independent transpositions to tet, 40 were in a single site, hot spot I, and 15 were in a second site, hot spot II. The only obvious sequence features common to TnS insertion hot spots are GC base pairs at each end of the 9 base pairs (bp) duplicated by TnS insertion. Mutation of these GC pairs to AT caused ^a sharp reduction of insertion into hot spot ^I (15). However, there are hundreds of GC pairs spaced ⁹ bp apart which are not hot spots, and therefore other features must also help guide Tn5 to its preferred insertion sites.

TnS insertion hot spots do not match any obvious consensus sequence, either in the internal base pairs, which are equivalent to the 6-bp consensus sequence for $Tn/0$ (9), or in the flanking sequences, such as those for IS4 (12); nor are there known protein binding sites adjacent to the hot spots equivalent to integration host factor binding sites found adjacent to an IS1 hot spot (8). Possibly, Tn5 prefers to insert into transcriptionally active DNA or into DNA with ^a particular topology.

The 9 bp duplicated by TnS insertion at hot spot ^I includes the -10 region of the *tet* promoter (P_{te}) and the transcription start site for the divergent *antitet* promoter (P_{antite}) (for a review, see reference 1) (Fig. 1). To determine whether the activities of these promoters affect TnS insertion into hot spot I, a series of deletion and insertion mutations in and near the promoters were made and the effects of these mutations on TnS insertion into hot spot ^I were assessed. The results presented here suggest that transcription or coupled transcription and translation from the tet promoter is important in hot spot ^I selection.

Negative supercoiling of the target DNA is needed for efficient TnS transposition (10), which implies that Tn5 insertion specificity is affected by DNA superhelical density. Findings that pBR322 and many of its derivatives are highly negatively supercoiled in cells lacking topoisomerase ^I led to a model in which the superhelical density along a single DNA molecule is not uniform because paired domains of negative and positive supercoiling are transiently created by transcription (13, 22, 25). Furthermore, translation of the membrane protein encoded by tet, in the case of pBR322, helps maintain these paired domains by impeding the equilibration of supercoils by preventing free rotation of the DNA (14). Hot spot ^I is in the proposed negatively supercoiled domain (13). An amber codon in tet, 345 bp distal to hot spot I, eliminates hypernegative supercoiling because it stops translation of the TetA protein (14). The results presented here show that this amber mutation reduced insertion of Tn5 into hot spot ^I and suggest that the preference for hot spot ^I reflects, in part, its inclusion in this negatively supercoiled domain.

MATERIALS AND METHODS

Strains. The bacterial strains used are derivatives of Escherichia coli K-12. DB1572 (15) is lacZ124::TnS and $sup⁺$. MC1061 (5) is $sup⁰$ and was obtained from Henry Huang.

General procedures. Standard techniques were used for bacterial transformation, plasmid isolation, and restriction mapping (18). Expression of *tet* was crudely measured on the basis of plating efficiency and colony size on LN agar plates (15) containing 12.5 μ g of tetracycline. DNA sequencing was performed by using either the Maxam-Gilbert (19) or dideoxy (23) method as modified for double-stranded DNA (26). Oligonucleotides were made on an Applied Biosystems model 380A DNA synthesizer.

Construction of mutant plasmids. pBR322 derivatives were constructed in vitro. Schematic diagrams and the details of the construction for most of these derivatives are shown in Fig. 2. Plasmid pBX has an in-frame amber (UAG) codon 345 bp from hot spot I, due to insertion of an XbaI linker (CTCTAGAG) in the filled-in BamHI site of pBR322, which also recreates the BamHI site.

Isolation and mapping of TnS insertions. Independent TnS transpositions from a chromosomal $lacZ$ gene to the tet gene in pBR322 derivatives were obtained as described previously (15). Insertions of $Tn5$ in the *tet* gene (Kan^r Tet^s phenotype) were first mapped by digestion with restriction enzymes

^{*} Corresponding author.

^t Present address: Monsanto Company, St. Louis, MO 63198.

FIG. 1. Sequence around hot spot I in pBR322. The pBR322 sequence from the EcoRI site at bp 0 to bp 80 is shown (24). Hot spot I (bp 31 to 39) and hot spot II (bp 72 to 80) are represented with larger letters. The *tet* and *antitet* promoters are shown (4). The -35 and -10 regions of P_{tet} are designated with overbars, and tet transcription starts at the G at bp 45. The -35 and -10 regions of $P_{antiter}$ are underscored, and antitet transcription begins at either of two T's at bp 36 and 37. The EcoRI and HindIII sites are indicated by the bars between the DNA strands.

EcoRI (which cleaves once in pBR322, 33 bp from hot spot I) and XhoI (which cleaves 486 bp from the ends of TnS) and by agarose gel electrophoresis. Inserts at or near hot spot ^I were identified on the basis of a characteristic \sim 525-bp band on an agarose gel and were then analyzed with single-basepair resolution by end labeling the plasmid DNA at either the ClaI or EcoRI site, subcutting with Hinfl (which cleaves ⁵ bp from the ends of TnS), and polyacrylamide-urea gel electrophoresis.

Tn5 transpositions from bacteriophage λ to pBR322 and pBX were obtained by infecting MC1061 carrying the plasmid with λ c1857 b221 Oam29 Pam80 rex:: Tn5 (21) at a multiplicity of infection of 4 bacteriophages per cell. The cells were grown for 3 h at 37°C after infection and spread on LN agar plates containing 60μ g of kanamycin per ml. After overnight growth, each plate was scraped, plasmid DNA was isolated and transformed into MC1061, and Amp^r Kan^r transformants were selected. Plasmid DNA was isolated, and insertions within the 375-bp EcoRI-BamHI fragment (which contains hot spots ^I and II) were identified by cleavage with EcoRI and BamHI and agarose gel electrophoresis. Tn5 inserts in hot spots ^I and II were then identified by using EcoRI and XhoI digestions and endlabeled EcoRI-Hinfl fragments as described above. To ensure independence, only one plasmid with an insertion in the EcoRI-BamHI fragment was examined from each plate.

RESULTS

Hot spot ^I is recognized in another plasmid. To determine whether the use of hot spot ^I depends on a particular overall arrangement of plasmid sequences, we mapped TnS insertions in pACYC184 (6), a plasmid which has the same tet gene as pBR322 and which is also hypernegatively supercoiled in topA strains (14) but is otherwise unrelated. Six of 35 Tets pACYC184::TnS plasmids contained TnS at hot spot I; the others contained Tn5 at many other sites. This suggests that hot spot ^I use is determined by sequences within the *tet* gene or its promoter.

Importance of sequences within the 9 bp duplicated by insertion. The internal 7 bp of the 9-bp target duplications differ among hot spots (2), and we tested whether these sequences contributed to insertion into hot spot ^I by replacing the central 7 bp of hot spot ^I with the corresponding 7 bp from hot spot II (plactet-3 and plactet-4) in each orientation. These substitutions eliminated *tet* promoter function, and so a lac promoter was added in the EcoRI site to restore transcription of tet and thereby allow tet::Tn5 inserts to be identified on the basis of a Tet^s phenotype. These plasmids expressed tet at levels similar to those expressed by pBR322.

We found that Tn5 did not insert into either hybrid hot spot (Fig. 3). Thus, the use of hot spot ^I depends on sequences within the 9 bp duplicated by Tn5 insertion.

Mutations ⁵' to the tet promoter. The plasmids ptet-1 and pEX have sequence changes immediately adjacent to and ¹¹ bp ⁵' to the tet promoter, respectively, which did not affect tet expression and did not affect insertion into hot spot ^I (Fig. 3). Plasmids plactet-1 and placE were derived from ptet-1 and pEX, respectively, by insertion of a lac promoter 5' to tet, and tests with these plasmids showed that transcription through hot spot ^I had little, if any, effect (less than twofold) on insertion into hot spot ^I (Fig. 3). Therefore, sequence changes which do not affect the tet promoter did not have large effects on insertion into hot spot I.

Mutations in the tet promoter. Deletion of all (placEH) or part (plactet-2) of the -35 region of the *tet* promoter or an increase in the spacing between the -35 and the -10 regions (placHX) reduced the activity of the tet promoter, but tet expression was restored to near normal levels in placEH and placHX by addition of a lac promoter. tet expression was reduced in plactet-2 even though the same lac promoter was used. The frequency of insertions into hot spot ^I relative to all insertions in tet, was reduced at least fivefold by these sequence changes (Fig. 3). These data show that mutations that affected P_{tet} also affected insertion into hot spot I.

Mutations in P_{antitet}. To determine whether the antitet promoter (Fig. 1) was important for insertion into hot spot I, two plasmids with sequence changes ³' to hot spot ^I were tested. Plasmid pl(39) has a 13-bp insertion between the tet and antitet promoters that should change the start site of transcription from both promoters, and pI(39-49) has a substitution that removes the -10 region of P_{antite} and alters the start site of tet transcription. These changes caused a small, if any, effect (less than twofold) on the frequencies of insertion into hot spot ^I (Fig. 3). Thus, transcription from $P_{antiter}$ is not likely to be a major factor in Tn5 insertion into hot spot I.

A second copy of hot spot I. Plasmid pI(39)-HSI contains ^a second copy of hot spot ^I at the end of the inserted sequence in pl(39). The tet promoter is intact in this plasmid and overlaps the first copy of hot spot ^I (as in wild-type pBR322); the second copy is not part of a promoter. TnS inserted efficiently into both copies of hot spot ^I (Fig. 3), showing that hot spot ^I can function well even if it is not part of a promoter.

Effects of ^a small distant mutation which affects DNA supercoiling. Plasmid $pBX(14)$ (Fig. 4) has an amber codon inserted into the BamHI site of pBR322, 345 bp from hot spot I. TnS insertion specificity in this amber-mutant plasmid

J. BACTERIOL.

FIG. 2. Diagram of mutant plasmids. The pBR322 sequence extending from the Ec_0 RI site to position 70, which includes hot spot I, P_{ref} , and P_{antiter}, is shown schematically. E, *Eco*RI (bp 4361); H, *HindIII* (bp 29); X, *XhoI*; P_{lac}, 95-bp AluI fragment containing the lacUV5
promoter (3); insert, 5'ATTGGAGGTACCC; XXXXX, 5'TGCCTGA or 5'TCAGGCA. The -35 promoters are represented as boxes. Hot spot I is part of the -10 region of P_{ref} plactet-3 and plactet-4 have a lac promoter inserted in the EcoRI site, and the internal ⁷ bp of hot spot ^I (pBR322 positions ³² to 38) is replaced with the sequences 5'TGCCTGA and 5'TCAGGCA, respectively (internal base pair of hot spot II, both orientations made by oligomutagenesis, as in reference 17). pEX was made by inserting an XhoI linker (CCTCGAGG) into the filled-in EcoRI site of pBR322. placE was made by insertion of the lac promoter into the EcoRI site of pBR322. ptet-1 has a deletion between the EcoRI site and bp 9 of pBR322 and was created by Bal 31 digestion followed by insertion of an XhoI linker. plactet-1 has a lac promoter inserted between the EcoRI and XhoI sites of ptet-1. plactet-2 has a deletion created by Bal 31 digestion between the EcoRI site and bp 11 of pBR322, combined with insertion of a *lac* promoter. placHX has an XhoI linker inserted into the filled-in HindIII site of placE. placEH has a lac promoter inserted between the Ec oRI and HindIII sites of pBR322. pI(39) has an insertion of ¹³ bp (5' ATTGGAGGTACCC) between bp ³⁹ and bp ⁴⁰ of pBR322 and was created by ^a combination of Bal ³¹ digestion and oligonucleotide mutagenesis. pI(39-49) has the same sequence inserted as pI(39), but between bp 39 and bp 49. pI(39)-HSI has an insertion
of 5'AATTGGAGGTCACCAGCTTTAAG, which contains a copy of hot spot I, between bp 39 and mutagenesis.

FIG. 3. Frequency of Tn5 insertion into hot spot ^I of pBR322 derivatives. Each bar is labeled with the name of the plasmid (below) and the number of insertions into hot spot ^I relative to the total number of tet::Tn5 isolates (above). The bar height represents this fraction as a percent. There are two copies of hot spot ^I in pI(39)-HSI, and they are reported as copy ¹ (copy which is in normal position relative to the tet promoter) and copy 2 (inserted copy which is not part of the tet promoter).

was analyzed by using transposition from λ in the sup⁰ strain MC1061, because DB1572, the host for other transpositions, contains an amber suppressor. Because the amber mutation makes pBX Tet^s, restriction mapping was used to identify TnS inserts in the 375-bp EcoRI-BamHI fragment. None of ²¹ inserts into this fragment of pBX was at hot spot I, whereas 10 of 22 inserts into the corresponding EcoRI-BamHI fragment were at hot spot ^I of the pBR322 wild-type control. Thus, a small insertion, 345 bp distal to hot spot I, which affects plasmid supercoiling also prevents insertion into hot spot I.

DISCUSSION

The broad distribution of Tn5 insertion sites and the absence of a strong consensus sequence, even among insertion hot spots (2, 16), indicate that target specificity is complex and probably affected by several features of target DNAs. On the basis of the present results, we propose that local high negative supercoiling of target DNA contributes strongly to Tn5 insertion into hot spot I, the most preferred site in plasmid pBR322. The most compelling evidence is that an amber mutation in tet, which affects plasmid supercoiling, also dramatically reduced insertion into hot spot I.

Recent studies with cells lacking DNA topoisomerase ^I and with cells treated with gyrase inhibitors have indicated that transcription and translation of the tet gene creates paired domains of positive and negative supercoiling in plasmid pBR322 (13, 14, 22, 25). Although these domains of supercoiling have not been physically mapped, they are thought to exist also in wild-type cells because very strong transcription alters plasmid supercoiling (7) and the supercoiling-induced transition of the DNA from the B to the Z form DNA in vivo depends on where in the plasmid the Z-forming DNA segment is located (11). Hot spot ^I is in the

FIG. 4. Tn5 insertion into hot spots ^I and II of pBR322 and pBX. (a) Schematic representation of plasmids pBR322 and pBX. The EcoRI-to-BamHI region of pBR322 is shown. The positions of hot spots ^I and II are indicated with boxes. pBX contains an insertion of an 8-bp sequence (XbaI linker) which recreates the BamHI site and contains an amber codon which is in frame for the tet gene. (b) Frequency of Tn5 insertion into hot spots ^I and II of pBR322 and pBX. Each bar is labeled with the name of the plasmid (below) and the number of tet::Tn5 insertions into the hot spot relative to the number of insertions into the EcoRI-BamHI fragment (above). The bar height represents this fraction as a percent.

plasmid segment of pBR322 that these experiments suggest is highly negatively supercoiled.

The clearest evidence for a causal relationship between highly negative supercoiling and insertion into hot spot ^I is shown by tests with plasmid pBX. Its tet amber mutation reduced the normal highly negative supercoiling of pBR322 in cells lacking DNA topoisomerase ^I (14), and the present experiments showed that this amber mutation also blocked TnS insertion into hot spot I, even though it is more than 300 bp from hot spot I. Furthermore, TnS insertion is reduced in cells which have been treated with gyrase inhibitors, suggesting that TnS requires a negatively supercoiled target (10). Finally, mutations that decreased transcription from the tet promoter and decreased insertion into hot spot ^I also reduced negative supercoiling in cells lacking DNA topoisomerase ^I (J. K. Lodge, Ph.D. thesis, Washington University, St. Louis, Mo., 1989).

The 9-bp segment duplicated by insertion at hot spot ^I includes the -10 region of the *tet* promoter (4). Mutations which decreased expression from the tet promoter (including substitutions consisting of the central 7 bp of hot spot II) also decreased insertion into hot spot I, suggesting that insertion into hot spot I may depend on use of this site as a -10 promoter region. However, when a second copy of hot spot ^I was present a short distance from the intact tet promoter but was not part of the promoter, this second copy was also recognized as a hot spot. These results indicate that although transcription from tet is important, the targeted site need not overlap a promoter for it to be used preferentially.

High negative supercoiling is only one of several factors contributing to TnS insertion specificity. Tests with base substitutions at each end of hot spot ^I showed that the GC pairs at these positions are also important for Tn5 insertion at that hot spot and that their importance does not stem from an effect on plasmid superhelical density (15). Additionally, placement of a *lac* promoter upstream from the *tet* promoter reduced insertion into hot spot ^I by about twofold. Transcription through hot spot ^I might disrupt the transposition complex or perturb the high negative supercoiling at hot spot I. In contrast, hot spot II, which is 41 bp downstream from hot spot I, is normally transcribed from the tet promoter, and separate studies have suggested that insertion into certain other preferred regions actually depends on their being transcribed from an upstream promoter (20). Finally, insertion into hot spot II was not affected by the amber mutation in plasmid pBX. This implies that the use of hot spot II, unlike the use of hot spot I, does not depend on high negative supercoiling.

In summary, the specificity of TnS insertion appears to be quite complex. Several determinants, including DNA supercoiling, GC pairs at the ends of preferred sites, and probably other aspects of the DNA sequence as well, contribute quantitatively to insertion into various preferred sites. The particular features used at one site differ from those used at other sites, and the tens or hundreds of sites in any typical gene that are used at lower frequencies may also be part of this continuum, reflecting positions where the constellation of potentiating features is less optimal. For very frequent insertion into hot spot I, high negative supercoiling may facilitate local DNA melting, which in turn might increase recognition or cleavage of target sequences during transposition.

ACKNOWLEDGMENTS

We are grateful to K. Dodson and K. Weston-Hafer for stimulating discussions.

This work was supported by a grant from the Lucille P. Markey Charitable Trust, by Public Health Service grant GM-37138 from the National Institutes of Health, by grant DMB-8608193 from the National Science Foundation, and by grant DE-FG02-89ER-60862 to D.E.B. J.K.L. was supported by a Graduate and Professional Opportunities Program fellowship.

LITERATURE CITED

- 1. Balbas, P., X. Soberon, E. Merino, M. Zurita, H. Lomeli, F. Valle, N. Flores, and F. Bolivar. 1986. Plasmid vector pBR322 and its special-purpose derivatives-a review. Gene 50:3-40.
- 2. Berg, D. E., M. A. Schmandt, and J. B. Lowe. 1983. Specificity of transposon Tn5 insertion. Genetics 105:813-828.
- 3. Bogosian, G., and J. F. Kane. 1987. Nucleotide sequence of the EcoRI fragment from pLJ3 bearing two tandem lac promoters. Nucleic Acids Res. 15:7185.
- 4. Brosius, J., R. L. Cate, and A. P. Perlmutter. 1982. Precise location of two promoters for the beta-lactamase gene on pBR322. J. Biol. Chem. 257:9205-9210.
- 5. Casadaban, M., and S. Cohen. 1980. Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. J. Mol. Biol. 138:179-207.
- 6. Chang, A. C. Y., and S. N. Cohen. 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P1SA cryptic miniplasmid. J. Bacteriol. 134: 1141-1156.
- 7. Figueroa, N., and L. Bossi. 1988. Transcription induces gyration of the DNA template in Escherichia coli. Proc. Natl. Acad. Sci. USA 85:9416-9420.
- 8. Gamas, P., M. G. Chandler, P. Prentki, and D. J. Galas. 1987. Escherichia coli integration host factor binds specifically to the ends of the insertion sequence IS1 and to its major insertion hot-spot in pBR322. J. Mol. Biol. 195:261-272.
- Halling, S. M., and N. Kleckner. 1982. A symmetrical six-base target site sequence determines TnlO insertion specificity. Cell 28:155-163.
- 10. Isberg, R. R., and M. Syvanen. 1982. DNA gyrase is ^a host factor required for transposition of Tn5. Cell 30:9-18.
- 11. Jaworski, A., J. A. Blaho, J. E. Larson, M. Shimizu, and R. D. Wells. 1989. Tetracycline promoter mutations decrease non-B DNA structural transition, negative linking differences and deletions in recombinant plasmids in Escherichia coli. J. Mol. Biol. 207:513-526.
- 12. Klaer, R., S. Kuhn, H. Fritz, E. Tillman, I. Saint-Girons, P. Haberman, D. Pfeifer, and P. Starlinger. 1980. Studies on transposition mechanisms and specificity of IS4. Cold Spring Harbor Symp. Quant. Biol. 45:215-224.
- 13. Liu, L. F., and J. C. Wang. 1987. Supercoiling of the DNA template during transcription. Proc. Natl. Acad. Sci. USA 84:7024-7027.
- 14. Lodge, J. K., T. Kazic, and D. E. Berg. 1989. Formation of supercoiling domains in plasmid pBR322. J. Bacteriol. 171: 2181-2187.
- 15. Lodge, J. K., K. Weston-Hafer, and D. E. Berg. 1988. Transposon Tn5 target specificity: preference for G/C pairs. Genetics 120:645-650.
- 16. Lodge, J. K., K. Weston-Hafer, J. B. Lowe, K. W. Dodson, and D. E. Berg. 1985. Determinants of TnS insertion specificity, p. 37-42. In M. Simon and I. Herskowitz (ed.), Genome rearrangement. Alan R. Liss, New York.
- 17. Mandecki, W. 1986. Oligonucleotide-directed double-strand break repair in plasmids of Escherichia coli: a method for site-directed mutagenesis. Proc. Natl. Acad. Sci. USA 83:7177- 7181.
- 18. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- 19. Maxam, A., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. 65: 499-580.
- 20. McKinnon, R. D., J. S. Waye, D. S. Bautista, and F. L. Graham. 1985. Nonrandom insertion of TnS into cloned human adenovirus DNA. Gene 40:31-38.
- 21. Phadnis, S. H., and D. E. Berg. 1987. Identification of base pairs in the IS50 0 end needed for IS50 and TnS transposition. Proc. Natl. Acad. Sci. USA 84:9118-9122.
- 22. Pruss, G. J., and K. Drlica. 1986. Topoisomerase ^I mutants: the gene on pBR322 that encodes resistance to tetracycline affects plasmid DNA supercoiling. Proc. Natl. Acad. Sci. USA 83: 8952-8956.
- 23. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
- 24. Sutcliffe, J. G. 1978. Complete nucleotide sequence of the Escherichia coli plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. 42:77-90.
- 25. Wu, H., S. Shyy, J. C. Wang, and L. F. Liu. 1988. Transcription generates positively and negatively supercoiled domains in the template. Cell 53:433-440.
- 26. Zagurski, R. J., K. Baumeister, N. Lomax, and M. L. Berman. 1985. Rapid and easy sequencing of large linear double-stranded DNA and supercoiled plasmid DNA. Gene Anal. Technol. 2:89-94.