JOURNAL OF BACTERIOLOGY, Jan. 1988, p. 352–358
0021-9193/88/010352-07$02.00/0

Vol. 170, No. 1

Sequence Requirements of *Escherichia coli attTn7*, a Specific Site of Transposon Tn7 Insertion

ROBERT L. McKOWN,[1,2] KARINA A. ORLE,[1,2] THOMAS CHEN,[1,2] AND NANCY L. CRAIG[1,2,3]*

Departments of Microbiology and Immunology[1] and Biochemistry and Biophysics,[3] and G. W. Hooper Foundation,[2] University of California at San Francisco, San Francisco, California 94143

Received 30 September 1987/Accepted 8 October 1987

Transposon Tn7 transposes at high frequency to a specific site, *att*Tn7, in the *Escherichia coli* chromosome. We devised a quantitative assay for Tn7 transposition in which Tn7-end derivatives containing the *cis*-acting transposition sequences of Tn7 transpose from a bacteriophage lambda vector upon infection into cells containing the Tn7-encoded transposition proteins. We used this assay to identify a 68-base-pair DNA segment containing the sequences essential for *att*Tn7 target activity. This segment is positioned asymmetrically with respect to the specific point of Tn7 insertion in *att*Tn7 and lacks obvious homology to the sequences at the ends of Tn7 which participate directly in transposition. We also show that some sequences essential for *att*Tn7 target activity are contained within the protein-coding sequence of a bacterial gene.

Transposons are precise DNA segments which can translocate from place to place in the genome. Most procaryotic elements transpose at low frequency and show little target site specificity upon insertion into large DNA molecules such as bacterial chromosomes (for reviews, see references 22 and 41). Tn7 is a transposon which provides resistance to trimethoprim and to streptomycin and spectinomycin (3) (Fig. 1A). Tn7 is of particular interest because it can transpose at high frequency to a specific target site. The capacity of Tn7 for high-frequency transposition to the *Escherichia coli* chromosome was revealed by the finding that after incompatibility exclusion of a plasmid containing Tn7 in the absence of selection for Tn7, 5 to 100% of the chromosomes contain Tn7 (2, 3, 18, 24). Similarly, we have observed that after temperature exclusion of F'ts *lac*$^+$::Tn7, about 5 to 10% of the plasmid-free cells contain chromosomal Tn7 (C. Waddell and N. L. Craig, unpublished observation). When Tn7 transposes to the *E. coli* chromosome, it inserts in a specific site at about min 84 called *att*Tn7 (3, 24, 25); this chromosomal site is referred to hereafter as *att*Tn7$_{84}$. The specific point of Tn7 insertion lies between *phoS*, a gene involved in phosphate transport, and *glmS*, a gene involved in cell wall biosynthesis (25, 47) (Fig. 1B). In *att*Tn7$_{84}$::Tn7, the left end of Tn7 (Tn7L), which encodes its drug resistance determinants, is adjacent to *phoS*, and the right end of Tn7 (Tn7R), which encodes its transposition proteins, is adjacent to *glmS* (24). Thus, Tn7 insertion into the *E. coli* chromosome is both site and orientation specific. Site-specific insertion of Tn7 into the chromosomes of other bacteria has also been observed; the organisms include *Agrobacterium tumefaciens* (19), *Vibrio* species (45), *Caulobacter crescentus* (10), *Pseudomonas aeruginosa* (6), *Rhodopseudomonas capsulata* (49), *Rhizobium meliloti* (5), *Xanthomonas campestris* pv. *campestris* (46), and *Pseudomonas fluorescens* (1).

It has previously been shown that when large *E. coli* chromosomal fragments containing the specific insertion point are introduced into plasmids, Tn7 inserts at high frequency into these DNA fragments in a site- and orientation-specific fashion (24, 36), demonstrating that *att*Tn7 activity is determined by sequences within these segments. In this work, we examined the functional extent of *att*Tn7 by comparing the target activities of DNA fragments containing the specific point of Tn7 insertion and various extents of flanking sequence. We demonstrated that a small (68-base-pair [bp]) DNA fragment positioned asymmetrically with respect to the specific point of Tn7 insertion has *att*Tn7 activity; i.e., it is a target for high-frequency, site-specific insertion of Tn7.

* Corresponding author.

MATERIALS AND METHODS

Media, chemicals, and enzymes. LB broth and agar were as described by Miller (30) except that 1 mg of glucosamine per ml was added to the agar. With trimethoprim, Iso-Sensitest agar (Oxoid Ltd.) was used. When used, supplements were as follows: carbenicillin, 100 μg/ml; chloramphenicol, 30 μg/ml; kanamycin, 50 or 100 μg/ml; tetracycline, 5 μg/ml; and trimethoprim, 100 μg/ml. DNA-modifying enzymes were obtained from commercial sources and used as recommended by the manufacturer.

Manipulation and analysis of DNA. Plasmid growth, isolation, transformation, and restriction enzyme analyses were performed as described by Maniatis et al. (27). Cloning techniques were performed as described by Maniatis et al. (27) except that usually DNA fragments contained in slices excised from low-melting-temperature agarose gels (SeaPlaque; FMC Corp., Marine Colloids Div.) were used directly in the assembly of recombinant molecules as described by Struhl (43). Bacteriophage lambda growth was performed as described by Maniatis et al. (27). DNA sequence analysis was performed as described by Sanger et al. (40), using double-stranded plasmid DNA as a template (7). To circumvent difficulties with sequencing the potential G+C-rich stem-loop structure in *att*Tn7, we included T4 gene 32 protein (B. Alberts, University of California, San Francisco) (33) in the sequencing reaction mixtures and used either Klenow fragment with deoxy-7-deazaguanosine triphosphate (American Bionetics) (32) or Sequenase (United States Biochemical Corp.) with dITP (31) in place of dGTP. In addition, we sequenced both strands of some novel junctions, included 25% formamide in the sequencing gels, and ran the gels at high temperature.

Strains. NLC51 carries F$^-$ *araD139* Δ(*argF-lac*)*U169* *rpsL150* *relA1* *flbB5301* *deoC1* *ptsF25* *rbsR* Val^r *recA56*

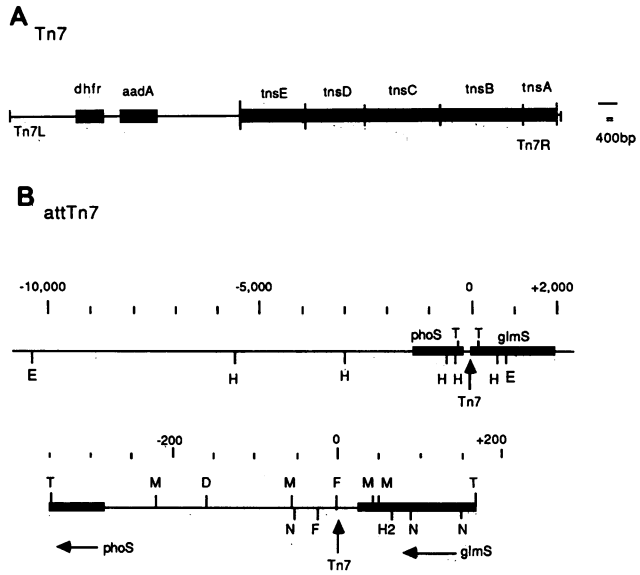

FIG. 1. Tn7 and *att*Tn7. (A) Tn7. The left end of Tn7 (Tn7L) encodes its drug resistance genes: *dhfr* (dihydrofolate reductase), which confers trimethoprim resistance (12), and *aadA* [3″(9)-*O*-nucleotidyltransferase], which confers streptomycin and spectinomycin resistance (11). The right end of Tn7 (Tn7R) encodes the Tn7 transposition genes: *tnsA tnsB*, *tnsC*, *tnsD*, and *tns*E (36; C. Waddell and N. L. Craig, Genes Dev., in press). (B) *att*$Tn7_{84}$ and an *att*Tn7-containing plasmid. (Top) Physical map of the *att*$Tn7_{84}$ region of *E. coli* chromosome as determined by Lichtenstein and Brenner (25), Walker et al. (47), Surin et al. (44), and Magota et al. (26). The Tn7 insertion point is indicated by a vertical arrow. The *att*Tn7 region is numbered as described in Results. The coding sequences for *phoS* and *glmS* are indicated by closed bars. Only the *Taq*I sites immediately flanking the specific insertion point are shown. (Bottom) *att*Tn7 segment of pRM2 (see Materials and Methods). The Tn7 insertion point is indicated by a vertical arrow. Portions of the coding sequences of *phoS* and *glmS* contained in this fragment are indicated, and the direction of transcription of these genes is also shown. Restriction endonuclease sites: D, *Dra*I; E, *Eco*RI; F, *Fok*I; H, *Hpa*I; H2, *Hin*cII; M, *Mae*III; N, *Nla*III; T, *Taq*I.

(29a). The Tn7 derivative used in this work, Tn7S, contains an IS*1* insertion near the drug resistance determinants of Tn7; however, the transposition properties of Tn7S are indistinguishable from those of canonical Tn7 (18). LA3 is NLC51 *att*$Tn7_{84}$::Tn7 (29a). NK5012 (from N. Kleckner, Harvard University) carries *thr leu thi*$^+$ *tonA lacY1 supE44* (strain is also known as C600).

Construction of lambda donor phages containing Tn7 end derivatives. The Tn7-end derivative Tn7S::Tn9Δ*Pst*I (18) (J. Shapiro, University of Chicago) contains a segment of the left end of Tn7 (from the terminus Tn7L1 to about Tn7L1900) and a segment of the right end of Tn7 (from the terminus Tn7R1 to Tn7R537) flanking a segment of Tn9 encoding chloramphenicol resistance. This Tn7-end derivative is hereafter referred to as Tn7-Cm. The Tn7-end derivative Tn7-Km, which contains Tn7L166 and Tn7R199 flanking a segment encoding kanamycin resistance, was constructed by in vivo transposition of Tn7S::Tn9Δ*Pst*I into pRM2 (see below), digestion of the resulting plasmid with *Hpa*I and *Bal*I, which cut in Tn7L and Tn7R, respectively (25), ligation in the presence of a *Sal*I linker, and subsequent introduction of a fragment, with terminal *Sal*I sites, encoding kanamycin resistance (Pharmacia). This Tn7-end derivative contains the necessary *cis*-acting sequences for Tn7 transposition (L. Arciszewska and N. L. Craig, manuscript in preparation). The *att*Tn7 segment of pRM4 (see below) lies within a transposition-defective, tetracycline-resistant Tn*10* derivative flanked by *hisOGD* DNA. Insertions of Tn7-Cm and Tn7-Km into pRM4 *att*Tn7 were isolated by transformation of plasmid DNA isolated from strains containing *att*$Tn7_{84}$::Tn7, pRM4, and pOX38::Tn7-end derivatives, selecting for Tn7-end-derivative drug resistance. The site- and orientation-specific insertion of the Tn7-end derivatives into the pRM4 *att*Tn7 site was verified by restriction enzyme digestion (data not shown). Derivatives of phage lambda 780 (*b2*::*hisOGD b522 c*I857 *P*am80 *nin5*) (13) that acquired the Tn*10*::*att*Tn7::Tn7-end-derivative segments by homologous recombination within *hisOGD* were obtained by isolating tetracycline-resistant phages from phage stocks grown in NK5012(pRM4::Tn7-end-derivative) strains by using the selective plaque assay (28). Acquisition of the Tn7-end derivatives by the phages was verified by demonstrating that lysogens of the phages also expressed Tn7-end-derivative drug resistance. Lambda RM7 is lambda 780 *hisG9424*::Tn*10* del16 del17::*att*Tn7(−342 to +165)::Tn7 L1900 Cm^r Tn7R537, and lambda KK1 is lambda 780 *hisG9424*::Tn*10* del16 del17::*att*Tn7(−342 to +165)::Tn7 L166 Km^r Tn7R199 (see Results for *att*Tn7 numbering).

Lambda hop assays. A culture of cells to be assayed was grown to about 4×10^8 cells per ml in LB broth plus 0.2% maltose, concentrated by centrifugation, and suspended at 1.6×10^9 cells per ml in 0.01 M $MgSO_4$. Phage was added to 0.1 ml of cells at a multiplicity of 0.1 per cell. The resulting 0.2-ml mixture was incubated at 37°C for 15 min, 0.8 ml of LB broth plus 0.01 M sodium citrate was added, and incubation was continued at 37°C with shaking for 60 min. Samples from the mixture (or dilutions thereof in LB broth plus 0.01 M sodium citrate) were spread on plates containing appropriate antibiotics (chloramphenicol or chloramphenicol-trimethoprim for Tn7-Cm or kanamycin or kanamycin-trimethoprim for Tn7-Km), and the plates were incubated at 37°C for 1 day before counts were made.

***att*Tn7-containing plasmids.** In pAR6 (35), the chromosomal *Eco*RI fragment containing the specific Tn7 insertion point is inserted into the *Eco*RI site of pBR322 with *att*Tn7R adjacent to vector β-lactamase; this plasmid contains *att*Tn7(−10,200 to +869).

We constructed the following *att*Tn7 plasmids. For pRM1, the *Hpa*I-*Hpa*I fragment from pAR6 containing the specific Tn7 insertion point was inserted into the *Hin*cII site of pUC18 with *att*Tn7R adjacent to vector p*lac*; this plasmid contains *att*Tn7(−459 to +623). To construct pRM2, the *Taq*I-*Taq*I fragment from pRM1 containing the specific Tn7 insertion point was inserted into the *Acc*I site of pUC18 with *att*Tn7R adjacent to vector p*lac*; this plasmid contains *att*Tn7(−342 to +165) (Fig. 1B). To construct pRM4, the *Sma*I site of pRM2 was converted to a *Hin*dIII site by linker insertion and the resulting *Hin*dIII-*Hin*dIII *att*Tn7(−342 to +165) fragment was inserted into the unique *Hin*dIII site of pNK217 (13) with *att*Tn7R adjacent to vector-encoded tetracycline resistance.

To construct pKAO1-1 and pKAO1-2, the *Dra*I-*Hin*cII fragment from pRM2 containing the specific insertion point was inserted into the *Hin*cII site of pUC18 with *att*Tn7R adjacent to vector p*lac* (pKAO1-1) or with *att*Tn7L adjacent to vector p*lac* (pKAO1-2); these plasmids contain *att*Tn7 (−158 to +64).

For pKAO3, the *Nla*III-*Hin*cII fragment from pRM2 containing the specific insertion point was inserted into the

*Hin*cII-*Sph*I sites of pUC18 such that *att*Tn*7*R is adjacent to vector p*lac*; this plasmid contains *att*Tn*7*(−52 to +64).

For pKAO4-1, pKAO4-2, and pKAO4-3, the *Pvu*II-*att*Tn*7*(+64 to −158)-*Pst*I fragment of pKAO1-1 was isolated and partially digested with *Fok*I, the termini were filled in with Klenow fragment, and the fragment extending from *Pvu*II to the *Fok*I site distal to the specific point of Tn*7* insertion was inserted into the *Sma*I site of pKO500 with *att*Tn*7*L adjacent to vector *galK* to give pKAO4-1. pKO500 (from K. McKenney, National Institutes of Health) is pKO1 (29) derivatized by replacement of the *Eco*RI site by a *Sal*I linker, insertion of the 300-bp *Pvu*II-*Pvu*II fragment containing p*lac* and the polylinker from m13mp11 into the *Sma*I site of pKO1, and subsequent deletion of the p*lac*-containing *Hin*dIII fragment to give *Sal*I-*Hin*dIII(polylinker)*Eco*RI-*galK*. pKAO4-2 and pKAO4-3 contain the *Eco*RI *att*Tn*7* fragment from pKAO4-1 inserted into the *Eco*RI site of pUC19 with *att*Tn*7*R adjacent to p*lac* (pKAO4-2) or with *att*Tn*7*L adjacent to p*lac* (pKAO4-3). All of these plasmids contain *att*Tn*7*(−25 to +64).

To construct pKAO16, pEG10, and pEG11, the *Pvu*II-*att*Tn*7*(+64 to −158)-*Pst*I fragment of pKAO1-1 was treated as described above and the fragment extending from *Pvu*II to the *Fok*I site proximal to the specific insertion point was inserted into the *Sma*I site of pKO500 with *att*Tn*7*L adjacent to vector *galK* to give pKAO16. pEG10 and pEG11 contain the *Sst*I *att*Tn*7* fragment from pKAO16 inserted into the *Sst*I site of pKL600 with *att*Tn*7*R adjacent to p*lac* (pEG10) or with *att*Tn*7*L adjacent to p*lac* (pEG11). pKL600 (K. McKenney) is pKO1 (29) derivatized by replacement of the *Eco*RI site by a *Sal*I linker, introduction of the 300-bp *Pvu*II-*Pvu*II fragment containing the p*lac* promoter and polylinker from m13mp10 into the *Sma*I site, and introduction of a translation stop codon in place of the polylinker *Eco*RI site. All of these plasmids contain *att*Tn*7*(−4 to +64).

To construct pKAO15, the *Pvu*II-*att*Tn*7*(+64 to −52)-*Pst*I fragment of pKAO1-1 was treated as described above and the fragment extending from *Pvu*II to the *Fok*I site proximal to the specific insertion point was inserted into the *Sma*I site of pKO500 with *att*Tn*7*R adjacent to vector *galK*; this plasmid contains *att*Tn*7*(−3 to +64).

For pTC1, the *Eco*RI-*Hin*dII *att*Tn*7* fragment from pRM2 was isolated and digested with *Mae*III, the termini were filled in with Klenow fragment, and the *Mae*III fragment containing the specific insertion point was inserted into the *Hin*cII site of pUC18 with *att*Tn*7*R adjacent to vector p*lac*; this plasmid contains *att*Tn*7*(−55 to +38).

RESULTS

A small DNA segment has *att*Tn*7* activity. In our nomenclature for *att*Tn*7*, the middle base of the 5-bp chromosomal sequence duplicated upon Tn*7* insertion (15, 25) is designated 0, sequences extending leftward toward *phoS* from the point of insertion (*att*Tn*7*L) are given a minus value, and sequences extending rightward toward *glmS* are given a plus value (Fig. 2). pAR6 (35) contains *att*Tn*7*(−10,200 to +869), and we constructed plasmids containing the specific insertion point and various extents of flanking *att*Tn*7*L and *att*Tn*7*R sequences. We evaluated the capacity of these plasmids to be targets for Tn*7* insertion by introducing them into *recA* *att*Tn7_{84}::Tn*7* hosts and then determining the fraction of the plasmids occupied by Tn*7*. The fractional occupancies by Tn*7* of several plasmids containing *att*Tn*7* sequences, ranging from *att*Tn*7*(−10,200 to +869) (pAR6) to *att*Tn*7*(−3 to +64) (pKAO15), is several thousandfold higher than that of a non-*att*Tn*7* plasmid (Table 1). We also examined the position of Tn*7* insertion in these plasmids by restriction enzyme digestion and found that Tn*7* inserted site and orientation specifically into the tested plasmids (data not shown). Thus, *att*Tn*7*(−3 to +64) is a target for efficient site- and orientation-specific insertion of Tn*7*.

Quantitative assay for Tn*7* transposition. Although the fractional occupancy assay described above demonstrated that sequences immediately surrounding the specific insertion point can direct insertion of Tn*7*, we were concerned that the observed fractional occupancies might be influenced by differences between the unoccupied and occupied plasmids other than the target activity of the *att*Tn*7* segment, such as plasmid stability and the efficiency of DNA isolation, and thus might not be an accurate reflection of the target activity of the *att*Tn*7* segment. Therefore, we developed an acute assay for Tn*7* transposition. We devised a lambda hop assay (23) in which a Tn*7*-end derivative transposes from a replication- and integration-defective variant of bacteriophage lambda upon phage infection. The Tn*7*-end derivatives Tn*7*-Cm (18) and Tn*7*-Km (Arcisewska and Craig, in preparation) contain the essential *cis*-acting sequences for Tn*7* transposition and transpose when complemented in *trans* by Tn*7*-encoded transposition proteins. We constructed phage lambda vectors containing these Tn*7*-end derivatives as described in Materials and Methods. Upon infection of phages carrying the Tn*7*-end derivatives into strains containing Tn*7* in the chromosome and the *att*Tn*7*-containing plasmid pAR6, high-frequency transposition of the Tn*7*-end derivatives was observed (Table 2, row 7). The Tn*7*-end derivatives inserted into pAR6 in site- and orientation-specific fashion (data not shown). Product colonies were lambda and tetracycline sensitive and thus contained no information from the donor phage other than the Tn*7*-end derivative (data not shown). High-frequency transposition required Tn*7* (Table 2, rows 1 and 2) and a vacant plasmid *att*Tn*7* site (Table 2, rows 3 to 6). (Although Tn*7* in *att*Tn7_{84}::Tn*7* can transpose to the *att*Tn*7* plasmid as described above, the fraction of plasmids actually occupied by

TABLE 1. Evaluation of *att*Tn*7* target activity by analysis of Tn*7* insertion[a]

Plasmid	Plasmid *att*Tn*7* sequence	Trial	Tp^r colonies/ 5 μl of DNA[b]	Cb^r colonies/ 0.1 μl of DNA[c]	Fractional occupancy by Tn*7*
pAR6	−10,200 to +869	1	111	2,384	9.3×10^{-4}
		2	96	2,600	7.4×10^{-4}
		3	106	1,824	1.2×10^{-4}
pRM2	−342 to +165	1	16	2,496	1.3×10^{-4}
		2	52	3,144	3.8×10^{-4}
		3	4	2,520	3×10^{-5}
pKAO4-1	−28 to +64	1	43	160	5.3×10^{-3}
		2	1,384	312	8.9×10^{-2}
pKAO16	−4 to +64	1	3,584	1,120	6.4×10^{-2}
		2	2,848	850	6.7×10^{-2}
		3	3,528	1,830	3.9×10^{-2}
pKAO15	−3 to +64	1	1,144	2,824	8.1×10^{-3}
		2	1,000	3,528	5.7×10^{-3}
pUC19		1	0	288	$<2 \times 10^{-6}$

[a] The plasmids to be tested were first introduced into the Tn*7*-containing strain LA3; plasmid DNA was then isolated and transformed, selecting for either trimethoprim resistance (Tp^r) to assay Tn*7*-containing plasmids or carbenicillin resistance (Cb^r) to assay total plasmids. Each trial represents an independent introduction of the plasmid into LA3.

[b] Per 20 μl of DNA for pUC19.

[c] Per 0.01 μl of DNA for pUC19.

TABLE 2. Evaluation of *att*Tn*7* target activity by Tn*7* lambda hop assays[a]

Tn*7* in *att*Tn7_{84}	Plasmid	Plasmid *att*Tn*7* sequence	Tn*7* end derivative	Transposition/PFU (mean ± SEM [*n*])
−	pUC18		Tn*7*-Km	$<10^{-6}$ (4)
−	pRM2	−342 to +165	Tn*7*-Km	$<10^{-6}$ (4)
+	pUC18		Tn*7*-Cm	$<10^{-6}$ (7)
+	pUC18		Tn*7*-Km	$<10^{-6}$ (5)
+	pAR6::Tn*7*	−10,200 to +869::Tn*7*	Tn*7*-Km	$<10^{-6}$ (3)
+	pRM2::Tn*7*	−342 to +165::Tn*7*	Tn*7*-Km	$<10^{-6}$ (3)
+	pAR6	−10,200 to +869	Tn*7*-Km	$(7.3 \pm 0.7) \times 10^{-3}$ (5)
+	pRM1	−459 to +623	Tn*7*-Km	$(3.0 \pm 1.2) \times 10^{-3}$ (4)
+	pRM2	−342 to +165	Tn*7*-Cm	$(11.0 \pm 2.3) \times 10^{-3}$ (10)
+	pRM2	−342 to +165	Tn*7*-Km	$(7.9 \pm 1.9) \times 10^{-3}$ (9)
+	pKAO1-1	−158 to +64	Tn*7*-Cm	$(9.9 \pm 2.1) \times 10^{-3}$ (10)
+	pKAO1-2	−158 to +64	Tn*7*-Cm	$(16.1 \pm 3.2) \times 10^{-3}$ (7)
+	pKAO1-1	−158 to +64	Tn*7*-Km	$(2.8 \pm 0.6) \times 10^{-3}$ (4)
+	pKAO3	−52 to +64	Tn*7*-Cm	$(6.6 \pm 1.3) \times 10^{-3}$ (9)
+	pKAO3	−52 to +64	Tn*7*-Km	$(6.9 \pm 2.5) \times 10^{-3}$ (8)
+	pKAO4-1	−25 to +64	Tn*7*-Km	$(12.9 \pm 1.9) \times 10^{-3}$ (13)
+	pKAO4-2	−25 to +64	Tn*7*-Km	$(15.0 \pm 1.7) \times 10^{-3}$ (4)
+	pKAO4-3	−25 to +64	Tn*7*-Km	$(6.1 \pm 1.0) \times 10^{-3}$ (4)
+	pKAO16	−4 to +64	Tn*7*-Km	$(3.2 \pm 1.6) \times 10^{-3}$ (4)
+	pEG10	−4 to +64	Tn*7*-Km	$(6.0 \pm 0.8) \times 10^{-3}$ (5)
+	pEG11	−4 to +64	Tn*7*-Km	$(8.8 \pm 3.0) \times 10^{-3}$ (5)
+	pKAO15	−3 to +64	Tn*7*-Km	$(5.6 \pm 1.2) \times 10^{-3}$ (9)
+	pTC1	−55 to +38	Tn*7*-Cm	$<10^{-6}$ (10)
+	pTC1	−55 to +38	Tn*7*-Km	$<10^{-6}$ (4)

[a] Tn*7* lambda hop assays using either lambda KK1 as the Tn*7*-Km donor phage or lambda RM7 as the Tn*7*-Cm donor phage and either NLC51 (rows 1 and 2) or LA3 (rows 3 to 24) containing the indicated plasmids were carried out as described in Materials and Methods.

Tn*7* is small, so that most of the plasmid *att*Tn*7* sites are vacant [Table 1]). At low frequency (about 2×10^{-7} transpositions per PFU), the Tn*7*-end derivatives do transpose from the donor phage to chromosomal sites (K. Kubo and N. L. Craig, unpublished observation).

Target activity of plasmids containing various *att*Tn*7* segments. We used the Tn*7* lambda hop assay to directly and quantitatively compare the *att*Tn*7* target activities of plasmids containing the specific Tn*7* insertion point and various extents of flanking nucleotide sequence. We found that the target activity of *att*Tn*7*(−3 to +64) (Table 2, row 22) was equivalent to that of larger segments containing the specific insertion point and more extensive flanking sequence (rows 7 to 21). These results demonstrate that the sequence information required for *att*Tn*7* target activity, i.e., to act as a target for high-frequency insertion of Tn*7*, is contained within *att*Tn*7*(−3 to +64). We also found that a DNA segment containing less *att*Tn*7*R information, *att*Tn*7*(−55 to +38), was not a reactive target for Tn*7* insertion (lines 23 and 24), demonstrating that sequence information within the segment *att*Tn*7*(+39 to +64) is essential for *att*Tn*7* target activity.

In these experiments, the *att*Tn*7* segment was located downstream of a strong promoter provided by the vector backbone (see Materials and Methods). We directly compared the target activity of *att*Tn*7* segments located in either orientation with respect to this external promoter, i.e., with transcription entering *att*Tn*7* from *att*Tn*7*R or from *att*Tn*7*L. We found that the orientation of *att*Tn*7* with respect to an external promoter does not affect its target activity (Table 2, compare rows 11 and 12, 17 and 18, and 20 and 21).

Insertion of Tn*7*-end derivatives into *att*Tn*7* is site and orientation specific. To evaluate the insertion sites of the Tn*7*-end derivatives, we isolated plasmid DNA from the products of Tn*7* lambda hops and analyzed it by restriction enzyme digestion. These experiments revealed that the Tn*7*-end derivatives inserted site and orientation specifically into the active plasmid *att*Tn*7* sites, independent of whether transcription entered *att*Tn*7* from *att*Tn*7*R or from *att*Tn*7*L (data not shown). Thus, sequence information within *att*Tn*7*(−3 to +64) provides the fundamental characteristics of *att*Tn*7*: the capacity to be a high-frequency target for the site- and orientation-specific insertion of Tn*7*.

Nucleotide sequence of *att*Tn*7* and *att*Tn*7*::Tn*7*. Several other investigators have sequenced the *att*Tn*7* region, and the reported sequences immediately surrounding the specific insertion point differ (15, 25, 47). We also sequenced *att*Tn*7*(−52 to +64), and our results agree with those of Gay et al. (15). We used several methods to overcome the difficulties of sequencing the G+C-rich potential stem-loop region in *att*Tn*7* (see Materials and Methods). The sequence of the *att*Tn*7* region obtained by Gay et al. (15) and by us is shown in Fig. 2. We also sequenced a number of independent insertions of Tn*7*-Km into several different *att*Tn*7* segments. We sequenced both the left and right novel junctions of nine independent insertions into *att*Tn*7*(−25 to +64), four in pKAO4-2 and five in pKAO4-3. In eight instances, duplication of the 5-bp target sequence CCCGC was observed (Fig. 2). Lichtenstein and Brenner (25) observed the same duplication in their analysis of two independent Tn*7* insertions into an *att*Tn*7*-containing plasmid. In one instance, with pKAO4-2, we observed duplication of the sequence CCGCT which is displaced by 1 nucleotide toward *glmS* from the CCCGC sequence. We also sequenced 10 independent insertions into *att*Tn*7*(−4 to +64), 5 in pEG10 and 5 in pEG11. In nine instances, the target sequence CCCGC was also duplicated. In one instance, with pEG10, the sequence CCCCG which is displaced by 1 nucleotide toward *phoS* from the CCCGC sequence was duplicated. Because only two Tn*7* insertions into larger *att*Tn*7* segments have been sequenced (25), it is not clear whether the variation in the point of insertion we observed reflects the small *att*Tn*7* and Tn*7*-end derivative we used or authentic variation in the position of Tn*7* insertion. The alternate

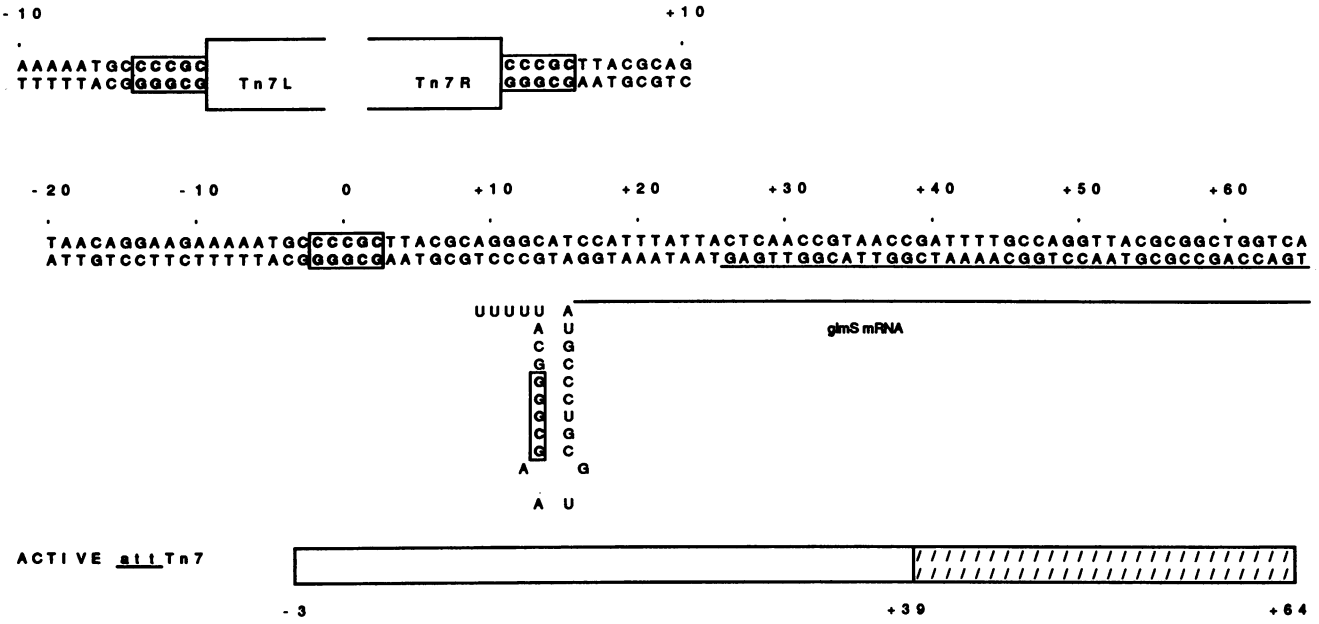

FIG. 2. Features of *att*Tn*7*. (Top) Sequences of the novel junctions created by Tn*7* insertion in *att*Tn*7* as directly determined by us for *att*Tn*7*L/Tn*7*L and by Gay et al. (15) and us for Tn*7*R/*att*Tn*7*R. The boxed bases are those duplicated upon Tn*7* insertion. (Middle) Sequence of *att*Tn*7* as determined by Gay et al. (15) and by us. The boxed bases are those duplicated upon Tn*7* insertion. The underlined bases are the protein-coding sequences at the 3′ end of *glmS*. (Bottom) Potential secondary structure at the 3′ end of *glmS* transcripts suggested by Gay et al. (15). *att*Tn*7*(−3 to +64) (boxed) provides *att*Tn*7* activity, and *att*Tn*7*(+39 to +64) (hatched region) contains information essential to *att*Tn*7* activity.

positions of insertion were both observed under conditions in which transcription enters *att*Tn*7* from *att*Tn*7*R, a condition which reflects the configuration of $attTn7_{84}$ (47).

DISCUSSION

We showed that the 68-bp *att*Tn*7*(−3 to +64) DNA fragment which includes the specific Tn*7* insertion point and extends rightward about 65 bp contains sufficient sequence information to promote the efficient site- and orientation-specific insertion of both intact Tn*7* and several Tn*7*-end derivatives. Our results are consistent with the hypothesis that *att*Tn*7* is contained entirely within this fragment, although we have not excluded the possibility that there are other important *att*Tn*7* sequences whose activity we artificially compensated for by examining the target activity of *att*Tn*7* fragments in multicopy plasmids. We also found that *att*Tn*7*(− 55 to +38) does not promote Tn*7* insertion. Thus, information within *att*Tn*7*(+39 to +64) is essential for *att*Tn*7* activity and at least some *att*Tn*7* information actually lies at a considerable distance from the specific insertion point. Our analysis did not establish whether the point of insertion itself is critical to *att*Tn*7* activity. It is possible that for Tn*7* insertion into *att*Tn*7*, the sites of specific recognition and action are distinct, as occurs with certain restriction enzymes (20).

An interesting result of our analysis is that at least some information required for *att*Tn*7* function actually lies within the coding sequence of the *glmS* gene, although the specific insertion point itself is downstream of *glmS* (Fig. 2). Thus, some nucleotides in this region have dual roles: they provide template information for mRNA synthesis and provide *att*Tn*7* activity. There are other nucleotide sequences which have been shown to have multiple functions; for example, the origin of phage lambda replication lies within the coding sequence of the lambda replication protein O (8, 14). Specific Tn*7* insertion sites have been also been observed in the chromosomes of other bacteria (1, 5, 6, 10, 19, 45, 46, 49). One interesting possibility is that these sites share considerable sequence homology with *E. coli att*Tn*7* and that these sequences are highly conserved because they also provide information for the *glmS* product, an enzyme involved in an important cellular function, cell wall biosynthesis. While this manuscript was in preparation, we learned that Lichtenstein and coworkers have found that specific Tn*7* insertion sites in several bacteria are adjacent to sequences homologous to the *E. coli glmS* terminus (C. Lichtenstein, personal communication).

As initially pointed out by Gay et al. (15), the specific point of Tn*7* insertion lies in a region of potential secondary structure: a G+C-rich stem-loop followed by T residues (Fig. 2). This sequence motif is characteristic of both Rho-independent transcription terminators (37) and secondary structures which stabilize the 3′ ends of mRNA (17). Gay et al. (15) showed that the 3′ termini of *glmS* transcripts do occur in this region. By introducing the *att*Tn*7*(−52 to +64) fragment into a terminator probe vector, we have shown directly that *att*Tn*7* does act as a transcription terminator (E. Gringauz, K. Orle, C. Waddell, and N. L. Craig, submitted for publication). A direct role for transcription has been implicated in other genomic rearrangements such as *Saccharomyces cerevisiae* mating-type switching (21) and immunoglobulin gene assembly (4, 42) and in the activation of certain origins of replication (8, 9). We show here that *att*Tn*7* activity is independent of the direction of transcription entering *att*Tn*7*. In other experiments, we have found that *att*Tn*7* segments promote efficient insertion of the Tn*7*-end derivatives when located in vectors which do not provide high-level transcription into the *att*Tn*7* segment (Gringauz et al., submitted). Thus, it is unlikely that transcription from *glmS* is a regulator of *att*Tn*7* activity.

Why is *att*Tn*7* such a specific and reactive site for Tn*7* insertion? There are no considerable homologies between the DNA segments which participate directly in Tn*7* site-specific insertion, i.e., *att*Tn*7* (Fig. 2) and the ends of Tn*7* (15, 16, 25). Other mobile genetic elements such as bacteriophage lambda which recognize specific target sites employ extensive structural similarities which include regions of precise DNA homology between the target site and the mobile element to direct insertion (for reviews, see references 39 and 48). We do not yet know how the ends of Tn*7* specifically recognize *att*Tn*7* in the absence of obvious structural similarity between these DNA segments. One attractive hypothesis is that a protein(s) recognizes the ends of Tn*7* and that this protein interacts with another protein which recognizes *att*Tn*7*. Alternatively, a single protein which recognizes two different nucleotide sequences (34, 38) could mediate this interaction. Further work is required to define the nucleotides within the *att*Tn*7*(−3 to +64) segment which are essential to *att*Tn*7* target activity and to determine how the ends of Tn*7* recognize *att*Tn*7*.

ACKNOWLEDGMENTS

We thank Andrew Wright, Jim Shapiro, Nancy Kleckner, Keith McKenney and John Clark for their generous gifts of strains, plasmids, and phage; Lidia Arciszewska for Tn*7*-Km; Ken Kubo for lambda KK1, and Ella Gringauz for pEG10 and pEG11. We also thank Conrad Lichtenstein for communicating results before publication. The other members of the laboratory (Lidia Arciszewska, Ken Kubo, Ella Gringauz, and Candy Waddell) also provided many helpful discussions and comments concerning the manuscript. We are also grateful to Susan Michaelis for many helpful discussions and for comments concerning the manuscript. We thank Karyl Nakamura for excellent preparation of the manuscript.

R.L.M. was supported by a National Research Service Award from the National Cancer Institute, and the work was supported by a Public Health Service grant from the National Institute of Allergy and Infectious Diseases to N.L.C.

LITERATURE CITED

1. **Barry, G.** 1986. Permanent insertion of foreign genes into the chromosomes of soil bacteria. Bio/Technology **4**:446–449.
2. **Barth, P. T., and N. Datta.** 1977. Two naturally occurring transposons indistinguishable from Tn*7*. J. Gen. Microbiol. **102**: 129–134.
3. **Barth, P. T., N. Datta, R. W. Hedges, and N. J. Grinter.** 1976. Transposition of a deoxyribonucleic acid sequence encoding trimethoprim and streptomycin resistances from R483 to other replicons. J. Bacteriol. **125**:800–810.
4. **Blackwell, T., M. Moore, G. Yancopoulos, H. Suh, S. Lutzker, E. Selsing, and F. Alt.** 1986. Recombination between immunoglobulin variable region gene segments is enhanced by transcription. Nature (London) **324**:585–589.
5. **Bolton, E., P. Glynn, and F. O'Gara.** 1984. Site specific transposition of Tn*7* into a *Rhizobium meliloti* megaplasmid. Mol. Gen. Genet. **193**:153–157.
6. **Caruso, M., and J. Shapiro.** 1982. Interactions of Tn*7* and temperate phage F116L of *Pseudomonas aeruginosa*. Mol. Gen. Genet. **188**:292–298.
7. **Chen, E., and P. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4**:165–170.
8. **Denniston-Thompson, K., D. Moore, K. Kruger, M. Furth, and F. Blattner.** 1977. Physical structure of the replication origin of bacteriophage lambda. Science **198**:1051–1056.
9. **Dove, W., H. Inokuchi, and W. Stevens.** 1971. Replication control in phage lambda, p. 747–771. *In* A. D. Hershey (ed.), The bacteriophage lambda. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
10. **Ely, B.** 1982. Transposition of Tn*7* occurs at a single site on the *Caulobacter crescentus* chromosome. J. Bacteriol. **151**:1056–1058.
11. **Fling, M., J. Kopf, and C. Richards.** 1985. Nucleotide sequence of the transposon Tn*7* gene encoding an aminoglycoside-modifying enzyme, 3"(9)-O-nucleotidyltransferase. Nucleic Acids Res. **13**:7095–7106.
12. **Fling, M., and C. Richards.** 1983. The nucleotide sequence of the trimethoprim-resistant dihydrofolate reductase gene harbored by Tn*7*. Nucleic Acids Res. **11**:5147–5158.
13. **Foster, T., M. Davis, D. Roberts, K. Takeshita, and N. Kleckner.** 1981. Genetic organization of transposon Tn10. Cell **23**:201–213.
14. **Furth, M., F. Blattner, C. McLeester, and W. Dove.** 1977. Genetic structure of the replication origin of bacteriophage lambda. Science **198**:1046–1051.
15. **Gay, N., V. Tybulewicz, and J. Walker.** 1986. Insertion of transposon Tn*7* into the *Escherichia coli glmS* transcriptional terminator. Biochem. J. **234**:111–117.
16. **Gosti-Testu, F., and J. Brevet.** 1982. Determination des sequences terminales du transposon Tn*7*. C.R. Acad. Sci. **294**: 193–196.
17. **Guarneros, G., C. Montanez, T. Hernandez, and D. Court.** 1982. Posttranscriptional control of bacteriophage λ *int* gene expression from a site distal to the gene. Proc. Natl. Acad. Sci. USA **79**:238–242.
18. **Hauer, B., and J. Shapiro.** 1984. Control of Tn*7* transposition. Mol. Gen. Genet. **194**:149–158.
19. **Hernalsteens, J., F. Van Vliet, M. De Beuckeleer, A. De Picker, G. Engler, M. Lemmers, M. Holsters, M. Van Montagu, and J. Schell.** 1980. The *Agrobacterium tumefaciens* Ti plasmid as a host vector system for introducing foreign DNA in plant cells. Nature (London) **287**:654–656.
20. **Kessler, C., P. Neumaier, and W. Wolf.** 1985. Recognition sequences of restriction endonucleases and methylases—a review. Gene **33**:1–102.
21. **Klar, A., J. Strathern, and J. Abraham.** 1984. Involvement of double-strand chromosomal breaks for mating-type switching in *Saccharomyces cerevisiae*. Cold Spring Harbor Symp. Quant. Biol. **49**:77–88.
22. **Kleckner, N.** 1981. Transposable elements in prokaryotes. Annu. Rev. Genet. **15**:341–404.
23. **Kleckner, N., D. Barker, D. Ross, and D. Botstein.** 1978. Properties of the translocatable tetracycline-resistance element Tn10 in *Escherichia coli* and bacteriophage lambda. Genetics **90**:427–461.
24. **Lichtenstein, C., and S. Brenner.** 1981. Site-specific properties of Tn*7* transposition into the *E. coli* chromosome. Mol. Gen. Genet. **183**:380–387.
25. **Lichtenstein, C., and S. Brenner.** 1982. Unique insertion site of Tn*7* in the *E. coli* chromosome. Nature (London) **297**:601–603.
26. **Magota, K., N. Otsuji, T. Miki, T. Horiuchi, S. Tsunasawa, J. Kondo, F. Sakiyama, M. Amemura, T. Morita, H. Shinagawa, and A. Nakata.** 1984. Nucleotide sequence of the *phoS* gene, the structural gene for the phosphate-binding protein of *Escherichia coli*. J. Bacteriol. **157**:909–917.
27. **Maniatis, T., E. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
28. **Maurer, R., B. Osmond, E. Shekhtman, A. Wong, and D. Botstein.** 1984. Functional interchangeability of DNA replication genes in *Salmonella typhimurium* and *Escherichia coli* demonstrated by a general complementation procedure. Genetics **108**:1–23.
29. **McKenney, K., H. Shimatake, D. Court, U. Schmeissner, and M. Rosenberg.** 1981. A system to study promoter and terminator signals recognized by *Escherichia coli* RNA polymerase, p. 383–415. *In* J. Chirikjian and T. Papas (ed.), Gene amplification and analysis, vol. 2. Elsevier/North-Holland Publishing Co., New York.

29a. **McKown, R. L., C. S. Waddell, L. K. Arciszewska, and N. L. Craig.** 1987. Identification of a transposon Tn*7*-dependent DNA binding activity that recognizes the ends of Tn*7*. Proc. Natl. Acad. Sci. USA **84**:7807–7811.

30. **Miller, J.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

31. **Mills, D., and F. Kramer.** 1979. Structure-independent nucleotide sequence analysis. Proc. Natl. Acad. Sci. USA **76:** 2232–2235.
32. **Mizusawa, S., S. Nishimura, and F. Seela.** 1986. Improvement of the dideoxy chain termination method of DNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. Nucleic Acids Res. **14:**1319–1324.
33. **O'Farrell, P.** 1987. Dideoxy DNA sequencing using T4 gene 32 protein. B.M. Biochemica **4**(2):13.
34. **Pfeifer, K., T. Prezant, and L. Guarente.** 1987. Yeast HAP1 activator binds to two upstream activation sites of different sequence. Cell **49:**19–27.
35. **Reynolds, A., J. Felton, and A. Wright.** 1981. Insertion of DNA activates the cryptic *bgl* operon in *E. coli* K12. Nature (London) **293:**625–629.
36. **Rogers, M., N. Ekaterinaki, E. Nimmo, and D. Sherratt.** 1986. Analysis of Tn*7* transposition. Mol. Gen. Genet. **205:**550–556.
37. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13:**319–353.
38. **Ross, W., and A. Landy.** 1982. Bacteriophage λ int protein recognizes two classes of sequence in the phage *att* site: characterization of arm-type sites. Proc. Natl. Acad. Sci. USA **79:**7724–7728.
39. **Sadowski, P.** 1986. Site-specific recombinases: changing partners and doing the twist. J. Bacteriol. **165:**341–347.
40. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
41. **Shapiro, J.** 1983. Mobile genetic elements. Academic Press, Inc., Orlando, Fla.
42. **Stavnezer-Nordgen, J., and S. Sirlin.** 1986. Specificity of immunoglobulin heavy chain switch correlates with activity of germline heavy chain genes prior to switching. EMBO J. **5:**95–102.
43. **Struhl, K.** 1983. Direct selection for gene replacement events in yeast. Gene **26:**231–242.
44. **Surin, B. P., D. A. Jans, A. L. Fimmel, D. C. Shaw, G. B. Cox, and H. Rosenberg.** 1984. Structural gene for the phosphate-repressible phosphate-binding protein of *Escherichia coli* has its own promoter: complete nucleotide sequence of the *phoS* gene. J. Bacteriol. **157:**772–778.
45. **Thomson, J. A., M. Hendson, and R. M. Magnes.** 1981. Mutagenesis by insertion of drug resistance transposon Tn*7* into a vibrio species. J. Bacteriol. **148:**374–378.
46. **Turner, P., C. Barber, and M. Daniels.** 1984. Behaviour of the transposons Tn*5* and Tn*7* in *Xanthomonas campestris* pv. *campestris*. Mol. Gen. Genet. **193:**153–157.
47. **Walker, J., N. Gay, M. Saraste, and A. Eberle.** 1984. DNA sequence around the *Escherichia coli unc* operon. Biochem. J. **224:**799–815.
48. **Weisberg, R., and A. Landy.** 1983. Site-specific recombination in phage lambda, p. 211–250. *In* R. Hendrix, J. Roberts, F. Stahl, and R. Weisberg (ed.), Lambda II. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
49. **Youvan, D., J. Elder, D. Sandlin, K. Zsebo, D. Alder, N. Panopoulos, B. Marrs, and J. Hearst.** 1982. R-prime site-directed transposon Tn*7* mutagenesis of the photosynthetic apparatus in *Rhodpseudomonas capsulata*. J. Mol. Biol. **162:**17–41.