Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 3958–3962, June 1989
Biochemistry

Tn*7* transposition: Recognition of the *attTn7* target sequence

(recombination/DNA binding protein/*tnsD* gene)

CANDACE S. WADDELL[†‡§] AND NANCY L. CRAIG[†‡¶‖]

Departments of [†]Biochemistry and Biophysics and [¶]Microbiology and Immunology, and the [‡]G. W. Hooper Foundation, University of California, San Francisco, San Francisco, CA 94143

Communicated by J. Michael Bishop, February 3, 1989

ABSTRACT **The bacterial transposon Tn*7* encodes two distinct but overlapping transposition pathways. *tnsABC* + *tnsD* promote transposition to a specific site, *attTn7*, while *tnsABC* + *tnsE* promote transposition to many other sites unrelated to *attTn7*. We have identified a *tnsD*-dependent DNA binding activity that specifically recognizes *attTn7*. We have localized the recognition sequences for this activity to a 28-base-pair region and have shown that this same region can provide specific properties of an *attTn7* target *in vivo*. Interestingly, these sequences are positioned more than 25 base pairs from the specific point of Tn*7* insertion.**

The transposon Tn*7* inserts at high frequency into a specific site in the genomes of many different bacteria (1). In *Escherichia coli* this site, called *attTn7*, is located at minute 84 of the chromosome (2–4). The actual point of Tn*7* insertion lies within the transcriptional terminator of *glmS* (5, 6), a gene involved in cell wall biosynthesis. Insertions at this site are orientation specific such that the right end of Tn*7* is adjacent to the carboxyl terminus of *glmS* (ref. 3; Fig. 1). Tn*7* transposition is accompanied by a 5-base-pair (bp) duplication of target sequences at the point of insertion (5, 7, 8). We designate the central base pair of this duplicated sequence 0; sequences toward *glmS* are given a plus (+) value and sequences away from *glmS* are given a minus (−) value (Fig. 1). Results of experiments performed *in vivo* suggest that the sequences required for Tn*7* transposition lie to the *glmS* side of the point of insertion and are within the region *attTn7* (+7 to +64) (6). These studies revealed two interesting features of the *attTn7* target: (*i*) the nucleotide sequence at the point of insertion can be altered with no apparent effect on the efficiency or specificity of transposition, and (*ii*) some essential nucleotides are distant from the point of Tn*7* insertion and lie within the *glmS* coding frame (8).

Tn*7* encodes five transposition genes (*tnsABCDE*) that mediate transposition via two pathways that differ in their choice of target site (9, 10). Both pathways share a requirement for *tnsA*, *tnsB*, and *tnsC*. One of these genes, *tnsB*, encodes a specific DNA-binding protein that recognizes sequences within the ends of Tn*7* (ref. 11; R. McKown, L. Arciszewska, and N.L.C., unpublished data). The functions of the *tnsA* and *tnsC* gene products are not yet known. In one Tn*7* pathway, *tnsABC* + *tnsD* promote transposition to target sites of a related sequence. These sites include *attTn7*, to which Tn*7* transposes at high frequency, and pseudo-*attTn7* sites, to which Tn*7* transposes at low frequency (K. Kubo and N.L.C., unpublished data). Pseudo-*attTn7* sites are specific chromosomal targets that display significant homology to *attTn7* in the region from positions +28 to +60 (K. Kubo and N.L.C., unpublished data). In the second transposition pathway, *tnsABC* + *tnsE* promote low frequency Tn*7* transposition to many different sites that bear no obvious relationship to *attTn7* (ref. 7; K. Kubo and N.L.C., unpublished data). Based on the pathway-specific requirement for *tnsD* and *tnsE*, we proposed that the proteins encoded by these genes have direct roles in Tn*7* target selection (10).

We report here the identification of a *tnsD*-dependent DNA binding activity that specifically recognizes *attTn7* sequences. This finding provides biochemical support for our model of Tn*7* transposition and further refines our understanding of *attTn7*.

MATERIALS AND METHODS

Bacteria and Plasmids. LB and broth were as described by Miller (12) except that glucosamine at 1 mg/ml was added to agar. Trimethoprim selection was on Iso-Sensitest agar (Oxoid, Basingstoke, U.K.). Antibiotics were carbenicillin, kanamycin, and trimethoprim at 100 μg/ml and tetracycline at 20 μg/ml. NLC51 is *E. coli* F⁻ *araD139* Δ(*argF-lac*)*U169* *rpsL150* *relA1* *flbB5301* *deoC1* *ptsF25* *rbsR* *val*[R] *recA56* (11). LA3 is NLC51 *attTn7*::Tn*7* (11). pCW4 contains *tnsABCDE* and pCW4::miniMuΩ72 carries an insertion mutation that disrupts *tnsD* (10). pCW23 and pCW25 contain *tnsD* in opposite orientations with respect to P*lac* in pUC18 (10). In pCW23, the 5′ end of *tnsD* is adjacent to P*lac*. pKAO4-3 contains *attTn7* (−25 to +64) (8). pEG61 contains *attTn7* (−52 to +64) and pEG61/1-1 has a C·G to T·A point mutation at nucleotide position *attTn7* (+37) (6). pCW80 and pCW85 contain *attTn7* (+28 to +55). The synthetic oligodeoxynucleotide

```
      +30        +40        +50
       |          |          |
5′-AACCGTAACCGATTTTGCCAGGTTACG
     GCATTGGCTAAAACGGTCCAATGCTTAA–5′
```

(Biomolecular Resource Center, University of California, San Francisco) underwent reaction with Klenow fragment in the presence of dGTP and TTP to repair the 3-bp overhang and was inserted into *Sma* I/*Eco*RI-digested pUC18 (pCW80). The *attTn7* sequence was verified by DNA sequencing. The *Hin*dIII/*Eco*RI fragment from pCW80 containing *attTn7* was inserted into *Hin*dIII/*Eco*RI-digested pUC19 (pCW85). pCW77 contains *attTn7* (+32 to +55) and was obtained by insertion of the synthetic oligodeoxynucleotide

```
                +40        +50
                 |          |
5′-CCGGCGTAACCGATTTTGCCAGGTTACG
     GCATTGGCTAAAACGGTCCAATGCTTAA–5′
```

(Biomolecular Resource Center) into *Xma* I/*Eco*RI-digested pUC19. The *attTn7* sequence was verified by DNA sequencing.

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.

[§]Present address: U.S. Department of Agriculture Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710.
[‖]To whom reprint requests should be addressed.

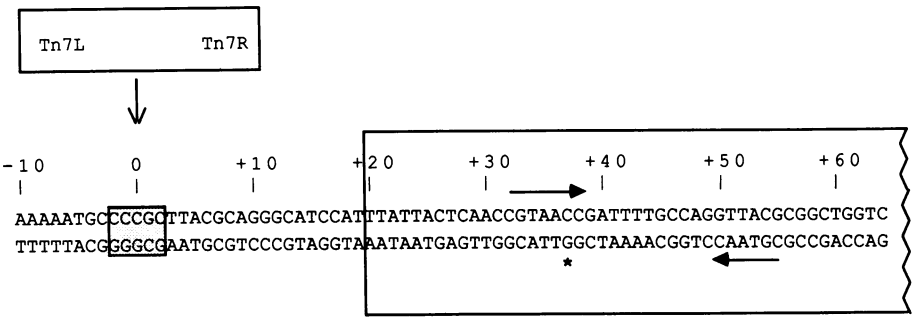

FIG. 1. *attTn7*. The nucleotide sequence of the *attTn7* region is shown and is numbered as described in the text. The 5-bp sequence duplicated upon Tn*7* insertion (shaded box) and the carboxyl terminus of *glmS* (large box) are indicated. Horizontal arrows indicate the perfect inverted repeat sequence in *attTn7* (+32 to +55). *, Nucleotide position +37. *attTn7* (+46 to +50) is a recognition sequence for the *dcm* methylase. Tn*7* (upper box) inserts into *attTn7* so that the left (Tn*7L*) and right (Tn*7R*) ends are oriented as shown.

Manipulation and Analysis of DNA. Plasmid growth, isolation, restriction enzyme analysis, and transformation were performed as described by Maniatis *et al.* (13). Recombinant molecules were made by standard cloning techniques, except that vector DNAs were contained in slices from low-melting-point agarose gels (Sea Plaque; FMC) as described by Struhl (14). DNA-modifying enzymes were obtained from commercial sources and used as recommended by the manufacturer. DNA fragments used in the binding assay were generated by digestion of plasmids with appropriate restriction enzymes, purified by separation in a nondenaturing polyacrylamide gel and elution from a gel slice, and 5'-end-labeled with deoxynucleoside 5'-[α-^{32}P]triphosphate and the Klenow fragment of DNA polymerase I. The oligodeoxynucleotide was 5'-end-labeled directly as described above.

DNA Fragments Used in Binding Assays. Fragments containing the following sequences were tested: pUC18 (141), the 141-bp pUC18 *Pvu* II/*Eco*RI fragment; *attTn7* (−52 to +64), the 143-bp pEG61 *Sst* I/*Sal* I fragment; +37 point mutation in *attTn7* (−52 to +64), the 143-bp pEG61/1-1 *Sst* I/*Sal* I fragment; *attTn7* (−25 to +64), the 135-bp pKAO4-3 *Eco*RI fragment; *attTn7* (+28 to +55), the 63-bp pCW80 *Hin*dIII/*Eco*RI fragment; *attTn7* (+32 to +55), the 62-bp pCW77 *Hin*dIII/*Eco*RI fragment; *attTn7* (+10 to +64) is a synthetic 55-bp oligodeoxynucleotide (6).

Preparation of Cell Extracts. Extracts were prepared from NLC51 and plasmid-containing derivatives of NLC51. Cultures were grown to an OD_{600} of 0.8 in LB broth with appropriate antibiotic at 37°C, harvested by centrifugation, washed, and resuspended (2 ml per g of cells) in a lysis buffer of 50 mM Tris·HCl, pH 8.0/100 mM KCl/1 mM EDTA/2 mM dithiothreitol/10% (vol/vol) glycerol. Cells were lysed by sonication and cell debris was removed by centrifugation at 48,400 × *g* for 15 min at 4°C. The resulting supernatant was dialyzed against lysis buffer overnight at 4°C and used directly or stored at −80°C. Preparations yielded ≈20 mg of total protein per ml, determined as described by Bradford (15) using bovine serum albumin as a standard.

DNA Binding Assays. Reaction mixtures (20 μl) contained 48 mM Tris·HCl (pH 8.0), 95 mM KCl, 1 mM EDTA, 1.9 mM dithiothreitol, 9.5% (vol/vol) glycerol, 0.1–0.5 ng of end-labeled DNA fragment, and 1 μl of crude extract unless otherwise noted. Reaction mixtures were incubated at 23°C for 10 min. Some reactions, noted in figure legends, also contained poly(dI-dC)-poly(dI-dC) (Pharmacia) to prevent nonspecific binding (16). These reaction mixtures, minus end-labeled DNA, were preincubated at 23°C for 10 min followed by a 10-min 23°C incubation upon the addition of end-labeled DNA. All fragments were tested with and without poly(dI-dC)-poly(dI-dC) and the same specific bands of retarded mobility were observed. However, the amount of protein-bound DNA relative to free DNA was greater when poly(dI-dC)-poly(dI-dC) was omitted. We also observed that the proportion of free versus bound DNA varied between lots of poly(dI-dC)-poly(dI-dC). All reaction mixtures were loaded onto gels containing 7.5% acrylamide, 0.2% *N,N'*-methylenebisacrylamide and were electrophoresed in Tris borate buffer for 3 hr at 10 V/cm. Gels were dried under vacuum and autoradiographed.

λ Hop Assays. *attTn7* target activity was evaluated using a Tn*7* λ hop assay (8) with sodium citrate omitted from the protocol. Briefly, *attTn7*-containing plasmids to be assayed were introduced into the Tn*7*-containing strain LA3. The chromosomal Tn*7* supplies the Tn*7*-encoded transposition proteins. The Tn*7*-end derivative, mini-Tn*7*Km, contains Tn*7*'s essential cis-acting transposition sequences and transposes from an integration- and replication-defective λ derivative upon infection into the cells. The λ donor phage was KK1 [λ 780 *hisG9424*::Tn*10 dell6 dell7*::*attTn7* (−342 to +165)::Tn*7L166*-kanamycin resistance-Tn*7R199*] (8).

Transposition products were selected on kanamycin-trimethoprim plates and purified by repeated rounds of selective growth. The plasmid DNA was recovered by an alkaline lysis "mini-prep" procedure, diluted, and transformed into DH5, a derivative of DH1 cells (17), selecting for kanamycin resistance. After purification of transformants, DNA was isolated and analyzed by restriction enzyme digestion. Both site and orientation of the mini-Tn*7*Km insertions were determined. pCW80 and pCW85 contain *attTn7* (+28 to +55) within a polylinker sequence so that the spatially equivalent *attTn7* (0) position lies within a unique *Sph* I restriction site. This facilitated accurate mapping of mini-Tn*7* insertions within the polylinker to a resolution of 6 bp.

RESULTS

The *tnsD* Gene Product Directs Specific Binding to *attTn7* Sequences. Genetic analysis has shown that *tnsD* is required to promote Tn*7* transposition to *attTn7* (9, 10). Using a gel-retardation assay (18, 19), we have identified a *tnsD*-dependent DNA binding activity that specifically recognizes *attTn7*. Crude extracts derived from cells bearing the *tnsD* plasmid pCW23 contain an activity that binds to and retards the migration of an *attTn7* (−25 to +64) fragment (complex D in Fig. 2, lane 4). No such activity is observed in extracts derived from cells lacking the *tnsD* plasmid (lane 3). Each of these extracts contains a host activity that promotes formation of a complex of different retarded mobility (complex H in Fig. 2, lanes 3 and 4). *tnsD*-dependent complex formation is complete within 10 min and remains unchanged after prolonged incubation (up to 90 min; data not shown). The *tnsD*-dependent complex is not observed with boiled extracts, whereas the host-dependent complex is unaltered by heat treatment (data not shown).

We examined the sequence specificity of the *tnsD*-dependent binding activity in several ways. This activity does

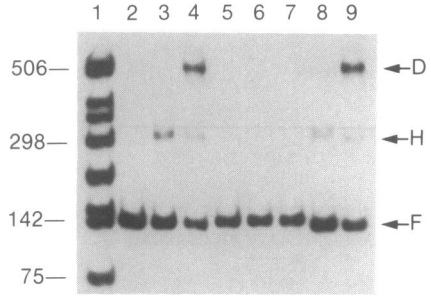

FIG. 2. *tnsD*-dependent binding to *attTn7* (−25 to +64). End-labeled DNA fragments reacted with crude extract and competitor DNA as indicated below and as described in *Materials and Methods*. After incubation, reaction mixtures were analyzed by gel electrophoresis and autoradiography. Arrows indicate free DNA (F), host-dependent complex (H), and *tnsD*-dependent complex (D). Labeled DNA: lane 1, *Hin*fI-digested pBR322 DNA standards, the sizes of selected bands are shown in bp; lanes 2–4, 8, and 9, *attTn7* (−25 to +64); lanes 5–7, pUC18 (141). Crude extract: lanes 2 and 5, no extract; lanes 3 and 6, host (20 μg of extract protein); lanes 4 and 7–9, pCW23 (5 μg of pCW23 extract protein plus 15 μg of host extract protein). Unlabeled competitor DNA: lane 8, 25 ng of unlabeled *attTn7* (+10 to +64); lane 9, 25 ng of unlabeled pUC18 (141).

not recognize a DNA fragment of similar length that lacks *attTn7* sequences (Fig. 2, lanes 5–7); no bands of retarded mobility are observed. In addition, a large excess of an unlabeled *attTn7* fragment successfully competes with the labeled *attTn7* fragment for the *tnsD*-dependent DNA binding activity (lane 8), whereas the same amount of an unlabeled pUC18 fragment has no effect on complex formation (lane 9). We observe specific competition when extracts are preincubated with unlabeled DNA (as shown) or when both labeled and unlabeled DNAs are added to the reaction mixture simultaneously (data not shown). Addition of $NaDodSO_4$ (final concentration, 0.1%) to a reaction mixture disrupts preformed complexes so that a retarded band of protein-bound DNA is not observed (data not shown). These results demonstrate that pCW23 extracts contain a *tnsD*-dependent DNA binding activity that specifically recognizes *attTn7* and that sequence recognition by this activity results in a non-covalent protein–DNA interaction.

Do other *tns*-encoded proteins affect or alter the *tnsD*-dependent *attTn7* complex? A single *tns*-dependent band of retarded mobility is observed when extracts from cells containing the multicopy *tnsABCDE* plasmid pCW4 (complex D in Fig. 3, lane 3) are assayed. This complex has the same mobility as the *tnsD*-dependent complex (lane 6).

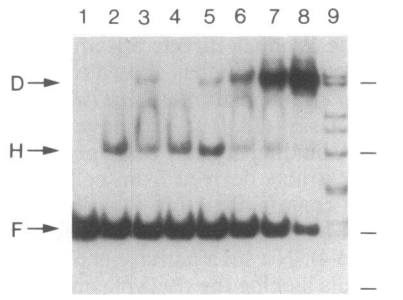

FIG. 3. Binding analyses of different *tns* extracts. Reaction mixtures containing crude extract protein as indicated below and end-labeled *attTn7* (−25 to +64) were analyzed as described in Fig. 2. Arrows indicate free DNA (F), host-dependent complex (H), and *tnsD*-dependent complex (D). Dashes indicate the same size markers as in Fig. 2. Lanes: 1, no extract; 2, host (20 μg); 3, pCW4 (20 μg); 4, pCW4::miniMuΩ72 (20 μg); 5, pCW25 (20 μg); 6, pCW23 (0.8 μg of pCW23 extract protein plus 19.2 μg of host extract protein); 7, pCW23 (4 μg of pCW23 extract protein plus 16 μg of host extract protein); 8, pCW23 (20 μg); 9, DNA standards.

Moreover, extracts from pCW4 derivatives in which the *tnsD* gene is disrupted have lost the *attTn7* specific DNA binding activity (lane 4). Thus, we have no evidence suggesting that another *tns*-dependent protein binds the *attTn7* fragment. We have also examined the *attTn7* binding properties of extracts from cells containing a single chromosomal copy of Tn*7*. Such Tn*7* extracts do contain the host-encoded DNA binding activity but *tnsD*-dependent complex formation with the *attTn7* fragment is vastly reduced, if detectable at all (data not shown). These results suggest that the amount of *tnsD*-dependent binding activity present in Tn*7* extracts is below that required to detect activity in our gel assays.

pCW23 extracts contain at least 25-fold more *attTn7* binding activity than do extracts derived from cells containing either pCW4 or a different *tnsD* plasmid, pCW25 (Fig. 3, compare lanes 6–8 to lanes 3 and 5). In the pCW23 plasmid, the *tnsD* gene is downstream of a strong vector promoter (see *Materials and Methods*). This high level of binding activity correlates with our *in vivo* transposition results. A mini-Tn*7* element displays elevated levels of transposition to an *attTn7*-containing plasmid when provided with pCW23 and a *tnsABC* plasmid in trans (10). It is tempting to speculate that the increased transposition frequency reflects an increased amount of *tnsD*-dependent *attTn7* binding activity.

In summary, we have identified an *attTn7* binding activity that is dependent on *tnsD* and neither requires nor is altered by any other *tns* gene function. While pCW23 extracts are a particularly rich source of this activity, it remains to be rigorously determined whether or not the *tnsD*-encoded protein is itself the DNA binding protein.

***attTn7* (+28 to +55) Can Participate in *tnsD*-Dependent Complex Formation.** What sequences in *attTn7* (−25 to +64) are required for *tnsD*-dependent binding? We have previously shown that a plasmid containing *attTn7* (+7 to +64) is a target for high frequency site- and orientation-specific Tn*7* transposition *in vivo* (6). An end-labeled synthetic oligonucleotide of the sequence *attTn7* (+10 to +64) is also an efficient substrate for *tnsD*-dependent complex formation when incubated with pCW23 extracts (complex D in Fig. 4*A*, lane 4). Therefore, the sequences that direct *tnsD*-dependent binding must lie within this 55-bp *attTn7* segment. [We note that in competition experiments described above, the *attTn7* (+10 to +64) oligonucleotide served as the specific cold competitor (see Fig. 2, lane 8).] Further evidence that sequences within this region contribute to *attTn7* activity comes from an *in vivo* analysis of mutant *attTn7* fragments (6). A C·G to T·A transition at nucleotide position +37 reduces Tn*7* transposition to *attTn7* by a factor of ≈100. This mutation also affects *tnsD*-dependent complex formation. An *attTn7* DNA fragment containing the +37 point mutation required ≈5-fold more pCW23 extract to detect complex formation than did a fragment of the wild-type *attTn7* sequence (complex D in Fig. 4*B*, compare lanes 6 and 7 to lane 3). This suggests that the nucleotide change at position +37 alters the affinity of the DNA binding activity to the mutant *attTn7*. Point mutations at positions +2 and +13 had no effect on complex formation *in vitro* (data not shown) or on *attTn7* target activity *in vivo* (6).

To further localize the DNA recognition sequences, we examined the activity of smaller *attTn7* segments in the gel retardation assay. A DNA fragment containing a 28-bp segment, *attTn7* (+28 to +55), did promote *tnsD*-dependent complex formation when reacted with pCW23 extracts (complex D in Fig. 4*A*, lane 7). The *attTn7* (+28 to +55) fragment displayed the same relative mobility shift between free and bound DNA (retardation equivalent to a size increase of ≈375 bp of DNA) as did the *attTn7* (−25 to +64) fragment and the *attTn7* (+10 to +64) oligonucleotide. Note that the host-dependent complex observed with larger *attTn7* fragments is not detectable with the *attTn7* (+28 to +55) fragment (com-

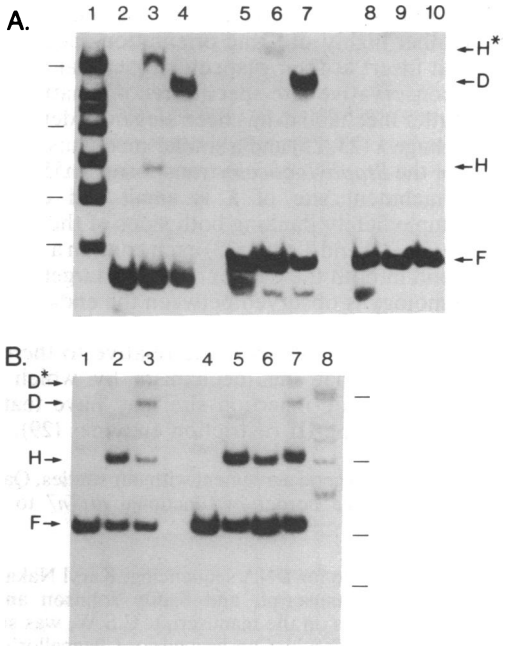

FIG. 4. Binding analyses of different *attTn7* segments. Reaction mixtures containing crude extract protein and end-labeled *attTn7* fragments as indicated below were analyzed as described in Fig. 2. Arrows indicate free DNA (F), host-dependent complexes (H and H*), and *tnsD*-dependent complexes (D and D*). Dashes indicate the same size markers as in Fig. 2. (*A*) Lanes: 1, DNA standards; 2–4, *attTn7* (+10 to +64); 5–7, *attTn7* (+28 to +55); 8–10, *attTn7* (+32 to +55); with no extract, host extract (20 μg), and *tnsD* extract (20 μg), respectively. The host-dependent complex H* (lanes 3 and 6) is not observed with larger *attTn7* fragments. The presence and amount of the H* complex varies substantially between reactions (data not shown) and the significance of this complex is not clear. (*B*) All reactions include poly(dI-dC)-poly(dI-dC). Lanes: 1–3, *attTn7* (−52 to +64) with no extract, host extract (20 μg), and *tnsD* extract (20 μg), respectively; 4–7, +37 mutation in *attTn7* (−52 to +64) with no extract, host extract (20 μg), *tnsD* extract (20 μg), and *tnsD* extract (100 μg), respectively; 8, DNA standards. A second *tnsD*-dependent complex is seen in these reactions (complex D*; lanes 3 and 7). A similar D* complex is occasionally observed in reactions with *attTn7* (−25 to +64) (data not shown). This complex may contain both the *tnsD*- and the host-dependent binding activities.

pare complex H in Fig. 4*A*, lane 6 to lane 3 and to Fig. 2, lane 3). Analysis of other *attTn7* fragments suggests that removal of nucleotides between positions +10 and +27 destroys this apparent host protein binding site (data not shown).

Inspection of the nucleotide sequence of *attTn7* (+28 to +55) reveals a large region of dyad symmetry extending from position +32 to position +55 (see Fig. 1). Because many DNA binding proteins utilize inverted repeat sequences as recognition sites (20), we asked whether this region is sufficient for *tnsD*-dependent binding. When an end-labeled DNA fragment containing the inverted repeat sequence *attTn7* (+32 to +55) was incubated with pCW23 extracts, no discrete band of retarded mobility was observed (Fig. 4*A*, lane 10). Increased amounts of pCW23 extract (up to 20-fold) also failed to produce detectable complex formation (data not shown). Thus, this perfect inverted repeat is not, or is not the sole component of, the *tnsD*-dependent recognition site, and some nucleotides between positions +28 and +31 are essential for DNA binding.

***attTn7* (+28 to +55) Is a Target for Tn7 Transposition *in Vivo*.** Are the sequences required for *tnsD*-dependent DNA binding *in vitro* the same sequences required for *attTn7* target activity *in vivo*? Using the Tn7 λ hop assay (see *Materials and Methods*), we evaluated the target activity of plasmids containing *attTn7* (+32 to +55) and *attTn7* (+28 to +55). The plasmid containing the binding-deficient *attTn7* (+32 to +55) segment has no apparent activity as a target for Tn7 transposition. Transposition in cells containing this plasmid occurs at the same low frequency as in cells lacking an *attTn7* plasmid (Table 1, lines 1 and 2). The rare transposition products are likely insertions in the *E. coli* chromosome because the mini-Tn7 element did not cosegregate with the *attTn7* plasmid (data not shown). By contrast, a plasmid containing the binding-proficient *attTn7* (+28 to +55) segment does act as a target for Tn7 transposition (Table 1, lines 3 and 4). Moreover, *attTn7* (+28 to +55) encodes sufficient information to direct site- and orientation-specific insertion of the Tn7 element. Restriction enzyme analysis of 24 independent transposition products revealed that all of the insertions had occurred in the expected orientation with the right end of the element nearest to *attTn7* (+28) (data not shown). Recall that Tn7 transposition results in a 5-bp target duplication with the central nucleotide of the duplication designated *attTn7* (0) (see Fig. 1). Twenty-one of the 24 transposition products had inserted so that the central base pair of the target duplication was located in the 6-bp region spatially equivalent to *attTn7* (0 to +5) (data not shown). The resolution of our mapping precludes comment on the fidelity of the exact insertion point. A small degree of wobble, as much as 3 bp, at the point of insertion has been observed with larger *attTn7* fragments (6, 8). Three of the 24 insertions occurred at positions approximately equivalent to *attTn7* (−18), (−55), and (−128); the frequency of such events for wild-type *attTn7* transposition has not been determined.

Although *attTn7* (+28 to +55) confers the specificity properties of an *attTn7* target, it does not provide the high activity level observed with larger *attTn7* segments. Transposition to the *attTn7* (+28 to +55) plasmids pCW80 and pCW85 is 100–1000 times lower than to the *attTn7* (−25 to +64) plasmid pKAO4-3 (Table 1, compare lines 3 and 4 to line 5). Another indication that the 28-bp *attTn7* does not confer wild-type activity is its sensitivity to vector-encoded asymmetries. The 6-fold difference in transposition frequency between pCW80 and pCW85 correlates with the orientation of *attTn7* (+28 to +55) with respect to vector sequences. Large *attTn7* fragments, such as *attTn7* (−25 to +64), are insensitive to the orientation of vector sequences (6, 8).

In conclusion, the *in vivo* transposition results demonstrate that *attTn7* (+28 to +55) encodes sufficient information to promote site- and orientation-specific Tn7 transposition. The reduced transposition frequency suggests that *attTn7* (+28 to +55) contains most, but not all, of the information required for high-frequency wild-type *attTn7* target activity.

DISCUSSION

Genetic analysis has shown that Tn7 encodes two target-specific genes, *tnsD* and *tnsE*; *tnsD* promotes transposition to a specific chromosomal site, *attTn7*, and *tnsE* promotes transposition to many unrelated sites (9, 10). Based on these

Table 1. Evaluation of *attTn7* target activity

Plasmid	*attTn7* sequence	Transposition/PFU, mean ± SEM (*n*)
pUC18*	—	$3.6 \pm 3.6 \times 10^{-7}$ (8)
pCW77	+32 to +55	$8.3 \pm 5.2 \times 10^{-7}$ (7)
pCW80	+28 to +55	$2.0 \pm 1.5 \times 10^{-5}$ (13)
pCW85	+28 to +55	$1.2 \pm 0.7 \times 10^{-4}$ (8)
pKAO4-3	−25 to +64	$2.2 \pm 0.95 \times 10^{-2}$ (8)

Tn7 λ hop assays were performed in LA3 cells containing the indicated plasmids. PFU, plaque-forming units.
*A similar transposition frequency ($6.0 \pm 0.7 \times 10^{-7}$) is observed when LA3 cells do not contain a plasmid (K. Kubo and N.L.C., unpublished data).

data, we proposed a model for Tn7 transposition in which TnsD and TnsE act as DNA binding proteins that interact with Tn7 target sequences. As a first step in testing this model, we have used a gel-retardation assay to identify a Tn7-dependent DNA binding activity in crude extracts that specifically recognizes *attTn7*. This binding activity is *tnsD*-dependent and does not require any other *tns* gene function. Our results provide biochemical evidence that the unique requirement for *tnsD* in *attTn7* transposition reflects a direct role of TnsD in *attTn7* sequence recognition. Our findings suggest, but do not rigorously establish, that TnsD is itself the active DNA binding protein in our assays.

We have localized the *tnsD*-dependent binding site and have asked whether these same DNA sequences are sufficient for *attTn7* target activity. We observe a correlation between *tnsD*-dependent *attTn7* sequence recognition *in vitro* and *attTn7* target activity *in vivo* (see Fig. 1 for *attTn7* organization). The *tnsD*-dependent binding activity recognizes an *attTn7* segment extending from position +10 to position +64; this segment provides wild-type transposition activity. Moreover, the binding activity is sensitive to a mutation at position +37, a nucleotide previously identified as critical for *attTn7* target activity. In addition, we have shown that the DNA binding protein recognizes the 28-bp *attTn7* (+28 to +55) yet fails to bind the slightly smaller *attTn7* (+32 to +55). We conclude that nucleotides between +28 and +31 are essential for *attTn7* sequence recognition. This interpretation is supported by the failure of the *attTn7* (+32 to +55) plasmid to function as a Tn7 transposition target *in vivo* and the ability of Tn7 to transpose to *attTn7* (+28 to +55) plasmids in a site- and orientation-specific manner. We know from previous *in vivo* studies (8) that some nucleotides between positions +39 and +64 are also essential for target activity. Interestingly, a *dcm* methylation site is encoded in this region so that the cytosines at positions +47 and +49 are methylated *in vivo*. Because there is no detectable difference in *tnsD*-dependent binding with a synthetic oligonucleotide versus DNA recovered from cells, we conclude that methylation at these positions is not required for *attTn7* recognition. It will be interesting to determine what role, if any, *dcm* methylation might have in the specificity and frequency of *attTn7* insertion *in vivo*.

Although the 28-bp *attTn7* (+28 to +55) does provide *attTn7* target activity *in vivo*, the reduced transposition frequency (down by a factor of ≈100 from wild type) suggests that some target information is lacking. Several possible explanations can be considered. Required sequences may include additional specific nucleotides to either side of the +28 to +55 *attTn7* region that contribute directly to the *tnsD*-dependent recognition site. If this is the case, the effect of such nucleotides on the protein–DNA interaction is not detectable under the conditions of our gel assays. The decreased transposition frequency may reflect the deletion of a host protein binding site. We have observed that larger *attTn7* segments contain sequences recognized by a heat-stable *E. coli* protein. The loss of this binding site correlates with decreased target activity. Several transposons and site-specific recombinases utilize small heat-stable host proteins as accessory factors (21–25). To date, there is no genetic evidence suggesting a role for host proteins in Tn7 transposition (1). Alternatively, nucleotides that are not part of a protein recognition site may affect transposition. Such nucleotides could, for example, contribute to the construction, or stability, of a higher-order structure required for transposition or could affect the efficiency at which the DNA strands are cleaved at the point of insertion.

These studies have confirmed and extended our earlier view of *attTn7*: sequences essential for activity are distinct and separate from the actual point of insertion. This feature is one of several that distinguishes Tn7 site-specific transposition from other highly site- and orientation-specific mobile elements that insert at their respective bacterial attachment sites via a conservative site-specific recombination mechanism. This latter mechanism has been elegantly determined in the case of phage λ (23, 25) and a similar mechanism has been suggested for the *Staphylococcus* transposon Tn*554* (26). The bacterial attachment site of λ is small and consists of sequences immediately flanking both sides of the cross-over point (27). λ (23, 25) and Tn*554* (28) each contain a short DNA segment found in both the element and the target sequence. No such homology is observed between the ends of Tn7 and its chromosomal insertion site (5, 7, 8). The displacement of the *tnsD*-dependent recognition site relative to the point of insertion suggests that the mechanism by which Tn7 is directed to its specific insertion site may have features in common with the type III restriction enzymes (29).

Note Added in Proof. In good agreement with our studies, Qadri *et al.* (30) have localized one border of functional *attTn7* to between positions +54 and +58.

We thank Karina Orle for DNA sequencing, Karyl Nakamura for preparation of the manuscript, and Sandy Johnson and Susan Michaelis for comments on the manuscript. C.S.W. was supported by a University of California, San Francisco, Chancellor's Fellowship and the work was supported by an award from the National Institutes of Health to N.L.C.

1. Craig, N. (1989) in *Mobile DNA*, eds. Berg, D. & Howe, M. (Am. Soc. Microbiol., Washington, DC), pp. 211–225.
2. Barth, P., Datta, N., Hedges, R. & Grinter, N. (1976) *J. Bacteriol.* **125**, 800–810.
3. Lichtenstein, C. & Brenner, S. (1981) *Mol. Gen. Genet.* **183**, 380–387.
4. Walker, J., Gay, N., Saraste, M. & Eberle, A. (1984) *Biochem. J.* **224**, 799–815.
5. Gay, N., Tybulewicz, V. & Walker, J. (1986) *Biochem. J.* **234**, 111–117.
6. Gringauz, E., Orle, K., Waddell, C. & Craig, N. (1988) *J. Bacteriol.* **170**, 2832–2840.
7. Lichtenstein, C. & Brenner, S. (1982) *Nature (London)* **297**, 601–603.
8. McKown, R., Orle, K., Chen, T. & Craig, N. (1988) *J. Bacteriol.* **170**, 352–358.
9. Rogers, M., Ekaterinaki, N., Nimmo, E. & Sherratt, D. (1986) *Mol. Gen. Genet.* **205**, 550–556.
10. Waddell, C. & Craig, N. (1988) *Genes Dev.* **2**, 137–149.
11. McKown, R., Waddell, C., Arciszewska, L. & Craig, N. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 7807–7811.
12. Miller, J. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
13. Maniatis, T., Fritsch, E. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
14. Struhl, K. (1983) *Gene* **26**, 231–242.
15. Bradford, M. (1976) *Anal. Biochem.* **72**, 248–254.
16. Singh, H., Sen, R., Baltimore, D. & Sharp, P. (1986) *Nature (London)* **319**, 154–158.
17. Hanahan, D. (1983) *J. Mol. Biol.* **166**, 557–580.
18. Garner, M. & Revzin, A. (1981) *Nucleic Acids Res.* **9**, 3047–3060.
19. Fried, M. & Crothers, D. (1981) *Nucleic Acids Res.* **9**, 6505–6525.
20. Pabo, C. & Sauer, R. (1984) *Annu. Rev. Biochem.* **53**, 293–321.
21. Craigie, R., Arndt-Jovin, D. & Mizuuchi, K. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7570–7574.
22. Morisato, D. & Kleckner, N. (1987) *Cell* **51**, 101–111.
23. Weisberg, R. & Landy, A. (1983) in *Lambda II*, eds. Hendrix, R., Roberts, J., Stahl, F. & Weisberg, R. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 211–250.
24. Johnson, R. & Simon, M. (1987) *Trends Genet.* **3**, 262–267.
25. Landy, A. (1989) *Annu. Rev. Biochem.*, in press.
26. Bastos, M. & Murphy, E. (1988) *EMBO J.* **7**, 2935–2941.
27. Mizuuchi, M. & Mizuuchi, K. (1985) *Nucleic Acids Res.* **13**, 1193–1208.
28. Murphy, E. & Lofdahl, S. (1984) *Nature (London)* **307**, 292–294.
29. Yuan, R. (1981) *Annu. Rev. Biochem.* **50**, 285–315.
30. Qadri, M., Flores, C., Davis, A. & Lichtenstein, C. (1989) *J. Mol. Biol.*, in press.