Volume 13 Number 19 1985 Nucleic Acids Research

Nucleotide sequence of the transposon Tn*7* gene encoding an aminoglycoside-modifying enzyme, 3"(9)-*O*-nucleotidyltransferase

Mary E.Fling, Jan Kopf and Cindy Richards

Wellcome Research Laboratories, Research Triangle Park, NC 27709, USA

Received 6 May 1985; Revised and Accepted 5 September 1985

ABSTRACT

The nucleotide sequence of a transposon Tn*7* DNA fragment encoding a 3"(9)-*O*-nucleotidyltransferase, an aminoglycoside-modifying enzyme, which mediates bacterial resistance to spectinomycin and streptomycin, was determined. The *aad*A structural gene was 786 bases long and predicted a polypeptide of 262 amino acids with a calculated molecular weight of 29,207. Comparison of the DNA sequences of Tn*7* and plasmid R538-1 indicated that their *aad*A genes were nearly identical. Comparison of the polypeptides predicted by the *aad*A genes of Tn*7* and Tn*554* indicated that the genes were related.

INTRODUCTION

Bacterial resistance to spectinomycin (Sp) and streptomycin (Sm) can be mediated by plasmid-specified genes that code for aminoglycoside-modifying enzymes (1). Streptomycin adenylyltransferase is a 3"(9)-*O*-nucleotidyltransferase, AAD(3")(9), which adenylates either the 3"-hydroxyl on the amino-hexose III ring of streptomycin or the 9-hydroxyl on the actinamine ring of spectinomycin (2). The AAD(3")(9) enzyme has been found in both Gram-negative and Gram-positive genera (3). Spectinomycin adenylyltransferase, a 9-*O*-nucleotidyltransferase, AAD(9), adenylates only the 9-hydroxyl on the actinamine ring of spectinomycin (2). The AAD(9) enzyme is the product of a gene that is present on transposon Tn*554*, which is found in *Staphylococcus aureus* (4).

The dual substrate profile of the AAD (3")(9) enzyme is common, suggesting that the *aad*A gene is widespread. One source of the *aad*A gene, which will be considered in this study, is plasmid R538-1, a relative of plasmid R100 (5). On plasmid R538-1 or R100, the *aad*A gene is linked to a sulfonamide-resistance determinant. It is disseminated as part of Tn*21* or any of several related complex transposons that are present in plasmids R100, R1, R5, R6, or NR1 (6,7,8). Another distinct source of the *aad*A gene is the transposon Tn*7*, originally isolated from the R-plasmid

© IRL Press Limited, Oxford, England.

R483 (9). On Tn*7* the *aadA* gene is linked to the Type I dihydrofolate reductase gene, which confers resistance to the antibiotic trimethoprim.

To study the *aadA* gene of transposon Tn*7*, the nucleotide sequence was determined. The nucleotide sequences and deduced polypeptides of the Tn*7*, Tn*554*, and R538-1 *aadA* genes were compared. The evolution of the *aadA* genes of both Gram-negative and Gram-positive origin is discussed.

MATERIALS AND METHODS

Bacterial Strains and Plasmids

The bacterial strain *E coli K12*, C600, containing plasmid pFE506 (10) was used to make all deletion constructions and to clone the *aadA* gene. The cloning vectors, pUC8 and pUC9, their host JM83, the bacteriophages M13mp8 and M13mp9, and their host JM103 were acquired through Bethesda Research Laboratories, Inc., Gaithersburg, MD, and propagated as previously described (11,12).

Antibiotic Sensitivity Determination

Minimal inhibitory concentrations (MIC) for Sp were determined by a standard agar dilution method (13).

Molecular Cloning, Clone Analysis, and Gel Electrophoresis

Hybrid plasmids and bacteriophages were constructed using techniques previously described (14). DNA fragments were isolated from agarose gels by electrophoresis onto NA45 paper (15).

Nucleotide Sequence Determination

The chain termination technique of DNA sequencing (16) was used. Templates were generated as previously described (14). Sequencing reagents supplied as a kit from New England Nuclear Corp., Boston, MA, were supplemented by a 17-mer primer (Pharmacia P.L. Biochemical, Milwaukee, WI) and the Klenow fragment of polymerase I (Bethesda Research Laboratories, Inc.). Buffer gradient sequencing gels were prepared as described by Biggin *et al.* (17). Compressions of the sequencing products during gel electrophoresis were resolved by substituting dITP for dGTP (18). The concentration of dITP was 50 µM in the I and A reactions, and 65 µM in the C and T reactions.

RESULTS

Cloning and Sequencing the Spectinomycin-Resistance Determinant of Tn*7*

Evidence of the linkage of the Type I dihydrofolate reductase and the *aadA* genes on the left end of Tn*7* was established by forming the

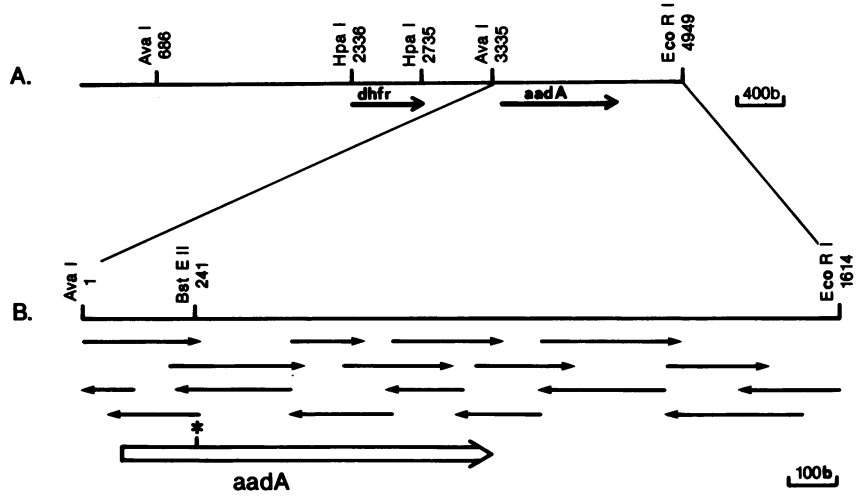

Figure 1. Restriction endonuclease map of plasmid pFE506 and nucleotide sequencing strategy. The upper line represents the 4949 bases of the left end of Tn7. The ColE1-derived vector is not shown. Arrows show the positions of the *dhfr* gene (2354-2827) and the *aadA* gene (3417-4203). The lower line represents the 1614 bp *Ava*I-*Eco*RI fragment which was sequenced (expanded scale). The closed arrows indicate the direction and extent of the dideoxy sequencing reactions. The *Bst*EII site at 241 is indicated by (*).

ColE1::Tn7 derivative, plasmid pFE506 (10), shown in Figure 1. Expression of the Sm/Sp resistance determinant was not affected by deletion of the 2.7 kilobase pair (kbp) *Ava*I fragment from plasmid pFE506. The gene was localized to the 1.6 kbp *Ava*I-*Eco*RI fragment (Figure 1, positions 3335-4949) by cloning it into pUC9, creating plasmid pFE1229. This plasmid construction conferred Sp and Sm resistance upon its host. Resistance was also seen when the fragment was in the opposite orientation to the *lac* promoter in pUC8, indicating that the gene and its regulatory sequences were intact within the cloned fragment.

The sequence data, obtained from both strands of the 1614 bp *Ava*I-*Eco*RI fragment as shown in Figure 1, are displayed in Figure 2. A single, open reading frame large enough to encompass the structural gene had an ATG initiation codon at base 82 and a TAA terminator codon at base 868. There was a stop codon upstream at base 22 in phase with the open reading frame. The polypeptide deduced from this open reading frame had a calculated molecular weight of 29,207. There are two potential initiation codons, the ATG at base 82 or the GTG at base 94. Some homology with the

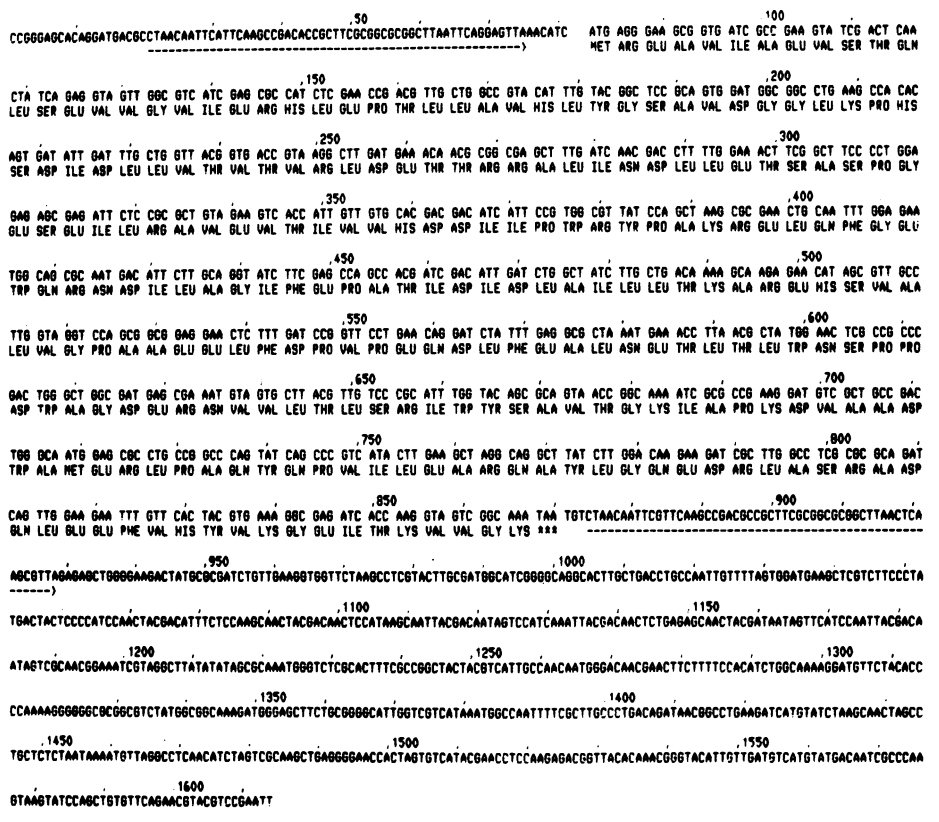

Figure 2. Nucleotide sequence of the streptomycin adenylyltransferase gene. The DNA sequence of the 1614 bp *Ava*I-*Eco*RI fragment of Tn7 is shown as the equivalent of the mRNA sequence in which U is replaced by T. The sequence is numbered from the first base of the cloned DNA and dots are positioned every 10 bp. The amino acid sequence is displayed under the DNA sequence and is translated from the possible initiation codon, ATG, (position 82) and terminated at a TAA (position 868). A second possible initiation codon is GTG at position 94. Sequences with some homology to the *E. coli* consensus promoter are 5'-GACA (base 41) and 5'-TAATTC (base 61). Possible Shine-Dalgarno sequences are 5'-GGAG (base 69) and 5'-GAGG (base 84). The dashed arrows indicate 2 directly-repeated DNA sequences, one proximal (bases 21-75) to and the other distal (bases 874-928) to the proposed structural gene. The directly-repeated DNA sequences are also palindromic sequences centered about base 49 or base 903.

E. coli consensus -35 and -10 promoter regions (19,20) was detected at base 41 (GACA) and base 61 (TAATTC). Sequences complementary to the 3' end of 16S rRNA (21) were GGAG at base 69 and GAGG at base 84. Neither of these potential Shine-Dalgarno sequences was optimally spaced from the putative promoter or from the two possible initiation codons (22).

To verify the reading frame assignment, a lesion was introduced into the putative *aadA* gene *in vitro* using a *Bst*EII cleavage site at position 241 (Figure 1). Plasmid pFE1229 DNA was restricted with *Bst*EII, the ends were filled in with the Klenow fragment of DNA polymerase I, and the flush ends were religated. Strain JM83, transformed with the mutated DNA, was no longer resistant to Sp or Sm. The Sp MIC was 10 µg/ml, the same as for sensitive *E. coli*. The MIC of the strain containing the wild type determinant on a fragment in pUC9 (pFE1229) was >1000 µg/ml. When the wild type insert was oriented in the opposite direction, the MIC was reduced to 500 µg/ml. Although the cloned fragment apparently contained its own promoter, expression was increased when the *aadA* promoter was placed in tandem with the *lac* promoter.

Sequencing products were occasionally compressed during gel electrophoresis, causing reading ambiguity (Figure 3). Compressions first appeared where the single-stranded DNA products extended into palindromic sequences with a high G+C content. Several troublesome areas had compressions within the same region on both strands. Increasing the temperature at which the gel was run or adding formamide to the sequencing gel had no effect. A uniformly-spaced DNA ladder was produced by substituting dITP for dGTP in the four sequence reactions as shown in Figure 3 for the complement of bases 605-613. Although the G+C content of these compressed regions was about 87-90%, the overall G+C content of the gene was 52.5%, and that of the entire 1614 bp fragment was 50.9%. The nearby fragment containing the dihydrofolate reductase gene had an average G+C composition of 40.5% (14).

One striking feature of the primary sequence was the repeated DNA at each end of the putative structural gene. The first of two long, directly-repeated DNA sequences was at bases 21 through 75 (Figure 2). The second repeat began four bases distal to the putative gene at base 874 and ended at base 928. Of the 54 bases in the repeated DNA, 49 were identical.

Each repeat was a palindromic sequence, centered at base 49 for the first repeat, and at base 902 for the second repeat. The potential hairpins included all 54 bases, and they were similar to each other. For the proximal sequence, the hairpin had a stem of 22 base-paired residues and 10 unpaired bases, with a loop of three bases. Of five base differences in the second repeat, four are substitutions of G or C for A or T showing a net increase of three GC base pairs. Theoretically, these

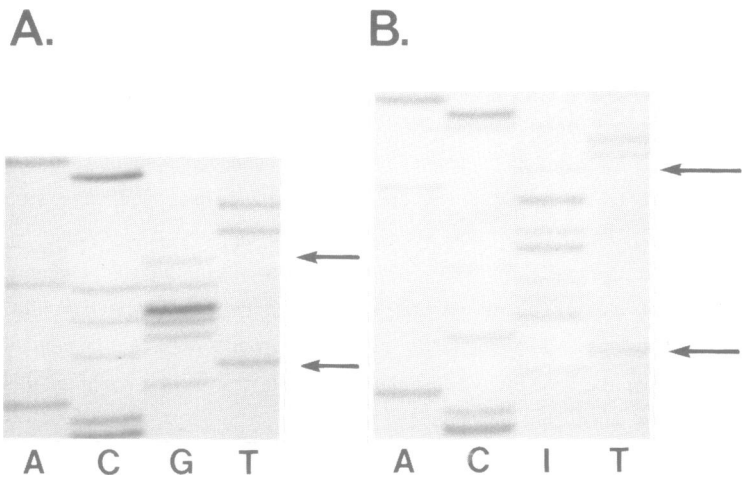

Figure 3. Autoradiographs of DNA sequencing gels showing compression and dITP incorporation. The sequence ladder shows the complement of bases 598-618. A: Sequence reactions were performed as usual and separated by electrophoresis on 0.4mm, 8.0% denaturing polyacrylamide gels. The compressed region, indicated by the arrows, can be read as 5'-TCGGCGZAG (Z = C or G). B: Sequence reactions were performed substituting dITP for dGTP. The A and I reactions contained 50μM dITP and the C and T reactions contained 65μM dITP. The region indicated by the arrows is no longer compressed. It can be read as 5'-TCGGGCGGCGAG.

changes would stabilize the stem of the gene-distal hairpin structure. The sequencing products from the second, repeated sequence migrated anomalously, in the same manner as those shown in Figure 3, suggesting that intrastrand DNA bonding had occurred.

Comparison of the *aadA* genes of Tn7, Tn*554* and plasmid R538-1

The amino acid sequences predicted by the Tn7 and the Tn*554* (23) *aadA* genes were compared by introducing two gaps in each sequence to maximize homology (Figure 4). In this optimal alignment, 93 of the 262 Tn7 residues were identical to Tn*554* residues, resulting in a correspondence of 35% identity. Although the overall amino acid homology was 35%, three regions representing almost half of the residues in the polypeptides were 50-60% homologous. The homology was 50% for region A, residues 31-70; 55% for region B, residues 98-145; and 66% for region C, residues 181-212. Nucleotide sequence homology was as great as amino acid sequence homology in the regions designated A, B, and C.

The nucleotide sequence of the fragment containing the Tn7 *aadA* gene was compared to the sequence of a fragment containing the plasmid R538-1

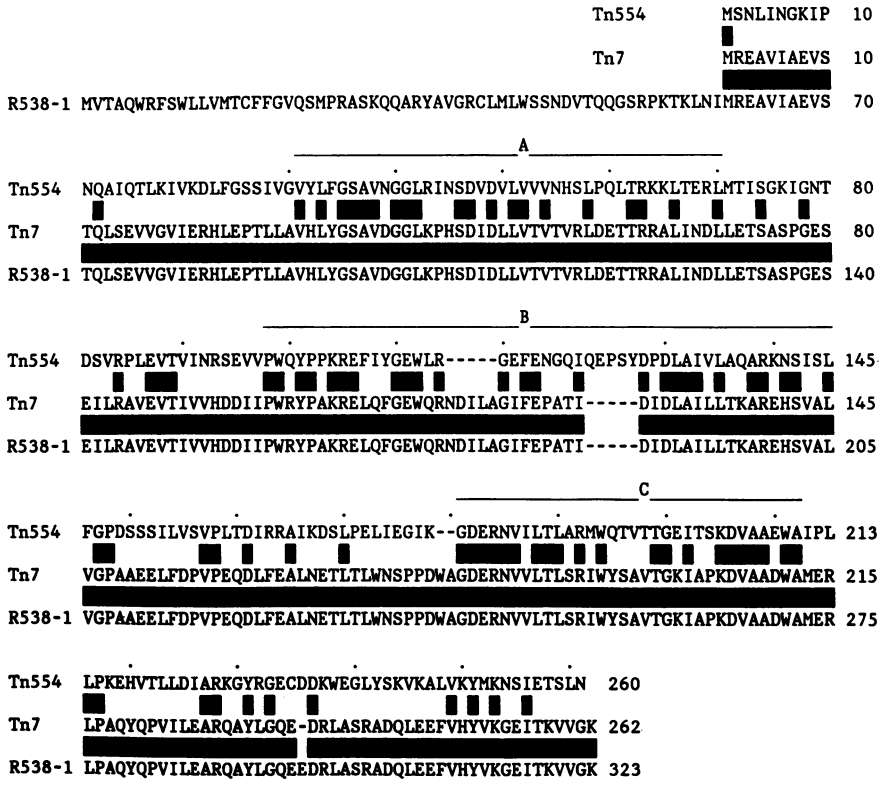

Figure 4. Comparison of amino acid sequences. The amino acid sequence deduced from the Tn7 *aad*A gene was compared to those deduced from the Tn*554* and R538-1 *aad*A genes. The sequences were aligned with two gaps in the Tn7 sequence, two gaps in the Tn*554*, and one gap in the R538-1 sequence to increase homology. Dots positioned above the aligned residues correspond to every 10 Tn7 residues. Identical residues are denoted with a (■). Three regions of greater homology, A, B, and C, are indicated by overlining. No penalties were imposed for gaps.

*aad*A gene, which was cloned and sequenced by Hollingshead and Vapnek (24). The Tn7 bases 71-930, which included all of the *aad*A structural gene, corresponded almost perfectly to bases 391-1253 of the R538-1 sequence. Outside of this region, there was no significant sequence homology. Although the DNA sequences of R538-1 and Tn7 were nearly identical over the entire Tn7*aad*A gene, the R538-1 open reading frame described by Hollingshead and Vapnek begins 180 bases upstream of the proposed Tn7 initiation codon, ATG. The cloned Tn7 DNA did not extend this far proximal to the structural gene, and did not correspond to the R538-1 DNA sequence. The R538-1 *aad*A gene predicted a polypeptide nearly identical

to that of the Tn7*aadA* gene with 60 additional amino acids at the amino terminus (Figure 4). The calculated molecular weight of the R538-1 polypeptide was 36,048, considerably larger than the 27,500 estimated from maxicell proteins in SDS-polyacrylamide gels (24). Within the structural genes there were three other differences. Two of these differences were single base changes that did not change the specified amino acid (Figure 4, Tn7 residues 227, 249). The other difference was an additional triplet in the R538-1 sequence coding for glutamate (Figure 4, R538-1 residue 296). Introducing a single residue gap in the Tn7 sequence to accommodate this extra amino acid increased the homology between the Tn7 and the Tn*554* proteins.

The region of identity shared by the sequences of Tn7 and plasmid R538-1 included 11 bases proximal to and 62 bases distal to the Tn7*aadA* structural gene. The identity ceased at, or near, each of the 3' ends of the direct repeats, so that the R538-1 gene, as proposed here, would include a few bases of the repeated sequence proximal to and all of the bases of the repeated DNA distal to the Tn7 gene (corresponding to Tn7 bases 72-75, 874-928). The distal repeat of R538-1 differed from that of Tn7 at a stretch of 11 Gs and Cs (Tn7, bases 903-913). These bases were compressed in our gels, and after dITP was incorporated, the sequence of 5'-GGCGCGGC was read where Hollingshead and Vapnek had read 5'-CGGCG.

DISCUSSION

The *aadA* gene was located 700 bases distal to the dihydrofolate reductase gene on Tn7, and the nucleotide sequence was determined. The calculated mass of the AAD(3")(9) enzyme deduced from this sequence was 29,207 daltons, in excellent agreement with the mass of 30,000 daltons estimated from SDS-polyacrylamide gel electrophoresis of proteins synthesized in minicells containing the cloned gene (10). A lesion in the putative gene rendered the host cells sensitive to spectinomycin and streptomycin.

The primary sequence of the Tn7-borne *aadA* gene might be described as a discrete DNA sequence bounded by direct or inverted repeats. The sequence of the long palindromes bounding the Tn7 *aadA* gene was not obviously related to any of the known sequences at the ends of transposons or IS elements. However, the structural gene with its boundary repeats is reminiscent of the structures described by Heffron as degenerate transposons (6). These elements may encode phenotypic traits, but require

transposition functions supplied by another mobile element in order to transpose. Although the structural genes of the Tn*7* and R538-1 *aadA* loci have diverged only slightly, considerable divergence in the R538-1 region corresponding to the proximal, directly-repeated Tn*7* DNA has occurred. This region corresponded to bases 21-69 of the postulated hairpin. There was no base-pairing potential remaining in the R538-1 sequence. The extensive changes in this region may be a consequence of pressure to devise a different, functional promoter in the environment of plasmid R538-1 or an ancestral plasmid. Interestingly, the distal, repeated DNA has remained nearly intact.

We concluded from our comparison of the nucleotide sequences of Tn*7* and plasmid R538-1 that their *aadA* genes were nearly identical. Although the R538-1 protein may be longer than its Tn*7* or Tn*554* counterparts, we propose that its amino terminus is at met-61 which corresponds to met-1 of the Tn*7* protein. The longer R538-1 reading frame was rationalized on the basis of the supposed periplasmic location of the R538-1 AAD(3")(9) enzyme (24). Although the early work localizing the aminoglycoside-modifying enzymes supported a periplasmic location, more recent work (25) suggests that the previously used osmotic shock procedures released substantial amounts of some cytoplasmic proteins. Using a more gentle spheroplasting method of localizing cellular proteins, the AAD(3")(9) enzyme was shown to be retained by the spheroplasts, suggesting that it is a cytoplasmic, rather than periplasmic, protein (26). Thus, the AAD(3")(9) enzyme would not be expected to have a signal sequence as postulated (24). The inferred R538-1 sequence predicted a significantly larger gene product than was determined experimentally (24). Since the assignment of the R538-1 protein N-terminus was largely dependent on the disputed hypothesis that the gene product was exported, it is realistic to presume that the N-terminus coincides with those of Tn*7* and Tn*554*. In fact, the position and size of the Tn*7* structural gene was nearly unambiguous, since a reading frame of the correct size was preceded by a stop codon. The Tn*554* structural gene was also unambiguously located using a strong Shine-Dalgarno sequence with excellent positioning to establish the 5' end and transcriptional analysis to establish the 3' end (23). The excellent correspondence between Tn*554*- and Tn*7*-deduced polypeptides also provides convincing support for designating Met 61 as the N-terminus of the R538-1 protein.

The extensive conservation between the *aadA* genes of Tn*7* and plasmid

R538-1 was striking because the genes have been residing on unrelated replicons subject to different selective pressure. The extensive homology contrasts with the variation found among other classes of enzymes in Gram-negative organisms, such as the TEM β-lactamases, the gentamycin acetyltransferases, and the 3'-aminoglycoside phosphotransferases (3'-APH). The TEM β-lactamases on different Tn*3* transposons have diverged enough to form two classes of electrophoretically different enzymes (27). Similarly, the genes for the gentamycin acetyltransferases, AAC3-1 and AAC3-3, have diverged enough to be phenotypically different with respect to substrate profile (28). Yet, they reside on similar, if not identical, transposons or plasmids. The 3'-APH enzymes, which modify antibiotics in the kanamycin and neomycin classes (3), are encoded by bacterial transposons and the neomycin-producer *Streptomyces fradiae* (29). They have diverged considerably to form three distinct classes, Types I, II, and III, based on substrate profiles, yet they probably have a common evolutionary origin (30,31,32).

Comparison of the amino acid sequences predicted by the *aad*A genes of Tn*7* and Tn*554* showed that the proteins were related. The extent of homology was slightly greater than that seen among the bacterial dihydrofolate reductases. Even though the sequence identity between the *E. coli* and *L. casei* dihydrofolate reductases is only about 27%, x-ray crystallographic studies have shown that these two bacterial enzymes have nearly identical structural backbones (33). Thus, they are highly conserved structurally without extensive sequence identity. The considerable homology between the Tn*7* and the Tn*554* *aad*A gene products indicated that they, too, were structurally conserved.

The homology between the Gram-negative AAD(3")(9) enzymes and the Gram-positive AAD(9) enzyme may reflect a requirement to retain the functionally important regions that these enzymes share. The AAD(3")(9) enzyme is also related to the AAD(2") enzyme, which is able to adenylylate antibiotics in the kanamycin and gentamycin classes, but not Sm or Sp (3). Small fragments of the cloned gene (*aad*B) encoding the AAD(2") enzyme were hybridized to plasmid DNAs coding for aminoglycoside resistance (34). One of two small fragments from within the gene was specific for the *aad*B gene determining AAD(2"), but the other could hybridize to an *aad*A gene as well. Neither fragment could hybridize to fragments containing the genes for AAD(6') or AAD(4') enzymes, representatives of two 2 classes of adenylyltransferases that are found in Gram-positive organisms (3). Since

the hybridization conditions were stringent, 80-90% homology was required for a positive result. Thus, the possibility that the AAD(6') or AAD(4') enzymes are related to the AAD(3")(9) enzyme remains. Based on the comparison of the *aad*A genes, it is reasonable to conclude that conservation of a functionally important structural domain, possibly that involved with adenylylation, is required. The variation that occurs may permit binding of different antibiotic substrates while preserving the tertiary structure necessary for function. Even more versatile nucleotidyltransferases may emerge as a result of selective pressure from current or new aminoglycosides. Dissemination should readily occur because all of the known *aad*A and *aad*B genes are present on transposons.

ACKNOWLEDGEMENTS

We thank E. Murphy for communicating her sequence data prior to publication and L. Elwell for helpful advice and discussions.

REFERENCES

1. Benveniste, R., Yamada, T., and Davies, J. (1970) Infect. and Immun. 1, 109-119.
2. Davies, J.E. and Benveniste, R.E. (1974) Ann. N.Y. Acad. Sci., 235, 130-136.
3. Davies, J. and Smith, D.I. (1978) Annu. Rev. Microbiol. 32, 469-518.
4. Murphy, E., Phillips, S., Edelman, I., and Novick, R.P. (1981) Plasmid 5, 292-305.
5. Alton, N.K. and Vapnek, D. (1978) Plasmid 1, 388-404.
6. Heffron, F. (1983) in Mobile Genetic Elements, Shapiro, J.A. Ed., pp. 223-260, Academic Press, New York.
7. Kopecko, D.J. (1980) in Plasmids and Transposons, Stuttard,C. and Rozee, K.R. Eds., p. 195, Academic Press, New York.
8. Tanaka, M., Yamamoto, T. and Sawai, T. (1983) J. Bacteriol. 153, 1432-1438.
9. Barth, P.T., Datta, N., Hedges, R.W. and Grinter, N.J. (1976) J. Bacteriol. 125, 800-810.
10. Fling, M.E. and Elwell, L.P. (1980) J. Bacteriol. 141, 779-785.
11. Vieira, J. and Messing, J. (1982) Gene 19, 259-268.
12. Messing, J., Crea, R. and Seeburg, P.H. (1981) Nucleic Acids Res. 9, 309-321.
13. National Committee for Clinical Laboratory Standards (1980) pp. 5-10, National Committee for Clinical Laboratory Standards, Villanova, Pennsylvania.
14. Fling, M.E. and Richards, C. (1983) Nucleic Acids Res. 11, 5147-5158.
15. Lazardi, P.M., Binder, R. and Short, S.A. (1984) Gene Anal. Tech. 1, 22-28.
16. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
17. Biggin, M.D., Gibson, T.J. and Hong, G.F. (1983) Proc. Natl. Acad. Sci. 80, 3963-3965.
18. Sanger, F., Coulson, A.R., Hong, G.F., Hill, D.F. and Petersen, G.B. (1982) J. Mol. Biol. 162, 729-773.
19. Rosenberg, M. and Court, D. (1979) Annu. Rev. Genet. 13, 319-353.

20. Hawley, D.K. and McClure, W.R. (1983) Nucleic Acids Res. 11, 2237-2255.
21. Shine, J. and Dalgarno, L. (1974) Proc. Natl. Acad. Sci. USA 71, 1342-1346.
22. Kozak, M. (1983) Microbiol. Rev. 47, 1-45.
23. Murphy, E. (1985) Mol. Gen. Genet. in press.
24. Hollingshead, S. and Vapnek, D. (1985) Plasmid 13, 17-30.
25. Perlin, M.H. and Lerner, S.A. (1981) J. Bacteriol. 147, 320-325.
26. Dickie, P., Bryan, L.E. and Pickard, M.A. (1978) Antimicrob. Agents Chemother. 14, 569-580.
27. Matthew, M. and Hedges, R.W. (1976) J. Bacteriol. 125, 713-718.
28. Rubens, C.E., McNeill, W.F. and Farrar, J.R. (1979) J. Bacteriol. 140, 713-719.
29. Thompson, C.J. and Gray, G.S. (1983) Proc. Natl. Acad. Sci. 80, 5190-5194.
30. Trieu-Cuot, P. and Courvalin, P. (1983) Gene 23, 331-341.
31. Beck, E., Ludwig, G., Auerswald, E.A., Reiss, B. and Schaller, H. (1982) Gene 19, 327-336.
32. Oka, A., Sugisaki, H. and Takanami, M. (1981) J. Mol. Biol. 147, 217-226.
33. Volz, K.W., Matthews, D.A., Alden, R.A., Freer, S.T., Hansch, C., Kaufman, B.T. and Kraut, J. (1982) J. Biol. Chem. 257, 2528-2536.
34. Tenover, F.C., Gootz, T.D., Gordon, K.P., Tompkins, L.S., Young, S.A., and Plorde, J.J. (1984) J. Infect. Dis. 150, 678-687.