THE DNA SEQUENCE OF AN IS*1*-FLANKED TRANSPOSON CODING FOR RESISTANCE TO CHLORAMPHENICOL AND FUSIDIC ACID

Roberto MARCOLI, Shigeru IIDA and Thomas A. BICKLE
Microbiology Department, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Received 12 November 1979

1. Introduction

We have isolated a bacteriophage P1Cm carrying a chloramphenicol transposon derived from the R plasmid NR1 (=R100) [1–3]. The transposon, Tn*cam*-204, consists of about 900 basepairs of DNA flanked by direct repeats of IS*1* and it codes for resistance to both chloramphenicol and fusidic acid. We decided to determine the DNA sequence of this transposon for two reasons:

(i) The transposon appeared too small to code for the chloramphenicol resistance gene product ($22–24 \times 10^3$ mol. wt [4]), a protein for fusidic acid resistance and a putative 'transposase' [5] which would be involved in transposition.

(ii) Since the expression of chloramphenicol resistance in *Enterobacteriaceae* is under the control of catabolite repression [6], the sequence of the non-coding regions of the DNA should contain a catabolite repressor binding site (CAP site) which could be compared with the 3 published CAP sites controlling the *lac, gal* and *ara* operons of *Escherichia coli* [7–9].

The results indicate that:

(a) A gene for a putative 'transposase' does not exist;

(b) The gene(s) for resistance to chloramphenicol and fusidic acid are either identical or overlap in the same reading frame;

(c) A probable CAP site was found ~120 basepairs preceding the initiation of transcription.

2. Materials and methods

DNA of phage P1Cm204 carrying the transposon Tn*cam*204 [2] was digested with *Pst*I and the digest was cloned in the single *Pst*I site of the plasmid pBR322 [10] selecting for chloramphenicol resistance. Two plasmids, pShI41 and pShI44 in which the *cam* segment was inserted in opposite orientation were used for the sequence analysis. The results obtained with both plasmids were the same. DNA sequencing was done using minor modifications to the method in [11]. Reaction products were separated on either 8% or 20% polyacrylamide gels 0.5 mm thick at 50 V/cm.

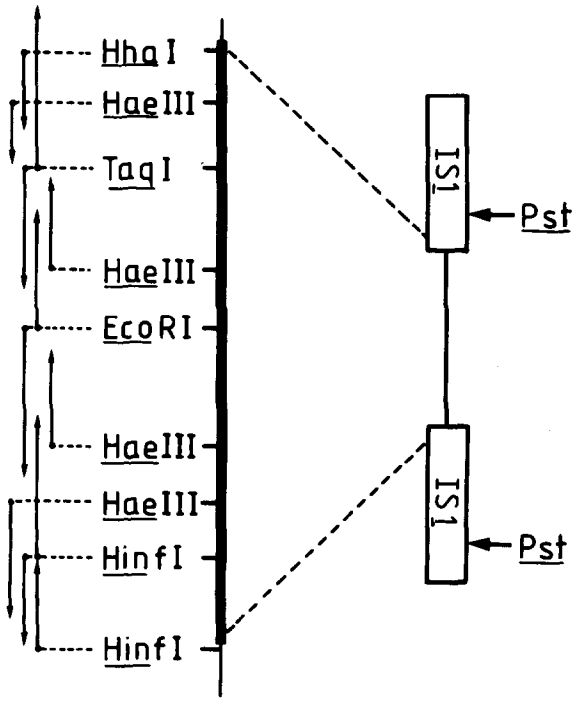

Fig.1. The strategy for sequencing Tn*cam*204. The restriction map shows only those sites that were used for sequencing. The arrows show the extent of the sequence determined from each site.

3. Results

The strategy employed in the determination of the sequence of Tn*cam*204 is shown in fig.1 and the sequence itself in fig.2. The DNA between the two IS*1* elements contains 921 basepairs. About 90% of the sequence has been determined in both strands. This was important because the presence of 5-methylcytosine residues and secondary structure in the DNA led to sequencing artifacts that only became apparent when the sequence from the two strands was compared. A sequencing gel showing some of these artifacts is shown in fig.3.

A computer analysis of the sequence, kindly conducted by Dr J. Shepherd, showed that only one protein >50 amino acids could be coded by this DNA. The amino acid sequence of this protein read from the DNA sequence agrees well with that of chloramphenicol transacetylase determined by Dr W. V. Shaw and his collaborators (personal communication). This protein sequence is also shown in fig.2.

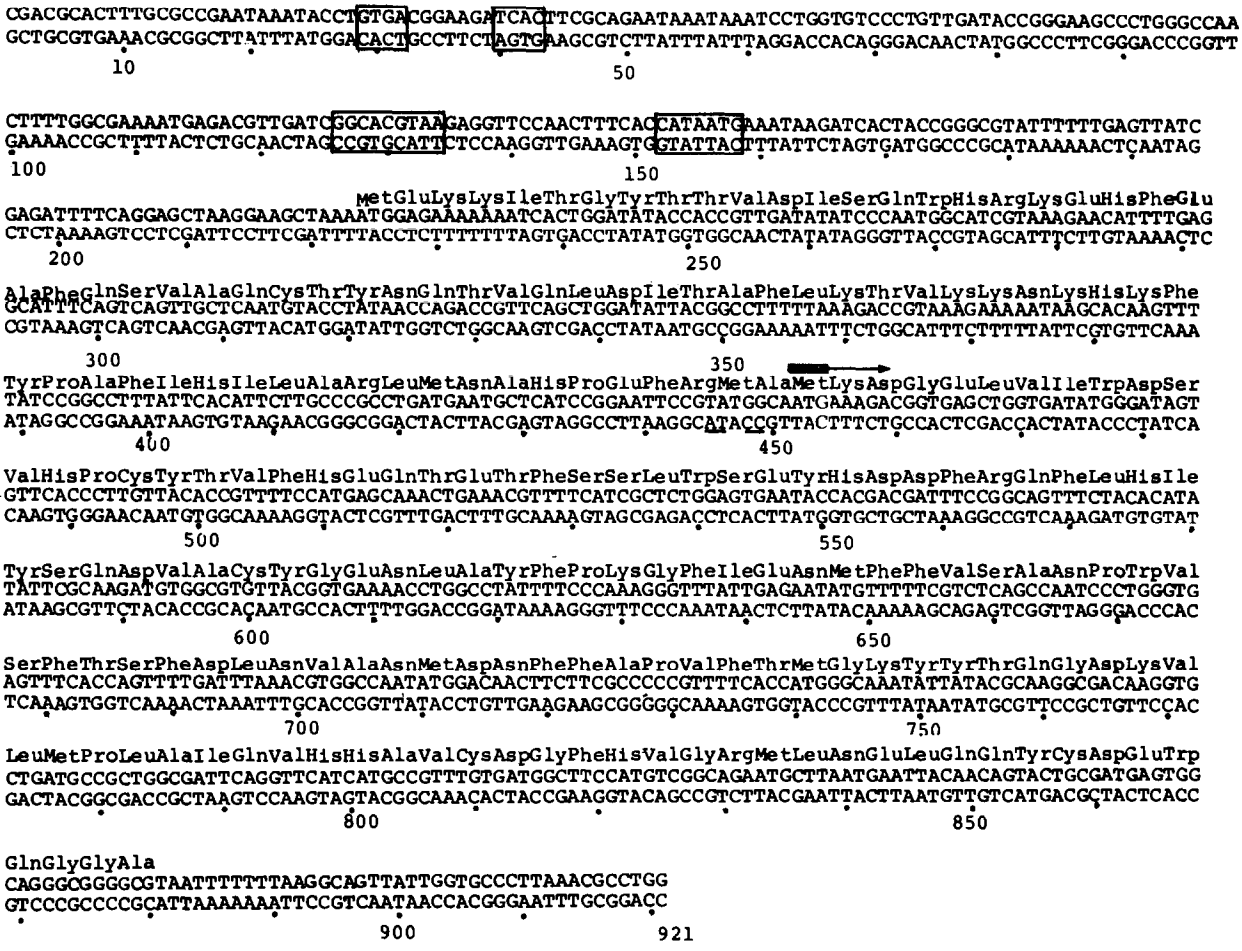

Fig.2. The DNA sequence of Tn*cam*204. The sequence shown is that of the DNA between the two IS*1* elements. The limits were determined by sequencing into the known sequence of IS*1* [17]. The protein sequence of the CAT protein is also shown. The positions of the probable CAP binding site beginning at position 29 and the promoter between positions 126–158 are indicated by boxes. The possible internal start of a protein for fusidic acid resistance discussed in the text is shown by the arrow beginning at position 450 and the nucleotide 5′ to this that show complementarity with the 3′-end of 16 S ribosomal RNA are underlined.

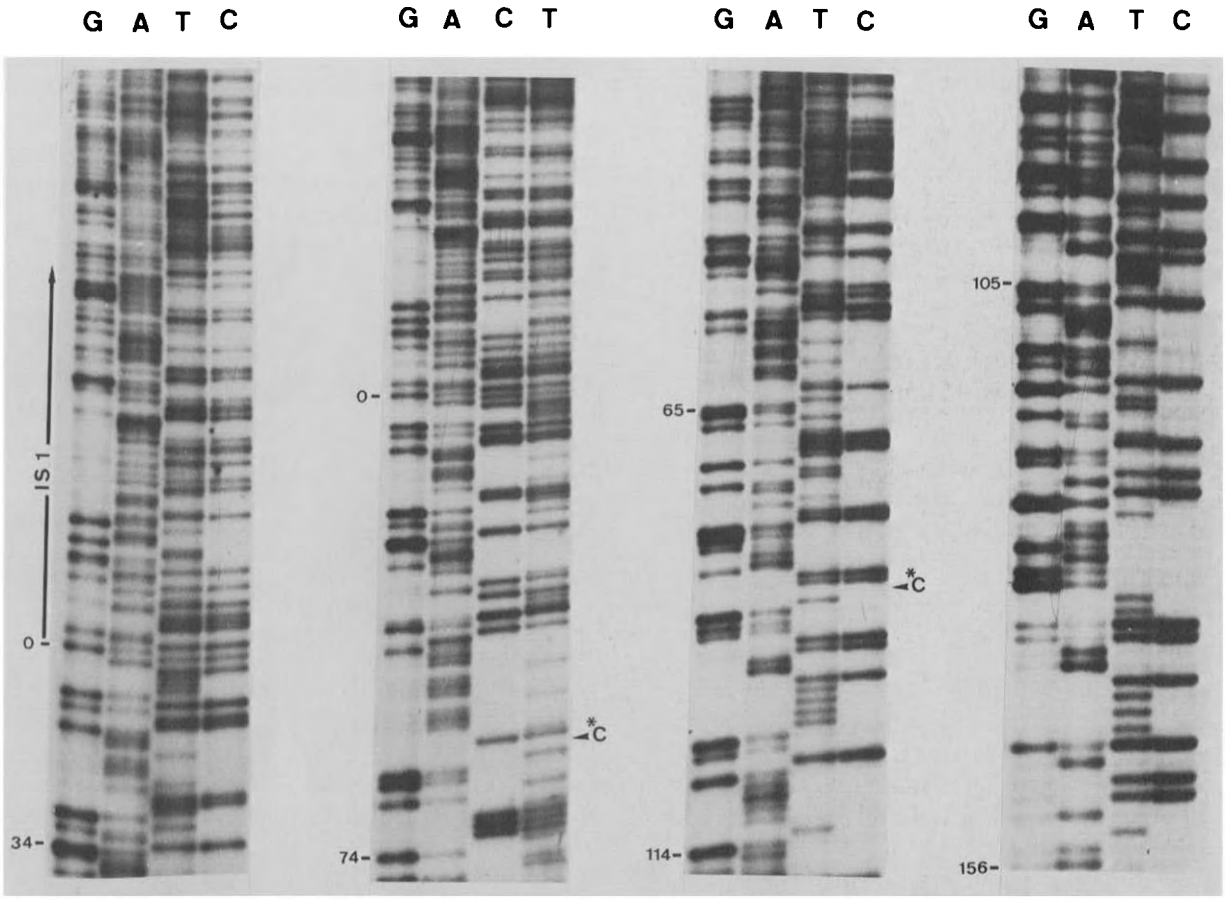

Fig.3. Part of an 8% polyacrylamide sequence gel. The DNA was labelled by repair synthesis at the *Taq*I site at position 194. The sequence shown begins at position 156 and continues into the IS*1* element [17]. The symbol *C shows the position of 5-methyl cytosine residues which lead to a gap in the sequencing ladder.

The most likely promoter for the expression of this gene is shown boxed in at positions 126 and 152. The Pribnow box of this promoter has 6 out of 7 basepairs in common with that of the bacteriophage $\lambda\, p_R$ promoter [12], and the other boxed region has much homology with the corresponding regions of other *E. coli* promoters [13]. The conclusion that this is the promoter for the *cat* gene is strengthened by the finding of Drs S. F. J. Le Grice and H. Matzura (personal communication) that the RNA transcribed from this promoter towards the *Eco*RI site (fig.1) has exactly the size that we would predict.

We have found only one sequence in the region preceding the *cat* gene that shows any homology with other CAP binding sites. This site (shown boxed in fig.2 between positions 29 and 43) is situated some 120 basepairs 5′ to the probable initiation site for transcription. Its sequence is compared with the other published CAP sites [7–9] in fig.4.

cat	GTGACGGAAGATCAC
lac	GTGAGTTAGCTCAC
ara	GTGATTATAGACAC
gal	GTGTAAACGATTCCAC

Fig.4. A comparison of the DNA sequence of the CAP binding sites of *cat* described here, *lac* [7], *gal* [8] and *ara* [9].

4. Discussion

The DNA sequence of Tn*cam*204 that we describe here contains the gene for chloramphenicol transacetylase beginning 220 basepairs from one of the flanking IS*1* elements and extending to within 40 basepairs of the other. No other protein >50 amino acids could be coded by this sequence, including proteins whose genes might overlap with that of chloramphenicol transacetylase. Where then are the genes for resistance to fusidic acid and the putative transposase?

Functions coded by transposons that are involved in the transposition process have been demonstrated for some transposons such as Tn*3* [5] but never for those flanked by IS*1* elements. Since IS*1* itself is able to transpose, it is likely that this class of transposon translocates by virtue of the flanking IS*1* elements and host functions and that any DNA sequence flanked by direct repeats of IS*1* is a transposon ([14,15], H. Reif, personal communication).

We have analysed mutants that have simultaneously lost their resistance to both chloramphenicol and fusidic acid (S.I., unpublished) shown by DNA sequence analysis to affect the extreme C-terminal part of the chloramphenicol transacetylase (unpublished). We are thus left with two possibilities for fusidic acid resistance:

1. It might be due to a second activity of chloramphenicol transacetylase;
2. A more interesting possibility is that fusidic acid resistance might be mediated by a truncated protein containing the C-terminal 2/3rds of chloramphenicol transacetylase.

The internal Met in the nucleotide sequence (452) is preceded by a sequence with homology to the 3′-end of the 16 S ribosomal RNA (shown underlined in fig.2) [16] raising the possibility that protein synthesis might initiate at this point to produce a protein in the same reading frame as chloramphenicol transacetylase but lacking the N-terminal 77 amino acids. Experiments are under way to distinguish between these two possibilities.

Acknowledgements

We would like to thank Dr W. Arber for encouragement and advice, Drs S. F. J. le Grice and W. V. Shaw for making available their results prior to publication and Dr J. C. W. Shepherd for computer analysis. This work was supported by grants from the Swiss National Foundation for Scientific Research.

References

[1] Iida, S. and Arber, W. (1977) Mol. Gen. Genet. 153, 259–269.
[2] Arber, W., Iida, S., Jütte, H., Caspers, P., Meyer, J. and Hänni, C. (1978) Cold Spring Harbor Symp. Quant. Biol. 43, 1197–1208.
[3] Rownd, R., Nakaya, R. and Nakamura, A. (1966) J. Mol. Biol. 17, 376–393.
[4] Shaw, W. V. (1975) Methods Enzymol. 43, 737–755.
[5] Heffron, F., So, M. and McCarthy, B. J. (1978) Cold Spring Harbor Symp. Quant. Biol. 43, 1279–1285.
[6] DeCrombrugghe, B., Pastan, I., Shaw, W. V. and Rosner, J. (1973) Nature New Biol. 241, 237–239.
[7] Dickson, R. C., Abelson, J., Barnes, W. and Reznikoff, W. S. (1975) Science 187, 27–35.
[8] Musso, R., DiLauro, R., Adhya, S. and DeCrombrugghe, B. (1977) Cell 12, 847–854.
[9] Greenfield, L., Boone, T. and Wilcox, G. (1978) Proc. Natl. Acad. Sci. USA 75, 4724–4728.
[10] Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. L., Heyneker, H. L. and Boyer, H. W. (1977) Gene 2, 95–113.
[11] Maxam, A. M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560–564.
[12] Walz, A. and Pirrotta, V. (1975) Nature 254, 118–127.
[13] Gilbert, W. (1976) in: RNA Polymerase (Losick, R. and Chamberlin, M. eds) pp. 193–205, Cold Spring Harbor Lab., NY.
[14] Iida, S. and Arber, W. (1980) Mol. Gen. Genet. in press.
[15] Rosner, J. L. and Guyer, M. S. (1980) Mol. Gen. Genet. in press.
[16] Shine, J. and Dalgarno, L. (1974) Proc. Natl. Acad. Sci. USA 71, 1342–1346.
[17] Ohtsubo, H. and Ohtsubo, E. (1978) Proc. Natl. Acad. Sci. USA 75, 615–619.