Gene, 25 (1983) 83–92
Elsevier

GENE 858

Sequence homology between the tetracycline-resistance determinants of Tn*10* and pBR322

(Membrane protein; transposon; plasmid vector; recombinant DNA; *Escherichia coli*)

Toai T. Nguyen, Kathleen Postle and Kevin P. Bertrand *

Department of Microbiology, College of Medicine, University of California, Irvine, CA 92717 (U.S.A.) Tel. (714) 856-6115

(Received June 16th, 1983)
(Accepted June 24th, 1983)

SUMMARY

The Tn*10* tetracycline resistance gene, *tetA*, encodes a tetracycline-inducible protein with an apparent M_r of 36×10^3. We have determined the nucleotide sequence of the Tn*10* *tetA* gene. The extent of the *tetA* gene was determined by analysis of amino-terminal and carboxy-terminal deletion mutants. We conclude that a single Tn*10* gene, the *tetA* gene, is sufficient to confer tetracycline resistance. The predicted M_r of the *tetA* protein is 43.2×10^3. The sequence homology between the Tn*10* *tetA* gene and the pBR322 tetracycline resistance determinant (49% nucleotide homology, 44% amino acid homology) indicates that these phenotypically distinct tetracycline-resistance determinants must have evolved from a common ancestral sequence. The markedly hydrophobic character of the predicted amino acid sequences of the Tn*10* *tetA* and pBR322 *tet*-coded proteins suggests that a substantial portion of these proteins may be embedded within the cytoplasmic membrane.

INTRODUCTION

The transposon Tn*10* confers high-level tetracycline resistance in *Escherichia coli* (Foster et al., 1975) and other enteric bacteria (Kleckner et al., 1975). Active efflux of tetracycline is a major component of the resistance mechanism (Ball et al., 1980; McMurray et al., 1980); however, the involvement of additional components has also been suggested (Levy et al., 1977; Shales et al., 1980). The only Tn*10* gene product that is unequivocally implicated in the mechanism of resistance is the 36-kDal Tet protein (Coleman et al., 1981; 1983; Jorgensen et al., 1979; Zupancic et al., 1980), first described by Levy and McMurry (1974). We have designated the structural gene for this protein, *tetA*. Mutations in the *tetA* gene lead to increased intracellular accumulation of tetracycline (Coleman et al., 1983), raising the possibility that the TetA protein is directly involved in the efflux of tetracycline. Cell fractionation experiments indicate that the TetA protein is associated with the cytoplasmic membrane of *E. coli* minicells (Levy et al., 1974; 1977; Yuang et al., 1976), and presumably whole cells as well.

Expression of Tn*10* tetracycline resistance is regulated; subinhibitory levels of tetracycline induce maximal expression of resistance (Izaki et al., 1966). Synthesis of the TetA protein is likewise induced by tetracycline (Levy et al., 1974; Yang et al., 1976); regulation appears to occur at the level of *tetA* trans-

* To whom correspondence and reprint requests should be addressed.

Abbreviations: bp, base pairs; kb, kilobase pairs; RBS, ribosome binding site; SDS, sodium dodecyl sulfate; Tn, transposon.

cription initiation and involves the protein product of the *tetR* repressor gene as well as tetracycline (Beck et al., 1982; Bertrand et al., 1983a; 1983b; Hillen et al., 1982; Wray et al., 1981; Yang et al., 1976). The location and orientation of the *tetA* and *tetR* genes have been deduced by analysis of Tn*10 tet* deletion mutants (Coleman et al., 1983; Jorgensen et al., 1979; Wray et al., 1981) and *tet::lacZ* gene fusions (Beck et al., 1982; Bertrand et al., 1983b); *tetA* and *tetR* are transcribed from divergent promoters located in a regulatory region between the two structural genes (Fig. 1). The DNA sequence of the regulatory region and S1 nuclease mapping of the 5′ ends of *tetA* and *tetR* RNA indicate that the *tetA* and *tetR* promoters overlap; the *tetA* and *tetR* transcription start sites are separated by 36 bp (Bertrand et al., 1983a).

Considerable discussion has focused on whether or not there is an additional gene located between the *tetA* promoter and the *tetA* structural gene, or perhaps overlapping the *tetA* structural gene (Coleman et al., 1983; Curiale et al., 1982; Jorgensen et al., 1979). This speculation is based on several findings: (i) Tetracycline-sensitive mutations that map between the *tetA* promoter and the 3′ end of the *tetA* coding region comprise two genetic complementation groups (Coleman et al., 1983; Curiale et al., 1982). (ii) The apparent M_r of the TetA protein (36 kDal) is somewhat smaller than the coding capacity of the 1325-bp region to which the *tetA* gene has been mapped (Jorgensen et al., 1979). (iii) A 15-kDal tetracycline-inducible protein has been observed in minicells containing Tn*10* on the plasmid R222 (Yang et al., 1976). (iv) Finally, the apparent complexity of the tetracycline-resistance mechanism has suggested the involvement of multiple gene products (Levy et al., 1977; Shales et al., 1980). The results presented here eliminate the possibility that there is an additional *tet* gene located between the *tetA* promoter and the *tetA* structural gene, and lead us to conclude that the TetA protein can confer tetracycline resistance in the absence of other Tn*10*-encoded proteins.

The Tn*10* and pBR322 tetracycline-resistance determinants represent two of at least four phenotypically distinct classes of plasmid-encoded tetracycline-resistance determinants (Mendez et al., 1980); Tn*10 tet* DNA and pBR322 *tet* DNA do not cross-hybridize (Mendez et al., 1980). In spite of the widespread use of pBR322 as a cloning vector, the genetic structure of the pBR322 *tet* region, like that of the Tn*10 tet* region, has remained uncertain. The nucleotide sequence of pBR322 reported by Sutcliffe (1979) suggests that there may be two overlapping genes in the pBR322 *tet* region. However, Peden (1983) has recently reported a revision in the DNA sequence of the pBR322 *tet* region, which indicates that there is a single extended open translational reading frame in the *tet* region. Comparison of the nucleotide sequence of the Tn*10 tetA* gene presented here and the revised nucleotide sequence of the pBR322 *tet* gene shows that there is substantial sequence homology between the Tn*10* and pBR322 tetracycline-resistance determinants.

MATERIALS AND METHODS

(a) Enzymes and biochemicals

Restriction enzymes, bacterial alkaline phosphatase, polynucleotide kinase, DNA polymerase I (Klenow fragment) and T4 DNA ligase were purchased from Bethesda Resarch Laboratories. Radiochemicals were purchased from Amersham-Searle. 5a,6-Anhydrotetracycline was kindly provided by N. Belcher, Charles H. Pfizer Co.

(b) Plasmid construction

pBT107 (Fig. 1 and Moyed et al., 1983) consists of the 2790-bp Tn*10 Bgl*II fragment inserted into the *Bam*HI site of pACYC177. pBT127 and pBT137 (Fig. 1) were constructed by digesting pBT107 DNA with *Hin*dIII and *Cla*I, respectively, and religating the digestion products with T4 DNA ligase. pBT147 (Fig. 1) was constructed by digesting pBT107 DNA with *Nco*I and *Eco*RV followed by sequential incubation with DNA polymerase I (Klenow fragment) and T4 DNA ligase.

(c) Minicell analysis of plasmid proteins

Plasmids were introduced into the minicell-producing strain P678-54; and minicells were prepared, labeled with [^{35}S]methionine, and processed for electrophoresis in SDS-polyacrylamide gels as described by Moyed et al. (1983). Samples to

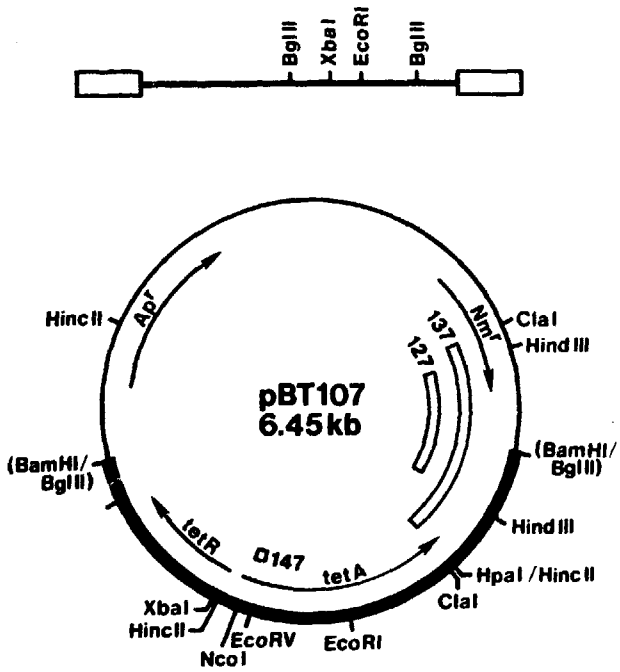

Fig. 1. Restriction maps of the Tn*10 tet* genes. Top: The transposon Tn*10* consists of 1400-bp IS*10* sequences (open bars) in inverted orientations flanking a 6400-bp central region (solid line). The location of the 2790-bp *Bgl*II restriction fragment that spans the *tetA* resistance gene and the *tetR* repressor gene is shown. Bottom: The plasmid pBT107 consists of the 2790-bp *Bgl*II fragment of Tn*10* (heavy line) inserted into the *Bam*HI site of the plasmid vector pACYC177 (thin line). The location and orientation (see arrows) of the *tetA* and *tetR* genes relative to the ampicillin resistance (Ap^r) and neomycin resistance (Nm^r) genes of pACYC177 are shown. Open bars indicate the extent of the deletions in plasmids pBT147, pBT137 and pBT127.

be induced received 1 μg/ml 5a,6-anhydrotetracycline immediately prior to labeling.

(d) Nucleotide sequence analysis

Overlapping restriction fragments were isolated from the *tetA* region of pBT107 DNA and sequenced by the Maxam and Gilbert method (1980). Fragments were either 5′-end-labeled with [γ-^{32}P]ATP and polynucleotide kinase or 3′-end-labeled with an appropriate [α-^{32}P]dNTP and DNA polymerase I (Klenow fragment). The sequence was determined on both strands and all restriction sites were overlapped (Fig. 2).

(e) Computer analysis of sequence data

A collection of programs provided by R. Staden and implemented on a DEC-10 computer was used to analyze DNA sequence data. In addition, the MOLGEN group at Stanford University generously provided access to the SEQ program. A version of the SOAP program described by Kyte and Doolittle (1982), and implemented on a DEC-10 computer and Tektronix 4662 plotter, was used to analyze the Tet protein sequences.

RESULTS

(a) Deletion analysis

The extent of the *tetA* gene was determined by constructing deletion derivatives of the plasmid

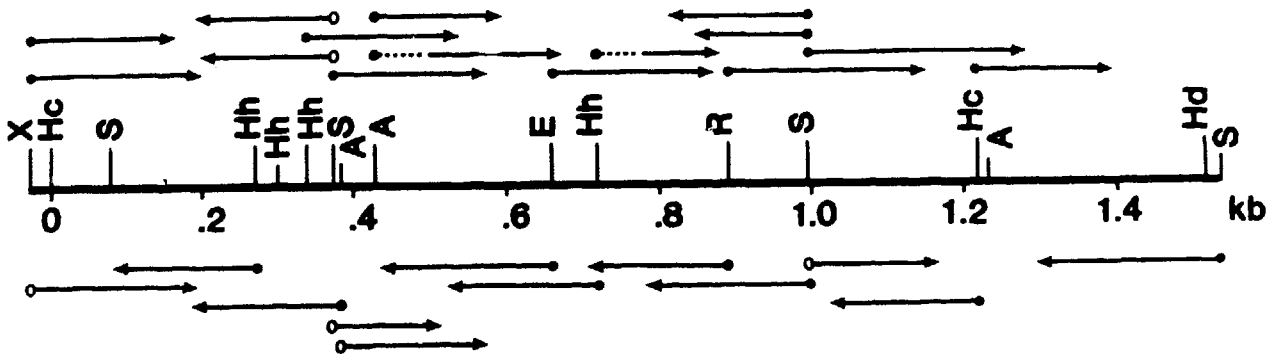

Fig. 2. Partial restriction map and sequencing strategy for the Tn*10 tetA* region. The scale is in kilobase pairs (kb) as measured from the *Hin*cII recognition site on the left. The filled circles (5′-end-labelings) and open circles (3′-end-labelings) indicate restriction sites used for sequencing: X, *Xba*I; Hc, *Hin*cII; S, *Sau*3A; Hh, *Hha*I; A, *Alu*I; E, *Eco*RI; Hd, *Hin*dIII. The arrows indicate the direction of sequencing and the length of the sequence determined.

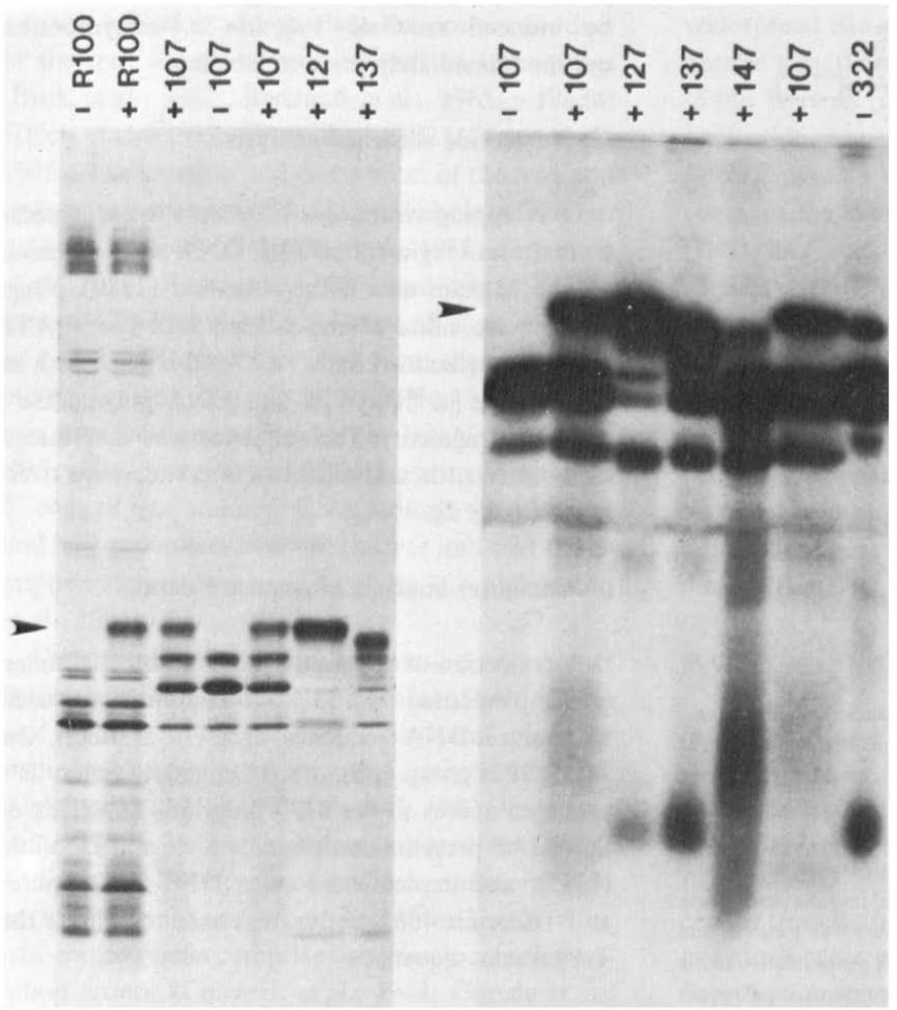

Fig. 3. Minicell analysis of Tn*10* Tet proteins. [^{35}S]methionine-labeled proteins synthesized in minicells containing the plasmids R100, pBT107, pBT127, pBT137, pBT147 and pBR322 were electrophoresed in SDS/polyacrylamide gels. Minicells were labeled either in the presence (+) or absence (-) of inducer (5a,6-anhydrotetracycline). The arrow indicates the position of wild type Tn*10* TetA protein. The gel on the left is a 5–20% polyacrylamide gradient gel; the gel on the right is a 12.5% polyacrylamide gel.

pBT107 (Fig. 1) and examining the sizes of the plasmid-encoded proteins in minicells (Fig. 3). The pBT107 *tet* region contains unique *Nco*I, *Eco*RV, *Cla*I, *Hpa*I and *Hin*dIII sites that are 110, 175, 1209, 1224, and 1528 bp, respectively, downstream from the *Hin*cII site in the *tetA* promoter (Fig. 1). The pBT107 TetA protein has an apparent M_r of $37 \pm 1 \times 10^3$ in the Laemmli SDS-polyacrylamide gel system used here. Deletion of the region between the *Nco*I and *Eco*RV sites (pBT147) leads to the synthesis of a 34 ± 1 kDal mutant TetA protein. This result suggests that the 5′ end of the *tetA* coding region is promoter-proximal to the *Nco*I site. Deletion of Tn*10* sequences distal to the *Cla*I site (pBT137) also results in the synthesis of a 34 ± 1 kDal mutant TetA protein, whereas deletion of sequences distal to the *Hin*dIII site (pBT127) has no apparent effect on the gel mobility of the TetA protein. These results are consistent with the results of other investigators (Coleman et al., 1981; 1983; Jorgensen et al., 1979), which demonstrate that the 3′ end of the *tetA* coding region is between the *Hpa*I and *Hin*dIII sites.

The tetracycline-resistance phenotypes of the deletion plasmids are somewhat anomalous. Minimal inhibitory concentrations of tetracycline for *E. coli* JA221 containing pACYC177, pBT107 (*tetA*$^+$), pBT147(*ΔtetA*), pBT137(*ΔtetA*) and pBT127(*tetA*$^+$) are 2.5, 30, 0.2, 80 and 70 μg/ml tetracycline, respectively (Moyed et al., 1983). Thus the amino-terminal *Nco*I-*Eco*RV deletion results in hypersensitivity to tetracycline, whereas the carboxy-terminal *Cla*I deletion results in increased resistance to tetracycline. We are currently examining the basis of these unexpected phenotypes.

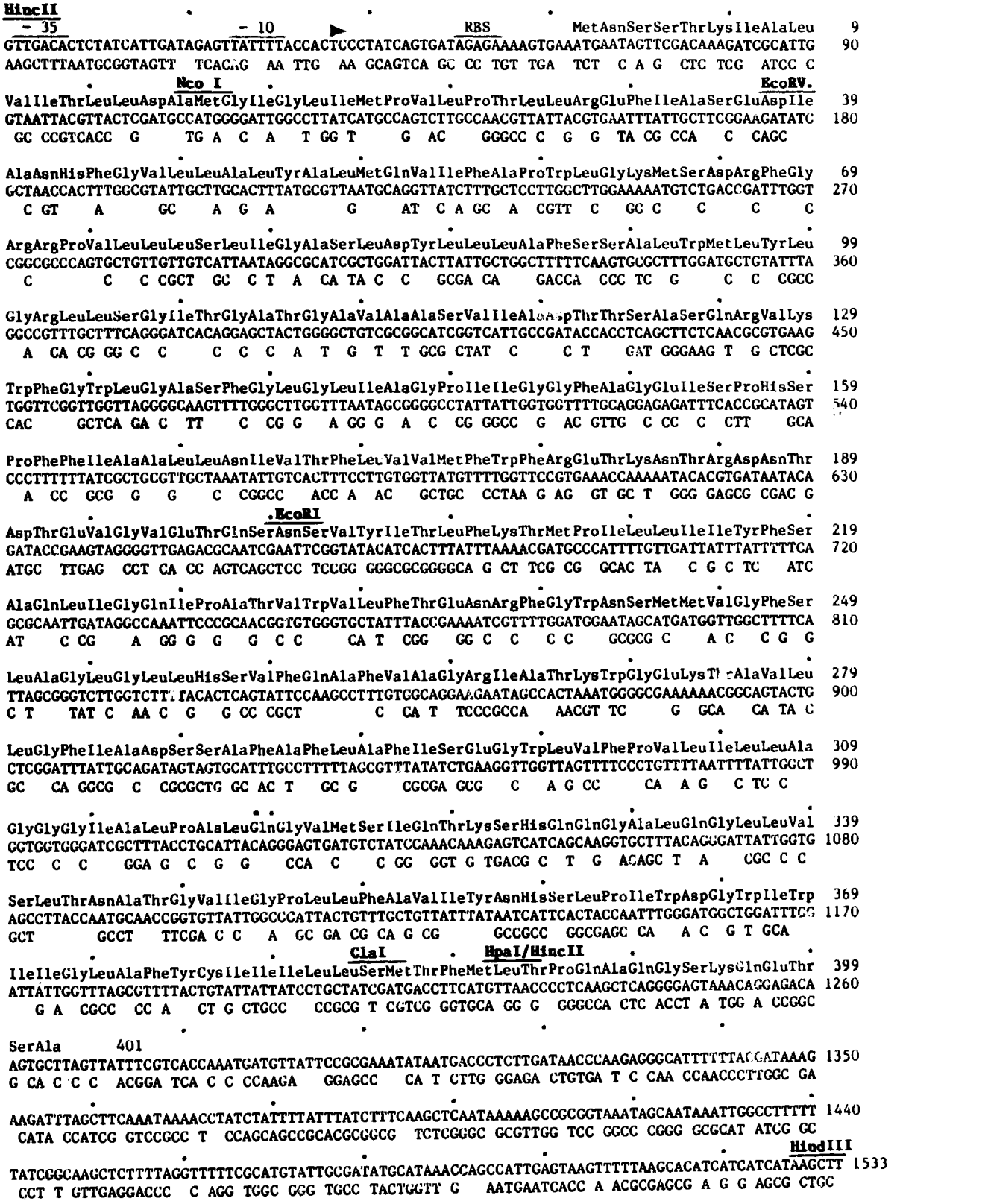

Fig. 4. Comparison of Tn*10 tetA* and pBR322 *tet* nucleotide sequences. The complete nucleotide sequence of the Tn*10 tetA* region is numbered from the *Hin*cII recognition site. Only those nucleotides in the pBR322 *tet* sequence that differ from the Tn*10 tetA* sequence are indicated below the Tn*10* sequence. The pBR322 sequence corresponds to bp 29-1560 as reported by Sutcliffe (1979) with the insertion of a C between bp 525 and bp 529 as described by Peden (1983). The predicted amino acid sequence of the Tn*10* TetA protein is numbered from the probable ATG initiation codon at bp 64. Selected restriction sites, the − 35 and − 10 sequences of the *tetA* promoter, the *tetA* transcription initiation site (arrow) and the probable *tetA* ribosome binding site (RBS) are indicated above the sequence (Bertrand et al., 1983a).

(b) Nucleotide sequence

The nucleotide sequence of the 1533-bp *Hin*cII-*Hin*dIII region spanning the *tetA* gene was determined (Fig. 4). A 407-codon open-translation reading frame extends from bp 46 to 1266 (Fig. 4). As discussed below, the probable translation start site for the TetA protein is at bp 64 within this reading frame. The deduced amino acid sequence for the TetA protein is also presented in Fig. 4. The next longest open reading frames in the proper *tetA* orientation extend 37, 36, and 29 codons (bp 107–217, 1308–1415, and 450–536, respectively; Fig. 4). The longest open reading frames in the opposite orientation extend 86, 53 and 44 codons (bp 441–698, 714–872, and 1396–1527, respectively; Fig. 4).

DNA sequence analysis of the pBT147 deletion mutant showed that the *Nco*I–*Eco*RV deletion removes 63 bp extending from bp 115 to 177 (Fig. 4). This corresponds to an in-frame deletion of 21 amino acids extending from amino acids 16 to 36 in the deduced TetA protein sequence.

DISCUSSION

(a) Tn*10* *tet*A structural gene

Several investigators have suggested that the 1500-bp region of Tn*10* downstream from the *tetA* promoter might encode a small protein, in addition to the 36 kDal TetA protein (Coleman et al., 1983; Curiale et al., 1982; Jorgensen et al., 1979). The orientation of the *tetA* gene (Jorgensen et al., 1979), the location of the *tetA* promoter (Bertrand et al., 1983a; 1983b; Jorgensen et al., 1979; Wray et al., 1981) and the approximate location of the 3′ end of the *tetA* coding region (Coleman et al., 1981; 1983; Jorgensen et al., 1979) have been defined previously and are confirmed by additional data presented here. However, the 5′ end of the *tetA* coding region has not been well-defined, leaving open the possibility of a small promoter-proximal gene upstream of the *tetA* gene. Our finding that the 63-bp *Nco*I-*Eco*RV deletion (pBT147) reduces the apparent M_r of the TetA protein by about 3000 indicates that the *tetA* translation start site is promoter-proximal to the *Nco*I site. The major open reading frame contains only three possible translation start codons between the 5′ end of *tetA* RNA (bp 37, Fig. 4) and the *Nco*I-*Eco*RV deletion (bp 115–177): one GTG codon (bp 46–48) and two ATG codons (bp 64–66 and 112–114). Several considerations support the choice of the ATG codon at bp 64–66 as the probable translation initiation codon for *tetA*: (1) The ATG at bp 64–66 is preceded by the sequence TAGAGAA which is similar to the ribosome recognition sites of many *E. coli* genes (Gold et al., 1981) and identical to a sequence in the ribosome recognition site of the *E. coli trpE* gene (Platt et al., 1976). (2) Insertion of the 164-bp *Taq*I restriction fragment spanning the *tetA* promoter (including bp 1–75, Fig. 4) into the promoter expression vector λRS205 leads to high level expression of β-galactosidase, a result which indicates that the *Taq*I fragment contains a translation initiation site as well as the *tetA* promoter (Bertrand et al., 1983b). (3) Finally, the *tetA* gene appears to be reasonably well expressed (cf. Fig. 3); translation initiation at the GTG codon (bp 46–48) would almost certainly be inefficient due to its proximity to the 5′ end of *tetA* RNA and the absence of a ribosome recognition site. Translation initiation at the probable ATG start codon (bp 64–66) would result in the synthesis of a 401 amino acid TetA protein having a calculated M_r of 43.2×10^3.

The data presented here appear to eliminate the possibility that an additonal gene is located between the *tetA* promoter and the *tetA* structural gene. We note that the pBT107 TetA protein is indistinguishable in size from the TetA protein encoded by R100, the R plasmid from which Tn*10* was isolated (Fig. 3). Therefore, it is unlikely that the pBT107 *tetA* gene is a fusion of two Tn*10* *tet* genes. Moreover, quantitation of *tet* RNA by hybridization to single-strand M13 *tet* DNA probes indicates that there is negligible "wrong-strand" transcription through the *tetA* region (unpublished observations); therefore, it is unlikely that any of the short open translation reading frames in this orientation are utilized. Since plasmids that contain the region from the *Hin*cII site in the *tetA* promoter to near the *Hin*cII site in the 3′ coding region of the *tetA* gene can confer tetracycline resistance (Wray et al., 1981), we conclude that the 36-kDal TetA protein is sufficient to confer tetracycline resistance in the absence of other Tn*10*-encoded proteins. We speculate that the TetA protein constitutes the tetracycline efflux pump

implicated in the mechanism of Tn*10* tetracycline resistance (Ball et al., 1980; McMurry et al., 1980). Alternatively, the TetA protein may function in concert with one or more host proteins to effect tetracycline efflux. The data presented here indicate that the partial genetic complementation that has been observed among tetracycline-sensitive mutations that map within the *tetA* region (Coleman et al., 1983; Curiale et al., 1982) is intracistronic complementation. This conclusion and the observation that some tetracycline-sensitive mutations exhibit a dominant phenotype (Coleman et al., 1983; Foster et al., 1976) suggest that functional TetA protein may exist as a homomultimer.

(b) Sequence homology between the *tet* genes in Tn*10* and pBR322

The Tn*10* and pBR322 tetracycline-resistance determinants are phenotypically and genetically distinct. They confer different levels of resistance to tetracycline and tetracycline analogs, and they fail to cross-hybridize in DNA-DNA hybridization experiments (Mendez et al., 1980). In spite of these differences we find that the Tn*10* *tetA* and pBR322 *tet* DNA sequences show significant homology (Fig. 4). Within the coding region for Tn*10* *tetA* (bp 64–1266) there is 49% (588/1203 bp) nucleotide homology with the pBR322 *tet* region. There are shorter segments that have somewhat higher homology (e.g. 76% homology, bp 206–257), however, there are no regions of sequence identity longer than eight nucleotides. As shown in Fig. 4, the homology outside the coding region is not significant (22% homology, bp 1–63 and 1267–1533); however, no attempt was made to maximize the homology outside the coding region by shifting the sequences and thus introducing insertions or deletions. In fact, the Tn*10* *tetA* and pBR322 *tet* promoter regions show significant sequence homology (53%) when they are aligned (Bertrand et al., 1983a). Whereas the Tn*10* *tetA* transcription start site is at bp 37 (Fig. 4), the

```
Tn10       M N S S T K I A L V I T L L D A M G I G L I M P V L P T L   29
pBR322 M K S N N A L I V I L G T V T L D A V G I G L V M P V L P G L   31

L R E F I A S E D I A N H F G V L L A L Y A L M Q V I F A P W L G K M   64
L R D I V H S D S I A S H Y G V L L A L Y A L M Q F L C A P V L G A L   66

S D R F G R R P V L L L S L I G A S L D Y L L L A F S S A L W M L Y L   99
S D R F G R R P V L L A S L L G A T I D Y A I M A T T P V L W I L Y A  101

G R L L S G I T G A T G A V A A S V I A D T T S A S Q R V K W F G W L  134
G R I V A G I T G A T G A V A G A Y I A D I T D G E D R A R H F G L M  136

G A S F G L G L I A G P I I G G F A G E I S P H S P F F I A A L L N I  169
S A C F G V G M V A G P V A G G L L G A I S L H A P F L A A A V L N G  171

V T F L V V M F W F R E T K N T R D N T D T E V G V E T Q S N S V Y I  204
L N L L L G C F L M Q E S H K G E R R P M P L R A F N P V S S F R W A  206

T L F K T M P I L L I I Y F S A Q L I G Q I P A T V W V L F T E N R F  239
R G M T I V A A L M T V F F I M Q L V G Q V P A A L W V I F G E D R F  241

G W N S M M V G F S L A G L G L L H S V F Q A F V A G R I A T K W G E  274
R W S A T M I G L S L A V F G I L H A L A Q A F V T G P A T K R F G E  276

K T A V L L G F I A D S S A F A F L A F I S E G W L V F P V L I L L A  309
K Q A I I A G M A A D A L G Y V L L A F A T R G W M A F P I M I L L A  311

G G G I A L P A L Q G V M S I Q T K S H Q Q G A L Q G L L V S L T N A  344
S G G I G M P A L Q A M L S R Q V D D D H Q G Q L Q G S L A A L T S L  346

T G V I G P L L F A V I Y N H S L P I W D G W I W I I G L A F Y C I I  379
T S I T G P L I V T A I Y A A S A S T W N G L A W I V G A A L Y L V C  381

I L L S M T F M L T P Q A Q G S K Q E T S A * 401
L P A L R R G A W S R A T S T * 396
```

Fig. 5. Comparison of Tn*10* TetA and pBR322 Tet protein sequences. The predicted amino acid sequences are numbered from their probable amino-terminal methionine residues (Sutcliffe, 1979; Peden, 1983). Identical residues are boxed.

location of the probable pBR322 *tet* transcription start site corresponds to Tn*10* bp 16 or 17 (Bertrand et al., 1983a; Sutcliffe, 1979).

The amino acid sequence homology between the predicted Tn*10* TetA and pBR322 Tet protein sequences is more readily apparent that the nucleotide homology (Fig. 5). Relative to the Tn*10* TetA sequence, there is 44% (177/401) amino acid homology. The predicted lengths of the Tn*10* TetA and pBR322 Tet proteins are 401 and 396 amino acids, respectively (Fig. 5, Peden, 1983), which is consistent with our observation that the sizes of the Tn*10* and pBR322 *tet* proteins are indistinguishable in SDS-polyacrylamide gels (Fig. 3). A striking feature of the amino acid homology is that it involves no net insertions or deletions. Among the 177 identical amino acids, only 53 are specified by the same codon in both Tn*10* and pBR322. As expected, the majority (86%) of the synonymous codon changes are third-position changes. We conclude that the Tn*10* *tetA* gene and the pBR322 *tet* gene diverged from a common sequence by successive nucleotide substitutions.

The nucleotide substitutions that have occured during the divergence of the Tn*10* *tetA* and pBR322 *tet* genes have resulted in a significant difference in their G + C contents. Whereas the Tn*10* *tetA* coding region is 43.4% G + C, the pBR322 *tet* coding region is 61.6% G + C. The IS*10* left and IS*10* right elements that flank the *tetA* region are 44.3% and 43.9% G + C, respectively (Halling et al., 1982). Thus it appears that the transposon Tn*10* is, in general, quite A + T-rich. In this context, we note that the Tn*10* *tetA* gene also shows some pecularities in codon usage that are not characteristic of *E. coli* genes. For example, the leucine codon usage in Tn*10* *tetA* is: TTA (23), TTG (13), CTT (11), CTC (2), CTA (4) and CTG (7). In contrast, virtually all *E. coli* genes examined appear to have a preference for the leucine codon CTG (Konigsberg et al., 1983). The pattern of codon usage in pBR322 *tet* is generally more like that of moderately expressed *E. coli* genes.

An interesting feature of the predicted amino acid sequence homology is that it is not randomly distributed over the lengths of the two proteins. There are short regions of relatively low homology at the amino-terminus (15% homology, Tn*10* amino acids 1–13), in the middle (12% homology, amino acids 169-220), and at the carboxy-terminus (0% homology, amino acids 377–401), which bound two longer regions of relatively high homology (60% homology, amino acids 14–168; 48% homology, amino acids 221–376). Moreover, the amino acid substitutions that have occurred within the two regions of high homology tend to be more conservative (i.e., similar in size and polarity) than the substitutions that have occurred in the regions of low homology. In this context, it is interesting that the carboxy-terminal 17 amino acids of the Tn*10* TetA protein can be deleted without significant loss of function (i.e., the *Cla*I deletion in pBT137). It is possible that the two relatively more conserved regions of the Tn*10* and pBR322 Tet proteins reflect the existence of structural and, perhaps, functional domains in the Tet proteins.

With the exception of the filamentous phage coat proteins, there are few if any examples of *E. coli* cytoplasmic membrane proteins that are synthesized with proteolytically removed amino-terminal signal sequences (Michaelis et al., 1982). The predicted amino-termini of the Tn*10* and pBR322 Tet proteins resemble known signal sequences in that they contain a basic amino acid (lysine 6 in Tn*10*; lysine 2 in pBR322) followed by a 24–27 amino acid hydrophobic region (Fig. 5). However, the acidic amino acid (aspartate 15 in Tn*10*; aspartate 17 in pBR322) within the hydrophobic region is uncharacteristic of *E. coli* signal sequences. Mutations that introduce acidic amino acids into the hydrophobic portions of the signal sequences of lipoprotein (Lin et al., 1978), λ receptor (Emr et al., 1980) and maltose-binding protein (Bedouelle et al., 1980) all result in failure to proteolytically process the precursors for these proteins. The properties of the *Nco*I-*Eco*RV deletion mutant (pBT147) also tend to support the hypothesis that the Tn*10* TetA protein does not have a proteolytically removed signal sequence. If a 20–25 amino acid signal peptide was cleaved from the amino-terminus of the TetA protein, then deletion of amino acids 16–36 in the predicted TetA protein sequence would be expected to remove the proteolytic cleavage site, yielding a protein as long or longer than the normally processed species.

The predicted amino acid sequences of the Tn*10* and pBR322 Tet proteins indicate that these proteins are markedly hydrophobic. Kyte and Doolittle (1982) have proposed an amino acid "hydropathy" scale; the values range from −4.5 for arginine to

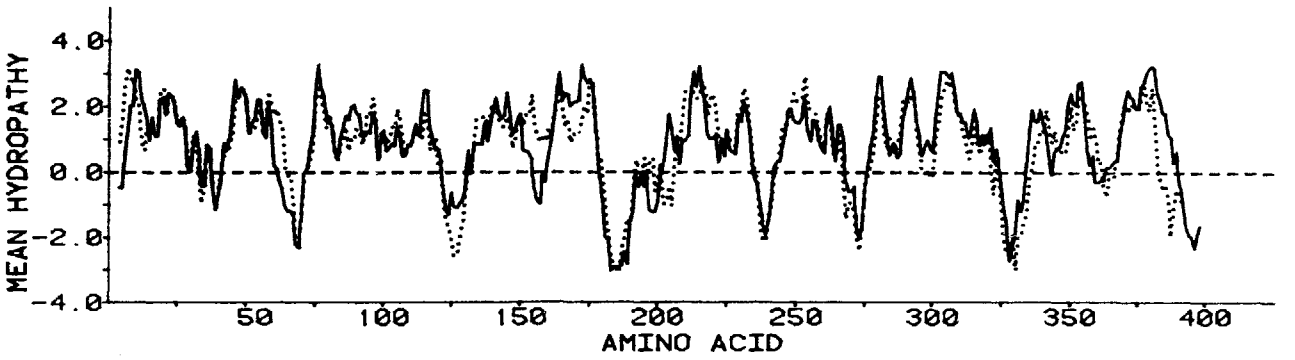

Fig. 6. Hydropathy profiles of the Tn*10* and pBR322 Tet proteins. A version of the SOAP program described by Kyte and Doolittle (1982) was used to determine the mean hydropathy of a moving segment of seven amino acid residues. The values are plotted at the midpoint of each segment, proceeding from the amino-terminus on the left to the carboxy-terminus on the right. Relatively hydrophobic regions appear above the midline. Tn*10* TetA amino acids 1–401, solid line; pBR322 Tet amino acids 3–396, dotted line.

+4.5 for isoleucine. When the hydropathies of the Tn*10* and pBR322 Tet proteins are evaluated along their amino acid sequences by a moving segment procedure (Kyte et al., 1982), it is apparent that both Tet proteins contain a number of relatively long hydrophobic regions separated by shorter hydrophilic regions (Fig. 6). Moreover, the lengths and distribution of the hydrophobic regions are remarkably similar in the two Tet proteins; the hydrophobic regions comprise about 70% of the amino acid residues in each case. By analogy with the observed membrane-spanning segments of bacteriorhodopsin (Engelman et al., 1980; Foster et al., 1983; Kyte et al., 1982), and the proposed membrane-spanning segments of lactose permease (Foster et al., 1983), we speculate that many or all of the major hydrophobic regions of the Tn*10* and pBR322 Tet proteins may be membrane-spanning segments. Thus a major portion of the Tet proteins may be embedded within the lipid bilayer of the cytoplasmic membrane.

During the preparation of this manuscript, Hillen et al. (1983) reported the nucleotide sequence of the same region of the Tn*10* tetracycline-resistance determinant. The sequence presented here differs from the sequence of Hillen et al. (1983) at four positions (nucleotides 905, 965, 1051, and 1123, Fig. 4); however, none of these differences significantly affects the predictions regarding the structures of the *tetA* gene and protein.

ACKNOWLEDGEMENTS

We thank George Gutman, Russell Doolittle, and Dean Goddette for helpful discussion and for assistance with the computer analysis of the sequence data, and Keith Peden for providing the revised sequence of pBR322 *tet* prior to publication. Stephanie Broderick, Laurie Smith and Rose Sekulovich constructed the deletion plasmids. This work was supported by Research Grant AI16735, Research Fellowship GM06702 (KP) and Research Career Development Award A100470 (KPB) from the National Institutes of Health.

REFERENCES

Ball, P.R., Shales, S.W. and Chopra, I.: Plasmid-mediated tetracycline resistance in *Escherichia coli* involves increased efflux of the antibiotic. Biochem. Biophys. Res. Commun. 93 (1980) 74–81.

Beck, C.F., Mutzel, R., Barbe, J. and Muller, W.: A multifunctional gene (*tetR*) controls Tn*10*-encoded tetracycline resistance. J. Bacteriol. 150 (1982) 633–642.

Bedouelle, H., Bassford, Jr., P.J., Fowler, A.V., Zabin, I., Beckwith, J. and Hofnung, M.: Mutations which alter the function of the signal sequence of the maltose binding protein of *Escherichia coli*. Nature 285 (1980) 78–81.

Bertrand, K.P., Postle, K., Wray, L.V. and Reznikoff, W.S.: Overlapping divergent promoters control expression of Tn*10* tetracycline resistance. Gene 23 (1983a) 149–156.

Bertrand, K.P., Postle, K., Wray, L.V. and Reznikoff, W.S.: Construction of a lambda phage promoter vector and its use in the analysis of the regulation of the Tn*10* tetracycline resistance determinant. J. Bacteriol. (1983b) in press.

Coleman, D.C. and Foster, T.J.: Analysis of the reduction in expression of tetracycline resistance determined by transposon Tn*10* in the multicopy state. Mol. Gen. Genet. 182 (1981) 171–177.

Coleman, D.C., Chopra, I., Shales, S.W., Howe, T.G.B. and Foster, T.J.: Analysis of tetracycline resistance encoded by transposon Tn*10*: Deletion mapping of tetracycline-sensitive point mutations and identification of two structural genes. J. Bacteriol. 153 (1983) 921–929.

Curiale, M.S. and Levy, S.B.: Two complementation groups mediate tetracycline resistance determined by Tn*10*. J. Bacteriol. 151 (1982) 209–215.

Emr, S.D., Hedgepeth, J., Clement, J.-M., Silhavy, T.J. and Hofnung, M.: Sequence analysis of mutations that prevent export of λ receptor, and *Escherichia coli* outer membrane protein. Nature 285 (1980) 82–85.

Engelman, D.M., Henderson, R., McLachlan, A.D. and Wallace, B.A.: Path of the polypeptide in bacteriorhodopsin. Proc. Natl. Acad. Sci. USA 77 (1980) 2023–2027.

Foster, T.J., Howe, T.G.B. and Richmond, K.M.V.: Translocation of the tetracycline resistance determinant from R100-1 to the *Escherichia coli* K-12 chromosome. J. Bacteriol. 124 (1975) 1153–1158.

Foster, T.J.: R factor-mediated tetracycline resistance in *Escherichia coli* K-12: Dominance of some tetracycline sensitive mutants and relief of dominance by deletion. Mol. Gen. Genet. 143 (1976) 339-344.

Foster, D.L., Boublik, M. and Kaback, H.R.: Structure of the *lac* carrier protein of *Escherichia coli*. J. Biol. Chem. 258 (1983) 31–34.

Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B.S. and Stormo, G.: Translation initiation in prokaryotes. Annu. Rev. Microbiol. 35 (1981) 365–403.

Halling, S.M., Simons, R.W., Way, J.C., Walsh, R.B. and Kleckner, N.: DNA sequence organization of IS*10*-right of Tn*10* and comparison with IS*10*-left. Proc. Natl. Acad. Sci. USA 79 (1982) 2608-2612.

Hillen, W., Klock, G., Kaffenberger, I., Wray, L.V. and Reznikoff, W.S.: Purification of the *tet* repressor and *tet* operator from the transposon Tn*10* and characterization of their interaction. J. Biol. Chem. 257 (1982) 6605–6613.

Hillen, W. and Schollmeier, K.: Nucleotide sequence of the Tn*10* encoded tetracycline resistance gene. Nucl. Acids Res. 11 (1983) 525–539.

Izaki, K., Kiuchi, K. and Arima, K.: Specificity and mechanism of tetracycline resistance in a multiple drug resistant strain of *Escherichia coli*. J. Bacteriol. 91 (1966) 628–633.

Jorgensen, R.A. and Reznikoff, W.S.: Organization of structural and regulatory genes that mediate tetracycline resistance in transposon Tn*10*. J. Bacteriol. 138 (1979) 705–714.

Kleckner, N., Chan, R.K., Tye, B.-K. and Botstein, D.: Mutagenesis by insertion of a drug-resistance element carrying an inverted repetition. J. Mol. Biol. 97 (1975) 561–575.

Konigsberg, W. and Godson, G.N.: Evidence for use of rare codons in the *dnaG* gene and other regulatory genes of *Escherichia coli*. Proc. Natl. Acad. Sci. USA 80 (1983) 687–691.

Kyte, J. and Doolittle, R.F.: A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157 (1982) 105–132.

Levy, S.B. and McMurry, L.: Detection of an inducible membrane protein associated with R-factor-mediated tetracycline resistance. Biochem. Biophys. Res. Commun. 56 (1974) 1060–1068.

Levy, S.B., McMurry, L., Onigman, P. and Saunders, R.M.: Plasmid-mediated tetracycline resistance in *E. coli*. in Drews, J. and Hogenauer, G. (Eds.) Topics in infectious diseases, Vol. 2–1977. Springer-Verlag, Vienna, 1977, pp. 181–206.

Lin, J.J.C., Kanazawa, H., Ozols, J. and Wu, H.C.: An *Escherichia coli* mutant with an amino acid alteration within the signal sequence of outer membrane prolipoprotein. Proc. Natl. Acad. Sci. USA 75 (1978) 4891–4895.

McMurry, L., Petrucci Jr., R.E. and Levy, S.B.: Active efflux of tetracycline encoded by four genetically different tetracycline resistance determinants in *Escherichia coli*. Proc. Natl. Acad. Sci. USA 77 (1980) 3974–3977.

Maxam, A.M. and Gilbert, W.: Sequencing end-labeled DNA with base-specific chemical cleavage. in Grossman, L. and Moldave, K. (Eds.), Methods in Enzymology, Vol. 65. Academic Press, New York, 1980, pp. 499–560.

Mendez, B., Tachibana, C., and Levy, S.B.: Heterogeneity of tetracycline resistance determinants. Plasmid 3 (1980) 99–108.

Michaelis, S. and Beckwith, J.: Mechanism of incorporation of cell envelope protein in *Escherichia coli*. Annu. Rev. Microbiol. 36 (1982) 435–465.

Moyed, H.S., Nguyen, T.T. and Bertrand, K.P.: Multicopy Tn*10* *tet* plasmids confer sensitivity to induction of *tet* gene expression. J. Bacteriol. 155 (1983) 549–556.

Peden, K.W.C.: Revised sequence of the tetracycline resistance gene of pBR322. Gene 22 (1983) 277–280.

Platt, T., Squires, C. and Yanofsky, C.: Ribosome-protected regions in the leader-*trpE* sequence of *Escherichia coli* tryptophan operon messenger RNA. J. Mol. Biol. 103 (1976) 411–420.

Shales, S.W., Chopra, I., and Ball, P.R.: Evidence for more than one mechanism of plasmid-determined tetracycline resistance in *Escherichia coli*. J. Gen. Microbiol. 121 (1980) 221–229.

Sutcliffe, J.G.: Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. 43 (1979) 77–90.

Wray Jr., L.V., Jorgensen, R.A. and Reznikoff, W.S.: Identification of the tetracycline resistance promoter and repressor in transposon Tn*10*. J. Bacteriol. 147 (1981) 297–304.

Yang, H.-L., Zubay, G. and Levy, S.B.: Synthesis of an R plasmid protein associated with tetracycline resistance is negatively regulated. Proc. Natl. Acad. Sci. USA 73 (1976) 1509–1512.

Zupancic, T.J., King, S.R., Pogue-Geile, K.L. and Jaskunas, S.R.: Identification of a second tetracycline-inducible polypeptide encoded by Tn*10*. J. Bacteriol. 144 (1980) 346–355.

Communicated by H.O. Smith.