Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 5975–5979, October 1984
Biochemistry

Mercuric ion-resistance operons of plasmid R100 and transposon Tn*501*: The beginning of the operon including the regulatory region and the first two structural genes

(heavy metal resistance/bacterial plasmids/DNA sequencing)

TAPAN K. MISRA*, NIGEL L. BROWN†, D. C. FRITZINGER†, R. D. PRIDMORE†, W. M. BARNES‡, LINDA HABERSTROH*, AND SIMON SILVER*

*Department of Biology, Washington University, St. Louis, MO 63130; †Department of Biochemistry, University of Bristol, Bristol BS8 1TD, United Kingdom; and ‡Department of Biological Chemistry, Washington University School of Medicine, St. Louis, MO 63110

Communicated by Boris Magasanik, June 22, 1984

ABSTRACT **The mercuric ion-resistance operons of plasmid R100 (originally from *Shigella*) and transposon Tn*501* (originally from a plasmid isolated in *Pseudomonas*) have been compared by DNA sequence analysis. The sequences for the first 1340 base pairs of Tn*501* are given with the best alignment with the comparable 1319 base pairs of R100. The homology between the two sequences starts at base 58 after the end of the insertion sequence IS-1 of R100. The sequences include the transcriptional regulatory region, and the homology is particularly strong in regions just upstream from potential transcriptional initiation sites. The *trans*-acting regulatory gene *merR* consists of 180 base pairs in both cases and codes for a highly basic polypeptide of 60 amino acids, which is also rich in serine. The Tn*501* and R100 *merR* genes differ in 25 of the 180 base positions, and the resulting polypeptides differ in seven amino acids. The regulatory region before the major transcription initiation site contains potential −35 and −10 sequences and dyad symmetrical sequences, which may be the *merR* binding sites for transcriptional regulation. The first structural gene, *merT*, encodes a highly hydrophobic polypeptide of 116 amino acids. The R100 and Tn*501* *merT* genes differ in 17% of their positions, leading to 14 (12%) amino acid changes. This region had previously been shown to encode a protein governing membrane transport of mercuric ions. The second structural gene, *merC*, would give a 91 amino acid polypeptide with a hydrophobic amino-terminal segment. The Tn*501* and R100 *merC* genes differ at 37 base positions, leading to 10 amino acid changes.**

Mercuric ion resistance, like many other resistances to toxic heavy metals in prokaryotes, is governed by genes on plasmids and transposons (1–3). Plasmid R100 is the initial antibiotic-resistance plasmid that appeared in 1956 in a *Shigella flexneri* strain in Japan (4, 5), and its total size is 90 kilobase pairs (kb). It encodes resistance to chloramphenicol, fusidic acid, tetracycline, sulfonamide, and streptomycin, as well as to mercuric ions (6). The last three resistances are encoded on the 19.9-kb transposon Tn*21* (7, 8) that occurs between flanking 768-base-pair insertion elements IS-1a and IS-1b (7, 9), which are not part of the transposon (7). Transposon Tn*501* is an 8.5-kb transposable element that encodes resistance to mercuric ions in addition to transposition functions (10). It was first found in a *Pseudomonas aeruginosa* strain (11), and it was moved into *Escherichia coli* for subsequent study (10, 12). The organization of the mercuric ion resistance determinant of plasmid R100 is well known. Transposon mutagenesis studies (13–15) and cloning studies (16, 17) determined the orientation of the operon and that the mercuric ion-resistance determinant starts close to the right boundary of IS-1. The operon was initially (13, 16) shown to consist of a regulatory gene (*merR*) that encodes a *trans*-acting positive/negative regulatory product, followed by an operator–promoter region, and then by a large transcriptional unit for a gene (*merT*) governing a membrane transport function, followed by the 1.7-kb gene for the enzyme mercuric reductase (*merA*). More recent studies (14) with Tn*5* transposon mutagenesis of the *mer* region cloned from plasmid R100 established two additional genes: *merC*, of unknown function, but located between *merT* and *merA*; and *merD*, also of unknown function and located operator–distal from *merA*. Additional transposon mutagenesis and cloning studies of Tn*501* (12, 17) and R100 (15) gave results consistent with a similar overall organization of the genetic determinants.

We report here the sequences of the beginning of the *mer* operons in both Tn*501* and Tn*21*. The two systems are strikingly similar, having the same gene orders, and having identical bases in ≈85% of the positions in the first three genes. Most base changes between the two systems occur in noncoding regions or in third "wobble" positions. This remarkable degree of conservation indicates that the two systems have a common and recent evolutionary ancestry. The nucleotide sequence and the predicted amino acid sequences of the gene products allow us to suggest a more detailed molecular mechanism for mercury detoxification by these systems.

MATERIALS AND METHODS

The chain-termination DNA sequencing methods of Sanger *et al.* (18) were used for both sequences. The Tn*501* sequence was determined with randomly cloned restriction enzyme fragments in bacteriophage M13mp7 (19), using the methods described (17, 20). The R100 sequences were determined from larger 2- to 4-kb fragments cloned into M13mWB2344 by the method of generating ordered deletions from the primer location into the cloned sequence (21, 22). Both sequences have been determined completely from both DNA strands. The Tn*501* sequence was cloned from the plasmids described in ref. 17, and the R100 sequence was from plasmid pDU1003 (14).

RESULTS

DNA Sequence. Fig. 1 shows the first 1319 bases of plasmid R100 plus the terminal 22 right-end nucleotides of IS-1 (9). The R100 sequence is numbered from the end of IS-1. Only that strand equivalent to the messenger RNA is shown. In determining this sequence, we sequenced from a cloned fragment starting at the *Tth*111I site in IS-1 and, therefore,

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.

Abbreviations: kb, kilobase pair(s); IR, inverted repeat.

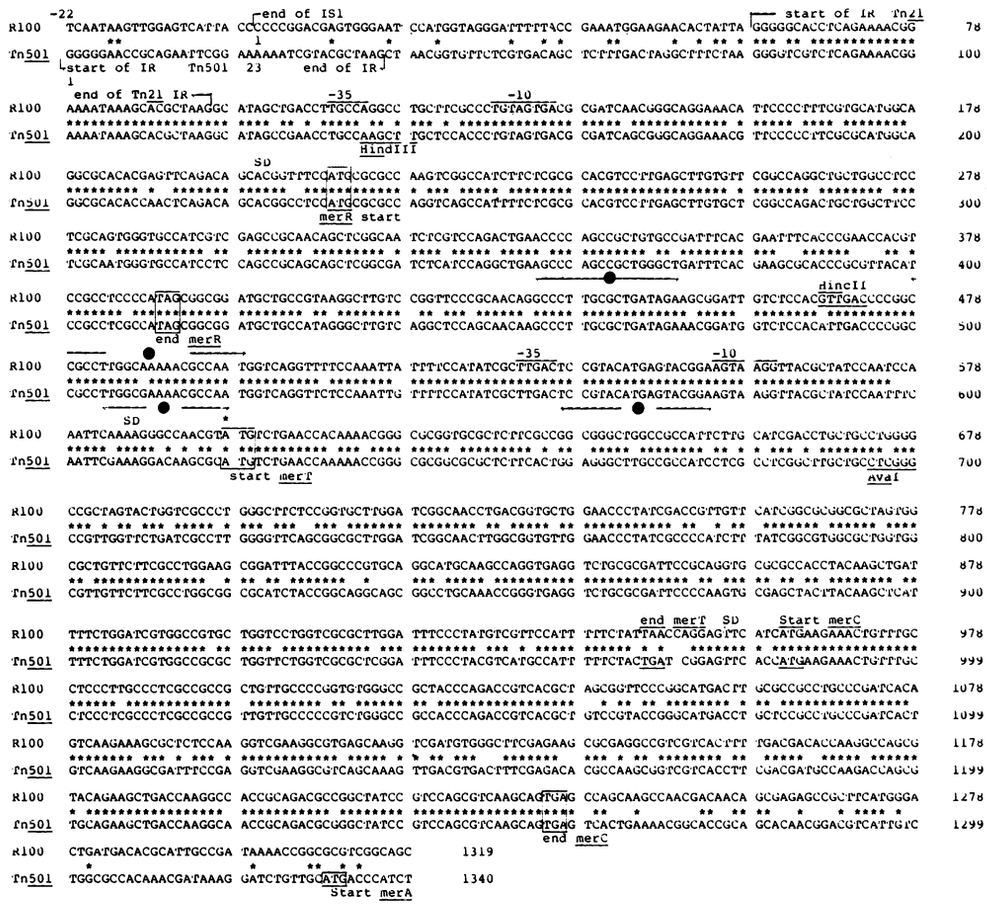

FIG. 1. DNA sequences of the beginning of transposon Tn*501* and the corresponding region of plasmid R100. The sequences start at the first nucleotide of Tn*501* (20) and with the final 22 bases of IS-1 (9). Only the strand equivalent to the mRNA (anti-coding strand) is shown. Asterisks between the sequences indicate identical bases in those positions. The end of IS-1 and the limits of the IRs of Tn*501* and Tn*21* are indicated. Possible −35 and −10 transcriptional initiation and ribosomal binding signals [Shine–Dalgarno sequence (SD)] from comparison with canonical consensus sequences (24–26) are shown. The dyad symmetrical sequences (——●——) are discussed in the text. The initiation and termination codons of the proposed reading frames are marked. The substrate sequences for three restriction endonuclease enzymes *Hin*dIII, *Hin*cII, and *Ava* I are indicated.

the terminal 56 base pairs of IS-1. This is compared with the first 1340 bases of Tn*501*. The sequences are aligned starting with base 58 of R100, approximately the start of the left inverted repeat (IR) of Tn*21* in the R100 sequence. There is no significant sequence homology before position 58 of the R100 sequence and the corresponding position 80 in the Tn*501* sequence. Both sequences are from mercuric ion-resistance transposons, yet the terminal 38-base-pair IRs of the two transposons do not align when the remainder of the sequence is aligned; rather, the functional IRs are displaced by 80 base pairs. Once aligned after base position 58 of R100, the two sequences are of identical length until position 949 of the R100 sequence, which has no counterpart in the Tn*501* sequence. The sequence alignment then continues with identical lengths through the end of the *merC* gene (at position 1237 of R100). From that position, the intragene sequences between *merC* and the next gene *merA* show no homology, and they differ in length. The lengths of the three significant reading frames are identical in the two sequences. The three restriction endonuclease sites shown (*Hin*dIII at position 136 of Tn*501*, *Ava* I at position 694 of Tn*501*, and *Hin*cII at position 467 of R100), which have been used in mapping gene functions (12, 14), are absent from the alternative sequence because of single base changes. The R100 sequence lacks restriction endonuclease sites for the 14 enzymes that were experimentally shown not to cut the mercuric ion-resistance region of R100 (14).

Likely transcriptional starts have been identified on the sequence. RNA polymerase binding data (unpublished observations) suggest that there are two transcriptional starts in Tn*501*: one in the first 200 nucleotides and one between nucleotides 499 and 700. Sequences corresponding to the canonical −35 and −10 transcriptional initiation sequences (24) are shown within these two regions in Fig. 1. The first such sequence (positions 110–135 of the R100 sequence) precedes an open reading frame (for the *merR* gene; see below) at the end of which is a dyad symmetrical sequence in both systems (positions 400–421 of the Tn*501* sequence). This could form a stem–loop structure in mRNA and may be involved in transcription termination, although the uracil-rich sequence typical of a ρ independent terminator (25) is absent. Other dyad symmetries in one system in this region (e.g., positions 356–372 and 431–460 in Tn*501*) are not fully conserved in the other system; the significance of this is not known. The second potential transcriptional initiation sequence is at positions 554–583 and is part of the regulated *mer* promotor. The −35 and −10 sequences are within a 57-nucleotide sequence that is identical in both Tn*501* and R100. In this region are two dyad symmetries. One of these is at positions 505–520 (six symmetrical nucleotides in Tn*501* and five in R100), which is approximately 80–95 nucleotides before the likely transcriptional start point; the other is at positions 559–576 (seven symmetrical nucleotides) between the −35 and −10 sequences. With one unpaired base in each strand, this sequence extends to 12 matching nucleotides. Following the *merR* coding sequence and preceding the −35 region for the next transcriptional initiation start, there is a region including an approximate inverted repeat, a series of guanines and cytosines, and a stretch rich in thymines (between positions 483 and 533 of the R100 sequence) consistent with a transcriptional termination site (25).

There are three significant open reading frames in Fig. 1 (the initiation and stop codons are marked), and these appear from their positions to correspond to the genes *merR*, *merT*, and *merC* of Ni'Bhriain *et al.* (14). The three reading frames are the same length (180, 348, and 273 base pairs) in both R100 and Tn*501*. The first (*merR*) reading frame is not preceded by a strong consensus translation initiation sequence

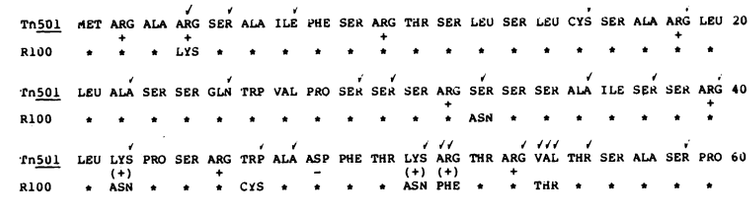

FIG. 2. Amino acid sequences of the *merR* polypeptides of R100 and Tn*501* deduced from the DNA sequences. Asterisks represent identical amino acids in a given position; only differences in amino acids are noted. Ticks above a given position represent a nucleotide difference in the first, second, or third position in the corresponding codon in the DNA. Charged amino acids are marked; parentheses indicate a difference in charge between the two polypeptides.

(26), which suggests that the gene product may be produced only in low amounts, compatible with a regulatory function. The other two reading frames (*merT* and *merC*) have better translational initiation consensus sequences (Fig. 1). Additional potential reading frames in the sequences are unlikely to be biologically functional, because they do not contain suitable translational initiation sequences (26) and they do not fit with the available genetic (14) and transcriptional data. For Tn*501*, the *merA* gene starts at position 1330 (which corresponds to position 29 in figure 3 of ref. 17). The region between *merC* and *merA* in Tn*501* contains 71 nucleotides. In R100, the region between the end of *merC* and the start of *merA* contains 507 bases, which are not homologous to those in Tn*501*. Note that the transcriptional direction both for *merR* and for the long structural gene transcript is from left to right in Fig. 1. This agrees with previous conclusions from *mer–lac* fusion studies (14) and from cloning and transposition polarity studies (13–16).

The Predicted Amino Acid Sequences. The first long open reading frame of 180 base pairs in each sequence corresponds to the position and orientation for the gene for the *trans*-acting *merR* regulatory protein (13, 14). The predicted *merR* gene product of 60 amino acids (Fig. 2) has an unusual composition. After the arginine→lysine substitution at position 4, there are 28 amino acids that are completely identical. The sequences contain a total of 17 (R100) or 18 (Tn*501*) serines out of the 60 amino acids, which makes it difficult to predict secondary or tertiary structure. The sequences contain 7 (R100) or 9 (Tn*501*) arginine residues, 1 or 2 lysines, and a net positive charge of 7 or 10 in R100 and in Tn*501*, respectively. There is a common cysteine at position 16, of potential interest for a regulatory protein whose regulatory cofactor is ionic Hg^{2+} (2, 13–16) and a second cysteine at position 46 of the R100 polypeptide sequence. It is apparent from inspection that the *merR* polypeptide has the potential of a highly positively charged DNA-binding protein. Comparing the R100 and Tn*501* sequences, one sees only 7 amino acid changes out of 60 amino acids and only 25 base pair changes (marked with ticks in Fig. 2) out of 180 base pairs, giving 86% nucleotide identity between R100 and Tn*501*.

merT. The second long open reading frame, corresponding in position to the *merT* gene (13, 14, 16), whose product functions to transport Hg^{2+} across the cell membrane, would give a translation product of 116 amino acids (Fig. 3) for both R100 and Tn*501*. Even by visual inspection, this polypeptide is extraordinarily hydrophobic. The distribution of charged and hydrophobic residues has enabled predictions of membrane-spanning regions (27) and a functional model (see *Discussion*). From the sequences in Fig. 3, there are 14 amino acid changes of 116 amino acids (88% conserved) and 59 nucleotide-pair changes of 348 (83% conserved).

merC. The third reading frame in the sequence shown (Fig. 1) corresponds in position to *merC*, which was found in R100 by Tn*5* insertional mutagenesis studies (14). The predicted translation product of the *merC* gene is a polypeptide of 91 amino acids (Fig. 4), starting with a charged NH_2-terminal tripeptide, followed by a very hydrophobic region of 37 amino acids, and then a region containing many charged and hydrophobic amino acids continuing to the carboxyl terminus. This NH_2-terminal sequence is homologous to the "leader sequences" of known membrane-associated and periplasmic proteins (28). Only 10 amino acids differ between the *merC* polypeptides of R100 and Tn*501*, and these include the compensating Lys→Glu change at position 46 and the Glu→Gln change at position 61. Only 37 base changes are seen between the two reading frames (86% conserved), while the polypeptides are 89% conserved in amino acid residues. Although the function of *merC* is unknown (14), this high degree of evolutionary conservation suggests that the gene is essential.

DISCUSSION

The DNA sequences shown in Fig. 1 provide two useful and interesting sets of information. First, the sequences and their analysis further our understanding of the structure and function of the mercuric ion-resistance determinants, which are the most thoroughly understood of the plasmid-borne heavy metal resistance systems (2, 3, 13–16). Second, the comparison of homologous sequences originating from *Shigella* (R100) and *Pseudomonas* (Tn*501*) affords an opportunity to study questions of gene evolution and codon usage with functional genes under tight evolutionary constraints.

The DNA sequence analysis of the mercuric ion-resistance determinants of R100 and Tn*501* confirm the positions and orientations of the first three genes (*merR*, *merT*, and *merC*) deduced from genetic and cloning experiments (13, 14, 16), as well as an operator promoter region between *merR* and *merT* (Fig. 1). From Tn*5* mutagenesis experiments and restriction endonuclease digestion analysis, Barrineau

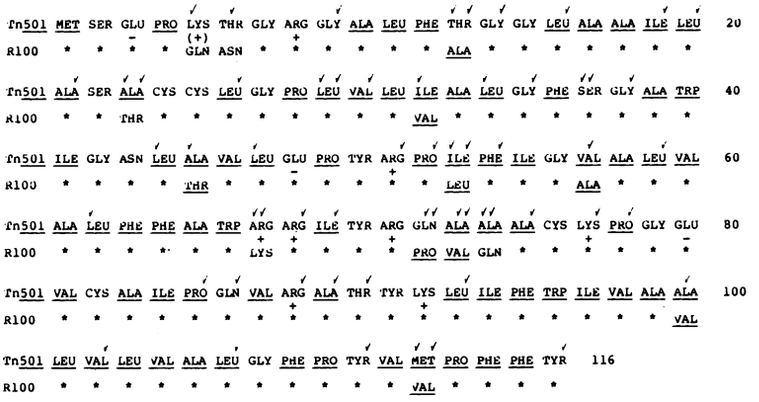

FIG. 3. Amino acid sequences of the *merT* polypeptides of R100 and Tn*501* deduced from the DNA sequences. Symbols are as in Fig. 2. Hydrophobic amino acids are underlined.

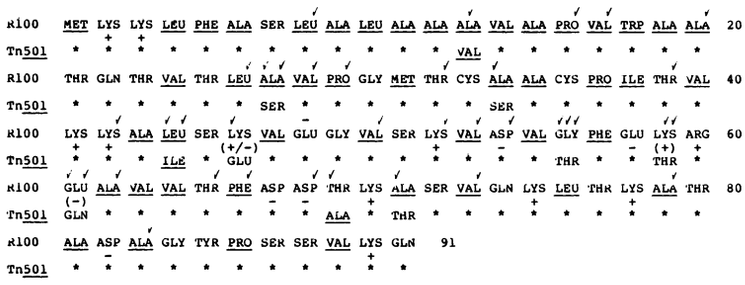

FIG. 4. Amino acid sequences of the *merC* polypeptides of R100 and Tn*501* deduced from the DNA sequences. Symbols are as in Fig. 2.

and Summers (15) have generated a slightly different picture of the beginning region of R100 from that developed also from Tn*5* insertions and restriction endonuclease analysis (14). The reading frames (genes) in Fig. 1 and the amino acid sequences (Figs. 2–4) derived from them are entirely consistent with the genetic results described in ref. 14.

There have been attempts to correlate the Hg^{2+}-inducible synthesis of polypeptides with λ *mer* phage (29), minicell (refs. 14 and 30; unpublished results) and maxicell (17) experiments, with segments of cloned DNA and with the positions of polar insertion mutants. Only the *merA* polypeptide, the subunit of the mercuric reductase enzyme, has been unambiguously identified (14, 17, 29, 30). The molecular weights of other Hg^{2+}-inducible polypeptides on polyacrylamide gels (14, 17, 29, 30) are somewhat higher than those from translation of the DNA sequences (M_r, 12,500, 9,500, and 6,500 for the primary translation products of the *merT*, *merC*, and *merR* genes, respectively). This may be accounted for by the hydrophobicity of the *merT* and *merC* products, which may cause them to run anomalously in NaDodSO$_4$/polyacrylamide gels, and by the small size of the *merR* protein. The DNA sequences presented here should allow these proteins to be identified more easily by gene cloning and by site-directed mutagenesis.

The predicted primary structures of the *merR*, *merT*, and *merC* gene products (Figs. 2–4) and the sequence of the regulatory region (Fig. 1) enable us to propose a detailed model for the mechanism of action of the mercuric ion-resistance determinants. The highly basic *merR* gene product represses transcription of the *mer* operon (13), and we suggest that it does this by binding to the dyad symmetrical sequence at positions 537–554 of the R100 sequence. Mercuric ions may bind to the cysteine residue (Cys-16) in the Tn*501* and R100 *merR* gene products, altering the specificity of DNA binding of these proteins and causing derepression of the *mer* operon. In addition, we hypothesize that the Hg^{2+}–protein complex binds to the dyad symmetrical sequence at positions 483–490 of the R100 sequence and causes increased transcription from the *mer* promoter. The *merR* gene of R100 is known to have both positive and negative regulatory functions (13, 14). The differences in net charge (+10 or +7) and in the number of cysteine residues (1 or 2) of the *merR* gene products of Tn*501* and R100, respectively, and the different symmetries of the proposed binding sites on DNA for positive regulation (nucleotides 483–498) suggest that there may be subtle differences in the regulation of the two systems.

Calculations of the hydropathic index (27) along the predicted amino acid sequences of the primary translation products of the *merT* gene of Tn*501* and R100 suggest that these polypeptides could form three transmembrane helices with two highly polar regions (residues 1–8 and 67–92; Fig. 3) at the membrane surface. This would place Cys-76 and Cys-82 on one face of the membrane and Cys-24 and Cys-25 within the transmembrane helix.

The predicted amino acid sequence of the primary translation product of the *merC* gene has a good "leader sequence" (28), suggesting that the polypeptide is likely to be vectorially synthesized across the cell membrane and may be processed by signal peptidase (31); the absence of a second hydrophobic region suggests that the polypeptide may move through the membrane into the periplasmic space. Barrineau and Summers (15) have reported from unpublished experiments the presence of a Hg^{2+}-inducible small polypeptide that is 50% released into the periplasm. It is tempting to suggest that the role of the *merC* gene product is to scavenge Hg^{2+} ions in the periplasmic space.

Table 1. Codon usage frequencies in *merR*, *merT*, and *merC* genes*

	R100	Tn*501*		R100	Tn*501*		R100	Tn*501*		R100	Tn*501*
UUU Phe	4	4	UCU Ser	1	1	UAU Tyr	4	2	UGU Cys	2	0
UUC Phe	10	10	UCC Ser	9	11	UAC Tyr	2	4	UGC Cys	6	7
UUA Leu	0	0	UCA Ser	0	2	UAA Ter	1	0	UGA Ter	1	2
UUG Leu	4	7	UCG Ser	9	6	UAG Ter	1	1	UGG Trp	5	6
CUU Leu	3	1	CCU Pro	0	0	CAU His	0	0	CGU Arg	0	0
CUC Leu	3	7	CCC Pro	5	6	CAC His	0	0	CGC Arg	7	10
CUA Leu	3	0	CCA Pro	5	4	CAA Gln	1	3	CGA Arg	2	1
CUG Leu	15	10	CCG Pro	6	5	CAG Gln	6	4	CGG Arg	2	2
AUU Ile	4	4	ACU Thr	2	4	AAU Asn	1	0	AGU Ser	0	0
AUC Ile	6	9	ACC Thr	9	9	AAC Asn	4	1	AGC Ser	7	9
AUA Ile	0	0	ACA Thr	1	2	AAA Lys	2	4	AGA Arg	1	1
AUG Met	4	5	ACG Thr	5	3	AAG Lys	12	10	AGG Arg	2	2
GUU Val	3	6	GCU Ala	3	5	GAU Asp	2	2	GGU Gly	3	1
GUC Val	12	11	GCC Ala	24	20	GAC Asp	3	3	GGC Gly	8	7
GUA Val	2	1	GCA Ala	5	3	GAA Glu	3	3	GGA Gly	1	2
GUG Val	10	8	GCG Ala	8	12	GAG Glu	3	3	GGG Gly	3	4

*Total of 270 codons including 61 for *merR*, 117 for *merT*, and 92 for *merC*. The stop codons in each reading frame (Ter) are included.

The pairing of cysteine residues in the polypeptide products of the *merC* gene (Cys-33 and Cys-36; Fig. 4), the *merT* gene (Cys-24 and Cys-25; Cys-76 and Cys-82; Fig. 3), and in the *merA* mercuric reductase (Cys-10 and Cys-13; Cys-136 and Cys-141; Cys-558 and Cys-559; ref. 17) suggests a plausible pathway for Hg^{2+} during its detoxification. Mercuric ions might be initially sequestered by binding to the thiol groups of the *merC* gene product, and "handed-on" by a series of redox exchange reactions through the two pairs of thiols in the *merT* gene product (one pair on or near each face of the inner membrane) to the thiols of the mercuric reductase. The biological advantage of such a redox-transfer system is that toxic Hg^{2+} is sequestered and never released to interact with cellular constituents until it has been reduced to nontoxic Hg^0. Direct experimental evidence to support this model is lacking, but the availability of the nucleotide sequence facilitates the design of experiments to test the hypothesis.

The sequences presented here support the proposition that the two mercuric ion-resistance determinants have a close evolutionary relationship, although R100 originated in *S. flexneri* (4, 5), which has an overall chromosomal G+C content of 51% (32), while Tn*501* originated in *P. aeruginosa* (11), which has an overall G+C content of 67% (32). The G+C content of the Tn*501* and R100 sequences are both 59 mol %, which is reflected in a codon usage in which codons ending in C or G are preferred (78% in R100 and 79% in Tn*501*) to those ending in U or A (Table 1). Brown *et al.* (17) argued that this preference for G and C at the third codon position may reflect a typical codon usage for *Pseudomonas* genes, where the bulk DNA has a G+C content of 67 mol % (32). Recent descriptions of other *Pseudomonas* genes support this assertion (33, 34). The G+C content of *Shigella* genes would be expected to be closer to that of *Escherichia* than to that of *Pseudomonas*, and the *ompA* gene of *Shigella dysenteriae* has a G+C content of 50 mol % (23). The similarity between the sequences of the Tn*501* and R100 mercuric ion-resistance determinants argues for a common origin; and the base composition and predicted codon usage (Table 1) imply that this origin was in a G+C-rich genus, such as *Pseudomonas*.

The nucleotide sequences of the start of the Tn*501* and R100 *mer* regions presented in this paper show a remarkable degree of homology; yet they differ markedly in two regions in the sequence—namely, at the beginning of the sequences and between the *merC* and *merA* genes. The start of the homology between Tn*501* and R100 coincides with the Tn*21* inverted terminal repeat. It is natural to assume that Tn*501* and Tn*21* might have homologous IR regions. Indeed, the IRs of the two systems are identical in 30 of 38 positions (79%) for the left IRs (Fig. 1) and 28 of 38 positions (74%) for the right IRs (7, 35), but the two sequences do not occur in comparable positions for the left IRs (Fig. 1). There are 80 additional base pairs in Tn*501*. This additional sequence includes a virtual repeat of the left-most functional IR in Tn*501* with a sequence that is identical in 35 of 38 bases in Tn*501* with the functional IR for Tn*21* (Fig. 1; refs. 7 and 8). Although the sequence between positions 81 and 118 of Tn*501* has 37 of 38 identical bases to the left IR of Tn*501*, it does not generally function for the transposon as a terminal sequence in transposition. Nucleotides 81–118 of Tn*501* (which align in Fig. 1 with the inverted repeat of Tn*21*) constitute a functional inverted repeat for complementation of transposition by Tn*21* (unpublished results) and it is proposed that Tn*501* arose by the transposition of a Tn*21*-like mercury-resistance transposon into a site 80 nucleotides from the start of a transposable element, with subsequent deletion of the Tn*21* transposition genes.

Georg Jander, Annette Schmidt, and Joseph Winnie helped with the DNA sequencing. The free and frequent exchanges of information between our laboratories and N. Ni'Bhriain, T. J. Foster, J. Grinsted, and R. Schmitt, greatly facilitated our efforts. This work was supported by National Institutes of Health Grants AI15672 and GM24956 and U.K. Medical Research Council Grant G9791047CB. N.L.B. is a Royal Society E.P.A. Cephalosporin Fund Senior Research Fellow.

1. Novick, R. P. & Roth, C. (1968) *J. Bacteriol.* **95,** 1335–1342.
2. Summers, A. O. & Silver, S. (1978) *Annu. Rev. Microbiol.* **32,** 637–672.
3. Silver, S. (1981) in *Molecular Biology, Pathogenicity and Ecology of Bacterial Plasmids*, eds. Levy, S., Clowes, R. C. & Koenig, E. L. (Plenum, New York), pp. 179–189.
4. Nakaya, R., Nakamura, A. & Murata, Y. (1960) *Biochem. Biophys. Res. Commun.* **3,** 654–659.
5. Watanabe, T. (1966) *N. Engl. J. Med.* **275,** 888–894.
6. Miki, T., Easton, A. M. & Rownd, R. H. (1978) *Mol. Gen. Genet.* **158,** 217–224.
7. Zheng, Z. X., Chandler, M., Hipskind, R., Clerget, M. & Caro, L. (1981) *Nucleic Acids Res.* **9,** 6265–6278.
8. de la Cruz, F. & Grinsted, J. (1982) *J. Bacteriol.* **151,** 222–228.
9. Ohtsubo, H. & Ohtsubo, E. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 615–619.
10. Bennett, P. M., Grinsted, J., Choi, C. L. & Richmond, M. H. (1978) *Mol. Gen. Genet.* **159,** 101–106.
11. Stanisich, V. A., Bennett, P. M. & Richmond, M. H. (1977) *J. Bacteriol.* **129,** 1227–1233.
12. Grinsted, J., Bennett, P. M., Higginson, S. & Richmond, M. H. (1978) *Mol. Gen. Genet.* **166,** 313–320.
13. Foster, T. J., Nakahara, H., Weiss, A. A. & Silver, S. (1979) *J. Bacteriol.* **140,** 167–181.
14. Ni'Bhriain, N., Silver, S. & Foster, T. J. (1983) *J. Bacteriol.* **155,** 690–703.
15. Barrineau, P. & Summers, A. O. (1983) *Gene* **25,** 209–221.
16. Nakahara, H., Silver, S., Miki, T. & Rownd, R. H. (1979) *J. Bacteriol.* **140,** 161–166.
17. Brown, N. L., Ford, S. J., Pridmore, R. D. & Fritzinger, D. C. (1983) *Biochemistry* **22,** 4089–4095.
18. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
19. Messing, J., Crea, R. & Seeburg, P. H. (1981) *Nucleic Acids Res.* **9,** 309–321.
20. Diver, W. P., Grinsted, J., Fritzinger, D. C., Brown, N. L., Altenbuchner, J., Rogowsky, P. & Schmitt, R. (1983) *Mol. Gen. Genet.* **191,** 189–193.
21. Barnes, W. M. & Bevan, M. (1983) *Nucleic Acids Res.* **11,** 349–368.
22. Barnes, W. M., Bevan, M. & Son, P. (1983) *Methods Enzymol.* **101,** 98–122.
23. Braun, G. & Cole, S. T. (1982) *Nucleic Acids Res.* **10,** 2367–2378.
24. Hawley, D. K. & McClure, W. R. (1983) *Nucleic Acids Res.* **11,** 2237–2255.
25. Rosenberg, M. & Court, D. (1979) *Annu. Rev. Genet.* **13,** 319–353.
26. Stormo, G. D., Schneider, T. D. & Gold, L. M. (1982) *Nucleic Acids Res.* **10,** 2971–2996.
27. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157,** 105–132.
28. Michaelis, S. & Beckwith, J. (1982) *Annu. Rev. Microbiol.* **36,** 435–465.
29. Dempsey, W. B., McIntire, S. A., Willetts, N., Schottel, J., Kinscherf, T. G., Silver, S. & Shannon, W. A., Jr. (1978) *J. Bacteriol.* **136,** 1084–1093.
30. Jackson, W. J. & Summers, A. O. (1982) *J. Bacteriol.* **151,** 962–970.
31. Silhavy, T. J., Benson, S. A. & Emr, S. D. (1983) *Microbiol. Rev.* **47,** 313–344.
32. Normore, W. M. (1976) in *CRC Handbook of Biochemistry and Molecular Biology*, ed. Fasman, G. D. (CRC Press, Cleveland, OH), Vol. 2, 3rd Ed., pp. 65–184.
33. Nakai, C., Kagamiyama, H., Nozaki, M., Nakazawa, T., Inouye, S., Ebina, Y. & Nakazawa, A. (1983) *J. Biol. Chem.* **258,** 2923–2928.
34. Maeda, K., Kneale, G. G., Tsugita, A., Short, N. J., Perham, R. N., Hill, D. F. & Petersen, G. B. (1982) *EMBO J.* **1,** 255–261.
35. Brown, N. L., Choi, C. L., Grinsted, J., Richmond, M. H. & Whitehead, P. R. (1980) *Nucleic Acids Res.* **8,** 1933–1945.