Mol Gen Genet (1986) 202:143–151

MGG

The nucleotide sequence of the mercuric resistance operons of plasmid R100 and transposon Tn*501*: further evidence for *mer* genes which enhance the activity of the mercuric ion detoxification system

Nigel L. Brown[1], **Tapan K. Misra**[2], **Joseph N. Winnie**[1], **Annette Schmidt**[2], **Michael Seiff**[2], **and Simon Silver**[2]

[1] Department of Biochemistry, University of Bristol, Bristol BS8 1TD, United Kingdom
[2] Department of Biology, Washington University, St. Louis MO 63130, USA

Summary. The DNA sequences of the mercuric resistance determinants of plasmid R100 and transposon Tn*501* distal to the gene (*merA*) coding for mercuric reductase have been determined. These 1.4 kilobase (kb) regions show 79% identity in their nucleotide sequence, and in both sequences two common potential coding sequences have been identified. In R100, the end of the homologous sequence is disrupted by an 11.2 kb segment of DNA which encodes the sulfonamide and streptomycin resistance determinants of Tn*21*. This insert contains terminal inverted repeat sequences and is flanked by a 5 base pair (bp) direct repeat. The first of the common potential coding sequences is likely to be that of the *merD* gene. Induction experiments and mercury volatilization studies demonstrate an enhancing but non-essential role for these *merA*-distal coding sequences in mercury resistance and volatilization. The potential coding sequences have predicted codon usages similar to those found in other Tn*501* and R100 *mer* genes.

Key words: Resistance-determinants – *Pseudomonas* genes – *Shigella* genes – Plasmid evolution – Homologous genes

Introduction

The mercuric resistance system found on bacterial plasmids and transposons is both the most complex of the heavy metal resistance systems and also that best understood (Summers and Silver 1978; Robinson and Tuovinen 1984; Foster 1983). The system is basically the same for plasmid R100 and transposon Tn*501* (Misra et al. 1984, 1985). First, there is a regulatory gene, *merR*, which encodes a trans-acting diffusable regulatory molecule (Foster et al. 1979; Misra et al. 1984; Barrineau et al. 1984). This is followed by an operator-promoter region, at which site the *merR* gene product acts. There then follows an operon consisting of three small genes in R100 (or two of these in Tn*501*) followed by the 1.7 kilobase (kb) gene *merA* for the enzyme mercuric reductase (Foster et al. 1979; Brown et al. 1983; NíBhriain et al. 1983; Misra et al. 1984, 1985; Barrineau et al. 1984). The small promoter proximal genes are thought to function in the transport of mercury from the outside of the cell to the cytoplasmic mercuric reductase enzyme. Beyond the *merA* gene, there is an additional gene, *merD*, of yet unknown function (NíBhriain et al. 1983). The evidence for this gene consists of the isolation of transposon insertion mutations (NíBhriain et al. 1983) and nitrosoguanidine-induced point mutations (T.G. Kinscherf and S. Silver, unpublished data) in which the lesions mapped just promoter-distal to *merA*. Both groups of mutants were more sensitive to Hg^{2+} than were cells with an intact *mer* operon, and in neither group were the K_m and heat resistance properties of the mercuric reductase enyme affected (NíBhriain et al. 1983). As a functional role for the *merD* gene has not been established, the existence of the gene has been questioned by Barrineau et al. (1984), with the suggestion that it is an artifact of the cloning and mutagenesis procedures used.

In this paper we present the nucleotide sequence of the region of the mercuric resistance operons of plasmid R100 and transposon Tn*501* from the end of the *merA* mercuric reductase gene (Brown et al. 1983; Misra et al. 1985) through to and beyond the position at which the sequences diverge. In the 1.4 kb DNA after the *merA* gene both systems contain a plausible open reading frame corresponding to the position of the *merD* gene and a second common open reading frame. Growth and mercury volatilization studies with a group of Tn*501* deletion mutants (Itoh et al. 1984; Haas et al. 1984; Itoh and Haas 1985) and the R100 *merD* insertion mutants (NíBhriain et al. 1983) provide further support for the existence of a determinant which adds to the level of mercuric ion resistance and affects the rate of mercury volatilization, but which is not essential for a low level of operon function. The nucleotide homology between transposon Tn*501* and plasmid R100 is very high (79%) out to a point at which the two sequences abruptly become dissimilar. Immediately beyond this point the nucleotide sequence of Tn*501* is homologous with the transposition region of transposon Tn*21* (Diver et al. 1983), on which transposon the mercury resistance determinant of R100 lies. This region of Tn*21* is separated from the *mer* genes by an 11.2 kb region, which contains the sulfonamide and streptomycin resistance determinants (de la Cruz and Grinsted 1982). The nucleotide sequence data show that this 11.2 kb has terminal inverted partial repeats and is flanked by 5 bp direct repeats. The function and evolution of the mercuric resistance operons (and surrounding DNA) can now be discussed with the help of the complete nucleotide sequences.

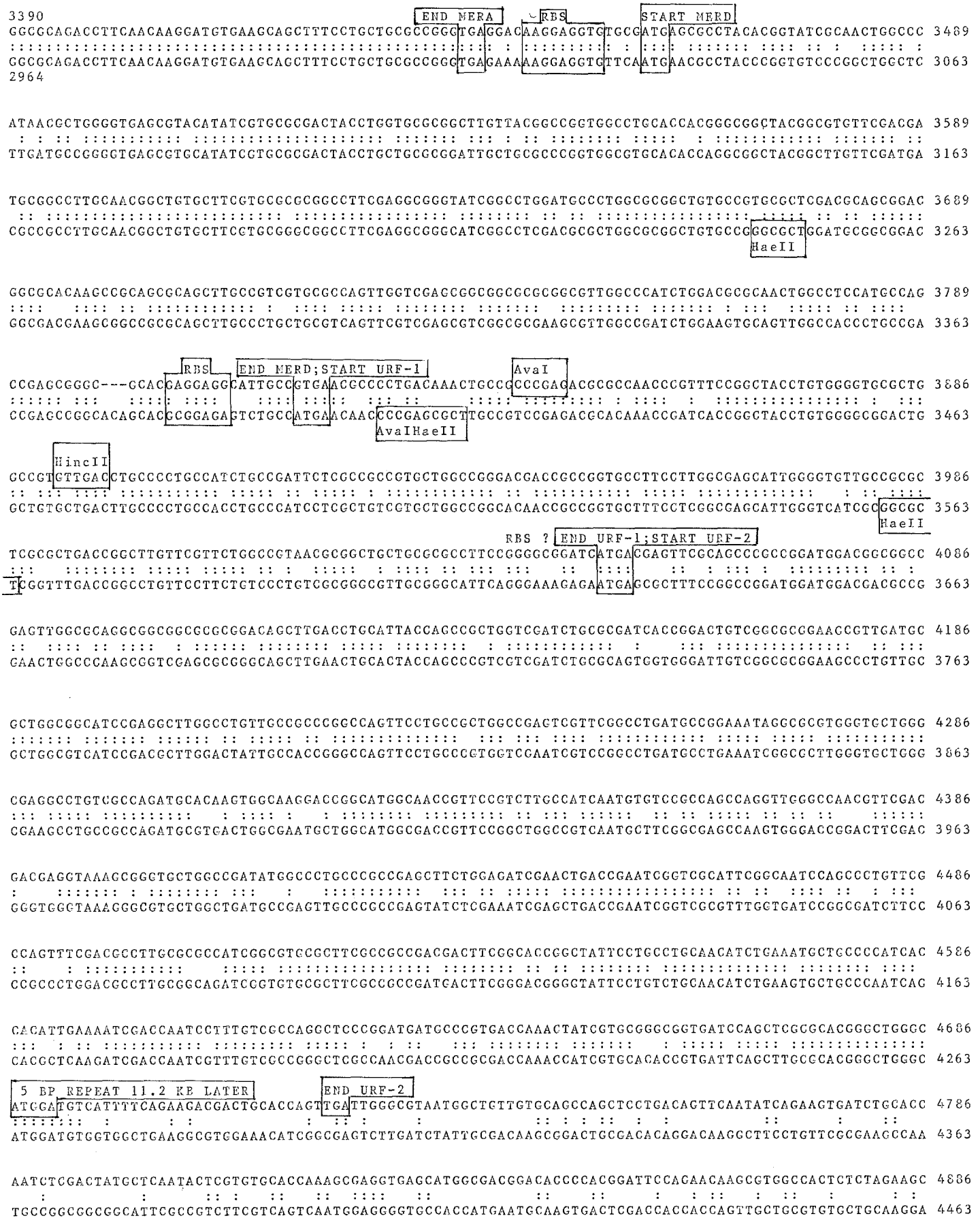
Materials and methods

The chain-termination nucleotide sequencing methods of Sanger et al. (1977) were used for both sequences. The

Offprint requests to: N.L. Brown

ALIGNMENT OF R100NERD SEQUENCE 1498 BP WITH TN501NERD SEQUENCE 1500 BP

K-tuple size=3;Window size=20;Gap penalty=20



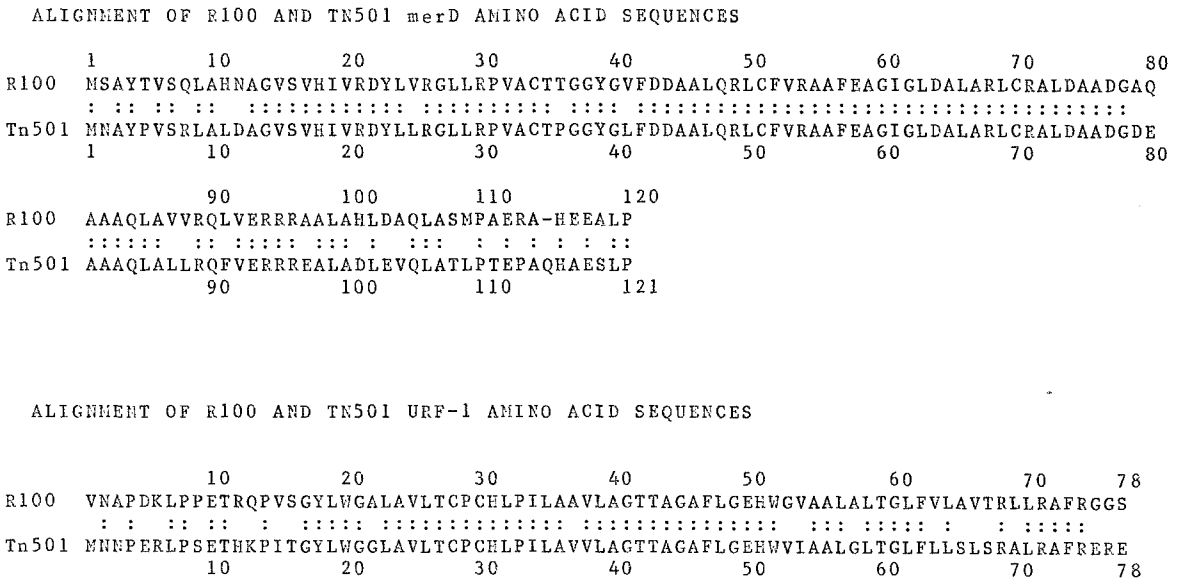


Fig. 2. The amino acid sequences deduced from translating the *merD* and URF-1 open reading frames marked in Fig. 1. The sequences were aligned using the program of Wilbur and Lipman (1983). The upper sequence in each case is that from R100. Standard one letter abbreviations for amino acids are used. Colons indicate identical amino acids. The dash indicates a missing amino acid in the R100 *merD* sequence when the two sequences were optimally aligned

Tn*501* sequence was determined from both strands with randomly cloned restriction enzyme fragments in bacteriophage M13mp7 (Messing et al. 1981) using methods previously described (Brown et al. 1983; Diver et al. 1983). The R100 sequences were determined on both strands by cloning 3 to 4 kb fragments from plasmid pDU1003 (NíBhriain et al. 1983) into M13mWB2344 and generating ordered deletions from the primer location (Barnes et al. 1983; Misra 1985).

Growth curves and ^{203}Hg volatilization assays were run as previously described (Clark et al. 1977; Foster et al. 1979; NíBhriain et al. 1983).

Results

Nucleotide sequence of the merD region

Figure 1 shows the nucleotide sequences of the last part of the mercuric resistance region of plasmid R100 and transposon Tn*501*, including the last 52 nucleotides of the *merA* gene, which were published previously (Brown et al. 1983; Misra et al. 1985). The first 1.3 kb of the two sequences shown in Fig. 1 have strong homologies, with identical base pairs at 79% of the positions. From position 4272 in Tn*501*, the sequence is entirely unrelated to that of R100. This position is preceded by the pentanucleotide sequence ATGGA (positions 4264–4268), which marks the endpoint of the homology between the transposition functions of Tn*501* and Tn*21* from the right hand end of both transposons (Diver et al. 1983). However, in Tn*21* the pentanucleotide appears twice, once at the end of the homology in the mercury resistance genes, and once at the start of the homology between the transposition regions; these two pentanucleotide sequences being separated by 11.2 kb. This is considered in more detail in the Discussion section.

Several restriction endonuclease sites which are important in delimiting functional regions of the mercury resistance determinants have been identified in the Tn*501* and R100 sequences. These are marked in Fig. 1, and are discussed in the section below on growth and mercury volatilization studies.

The *merA* gene termination codon is at position 3013 in the Tn*501* sequence (3439 in R100) and is followed by a potential ribosome binding site starting at position 3020 (3446 in R100), which has a run of 9 bases complementary to the 3′-end of 16S rRNA (Stormo et al. 1982); there is then an open reading frame that would code for a polypeptide of 121 amino acids ($M_r = 13014$ in Tn*501;* or of 120 amino acids, $M_r = 12723$ in R100; Fig. 2). These highly-conserved reading frames show 83% base pair identity and would encode polypeptides with 81% amino acid identity. From the available genetic data (NíBhriain et al. 1983), this is likely to be the coding sequence for the *merD* gene. The codon usage of the gene is very similar to that described for the other *mer* genes of Tn*501* and R100 (Brown et al. 1983; Misra et al. 1984, 1985), and the only base deletions

Fig. 1. The nucleotide sequences of the end of the mercuric resistance operon of plasmid R100 (*upper*) and transposon Tn*501* (*lower*). The sequences start toward the end of the 1.7 kb *merA* genes (Brown et al. 1983; Misra et al. 1985) and continue beyond the region of close homology between the sequences. The sequences were aligned using the program of Wilbur and Lipman (1983). Only the strand equivalent to the mRNA is shown. The numbering system for R100 starts immediately following the IS*1* insertion, while that for Tn*501* starts with the beginning of the inverted repeat (Misra et al. 1984). There is a difference of 426 bp in numbering at the end of *merA* due to an additional gene in R100 prior to the sequence shown. Colons between bases indicate identical nucleotides in those positions. Potential ribosomal binding sites (RBS), and start and stop codons are *boxed. URF,* unidentified reading frame. The three dashes indicate missing bases, when the two sequences were optimally aligned. The sites of some restriction nuclease cleavage targets (that have been used as described in the text) are indicated

```
  ALIGNMENT OF R100 MERD AND MERR AMINO ACID SEQUENCES

K-TUPLE SIZE=2;WINDOW SIZE=20;GAP PENALTY=3

                    10         20         30         40         50         60         70
merD       MSAYTVSQLAHNAGVSVHIVRDYLVRGLLRPVACTTGGYGVFDDAALQRLCFVRAAFEAGIGLDALARLCRALDAA
                :    :  ::: :   : :   ::::       :        :   :  ::  :   :  ::  : : :  :
merR  MENNLENLTIGVFAKAAGVNVETIRFYQRKGLLREPDKPYGSIRRYGEADVVRVKFVKSAQRLGFSLDEIAELLRLDDGT
                10         20         30         40         50         60         70

          80         90         100        110        120
merD  DGAQAAAQLAVVRQLVERRRAALAHLDAQLASMPAERAHEEALP
           :          :    : ::     :
merR  HCEEASSLAEHKLKDVREKMADLARMETVLSELVCACHARKGNVSCPLIASLQGEAGLARSAMP
                90         100        110        120        130        140




ALIGNMENT OF TN501 MERD AND MERR AMINO ACID SEQUENCES

K-TUPLE SIZE=2;WINDOW SIZE=20;GAP PENALTY=3

                    10         20         30         40         50         60         70
merD       MNAYPVSRLALDAGVSVHIVRDYLLRGLLRPVACTPGGYGLFDDAALQRLCFVRAAFEAGIGLDALARLCRALDAA
                :  ::: :   : :   :::       :        :   :  ::  :   :  ::  : : :  :
merR  MENNLENLTIGVFAKAAGVNVETIRFYQRKGLLLEPDKPYGSIRRYGEADVTRVRFVKSAQRLGFSLDEIAELLRLEDGT
                10         20         30         40         50         60         70

          80         90         100        110        120
merD  DGDEAAAQLALLRQFVERRREALADLEVQLATLPTEPAQHAESLP
           ::          :       ::  :  :  :
merR  HCEEASSLAEHKLKDVREKMADLARMEAVLSELVCACHARRGNVSCPLIASLQGGASLAGSAMP
                90         100        110        120        130        140
```

Fig. 3. Alignment of the amino acid sequences of the *merD* and the *merR* open reading frames for R100 and Tn*501* using the Wilbur and Lipman (1983) alignment program with parameters K-tuple = 2, window size 20 and gap penalty = 3

between the reading frames occur as a single group of three, thus maintaining the polypeptide sequence, as would be expected for a functional reading frame.

The amino acid sequence of *merD* was compared with those of all other *mer* operon genes using the protein alignment program of Wilbur and Lipman (1983). Surprisingly, the *merD* amino acid sequence showed a strong homology with the divergently-transcribed open reading frame now known to encode the *merR* regulatory DNA-binding protein (Foster and Brown 1985; Lund et al. 1985). These two amino acid sequences would be identical in 25 of the first 72 (34%) amino acids of the *merD* sequence (Fig. 3), which is a highly significant correlation. Some of the differences in additional positions are very conservative replacements of hydrophobic amino acids such as isoleucine to valine and other positions have amino acids that have been found in members of the family of DNA binding proteins thought to have a similar "helix-turn-helix" structure (Pabo and Sauer 1984) and recently studied by Ebright (1985). The invariant glycine that occurs at the turn position is Gly_{14} of *merD* and Gly_{18} of *merR* (Fig. 3); the hydrophobic bond would be between Ala_{14} in the first alpha helix and Ile_{24} in the second helix of *merR* (Ebright, personal communication); this would correspond to a hydrophobic bond between Ala_{10} and Val_{20} of the *merD* protein, for the same spacing interval. Valine is found in the comparable position to position 20 of *merD* in at least eight other members of this class of DNA-binding regulatory proteins (Fig. 12 of Pabo and Sauer 1984), including the repressors of the *lac* and *gal* operons. This possibility of a role for the *merD* polypeptide as a DNA binding regulatory protein is highly tentative, since it is based entirely on analysis of amino acid sequences of two open reading frames of the mercuric resistance system for which the corresponding polypeptides have not been directly identified. Furthermore, it is based largely on a theoretical analysis of sequences of other DNA-binding regulatory proteins (Pabo and Sauer 1984; Ebright 1985) for which there is only partial experimental support. The amino acid sequence homology in Fig. 3 is striking enough to justify presenting these data and analysis, but hardly convincing on this point.

The amino acid sequence predicted from the *merD* open reading frame (Fig. 2) shows no pronounced hydrophobic regions and no other striking structural features. The three cysteine residues are well separated on the linear sequence, rather than being closely paired as in other *mer* gene products. Cysteine pairing is suggested to be important in the role of these other gene products (Misra et al. 1984, 1985; Brown et al. 1983). Comparisons between the predicted amino acid sequence and the Protein Information Resource database revealed no significant homologies.

Overlapping the stop codon of the presumptive *merD* reading frame is the initiation codon for another potential reading frame, and this is preceded by a reasonable ribosomal binding site sequence (position 3381–3387 in Tn*501*; position 3804–3810 in R100). The reading frames in Tn*501* and R100 are highly conserved (76% base pair identity), and could encode polypeptides of 78 amino acids (Fig. 2) which show 72% amino acid identity. This reading frame has a codon usage similar to those of the known functional

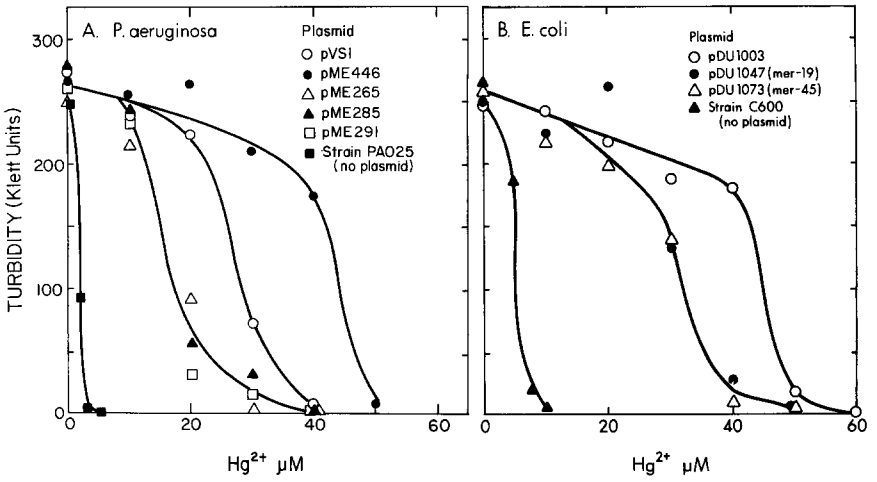


Fig. 4A, B. Inhibition by Hg^{2+} of growth of (**A**) *P. aeruginosa* strains without a plasmid or (■), with plasmids pVS1 (○), pME446 (●) pME265 (△), pME285 (▲) and pME291 (□) or (**B**) *E. coli* strains without a plasmid (▲) or with *merD* insertion mutants pDU1047 (●) or pDU1075 (△) of NíBhriain et al. (1983). Average turbidities for three separate growth experiments with each strain are shown. The standard deviations for turbidities of the uninhibited cells were ±4 or 5 Klett units. Turbidity was measured 7 h (*P. aeruginosa*) or 5 h (*E. coli*) after dilution of overnight cultures in nutrient broth into fresh broth and aeration by rotation of the tubes (late log phase of growth)

mer genes, although there is no direct evidence for a genetic function beyond *merD* which is involved in mercuric ion resistance. Again, homology comparisons and structural predictions yielded no information that was useful in determining a possible function for this polypeptide. Two cysteines are predicted, and these are close together, as observed in the predicted products of the promoter-proximal inducible *mer* genes. We will use the designation URF-1 for this *u*nidentified *r*eading *f*rame in this paper (Fig. 1).

Overlapping the termination codon for the URF-1 reading frame is the start of another potential reading frame (position 3628 in Tn*501*, and position 4051 in R100; Fig. 1). However, this reading frame goes beyond the break-point of homology between the Tn*501* and R100 systems. In Tn*501* the reading frame is 987 base pairs long, giving a potential polypeptide product of 329 amino acids; whereas in R100 the reading frame is 609 base pairs, giving a potential product of 203 amino acids. The reading frame in Tn*501* extends 151 nucleotides beyond the sequence presented in this paper, into the sequence presented by Diver et al. (1983); in that paper the TGA termination codon starts at nucleotide 493.

The first 369 bp of the R100 sequence shown in Fig. 1 have recently been published by Barrineau et al. (1984). The independent determinations agree, with the exception of the G at position 3721, which is absent from their sequence. We have carefully checked the sequence on both DNA strands, and consistently find a G at this position. The corresponding nucleotide in the Tn*501* sequence (a C) has also been checked as correct.

Growth and Mercury volatilization

Tn*501* variants with deletions of some of the sequence shown in Fig. 1 and R100 variants with transposon insertions in the *merD* region have been used in studies of inhibition of growth by Hg^{2+} and by organomercurials and of mercury volatilization. These build on the related studies by NíBhriain et al. (1983). Itoh and Haas (1985) constructed a series of smaller plasmids from the 29.2 kb plasmid pVS1 on which Tn*501* was first found (Stanisich et al. 1977). Plasmid pME265 (7.5 kb) contains the *mer* sequence from the left inverted repeat of Tn*501* (Misra et al. 1984) to the *Ava*I site at position 3404 of the Tn*501* DNA sequence (Fig. 1), i.e. to just beyond the end of *merD*. Plasmids pME291 (8.4 kb) and pME285 (10.6 kb) both contain *mer* DNA from the *Hin*dIII site at position 136 of Tn*501* (see Misra et al. 1984) to the *Hae*II site at position 3245 and thus do not contain a complete *merD* gene.

P. aeruginosa strains carrying the deletion plasmids pME265, pME285 and pME291 consistently showed a somewhat greater sensitivity to Hg^{2+} inhibition than the *P. aeruginosa* strain with plasmid pVS1 (Fig. 4A). There was no reproducible difference in sensitivity to Hg^{2+} between the three deletion plasmids. The slightly greater sensitivity to Hg^{2+} of these three plasmid-containing strains was similar to that with the R100 *merD* insertion mutants (Fig. 4B). This difference was reproducible and (although small) could lead to a two-to-five fold difference in cell mass (Fig. 4). The *P. aeruginosa* strain with plasmid pME446, which contains the entire mercuric resistance operon (Itoh and Haas 1985) was even more resistant to Hg^{2+} than was the isogenic strain with the wild type plasmid pVS1. As will be shown below, this appears to be due to constitutive expression of the mercury volatilization activity.

The *merD* insertion mutants of the R100 mercury resistance system in *E. coli* plasmid pDU1003 of NíBhriain et al. (1983) also showed an intermediate resistance to Hg^{2+} (Fig. 4B; see also NíBhriain 1985) and to the organomercurials merbromin, fluorescein mercuric acetate and *p*-hydroxymercuribenzoate (data not shown). The organomercurial resistance of *P. aeruginosa* with plasmids pME265, pME285, and pME291 was also intermediate between the strain with plasmid pVS1 and that with the plasmid-less sensitive strain (data not shown). Thus the *merD* region confers intermediate resistance toward both inorganic Hg^{2+} and toward the organomercurials toward which pVS1 confers resistance (Clark et al. 1977). Uninduced cells were used in the experiments in Fig. 3; when a similar experiment was run with *P. aeruginosa* strains that had been induced for mercury volatilization by exposure to 5 μM Hg^{2+} for 1 h prior to addition of higher Hg^{2+} concentrations, the differences in resistance levels persisted and the strain with plasmid pME446 was more resistant than that with pVS1; and that strain was more resistant to Hg^{2+} than were those with plasmids pME265, pME285 and pME291 (data not shown).

In order to understand better the basis for the somewhat greater Hg^{2+} sensitivity of the three deletion plasmids,

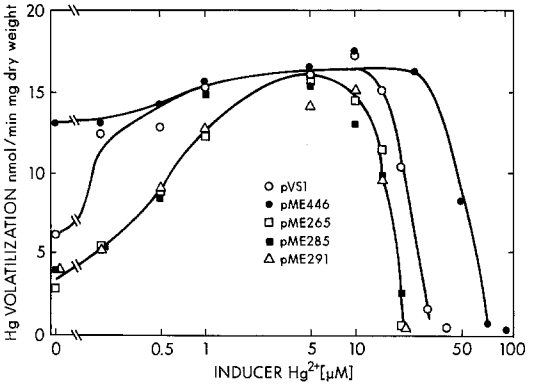


Fig. 5. Induction of mercury volatilization activity with Tn*501* deletion variants. *P. aeruginosa* strains with the indicated plasmids were grown into log phase of growth in nutrient broth and induced for 1 h at 37° C with the indicated concentration of Hg^{2+}. The cells were harvested by centrifugation and added (0.27 mg/ml dry weight) to assay buffer (NíBhriain et al. 1983) containing 5 μM $^{203}Hg^{2+}$. Samples were removed periodically and remaining radioactivity was measured to determine the initial rates of mercury volatilization

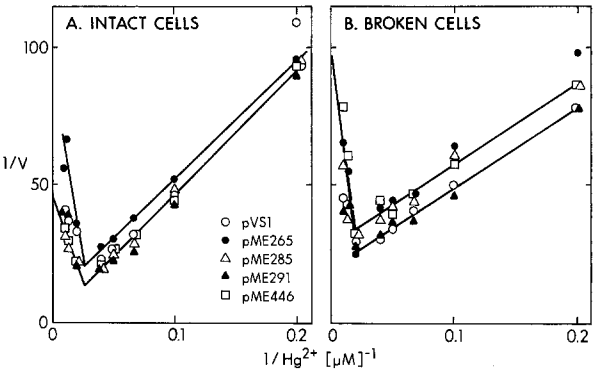


Fig. 6A, B. Mercury volatilization by (**A**) intact cells and (**B**) disrupted cells of *P. aeruginosa* with plasmid pVS1 (○) or deletion variants pME265 (●), pME285 (△), pME291 (▲), or pME446 (□). **A** The cells were grown, induced (with 5 μM Hg^{2+} in each case), harvested and assayed with the conditions in Fig. 5 but with the $^{203}Hg^{2+}$ concentration varying from 5 to 100 μM. **B** Portions of the same cells used intact in part A were disrupted by passage through a French Pressure Cell at 16,000 pounds per inch^2 (110 MPa). The assays of disrupted cells were run under the same conditions as with intact cells except 200 μM NADPH was added

pME265, pME285 and pME291 (Fig. 4A), we studied the effect of different inducer concentrations on the rate of volatilization of mercury by the cells under conditions of constant substrate concentration (5 μM). The results (Fig. 5) show that high induction occurred in the range 1 to 10 μM Hg^{2+} for the strains with the deletion plasmids pME265, pME285 and pME291, as well as with the wild type plasmid pVS1. There was a notable correlation between the mercury volatilization activity of cells previously exposed to high inducer concentrations (more than 20 μM Hg^{2+}; Fig. 5) and mercury resistance (Fig. 4A). Plasmid pME446 conferred the highest volatilization activity at inducer concentrations above 20 μM and protected against the highest concentrations of Hg^{2+} in growth studies; whereas pME265, pME285 and pME291 conferred the lowest activities and resistance to lower maximum Hg^{2+} concentrations. The strain with plasmid pME446 showed a partially constitutive mercury volatilization activity that was increased only by 30% on induction. The uninduced volatilization rates in the current experiment were higher than those observed in some earlier experiments (Nakahara et al. 1979; NíBhriain et al. 1983) but higher uninduced levels have been seen before (Clark et al. 1977). Cells containing pVS1 consistently showed a higher uninduced cell activity (about 30% of the optimally induced rate; Fig. 5) than cells containing the deletion plasmids. Net inhibition of whole cell volatilization activity was found after exposure for 1 hour to high Hg^{2+} concentrations. There was no difference between the induction kinetics of an *E. coli* strain with plasmid pDU1003 and strains with the *merD* mutant plasmids pDU1047 and pDU1073 in an experiment similar to that in Fig. 5 (data not shown).

As a first approach to establish whether the somewhat greater sensitivity of the Tn*501* deletion variants (Fig. 4A) was entirely due to a difference in induction (Fig. 5) or might also reflect a difference in volatilization activity, the appropriate *P. aeruginosa* strains with the plasmids were grown and induced at a constant and optimum mercuric ion concentration, and then assayed for volatilization activity of both intact whole cells and of disrupted cells (Fig. 6). There was no significant difference between the cells with different plasmids (Fig. 6A). The slightly lower rates for the cells with plasmid pME265 were not reproducible and probably reflected slightly lower induction in this experiment. Although there were no significant differences between the strains, the intact cells showed both a higher V_{max} and a higher K_m at lower substrate concentrations (i.e. before substrate inhibition; Fig. 6A) than did the disrupted cells (Fig. 6B). The ratio of the K_m values was approximately 8 and that of the V_{max} values about 3.

Discussion

To advance our understanding of the mechanism of mercury resistance, we have sequenced the mercuric resistance determinants of plasmid R100 and transposon Tn*501* from the beginning of the mercury resistance operon (Brown et al. 1983; Misra et al. 1984, 1985). There is still some question as to how far the mercuric resistance determinant extends in each of these systems. There is no doubt that the DNA from the *Hin*dIII site at position 136 of Tn*501* to at least the end of the *merA* gene at position 3015 is required, but the extent of the operon distal to this point is not certain. The sequence of this region of the DNA is presented for the first time in this paper. Together with the recently published sequence of the Tn*501 tnpA* gene (Brown et al. 1985), this completes the 8356 base pair sequence of Tn*501*. Although published in different sections, the Tn*501* sequence has been deposited in the EMBL library as a continuous sequence.

There is evidence from three reports (NíBhriain et al. 1983; Itoh and Haas 1985; and this paper) that mutations beyond the end of the *merA* gene result in a lowering of the mercuric resistance level, but not in a total loss of resistance nor of volatilization activity. The initial results of NíBhriain et al. (1983) with R100 led them to propose an additional gene, *merD*, that lay beyond the *merA* gene; and Itoh and Haas (1985) isolated *merA*-distal deletion mutants of Tn*501* which showed greater sensitivity to Hg^{2+}

toxicity. The nucleotide sequences of R100 and Tn*501* shown in Fig. 1 each contain a good candidate for the *merD* gene, with an open reading frame starting at position 3033 (in Tn*501*; or 3459 in R100) preceded by a potential ribosome binding sequence. Following this reading frame there is another unidentified open reading frame common to both sequences, URF-1, starting at nucleotide 3395 (in Tn*501*; and 3818 inR100), and preceded by a potential ribosome binding sequence. This reading frame overlaps by two base pairs the stop codon for *merD*. The polypeptide sequence obtained from translation of URF-1 (Fig. 2) showed no homology to other *mer* sequences nor to sequences in the PIR database. If the URF-1 sequence does not have a current function, one can wonder about the sequence conservation between R100 and Tn*501* in this region. It may also be involved in mercuric ion resistance (see below). We provisionally reserve the mnemonic "*merE*" for the gene represented by this reading frame, pending a more convincing elucidation of its role. A third reading frame found (URF-2; nucleotides 3628–4617 of Tn*501*) occupies all the remaining coding capacity of the DNA up to the site of action of the transposition resolvase (Rogowsky and Schmitt 1984), but the analogous reading frame in R100 (nucleotides 4051–4723) is disrupted by 11.2 kb of DNA starting at position 4692. URF-2 lacks a suitable ribosomal binding site preceeding it. The URF-2 reading frames of R100 and Tn*501* become unrelated after R100 position 4694 and terminate at different positions. That for R100 is shown in Fig. 1. There are other differences between the mercuric resistance systems of R100 and Tn*501* (for example, there is an additional reading frame in R100 immediately prior to the *merA* gene, which is not found in Tn*501*; Misra et al. 1985). The extra reading frame in R100 does not show sequence homology to URF-2 of Tn*501* or R100, either in the nucleotide sequence or between the predicted gene products.

The mercuric resistance operon of R100 is present on the transposon Tn*21*, the *tnpR* gene of which has been sequenced and compared with that of Tn*501* (Diver et al. 1983). The homology between the transposition functions starts at exactly the same place in Tn*501* at which the homology between the mercuric ion determinants stops. The two homology breakpoints in Tn*21* are 11.2 kb apart. Inspection of the DNA sequence from R100 (Tn*21*) at ends of the nonhomologous region (Fig. 7) reveal an inverted repeat sequence with 23 out of 25 base pairs matching, flanked by a direct repeat of five base pairs. This suggests an evolutionary scheme in which the sulfonamide and streptomycin resistance determinants of Tn*21* were present on a DNA fragment, with inverted repeats, which transposed into the URF-2 region of an ancestral mercury resistance transposon, and generated five base pair direct repeats as part of the transposition process. This ancestral mercuric resistance transposon may be the progenitor of both Tn*21* and Tn*501*, as envisaged by Grinsted and Brown (1984).

The results presented in this paper of growth and mercury volatilization experiments with a number of deletion mutants of Tn*501* have further demonstrated that sequences distal to the *Hae*II site at position 3246 of Tn*501* are not essential to mercuric ion resistance. However, mutants (pME285 and pME291) in which the deletions extend from this site (Haas et al. 1984; Itoh and Haas 1985) showed a reduced level of mercuric ion resistance as measured by their ability to grow on high levels of mercuric salts

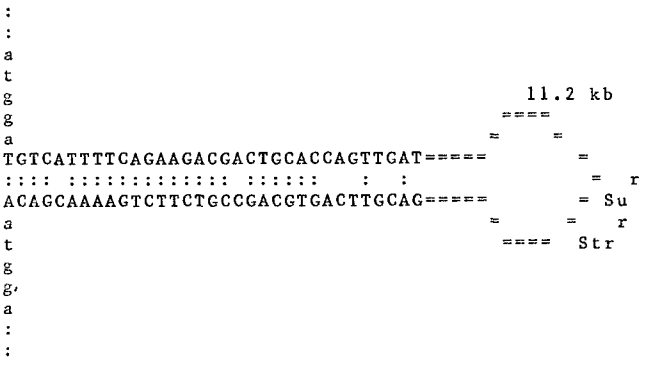


Fig. 7. Nucleotide sequence of R100 (Tn*21*) between the region of homology with the mercury resistance genes of Tn*501* (Fig. 1) and the region of homology with the transposition functions of Tn*501* (Diver et al. 1983). The sequence data are drawn as a single strand to emphasize the 5 bp direct repeat (*lower case*) at the homology breakpoint from both directions, and to show the partial inverted repeat structure at the ends of the non-homologous region, which is 11.2 kb in length and contains the determinants for sulfonamide (Su) and streptomycin (Str) resistances

(Fig. 4A), and in their tolerance to high concentrations of inducing mercuric ion (Fig. 5). In our experiments, the mutant pME265 was indistinguishable from the mutants pME285 and pME291 in its response to mercuric ion. The restriction data of Haas and coworkers (D. Haas, personal communication) suggest that the deletion in this mutant extends from the *Ava*I site at position 3404 (Fig. 1); pME265 would therefore contain the complete coding sequence designated *merD*. The relative roles of the *merA*-distal reading frames in this effect is not clear, but these data suggest that the reading frame designated URF-1 may play a role in the optimal expression of mercury resistance.

The designation of the number and function of genes involved in mercuric ion resistance requires the construction of well-defined mutants and the identification of the gene products involved in mercury resistance. The availability of the nucleotide sequences of two separate mercuric resistance operons will facilitate such efforts.

Nomenclature of the mer genes

This paper completes the nucleotide sequence of the *mer* operons of the R100 and Tn*501* systems. NíBhriain et al. (1983) identified two genes in R100 (*merC* and *MerD*) additional to those (*merR, merT* and *merA*) originally identified by Foster et al. (1979), and these genes were initially correlated with the available nucleotide sequence of Tn*501*. The R100 system was subsequently shown to have an extra reading frame compared with Tn*501* (Misra et al. 1985), and this has allowed N. NíBhriain to directly compare her genetic and physical mapping data with the nucleotide sequence. She concluded (personal communication) that the reading frame designated *merC* by Misra et al. (1984) is not the reading frame in which the *merC* mutants of NíBhriain et al. (1983) are located. Instead, the mutants are in the extra reading frame of R100. Of her 4 mutants designated *merT*, only 1 was in the reading frame designated *merT* by Misra et al. (1984) and Barrineau et al. (1984); the remainder are in the reading frame designated *merC* in the nucleotide sequence (Misra et al. 1984; Barrineau et al. 1984). Because of this confusion we propose a new nomenclature of the mercury resistance genes of R100

and Tn*501*. This is with the agreement of the other groups who are studying these systems (T.J. Foster in Dublin and A.O. Summers in Georgia).

The regulatory gene retains the name *merR*, and is the first gene of the mercuric resistance determinant. It should be noted that there is an increasing body of evidence for this gene being transcribed divergently from the rest of the operon (Foster and Brown 1985; Lund et al. 1985), and not in the same direction as suggested by Misra et al. (1984) and Barrineau et al. (1984). The operator-promoter region precedes a polycistronic region, which is likely to form a single transcriptional unit. The genes are in order *merT, merP, merC, merA, merD* and URF-1. The *merT* gene encodes a very hydrophobic polypeptide which is likely to be a membrane protein; the *merP* (for possibly periplasmic) gene in R100 contains mutants defective in membrane uptake, which were originally mapped in *merT* (NíBhriain et al. 1983) and is the reading frame designated *merC* by Misra et al. (1984). Jackson and Summers (1982) and Barrineau et al. (1984) have preliminary evidence that this gene may code for a periplasmic polypeptide. The *merC* gene is not present in Tn*501* and the function of this gene is unknown, although there are transposon-insertion mutants of this gene (NíBhriain et al. 1983) and Barrineau et al. (1984) suggested that it may be the gene for a second inner membrane protein found in minicell experiments. Barrineau and Summers (1983) independently isolated a series of mutations in this region, which they also called *merC*. This is followed by the *merA* gene (Brown et al. 1983; Misra et al. 1985). The *merA*-distal region is the subject of this paper, and is proposed to consist of gene *merD*, for which point and transposon-insertion mutants are available in R100, and, possibly another gene, URF-1, for which there is some evidence of a role in mercuric ion resistance, since the Tn*501* deletion mutant pME265 has altered Hg^{2+} resistance (Fig. 4) while it has an intact *merD* gene and a disrupted URF-1 reading frame.

The R100 sequence is part of a transposon, Tn*21* (de la Cruz and Grinsted 1982). The start of the nucleotides of Tn*501* and Tn*21* include closely related inverted repeated sequences (Misra et al. 1984; Grinsted and Brown 1984). The close (over 80% base identities) homology between the two systems continues from the inverted repeats for more than 4 kb, except for the region of *merC* (which is the reading frame found in R100 but not in Tn*501*; Misra et al. 1985). After the sequences shown in Fig. 1, and the 1.2 kb sequence shown in Fig. 2 of Diver et al. (1983) (which overlaps the Tn*501* sequence shown here by 233 base pairs), there remains only the sequence of the *tnpA* transpose gene of Tn*501* to complete the entire 8356 bp of Tn*501*. That sequence was recently reported (Brown et al. 1985). Direct comparison of the two sequences allowed us to localize a 11.2 kb insert in Tn*21* beginning at position 4692, before which and after which the homologies between the two systems are very close. The presumedly ancestral tansposition event led to a 5 bp repeat bounding inverted repeated sequences (Fig. 7). This 11.2 kb element is probably no longer functional as a transposon, or at least there is no evidence for transposition of this segment, which includes genes determining resistances to sulfonamides and streptomycin (de la Cruz and Grinsted 1982).

Acknowledgements. We thank W.M. Barnes for lessons on the ways and lore of dideoxyribonucleotide sequencing and data handling. Georg Jander, Lisa Baier, D.C. Fritzinger and R.D. Pridmore helped with the nucleotide sequencing and Jerome Harville with the growth and volatilization experiments. The frequent exchanges of preliminary information between our laboratories and N. NíBhriain, T.J. Foster, J. Grinsted, and D. Haas greatly facilitated our efforts. This work was supported by U.S.A. National Institutes of Health grants AI15672 and GM24956 and U.K. Medical Research Council grants G9791047CB and G8218936CB. N.L.B. is a Royal Society E.P.A. Cephalosporin Fund Senior Research Fellow.

References

Barnes WM, Bevan M, Son PH (1983) Kilo-sequencing: Creation of an ordered nest of asymmetric deletions across a large target sequence carried on phage M13. Methods Enzymol 101:98–122

Barrineau PJ, Summers AO (1983) A second positive regulatory function in the *mer* (mercury resistance) operon. Gene 25:209–221

Barrineau P, Gilbert P, Jackson WJ, Jones CS, Summers AO, Wisdom S (1984) The DNA sequence of the mercury resistance operon of the IncFII plasmid NR1. J Mol Appl Genet 2:601–619

Brown NL, Ford SJ, Pridmore RD, Fritzinger DC (1983) Nucleotide sequence of a gene from the *Pseudomonas* transposon Tn*501* encoding mercuric reductase. Biochemistry 22:4089–4095

Brown NL, Winnie JN, Fritzinger D, Pridmore RD (1985) The nucleotide sequence of the *tnpA* gene completes the sequence of the *Pseudomonas* transposon Tn*501*. Nucleic Acids Res 13:5657–5669

Clark DL, Weiss AA, Silver S (1977) Mercury and organomercurial resistance determined by plasmids in *Pseudomonas*. J Bacteriol 132:186–196

de la Cruz F, Grinsted J (1982) Genetic and molecular characterization of Tn*21*, a multiple resistance transposon from R100.1. J Bacteriol 151:222–228

Diver WP, Grinsted J, Fritzinger DC, Brown NL, Altenbuchner J, Rogowsky P, Schmitt R (1983) DNA sequences of and complementation by the *tnpR* genes of Tn*21*, Tn*501* and Tn*1721*. Mol Gen Genet 191:189–193

Ebright RH (1985) Proposed amino acid-base pair contacts for thirteen sequence-specific DNA binding proteins. In: Oxender D (ed) Protein structure, folding, and design. Alan R. Liss, New York, in press

Foster TJ (1983) Plasmid-determined resistance to antimicrobial drugs and toxic metal ions in bacteria. Microbiol Rev 47:361–409

Foster TJ, Brown NL (1985) Identification of the *merR* gene of R100 by using *mer-lac* gene and operon fusions. J Bacteriol 163:1153–1157

Foster TJ, Nakahara H, Weiss AA, Silver S (1979) Transposon *A*-generated mutations in the mercuric resistance genes of plasmid R100-1. J Bacteriol 140:167–181

Grinsted J, Brown NL (1984) A Tn*21* terminal sequence in Tn*501*: Complementation of *tnpA* gene function and transposon evolution. Mol Gen Genet 197:497–502

Haas D, Cryz SJ Jr, Itoh Y, Leisinger T, Luthi E, Mercenier A, Reimmann C, Rella M, Soldati L, Watson JM, Wretlind B (1984) Some applications of transposon insertion mutagenesis in *Pseudomonas*. In: Heslot H (ed) Genetique des micro-organismes industriels. Societe Francaise de Microbiologie, Paris, pp 91–111

Itoh Y, Haas D (1985) Cloning vectors derived from the *Pseudomonas* plasmid pVS1. Gene 36:27–36

Itoh Y, Watson JM, Haas D, Leisinger T (1984) Genetic and molecular characterization of the *Pseudomonas* plasmid pVS1. Plasmid 11:206–220

Jackson WJ, Summers AO (1982) Biochemical characterization of $HgCl_2$-inducible polypeptides encoded by the *mer* operon of plasmid R100. J Bacteriol 151:962–970

Lund PA, Ford SJ, Brown NL (1985) Transcriptional regulation of the mercury-resistance genes of transposon Tn*501*. J Gen Microbiol, in press

Messing J, Crea R, Seeburg PH (1981) A system for shotgun DNA sequencing. Nucleic Acids Res 9:309–321

Misra TK (1985) A new strategy to create ordered deletions for rapid nucleotide sequencing. Gene 34:263–268

Misra TK, Brown NL, Fritzinger DC, Pridmore RD, Barnes WM, Haberstroh L, Silver S (1984) 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. Proc Natl Acad Sci USA 81:5975–5979

Misra TK, Brown NL, Haberstroh L, Schmidt A, Goddette D, Silver S (1985) Mercuric reductase structural genes from plasmid R100 and transposon Tn*501*: functional domains of the enzyme. Gene 34:253–262

Nakahara H, Silver S, Miki T, Rownd RH (1979) Hypersensitivity to Hg^{2+} and hyperbinding activity associated with cloned fragments of the mercurial resistance operon of plasmid NR1. J Bacteriol 140:161–166

NíBhriain N (1985) Molecular genetic analysis of the mercury resistance (*mer*) operon of plasmid R100. Ph D Thesis, Trinity College, Dublin

NíBhriain N, Silver S, Foster TJ (1983) Tn*5* insertion mutations in the mercuric ion resistance genes derived from plasmid R100. J Bacteriol 155:690–703

Pabo CO, Sauer RT (1984) Protein-DNA recognition. Annu Rev Biochem 53:293–321

Robinson JB, Tuovinen OH (1984) Mechanisms of microbial resistance and detoxification of mercury and organomercury compounds: Physiological, biochemical and genetic analyses. Microbiol Rev 48:95–124

Rogowsky P, Schmitt R (1984) Resolution of a hybrid cointegrate between transposons Tn*501* and Tn*1721* defines the recombination site. Mol Gen Genet 193:162–166

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Stanisich VA, Bennett PM, Richmond MH (1977) Characterization of a translocation unit encoding resistance to mercuric ions that occurs on a nonconjugative plasmid in *Pseudomonas aeruginosa*. J Bacteriol 129:1227–1233

Stormo GD, Schneider TD, Gold LM (1982) Characterization of translational initiation sites in *E. coli*. Nucleic Acids Res 10:2971–2996

Summers AO, Silver S (1978) Microbial transformations of metals. Annu Rev Microbiol 32:637–672

Wilbur WJ, Lipman DJ (1983) Rapid similarity searches of nucleic acid and protein data banks. Proc Natl Acad Sci USA 80:726–730

Communicated by D. Goldfarb

Received September 3, 1985