Molecular Microbiology (1988) 2(1), 109–119

# Identification and sequencing of the *Escherichia coli cet* gene which codes for an inner membrane protein, mutation of which causes tolerance to colicin E2

**L. S. Drury and R. S. Buxton***
*National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK.*

## Summary

**Dominant mutations of the *cet* gene of *Escherichia coli* result in tolerance to colicin E2 and increased amounts of an inner membrane protein with an $M_r$ of 42 000. We have cloned the *cet*$^+$ gene and sequenced its DNA, revealing that the gene product, coded by the longest open-reading frame, has an $M_r$ of 49 772, with five predicted transmembrane structures towards its carboxy terminus and one at its amino terminus. We have demonstrated that the *cet* locus does in fact code for the inner membrane protein that is present in increased amounts in *cet* mutants, and we have shown that this increased amount of Cet protein is the result of enhanced transcription. The *cet* gene is shown to be in the same operon as the *phoM* gene, which is required in a *phoR* background for expression of the structural gene for alkaline phosphatase, *phoA*. Although the Cet protein is not required for *phoA* expression, our experiments suggest that the Cet protein has an enhancing effect on the transcription of *phoA*. No effect of phosphate concentration on *cet* or *phoM* gene expression could be found and thus their primary function may not be connected to the phosphate regulon.**

## Introduction

Colicin E2 is a protein antibiotic, produced by strains of the Enterobacteriaceae carrying the plasmid ColE2. It kills other sensitive strains by causing rapid and extensive degradation of DNA (Nomura, 1983) and is itself an endonuclease (Schaller and Nomura, 1976). We have isolated and investigated colicin-E2-tolerant mutants of *E. coli* (Cet$^-$) which still adsorb colicin E2 to the outer membrane receptor but are blocked at some stage in the transport of the colicin to its intracellular target, in the hope that these mutants would be impaired in new and interesting membrane functions (Buxton and Holland, 1973; 1974).

Received 7 July, 1987; revised 10 September, 1987. *For correspondence.

Genetical and biochemical analyses revealed that certain point mutations in *cet* were dominant (Buxton and Holland, 1973) and resulted in an increased amount of an inner membrane protein (Samson and Holland, 1970). This presumably blocked colicin E2 from reaching its intracellular target without affecting the colicin receptor functions. On the other hand, deletion of the *cet* gene did not result in colicin-E2-tolerance (unpublished data). The dominant nature of the *cet* point mutations was compatible with *cet* being either a regulatory gene, or coding for an envelope protein, or being a protein normally present in the cytoplasm, a mutation of which caused it to bind to the envelope (Buxton and Holland, 1973). To resolve these possibilities, we have now undertaken a molecular analysis of this gene by cloning it and sequencing its DNA.

## Results

### *Cloning of the* cet *gene and identification of its gene product*

The *cet* gene was identified by the isolation of mutants tolerant to colicin E2, i.e. still adsorbing this colicin and therefore not blocked in outer membrane receptor functions (Hill and Holland, 1967). The genetic location of *cet* was found to be between *serB* and *thr*, the gene order being *serB-trpR-cet-thr* (Buxton and Holland, 1973). We have cloned a 6 kb *Sal*I fragment from the region between *trpR* and *thr* (Buxton and Drury, 1983b) and found that it codes for, amongst others, a protein of $M_r$ 42 000 (Buxton and Drury, 1984). This was a possible candidate for the Cet protein since a strain with a dominant *cet* mutation has an increased amount of an inner membrane protein with an estimated $M_r$ of 43 000 (Samson and Holland, 1970). This protein disappeared when a *Kpn*I fragment was removed from the 6 kb *Sal*I fragment (Buxton and Drury, 1984). Tn*1000* insertions into this region of the pRB38 plasmid have been analysed using the maxicell system, and the results, presented in Fig. 1, show that the $M_r$ 42 000 band disappeared when Tn*1000* was inserted in a region between the *dye* gene and the nearest *Kpn*I site.

We have attempted directly to correlate this region with the colicin-E2-tolerant phenotype. To do this we transferred the *cet2* allele from strain RB208 (*cet2*) onto the pRB38 plasmid (*cet*$^+$) by homogenotization. This was done by transforming RB208 with pRB38, pooling these

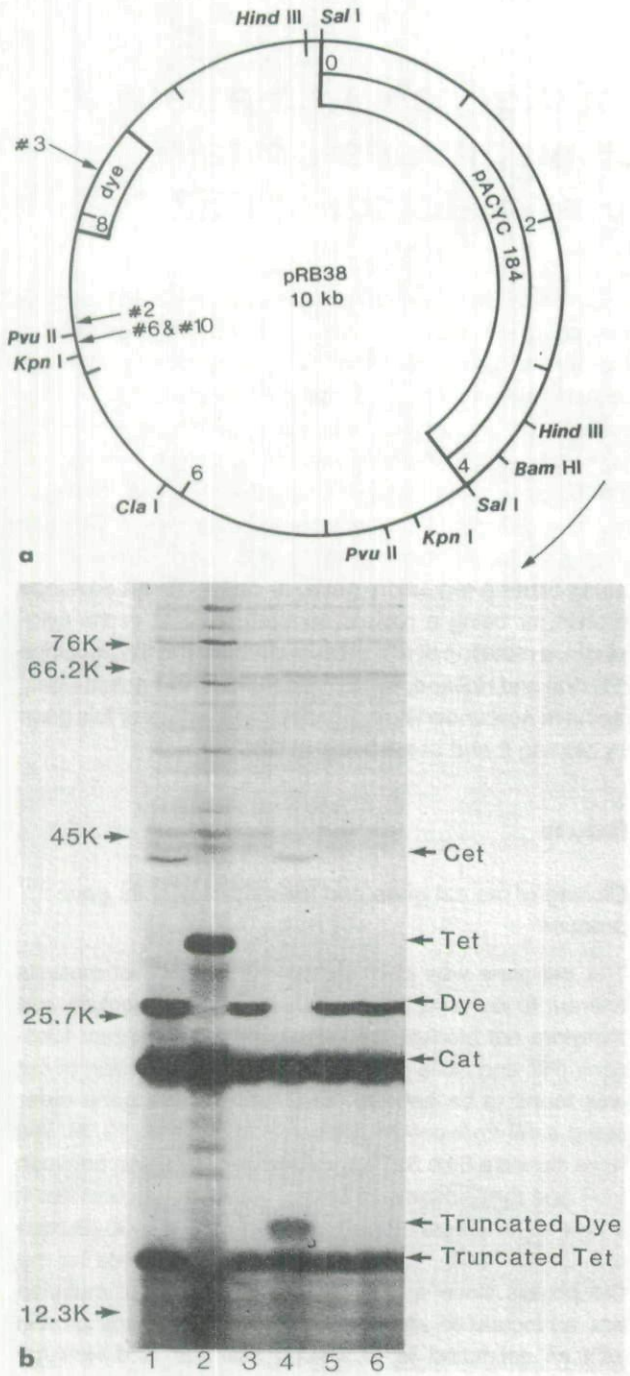


**Fig. 1.** Restriction map of plasmid pRB38 (a) and SDS-PAGE of maxicell experiments using this plasmid and its derivatives. (b) Numbers are sizes in kb. The positions of Tn*1000* insertions into the plasmid are marked (#). All the *Hin*dIII, *Kpn*I, *Sal*I and *Bam*HI restriction sites are shown, as well as the *Cla*I and *Pvu*II sites within the chromosomal DNA. The arrow shows the direction of transcription from the strong *tet* promoter (Stüber and Bujard, 1981). These plasmids were transformed into strain CSR603 and used in a maxicell experiment as described in *Experimental procedures*. The autoradiogram (b) shows the [$^{35}$S]-methionine-labelled polypeptides from this experiment after separation by SDS-PAGE.

The acrylamide concentration was 11%: track 1 was pRB38; 2, pACYC184; 3, Tn*1000* insert #2 in pRB38; 4, insert #3; 5, insert #6; 6, insert #10. The molecular-weight markers were: 77 K, ovotransferrin (hen egg); 66.2 K, albumin (bovine serum); 45 K ovalbumin (hen egg); 25.7 K, chymotrypsinogen A (bovine); 12.3 K, cytochrome C (equine).

transformants, extracting the plasmid DNA and using this to transform RB2181 (Δ *serB-thr*). After plating onto agar plates containing chloramphenicol (to retain pRB38) and colicin E2, three transformant clones were found which had the Cet-tolerant phenotype, i.e. E2-tolerant and colicin-E3-sensitive. These carried the *cet2* allele on the pRB38 plasmid and one such plasmid was named pRB62. Its presence conferred a very slow growth rate on the host strain. Restriction endonuclease digestion showed pRB62 to be the same size and contain the same restriction sites as those known to be in pRB38 (data not shown). In a maxicell experiment, pRB62 produced an increased amount of the $M_r$ 42 000 protein (data not shown).

Removal of the *Kpn*I fragment from pRB62 resulted in a colicin-E2-sensitive phenotype in a Δ(*serB-thr*) strain. This is because although *cet2*/*cet*$^+$ partial diploids are E2-tolerant, i.e. the *cet2* allele is dominant to *cet*$^+$, Δ *cet* strains are E2-sensitive (Buxton and Holland, 1973 and unpublished data). Attempts to isolate Tn*1000* insertions in the *cet* gene of pRB62 were unsuccessful since the *cet2* allele appeared to revert very readily. This is perhaps understandable since pRB62 carrying *cet2* had such a dramatic effect on reducing growth rate.

We have also looked at the envelope proteins from strains carrying pRB38 (*cet*$^+$) and pRB62 (*cet2*) (Fig. 2). It can be seen that there is a large increase in an inner membrane protein of $M_r$ 42 000 in the strain carrying pRB62 when compared with the strain carrying pRB38. This is likely to be the Cet protein, since it co-electrophoreses with the $M_r$ 42 000 protein seen in maxicells (data not shown). Thus, we have demonstrated that part, at least, of the *cet* gene lies between the two *Kpn*I sites, that this gene codes for a protein of $M_r$ 42 000, and that the latter is probably the same inner membrane protein that is present in increased amounts in *cet* mutants.

### *DNA sequencing of the* cet *gene*

A 0.7 kb *Cla*I-*Kpn*I fragment (see Fig. 3) was subcloned from the 6 kb *Sal*I fragment into a pair of complementary M13 vectors, mp18 and mp19, cleaved with *Acc*I and *Kpn*I. Shotgun cloning from the *Cla*I-*Kpn*I fragment was also used to generate clones. Dideoxy sequencing of these was used to generate the sequence shown in Fig. 4. Using a *Cla*I-*Pvu*II fragment, we were able to join this sequence to adjacent DNA sequences, and using a *Kpn*I-*Sst*I fragment from pRB50, which has Tn*1000* inserted into the 5′ end of the *dye* gene and hence a new *Sst*I site within the transposon, we were able to read across the *Pvu*II site into DNA which was sequenced previously (Drury and Buxton, 1985).

Analysis of this sequence revealed an open-reading frame (ORF) from bp 344 to bp 1693 coding for a protein of $M_r$ 49 772. This ORF ends at a stem-and-loop structure

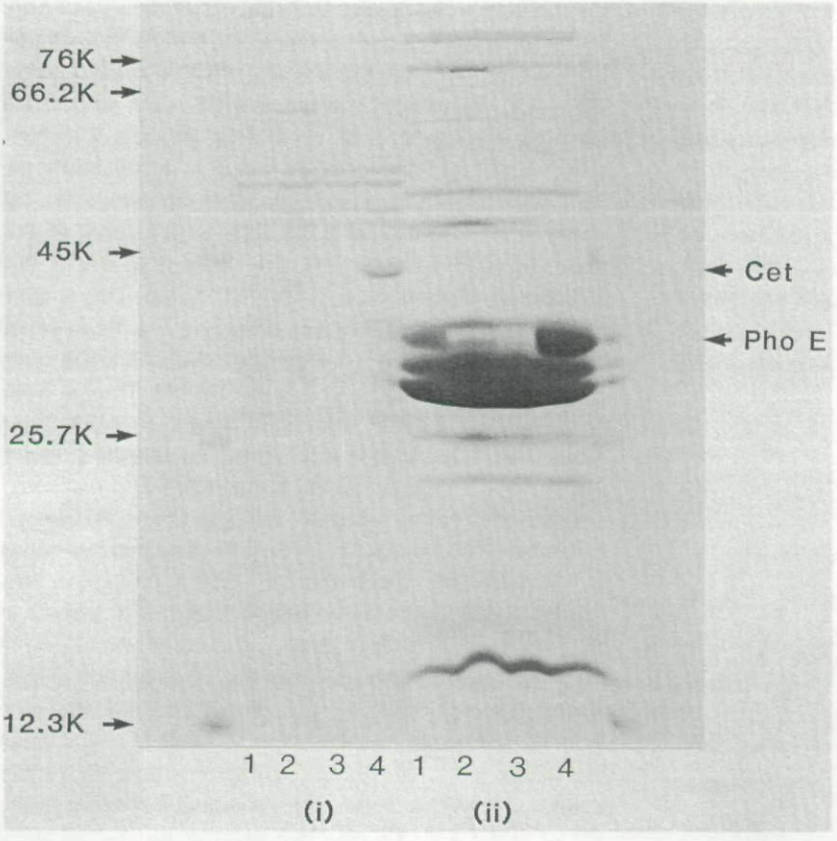


**Fig. 2.** SDS-PAGE of inner (i) and outer membrane (ii) proteins, extracted with 0.5% Sarkosyl, and stained with Coomassie blue. The tracks were: 1, strain BW1308; 2, strain BW3218; 3, strain BW3218/pRB38; 4, strain BW3218/pRB62. The acrylamide concentration was 11%. The molecular-weight markers were the same as in Fig. 1. The position of the Cet protein in the inner membrane fraction is marked, together with the position of an outer membrane protein, possibly PhoE, which is also increased in strain BW3218/pRB62.

typical of a rho-independent transcriptional terminator (Drury and Buxton, 1985). This terminator appears also to be the terminator for the *dye* gene, transcribed in the opposite direction, which we have also recently sequenced (Drury and Buxton, 1985). The start of translation for the ORF could be at either one of three AUG triplets: the first, at bp 344, is preceded by a good Shine-Dalgarno ribosome-binding site (Shine and Dalgarno, 1975), namely AGGAGA at bp 333-338 (Fig. 4).

The size of the protein over-produced in *cet* mutants ($M_r$ 42 000) was compatible either with it being encoded by the longest possible ORF, or by a shorter ORF. There are also other possible start sites if the less common initiator codons GUG, UUG and AUU are considered (see review by Kozak, 1983). Thus without protein sequence data, we cannot be certain which start site is correct.

There is another ORF open from the *Cla*I site and terminating at a stop codon at bp 284–286. This is likely to be the 3′ end of *phoM*, which is known to map upstream from *cet* (Tommassen *et al.*, 1984; Makino *et al.*, 1984). *phoM* is one of three positive regulatory genes for the *pho* regulon, which is a group of genes involved in the transport and metabolism of phosphate. We cannot find a very satisfactory promoter sequence for *cet* in this region. A likely possibility is that *phoM* and *cet* are transcribed co-ordinately on a polycistronic transcript.

### *Predicted amino acid sequence of the Cet protein*

We compared the amino acid sequence of the Cet protein deduced from the DNA sequence using the Dayhoff protein sequences data base. No very convincing homologies were found.

The protein is quite hydrophobic, with 60% of the residues being non-polar. The number of charged amino acids is relatively low, being 16–17%; the arginine plus lysine content is approximately 7% and the glutamic plus aspartic acid content is 9%. The protein is therefore acidic, having a net charge of −8.

The putative transmembrane sequences have been identified by their hydropathy. Kyte and Doolittle (1982) reported that the hydropathy of membrane-spanning domains generally averages greater than +1.6 over a 19-residue segment. From the Kyte and Doolittle plot shown in Fig. 5, there appear to be six possible transmembrane sequences. There is one such possible sequence between residues 3 and 28 and there are 5 such sequences towards the carboxy-terminus of the protein.

### *Northern blot analysis of the* cet *gene*

As mentioned before, *cet* mutants have an increased amount of an inner membrane protein of $M_r$ 43 000, as

estimated by Samson and Holland (1970). *cet* mutations were found to be dominant to the *cet*$^+$ wild-type gene (Buxton and Holland, 1973), which suggested that these mutations were of a regulatory nature resulting in increased amounts of the Cet protein that in some way blocked entry of the colicin E2 molecule into the cell.

To test whether there was increased transcription of the *cet* gene in a *cet* mutant, we performed Northern blot analysis on total RNA isolated from RB85, a *cet*$^+$ strain, and RB208 carrying the *cet2* allele. This RNA was probed with the double-stranded *Pvu*II-*Taq*I DNA probe I, which, from the DNA sequencing data, should only cover the *cet* gene. Similar results were obtained with probe II (*Cla*I-*Kpn*I fragment) covering part of *phoM* as well as *cet*. As can be seen from Fig. 6a, the *cet2* mutant produces considerably more of a transcript approximately 1.5 kb long, under a variety of growth conditions. It seemed likely, therefore, that the *cet* mutation was a change in a regulatory sequence upstream from the start site of *cet* translation. *cet* mutants were more easily obtainable from some strains such as ASH10 and HfrH, than from others (Hill and Holland, 1967; Holland and Threlfall, 1969). This is compatible with the Northern blot shown in Fig. 6c in which there is a faint band at 1.5 kb in strain ASH10, at the same position as the major band in strain RB208. Thus, a weak promoter may exist in some strains such as ASH10 and HfrH which is mutated in a *cet* mutant to activate it, resulting in a great increase in *cet* transcription.

When the Northern-blot autoradiograms were exposed for longer periods (Fig. 6b), it could be seen that two bands at approximately 4.5 kb and 4.4 kb were present in both strains RB85 (*cet*$^+$) and RB208 (*cet2*), but absent in RB979 (Δ *cet*) and ASH10 (*cet*$^+$). This could be a transcript from the putative operon which includes *cet* and *phoM*, although we do not know why it was absent from strain ASH10. There was no increase in amounts of any of these transcripts when the bacteria were grown in low-phosphate media; if anything, there was an actual decrease of the 4.4 kb and 4.5 kb transcripts.

### *Effect of a* cet *mutation on expression of genes in the* pho *regulon*

*phoM* is required for the expression of *phoA* encoding alkaline phosphatase, in a *phoR* background (Wanner and

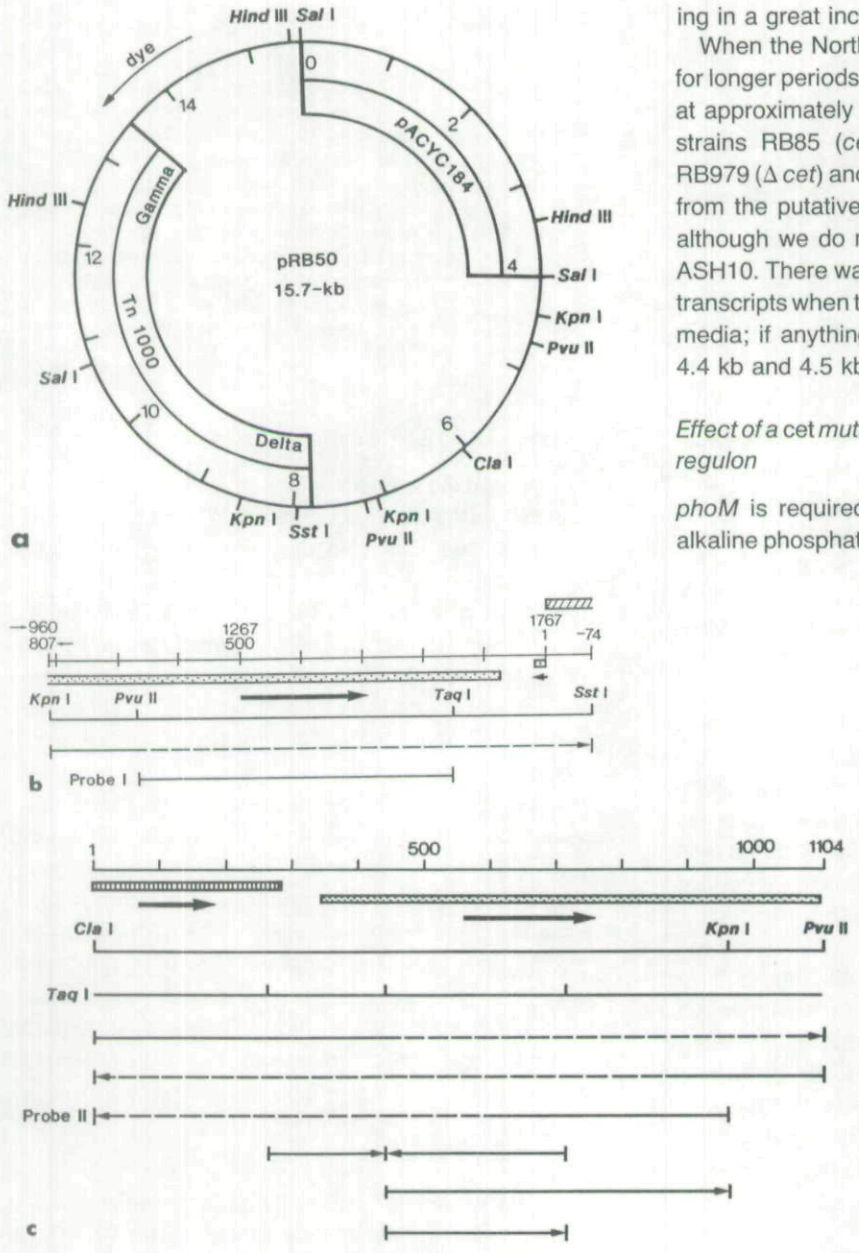


**Fig. 3.** Restriction map of plasmid pRB50 (a) and sequencing strategy of clones derived from this plasmid (b) and from plasmid pRB38 (c). Numbers in (a) are kb. The transposon Tn*1000* has been inserted in the 5′ end of the *dye* gene. All the *Hin*dIII, *Kpn*I, *Sal*I and *Sst*I restriction sites are shown, as are the *Cla*I and *Pvu*II sites within the cromosomal DNA. The *Sst*I-*Kpn*I fragment shown in (b) was derived from pRB850 and cloned into the M13 sequencing vectors, mp18 and mp19. The boxed areas comprise the *dye* 3′ coding region (·····), the *cet* 3′ coding region (·.·.·.·) and the δ end of Tn*1000* (////////). Nucleotides are numbered from the junction of bacterial and Tn*1000* DNA (marked ←), and from the *Cla*I site (marked →). The broad arrows represent the directions of transcription (⟶). The *Cla*I-*Pvu*II fragment was derived from pRB38 and cloned as above. The symbols are the same except that the 3′ end of the coding region of *phoM* is marked (||||||). The double-stranded DNA probes used in the Northern blots (Fig. 6) are shown.

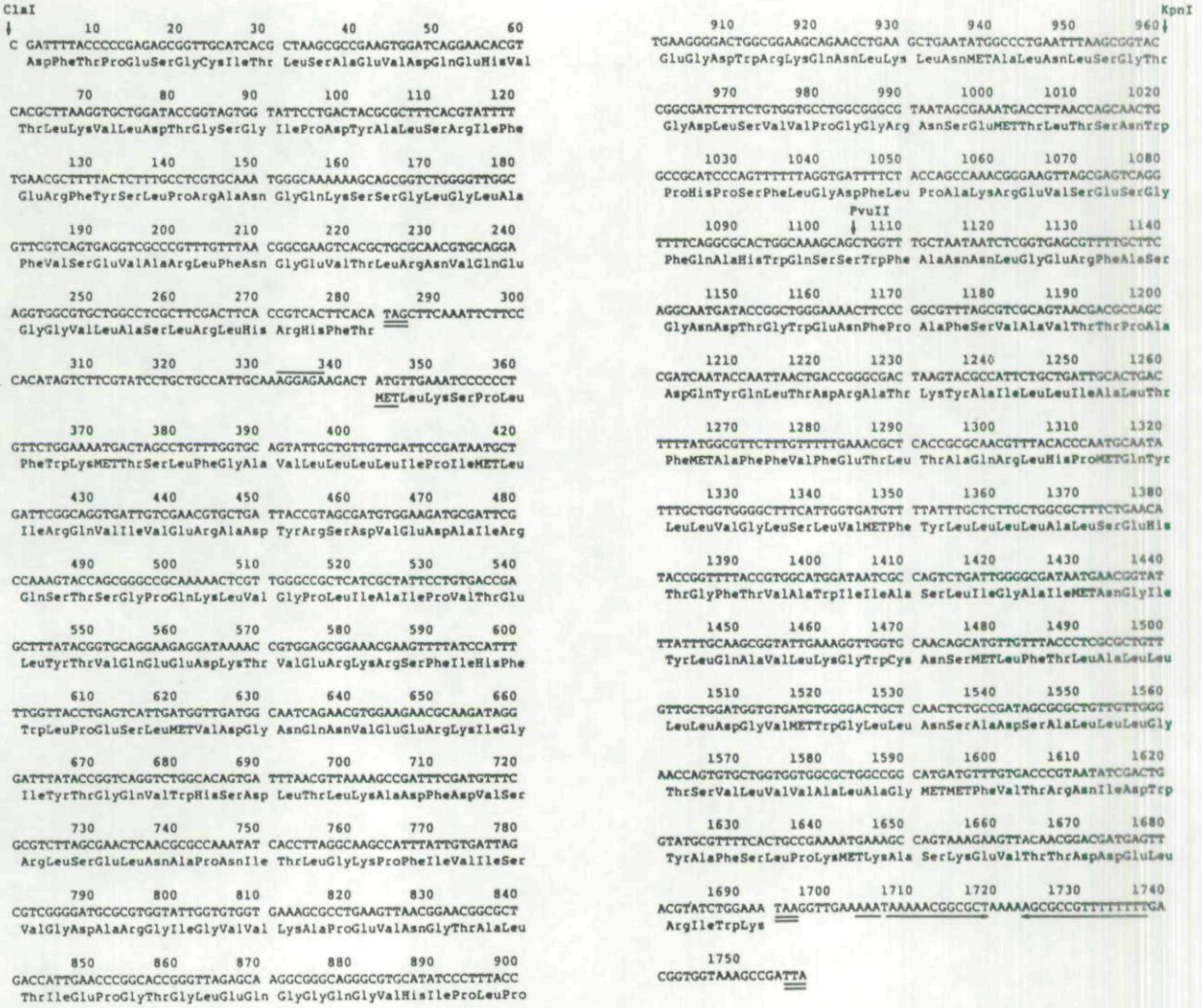


**Fig. 4.** Nucleotide and deduced amino acid sequence of the *E. coli cet* gene, together with the 3′ end of the *phoM* gene. The DNA sequence is numbered from the *ClaI* site. The potential ribosome binding site for *cet* is shown by overlining (‾‾‾), and the potential start amino acids are underlined. The stop codons for *phoM*, *cet* and *dye* are underlined (___), and the region of dyad symmetry representing the transcription termination site for *cet* and *dye* is shown (←). These sequence data have been submitted to the EMBL/GenBank Data Libraries under the accession number Y00538.

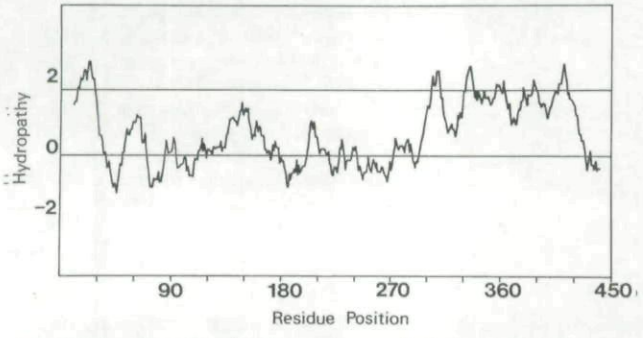


**Fig. 5.** Hydropathy plot for the Cet protein according to the method of Kyte and Doolittle (1982). The computer program continuously determines the average hydropathy of a 19-amino-acid segment as it advances through the protein from amino- to carboxy-terminus. The portions above the lower horizontal line (−0.4) indicate hydrophobic regions. The hydropathy of membrane-spanning domains generally averages greater than +1.6 (the upper horizontal line) over a 19-amino-acid segment (Kyte and Doolittle, 1982).

Latterell, 1980). We were inquisitive, therefore, about the effect *cet* had on alkaline phosphatase activity. We assayed the amount of alkaline phosphatase in a set of isogenic strains differing in their *cet* and *phoM* genes (Table 1). Deletion of the *serB-thr* region reduced alkaline phosphatase expression dramatically, whereas addition of pRB38 carrying $cet^+$ and $phoM^+$ restored expression, but only to 15% of the level in the $phoM^+$ $cet^+$ strain. However, introduction of the pRB62 ($phoM^+$ *cet2*) plasmid resulted in a five-fold increase of alkaline phosphatase activity over the $phoM^+$ $cet^+$ strain.

In order to determine whether this effect of the *cet2* mutation on alkaline phosphatase activity was an effect on *phoA* transcription, we introduced a series of plasmids into strain AS2*phoR*. This is a *phoR* strain lysogenized with a λ *phoA-lac* operon fusion, so that the *lacZ* gene is

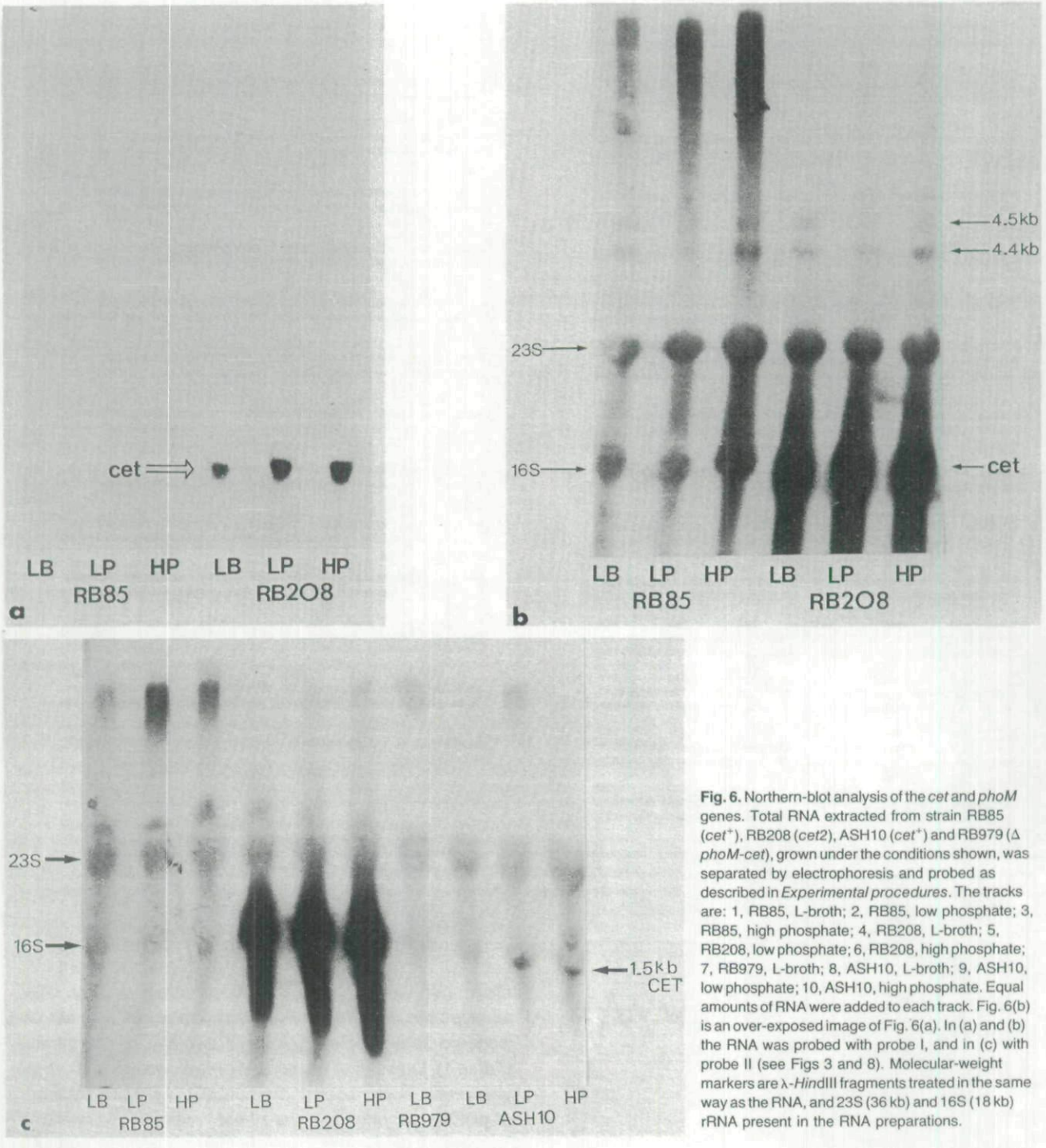


**Fig. 6.** Northern-blot analysis of the *cet* and *phoM* genes. Total RNA extracted from strain RB85 (*cet*$^+$), RB208 (*cet2*), ASH10 (*cet*$^+$) and RB979 (Δ *phoM-cet*), grown under the conditions shown, was separated by electrophoresis and probed as described in *Experimental procedures*. The tracks are: 1, RB85, L-broth; 2, RB85, low phosphate; 3, RB85, high phosphate; 4, RB208, L-broth; 5, RB208, low phosphate; 6, RB208, high phosphate; 7, RB979, L-broth; 8, ASH10, L-broth; 9, ASH10, low phosphate; 10, ASH10, high phosphate. Equal amounts of RNA were added to each track. Fig. 6(b) is an over-exposed image of Fig. 6(a). In (a) and (b) the RNA was probed with probe I, and in (c) with probe II (see Figs 3 and 8). Molecular-weight markers are λ-*Hin*dIII fragments treated in the same way as the RNA, and 23S (36 kb) and 16S (18 kb) rRNA present in the RNA preparations.

under *phoA* promoter control. We then assayed both alkaline phosphatase and β-galactosidase on sonic extracts. From the results presented in Table 1, again it can be seen that the introduction of the pRB62 plasmid carrying *cet2* resulted in a three-fold increase in both alkaline phosphatase and β-galactosidase. It should also be noted that the presence of the pACYC184 plasmid had the effect of actually reducing the specific activities of both enzymes, for some unknown reason, so the enhancing effect of the *cet2* mutation actually may be higher.

Direct Northern-blot analysis of *phoA* expression was carried out using, as probe, the 1.0 kb *Eco*RI-*Hin*dIII fragment from plasmid pJP91 (Tommassen *et al.*, 1984), which carries part of *phoA* (see Fig. 7). This was used to

**Table 1.** Effect of *cet* and *phoM* mutations on expression of *phoA*.

| Strain | Genotype | Alkaline phosphatase activity[a] | β-galactosidase activity |
|---|---|---|---|
| BW1308 | *phoR* | 187.4 | ND |
| RB2181 | *phoR*Δ(*serB-thr*) | 2.6 | ND |
| RB2181/pRB38 | *phoR*Δ(*serB-thr*)/*phoM*$^+$ cet$^+$ | 29.3 | ND |
| RB2181/pRB62 | *phoR*Δ(*serB-thr*)/*phoM*$^+$ cet2 | 1015.3 | ND |
| AS2*phoR* | *phoR*ϕ(*phoA-lac*) | 21.3 | 295.5 |
| AS2*phoR*/pACYC184 | *phoR*ϕ(*phoA-lac*)/pACYC184 | 4.8 | 63.6 |
| AS2*phoR*/pRB38 | *phoR*ϕ(*phoA-lac*)/*phoM*$^+$ cet$^+$ | 13.8 | 187.1 |
| AS2*phoR*/pRB62 | *phoR*ϕ(*phoA-lac*)/*phoM*$^+$ cet2 | 65.8 | 863.0 |

**a.** Bacteria were grown in L broth to mid-exponential phase, and cell extracts made as described in *Experimental procedures*. Activities are, for alkaline phosphatase, nmole *p* nitrophenol min$^{-1}$ mg$^{-1}$ protein, and for β-galactosidase, nmole *o* nitrophenol min$^{-1}$ mg$^{-1}$ protein. ND = not determined.

probe RNA extracted from strains RB85 (*cet*$^+$) and RB208 (*cet2*) grown in either high- or low-phosphate medium. It can be seen that in low-phosphate medium there were two bands of approximately 1.9 kb and 2.2 kb. We think these are *phoA*-specific transcripts since they were absent in cells grown in high-phosphate media, although in RB85 grown in such media there was a small amount of a larger (approximately 2.5 kb) transcript. The 1.9 kb and 2.2 kb transcripts were certainly increased in RB208 relative to RB85. Thus, enhanced amounts of the Cet protein appear to result in increased transcription of *phoA*.

## Discussion

*cet* mutants were originally isolated as a class of mutants tolerant to colicin E2, whilst adsorption of E2 was unimpaired relative to wild-type cells (Hill and Holland, 1967). These *cet* mutants differed from other E2-tolerant mutants, e.g. *tolA* and *tolB*, in retaining complete sensitivity to the closely related colicin E3, which has a different target, viz. 16S rRNA rather than DNA; thus they were mutants blocked at a late stage in the process of transmission of the colicin to its target.

*cet* mutants had increased amounts of an $M_r$ 43 000 inner membrane protein (Samson and Holland, 1970). In the present work, we have cloned the *cet* locus and shown that it codes for this inner membrane protein. Presumably this increased amount of Cet protein actually blocks colicin E2 from reaching its intracellular target, although we still have no idea how this occurs.

Our sequencing of the *cet* gene has revealed the presence of six possible transmembrane sequences typical of a membrane protein: five towards the carboxy-terminus and one at the amino terminus, thus suggesting that the ends of the protein are anchored in the membrane. Amemura *et al.* (1986), in their recent report of the sequence of the *phoM* gene, also report the sequence of an ORF downstream from *phoM* which is in fact the *cet* gene, although these authors had no information regarding its function or correlation with *cet*. Their sequence agrees with ours except that the G at bp 406 is actually an A in their strain. This could be due to the use of different strains which have undergone evolutionary divergence. This difference does not affect the amino acid sequence of Cet as it is in the wobble base position of a leucine residue.

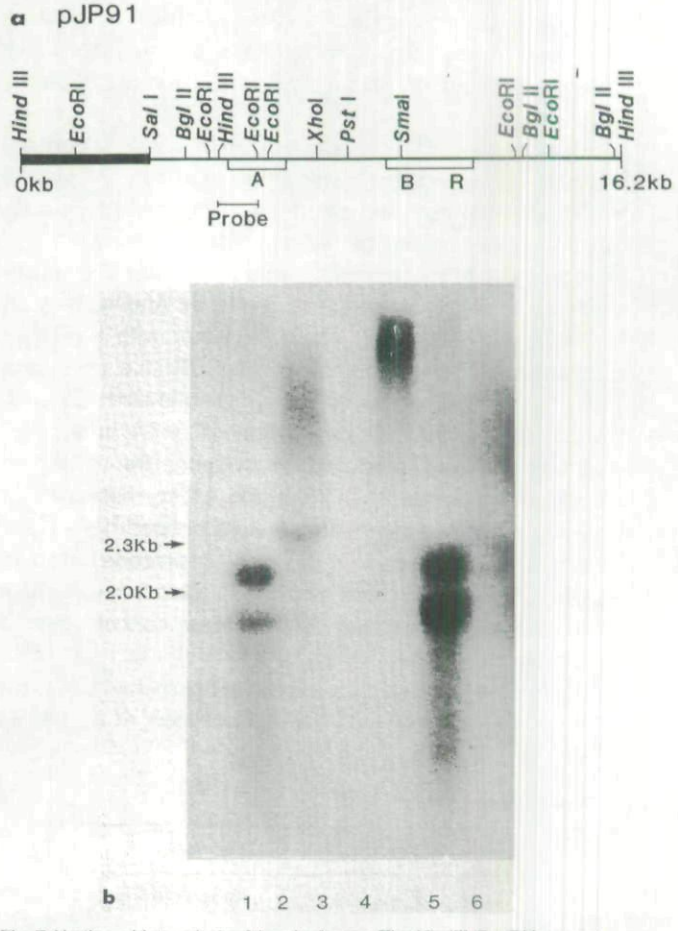


**Fig. 7.** Northern-blot analysis of the *phoA* gene. The *Hin*dIII-*Eco*RI fragment from plasmid pJP91 (a) (Tommassen *et al.*, 1984) was used to probe total RNA extracted from strain RB85 (*cet*$^+$) and RB208 (*cet2*) grown in either high- or low-phosphate medium. (b) The tracks are: 1, RB85, L-broth; 2, RB85, low phosphate; 3, RB85, high phosphate; 4, RB208, L-broth; 5, RB208, low phosphate; 6, RB208, L-broth. Molecular-weight markers are λ-*Hin*dIII fragments.

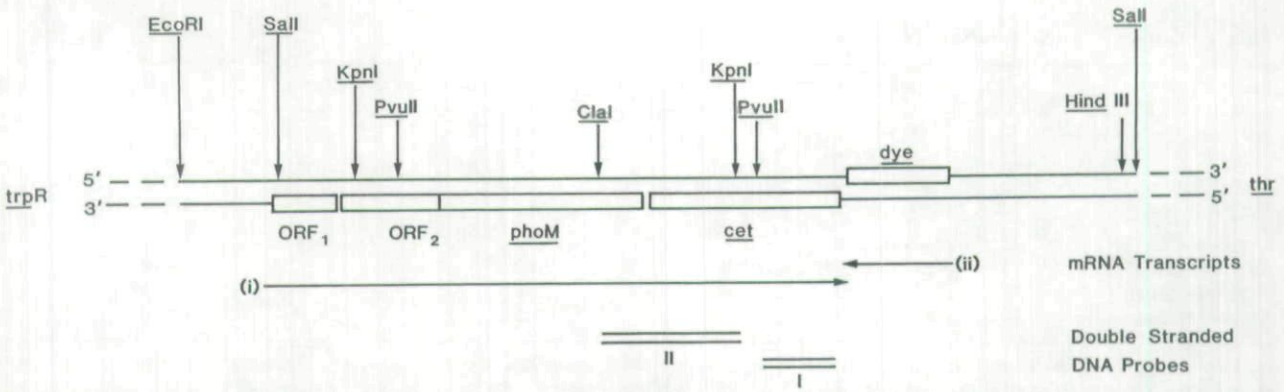


**Fig. 8.** Restriction map of the *phoM-cet* operon and the *dye* gene, and their orientation on the chromosome of *E. coli* K-12.
This figure is based on the results presented in the present paper and the results of Amemura *et al.* (1986). The mRNA transcripts of the $ORF_1$-$ORF_2$-*phoM-cet* operon (i), and the shorter *cet*-specific transcript in the *cet2* mutant (ii), are shown, together with the transcript from the *dye* gene (iii), based on the Northern-blot analysis presented in Fig. 6, and unpublished data (for *dye*). The extent of the DNA probes used for the Northern analysis shown in Fig. 6 are indicated.

Mutational activation of a weak promoter for *cet* could probably be the reason for the increased *cet*-specific mRNA production which we observe in Northern blots. This type of regulatory mutation, leading to increased production of Cet protein and blocking of colicin E2 entry to the cell, would explain why these *cet* mutations are dominant to $cet^+$ in partial diploids (Buxton and Holland, 1973).

Besides the 1.5 kb *cet*-specific mRNA, we also observed two bands in Northern blots at approximately 4.4 kb and 4.5 kb in both $cet^+$ and *cet* strains. This could be a polycistronic message encompassing both *phoM* and *cet*, and the size would be compatible with the results of DNA sequence analysis upstream of *phoM* by Amemura *et al.* (1986) who found that *phoM*, together with two other upstream ORFs and one downstream ORF (i.e. *cet*), probably formed an operon of about 4.0 kb in size. Although *phoM* is required for expression of *phoA* in a *phoR* background, we could find no evidence from Northern analysis that either the 4.4 kb and 4.5 kb messages, or the 1.5 kb *cet*-specific message, were induced in low-phosphate media. Amemura *et al.* (1986) have constructed an $ORF_1$-$ORF_2$-*phoM'-'lacZ* fusion and could find no regulation of *phoM* expression by phosphate concentration in the media.

Increased expression of *cet* in a *cet* mutant did, however, result in increased *phoA* transcription, as evidenced using *phoA-lac* fusions and Northern blot analysis. Nevertheless, unlike *phoM*, *cet* does not actually seem to be required for *phoA* expression, since Ludkte *et al.* (1984), who constructed various deletion plasmids in this region, found that the gene distal to *phoM* was not required for alkaline phosphatase synthesis in a $phoM^+$ *phoR* strain. As pointed out by Amemura *et al.* (1986), since the expression of this operon is not regulated by phosphate, the primary function of the operon may not be the regulation of the phosphate regulon. It could be concerned with some other physiological condition, since Wanner and McSharry (1982) showed that expression of several unknown phosphate-regulated promoters was also induced by carbon or nitrogen starvation. The effect that increased production of Cet protein has on the expression of genes of the *pho* regulon could be a direct effect on transcription; an analogy could be with ToxR, which also is a membrane protein, and which has been shown to be a transcriptional activator for cholera toxin (Miller *et al.*, 1987). On the other hand, it could be, for example, that increased amounts of the inner membrane Cet protein result in increased uptake or loss of some compound which affects *phoA* transcription, or it could be affecting the stability of the mRNA.

The *cet* gene ends at a stem-and-loop structure typical of a rho-independent transcription termination site. This appears to be shared with the *dye* gene, which is transcribed in the opposite direction to *cet* on the other DNA

**Table 2.** Strains of *E. coli* K-12.

| Strain | Mating type | Genotype | Origin/Reference |
|---|---|---|---|
| RB85 | $F^-$ | *thr leu thi lacY rpsL tonA supE44* ($\lambda^-$) | Buxton and Holland (1973) |
| RB979 | $F^-$ | RB85, Δ(*deoD-serB-trpR-phoM-cet-dye-thr*) | Buxton and Drury (1984) |
| ASH10 | $F^-$ | *leu metB thyA lacZ rpsL* ($\lambda^+$) *supE* | Buxton and Holland (1975) |
| RB208 | $F^-$ | *metB thyA lacZ rpsL* ($\lambda^+$) *supE ara cet2* | RB30×ASH10; Buxton (1973) |
| BW1308 | $F^-$ | *lac-169 proC*::Tn*5 phoR68 rpsL thi zaa*::Tn*10 crp200* | B.L. Wanner |
| BW3218 | $F^-$ | As BW1308, Δ(*serB-trpR-phoM-cet-dye-thr*) | B.L. Wanner |
| CSR603 | $F^-$ | *thr-1 leuB6 proA2 phr-1 recA1 argE3 thi-1 uvrA6 lacY1 rpsL31 supE44* | *A. Sancar via B.J. Bachmann* |
| *AS2phoR* | $F^-$ | *lac-169 araD139 rpsL* relA *thi phoR*/λφ (*phoA-lac*) | P.J. Bassford, Jr. |

strand (Drury and Buxton, 1985). Mutation of *dye* results in sensitivity to dyes, envelope protein changes and loss of expression of the sex factor F (Buxton and Drury, 1983; Buxton *et al.*, 1983). Similar bi-directional terminators have been described for *tonB*/P14 (Postle and Good, 1985), *his* (Carlomagno *et al.*, 1985) and *tetA*/*orfL* (Schollmeier *et al.*, 1985). There may be a functional reason for such a structure; obviously anti-termination at this site would result in anti-sense mRNA which could interfere with transcription of the other gene. We do not yet know whether there is in fact a functional relationship between *cet* and *dye*.

A map of the *phoM-cet-dye* region deduced from the results in the present paper and those of Drury and Buxton (1985) and Amemura *et al.* (1986) is presented in Fig. 8.

## Experimental procedures

### *Bacteria, phage and plasmids*

The bacterial strains, all of which are derivatives of *E. coli* K-12, are listed in Table 2. The plasmids used were: pRB38, which is pACYC184 carrying the 6 kb *Sal*I insert covering the *phoM-cet* region (Buxton and Drury, 1984); pRB62, which is pRB38 but carrying the *cet2* allele (see text); pRB50, which is pRB38 with Tn*1000* inserted at the 3′ end of the *dye* gene (*dye* 2197) (Drury and Buxton, 1985); and pJP91, which is a derivative of pACYC184 carrying the *phoA, B* and *R* genes, kindly donated by J. Tommassen (Tommassen *et al.*, 1984).

### *Tn*1000 *insertions, and the maxicell technique*

Insertions of the Tn*1000* transposon (γδ) into plasmids were performed as described previously (Buxton and Drury, 1983b) and proteins produced by plasmids were visualized using the maxicell technique (Buxton and Drury, 1983a; Buxton and Drury, 1984).

### *Inner and outer membrane analysis on SDS-PAGE*

Cells were grown in L-broth at 37°C to an $A_{600}$ of 0.5, and envelope proteins were prepared essentially as described by Boyd and Holland (1979). The separation of outer and inner membrane proteins was achieved using Sarkosyl NL97. Outer membranes, which remain insoluble after this procedure, were recovered by centrifugation at 100 000 g for 2 h. The outer membrane pellet was dissolved in 500 μl lysis buffer [0.05 M Tris-HCl, pH 6.8, 1% (w/v) SDS (sodium dodecyl sulphate), 0.02 M EDTA, 1% (w/v) β-mercaptoethanol, 10% (w/v) glycerol, 0.05% (w/v) bromophenol blue, and the supernatant which contains the inner membrane proteins was mixed with an equal volume of double-strength lysis buffer. After heating (100°C for 5 min), the samples were separated on SDS-PAGE as described by Laemmli (1970). Samples of $^{35}$S-methionine-labelled proteins from a maxicell experiment were loaded on the same gel for the identification of pRB38-coded proteins.

### *DNA handling techniques*

These have been described previously (Buxton and Drury, 1983b). Enzymes were purchased from Anglian Biotechnology Ltd., NBL Enzymes Ltd., or Boehringer Corporation Ltd.

### *DNA sequencing*

We have previously sequenced the *dye* gene, which is adjacent to *cet* (Drury and Buxton, 1985). We sequenced and identified an open-reading frame, now known to be part of *cet*, downstream from *dye* but translated towards *dye* on the opposite DNA strand. This sequence ended at a *Pvu*II site; to join this to the next DNA fragment, a *Kpn*I-*Cla*I piece, we cloned (into phage M13mp19) the *Kpn*I-*Sst*I fragment from pRB50, which has Tn*1000* inserted into the 3′ end of *dye* (see Fig. 3a) thus providing a convenient *Sst*I site. This was used to read across the *Pvu*II site. The *Cla*I-*Kpn*I fragment was cloned from pRB38 into the M13 vectors, mp10 and mp11, together with the *Taq*I clones shown in Fig. 3.

Cloning in M13 was as described previously (Drury and Buxton, 1985). We also used M13 mp18 and mp19, but had some trouble with these vectors since fragments of unknown molecular weights seemed to be present in phage DNA preparations.

Single-stranded M13 DNA was prepared (Schreier and Cortese, 1979) and used as template in the dideoxy chain-terminating sequencing method (Sanger *et al.*, 1977; 1980) with a 17-nucleotide synthetic primer (Pharmacia Ltd.), and deoxyadenosine 5′-α[$^{35}$S]thiotriphosphate (400 Ci mmol$^{-1}$) (Amersham International) as label. All the sequence was read in both directions except for the first 438 bp, since we could not isolate the *Cla*I-*Taq*I clone despite repeated attempts and the 262–438 bp *Taq*I fragment was found to be difficult to sequence when reading towards the 262 bp end. However, our sequence does agree with that of Amemura *et al.*, (1986), with the exception noted at 406 bp (see text).

Homology searches were performed using the Dayhoff data base and a Dec 20/60 computer using MGS software (Greer *et al.*, 1985).

### Extraction of RNA

Bacteria (20 ml) in exponential phase were harvested and washed once in 2 ml of ice-cold buffer containing 20 mM sodium acetate, pH 5.2, 1 mM EDTA and finally re-suspended in 200 μl of 10× buffer. SDS was added to a concentration of 1% (w/v) and the suspension was extracted with 400 μl of phenol (equilibrated with sodium acetate buffer) at 65°C for 10 min. Chloroform (200 μl) was added and mixed by inversion and the phases separated by centrifugation. The aqueous phase was extracted again with phenol/chloroform as above. Nucleic acids were then precipitated from the aqueous phase at −70°C with sodium acetate (0.3 M final concentration) and 2.5 volumes of ethanol. After centrifugation, the nucleic acid was resuspended in 100 μl of sterile distilled water and stored at −70°C.

### *Electrophoretic separation of RNA and Northern blotting*

Separation of RNA was performed on glyoxal or formaldehyde gels as described in Maniatis *et al.* (1982). To facilitate transfer of larger RNA species, gels were treated with 50 mM NaOH and

then neutralized before blotting. RNA was transferred to Hybond-N nylon membrane (Amersham International) and hybridization was carried out according to the Hybond manual.

### *DNA probes for Northern blotting*

Double-stranded DNA probes were restriction fragments separated in low-melting-temperature agarose (Sigma Chemical Co.) and were oligo-labelled using deoxycytidine 5′-[$\alpha^{32}$P]triphosphate (3000 Ci mmol$^{-1}$; Amersham) and hexa-nucleotide primers (Pharmacia) (Feinberg and Vogelstein, 1983).

### *Alkaline phosphatase assays*

These were performed essentially as described by Torriani (1966). Cell extracts were prepared from exponential cell cultures grown in low-phosphate medium or L-broth. Sonication was used to disrupt the cell envelope and, after centrifugation to remove cell debris, the supernatant was stored in aliquots at −20°C.

Phosphatase substrate was prepared (1 mg ml$^{-1}$) in 1.2 M Tris-HCl, pH 8.2 and allowed to equilibrate at 30°C. Equal volumes (0.3 ml) of cell extract and phosphatase substrate were incubated at 30°C for 15 min. The reaction was stopped with ½ volume of 2 N NaOH, the $A_{410}$ of the samples read, and the specific activity of the enzyme calculated. A molar extinction coefficient of *p*-nitrophenol of $1.62\times10^4$ was used (Wanner and McSharry, 1982).

## Acknowledgements

We thank P. J. Bassford, Jr., J. Tommassen, A. Sancar and B. L. Wanner for kindly donating strains and plasmids, N. M. Green for advice on identification of membrane-spanning domains, and P. Gillett for help with computing.

## References

Amemura, M., Makino, K., Shinagawa, H., and Nakata, A. (1986) Nucleotide sequence of the *phoM* region of *Escherichia coli*: four open reading frames may constitute an operon. *J Bacteriol* **168:** 294–302.

Boyd, A., and Holland, I.B. (1979) Regulation of the synthesis of surface protein in the cell cycle of *Ecoli* B/r. *Cell* **18:** 287–296.

Buxton, R.S. (1973) A genetic analysis of colicin E2 sensitivity in Escherichia coli K-12. PhD Thesis, University of Leicester.

Buxton, R.S., and Drury, L.S. (1983a) The transposon γδ (Tn*1000*) codes for a polypeptide with an apparent $M_r$ of 72 000, which is visible in maxi-cells. *FEMS Microbiol Lett* **17:** 287–290.

Buxton, R.S., and Drury, L.S. (1983b) Cloning and insertional inactivation of the *dye* (*sfrA*) gene, mutation of which affects sex factor F expression and dye sensitivity of *Escherichia coli* K-12. *J Bacteriol* **154:** 1309–1314.

Buxton, R.S., and Drury, L.S. (1984) Identification of the *dye* gene product, mutational loss of which alters envelope protein composition and also affects sex factor F expression in *Escherichia coli* K-12. *Mol Gen Genet* **194:** 241–247.

Buxton, R.S., Drury, L.S., and Curtis, C.A.M. (1983) Dye sensitivity correlated with envelope protein changes in *dye* (*sfrA*) mutants of *Escherichia coli* K12 defective in the expression of the sex factor F. *J Gen Microbiol* **129:** 3363–3370.

Buxton, R.S., and Holland, I.B. (1973) Genetic studies of tolerance to colicin E2 in *Escherichia coli* K-12. I. Re-location and dominance relationships of *cet* mutations. *Mol Gen Genet* **127:** 69–88.

Buxton, R.S., and Holland, I.B. (1974) Genetic studies of tolerance to colicin E2 in *Escherichia coli* K-12. II. Multiple mutations as a cause of the various phenotypic properties of Cet$^-$ mutants. *Mol Gen Genet* **131:** 159–171.

Carlomagno, M.S., Riccio. A., and Bruni, C.B. (1985) Convergently functional, rho-independent terminator in *Salmonella typhimurium. J Bacteriol* **163:** 362–368.

Drury, L.S., and Buxton, R.S. (1985) DNA sequence analysis of the *dye* gene of *Escherichia coli* reveals amino acid homology between the Dye and OmpR proteins. *J Biol Chem* **260:** 4236–4242.

Feinberg, A.P., and Vogelstein, B. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* **132:** 6–13.

Greer, W., Gillett, P., Bygrave, A., Weaver, N., and Khepar, J. (1985) *MGS Reference Manual,* Version 3D, Computing Laboratory, National Institute for Medical Research, Mill Hill, London.

Hill, C., and Holland, I.B. (1967) Genetic basis of colicin E susceptibility in *Escherichia coli*. I. Isolation and properties of refractory mutants and the preliminary mapping of their mutations. *J Bacteriol* **94:** 677–686.

Holland, I.B., and Threlfall, E.J. (1969) Identification of closely linked loci controlling ultraviolet sensitivity and refractivity to colicin E2 in *Escherichia coli. J Bacteriol* **97:** 91–96.

Kozak, M. (1983) Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. *Microbiol Rev* **47:** 1–45.

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

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227:** 680–685.

Ludtke, D., Bernstein, J., Hamilton, C., and Torriani, A. (1984) Identification of the *phoM* gene product and its regulation in *Escherichia coli* K-12. *J Bacteriol* **159:** 19–25.

Makino, K., Shinagawa, H., and Nakata, A. (1984) Cloning and characterization of the alkaline phosphatase positive regulatory gene (*phoM*) of *Escherichia coli. Mol Gen Genet* **195:** 381–390.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (eds). (1982) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.

Miller, V.L., Taylor, R.L., and Mekalanos, J.J. (1987) Cholera toxin transcriptional activator ToxR is a transmembrane DNA binding protein. *Cell* **48:** 271–279.

Nomura, M. (1963) Mode of action of colicines. *Cold Spring Harbor Symp Quant Biol* **28:** 315–324.

Postle, K., and Good, R.F. (1985) A bidirectional rho-independent transcription terminator between the *E. coli tonB* gene and an opposing gene. *Cell* **41:** 577–585.

Samson, A.C.R., and Holland, I.B. (1970) Envelope protein changes in mutants of *Escherichia coli* refractory to colicin E2. *FEBS Lett* **11:** 33–36.

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

Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H., and Roe, B.A. (1980) Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J Mol Biol* **143:** 161–178.

Schaller, K., and Nomura, M. (1976) Colicin E2 is a DNA endonuclease. *Proc Natl Acad Sci USA* **73:** 3989–3993.

Schollmeier, K., Gartner, D., and Hillen, W. (1985) A bidirectionally active signal for termination of transcription is located between *tetA* and *orfL* on transposon Tn*10*. *Nucl Acids Res* **13:** 4227–4237.

Schreier, P.H., and Cortese, R. (1979) A fast and simple method for sequencing DNA cloned in the single-stranded bacteriophage M13. *J Mol Biol* **129:** 169–172.

Shine, J., and Dalgarno, L. (1975) Determinant of cistron specificity in bacterial ribosomes. *Nature* **254:** 34–38.

Stüber, D., and Bujard, H. (1981) Organisation of transcriptional signals in plasmids pBR322 and pACYC184. *Proc Natl Acad Sci USA* **78:** 167–171.

Tommassen, J., Hiemstra, P., Overduin, P., and Lugtenberg, B. (1984) Cloning of *phoM*, a gene involved in regulation of the synthesis of phosphate limitation inducible proteins in *Escherichia coli* K12. *Mol Gen Genet* **195:** 190–194.

Torriani, A. (1966) Alkaline phosphatase from *Escherichia coli*. In *Procedures in Nucleic Acid Research,* Vol. 1 Cantoni, C.L., and Davies, D.R. (eds). New York: Harper & Row, pp. 224–235.

Wanner, B.L., and Latterell, P. (1980) Mutants affected in alkaline phosphatase expression: evidence for multiple positive regulators of the phosphate regulon in *Escherichia coli*. *Genetics* **96:** 353–366.

Wanner, B.L., and McSharry, R. (1982) Phosphate-controlled gene expression in *Escherichia coli* K12 using *MudI*-directed *lacZ* fusions. *J Mol Biol* **158:** 347–363.