J. Mol. Biol. (1987) **198**, 633–641

Molecular Analysis of the Promoter Region of the *Escherichia coli* K-12 *phoE* Gene

Identification of an Element, Upstream from the Promoter, Required for Efficient Expression of PhoE Protein

Jan Tommassen†, Margot Koster and Piet Overduin

Department of Molecular Cell Biology and
Institute of Molecular Biology
State University of Utrecht
Transitorium 3, Padualaan 8
3584 CH Utrecht, The Netherlands

(Received 21 April 1987, and in revised form 5 August 1987)

The *phoE* gene of *Escherichia coli* codes for an outer membrane pore protein whose expression is induced under phosphate limitation. The promoter of this gene contains a 17 base-pair fragment, designated a *pho* box, which is present also in other phosphate-controlled promoters. The mRNA start site was determined and found to be located downstream from the *pho* box, such that this element is located in the −35 region of the *phoE* promoter.

A set of promoter deletions was generated *in vitro* and analysis of these deletions revealed that sequences upstream from the *pho* box are required for the efficient expression of *phoE*. The required upstream region is located (in part) between positions −106 and −121 relative to the mRNA start site, and contains sequences homologous to a *pho* box and a correctly spaced Pribnow box, but in the reversed orientation relative to the regular −35 and −10 regions. A proper spacing between this upstream region and the −35 region appears to be important, since an oligonucleotide insertion in the intervening region interferes with *phoE* expression. By cloning the upstream region in a *lacZ* operon fusion vector, a weak phosphate limitation-inducible promoter activity could be detected.

1. Introduction

Growth of *Escherichia coli* under phosphate limitation results in the induction of the synthesis of many proteins, which function in order to scavenge traces of phosphate-containing nutrients from the surrounding medium (for a review, see Tommassen & Lugtenberg, 1982). One of these proteins is PhoE protein, which forms pores in the outer membrane through which small hydrophilic solutes can pass. PhoE protein pores are more efficient in the uptake of phosphate and phosphate-containing solutes than the related but constitutively expressed OmpF and OmpC protein pores (Korteland *et al.*, 1982). The synthesis of PhoE protein is co-regulated with that of many other proteins in a single regulon designated the *pho* regulon (Argast & Boos, 1980; Tommassen & Lugtenberg, 1980). The structural genes for these proteins are scattered throughout the chromosome: *phoE* is located at 6 minutes (Tommassen & Lugtenberg, 1981); *phoA*, encoding the periplasmic enzyme alkaline phosphatase at 8 minutes (Nakata *et al.*, 1971); the *pst* operon, encoding a binding-protein dependent, high-affinity uptake system for phosphate at 83 minutes (Willsky *et al.*, 1973), and the *ugp* locus encoding a binding-protein dependent uptake system for glycerol 3-phosphate at 76 minutes (Schweizer *et al.*, 1982).

At least three genes, *phoB*, *phoR* and *phoM*, are implicated in the regulation of the *pho* regulon (Wanner & Latterell, 1980). According to a model for this regulatory circuit (Tommassen & Lugtenberg, 1982; Tommassen *et al.*, 1982*a*), the PhoB protein functions as an activator of transcription of the structural genes of the *pho* regulon, whereas *phoB* expression in its turn is modulated by the products of *phoR* and *phoM*. In addition, *phoB* expression appears to be auto-regulated (Guan *et al.*, 1983). However, the possibility cannot be excluded that the *phoR* and *phoM* gene products also directly modulate the expression of the *pho*

† Author to whom all correspondence should be sent.

0022-2836/87/240633-09 $03.00/0

Table 1
Bacterial strains

Strain	Characteristics	Reference, source†
AB1157	F⁻, *thr leu proA2* (del *proA-proB-phoE-gpt*) *his thi argE lacY galK xyl rpsL*	Adelberg *via* PC
CE1224	*ompR* derivative of AB1157	Tommassen *et al.*, 1983
K10	Hfr Cav, *relA1 tonA22 pit-10 spoT1 ompF627*	Garen *via* CGSC
C86	*phoS21* derivative of K10	Garen *via* CGSC
C9	*phoR18* derivative of K10	Garen *via* CGSC
CE1248	F⁻, *thi pyrF ilvA argG tonA rpsL deoC vtr glpR ompR471 phoR69 phoA8 Δ(proB-phoE) recA56*	Korteland *et al.*, 1985
CE1265	F⁻, *thi pyrF ilvA argG rpsL glpR ompR471 phoR18 recA56 Δ(proB-phoE)*	Korteland *et al.*, 1985
MC4100	F⁻, Δ *lacU169 araD139 rpsL thi relA*	Casadaban, 1976

† PC, Phabagen Collection, Department of Molecular Cell Biology, Section Microbiology, State University of Utrecht, Utrecht, The Netherlands; CGSC, *E. coli* Genetic Stock Center, Department of Human Genetics, Yale University School of Medicine, New Haven, CT (B. J. Bachmann, Curator).

structural genes (Wanner & Latterell, 1980; Guan *et al.*, 1983). In any case, the promoter regions of the structural genes are expected to contain common elements to which the regulatory protein(s) bind. Comparison of the nucleotide sequences of the promoter regions of *phoA* (Kikuchi *et al.*, 1981), *phoE* (Overbeeke *et al.*, 1983) and the *pst* operon (Magota *et al.*, 1984; Surin *et al.*, 1984) indeed revealed a common 17 bp† element, designated a *pho* box (see Fig. 1). Such a *pho* box was also detected in the *phoB* promoter (Makino *et al.*, 1986). In all cases, the *pho* box was followed at a distance of 11 bp downstream by a potential Pribnow box, suggesting that the *pho* box functions as an aberrant −35 region in *E. coli pho*-controlled promoters. Here, we demonstrate that this Pribnow box in the *phoE* gene is indeed functional, by mapping of the transcription start point. To determine whether the *pho* box is sufficient for phosphate-controlled expression of the *phoE* gene, a deletion analysis was carried out in the upstream region. An upstream element was detected that is required for the efficient expression of the *phoE* gene. This upstream element shows homology to the *pho* box, but is in the reversed orientation.

2. Materials and Methods

(a) *Bacterial strains and growth conditions*

All bacterial strains are derivatives of *E. coli* K-12. Their sources and relevant characteristics are listed in Table 1.

Cells were grown overnight at 37 °C in L-broth (Tommassen *et al.*, 1983) or in a synthetic medium in which the phosphate concentration can be varied (Tommassen & Lugtenberg, 1980). The latter medium was supplemented with 660 μM-K_2HPO_4 (high-phosphate medium), which is sufficient to repress the *pho* regulon. For derepression, an overnight culture in high-phosphate medium was diluted 1:12·5 in fresh medium without phosphate added and incubation at 37 °C was continued for 5 h (Tommassen *et al.*, 1985). For growth of strains containing plasmids, the medium was supplemented with chloramphenicol (25 μg/ml) or ampicillin (50 μg/ml).

(b) *DNA manipulations*

Plasmid DNA was isolated as described by Birnboim & Doly (1979), followed in large-scale purifications by CsCl/ethidium bromide isopycnic centrifugation. The conditions for restriction endonuclease reactions and for treatment of DNA with *Bal*31 exonuclease, bacteriophage T4 DNA ligase, bacteriophage T4 polynucleotide kinase and Klenow fragment of DNA polymerase were as described by the manufacturers of the enzymes. DNA fragments were analysed on 0·6% (w/v) agarose gels or 4·5% (w/v) polyacrylamide (acrylamide to bisacrylamide, 30:1, w/w) gels.

Oligonucleotides were synthesized on a Biosearch 8600 DNA synthesizer and purified on 20% polyacrylamide gels.

Nucleotide sequencing was performed by using the dideoxy chain termination method on linearized, double-stranded plasmid preparations, which were purified on CsCl/ethidium bromide gradients (Korneluk *et al.*, 1985). Two synthetic oligonucleotides, p26 and p32 (indicated in Fig. 2) were used as sequencing primers.

(c) *Plasmid constructions*

Plasmid pJP12 (Tommassen *et al.*, 1982*b*) consists of a 4·9 kb *Sal*I restriction fragment containing the *phoE*

E.coli	-35		PB
consensus promoter	TTGACA	- 17 bp -	TATAAT
phoE	CTGTAATATATCTTTAA	- 11 bp -	TAAAAA
phoA	CTGTCATAAAGTTGTCA	- 11 bp -	TATAGT
pst	CTTACATATAACTGTCA	- 11 bp -	TATTTT
phoB	TTTTCATAAATCTGTCA	- 11 bp -	CATAAT
pho box	CT$^{G}_{T}$TCATA$^{A}_{T}$A$^{A}_{T}$CTGTCA	- 11 bp -	TATAAT

Figure 1. Homologies in the nucleotide sequences of the regulatory regions of *phoE*, *phoA*, *pst* and *phoB* as described by Makino *et al.* (1986). The upper line shows the consensus sequences of the −35 and −10 (PB) regions of *E. coli* promoters (Rosenberg & Court, 1979). A consensus sequence for promoters of the *pho* regulon, consisting of a 17 bp long *pho* box and a putative Pribnow box 11 bp downstream, is shown in the lower line.

† Abbreviations used: bp, base-pair(s); kb, 10^3 base-pairs; SDS, sodium dodecyl sulphate.

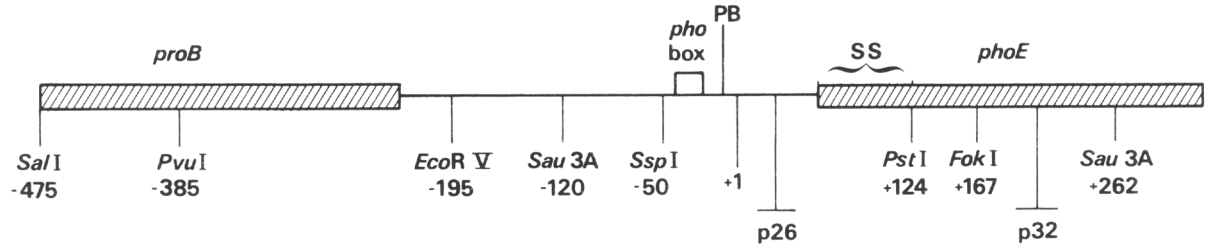

Figure 2. A representation of an 800 bp DNA fragment containing the promoter regions of the divergently transcribed *proB* and *phoE* genes. For the sake of clarity, only the restriction enzyme cleavage sites that were used in this study are depicted. The numbers underneath the cleavage sites refer to the distances (in bp) relative to the position of the transcription start point of *phoE* as determined in this work. *pho* box and PB indicate the positions of the *pho* box and putative Pribnow box from Fig. 1, respectively. SS corresponds to the DNA encoding the signal sequence of PhoE protein. p26 and p32 are synthetic oligonucleotides with the sequences dGCGCGAATATTCAGCGG and dCGCCATCTTTACTGGCG, respectively. They are complementary to the *phoE* mRNA.

gene, cloned in the vector pACYC184 (Chang & Cohen, 1978). A smaller derivative of this plasmid, pJP29, has been described (Bosch *et al.*, 1986). An 800 bp fragment of this plasmid, containing the promoter regions of the divergently transcribed *proB* and *phoE* genes is depicted in Fig. 2.

To construct pJP300, which contains a *Bam*HI linker insertion in the *Eco*RV site depicted in Fig. 2, 1 μg of pJP29 DNA was partially digested with 2 units of *Eco*RV by incubation for 5 min at 37 °C. The linear DNA was excised from an agarose gel and recovered by electroelution. The DNA was ligated with T4 DNA ligase to 100 ng of phosphorylated oligonucleotide dCGGATCCG (an 8-mer *Bam*HI linker). The excess of linkers was removed by digestion with *Bam*HI and electrophoresis on an agarose gel. After recovery of the linear plasmid DNA from the gel and re-ligation, the plasmid preparation was used to transform strain CE1224, and the plasmid content of chloramphenicol-resistant colonies was analysed by restriction enzyme digestions. Plasmid pJP300 is one of the plasmids obtained with the linker inserted at the intended *Eco*RV site.

A linker insertion in pJP29 in the *Ssp*I site indicated in Fig. 2 was obtained in a similar way, except that partial digestion with *Ssp*I was achieved by including 26 μg ethidium bromide/ml during the digestion and that a 12-mer *Bam*HI linker (dCCCGGATCCGGG) was inserted. A plasmid with the linker insertion at the intended *Ssp*I site was designated as pJP320.

Plasmid pMLB1010 (Travers *et al.*, 1983) is an operon fusion vector. It contains a complete *lacZ* coding sequence, including an ATG initiation codon and a ribosome binding site. A multiple cloning site is located upstream from this gene for insertion of foreign promoters. To construct pJP350 and pJP351, 1 μg of pJP29 was digested with *Sau*3A and restriction fragments were separated by polyacrylamide gel electrophoresis. The 382 bp *Sau*3A restriction fragment extending from position −120 to +262 (Fig. 2) was excised, purified as described by Maniatis *et al.* (1982) and ligated into *Bam*HI-digested pMLB1010. After re-digestion with *Bam*HI, the DNA preparation was used to transform strain MC4100, selecting for ampicillin-resistant colonies. Plasmid pJP350 contains the *phoE* promoter in its correct orientation in front of *lacZ*, pJP351 in the reversed orientation.

(d) S_1 *nuclease mapping*

S_1 nuclease mappings were performed as described by Berk & Sharp (1977). A DNA probe was obtained by digestion of pJP29 with *Sal*I and *Fok*I, and isolation of the *Sal*I–*Fok*I fragment indicated in Fig. 2 from a 5% preparative polyacrylamide gel. The fragment was 5′ end-labelled with T4 polynucleotide kinase and [γ-^{32}P]ATP and digested with *Pvu*I. The *Pvu*I–*Fok*I fragment, 5′ end-labelled at the *Fok*I site was purified from a polyacrylamide gel. RNA was isolated from cells as described by van den Elzen *et al.* (1980) and hybridized overnight with the DNA probe in the presence of 80% (v/v) formamide at 45 °C. Non-hybridized DNA was digested with S_1 nuclease for 30 min at 20 °C, and protected fragments were analysed by electrophoresis on a sequencing gel.

(e) *Primer extension*

As a primer, a synthetic oligonucleotide was used, indicated as p32 in Fig. 2. This oligonucleotide was 5′ end-labelled with T4 polynucleotide kinase and [γ-^{32}P]ATP. The primer was hybridized by mixing 50 ng of it with 10 μg of RNA in 100 mM-NaCl in a total volume of 10 μl. The samples were heated for 3 min at 70 °C, annealed for 10 min at 48 °C, and primer extension was carried out with 5 units of AMV reverse transcriptase as described (Giorgi *et al.*, 1983).

(f) *Cell envelope isolations*

Cell envelopes were isolated by differential centrifugation after ultrasonic disintegration of the cells (Lugtenberg *et al.*, 1975). The protein patterns of the cell envelopes were analysed by SDS/polyacrylamide gel electrophoresis.

(g) *β-Galactosidase assay*

For quantitative assays of β-galactosidase activity, cells were permeabilized as described by Villarejo *et al.* (1983) and β-galactosidase activity was measured as described by Miller (1972), using *O*-nitrophenyl-β-D-galactoside as substrate.

3. Results

(a) *Mapping of the phoE transcription start point*

Comparison of the promoter regions of the structural genes of the *pho* regulon revealed a 17 bp long homologous sequence (Fig. 1), which might be involved in the interaction with regulatory proteins. In all cases, this *pho* box was followed at a distance of 11 bp downstream by a putative

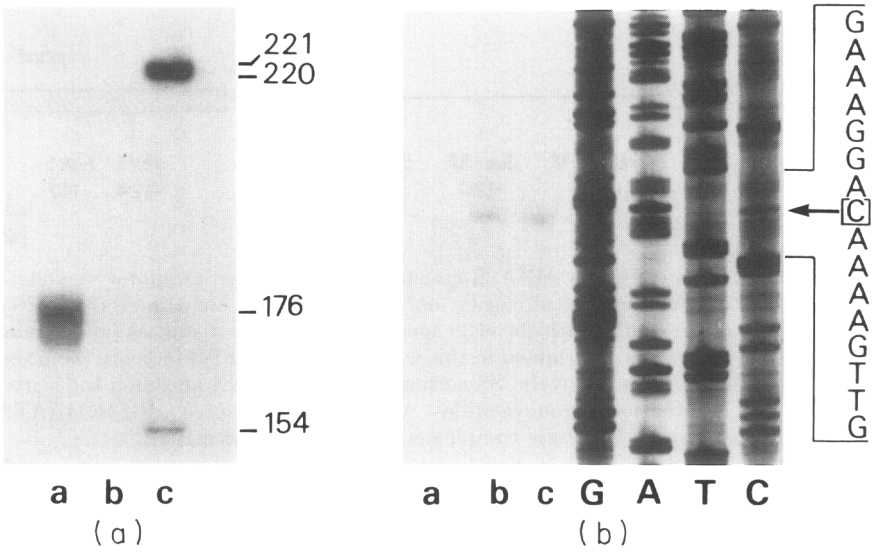

Figure 3. Determination of the *phoE* transcription start point by S_1 nuclease mapping and primer extension. (a) S_1 nuclease experiments. RNA was extracted from strain AB1157 containing pJP12 after growth under phosphate limitation (lane a) or in high-phosphate medium (lane b) and hybridized with a radioactively labelled DNA probe. After treatment with S_1 nuclease, protected DNA fragments were analysed on a 5% polyacrylamide sequencing gel next to molecular weight standard fragments (lane c), the length of which is indicated (in bp) at the right. (b) Primer extension. Oligonucleotide p32 (Fig. 2) was hybridized with mRNA from strain AB1157 containing pJP12, after growth of the strain in high-phosphate medium (lane a) or under phosphate limitation (lanes b and c) and the primer was extended with AMV reverse transcriptase and dNTPs. The DNA fragments were analysed on a 5% polyacrylamide sequencing gel in parallel with sequencing ladders obtained with the same primer and using the dideoxy chain termination method (Sanger *et al.*, 1977). Single-stranded DNA from an M13mp8 clone with the *Sal*I–*Bgl*II fragment of pJP29 inserted, was used as a template for the sequencing reactions. The transcription start point as deduced from this method is indicated by an arrow on the right.

Pribnow box. However, in the case of *phoE*, the highly conserved sixth base of the TATAAT hexamer is absent (Fig. 1). To investigate whether this putative Pribnow box is actually functional, the mRNA start site was determined by S_1 nuclease mapping. As a probe in these experiments, the *Pvu*I–*Fok*I fragment indicated in Figure 2, 5′ end-labelled at the *Fok*I site was used. mRNA was isolated from strain AB1157 containing *phoE*$^+$ plasmid pJP12. After hybridization of the RNA with the labelled DNA probe and digestion with S_1 nuclease, no bands were detectable in RNA isolated from cells grown with sufficient phosphate (Fig. 3(a), lane b). In contrast, eight to ten bands were found when mRNA extracted from phosphate-depleted cells was used (lane a). Similar results were obtained when mRNA was extracted from strain K10, containing a chromosomal *phoE* gene (data not shown). The same bands were observed with mRNA isolated from cells of strains C86 and C9 after growth in medium containing sufficient phosphate (data not shown). These strains produce PhoE protein constitutively due to *phoS* and *phoR* mutations, respectively. The bases corresponding to the bands observed in the S_1 mapping experiments are indicated in Figure 4. Although it is unlikely that all these bases serve as transcription start points, it is clear from these experiments that the actual transcription start point is located very close to the putative Pribnow box, which was deduced from the homology between *pho*-controlled promoters.

To localize the transcription start point more accurately, the primer extension method was used. A 17-mer synthetic oligonucleotide corresponding to a fragment of the *phoE* gene (indicated as p32 in Fig. 2) was 5′ end-labelled and annealed to mRNA extracted from strain AB1157 containing pJP12. The primer was extended with reverse transcriptase and dNTPs and applied on a gel next to an M13 dideoxy sequencing reaction in which the same primer was used. The autoradiograph (Fig. 3(b)) revealed a single band, which was present only when mRNA extracted from phosphate-depleted cells was used (lanes b and c). The corresponding startpoint is indicated in Figure 4. This guanine residue is numbered +1 throughout this work, as the mRNA start site of the *phoE* gene. The transcription start point is located 8 bp downstream from the putative Pribnow box, which was indicated on the basis of the homology with other *pho* promoters, showing that this −10 region is indeed functional. The +1 base is located 58 bp upstream from the ATG initiation codon.

(b) *Deletion analysis of the phoE promoter region*

To determine whether the presence of a *pho* box is sufficient for phosphate-controlled expression of the *phoE* gene, a deletion analysis was performed in the upstream region. The *phoE*$^+$ plasmid pJP29 (Bosch *et al.*, 1986) contains two *Eco*RV recognition sites, one of which is located at position −195

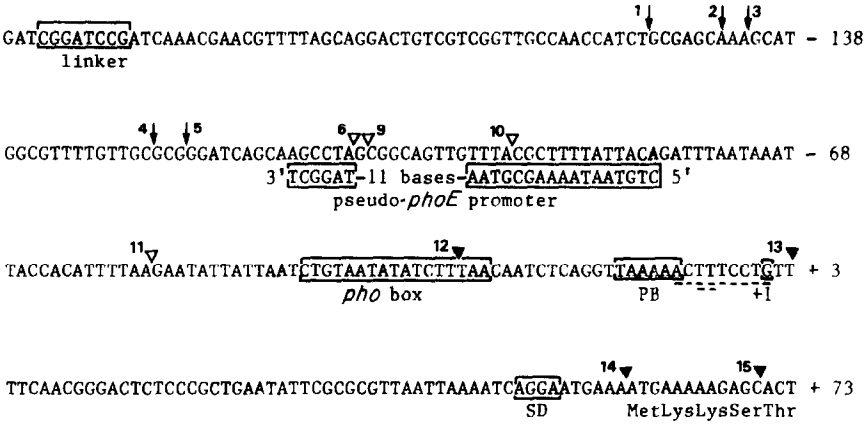

Figure 4. Nucleotide sequence of a DNA fragment of pJP300 encompassing the *phoE* promoter region. Boxed are the *Bam*HI linker inserted in the *Eco*RV site of pJP29 to yield pJP300, and the *pho* box, the Pribnow box (PB) and the ribosome-binding site (SD) preceding the *phoE* coding sequence. The putative mRNA start points deduced from the S_1 nuclease mapping experiments (Fig. 3(a)) are underlined, the strongest signals in these experiments being doubly underlined. The start point as deduced from the primer extension experiments is boxed and indicated as +1. Other nucleotides at the right of each line are numbered accordingly. Deletions were generated in the promoter region by digestion of pJP300 at the position of the linker with *Bam*HI, followed by limited digestion with *Bal*31 exonuclease. The end-points of the deletions as deduced from nucleotide sequencing are indicated by arrows and triangles above the sequence, and the numbers 1 to 15 correspond to plasmids pJP301 to pJP315, respectively. The end-points of the deletions in plasmids pJP307 and pJP308 are identical to that indicated for pJP306. Arrows indicate the end-points of deletions that do not affect *phoE* expression, open and filled triangles correspond to deletions that reduce or completely abolish *phoE* expression, respectively. The position of the pseudo-*phoE* promoter with homology to a consensus *pho* box and a correctly spaced Pribnow box (Fig. 1), but oriented opposite to the regular *pho* box and Pribnow box is indicated by showing the 3′ to 5′ sequence.

upstream from the *phoE* transcription start site (Fig. 2). This *Eco*RV site was converted into a unique *Bam*HI site by partial digestion of pJP29 with *Eco*RV and ligation of an 8-mer *Bam*HI linker into the linearized plasmid as described in Materials and Methods. The resulting plasmid is designated pJP300. The linker insertion does not influence *phoE* expression (data not shown).

To create deletions in the *phoE* promoter region, pJP300 was linearized with *Bam*HI and subsequently subjected to limited digestion with exonuclease *Bal*31. After ligation with T4 DNA ligase and redigestion with *Bam*HI, the DNA was used to transform *phoE* mutant strain CE1224, selecting for chloramphenicol-resistant colonies. Fifteen deletion mutants were chosen for detailed analysis and the corresponding plasmids were designated pJP301 to pJP315. Initial characterization of these plasmids showed that the *Bam*HI site was removed in all cases, whereas the *Pst*I site indicated in Figure 2 was retained.

To study the effect of the deletions on expression of PhoE protein, cell envelopes were isolated from derivatives of strain CE1224 containing pJP300 or either one of the mutant plasmids, after growth under high and low-phosphate conditions, respectively. The mutations could be divided into three classes: (1) the mutations on plasmids pJP301 to pJP305 did not affect the expression of PhoE protein as compared to pJP300; (2) expression of PhoE protein was drastically reduced in the case of plasmids pJP306 to pJP311, but still under phosphate control; and (3) in the case of pJP312 to pJP315, expression of the PhoE protein was completely abolished. Representative examples of each of these classes are shown in Fig. 5. These three different levels of expression were also observed upon transformation of *phoR* mutant strain CE1248 with the mutant plasmids (Fig. 5, lanes i to l).

To determine the extent of the deletions, nucleotide sequence analysis was performed using the dideoxy chain terminating method with linearized plasmid DNA as template and the oligonu-

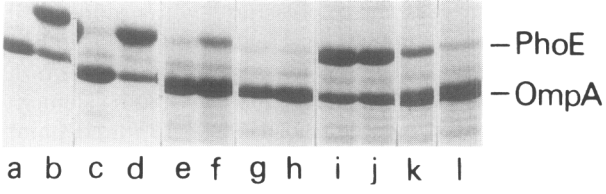

Figure 5. SDS/polyacrylamide gel electrophoresis patterns of cell envelopes of strains carrying plasmids with a wild-type *phoE* gene or with deletions in the promoter region. Cell envelope protein patterns are shown of derivatives of strain CE1224 carrying plasmids; pJP300, with a wild-type *phoE* gene (lanes a and b); pJP305, representative for a class 1 deletion (lanes c and d); pJP306, representative for a class 2 deletion (lanes e and f); and pJP312, representative for a class 3 deletion (lanes g and h). The strains were grown on high-phosphate medium (lanes a, c, e and g) or under phosphate limitation (lanes b, d, f and h). Also, the expression of the different *phoE* alleles in the constitutive *phoR* mutant strain CE1248 is shown (lanes i to l): pJP300 (lane i), pJP305 (lane j), pJP306 (lane k) and pJP312 (lane l). Only the relevant part of the gel is shown. The positions of PhoE protein and OmpA protein are indicated.

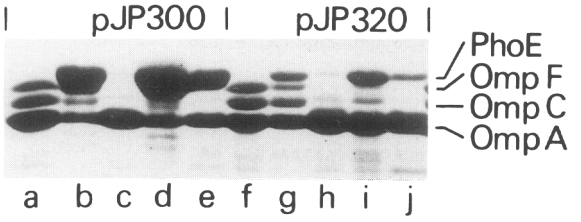

Figure 6. Effect of an insertion in the promoter region on the expression of PhoE protein. SDS/polyacrylamide gel electrophoresis patterns are shown of cell envelopes of pJP300 (lanes a to e) or pJP320 (lanes f to j) carrying derivatives of strains AB1157 (lanes a, b, f and g), CE1224 (lanes c, d, h and i) and CE1248 (lanes e and j). Cells of derivatives of AB1157 and CE1224 were grown in high-phosphate medium (lanes a, c, f and h) or low-phosphate medium (lanes b, d, g and i). Derivatives of *phoR* mutant strain CE1248 were grown on L-broth. Only the relevant part of the gel is shown. The positions of the PhoE, OmpF, OmpC and OmpA proteins are indicated at the right.

cleotides p26 and p32, indicated in Fig. 2, as primers. The results of these analyses are summarized in Figure 4, where the end-points of the deletions are indicated. The locations of the deletion end-points are in agreement with the three classes of mutations based on the PhoE expression levels, i.e. between positions −150 and −122 for class 1 mutations (pJP301 to pJP305), which do not affect *phoE* expression, between positions −106 and −54 for the class 2 mutations (pJP306 to pJP311), which reduce the level of expression, and between positions −27 and +70 for the class 3 mutations (pJP312 to pJP315), where PhoE expression is abolished. The observation that the expression of PhoE protein is abolished in the case of plasmid pJP312 is consistent with a role of the *pho* box in *phoE* expression. Sequences upstream from position −121 are apparently not required. Since the *pho* box and Pribnow box are preserved in the class 2 mutants, whereas the expression level is reduced, it appears that sequences upstream from the *pho* box are required for efficient expression of PhoE protein. At least part of this upstream element must be located between positions −121 (end-point of the deletion on pJP305) and −106 (end-point of the deletions on pJP306, 307 and 308). Comparison of the nucleotide sequence of this region with the sequences of the promoters of *phoA*, the *pst* operon and *phoB* did not reveal any significant homology. However, examination of the opposite strand revealed the presence of a potential *pho* box and correctly spaced Pribnow box in the reversed orientation (Fig. 4). This element will be designated the pseudo-*phoE*-promoter.

(c) *The region between the* pho *box and the upstream element*

The distance between the *pho* box and the pseudo-*phoE* promoter is 37 bp, 33 of which are A·T. We studied the importance of this region by inserting a G+C-rich 12-mer *Bam*HI linker. Plasmid pJP29 contains at least six *Ssp*I sites, one of which is located in this region at position −50 (Fig. 2). The plasmid was linearized by digestion with *Ssp*I in the presence of ethidium bromide and a 12-mer *Bam*HI linker was ligated into the linearized plasmid as described in Materials and Methods. After transformation and screening of the plasmid content of the transformants, a plasmid was elected with the *Bam*HI linker inserted in the desired position. This plasmid is designated pJP320.

Cell envelopes were isolated from *phoE* mutant strains AB1157 and CE1224 carrying pJP320 after growth under high and low-phosphate conditions, respectively (Fig. 6). The synthesis of PhoE protein was induced under phosphate limitation, but was reduced as compared to cells carrying pJP29. Similarly, the linker insertion affected the constitutive expression of the PhoE protein is a *phoR* mutant strain (Fig. 6, lanes e and j).

(d) *Is the pseudo-*phoE *promoter functional?*

Since the pseudo-*phoE* promoter is very similar to a regular promoter of the *pho* regulon, we wondered whether it could function as a promoter. To test this possibility, the *Sau*3A fragment extending from position −120 to +262 (Fig. 2) was cloned into the *Bam*HI site of the operon fusion vector pMLB1010 (Travers *et al.*, 1983) as described in Materials and Methods. The latter plasmid contains a *lacZ* gene without promoter but with its own ribosome-binding site and initiation codon. The *Sau*3A fragment was cloned in both orientations into this vector, yielding the plasmids pJP350 with the *phoE* promoter proximal to *lacZ* and pJP351 with the pseudo-*phoE* promoter proximal to *lacZ*.

The β-galactosidase activity was measured of derivatives of strain MC4100, carrying either one of the plasmids pMLB1010, pJP350 or pJP351. After growth in the presence of sufficient phosphate, all strains contained approximately equal β-galacto

Table 2
β-Galactosidase activity in strains carrying lacZ *operon fusions*

Strain	β-Galactosidase activity†	
	HP_i medium‡	LP_i medium‡
MC4100 (pMLB1010)	23·3	21·7
MC4100 (pJP350)	28·1	835·8
MC4100 (pJP351)	26·8	85·3
CE1265 (pMLB1010)	19·7	ND
CE1265 (pJP351)	89·9	ND

† β-Galactosidase activity is expressed as nmol of product formed per min per mg protein. ND, not determined.

‡ MC4100 derivatives were grown overnight in medium containing 660 μM-phosphate and then diluted 1 : 12·5 in fresh medium containing 660 μM-phosphate (HP_i) or no phosphate (LP_i) as described (Tommassen *et al.*, 1985). After 8 h subsequent incubation at 37 °C, β-galactosidase activity was measured. Derivatives of *phoR* mutant strain CE1265 were grown overnight in L-broth.

		pho box	PB
Consensus *pho* box		$CT^{G}_{T}TCATA^{A}_{T}A^{A}_{T}CTGTCA$ - 11 bp - TATAAT	
Pseudo-*phoE* promoters			
E.coli	-78	CTGTaATAAAAgcGTaA - 11 bp - TAggcT	-111
K.pneumoniae		CTGaCATATtTtTGTCA - 11 bp - cAgAAT	
E.cloacae		CgGTaATAAATCcGTCA - 11 bp - TAcgcT	

Figure 7. Comparison of the pseudo-*phoE* promoters in different *phoE* genes with the consensus *pho* box as shown in Fig. 1. In all cases, the pseudo-*phoE* promoter is separated by 37 bp from the primary *pho* box and is oriented in the opposite direction. The numbers −78 and −111 refer to the position relative to the transcription start point.

sidase activity (Table 2). As expected, growth of the pJP350-containing cells under phosphate starvation resulted in a drastic induction of β-galactosidase activity. Phosphate limitation also induced β-galactosidase activity in the pJP351-containing cells, although only three- to four-fold (Table 2). This was not the case in the pMLB1010-containing cells. Apparently, also in the case of pJP351, a phosphate limitation-inducible promoter has been cloned in front of the *lacZ* gene.

It was impossible to transform *phoS* or *phoR* mutant strains, which are constitutive for the *pho* regulon, with plasmid pJP350, probably because of the synthesis of a lethal fusion product under control of the *phoE* promoter. It was also impossible to transform *phoS* mutant strains with pJP351, but transformants were obtained from *phoR* mutant strain CE1265, although the colonies were relatively small. In this respect, it should be noted that the constitutive synthesis of PhoE protein is much lower in *phoR* mutant strains than in *phoS* mutant strains (Tommassen & Lugtenberg, 1980), probably because both the repressor function and the activator function of the PhoR protein is missing (Wanner & Latterell, 1980). After growth overnight in L-broth, cells of strain CE1265 carrying pJP351 exhibited a higher β-galactosidase level than CE1265 cells carrying pMLB1010 (Table 2). So, it appears that the promoter activity of the pseudo-*phoE* promoter is under control of the *pho* regulon.

4. Discussion

S_1 nuclease mapping and the primer extension method were used to identify the mRNA start site of the *phoE* gene. The former method revealed several potential start sites (Fig. 3(a)), which has been observed more often with this technique (e.g. Berman *et al.*, 1984; Hentschel *et al.*, 1980). However, the start site could be determined unambiguously by using the latter method. The sequenced promoters of the *pho* regulon contain a common 17 bp fragment, the *pho* box, which is followed at a distance of 11 bp by a putative Pribnow box (Surin *et al.*, 1984). The location of the mRNA start site of *phoE* appears to be consistent with the idea that this putative Pribnow box is indeed functional as the −10 region of the *phoE* promoter, and that the *pho* box functions as an aberrant −35 region. In this respect, the *pho*-controlled promoters resemble the promoters in other positively regulated systems, which often contain non-standard −35 regions (Raibaud and Schwartz, 1984). Therefore, these promoters are probably not properly recognized by RNA polymerase in the absence of the activator proteins.

Strains with mutations in the *phoR* gene produce PhoE protein constitutively, although at a much lower level than in fully induced strains or *phoS* mutant strains (Tommassen & Lugtenberg, 1980). This is probably due to the loss of both the repressor and the activator function of the phoR protein (Wanner & Latterell, 1980). In *phoR phoM* double mutants, the PhoE protein is not produced at all (Tommassen *et al.*, 1984), showing that the PhoM protein has an activator function (Wanner & Latterell, 1980). A possible explanation for these results is that the *phoE* promoter region actually contains two functional promoters, one stimulated by PhoR and PhoB, the other one by PhoM and PhoB. However, this possibility can now be ruled out, since identical results were obtained in S_1 mapping experiments, independent of the source of the mRNA.

Whereas the role of the *pho* box as an alternative −35 region in *pho*-controlled promoters is evident, it was not clear whether a *pho* box is sufficient for *pho*-controlled expression of PhoE protein. Therefore, deletions were created in the region upstream from the *pho* box. Deletions removing DNA upstream from position −121 relative to the mRNA start site did not influence *phoE* expression. However, deletions removing the DNA upstream from position −106 resulted in severely decreased expression. The residual expression was still *pho*-controlled, since it was inducible by phosphate starvation and constitutive in *phoR* mutant strains. These results show that an upstream element, which is located at least in part between residues −106 and −121, is involved in the expression of *phoE*. However, it shows also that this upstream element is not a unique binding site for an essential activator of the *pho* regulon.

What is the nature of this upstream element? Interestingly, the upstream region contains sequences homologous to a *pho* box and a correctly spaced Pribnow box but in the opposite orientation

relative to the primary *pho* box and −10 region of the *phoE* promoter. It seems very likely that this pseudo-*phoE* promoter is the required upstream element. This supposition is underscored by the observation that the pseudo-*phoE* promoter is conserved in *phoE* genes from other species. Recently, the *phoE* genes of *Klebsiella pneumoniae* and *Enterobacter cloacae* have been sequenced (van der Ley *et al.*, 1987) and, although the homology with the *E. coli phoE* gene is weak upstream from the primary *pho* boxes, a pseudo-*phoE* promoter can be discerned also in these promoters, in the same orientation as in the *E. coli phoE* gene (Fig. 7). Also, the spacing between the *pho* box and the pseudo-*phoE* promoter is identical, i.e. 37 bp, in all three promoters, which appears to be consistent with the observation that an oligonucleotide insertion in this region reduces *phoE* expression.

What is the function of the upstream element? The situation in the *phoE* promoter is reminiscent of the *tyrT* gene of *E. coli* encoding one of two major $tRNA^{Tyr}$ species. The region of DNA between −40 and −98 was shown to be required for the efficient expression of this gene (Lamond & Travers, 1983) but not for the regulatory response (Lamond & Travers, 1985). This region contains sequences homologous to the −35 and −10 regions of *E. coli* promoters, but in the opposite orientation to the primary RNA polymerase binding site. DNAase I protection experiments showed that this secondary site can actually bind RNA polymerase *in vitro* (Travers *et al.*, 1983). Similarly, it appears that the pseudo-*phoE* promoter can bind RNA polymerase *in vivo*, since a weak promoter activity could be demonstrated by cloning it in an operon fusion vector. In the case of *phoE*, binding of RNA polymerase to this secondary site appears to be dependent on the *pho* regulatory system, since promoter activity of the pseudo promoter could be detected only under phosphate limitation or in a constitutive *phoR* mutant strain, and since the pseudo promoter contains a *pho* box instead of a regular −35 region with a TTGACA consensus sequence. In the case of the *tyrT* gene, it was suggested that binding of RNA polymerase to this upstream binding site functions by activating initiation by polymerase bound at the primary binding site and/or by increasing the concentration of polymerase in the vicinity of the *tyrT* promoter (Travers *et al.*, 1983). Similar functions might be attributed to the pseudo-*phoE* promoter and are likely to be disturbed by insertion mutations in the intervening region.

The *phoE* gene is co-regulated with *phoA*, the *pst* operon and the *ugp* operon. Inspection of the nucleotide sequences of the promoter regions of these genes did not reveal any pseudo-*pho* promoters as in the *phoE* case, i.e. consisting of a *pho* box and a properly spaced Pribnow box oriented opposite to these genes. However, it should be noted that the promoters appear to be tandemly repeated in the *pst* operon (Surin *et al.*, 1984; Makino *et al.*, 1986) and in the *ugp* operon (Overduin *et al.*, unpublished results). In the former case, the Pribnow box belonging to the upstream *pho* box overlaps with the second *pho* box. Potential secondary RNA polymerase binding sites could not be detected upstream or downstream from the promoter of the *phoA* gene (Kikuchi et al., 1981). However, if the role of these secondary binding sites is to increase the concentration of RNA polymerase in the vicinity of the primary binding site, another mechanism may have evolved for *phoA*. In this respect, it should be noted that the *phoA* gene is preceded by an open reading frame in the same orientation (Kikuchi *et al.*, 1981). This putative gene is not followed by a G+C-rich dyad symmetry and a poly(T) sequence, which is characteristic for a Rho-independent transcription terminator (von Hippel *et al.*, 1984). So, transcription of this gene might result in an increased concentration of RNA polymerase bound near the *phoA* promoter. The promoter of this open reading frame has not been sequenced; it will be interesting to see whether it contains a *pho* box.

We thank C. Verhagen and W. van Duyn for assistance in some of the experiments, D. Bosch for providing plasmids and advice in S_1 mapping experiments, S. Langeveld for advice in primer extension experiments and E. Bremer for providing plasmids.

References

Argast, M. & Boos, W. (1980). *J. Bacteriol.* **143**, 142–150.
Berk, A. J. & Sharp, P. A. (1977). *Cell*, **12**, 721–732.
Berman, M. L., Jackson, D. E., Fowler, A., Zabin, I., Christensen, L., Fiil, N. P. & Hall, M. N. (1984). *Gene Anal. Techn.* **1**, 43–51.
Birnboim, H. C. & Doly, J. (1979). *Nucl. Acids Res.* **7**, 1513–1524.
Bosch, D., Leunissen, J., Verbakel, J., De Jong, M., Van Erp, H. & Tommassen, J. (1986). *J. Mol. Biol.* **189**, 449–455.
Casadaban, M. J. (1976). *J. Mol. Biol.* **104**, 541–555.
Chang, A. C. Y. & Cohen, S. N. (1978). *J. Bacteriol.* **134**, 1141–1156.
Giorgi, C., Blumberg, B. M. & Kolakofsky, D. (1983). *Cell*, **35**, 829–836.
Guan, C. D., Wanner, B. & Inouye, H. (1983). *J. Bacteriol.* **156**, 710–717.
Hentschel, C., Irminger, J.-C., Bucher, P. & Birnstiel, M. L. (1980). *Nature (London)*, **285**, 147–151.
Kikuchi, Y., Yoda, K., Yamasaki, M. & Tamura, G. (1981). *Nucl. Acids Res.* **21**, 5671–5678.
Korneluk, R. G., Quan, F. & Gravel, R. A. (1985). *Gene*, **40**, 317–323.
Korteland, J., Tommassen, J. & Lugtenberg, B. (1982). *Biochim. Biophys. Acta*, **690**, 282–289.
Korteland, J., Overbeeke, N., De Graaff, P., Overduin, P. & Lugtenberg, B. (1985). *Eur. J. Biochem.* **152**, 691–697.
Lamond, A. I. & Travers, A. A. (1983). *Nature (London)*, **305**, 248–250.
Lamond, A. I. & Travers, A. A. (1985). *Cell*, **40**, 319–326.
Lugtenberg, B., Meijers, J., Peters, R., Van der Hoek, P. & Van Alphen, L. (1975). *FEBS Letters*, **58**, 254–258.
Magota, K., Otsuji, N., Miki, T., Horiuchi, T., Tsunasawa, S., Kondo, J., Sakiyama, F., Amemura,

M., Morita, T., Shinagawa, H. & Nakata, A. (1984). *J. Bacteriol.* **157**, 909–917.

Makino, K., Shinagawa, H., Amemura, M. & Nakata, A. (1986). *J. Mol. Biol.* **190**, 37–44.

Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). Editors of *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Miller, J. H. (1972). *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Nakata, A., Petersen, G. R., Brooks, E. L. & Rothman, F. (1971). *J. Bacteriol.* **107**, 683–689.

Overbeeke, N., Bergmans, H., Van Mansfeld, F. & Lugtenberg, B. (1983). *J. Mol. Biol.* **163**, 513–532.

Raibaud, O. & Schwartz, M. (1984). *Annu. Rev. Genet.* **18**, 173–206.

Rosenberg, M. & Court, D. (1979). *Annu. Rev. Genet.* **13**, 319–353.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5463–5467.

Schweizer, H., Grussenmeyer, T. & Boos, W. (1982). *J. Bacteriol.* **150**, 1164–1171.

Surin, P., Jans, D. A., Fimmel, A. L., Shaw, D. C., Cox, G. B. & Rosenberg, H. (1984). *J. Bacteriol.* **157**, 772–778.

Tommassen, J. & Lugtenberg, B. (1980). *J. Bacteriol.* **143**, 151–157.

Tommassen, J. & Lugtenberg, B. (1981). *J. Bacteriol.* **147**, 118–123.

Tommassen, J. & Lugtenberg, B. (1982). *Annu. Microbiol. (Inst. Pasteur)*, **133A**, 243–249.

Tommassen, J., De Geus, P., Lugtenberg, B., Hackett, J. & Reeves, P. (1982*a*). *J. Mol. Biol.* **157**, 265–274.

Tommassen, J., Overduin, P., Lugtenberg, B. & Bergmans, H. (1982*b*). *J. Bacteriol.* **149**, 668–672.

Tommassen, J., Van Tol, H. & Lugtenberg, B. (1983). *EMBO J.* **2**, 1275–1279.

Tommassen, J., Hiemstra, P., Overduin, P. & Lugtenberg, B. (1984). *Mol. Gen. Genet.* **195**, 190–194.

Tommassen, J., Leunissen, J., Van Damme-Jongsten, M. & Overduin, P. (1985). *EMBO J.* **4**, 1041–1047.

Travers, A. A., Lamond, A. I., Mace, H. A. F. & Berman, M. L. (1983). *Cell*, **35**, 265–273.

van den Elzen, P. J. M., Konings, R. N. H., Veltkamp, E. & Nijkamp, H. J. J. (1980). *J. Bacteriol.* **144**, 579–591.

van der Ley, P., Bekkers, A., Van Meersbergen, J. & Tommassen, J. (1987). *Eur. J. Biochem.* **164**, 469–475.

Villarejo, M., Davis, J. L. & Granett, S. (1983). *J. Bacteriol.* **156**, 975–979.

von Hippel, P. H., Bear, D. G., Morgan, W. D. & McSwiggen, J. A. (1984). *Annu. Rev. Biochem.* **53**, 389–446.

Wanner, B. L. & Latterell, P. (1980). *Genetics*, **96**, 353–366.

Willsky, G. R., Bennett, R. L. & Malamy, M. H. (1973). *J. Bacteriol.* **113**, 529–539.

Edited by J. H. Miller