Molecular Biology of Carbon–Phosphorus Bond Cleavage

CLONING AND SEQUENCING OF THE phn (psiD) GENES INVOLVED IN ALKYLPHOSPHONATE UPTAKE AND C–P LYASE ACTIVITY IN ESCHERICHIA COLI B*

(Received for publication, April 28, 1989)

Chih-Ming Chen‡, Qi-Zhuang Ye‡, Zhimin Zhu§, Barry L. Wanner‡, and Christopher T. Walsh¶†
From the §Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115 and the ¶Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907

Whereas bacteria such as Escherichia coli have been known for some time to cleave carbon–phosphorus (C–P) bonds in unactivated alkylphosphonates, the enzymes responsible for C–P lyase activity have resisted detection or purification. Genes from E. coli B that support growth on alkylphosphonates as the sole phosphorus source have now been cloned (B. L. Wanner and J. A. Boline, unpublished data). Deletion analysis demonstrated that at least 13 kilobases of DNA information is required for E. coli to express the phosphonate utilization phenotype (Phn"). The complete nucleotide sequence of 15,611 bases has been determined, and the gene structure was examined. Seventeen open reading frames (phnA to phnQ) were identified in one transcriptional direction and five open reading frames in the divergent direction. Sequence homology searches identify PhnC, PhnK, PhnL, and, possibly, PhnN proteins as members of nucleotide-binding proteins of the binding protein-dependent transport systems. Candidates for other membrane components and regulatory proteins are also identified. A Pho box-like promoter sequence is also found upstream of the gene cluster starting at phnA, which is consistent with the observation of phosphate regulation of the Phn" response. Fourteen repetitive extragenic palindromic sequences are found in the phn DNA. 10 exist in the extragenic region between phnA and phnB, two between phnD and phnE, and two between phnK and phnL. An unusual finding is that one of the repetitive extragenic palindromic sequences actually overlaps with the reading frame of the phnE gene.

Whereas the great preponderance of phosphorus-containing molecules in biology are phosphate esters, with four oxygen substituents on phosphorus, there are both naturally occurring and synthetic phosphorus species with one (organophosphonate) or two (organophosphinate) carbon–phosphorus (C–P) bonds. Thus, phospholipids containing aminoethylphosphonate as head group rather than phosphatidylethanolamine are dominant constituents in Tetrahymena cell membranes (Hilderbrand and Henderson, 1983), whereas phosphoinothricin (2-amino-4-(methylphosphonyl)butanonic acid) is a potent inhibitor of bacterial and plant gluten synthetases (Colanduoni and Villafranca, 1986; Leeson et al., 1982). Synthetic organophosphonates and phosphinites have been widely used as insecticides, and N-phosphonomethylglycine (glyphosate) is a very widely used herbicide (Hilderbrand and Henderson, 1983).

The bacterial degradation of C–P-containing compounds appears to involve direct C–P cleavage (Cordeiro et al., 1986; Shinabarger and Braymer, 1986; Avila et al., 1987; Kishore and Jacob, 1987; Loo et al., 1987; PＦke et al., 1987). Whereas a bacterial phosphonatase degrading aminoethylphosphonate via phosphonoacetalddehyde and an enzyme-Lys-NH₂ Schiff’s base catalytic mechanism has been purified and initially characterized for structure and mechanism (La Nauze et al., 1970), the manner of enzymatic cleavage of unactivated alkylphosphonates (and dialkylphosphonates) has been obscure, as they are stable to both strong acid and base. Whereas many bacteria that can degrade alkylphosphonates (e.g. methylphosphonate, CH₃PO₂⁻ → CH₃ + HPO₂⁻) as a source of phosphorus for growth can be isolated by selective culture, only one successful attempt of detection of a C–P lyase activity has been reported in broken cells (Murata et al., 1988); in that case, only phosphate release in the presence of a phosphate (phosphonoacetate) has been demonstrated. The fate of the organic moiety is unknown. A number of chemical studies have suggested redox chemistry at the phosphorus center could be involved (Frost et al., 1987; Shames et al., 1987), and an intact membrane may be required for some membrane electron transport process to function.

We recently noted (Wackett et al., 1987b) that an Escherichia coli strain that could grow on methylphosphonate and yield quantitative amounts of methane gas as coproduct to the inorganic phosphate ceased to make CH₄ when exogenous P was added to the medium. This inorganic phosphate repression had all the characteristics of phosphate Pho regulon behavior (Wanner et al., 1981), including sensitivity to phoR and phoB gene effects (Wackett et al., 1987b). Thus, we screened 54 Muc(lacZ) fusions in phosphate starvation-inducible (psi) genes in E. coli and observed that an insertion in the psiD locus prevented growth in methylphosphonate as well as in propionylphosphonate and phosphophosphoninate. Accordingly, we have renamed the psiD locus as phnD for phosphonate assimilation. The Phn" phenotype could be complemented by a 15.6-kb DNA insert, and this clone was

1 The abbreviations used are: kb, kilobase(s); Rep, repetitive extragenic palindromic; MOPS, 4-morpholino propane sulfonic acid; ORF, open reading frame; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.
2 B. L. Wanner and J. A. Boline, unpublished data.
a reasonable candidate to harbor structural genes encoding C-P lyase components. The smallest subclone that still generates a Phn" phenotype is a 13.4-kb fragment. In this paper, we report the determination of the DNA sequence for this 15.6-kb DNA insert and encoded open reading frames of relevance to phosphonate utilization and C-P lyase activity. Preliminary evidence for complementation of the Phn" phenotype in an E. coli mutant by a genomic BamHI digest has been reported elsewhere by Loo et al. (1987).

**MATERIALS AND METHODS**

The bacterial strains and plasmids used are listed in Table I.

Restriction enzymes were purchased from various commercial sources and were used according to the manufacturers' specification.

Plasmid pBW120 contains a 15.6-kb BamHI fragment of E. coli B DNA inserted in pUC18.² Plasmid construction and transformation followed the procedures of Maniatis et al. (1982). Plasmids pIB10 and pIB13 were purchased from International Biotechnology, Inc., and were used for both plasmid construction and T7 RNA polymerase-based gene expression. 2 µg of pBW120 was digested with EcoRI and BamHI to completion and ligated to EcoRI- and/or BamHI-treated pIB1 vectors. The ligation mixture was introduced into strain BW1134 and HB101(pGPl-2) for study of methylphosphonate utilization and T7 RNA polymerase-dependent gene expression, respectively. The 15.6-kb BamHI fragment was also cloned into the transcriptional direction of the T7 promoter; and these plasmids were isolated. To determine the minimal size of the DNA required for the cellular expression of methylphosphonate utilization, terminal fragments of the 15.6-kb insert were removed by partial Sau3A digestion of pCMC1, partial SmaI digestion of pCMC1, KpnI digestion of pCMC1, or SphI digestion of pCMC2, generating plasmids pCMC10 to pCMC13, listed in Table I.

The utilization of methylphosphonate by DW11304 strains harboring different recombinant plasmids was assessed by growth in glucose/MOPS (Wanner, 1987) liquid or agar medium supplemented with 0.1 mM methylphosphonate as sole phosphorus source.

Plasmid pBW120 was digested with BamHI and EcoRI to complement the Phn" (15.6-kb BamHI fragment in pUC18) strain with the Klonozyme enzyme (Bethesda Research Laboratories), and the blunt-ended fragments were cloned into the HindII site of M13mp19. The recombinant phages were characterized by restriction digestion of the replicative forms, and the orientations were tested as described by Barnes et al. (1983). Sequencing clones were created by using the Cyclone kit (International Biotechnology, Inc.), and the extents of deletion were analyzed by GGB (180 mM Tris base, 80 mM sodium acetate, 0.8 mM Na₂EDTA, pH 8.3) agarose gel electrophoresis. DNA sequences were determined by the dideoxy chain termination method of Sanger et al. (1977) using 3⁵-dATP (Du Pont-New England Nuclear) and a Sequenase™ kit (United States Biochemical Corp.). Sequenase™ enzymes (versions I and II) were also obtained from Dr. S. Tabor of this department as gifts. Sequence data were analyzed by a DNASTAR sequence analysis package, the Wisconsin UWGCG program, and homology search against National Biomedical Research Foundation protein data base release 18.

The polypeptide chains encoded by the 15.6-kb BamHI fragment were analyzed using a T7 RNA polymerase expression system. Strain HB101 was first transformed by the T7 RNA polymerase-containing plasmid pOP1-2. The kanamycin-resistant strain was further transformed with plasmids pIB10 and pIB13 and their derivatives which contained the desired DNA fragments under T7 promoter control. The transformants were selected simultaneously by kanamycin and ampicillin. [5⁶S]Methionine-labeled cells were prepared by the method of Tabor and Richardson (1985). The labeled proteins were analyzed by SDS-PAGE according to Laemmli (1970). Protein gels were

**Table 1**

<table>
<thead>
<tr>
<th>Bacterial strains and plasmids</th>
<th>Name</th>
<th>Properties (genotypes and/or phenotypes)</th>
<th>Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacterial strains</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HB101</td>
<td></td>
<td>F&quot; hsdS2(r8° m16°) recA13 ara-14 proA2 lacY1 galK2 rpsL20(Sm') xylA5 metI-1 supE44 λ&quot; recA::cat(Cm') lac-169 Δ(psdD35-30 pho-510 thi-'(F'128-1270-11(tc) lacZΔ165) pheN(ΔB01) DE3(lac)X74 Δ(phoA82 PvuII) phoM(Δt)</td>
<td>Maniatis et al. (1982)</td>
</tr>
<tr>
<td>BW11334</td>
<td></td>
<td></td>
<td>Footnote 2</td>
</tr>
<tr>
<td>BW12720</td>
<td></td>
<td></td>
<td>Footnote 2</td>
</tr>
<tr>
<td>JM109</td>
<td></td>
<td>recA1 endA1 gyrA96 thi'- hsdR17(r6° m16°) supE44 relA1 λ&quot; Δ(lac-proAB) (F' traD36 proAB lacFΔZΔ15)</td>
<td>Yanisch-Perron et al. (1985)</td>
</tr>
<tr>
<td>Plasmids</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBW120</td>
<td></td>
<td>Phn&quot; (15.6-kb BamHI fragment in pUC18) Ap'</td>
<td>Footnote 2</td>
</tr>
<tr>
<td>pGP1-2</td>
<td></td>
<td>Km&quot;</td>
<td>Tabor and Richardson (1986)</td>
</tr>
<tr>
<td>pCMC1</td>
<td></td>
<td>Phn&quot; (15.6-kb BamHI fragment in pIB131) Ap'</td>
<td>This study</td>
</tr>
<tr>
<td>pCMC2</td>
<td></td>
<td>Phn&quot; (15.6-kb BamHI fragment in pIB31) Ap'</td>
<td>This study</td>
</tr>
<tr>
<td>pCMC4</td>
<td></td>
<td>Phn&quot; (4.2-kb BamHI-EcoRI fragment in pIB31) Ap'</td>
<td>This study</td>
</tr>
<tr>
<td>pCMC5</td>
<td></td>
<td>Phn&quot; (3.5-kb EcoRI fragment in pIB31) Ap'</td>
<td>This study</td>
</tr>
<tr>
<td>pCMC6</td>
<td></td>
<td>Phn&quot; (8-kb EcoRI-BamHI fragment in pIB30) Ap'</td>
<td>This study</td>
</tr>
<tr>
<td>pCMC7</td>
<td></td>
<td>Phn&quot; (4.2-kb BamHI-EcoRI fragment in pIB30) Ap'</td>
<td>This study</td>
</tr>
<tr>
<td>pCMC8</td>
<td></td>
<td>Phn&quot; (3.5-kb EcoRI fragment in pIB30) Ap'</td>
<td>This study</td>
</tr>
<tr>
<td>pCMC9</td>
<td></td>
<td>Phn&quot; (8-kb EcoRI-BamHI fragment in pIB31) Ap'</td>
<td>This study</td>
</tr>
<tr>
<td>pCMC10</td>
<td></td>
<td>Phn&quot; (KpnI deletion of pCMC1) Ap'</td>
<td>This study</td>
</tr>
<tr>
<td>pCMC11</td>
<td></td>
<td>Phn&quot; (SmaI deletion of pCMC2) Ap'</td>
<td>This study</td>
</tr>
<tr>
<td>pCMC12</td>
<td></td>
<td>Phn&quot; (SphI deletion of pCMC1) Ap'</td>
<td>This study</td>
</tr>
<tr>
<td>pCMC13</td>
<td></td>
<td>Phn&quot; (SphI deletion of pCMC2) Ap'</td>
<td>This study</td>
</tr>
<tr>
<td>pCMC17</td>
<td></td>
<td>Phn&quot; (SalI deletion of pCMC1) Ap'</td>
<td>This study</td>
</tr>
<tr>
<td>pQY1</td>
<td></td>
<td>Phn&quot; (3.5-kb SalI-BamHI fragment in pIB30) Ap'</td>
<td>This study</td>
</tr>
<tr>
<td>pQY2</td>
<td></td>
<td>Phn&quot; (4.3-kb EcoRI-SalI fragment in pIB30) Ap'</td>
<td>This study</td>
</tr>
</tbody>
</table>
stained with Coomassie Blue and then dried by gel dryer under vacuum. Autoradiography was performed using Kodak XAR-5 film.

RESULTS

Plasmid Constructions—Plasmids pCMC1 and pCMC2, like the parental plasmid pBW120, support the growth of BW11334 on methylphosphonate plates. This shows that the 15.6-kb insert has the full information for the Phn+ phenotype and that the methylphosphonate utilization genes are probably expressed under control of a promoter contained within the 15.6-kb fragment. To determine the minimum amount of DNA required for the Phn+ character, smaller clones were isolated. An analysis of six recombinant plasmids, pCMC4 to pCMC9 (Fig. 1), showed that none of these could support the growth of BW11334 on methylphosphonate medium. This shows that both EcoRI sites are located in the region required for the Phn+ phenotype and that more than 3.5 kb of DNA is needed. Clones were also made in which terminal fragments were removed from the 15.6-kb insert, from either end. Complete digestion of pCMC1 with KpnI and religation generated plasmid pCMC10, which had lost the terminal 2.2-kb KpnI fragment (Fig. 1). Plasmid pCMC1 partially digested with SmaI generated clones pCMC12 and pCMC11; and SphI-digested pCMC2 generated pCMC13, which had lost the indicated 4.4-kb fragment. Plasmids pCMC10 to pCMC13 retained the ability to grow on methylphosphonate agar, whereas pCMC13 did not when transformed into BW11334. This indicates that up to 2.2 kb of DNA on the right end is not necessary for the Phn+ phenotype, and the adjacent (as shown in Fig. 1) 2.2-kb KpnI–SphI fragment is required for the expression of the Phn+ phenotype. Terminal deletions of the other end of the 15.6-kb DNA insert were investigated by partial SalI digestion of pCMC1 DNA. Plasmid pCMC17, which has deleted the terminal 1.7-kb SalI–BamHI fragment, was isolated (Fig. 1) and shown unable to support the growth of BW11334 on methylphosphonate medium. These results indicate that at least 13 kb of the cloned DNA is required for methylphosphonate utilization and defines both the right and left ends of DNA required for complementation of Phn+ in BW11334. In addition to defining the minimal length of plasmid DNA required for complementation of the Phn+ phenotype, one needs to know the extent of the chromosomal deletion in strain BW11334 to assess how many of the phn genes are actually required for alkylphosphonate (such as methyl-, ethyl-, and aminoethylphosphonate) uptake and cleavage. Mapping studies in E. coli BW11394 show that the deletion encompasses phnD to phnQ and the downstream 10 kb of adjacent DNA but that intact copies of genes phnA to phnC are present. The only indication, then, that any of the genes phnA to phnC may be required for Phn+ phenotype is derived from the DNA sequence information and the detection of a Pho box upstream of the phnA reading frame (see below), which would account for our observed negative regulation of the phn genes (Wackett et al., 1987b) by inorganic phosphate.

DNA Sequence of phn Genes Involved in C–P Lyase Activity

DNA Sequencing and Open Reading Frames—DNA sequence determination was carried out by the dideoxy chain termination method (Sanger et al., 1977) using single-stranded M13 DNA as template. To do this, the 15.6-kb insert of pBW120 was digested with BamHI and EcoRI and blunt end-ligated to M13mp19 after the ends were filled in with the Klenow enzyme. Five out of the six expected clones were obtained. In more than 100 clones analyzed, only one recombinant contained the 8-kb EcoRI–BamHI fragment. Plasmid pCMC9, which also contains the 8-kb fragment, was digested with SalI, SalI plus BamHI, or SalI plus EcoRI and ligated to similarly treated M13mp18 or M13mp19 DNA. Clones suitable for DNA sequencing were generated by the T4 DNA polymerase-dependent single-strand deletion method (Dale et al., 1985) and then sequenced and analyzed as described under "Materials and Methods." The complete sequences for both strands of 15,611 base pairs was obtained and showed perfect agreement. The nucleotide sequence and the translated peptide sequences corresponding to each open reading frame (ORF) are shown in Fig. 2, where the BamHI sites for the initial cloning, act as reference points. The DNA sequence in Fig. 2A is oriented as shown for plasmid pCMC2 in Fig. 1.

Seventeen open reading frames were found on one DNA strand and are located from nucleotide 2,874 to the end of the 15.6-kb insert. The deletion data above show that much or all of this region is required for the Phn+ phenotype and the underlying C–P bond cleavage. These ORFs have therefore been designated as phn (phosphonate) genes, specifically phnA to phnQ (Fig. 3 and Table II). The first open reading frame of this strand (phnA) starts at nucleotide 2,874 and extends for 336 nucleotides. After an intergenic region of 557 nucleotides, the phnB gene starts at nucleotide 3,767 and ends at nucleotide 4,210. The third gene (phnC) begins at nucleotide 4,343 and stops at nucleotide 5,131. The intergenic length between phnB and phnC is 132 bases. phnD is an open reading frame of 1,017 nucleotides in length, and its initiation codon frame of 1,017 nucleotides in length, and its initiation codon is located 24 bases downstream from the termination of phnC. After 54 nucleotides of nontranslated region, the fifth gene (phnE) starts at nucleotide 6,227 and terminates at nucleotide 7,057. phnF is located 20 nucleotides away and extends for 726 nucleotides. The termination codon for phnF at nucleotide 7,903 overlaps the start of phnG at nucleotide 7,804. phnG extends for 453 nucleotides and terminates at nucleotide
FIG. 2. A, nucleotide sequence of *phn* genes and flanking region. A total of 15,611 bases are shown with translated peptide sequences. The complementary sequence of bases 1–2,800 is also presented and listed on the upper lines. The translated peptide sequences of *phnA* to *phnQ* are located below the DNA sequence. The peptide sequence of ORF 742 is situated above its DNA sequence and is written in reverse order (i.e. C to N terminus). The -35 and -10 sequences of proposed promoters are boxed, and the direction of transcription is indicated by thin arrows. Each Rep unit is marked with a boldface letter, and the direction is indicated by thick arrows. The predicted ribosomal binding sites of each open reading frames are underlined. 

B, DNA and deduced peptide sequences of ORFs 146, 269, 114, and 126. DNA sequences (written 5' to 3') corresponding to the mRNA sequences are shown with the translated peptide sequences. The numbers indicate the positions corresponding to the entire DNA sequence shown in Fig. 2A. Since these four ORFs are encoded divergently as compared to the *phn* genes, the numbers indicated are in reverse order.

FIG. 2—continued
DNA Sequence of phn Genes Involved in C-P Lyase Activity

The eighth gene of this cluster is phnH, which begins at nucleotide 8,253 and ends at nucleotide 8,837. This generates an overlapping situation where four bases (ATGA) are shared by phnG and phnH. The ninth open reading frame (phn1) is the longest in this group and has 1,065 nucleotides, ending at nucleotide 9,901. One base (A) is shared by phn1 and its preceding gene (phnH). Overlapping sequences are found for the next two ORFs in this cluster: the last eight nucleotides of phn1 (ATGGCTAA) are used by phnJ, and four nucleotides (ATGA) of phnJ overlap with phnK, which starts at nucleotide 10,736 and ends at nucleotide 11,494. An intergenic space of 110 bases follows phnK; and the next cluster of genes, as judging by the closeness of each ORF, begins at nucleotide 11,605. The first ORF of this final cluster (phnL) begins at nucleotide 11,605 and terminates at nucleotide 12,285. The pattern of gene overlapping is observed again in this gene cluster. By sharing four nucleotides (ATGA), phnM starts at nucleotide 12,282 and stops at nucleotide 13,418. This is the longest open reading frame (1137 bases) found in the phn genes. phnN is the third ORF of this cluster and is located at nucleotides 13,418-13,975. Both phnM and phnN share a single A base in their reading frames. A 14-base overlap between phnN and phnO is found in which the phnO gene begins at nucleotide 13,962 and ends at nucleotide 14,396. Separated by a single base (C), the initiation codon for the next gene (phnP) follows. phnP has a length of 759 nucleotides. The last open reading frame (phnQ) is located at nucleotides 15,131-15,496. The overlap of 26 nucleotides between phnP and phnQ is the largest in the entire phn DNA. Similar patterns in the length of overlapping genes are seen in both phnG to phnJ and phnL to phnO clusters as 4 bases and then 1 base followed by more than 8 bases. The properties of the 17 phn gene products are listed in Table II. Putative Shine-Dalgarno sequences exist for each of these 17 ORFs (Shine and Dalgarno, 1974). The distances between the Shine-Dalgarno sequence and initiation
codons were analyzed as shown in Table II. The distance between the Shine-Dalgarno sequence and the initiation codon ATG are from five to eight nucleotides, with the exception of phnP (9 bases), phnK (10 bases), and phnF (11 bases). The analysis of the termination codon used in these genes shows that the opal codon (TGA) is preferentially chosen, and only four ORFs use the ochre codon (TAA). None has an amber codon (TAG). These findings are consistent with the potential for functional expression of all 17 phn genes.

Eighteen open reading frames with lengths longer than 100 amino acids were identified on the strand with the transcriptional direction opposite to the phn genes. Only five ORFs have a putative Shine-Dalgarno sequence, however. The positions of these ORFs are shown in Fig. 3, and their properties are also included in Table II. These ORFs are designated by their number of amino acids and are well separated throughout the entire 15.6-kb DNA fragment. DNA sequences and the translated peptide sequences are shown in Fig. 2, A and B (ORFs 146, 269, 114, and 126). Since the removal of ~2.2 kb of DNA in the construction of plasmids pCMC10 to pCMC12 failed to abolish the Phn+ phenotype, ORF 742 is not involved in methylphosphonate utilization. Other plasmid constructions, e.g. pCMC13, indicate that none of the four other ORFs (ORFs 146, 269, 114, and 126) is alone sufficient for a Phn+ phenotype. All five ORFs on the opposite strand

<table>
<thead>
<tr>
<th>Gene Products</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>phnA</td>
<td></td>
</tr>
<tr>
<td>phnB</td>
<td></td>
</tr>
<tr>
<td>phnC</td>
<td></td>
</tr>
<tr>
<td>phnD</td>
<td></td>
</tr>
<tr>
<td>phnE</td>
<td></td>
</tr>
<tr>
<td>phnF</td>
<td></td>
</tr>
<tr>
<td>phnG</td>
<td></td>
</tr>
<tr>
<td>phnH</td>
<td></td>
</tr>
<tr>
<td>phnI</td>
<td></td>
</tr>
<tr>
<td>phnJ</td>
<td></td>
</tr>
<tr>
<td>phnK</td>
<td></td>
</tr>
<tr>
<td>phnL</td>
<td></td>
</tr>
<tr>
<td>phnM</td>
<td></td>
</tr>
<tr>
<td>phnN</td>
<td></td>
</tr>
<tr>
<td>phnO</td>
<td></td>
</tr>
<tr>
<td>phnP</td>
<td></td>
</tr>
<tr>
<td>phnQ</td>
<td></td>
</tr>
<tr>
<td>ORF 742</td>
<td></td>
</tr>
<tr>
<td>ORF 146</td>
<td></td>
</tr>
<tr>
<td>ORF 269</td>
<td></td>
</tr>
<tr>
<td>ORF 114</td>
<td></td>
</tr>
<tr>
<td>ORF 126</td>
<td></td>
</tr>
</tbody>
</table>

* Number of bases between the Shine-Dalgarno (SD) sequence (AAGGAGGT) and initiation codon.

- With potential helix-turn-helix sequence.
- With conserved sequences similar to the integral membrane proteins of the binding protein-dependent transport systems.
- With portion of nucleotide-binding sequences.
have TGA as a termination codon, and all except ORF 742 use GTC instead of ATG as an initiation codon.

**Promoters and Pho Box**—Since both pCMC1 and pCMC2 support the growth of BW11334 on methylphosphonate medium, it is likely that the phn genes are expressed by a promoter(s) within the cloned DNA. The DNA sequence was searched for sequences similar to the canonical -35 promoter sequence (TTGACA). Candidate(s) were then screened manually for the presence of a -10, TATAAT-like sequence located at the proper position, i.e., 16-21 bases away from the -35 region. The final results were compared to the list of 263 known promoters for *E. coli* genes (Harley and Reynolds, 1987). A similar procedure was performed using the -10 sequence first, then looking for the -35 region. In these ways, six promoter-like sequences were found. Rightward promoters (designated as PR-1 to PR-3) are potential candidates for expressing the phn genes. The promoter candidates (PL-1 to PL-3) are for the opposite direction (Fig. 4A). No ORF was found near PR-1, PR-2, or PL-2, whereas PR-3 is located within the phnB sequence, PL-1 is in ORF 742, and PL-3 is in ORF 114. Both PR-1 and PR-2 are located upstream of the phnA gene between nucleotides 2622 and 2650 (PR-1) and 2738 and 2776 (PR-2), suggesting that either or both could be used for expressing the phn genes. PL-2 is located between nucleotides 2715 and 2742 with -35 sequence (TTGAAA), and a spacing of 16 bases and could be used for expressing ORF 742. These analyses suggest, but do not prove, that the phn genes may be organized as an operon and transcribed as a polycistronic message.

Phosphonate assimilation is repressed by inorganic phosphate. Also, the phn (psiD) gene(s) has a Pho regulon promoter whose expression requires both phoB and phoR genes (Wackett et al., 1987b). We therefore searched for homology to the consensus Pho box sequence that precedes other phoB- and phoR-dependent promoters (Makino et al., 1986). Inspection of the promoter sequences revealed that the -35 region for PR-2 can qualify as the Pho box for the proposed phn operon (Fig. 4b). Nine of 11 bases at the -10 end match the Pho box consensus sequence. However, this putative Pho box and the -10 region are only seven nucleotides apart, whereas in other cases, a distance of 10 bases prevails.

**Rep Sequences**—During sequencing of the region at nucleotides 3,000-3,700, repeating units of ~100 bases were evident on the sequencing gel, and the sequence was difficult to delineate due to compression. The use of dITP in reaction mixtures and increased gel running temperature to 60–65 °C resolved this problem. Sequence analysis confirmed a repeating pattern for four units of 100 nucleotides (with at least 90% identity) which are linked in a head-to-tail fashion. An analysis showed that each ~100-base pair segment actually contains two Rep sequences, known from earlier studies (Stern et al., 1984). A total of 10 Rep units (denoted a to j) are seen between nucleotides 3,224 and 3,706 (Fig. 5) which occupy the intergenic region between phnA (which ends at nucleotide 3,210) and phnB (which begins at nucleotide 3,767). Two more pairs of Rep units, k-l, and m-n, are located at nucleotides 6,176–6,262 and 11,502–11,583, respectively. In the latter case, Rep units m and n occupy the intergenic space between phnK and phnL. Since the open reading frame of phnD ends at nucleotide 6,172 and phnE starts at nucleotide 6,227, it is clear that the 36 base long Rep unit k is situated between phnD and phnE; however, the 34-base Rep unit l is overlapped with nucleotides 3-36 at the 5'-end of the phnE open reading frame. This is unusual since, to date, in agreement with the definition of Rep sequences, all reported Reps are located extragenically. Although the sequences of Rep units k, l, and n are different from the Rep consensus sequence, similar hairpin structures can be maintained in each case. It will be interesting to see what regulatory function these Rep units as well as this proposed intragenic Rep unit I may have in the expression of phn genes. Rep sequences have been demonstrated very recently to be binding sites for DNA gyrase and might possibly be the attachment site of chromosome to the membrane (Yang and Ames, 1988).

**Homology Searches on PhnA to PhnQ**—DNA sequences of each open reading frame were checked for similarity to each other. The homology was indicated by a dash or a colon. A conserved region is shown. The conserved region is repeated between nucleotides 6,176-6,262 and 11,502-11,583, respectively. In the latter case, Rep units m and n occupy the intergenic space between phnK and phnL. Since the open reading frame of phnD ends at nucleotide 6,172 and phnE starts at nucleotide 6,227, it is clear that the 36 base long Rep unit k is situated between phnD and phnE; however, the 34-base Rep unit l is overlapped with nucleotides 3-36 at the 5'-end of the phnE open reading frame. This is unusual since, to date, in agreement with the definition of Rep sequences, all reported Reps are located extragenically. Although the sequences of Rep units k, l, and n are different from the Rep consensus sequence, similar hairpin structures can be maintained in each case. It will be interesting to see what regulatory function these Rep units as well as this proposed intragenic Rep unit I may have in the expression of phn genes. Rep sequences have been demonstrated very recently to be binding sites for DNA gyrase and might possibly be the attachment site of chromosome to the membrane (Yang and Ames, 1988).

**Fig. 5. Rep sequences.** The numbers indicate the positions of Rep unit clusters in the entire DNA sequence, and the arrows indicate the direction of each Rep unit. The consensus sequence generated by 14 Rep units in the phn operon are shown. For each Rep unit, identical bases are indicated by dots, and the bases different from those of the consensus sequence are listed. Sequence different from the canonical Rep sequence (Stern et al., 1984) is shown in lower-case letters, and a dash represents an inserted gap. The dashed arrows below the Rep canonical sequence indicate the regions in each Rep unit capable of base-pairing. All sequences are written 5' to 3'.
other by the Needleman and Wunsch method (1970). Three genes (\textit{phnC}, \textit{phnK}, and \textit{phnL}) coding for polypeptide chains with molecular sizes of 29,495, 27,833, and 24,792 daltons, respectively, showed a similarity score above 43\% between each pair (\textit{phnC}-\textit{phnK}, 45.4\%; \textit{phnK}-\textit{phnL}, 43.2\%; and \textit{phnL}-\textit{phnC}, 44.8\%). Protein sequence comparisons demonstrated 23.5\% similarity between \textit{PhnC} and \textit{PhnK}, 24.8\% between \textit{PhnK} and \textit{PhnL}, and 25.9\% between the encoded \textit{PhnL} protein and \textit{PhnC}. All three protein sequences showed a high degree of conservation at two regions that have been characterized previously as nucleotide-binding sequences. A search for homology with protein sequences in the National Biomedical Research Foundation data base revealed strong similarities between these three proteins and several nucleotide-binding membrane components of bacterial periplasmic enzymes. Fig. 6 shows an alignment of \textit{PhnC}, \textit{PhnK}, and \textit{PhnL}, with \textit{HisP}, \textit{MalK}, and \textit{OppD}, components of the histidine, maltose, and oligopeptide permeases, respectively (Gibson et al., 1982; Higgins et al., 1982; Hiles et al., 1987). These permease proteins are members of a family of nucleotide-binding proteins that includes \textit{PstB}, a component of the phosphate permease (Surin et al., 1985), the nodulation protein \textit{NodL} of \textit{Rhizobium} (Evans and Downie, 1986), and the human multidrug resistance F-glycoprotein (Gros et al., 1986). These data indicate that \textit{PhnC}, \textit{PhnK}, and \textit{PhnL} may belong to this family of nucleotide-binding proteins.

The \textit{phnH} gene, which codes for a polypeptide chain of 194 residues with a molecular size of 21,029 daltons, showed 26\% identity to human sex steroid-binding protein, 25\% to \textit{E. coli} primase, 22.5\% to bovine liver cytochrome-P450 steroid 21-\alpha hydroxylase, and 21\% to human androgen-binding protein. The biological significance of the similarity to steroid-binding proteins and DNA primase remains to be seen. Such proteins are thought to have evolved a binding domain for small nucleotide-binding membrane components of bacterial periplasmic transport systems. Comparison of the dinucleotide-binding fold with the peptide sequences of these 22 ORFs demonstrated that besides \textit{PhnC}, \textit{PhnK}, and \textit{PhnL}, \textit{PhnN} has the motif A sequence at the N-terminal end of the protein; but no motif B-like sequence was found in \textit{PhnN}. Since motif B in the nucleotide-binding protein family shows less sequence conservation compared to hydrophobic conservation and the ArsA protein of the arsenic transport system contains only the motif A type of sequence but still has the ATPase activity (Chen et al., 1986; Rosen et al., 1988), \textit{PhnN} may also show ATP (or XTP) hydrolytic activities.

The above observations that \textit{PhnC}, \textit{PhnK}, \textit{PhnL}, and \textit{PhnN} are likely members of the family of nucleotide-binding protein components of binding protein-dependent permease systems strongly suggest that an alkylphosphonate-dependent permease system may be involved in the expression of the methylphosphonate utilization phenotype. The counterparts of other members of the binding protein-dependent permease systems such as ligand-binding protein and integral membrane proteins would then also be expected to be present in the \textit{Phn} system. Homology search with the sequence EAA---G----------I-LP found a segment of EAA---G----------I-RG located 101 residues proximal from the C terminus of \textit{PhnM} (Fig. 7a). Along with \textit{PhnE}, for which the hydrophy profile predicted by the Kyte-Doolittle algorithm (Kyte and Doolittle, 1982) showed a significant level of hydrophobicity with the chance of forming at least six transmembrane helices (data not shown), both \textit{PhnE} and \textit{PhnM} may be membrane components of a \textit{Phn} transport system.

Search for homology with the helix-turn-helix motif indicated that \textit{PhnL}, \textit{PhnO}, and ORF 269 have the sequence pattern of A---G----------I/V (Fig. 7b). Further evaluation for each 22-residue segment according to the procedure of Brennan and Matthews (1989) generated scores between 0.81 and 0.85, suggesting that these proteins are at least candidates for DNA binding studies.

\textit{T7 Promoter-driven phn Gene Expression}—To determine whether the ORFs encoded by the 15.6-kb fragment are translated, the \textit{T7} RNA polymerase-dependent \textit{in vivo} transcription/translation system was used. Transformants with plasmids \textit{pCMC1} to \textit{pCMC9} and \textit{pIB}131 were labeled with \textsuperscript{35}S methionine according to the protocol of Tabor and Richardson (1985), and the labeled polypeptides were analyzed by 15\% SDS-polyacrylamide gel electrophoresis, as shown in Fig. 8. Both \textit{pIB}130 and \textit{pIB}131 direct expression of four proteins of between 26 and 30 kDa (Fig. 8, lane 1) which probably correspond to the precursor and mature forms of \textit{\beta}-lactamase encoded on these plasmids. Two additional proteins of 84 and 26 kDa are seen for \textit{pCMC1}-transformed cells (lane 8). The

![Fig. 6. Sequence homology among \textit{PhnC}, \textit{PhnK}, \textit{PhnL}, and \textit{PhnN} and nucleotide-binding proteins.](image-url)
brane proteins of binding protein-dependent permease systems. Con-}
proteins are between 0.65 and 0.80. 

inferred. Scores for the known DNA-binding 
with the identical residues in boldface. Calculated scores for these 22-}
base DNA (Brennan and Matthews, 1989) of DNA-binding proteins are shown 
b, potential helix-turn-helix motif of Phn proteins. The 22-residue 
standards are shown in kilodaltons. Peptide bands not seen in the 
lane 9, pCMC7; lane 5, pCMC6; lane 6, pQY1; lane 7, pCMC9; lane 8, pCMC1; 
match the general picture expected (Table II) for the phn 
gene products. Detailed studies were performed with con-
structs containing the 4.2-kb (pCMC4; lane 3), 3.5-kb 
(pCMC5; lane 4), and 8-kb (pCMC8; lane 5) inserts. A protein of 17 kDa is seen for pCMC4-containing cells that appears to 
match the calculated molecular sizes of PhnG to PhnQ. Fur-
there cloning of the pCMC6 insert with Sal1 digestion creates 
two plasmids (pQY1 and pQY2) that contain nucleotides 
3.5-kb (pCMC8; lane 10) nor the 8-kb (pCMC9; 
lanes 11) insert generates labeled protein other than p-lacta-
mases. The identity of the 26-kDa band is unclear since it was 
not seen in cells containing plasmids pCMC7 to pCMC9; 
a band with a slower mobility was found in pCMC7-trans-
formed cells (lane 9).

When the translation products were examined for the strand encoding the phnA to phnQ ORFs, more than 12 
radioisotope-labeled bands were seen on the autoradiograph produced by pCMC2-containing cells (Fig. 8, lane 2). The molecular sizes of these bands range from 14 to 45 kDa, which 
matches the general picture expected (Table II) for the phn 
gene products. Detailed studies were performed with con-
structs containing the 4.2-kb (pCMC4; lane 3), 3.5-kb 
(pCMC5; lane 4), and 8-kb (pCMC8; lane 5) inserts. A protein of 17 kDa is seen for pCMC4-containing cells that appears to 
match the calculated molecular sizes of PhnG to PhnQ. Fur-
there cloning of the pCMC6 insert with Sal1 digestion creates 
two plasmids (pQY1 and pQY2) that contain nucleotides 
3.5-kb (pCMC8; lane 10) nor the 8-kb (pCMC9; 
lanes 11) insert generates labeled protein other than p-lacta-
mases. The identity of the 26-kDa band is unclear since it was 
not seen in cells containing plasmids pCMC7 to pCMC9; 
a band with a slower mobility was found in pCMC7-trans-
formed cells (lane 9).

When the translation products were examined for the strand encoding the phnA to phnQ ORFs, more than 12 
radioisotope-labeled bands were seen on the autoradiograph produced by pCMC2-containing cells (Fig. 8, lane 2). The molecular sizes of these bands range from 14 to 45 kDa, which 
matches the general picture expected (Table II) for the phn 
gene products. Detailed studies were performed with con-
structs containing the 4.2-kb (pCMC4; lane 3), 3.5-kb 
(pCMC5; lane 4), and 8-kb (pCMC8; lane 5) inserts. A protein of 17 kDa is seen for pCMC4-containing cells that appears to 
match the calculated molecular sizes of PhnG to PhnQ. Fur-
there cloning of the pCMC6 insert with Sal1 digestion creates 
two plasmids (pQY1 and pQY2) that contain nucleotides 
3.5-kb (pCMC8; lane 10) nor the 8-kb (pCMC9; 
lanes 11) insert generates labeled protein other than p-lacta-
mases. The identity of the 26-kDa band is unclear since it was 
not seen in cells containing plasmids pCMC7 to pCMC9; 
a band with a slower mobility was found in pCMC7-trans-
formed cells (lane 9).

When the translation products were examined for the strand encoding the phnA to phnQ ORFs, more than 12 
radioisotope-labeled bands were seen on the autoradiograph produced by pCMC2-containing cells (Fig. 8, lane 2). The molecular sizes of these bands range from 14 to 45 kDa, which 
matches the general picture expected (Table II) for the phn 
gene products. Detailed studies were performed with con-
structs containing the 4.2-kb (pCMC4; lane 3), 3.5-kb 
(pCMC5; lane 4), and 8-kb (pCMC8; lane 5) inserts. A protein of 17 kDa is seen for pCMC4-containing cells that appears to 
match the calculated molecular sizes of PhnG to PhnQ. Fur-
there cloning of the pCMC6 insert with Sal1 digestion creates 
two plasmids (pQY1 and pQY2) that contain nucleotides 
3.5-kb (pCMC8; lane 10) nor the 8-kb (pCMC9; 
lanes 11) insert generates labeled protein other than p-lacta-
mases. The identity of the 26-kDa band is unclear since it was 
not seen in cells containing plasmids pCMC7 to pCMC9; 
a band with a slower mobility was found in pCMC7-trans-
formed cells (lane 9).

When the translation products were examined for the strand encoding the phnA to phnQ ORFs, more than 12 
radioisotope-labeled bands were seen on the autoradiograph produced by pCMC2-containing cells (Fig. 8, lane 2). The molecular sizes of these bands range from 14 to 45 kDa, which 
matches the general picture expected (Table II) for the phn 
gene products. Detailed studies were performed with con-
structs containing the 4.2-kb (pCMC4; lane 3), 3.5-kb 
(pCMC5; lane 4), and 8-kb (pCMC8; lane 5) inserts. A protein of 17 kDa is seen for pCMC4-containing cells that appears to 
match the calculated molecular sizes of PhnG to PhnQ. Fur-
there cloning of the pCMC6 insert with Sal1 digestion creates 
two plasmids (pQY1 and pQY2) that contain nucleotides 
3.5-kb (pCMC8; lane 10) nor the 8-kb (pCMC9; 
lanes 11) insert generates labeled protein other than p-lacta-
mases. The identity of the 26-kDa band is unclear since it was 
not seen in cells containing plasmids pCMC7 to pCMC9; 
a band with a slower mobility was found in pCMC7-trans-
formed cells (lane 9).

When the translation products were examined for the strand encoding the phnA to phnQ ORFs, more than 12 
radioisotope-labeled bands were seen on the autoradiograph produced by pCMC2-containing cells (Fig. 8, lane 2). The molecular sizes of these bands range from 14 to 45 kDa, which 
matches the general picture expected (Table II) for the phn 
gene products. Detailed studies were performed with con-
structs containing the 4.2-kb (pCMC4; lane 3), 3.5-kb 
(pCMC5; lane 4), and 8-kb (pCMC8; lane 5) inserts. A protein of 17 kDa is seen for pCMC4-containing cells that appears to 
match the calculated molecular sizes of PhnG to PhnQ. Fur-
there cloning of the pCMC6 insert with Sal1 digestion creates 
two plasmids (pQY1 and pQY2) that contain nucleotides 
3.5-kb (pCMC8; lane 10) nor the 8-kb (pCMC9; 
lanes 11) insert generates labeled protein other than p-lacta-
mases. The identity of the 26-kDa band is unclear since it was 
not seen in cells containing plasmids pCMC7 to pCMC9; 
a band with a slower mobility was found in pCMC7-trans-
formed cells (lane 9).

When the translation products were examined for the strand encoding the phnA to phnQ ORFs, more than 12 
radioisotope-labeled bands were seen on the autoradiograph produced by pCMC2-containing cells (Fig. 8, lane 2). The molecular sizes of these bands range from 14 to 45 kDa, which 
matches the general picture expected (Table II) for the phn 
gene products. Detailed studies were performed with con-
structs containing the 4.2-kb (pCMC4; lane 3), 3.5-kb 
(pCMC5; lane 4), and 8-kb (pCMC8; lane 5) inserts. A protein of 17 kDa is seen for pCMC4-containing cells that appears to 
match the calculated molecular sizes of PhnG to PhnQ. Fur-
there cloning of the pCMC6 insert with Sal1 digestion creates 
two plasmids (pQY1 and pQY2) that contain nucleotides 
3.5-kb (pCMC8; lane 10) nor the 8-kb (pCMC9; 
lanes 11) insert generates labeled protein other than p-lacta-
mases. The identity of the 26-kDa band is unclear since it was 
not seen in cells containing plasmids pCMC7 to pCMC9; 
a band with a slower mobility was found in pCMC7-trans-
formed cells (lane 9).
in size of some of the expected proteins as well as the large number of encoded polypeptides, it is difficult to assign unambiguously each polypeptide to a phn gene.

**DISCUSSION**

We report here the DNA sequence for *E. coli* B phn (psiD) genes which permit cells to grow on alkylphosphonates (e.g. methyl-, ethyl-, and 2-aminoethylphosphonate) as a sole phosphorus source. This growth ability as assessed with methylphosphonate leads to quantitative production of methane (analyzed by gas chromatography) and P, used for growth and so involves a C-P lyase enzymatic activity. This enzyme activity is not understood mechanistically since methylphosphonate can be boiled in strong acid and strong base without decomposition. The genetic ability of microbes to cleave these very stable C-P bonds is not only of relevance in understanding how naturally occurring C-P bonds in phosphonolipids are catalyzed, but also in the analysis of how widely used herbicides such as N-phosphonomethylglycine (glyphosate) are degraded in soil since C-P lyase activity is also involved for those enzymatic degradations (Shinabarger and Braymer, 1986; Kishore and Jacob, 1987; Wackett et al., 1987a).

Although a C-P lyase activity in crude extracts of *Entero bacter aerogenes* was reported and has been purified recently by Murata et al. (1988, 1989), it may be distinct since it is a soluble enzyme complex, and no report of the organic product has been made. The in vitro activity of the *E. coli* enzyme has not been detected (by us or others). Either very labile components (e.g. oxidant-sensitive species or components requiring an intact membrane) or proper submit interaction may be involved. To address this number and nature of genes and gene products required, we have turned to cloning and characterization of the phn genes as reported here.

One significant finding is the very large size of the DNA fragment needed to complement Phn⁺ phenotypes. The observation that a 15.6-kb complementing DNA fragment could substitute the phnA to phnQ operon (Fig. 3 and Table II). Initial searches for promoters and a phosphate box indicate a single candidate for a psi-regulatable phn promoter just upstream of expression studies showed gene products for phnB to phnQ. Candidates for an operon organization. T7 RNA polymerase sequences is the largest number yet reported. Particularly array of 10 Rep sequences (Stertn et al., 1984). The head-to-head Rep arrangement is characteristic, but the tandem array of 10 Rep sequences is the largest number yet reported. Particularly interesting is the interruption of the phnE reading frame by use of the Rep unit (Fig. 4B). The RNA structure prediction (Zuker and Striegler, 1981) of the region that includes Rep units k and l suggests that a hairpin structure (∆G° = −57.6 kcal/mol) can be formed with portions of the predicted Shine-Dalgarno sequence of the phnE reading frame hydrogen-bonding to its initiation codon; thus, interference of ribosome binding and subsequent translation are expected. Translational regulation of this kind has been proposed for the *ileH* and *uraB* genes (Nazos et al., 1986; San Francisco et al., 1988). Other hypotheses about Rep functions including translational regulation have been advanced. The recent observation of DNA gyrase recognition seems cogent, and this multiple Rep sequence array could be high affinity gyrase sites (Yang and Ames, 1988). The physiological consequences of interposition of such a dense array of Rep sequences between the phnA and phnB genes remains to be determined, as does the role of the paired Rep sequences between phnD and phnE, and phnK and phnL.

The discovery of up to 17 phn genes and encoded PhnA to PhnQ proteins reinforces our starting assumption that the ability of *E. coli* to grow on alkylphosphonates as phosphate source is a metabolically complex process with multiple gene involvement and under some type of Pho regulon control. The functional identification of so many protein components (and which ones are required for the Phn⁺ phenotype) is a bit daunting at the outset, and homology searches are only the preamble to subcloning, expression, and analysis by cellular localization and functional assay of specific protein components.

In work to be detailed elsewhere, subcloning of phn genes has yielded a 6.1-kb fragment, encompassing phnG to phnN (nucleotides 7,783-13,919), which permits cells to convert CH₃PO₂⁻ to CH₃ even though they grow only poorly or not at all on methylphosphonate. Some (or all) of these genes are C-P lyase candidates. Nonetheless, protein homology searches suggest that PhnC as well as PhnK and PhnL (and maybe PhnN also) show structural homologies to each other as well as to other bacterial proteins which share domains thought to encode binding domains for dinucleotides (such as ATP) and are localized to the bacterial cytoplasmic membrane. Thus, OppD, HisP, and PstB are parts of multicomponent permease or membrane transport systems (Ames, 1986). One can speculate that PhnC, PhnK, and PhnL (also PhnN) may have analogous roles in the uptake of alkylphosphonates, which may indicate that some of the other phn gene products are also elements in multicomponent transport systems (that could involve both periplasmic and intrinsic membrane components), as found in the oligopeptide permease, histidine permease, and, most intriguingly, the inorganic phosphate-specific transport systems (Ames, 1986; Higgins et al., 1982; Hiles et al., 1987; Suria et al., 1985). In this connection, PhnE shows high hydrophobicity content by Kyte-Doolittle analysis (Kyte-Doolittle, 1982), and PhnM has sequences similar to those found in other integral membrane proteins, making them reasonable initial candidates for intrinsic membrane components. So PhnK, PhnL, PhnM, and PhnN may all be the protein components for a phosphatase-specific uptake system, as in histidine, maltose, and phosphate-specific transport systems (Figs. 6 and 7). In turn, PhnC, PhnK, PhnL, and PhnN are candidates for overproduction and purification to assess nucleoside diphosphate and triphosphate binding and/or any nucleoside-triphosphate hydrolase activities.

In the search for regulatory protein(s) of the proposed phn operon, both PhnL and PhnO may have potential helix-turn-helix sequences in analogy to DNA-binding proteins. Judging from the fact that PhnL shows extensive homologies to nu-

---

5 G. F. Ames, personal communication.

4 Q.-Z. Ye et al., manuscript in preparation.
cleotide-binding proteins, it may be that PhnO is a regulatory protein of the phn operon.

In addition to anticipated protein components involved in alkylphosphonate uptake and membrane transport, one or more of the phn genes are expected to encode C–P lyase enzymatic activity. As noted above, this is a very difficult chemical cleavage step, and it would not be surprising if a multisubunit or multicomponent enzyme complex were required to yield alkane and P. In particular, one anticipates that acid/base enzyme chemistry will not suffice, and recent model chemistry is consistent with redox chemical involvement (Frost et al., 1987; Shames et al., 1987). We suspect it most likely that bacterial C–P lyase activity uses reduct ox chemistry on the bound organophosphonate, either by net one-electron reduction or by one-electron oxidation at phosphorus, transiently to facilitate C–P cleavage. Redox interconversions at phosphorus are exceedingly rare in biology; and thus, few mechanistic or protein structural precedents are available to guide expectation. It may be then that C–P lyase activity will require redox centers (inorganic or organic) in the enzyme itself and/or in reduct protein input and acceptor components and that an intact membrane capable of establishing a chemiosmotic potential coupled to electron flow may be required and explain the difficulty in detecting enzyme activity in broken cell extracts. Subsequent studies are directed to examine if any of the phn gene products have these characteristics and to determine the minimal number of components needed to reconstitute a functional C–P lyase activity, with initial focus on the eight plusO to plusN genes.

Acknowledgments—We wish to thank Dr. S. Tabor for the gifts of Sequenase enzymes. J. A. Boline and W. Metcalf also contributed to studies on phosphonate assimilation at Purdue University.

REFERENCES

Wanner, B. L. (1987) J. Bacterial. 169, 2026–2030
Molecular biology of carbon-phosphorus bond cleavage. Cloning and sequencing of the phn (psiD) genes involved in alkylphosphonate uptake and C-P lyase activity in Escherichia coli B.

C M Chen, Q Z Ye, Z M Zhu, B L Wanner and C T Walsh


Access the most updated version of this article at [http://www.jbc.org/content/265/8/4461](http://www.jbc.org/content/265/8/4461).

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at [http://www.jbc.org/content/265/8/4461.full.html#ref-list-1](http://www.jbc.org/content/265/8/4461.full.html#ref-list-1)