JOURNAL OF BACTERIOLOGY, Nov. 1988, p. 5110–5116
0021-9193/88/115110-07$02.00/0
Copyright © 1988, American Society for Microbiology
Vol. 170, No. 11

Membrane-Bound Phosphatases in *Escherichia coli*: Sequence of the *pgpA* Gene

TATEO ICHO†

Department of Biochemistry, College of Agricultural and Life Science, University of Wisconsin, Madison, Wisconsin 53706

Received 10 March 1988/Accepted 30 July 1988

One of the phosphatidyl glycerophosphate phosphatase genes of *Escherichia coli*, *pgpA*, was cloned, and its DNA sequence was determined. Its 507-base-pair open reading frame was consistent with the 18,000-molecular-weight product identified by a maxicell experiment. Between its possible promoter and methionine initiation codon, a repetitive extragenic palindromic sequence was found.

In *Escherichia coli*, there are three phospholipid phosphatase activities. As illustrated in Fig. 1, they dephosphorylate phosphatidyl glycerophosphate (PGP), phosphatidic acid (PA), and lysophosphatidic acid (LPA) (26). Previously, PA and LPA phosphatase were considered different from PGP phosphatase (4, 6). PGP phosphatase mutants were isolated, and two genes, *pgpA* and *pgpB*, were mapped at 18 and 28 min, respectively, on the *E. coli* chromosome (18). The former is responsible for a PGP-specific A phosphatase; the latter codes for a B phosphatase, which attacks at least two more substrates, PA and LPA. This paper describes the molecular cloning and DNA sequencing of the PGP A phosphatase. The study of the PGP B phosphatase appears in the accompanying paper (17).

Colonies of PGP phosphatase mutants accumulating [^{32}P]PGP were detected as characteristic dark spots on an X-ray film in an in vitro colony autoradiography assay of PGP phosphatase (18). The *pgpA* gene was cloned by using the same screening method. Colonies of a *pgpA* mutant carrying the plasmid with the wild-type *pgpA* allele appeared as light spots among the dark PGP A phosphatase-negative colonies. Starting from the isolation of an F′ factor carrying the 18-min region, the *pgpA* gene was located within a 704-base-pair (bp) *Eco*RV DNA fragment, whose sequence is presented in this paper. A maxicell experiment revealed a protein of 18 kilodaltons (kDa) specific for the *pgpA* gene.

MATERIALS AND METHODS

Bacteria and plasmids. Bacterial strains used are summarized in Table 1. Plasmids pBR322 (3), pACYC184 (6), and pTI10 were used as vectors for subcloning the *pgpA* gene. Plasmid pLL10 (28) was a derivative of pBR322 in whose *Bam*HI site a synthetic adapter (*Bam*HI-*Sma*I-*Eco*RI-*Sma*I-*Bam*HI) of 20 bp was inserted. Plasmid pTI10 was constructed by removing the *Sma*I-*Eco*RI-*Sma*I portion of the adapter fragment from pLL10. Consequently, a DNA fragment inserted into the *Sma*I site of pTI10 could be recovered by a single *Bam*HI digestion. Plasmid pBZ201 (*lon*) was a gift from B. A. Zehnbaur (34).

Materials. Authentic LPA was purchased from Serdary Research Laboratories, Ontario, Canada. Phospholipase A2 from *Naja naja* venom and phospholipase D from cabbage were from Sigma Chemical Co., St. Louis, Mo. Restriction endonucleases, nuclease *Bal*31, and DNA polymerase Klenow fragment were either from Bethesda Research Laboratories, Inc., Gaithersburg, Md., or New England Biolabs Inc., Beverly, Mass. Alkaline phosphatase from calf intestine, lyophilized, was from Boehringer Mannheim Biochemicals, Indianapolis, Ind. T4 polynucleotide kinase was from Pharmacia PL-Biochemicals Inc., Piscataway, N.J. [γ-^{32}P]ATP was from Amersham Co., Arlington Heights, Ill. ^{32}P$_i$, [^{35}S]methionine, L-α-[dipalmitoyl-1-^{14}C]phosphatidylcholine, and T4 DNA ligase were from New England Nuclear Corp., Boston, Mass.

Preparation of phosphatase substrates. The preparation procedure for [^{32}P]PGP from *sn*-glycero-3-[^{32}P]phosphate and CDP diglyceride by PGP synthetase and that of [^{32}P]PA from ^{32}P$_i$-labeled phospholipids by phospholipase D was described previously (18). For [^{32}P]LPA preparation, [^{32}P] phosphatidylglycerol and [^{32}P]phosphatidylethanolamine were fractionated from ^{32}P$_i$-labeled total *Escherichia coli* phospholipids by DEAE-cellulose chromatography (18) and digested with phospholipase A2 (32), followed by phospholipase D (25) and DEAE-cellulose chromatography. [^{14}C] LPA was made from [^{14}C]phosphatidylcholine by the same method. [^{14}C]PA was prepared from [^{14}C]phosphatidylcholine by phospholipase D digestion, followed by DEAE-cellulose chromatography as previously described for [^{32}P] PA purification (18).

Assays for membrane-bound phosphatases. The ^{32}P$_i$ release method for PGP phosphatase (method 1) and PA phosphatase (method 3), which measure trichloroacetic acid (TCA)-soluble ^{32}P counts in the water phase after acid chloroform extraction, were described previously (18). The ^{32}P$_i$ release method for LPA phosphatase was the same as that for PGP phosphatase (method 1) except that 100 μM [^{32}P]PGP was replaced by 100 μM [^{32}P]LPA.

The phosphatases were also assayed by separating products on thin-layer chromatography (TLC).

For PGP phosphatase, the reaction conditions were the same as those described for method 1 (18). After the reaction, a 2-μl portion of the mixture was immediately withdrawn and spotted on a cellulose chromatogram sheet (Eastman Kodak). [^{32}P]PGP and ^{32}P$_i$ were separated by using the solvent system 1 M ammonium acetate-ethanol (2:3). From the ratio of radioactivity, P$_i$/PGP, PGP phosphatase activity was calculated.

For LPA and PA phosphatase, the reaction conditions were the same as for method 3 (18) for LPA phosphatase and PA phosphatase except that [^{14}C]LPA and [^{14}C]PA were

† Present address: Laboratory of Genetics, Department of Biology, The University of Tokyo, Hongo, Tokyo 113, Japan.

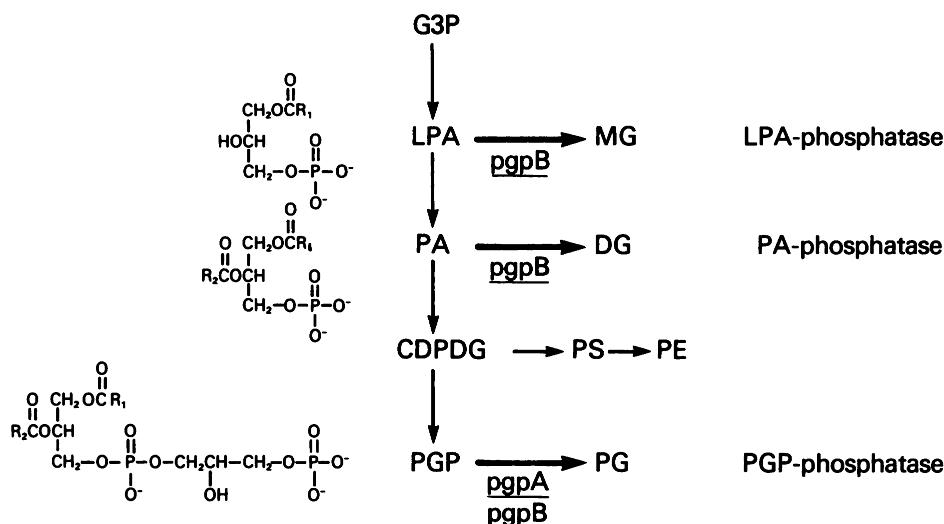

FIG. 1. Membrane-bound phosphatases that attack phospholipid molecules in *E. coli*. The reactions catalyzed by three phospholipid phosphatases are indicated by the thick arrows. Evidence for this scheme is discussed in references 18 and 26. G3P, *sn*-glycerol-3-phosphate; MG, monoglyceride; DG, diglyceride; CDPDG, CDP diglyceride; PS, phosphatidylserine; PG, phosphatidylglycerol; PE, phosphatidylethanolamine; R_1 and R_2, fatty acids.

replaced by [^{32}P]LPA and [^{32}P]PA, respectively. The reactions were terminated by adding 2.75 volumes of methanol (0.01 N HCl)–chloroform (2:1). After the extraction, neutral lipids were analyzed by TLC (27). LPA and PA phosphatase activities were calculated from the radioactive ratio of monoglyceride/total lipid and diglyceride/total lipid, respectively.

TABLE 1. Bacterial strains[a]

Strain	Relevant markers	Source or reference
KL16-99	Hfr KL16 *recA*	CGSC[b]
W3747	F13 (*argF*$^+$ *lac*$^+$ *tsx*$^+$ *purE*$^+$) *metB1 relA1*	CGSC
Hfr Boda8	Hfr *metB1 relA1*	CGSC
CSR603	F$^-$ *phr-1 recA1 uvrA6 gyrA98 rpsL31*	CGSC (9)
X478	F$^-$ *leuB6 proC32 purE42 trpE38 lysA23 metE70 rpsL109*	CGSC
R477	*rpsL-136 eda his-4 leu-6 thr-1*	J. Adler
E26	*pgpA2 pgpB26* derivative from strain 8	18
TI66	*proC32 pgpA2 pgpB26 recA* (*dnaZ*$^+$ *recA* derivative from TI63)	18
TI69	*proC32 purE42 recA* (*recA* derivative from X748)	
TI70	*phoA8 proC*$^+$ transductant of X478 (strain 8 donor)	18
TI71	*phoA8 pgpA2 proC*$^+$ transductant of X478 (E26 donor)	18
TI74	*phoA8 pgpB26 trp*$^+$ transductant of X478	
TI75	*phoA8 pgpA2 pgpB26 trp*$^+$ transductant of TI71	
TI80	*phoA8 recA* (*recA* derivative from TI70)	
TI83	*phoA8 pgpA2 pgpB26 recA* (*recA* derivative from TI75)	

[a] *pgpA* and *pgpB* mutations were introduced from strain E26 by P1 transductions. The *phoA* mutation was introduced from either strain 8 or E28 by selecting *proC*$^+$ transductants. All *recA* mutations were introduced by a cross with KL16-99, followed by the selection of *lysA*$^+$ *rpsL* recombinants.

[b] CGSC, *E. coli* Genetic Stock Center, Yale University, New Haven, Conn.

Colony autoradiography. The colony autoradiography to detect *pgpA*$^+$ clones was performed by using the A condition used for *pgpA* mutant isolation (18). To detect light spots among dark background colonies on X-ray film, the number of colonies per plate was kept under 100. Transformant colonies were transferred from a plate to a filter paper and lysed with the lysozyme EDTA. The filter paper was then dried by the air from a fan and put into a reaction mixture containing *sn*-glycerol-3-[^{32}P]phosphate and CDP diglyceride. The reaction was terminated by adding 20% TCA with 1 mM cold *sn*-glycerol-3-phosphate. The inclusion of *sn*-glycerol-3-phosphate made the background cleaner. [^{32}P]PGP accumulation in *pgpA* mutants was detected as a dark spot on the X-ray film. Colonies of *pgpA*$^+$ clones, which have high PGP A phosphatase activity, hydrolyzing [^{32}P]PGP to P_i and phosphatidylglycerol, appeared as bright spots on X-ray film. Either strain TI83 (*pgpA pgpB*) or strain TI81 (*pgpA pgpB*$^+$) was suitable for this assay.

DNA cloning and sequencing techniques. The DNA isolation and cloning techniques, including nuclease *Bal*31 digestion and the filling-in reaction by the DNA polymerase Klenow fragment, were performed according to published methods (10, 23). The DNA of F′ factor was isolated from cleared lysates (9). Plasmids were isolated by the sodium dodecyl sulfate (SDS) alkaline method (2). DNA fragments were isolated on either low-melting-point agarose or DEAE-cellulose paper (13). DNA sequencing was performed by the method of Maxam and Gilbert (24). The maxicell radiolabeling was performed by the protocol of Sancer et al. (29).

Computer programs. DNA and polypeptide sequences were analyzed by using the program library of the University of Wisconsin Genetics Computer Group (UWGCG) (11). For the prediction of mRNA secondary structure, the UWGCG version of program FOLD by Zuker and Stiegler (35), which was modified to use the energy file of Cech et al. (5), was used. The protein structure prediction program of Garnier et al. (14) was adapted to the VAX-VMS operating system by Michael Gribskov.

RESULTS

Comparison of PA and LPA phosphatase activities in PGP phosphatase-deficient mutants. Comparing PGP and PA phosphatase activities among PGP phosphatase-deficient mutants, it was reported that the *pgpA* gene was responsible specifically for a PGP phosphatase activity, whereas the *pgpB* gene encoded protein with both a PGP phosphatase and a PA phosphatase activity (18). Table 2 shows that the *pgpB* gene was also responsible for LPA phosphatase activity. The upper part of the table shows the results with the conventional $^{32}P_i$ release assays, while the lower part shows the results calculated from the separated radioactive counts of the reaction products identified on TLC plates, $^{32}P_i$ for PGP phosphatase, [^{14}C]diglyceride for PA phosphatase, and [^{14}C]monoglyceride for LPA phosphatase. The PGP and PA phosphatase activities obtained by these two methods were similar, confirming that real phosphatase activities are measured by the ^{32}P release assays. However, in the LPA phosphatase assay, the existence of lysophospholipase activity, which hydrolyzes LPA, generating *sn*-glycerol-3-[^{32}P]phosphate (12), must be considered. In the first assay, *sn*-glycerol-3-[^{32}P]phosphate can be extracted into the water phase, increasing apparent phosphatase activity. The LPA phosphatase activity determined by the second method is therefore more accurate.

There was about a 30% decrease in the PGP phosphatase activity in the *pgpA* mutant, whereas no significant difference was observed in PA and LPA phosphatase activities. On the other hand, in the *pgpB* mutant, all three phosphatase activities were decreased. PA phosphatase activity was reduced almost to the background level. The other two were reduced to half of their initial values. In the *pgpA pgpB* double mutant, PGP phosphatase activity dropped to its background levels, as did PA phosphatase activity. LPA phosphatase activity remained at the same intermediate level as that in the *pgpB* strain.

These results suggested the existence of three distinct phospholipid phosphatases in *E. coli*. One is designated PGP A phosphatase, which specifically hydrolyzes PGP. The second is PGP B phosphatase, which hydrolyzes PGP, PA, and LPA. The third is an LPA-specific phosphatase whose gene has not yet been identified. The following experiments proved that this assumption was correct for the *pgpA* gene product. The accompanying paper (18) clarifies the apparently complicated specificity of the *pgpB* gene product.

Isolation of F′ factors carrying the *pgpA* gene. *pgpA* mutations had been mapped to min 10, between *proC* and *dnaZ*, by P1 transduction (18). To clone the *pgpA* gene, several ColE1 *E. coli* hybrid plasmids of Clarke and Carbon (8) carrying the *dnaZ* or *acrA* gene, and one plasmid, pBZ201 (35), carrying the *lon* gene, were surveyed. However, none of the strains carrying these plasmids showed an increased level of PGP phosphatase. Therefore, smaller F′ factors carrying the *pgpA* gene were isolated by the method of Low (22). Hfr Boda8 was mated with TI69 (*proC purE rpsL recA*), and $proC^+$ streptomycin-resistant colonies were selected. Among the 500 clones tested, 15 were $proC^+$ *purE*. Seven of them conferred the $proC^+$ phenotype to TI66 (*proC pgpA pgpB recA*) at a high frequency. The in vitro PGP phosphatase assay of these F′ ductants indicated that two of them, TIF33 and TIF34, were carrying complementing activity for the PGP A phosphatase.

Subcloning of the *pgpA* gene. The cloning process is summarized in Fig. 2. About 2 μg of TIF33 DNA was recovered after repeated CsCl-ethidium bromide centrifugation from a cleared lysate of a 600-ml overnight LB culture of TI80 (TIF33). The F′ DNA was digested with *Hin*dIII, ligated with an *Hin*dIII digest of pACYC184 DNA, transformed into TI76 (*pgpA pgpB*), and plated on L agar containing 25 μg of chloramphenicol per ml. Among 1,000 colonies screened by the colony autoradiography assay, 3, pTIA1, pTIA2, and pTIA3, appeared as light spots on X-ray film (Fig. 3) and produced a high amount of PGP phosphatase activity in vitro. The plasmids pTIA1 and pTIA2 had the same *Hin*dIII insert of 19.6 kilobases (kb). The plasmid pTIA3 had an additional 2.1-kb *Hin*dIII fragment. Starting from plasmid pTIA1, the *pgpA* gene was subcloned down to a 704-bp *Eco*RV fragment in plasmid pTIA26.

The PGP phosphatase activity of the strains carrying the plasmids that were generated in the process of the subcloning are summarized in Table 3. The phosphatase activity increased when the plasmid became smaller, probably depending on the copy number of each plasmid. No stimulation of PA and LPA phosphatase activities by these plasmids was observed in these strains. These results indicated that the *pgpA* gene is the structural gene for the PGP-specific A phosphatase. To estimate the molecular weight of the *pgpA* gene product, a maxicell experiment was performed. Plasmids carrying the *pgpA* gene were introduced into strain CSR603, and [^{35}S]methionine-labeled products were analyzed by SDS-polyacrylamide gel electrophoresis. As shown in Fig. 4, a protein band of about 18,000 kDa unique to these plasmids was detected.

Generation of *Bal*31 deletion plasmids. For the precise definition of the *pgpA* gene, the *Eco*RV fragment isolated from plasmid pTIA26 was treated with nuclease *Bal*31 and filled in by DNA polymerase Klenow fragment, and the deleted fragments were inserted into the *Sma*I site of plasmid pTI10. Transformants having various sizes of inserts were isolated and their sonicates were assayed for PGP phosphatase activity. The result indicated that the minimum length of the *pgpA* gene was about 650 bp. Some of these plasmids whose ends were sequenced are summarized in Table 4. The strains carrying plasmids pTIA128 and pTIA155 showed reduced phosphatase activity, and the strain carrying plasmid pTIA156 showed no stimulation of phosphatase production. Therefore, these plasmids seems to have lost a part of the *pgpA* gene. Since the extent of PGP

TABLE 2. Comparison of membrane-bound phosphatase activities by two methods[a]

Strain	Phosphatase activity (nmol/min per mg of protein)		
	PGP	PA	LPA
A			
TI70	5.5 (0.9)	0.28 (0.04)	3.9 (0.9)
TI71 (*pgpA*)	4.9 (0.4)	0.27 (0.03)	3.9 (0.6)
TI74 (*pgpB*)	2.1 (0.7)	0.02 (0.00)	2.1 (0.0)
TI75 (*pgpA pgpB*)	0.9 (0.3)	0.02 (0.00)	2.2 (0.3)
B			
TI70	4.5	0.37	3.0
TI71	2.7	0.38	2.8
TI74	2.3	0.01	1.1
TI75	0.5	0.01	1.1

[a] Three phospholipid phosphatase activities of *pgp* mutants and the parent were compared by the $^{32}P_i$ release assay (part A) and the separated-products assay (part B) as described in Materials and Methods. In part A, the average values of three independent experiments are shown with the standard deviation in parentheses. In this particular set of experiments, the PGP phosphatase activities assayed by the first method had a higher background than normal (18).

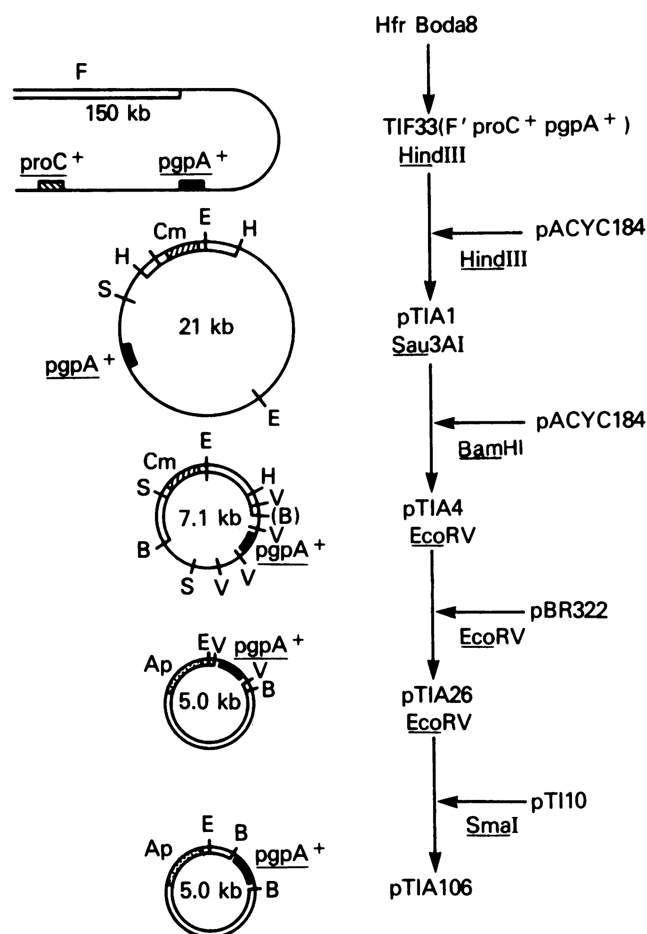

FIG. 2. Cloning of the *pgpA* gene. Plasmid pTIA1 DNA was partially digested with *Sau*3A and inserted into the *Bam*HI site of pACYC184 DNA. Transformants were screened for PGP phosphatase activity by the colony autoradiography assay. Plasmids pTIA4 and pTIA6 were among them; their DNA insert lengths were 2.8 and 5.2 kb, respectively. By inserting the 704-bp *Eco*RV fragment including the *pgpA* gene into the *Eco*RV site of pBR322, plasmid pTIA26 was obtained. Plasmid pTIA106, which was used for the sequencing, was constructed by inserting this *Eco*RV fragment into the *Sma*I site of pTI10. The portion of the *pgpA* gene and the vector in each plasmid are shown by the thick and double lines, respectively. H, *Hin*dIII; E, *Eco*RI; B, *Bam*HI; S, *Sph*I; V, *Eco*RV.

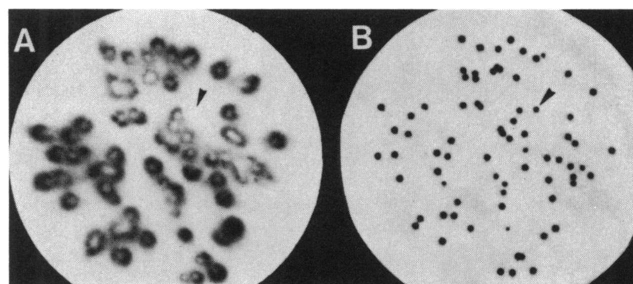

FIG. 3. Colony autoradiography to detect a $pgpA^+$ clone. Chloramphenicol-resistant transformant colonies of strain TI76 (*pgpA pgpB recA*) were transferred to filter paper, lysed, and incubated with a reaction mixture containing *sn*-glycerol-3-[^{32}P]phosphate and CDP diglyceride. The accumulation of [^{32}P]PGP was detected by autoradiography on X-ray film (A). The same filter paper was stained by Coomassie brilliant blue G (B). The arrows point to the initially isolated colony of strain TI83(pTIA1), which did not accumulate [^{32}P]PGP.

phosphatase activity was similar among the full-size clones, the effect of the fragment orientation on activity seemed to be negligible.

DNA sequencing. Figure 5 summarizes the fragments sequenced together with the restriction sites used. They cover the entire sequence of the *Eco*RV fragment on both strands. The open reading frame and the possible translational start for the *pgpA* gene are shown by a thick arrow at the bottom.

As shown in the DNA sequence in Fig. 6, the length of the *Eco*RV fragment was 704 bp. The *Bam*HI fragment derived from plasmid pTIA106, which carries the entire *Eco*RV fragment, had the extra sequence GATCCCC at each 5′ end, derived from the *Bam*HI-*Sma*I adapter fragment in plasmid PTI10. The ends of the chromosomal portion of the *Bal*31-generated plasmids, which are summarized in Table 4, are indicated by the arrows.

In this sequence, there was a 507-bp open reading frame starting from base 163 A(TG) and ending at base 669 (TA)A, which could make a protein of 167 amino acid residues, the molecular weight of which is 19,400. This value agrees well the molecular weight, 18,000, estimated from the maxicell experiment. There was no other open reading frame corresponding to this size in the sequence. As shown in Table 4, the strain carrying plasmid pTIA128 had reduced PGP phosphatase activity. This is explained by the loss of part of the C-terminal sequence. Although plasmid pTIA151 did not have the termination codon within the chromosomal DNA insert, the strain carrying this plasmid showed normal phosphatase activity.

About 100 bases upstream of the Met initiation codon, there were possible −10 and −35 regions for the *pgpA*

TABLE 3. Comparison of phosphatase activities among *pgpA* subclones[a]

Plasmid	PGP phosphatase activity (nmol/min per mg of protein)	Insert size (kb)	Parental plasmid	Insertion site(s)	Host vector
pTIA1	22.4	19.6	TIF33	*Hin*dIII	pACYC184
pTIA6	33.2	5.2	pTIA1	*Sau*3A-*Bam*HI	pACYC184
pTIA4	42.1	2.8	pTIA1	*Sau*3A-*Bam*HI	pACYC184
pTIA12	72.3	1.6	pTIA4	*Eco*RV	pACYC184
pTIA37	62.5	1.4	pTIA4	*Sau*3A-*Bam*HI	pBR322
pTIA26	87.8	0.7	pTIA12	*Eco*RV	pBR322

[a] The common host strain for these plasmids was TI83 (*phoA pgpA pgpB recA*). Phosphatase activities were assayed by measuring the release of $^{32}P_i$ from ^{32}P-labeled PGP by crude sonicates of this strain carrying the indicated plasmids. The size of the chromosomal DNA inserted into the host vector is shown together with the restriction site of its insertion.

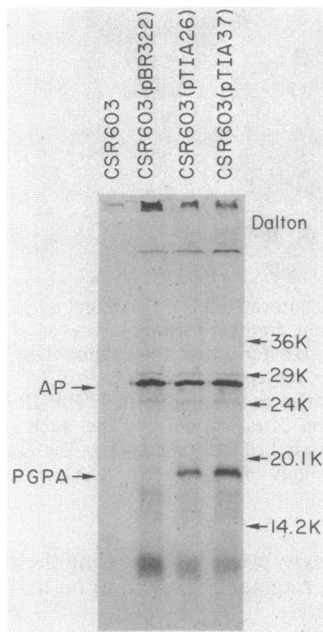

FIG. 4. Detection of an 18-kDa protein encoded by *pgpA*. Maxicell experiments were performed as described by Sancer et al. (29). UV-irradiated samples (1 ml) were labeled with 10 μCi of [^{35}S]methionine (>800 mCi/mmol) for 1 h, centrifuged, suspended in 60 μl of sample buffer, and boiled for 3 min, and half of the volume was loaded on a 12.5% SDS-polyacrylamide gel. After the run, the gel was dried and exposed to X-ray film. CSR603, CSR603(pBR322), CSR603(pTIA26), and CSR603(pTIA37) were tested. The bands corresponding to β-lactamase (AP) and the *pgpA* gene product are shown. Sizes are indicated in kilodaltons.

promoter, which resemble *E. coli* promoter consensus sequences (16, 31). The strain carrying plasmid pTIA156, which had lost part of this −10 region, produced no phosphatase activity, and that carrying plasmid pTIA155, which did not have the −35 sequence, showed reduced activity, supporting this tentative promoter assignment. Although the position of the ribosome binding site is not clear, at eight and seven bases before the Met codon, there was a GG pair which could possibly function in ribosome binding (15, 30).

TABLE 4. PGP phosphatase activity directed by plasmids carrying various lengths of the *pgpA* gene[a]

Plasmid	DNA insert Start (bp)	DNA insert End (bp)	DNA insert Length (bp)	PGP phosphatase activity (nmol/min per mg of protein)	Ratio (relative activity)
pTIA106	7	710	704	76	15.2
pTIA128	30	648	619	28	5.6
pTIA132	30	685	656	77	15.4
pTIA147	30	682	653	71	14.3
pTIA151	21	665	645	69	14.0
pTIA155	39	680	642	36	7.2
pTIA156	65	706	642	4	0.8
pTI10				5	1.0

[a] Plasmid pTIA106 contains the entire *Eco*RV fragment derived from plasmid pTIA26 in the *Sma*I site of plasmid pTI10. Other plasmids were generated by inserting the *Bal*31-treated *Eco*RV fragment into the same site. The start and end points of the sequences are shown with the coordinates given in Fig. 6. The actual sequences have seven base extensions, including the *Bam*HI cohesive ends of each end. Sonicates were prepared from strain R477 (*pgpA*$^+$ *pgpB*$^+$) carrying these plasmids and assayed for PGP phosphatase activity. The ratio of the specific activities was determined by taking the activity of R477(pTI10) as 1. Since *E. coli* has two PGP phosphatases, the actual extent of the overproduction is probably 2 times higher than this ratio.

Assuming that the transcriptional start is several bases after the promoter, there was an mRNA sequence of about 80 nucleotides before the Met codon. I checked the possible mRNA secondary structure around this leader mRNA region by using the program FOLD by Zucker and Stiegfield (35) and detected the structure of a 12-bp stem structure with a 7-bp loop, the energy of which was calculated as −21.9 kcal. This sequence is a repetitive extragenic palindromic (REP) sequence element, which is commonly found after coding sequence or in the untranslated region between two genes of an operon (33).

DISCUSSION

I have presented evidence that *pgpA* is a structural gene for the PGP A phosphatase, which specifically hydrolyzes PGP. As shown in the accompanying paper (18), PGP A phosphatase is localized in the inner membrane, like other phospholipid biosynthetic enzymes in *E. coli* (1). Although this protein has many hydrophobic residues, it is not as

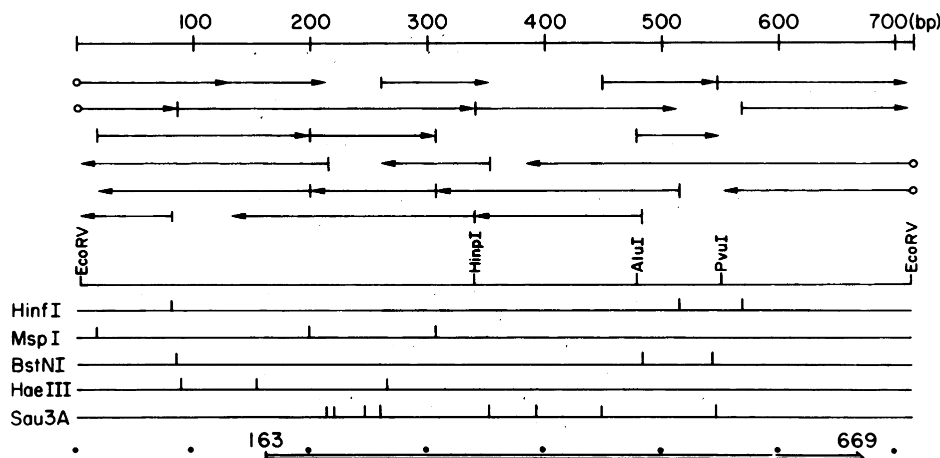

FIG. 5. Strategies for determination of the nucleotide sequence of the *pgpA* gene. The 5'-labeled fragments sequenced are shown by arrows with vertical bars or circles (*Bam*HI ends derived from the adaptor fragment in plasmid pTI10) at the end. Only the sites of the restriction enzymes used for cloning and sequencing are shown. The open reading frame for the *pgpA* gene is illustrated by the thick arrow at the bottom.

```
                                   <-147
                                   <-128
    <-106  <-26        .  <-151    <-132     <-155   .         .        <-156          .          .          .
  1 GATCCCCATCACTTTGAAGCCGGAACAGATTGCCGAACTGGAAACGCCGTATAAACCGCATCCTGTCGTAGGATTTAAATAAGAGTCCAGGCCTGATGAG  100
                                              tcTTGaaa                          ttcgTAtgaTt           GCCTGATGCG
                                                -35                                 -10

101 ACGTGACAAGCGTCACATCAGGCATCGGTGCACAACTACGACAGAATACCCAGCGGCCAGTGATGACCGATAAAATACAGGATGCCTGCGGAAATCACCC  200
    ACGT     GCGTCTTATCAGGC----                            CC      fMetThrAspLysIleGlnAspAlaCysGlyAsnHisPr
         REP                                               SD

201 CGGCGACAATATCGTCGATCATGATCCCCATGCCGCCATGCACATTGCGATCAAACCAGCGGATCGGCCACGGCTTCCACATATCCAGAATACGGAAAAT  300
    oGlyAspAsnIleValAspHisAspProHisAlaAlaMetHisIleAlaIleLysProAlaAspArgProArgLeuProHisIleGlnAsnThrGluAsn

301 CACAAACCCGGCGGCAACCCACTGCCAGTCATTGGTCGGCAGCGCCATGAGCGTGATCCACATACCAATAAATTCGTCCCAGACAATGCTGCCATGATCG  400
    HisLysProGlyGlyAsnProLeuProValIleGlyArgGlnArgHisGluArgAspProHisThrAsnLysPheValProAspAsnAlaAlaMetIleV

401 TGCACACCCATGTCTTTCGCCGTTTGATGACAAAGATAGACGCCGATACAGATCCCCAGCATCACCACCAGCGAGTAGAGCTGCCAGGGCAAAAAGGTCA  500
    alHisThrHisValPheArgArgLeuMetThrLysIleAspAlaAspThrAspProGlnHisHisHisGlnArgValGluLeuProGlyGlnLysGlyHi

501 TCAGATACCAGAACGGAATCGCTGCCAGCGGCCCATCGTCCCAGGAACGATCGGGCTTAATCCACTTCCGAATCCGACAGCAAGTAGATGCCACGGATTA  600
    sGlnIleProGluArgAsnArgCysGlnArgProIleValProGlyThrIleGlyLeuAsnProLeuProAsnProThrAlaSerArgCysHisGlyLeu

601 CTCATCTTCAGGCGACTTTTCGCGACATCTTTATGGCGTGGCAAAATGGTCATATCCTTTCCAGTCTAATGTAACAGGTTCGCCGTCACGAATAAAACAA  700
    LeuIlePheArgArgLeuPheAlaThrSerLeuTrpArgGlyLysMetValIleSerPheProValEnd        155->  ->132
                                                  128->           151->        147->

701 AGCCCTTCGATGGGGATC  718
      156->26->  106->
```

FIG. 6. DNA sequence of the *pgpA* gene. The termini of the chromosomal portion of DNA cloned into the *Sma*I site of plasmid pTI10, summarized in Table 1, are indicated by the arrows. These fragments have the sequence GATCCCC at their 5′ ends, which are derived from the *Bam*HI-*Sma*I adapter fragments in pTI10. These endpoints were either determined by DNA sequencing or deduced from the restriction site. The possible promoter region (−35, −10) for *pgpA* is shown together with the *E. coli* promoter consensus sequence (16, 31). The most conserved residues are capitalized. SD, the possible ribosome binding site (15, 30) is shown with the complementary sequence from the 5′ end of 16S rRNA. The REP is indicated with its consensus sequence (33). The starts and the ends of the DNA inserts listed in Table 4 are indicated by arrows with the plasmid numbers.

hydrophobic as other proteins involved in phospholipid biosynthesis, like CDP diglyceride synthetase (19), PGP B phosphatase (17), or diglyceride kinase (21). Although there are several short hydrophobic regions in this sequence, this protein has no large hydrophobic stretches which could span the membrane. The protein structure analysis carried out by the method of Garnier et al. (14) showed 24.3% α-helix and 11.2% β-sheet content without the decision constant (14). These values are rather low among the hydrophobic proteins involved in phospholipid biosynthesis.

According to Stern et al. (33), all copies of REP so far identified are in extragenic, nontranslated regions, either between two genes or within the 3′ untranslated region at the end of an operon. This is the first case in which an REP element was found between the promoter and the translational start in one gene. However, whether this REP is involved in the regulation of the *pgpA* gene is not known. Since the removal of REP in the case of the histidine operon does not affect downstream gene expression drastically but only reduces the expression of the genes about 50% (33), this might also be true in the case of the REP in the *pgpA* gene. However, the possibility that this structure is involved in some unknown regulation cannot be eliminated. Perhaps this structure could help to stabilize the unusually long leader *pgpA* mRNA and to optimize the expression of the *pgpA* gene. In fact, the insertion of the *Bam*HI fragment isolated from plasmid pTIA132 downstream of a *tac* promoter in runaway replication expression vector pTI5 (17) caused no large stimulation of PGP phosphatase production after isopropyl-β-D-thiogalactoside induction. The existence of REP may play some role, in this extreme case, in modulating gene expression. Among the phospholipid-biosynthetic genes, the *cds* gene has three tandem repeats of sequences similar to half of the REP sequence (33). Also, there is a copy of the REP sequence at the 3′ end of the *plsB* gene (19).

ACKNOWLEDGMENTS

I thank C. R. H. Raetz, Department of Biochemistry, University of Wisconsin at Madison, for his advice and support throughout this work. I thank Nobuyo Maeda, Laboratory of Genetics, University of Wisconsin at Madison, for her invaluable advice on DNA sequencing. I also thank R. B. Wickner for his encouragement and support in the preparation of the manuscript.

This work was supported by Public Health Service grant AM-19551 from the National Institutes of Health to C. R. H. Raetz.

LITERATURE CITED

1. **Bell, R. M., R. D. Mavis, M. J. Osborn, and P. R. Vagelos.** 1971. Enzymes of phospholipid metabolism: location in the cytoplasmic and outer membrane of the cell envelope of *Escherichia coli* and *Salmonella typhimurium*. Biochim. Biophys. Acta **249**:628–635.
2. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
3. **Bolivar, F.** 1978. Construction and characterization of new cloning vehicles. III. Derivatives of plasmid pBR322 carrying unique *Eco*RI sites: selection of *Eco*RI generated recombinant DNA molecules. Gene **4**:121–136.

4. **Bosch, H. V. D., and P. R. Vagelos.** 1970. Fatty acyl-CoA and fatty acyl-acyl carrier protein as acyl donors in the synthesis of lysophosphatidate and phosphatidate in *Escherichia coli*. Biochim. Biophys. Acta **218**:233–248.
5. **Cech, T. R., N. K. Tanner, I. Tinoco, Jr., B. R. Weir, M. Zuker, and P. S. Perlman.** 1983. Secondary structure of the *Tetrahymena* ribosomal RNA intervening sequence: structural homology with fungal mitochondrial intervening sequences. Proc. Natl. Acad. Sci. USA **80**:3903–3907.
6. **Chang, A. C. Y., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134**: 1141–1156.
7. **Chang, Y., and E. P. Kennedy.** 1967. Phosphatidyl glycerophosphate phosphatase. J. Lipid Res. **8**:456–462.
8. **Clarke, L., and J. Carbon.** 1976. A colony bank containing synthetic ColE1 hybrid plasmids representative of the entire *E. coli* genome. Cell **9**:91–99.
9. **Clewell, D. B., and D. R. Helinski.** 1970. Properties of a supercoiled deoxyribonucleic acid-protein complex and strand specificity of the relaxation event. Biochemistry **99**:4428–4440.
10. **David, R. W., D. Botstein, and J. R. Roth.** 1980. Advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
12. **Doi, O., and S. Nojima.** 1975. Lysophospholipase of *Escherichia coli*. J. Biol. Chem. **250**:5208–5214.
13. **Dretzen, G., M. Bellord, P. Sassone-Corsi, and P. Chambon.** 1981. A reliable method for the recovery of DNA fragments from agarose and acrylamide gels. Anal. Biochem. **112**:295–298.
14. **Garnier, J., D. J. Osguthorpe, and B. Robson.** 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. **120**: 97–120.
15. **Gold, L., D. Pribnow, T. Schneider, A. Schneider, B. S. Singer, and G. Stormo.** 1981. Translational initiation in prokaryotes. Annu. Rev. Microbiol. **35**:365–403.
16. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequence. Nucleic Acids Res. **11**:2237–2254.
17. **Icho, T.** 1988. Membrane-bound phosphatases in *Escherichia coli*: sequence of the *pgpB* gene and dual subcellular localization of the *pgpB* product. J. Bacteriol. **170**:5117–5124.
18. **Icho, T., and C. R. H. Raetz.** 1983. Multiple genes for membrane-bound phosphatases in *Escherichia coli* and their action on phospholipid precursors. J. Bacteriol. **153**:722–730.
19. **Icho, T., C. P. Sparrow, and C. R. H. Raetz.** 1985. Molecular cloning and sequencing of the gene for CDP-diglyceride synthetase of *Escherichia coli*. J. Biol. Chem. **260**:12078–12083.
20. **Ikemura, T.** 1981. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes. J. Mol. Biol. **146**:1–21.
21. **Lightner, V. A., R. M. Bell, and P. Mordrich.** 1983. The DNA sequences encoding *plsB* and *dgk* loci of *Escherichia coli*. J. Biol. Chem. **258**:10862–10866.
22. **Low, K. B.** 1972. *Escherichia coli* K-12 F-prime factors old and new. Bacteriol. Rev. **36**:587–607.
23. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
25. **Michaelis, S., and J. Beckwith.** 1982. Mechanism of incorporation of cell envelope proteins in *Escherichia coli*. Annu. Rev. Microbiol. **36**:435–464.
26. **Raetz, C. R. H.** 1986. Molecular genetics of membrane phospholipid synthesis. Annu. Rev. Genet. **20**:253–295.
27. **Raetz, C. R. H., and K. F. Newman.** 1978. Neutral lipid accumulation in the membranes of *Escherichia coli* mutants lacking diglyceride kinase. J. Biol. Chem. **253**:3882–3887.
28. **Rothstein, R. J., L. F. Low, C. P. Bahl, S. A. Narang, and R. Wu.** 1979. Synthetic adapters for cloning DNA. Methods Enzymol. **68**:98–109.
29. **Sancer, A., A. M. Hack, and W. D. Rupp.** 1979. Simple method for identification of plasmid-coded proteins. J. Bacteriol. **137**: 692–693.
30. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71**:1342–1346.
31. **Siebenlist, U., R. B. Simpson, and W. Gilbert.** 1980. E. coli RNA polymerase interacts homologously with two different promoters. Cell **20**:269–281.
32. **Slotta, K. H.** 1957. Phospholipases. Methods Enzymol. **4**:551–558.
33. **Stern, M. J., G. F.-L. Ames, N. H. Smith, E. C. Robinson, and C. F. Higgins.** 1984. Repetitive extragenic palindromic sequences: a major component of the bacterial genome. Cell **37**: 1015–1026.
34. **Zehnbauer, B. A., and A. Markovitz.** 1980. Cloning of gene *lon* (*capR*) of *Escherichia coli* K-12 and identification of polypeptides specified by the cloned deoxyribonucleic acid fragment. J. Bacteriol. **143**:852–863.
35. **Zuker, M., and P. Stiegler.** 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. **9**:133–148.