

Cloning, Nucleotide Sequence, and Expression in *Escherichia coli* of the Phospholipase D Gene from *Corynebacterium pseudotuberculosis*

ADRIAN L. M. HODGSON,[†] PHILLIP BIRD,[‡] AND IAN T. NISBET*

Research and Development Division, Commonwealth Serum Laboratories, 45 Poplar Road, Parkville 3052, Victoria, Australia

Received 20 June 1989/Accepted 23 November 1989

The phospholipase D (PLD) gene from *Corynebacterium pseudotuberculosis* has been cloned, sequenced, and expressed in *Escherichia coli*. Analysis of DNA sequence data reveals a major open reading frame encoding a 31.4-kilodalton protein, a size consistent with that estimated for the PLD protein by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Comparison of these data with the amino-terminal protein sequence indicates that the mature PLD protein is preceded by a 24-residue signal sequence. Expression of the PLD gene in *E. coli* is initiated from the corynebacterial promoter, and the resulting protein has sphingomyelinase activity. Primer extension mapping localized the 5' end of the PLD gene mRNA to a site 5 to 7 base pairs downstream of a region similar to the consensus sequence for *E. coli* promoters. Northern and Southern blot analyses suggest that the gene is transcribed from mRNA approximately 1.1 kilobases in length and that it is present in a single copy within the *C. pseudotuberculosis* genome.

Corynebacterium pseudotuberculosis is the gram-positive bacterium responsible for the disease in sheep known as caseous lymphadenitis or cheesy gland. The pathogenesis of caseous lymphadenitis usually involves entry of *C. pseudotuberculosis* through broken skin, following which the bacteria are carried to the local lymph node. Once established, they multiply to cause inflammation, necrosis, and abscessation of the node (3). An important factor in the dissemination of *C. pseudotuberculosis* within an infected animal is the production of an exotoxin, phospholipase D (PLD) (2). Sheep can be vaccinated against caseous lymphadenitis by using detoxified PLD prepared from *C. pseudotuberculosis* culture supernatants (2; D. G. Eggleton, J. A. Haynes, H. D. Middleton, and J. C. Cox, submitted for publication).

Studies with partially purified PLD indicate that it is a sphingomyelinase with a molecular mass of about 31 kilodaltons (kDa) and a pI of 9.2 to 9.6 (7). However, there have been no previous reports describing the purification of this toxin to homogeneity, and no amino acid sequence data are available. Here we report the cloning, expression, and complete nucleotide sequence of the *C. pseudotuberculosis* PLD structural gene.

MATERIALS AND METHODS

Bacterial strains and plasmids. *Escherichia coli* DH5 α (Bethesda Research Laboratories, Inc.) was the host for plasmids pUC12 and pUC118 (18). The Commonwealth Serum Laboratories (CSL) production strain of *C. pseudotuberculosis* (West Australian Department of Agriculture, strain 1030) and an isolate from New Zealand (isolate 107) were used in this study. The plasmids pCSL33, pCSL39, and pCSL40 were constructed during this study. pCSL33 is

pUC12 containing the PLD gene on a 2.7-kilobase (kb) fragment. pCSL39 and pCSL40 were derived by subcloning a 1.5-kb *SacI* fragment (Fig. 1) from pCSL33 into pUC118. In pCSL39 the PLD gene is transcribed in the opposite direction to the *lacZ* gene, whereas in pCSL40 it is transcribed in the same direction.

Media. *E. coli* strains were grown in Luria broth (10 g of tryptone, 5 g of yeast extract, and 10 g of NaCl per liter) containing 30 μ g of ampicillin per ml. *C. pseudotuberculosis* was grown in nutrient infusion broth (NIB; 500 ml of veal extract [CSL], 5 g of NaCl, and 10 g of proteose peptone per liter).

Toxin purification. Toxin was purified from *C. pseudotuberculosis* culture supernatants by cation-exchange chromatography (J. C. Cox, J. A. Haynes, D. Drane, E. Pietrzykowski, and I. Roberts, manuscript in preparation).

Assay for *C. pseudotuberculosis* PLD. *C. pseudotuberculosis* PLD activity was detected by using a modification of the Zaki assay (15, 17) and an assay to detect sphingomyelinase (4). For the Zaki assay 1.5-ml samples of a 15% (vol/vol) suspension of sheep erythrocytes in Alsevers solution were washed three times in 5 ml of buffered saline (0.01 M *N*-tris(hydroxymethyl)-methyl-*s*-aminoethane sulfuric acid, 0.01 M MgCl₂, 0.85% [wt/vol] NaCl, and 0.1% [wt/vol] bovine serum albumin [pH 7.3]). The washed cells were then diluted 10-fold in buffered saline. Doubling dilutions of test samples (100 μ l) in buffered saline were made in microdilution trays. Samples (100 μ l) of washed sheep erythrocytes were then added to each test sample and held at 37°C for 45 min. Beta-lysin (CSL special product) (50 μ l) was then added to each test well, and the trays were held at 37°C for 30 min and then on ice for 30 min. Microdilution trays were spun in a Beckman GPR bench-top centrifuge for 30 s at 400 \times *g*. The number of Zaki units present in a given sample was determined from the dilution of PLD test sample that no longer protected the blood cells from lysis by staphylococcal beta-hemolysin. Tests were standardized against a stock solution of PLD. The sphingomyelinase assay was performed at pH 7.4 as described by Gatt et al. (5) with the

* Corresponding author.

[†] Present address: Commonwealth Scientific and Industrial Research Organization Division of Animal Health, Parkville 3052, Victoria, Australia.

[‡] Present address: Department of Medicine, Monash Medical School, Alfred Hospital, Prahran 3181, Victoria, Australia.

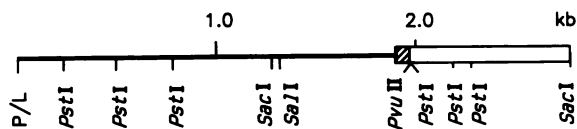


FIG. 1. Restriction map of the PLD gene region from *C. pseudotuberculosis*. ▨, PLD gene signal sequence; □, region encoding mature PLD protein.

chromogenic substrate trinitrophenylaminolauryl sphingomyelin (Sigma Chemical Co.). Because PLD requires Mg^{2+} for activity, 30 μM $MgCl_2$ was included in the reaction buffer.

N-terminal sequencing of the PLD protein. The N terminus of purified PLD (50 μg) was determined by using an Applied Biosystems 470A protein sequencer.

Preparation of oligonucleotides. PLD-specific oligonucleotides deduced from N-terminal PLD protein sequence data were synthesized by using an Applied Biosystems 380B DNA synthesizer. Oligonucleotides for screening gene libraries were prepared as mixtures to accommodate codon degeneracy or incorporated inosine residues at equivocal positions (13). Oligonucleotides to be used as sequencing primers were designed from derived DNA sequence data. Tritylated oligonucleotides were purified with oligonucleotide purification cartridges (Applied Biosystems).

Labeling oligonucleotides. Oligonucleotides were 5' end labeled (10) by using 50 μCi of γ -ATP (3,000 Ci per mmol; Amersham Corp.) and T4 polynucleotide kinase (5 U) to a specific activity of at least 10^7 dpm/ μg .

Isolation of genomic DNA from *C. pseudotuberculosis*. A 10-ml overnight culture of *C. pseudotuberculosis* was used to inoculate 200 ml of NIB containing 1.0 μg of penicillin per ml. After growth overnight at 37°C, the culture was centrifuged (14,000 $\times g$, 10 min), and the pellet was suspended in 10 ml of STET (8% sucrose, 0.5% Triton X-100, 50 mM EDTA [pH 8.0], 10 mM Tris hydrochloride [pH 8.0]) containing 10 mg of lysozyme per ml. The mix was held at 37°C for 2 h, and 3.0 ml of lysis solution (1.0% [wt/vol] sodium dodecyl sulfate, 0.2 M NaOH) was then added. After a further hour at 55°C the lysate was centrifuged at 47,800 $\times g$ for 30 min. Chromosomal DNA was isolated from the cleared lysates by using cesium chloride-ethidium bromide gradients (10).

Construction and screening of genomic library. Genomic DNA (22 μg) from *C. pseudotuberculosis* was partially digested with 6 U each of *Hae*III and *Alu*I. Digests were run on a 1.0% agarose gel; fragments between 2 and 4 kb were excised, purified with GeneClean (Bresatec, Adelaide), and cloned into the *Sma*I site of pUC12. Ligation mixes were transformed into *E. coli* DH5 α . Transformants were screened by using a rabbit anti-PLD serum (Cox et al., in preparation) and a SuperScreen immunoscreening system kit (Amersham).

Expression of PLD gene in *E. coli*. Cultures (1 ml) of *E. coli* harboring the toxin gene were pelleted and suspended in 100 μl of lysis buffer (12.5 μg of lysozyme per ml of STET). The mix was centrifuged, and 80 μl of supernatant was used in each assay. Lysates were tested for the production of PLD by using Western blotting (immunoblotting) and the modified Zaki test. For the sphingomyelinase assays, periplasmic extracts were prepared by the method of Kendall et al. (9).

Restriction mapping, subcloning, and Southern blotting. A restriction map of an antibody-reactive clone (pCSL33) was derived by using standard procedures (10). Plasmid DNA carrying the putative PLD gene was digested with *Pst*I or

*Sac*I, and the fragments were Southern blotted to nitrocellulose. Filters were hybridized (12) at room temperature with a toxin-specific end-labeled probe (5'-GTIGTICAC/TAAC/TCCIGC-3') for 2 h in 6 \times SSC (1 \times SSC is 0.15 M NaCl plus 0.015 M sodium citrate)-10 \times Denhardt solution (10), washed at increasing stringency as necessary (up to 45°C), and exposed to X-ray film (Fuji RX).

DNA sequence analysis. M13 clones were sequenced by using modified T7 DNA polymerase according to the instructions of the manufacturer (U.S. Biochemical Corp.) with either universal primer or synthetic oligonucleotides. Segments of overlapping DNA sequence were generated by using oligonucleotide primers designed from the PLD sequence as it became available. Sequence data were collated and analyzed by using the DNASIS software package (Pharmacia Fine Chemicals).

Primer extension. Total RNA was isolated from *E. coli* and *C. pseudotuberculosis* as described by Aiba et al. (1). Then 0.1 to 10 pmol of oligonucleotide primer (5'-TTGAGTGGT TAAAACGCGGTGGC-3') was end labeled (10), ethanol precipitated, and annealed to 20 μg of RNA. Primer extension was accomplished at 42°C for 1.5 h with 20 U of avian myeloblastosis virus reverse transcriptase (Pharmacia) under conditions specified by the manufacturer. Extension products were separated on a 6% polyacrylamide sequencing gel, which was fixed, dried, and exposed to X-ray film.

Analysis of *C. pseudotuberculosis* RNA. Total RNA was isolated from *C. pseudotuberculosis* as described above. A 950-base-pair (bp) *Pvu*II fragment containing most of the PLD gene plus a small portion of the *lacZ* gene was isolated from pCSL39 and labeled to a specific activity of 10^8 dpm/ μg with a random primer labeling kit (Bethesda Research Laboratories). RNA was electrophoresed on a 1% formaldehyde agarose gel, blotted to nitrocellulose, and then hybridized at 68°C overnight with the gene probe (10). Size estimations were made relative to RNA molecular weight markers (Bethesda Research Laboratories).

Analysis of *C. pseudotuberculosis* genomic DNA. Genomic DNA isolated from *C. pseudotuberculosis* was restricted, electrophoresed on a 1% agarose gel, and then Southern blotted (10) to a nitrocellulose filter (Amersham). The PLD gene-specific probe was produced as described above. Hybridization was at 68°C overnight in 6 \times SSC. The filter was then washed in 1 \times SSC for 20 min at 68°C and exposed to X-ray film.

RESULTS

Cloning of the PLD structural gene. *C. pseudotuberculosis* genomic DNA was partially digested with *Alu*I and *Hae*III, and fragments approximately 2 kb in size were purified and cloned into the *Sma*I site of pUC12. A library was prepared in DH5 α , and 15,000 clones were screened by using an anti-PLD antiserum. One positive clone was identified.

A 5.6-kb plasmid (pCSL33) was isolated from this clone and mapped by using a variety of restriction endonucleases (Fig. 1). To confirm that the 2.7-kb insert in pCSL33 contained at least part of the PLD gene, the restricted plasmid DNA was Southern blotted to nitrocellulose and hybridized with a PLD-specific oligonucleotide probe. A 1.5-kb *Sac*I fragment and a 200-bp *Pst*I fragment hybridized with this probe (data not shown). Since the hybridizing oligonucleotide was designed from protein sequence data derived from the PLD amino terminus, the 5' end of the PLD structural gene must be located on the 200-bp *Pst*I fragment (Fig. 1). Given the size of the mature protein (31 kDa) and hence a

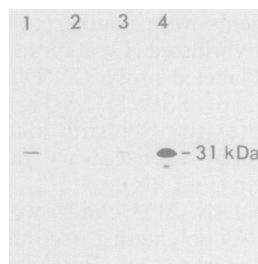


FIG. 2. Western blot analyses of *E. coli* lysates and PLD purified from *C. pseudotuberculosis*. Lanes: 1, lysate of *E. coli* harboring a 2.7-kb clone of the PLD gene region (pCSL33); 2, plasmid vector without PLD gene insert (pUC12); 3, 1.5-kb subclone of the PLD gene region (pCSL39); 4, purified PLD from *C. pseudotuberculosis*. The band beneath the 31-kDa protein in lane 4 is possibly a degradation product of PLD protein.

gene size of around 850 bp, it is likely that the entire structural gene is contained within the 1.5-kb *SacI* fragment (Fig. 1).

Expression of the PLD gene in *E. coli*. The 1.5-kb *SacI* fragment was subcloned into pUC118, and recombinants were identified by using the PLD-specific antiserum. Restriction analysis of a number of the clones that reacted with the antiserum revealed that the *SacI* fragment was present in both orientations. This result suggests that the PLD promoter is functional in *E. coli*. Two plasmids (pCSL39 and pCSL40) having the 1.5-kb fragment cloned in opposite orientations were chosen for further analysis.

Cell lysates were prepared from strains harboring pCSL33, pCSL39, or pCSL40 and electrophoresed on a 12% sodium dodecyl sulfate-polyacrylamide gel. Proteins were transferred to nitrocellulose by electroblotting and then tested for reactivity with the anti-PLD antiserum. Strains containing either the original clone (pCSL33) or one of the subclones (pCSL39 or pCSL40) produced a 31-kDa protein that reacted with PLD-specific antibodies (Fig. 2). This size is consistent with that of the mature PLD protein, indicating that the complete PLD structural gene is present on the 1.5-kb *SacI* fragment.

To determine the subcellular localization of PLD produced in *E. coli*, lysates of strains containing pCSL39 were separated into cytoplasmic and periplasmic fractions as described previously (9). Proteins from these fractions were separated on 12% sodium dodecyl sulfate-polyacrylamide gels and analyzed by silver staining or Western blotting (data not shown). These experiments demonstrated that PLD is predominately located in the periplasm, indicating that the PLD signal sequence functions in *E. coli*.

Activity of PLD produced in *E. coli*. PLD produced by *C. pseudotuberculosis* is commonly assayed by using the Zaki test, which measures the ability of PLD to block the lysis of sheep erythrocytes by staphylococcal beta-hemolysin (15, 17). Cell lysates from all PLD clones had demonstrable activity in this assay, whereas host cells harboring plasmids without insert DNA produced no detectable reaction (data not shown).

Previous studies have shown that PLD is a sphingomyelinase (5). In an assay that measures the hydrolysis of a synthetic substrate, trinitrophenylaminolauryl sphingomyelin, purified PLD showed reactivity similar to that of sphingomyelinase prepared from *Staphylococcus aureus* (Fig. 3). Periplasmic extracts prepared from *E. coli* harboring the PLD gene were also capable of hydrolyzing trinitrophenylaminolauryl sphingomyelin, whereas no activity was de-

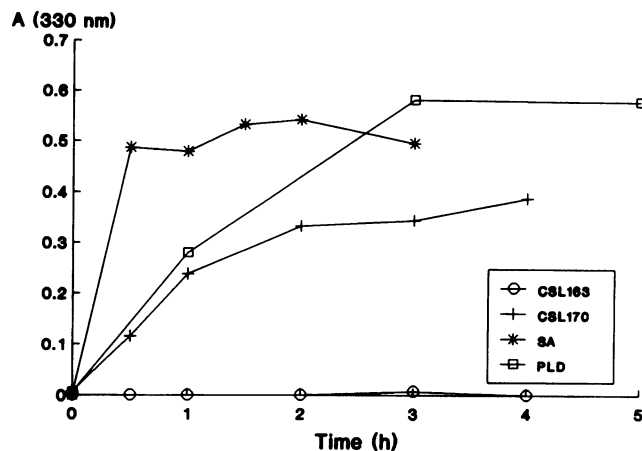


FIG. 3. Sphingomyelinase activity of PLD produced in *E. coli*. The assay (5) was performed with 120 nmol of trinitrophenylaminolauryl sphingomyelin as the substrate. Periplasmic extracts from *E. coli* DH5 α harboring pU118 (designated CSL163) or pCSL39 (designated CSL170) were used. Controls included 120 ng of sphingomyelinase purified from *S. aureus* (SA) and 200 ng of purified PLD (PLD). The samples from CSL163 and CSL170 contained approximately 100 μ g of protein. The CSL170 extract included 100 ng of PLD, estimated by enzyme immunoassay (data not shown).

tected from extracts of *E. coli* cells lacking the PLD gene. These data show that a functional protein is produced when the PLD gene is expressed in *E. coli*.

DNA sequence analysis. The 1.5-kb *SacI* fragment was subcloned in both orientations into M13mp18. Nucleotide sequence data from both strands were derived for approximately 1,100 bp of the *SacI* fragment (Fig. 4) by using a modified dideoxy-chain termination procedure (16).

Analysis of the DNA sequence reveals a major open reading frame beginning just before the *PvuII* site and terminating at the end of the cloned fragment (Fig. 1). This open reading frame is capable of coding for a 307-residue protein with a predicted size of 34.1 kDa. To confirm that this protein corresponds to PLD, the predicted polypeptide sequence was compared with the amino-terminal sequence data obtained from purified PLD. A stretch of 23 residues identical to the PLD protein sequence was identified 25 amino acids downstream from the predicted translational initiation site. The preceding 24-amino-acid sequence is strongly indicative of a signal sequence (19). On this basis the open reading frame encodes a polypeptide that includes a putative signal peptide (2.7 kDa) and the mature PLD protein (31.4 kDa). The size of mature PLD predicted from the sequence is therefore in close agreement with that estimated by SDS-polyacrylamide gel electrophoresis (Fig. 2).

The DNA and protein sequences of PLD were used to search the appropriate DNA (GENBANK, EMBL) or protein (NBRF) data bases for homologous sequences. Surveys of the nucleotide data bases failed to detect sequences showing significant similarity to PLD. Comparison of the PLD protein sequence with those of other phospholipases, however, revealed a small region of homology with the phospholipases A2 (Fig. 5).

Determination of the transcriptional start site. Expression studies in *E. coli* indicated that the *E. coli* RNA polymerase can use a *C. pseudotuberculosis* sequence as a promoter. Inspection of the DNA sequence upstream of the PLD structural gene revealed several regions that resemble the consensus *E. coli* promoter.

		(-30)	(-20)	(-10)	(1)	
1	CAA CTG TCA GCT TCT GAA ATA TAT GAG TAT AAA CAT ATG CAA TAA CCC	<u>TIG ATT</u>	TTA TTG TTA TTT AAG	<u>TTT CAT AAT</u>	GGG GAT ATA GCG CAG GGT	96
97	GTT CAC GCT GGT AAA GGG CTA TTT TTG GGC ACC TCT CCG TGG GAA ATG TAT CCG CAA ATT ATA GTG CGA CAT CCT TTT GTT GCT CTA TTT ATC GAA					192
		-20	-10	1		
193	ACT <u>TGG TGA</u> TAA AAA ATG AGG GAG AAA GTT GTT TTA TTC TTA TCA ATA ATT ATG GCG ATC ATG CTT CCG GTA GGG AAT GCA GCT GCA GCG CCT GTT	Met Arg Glu Lys Val Val Leu Phe Leu Ser Ile Ile Met Ala Ile Met Leu Pro Val Gly Asn Ala Ala Ala	<u>Ala Pro Val</u>			288
		10	20	30		
289	GTG CAT AAC CCA GCT TCT ACA GCA AAT CCG CCA GTC TAT GCG ATT GCC CAC CGC GTT TTA ACC ACT CAA GGC GTG GAT GAC GCA GTT GCG ATC GGT	Val His Asn Pro Ala Ser Thr Ala Asn Arg Pro Val Tyr Ala Ile Ala His Arg Val Leu	Thr Thr Gln Gly Val Asp Asp Ala Val Ala Ile Gly			384
		40	50	60		
385	GCG AAT GCG TTA GAA ATT GAC TTC ACT GCG TGG GGT CGT GGC TGG TGG GCA GAT CAT GAT GGT ATT CCT ACT AGC GCA GGT GCT ACT GCA GAG GAA	Ala Asn Ala Leu Glu Ile Asp Phe Thr Ala Trp Gly Arg Gly Trp Trp Ala Asp His Asp Gly Ile Pro Thr Ser Ala Gly Ala Thr Ala Glu Glu				480
		70	80	90		
481	ATT TTT AAG CAT ATA GCT GAT AAG CGT AAG CAG GGA GCA AAT ATT ACT TTC ACC TGG CTT GAC ATC AAG AAT CCA GAC TAC TGC AGG GAT GCT CGT	Ile Phe Lys His Ile Ala Asp Lys Arg Lys Gln Gly Ala Asn Ile Thr Phe Thr Trp Leu Asp Ile Lys Asn Pro Asp Tyr Cys Arg Asp Ala Arg				576
		100	110	120	130	
577	AGT GTG TGC TCC ATA AAT GCG TTG CGT GAT TTG GCA CGT AAA TAT CTT GAG CCG GCA GGG GTT CGA GTT CTC TAT GGG TTC TAT AAG ACA GTC GGC	Ser Val Cys Ser Ile Asn Ala Leu Arg Asp Leu Ala Arg Lys Tyr Leu Glu Pro Ala Gly Val Arg Val Leu Tyr Gly Phe Tyr Lys Thr Val Gly				672
		140	150	160		
673	GGA CCT GCC TGG AAG ACA ATC ACC GCT GAT CTT CGG GAT GGC GAG GCG GTA GCT CTT AGC GGC CCG GCG CAG GAC GTA TTA AAT GAT TTT GCA AGG	Gly Pro Ala Trp Lys Thr Ile Thr Ala Asp Leu Arg Asp Gly Glu Ala Val Ala Leu Ser Gly Pro Ala Gln Asp Val Leu Asn Asp Phe Ala Arg				768
		170	180	190		
769	TCT GAA AAT AAG ATC CTT ACT AAA CAA AAA ATC GCT GAC TAT GGT TAC TAC AAC ATT AAC CAA GGG TTT GGT AAC TGC TAT GGA ACC TGG AAT CCG	Ser Glu Asn Lys Ile Leu Thr Lys Gln Lys Ile Ala Asp Tyr Gly Tyr Tyr Asn Ile Asn Gln Gly Phe Gly Asn Cys Tyr Gly Thr Trp Asn Arg				864
		200	210	220		
865	ACT TGT GAT CAA CTC CGT AAG TCC AGC GAA GCT CGT GAC CAA GGA AAA CTC GGT AAA ACT TTT GGG TGG ACA ATC GCT ACA GGT CAG GAC GCG CGA	Thr Cys Asp Gln Leu Arg Lys Ser Ser Glu Ala Arg Asp Gln Gly Lys Leu Gly Lys Thr Phe Gly Trp Thr Ile Ala Thr Gly Gln Asp Ala Arg				960
		230	240	250		
961	GTT AAT GAT CTT TTA GGA AAA GCC AAC GTA GAT GGA CTG ATC TTT GGC TTT AAG ATT ACT CAC TTC TAC CGT CAT GCA GAC ACC GAA AAT TCT TTC	Val Asn Asp Leu Leu Gly Lys Ala Asn Val Asp Gly Leu Ile Phe Gly Phe Lys Ile Thr His Phe Tyr Arg His Ala Asp Thr Glu Asn Ser Phe				1056
		260	270	280		
1057	AAA GCC ATC AAG AGG TGG GTG GAT AAG CAC TCC GCT ACT CAC CAT CTG GCT ACC GTA GCG GAT AAC CCG TGG TGA GGG GCG AG	Lys Ala Ile Lys Arg Trp Val Asp Lys His Ser Ala Thr His His Leu Ala Thr Val Ala Asp Asn Pro Trp	***			1142

FIG. 4. Nucleotide sequence of the PLD gene from *C. pseudotuberculosis*. Underlined bases correspond to regions with complete or partial homology with the *E. coli* consensus sequences for -10 and -35 promoter regions. Numbers in brackets are based on thymine 84 as the transcriptional start site. rbs, Putative ribosome binding site. Underlined amino acid residues are those identified from sequencing purified PLD protein.

Primer extension experiments were conducted to determine the location of the PLD promoter. A labeled oligonucleotide designed to anneal to the 5' end of the PLD mRNA was hybridized to total RNA prepared from either *C. pseudotuberculosis* or *E. coli* carrying pCSL39 and then extended by using reverse transcriptase. The extension

products were separated on a 6% sequencing gel and compared with a DNA sequence ladder generated by using the same oligonucleotide as a primer (Fig. 6). These experiments suggested that the RNA polymerase in both *E. coli* and *C. pseudotuberculosis* initiates transcription at the thymine located at position 84 in the sequence. Based on the *E. coli* promoter consensus sequence (14), two overlapping -10 regions and a -35 box were identified upstream of this putative initiation site (Fig. 4).

Determination of PLD gene copy number in *C. pseudotuberculosis*. Separate samples of *C. pseudotuberculosis* genomic DNA were digested with restriction enzymes, electrophoresed on a 1% agarose gel, transferred to nitrocellulose, and hybridized to the same probe used for the RNA analysis (see above). With the exception of *Pst*I, the enzymes used in this experiment were chosen because they do not have recognition sites within the PLD-coding sequence. There-

	111
PLD	ARKYLEPAGV RVLYGFYKTV GGPAMKTITA DLRDGEAVAL SGPAQ
	* * * * * * * * *
PLA2	ADKGRPRWH YMDYGCYCGP GSGTVPDEL DRCCKTHDQC YAQAE
	12

FIG. 5. Amino acid sequence homology between *C. pseudotuberculosis* PLD and *Laticauda laticaudata* phospholipase A2 (PLA2). Underlined residues in phospholipase A2 are involved in calcium binding. In some phospholipase A2 molecules the glycine (G) marked with a line is an alanine (A) (19).

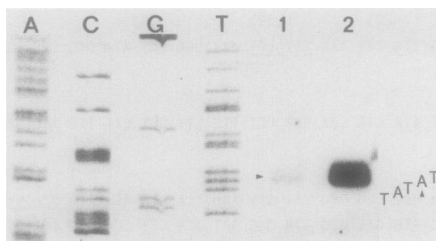


FIG. 6. Transcriptional initiation site of the PLD gene: autoradiogram showing results of primer extension reactions used to determine the position of the transcriptional start site. Primer extension products were generated by using *C. pseudotuberculosis* RNA (lane 1) and *E. coli* RNA (lane 2). The relevant DNA sequence is shown on the right. The arrowhead marks the position of the proposed start site at position 84 in the DNA sequence. Note that this sequence is complementary to that shown in Fig. 4.

fore, if the PLD gene is present more than once in the *C. pseudotuberculosis* genome, under stringent conditions it is likely that one or more of these enzymes will give rise to multiple hybridizing bands. Analysis of the PLD restriction map (Fig. 1) predicts that the probe should hybridize to 93- and 200-bp *Pst*I-generated fragments as well as to a fragment larger than 600 bp (depending on the position of the next site downstream from the coding sequence). Since fragments of less than 300 bp were not retained on the gel, only one fragment would be detected if there were a single PLD gene. Single hybridizing bands were observed in all the tracks (Fig. 7A), suggesting that the PLD gene is present as a single copy in the *C. pseudotuberculosis* genome.

Analysis of *C. pseudotuberculosis* RNA. Because the PLD-coding sequence finishes within 10 bp of the end of the cloned DNA, the fragment does not contain sequences resembling either transcriptional terminators for the PLD gene or the translational initiation signals for a downstream gene. To determine the size of the authentic PLD transcript, total RNA was prepared from *C. pseudotuberculosis*, separated on a 1% agarose denaturing gel, transferred to nitrocellulose, and hybridized to a PLD-specific probe labeled by random priming. The size (approximately 1100 bp) of the mRNA species that hybridized to the probe was similar to the size of the PLD gene (Fig. 7B). This suggests that the PLD transcript terminates just downstream of the coding sequence.

DISCUSSION

DNA sequence analysis of the *C. pseudotuberculosis* PLD structural gene reveals two possible open reading frames.

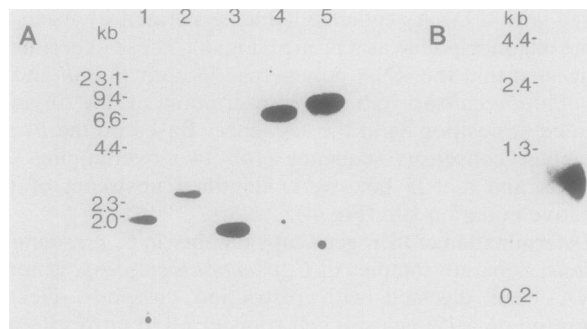


FIG. 7. (A) Southern blot analysis of *C. pseudotuberculosis* genomic DNA with a PLD gene-specific probe. Lanes: 1, *Sac*I; 2, *Pst*I; 3, *Sal*I; 4, *Bam*HI; 5, *Hind*III. (B) Northern blot analysis of *C. pseudotuberculosis* RNA with the same probe.

The first, beginning at base 142 (Fig. 4), is capable of encoding a polypeptide consisting of 18 amino acids. However, there is no evidence to suggest that this polypeptide is synthesized. The predicted translational start site for the second open reading frame is a methionine codon at position -24 (Fig. 4), which is preceded by a region similar to the *E. coli* consensus sequence for a ribosome-binding site (14) and also by two stop codons (Fig. 4). The major ORF encodes a 24-residue signal sequence (19) and a 259-amino-acid (31.4-kDa) polypeptide. This structure is consistent with previous data (7) identifying PLD as a 31.0-kDa protein secreted by *C. pseudotuberculosis*.

Comparison of the PLD protein sequence with other phospholipases reveals some similarity to the phospholipase A2 class (Fig. 5). It is interesting to note that this region is highly conserved among the phospholipase A2 enzymes and that both the calcium-binding and the active sites fall within it (20). Since PLD is a magnesium-dependent enzyme, we postulate that the homologous region in PLD may contain the PLD metal-binding site.

The sequence data show that the cloned fragment does not contain transcriptional terminators. It is therefore possible that other open reading frames exist on the same transcript downstream from the PLD gene. Analysis of *C. pseudotuberculosis* RNA indicates that this is probably not the case, since the mRNA is approximately the same size as the gene (Fig. 7B).

There are a number of regions upstream of the beginning of the PLD-coding sequence that resemble the consensus sequence for an *E. coli* promoter (14). To determine which one is the PLD promoter, the position of the transcriptional start site was determined. Primer extension data suggest that transcription of the PLD gene in both *C. pseudotuberculosis* and *E. coli* is initiated at thymine number 84 (Fig. 4 and 6). Initiation at a thymine does not follow the *E. coli* consensus model, which favors purines (11). However, this result may reflect the sensitivity of the primer extension technique.

Regions similar to the consensus sequences for the *E. coli* promoter -10 and -35 boxes are located close to this putative transcriptional start site (Fig. 4). As observed for the *Corynebacterium diphtheriae* toxin gene, the PLD promoter has two possible -10 regions (6, 8). In *C. diphtheriae* it has been suggested that the TAGGAT box showing optimal spacing (18 bp from the -35 region) would be used in preference to the possible alternative -10 region (TATAAT). Although the latter box is identical to the *E. coli* consensus sequence, it is only 12 bp from the -35 region. Recently both -10 regions were shown to function in *E. coli*; however, as predicted the optimally spaced -10 region was favored (4). In the case of the PLD gene, the spacing between the putative -10 (CATAAT) and -35 regions is 18 bp, thus fitting the *E. coli* consensus model. The other putative -10 region for the PLD gene (TTTCAT) does not closely resemble the *E. coli* consensus -10 box (TATAAT) and is suboptimally spaced. The question of which -10 region directs gene expression in *E. coli* and corynebacteria is currently being addressed.

ACKNOWLEDGMENTS

We thank J. Bennet, M. Black, J. Tkalcovic, and J. Morcom for expert assistance and B. Grego and Y. Ha (Baker Medical Research Institute) for performing protein sequencing and oligonucleotide syntheses. Thanks to Ross Hamilton for photography.

This work was conducted with the approval of the Australian Recombinant DNA Monitoring Committee.

LITERATURE CITED

1. Aiba, H., S. Adhya, and B. de Crombrughe. 1981. Evidence for two functional *gal* promoters in intact *Escherichia coli* cells. *J. Biol. Chem.* **256**:11905-11910.
2. Batey, R. G. 1986. Pathogenesis of caseous lymphadenitis in sheep and goats. *Aust. Vet. J.* **63**:269-272.
3. Beveridge, W. I. B. 1983. Animal health in Australia, vol 4. Bacterial diseases of cattle, sheep and goats. Australian Government Publishing Service, Canberra.
4. Boyd, J., and J. R. Murphy. 1988. Analysis of the diphtheria *tox* promoter by site-directed mutagenesis. *J. Bacteriol.* **170**:5949-5952.
5. Gatt, S., T. Dinur, and Y. Barenholz. 1978. A spectrophotometric method for determination of sphingomyelinase. *Biochim. Biophys. Acta* **530**:503-507.
6. Greenfield, L., M. J. Bjorn, G. Horn, D. Fong, G. A. Buck, R. J. Collier, and D. A. Kaplan. 1983. Nucleotide sequence of the structural gene for diphtheria toxin carried by corynebacteriophage beta. *Proc. Natl Acad. Sci. USA* **80**:6853-6857.
7. Hsu, T. Y., H. W. Renshaw, C. W. Livingston, J. L. Augustine, D. L. Zink, and B. B. Gauer. 1985. *Corynebacterium pseudotuberculosis* exotoxin: fatal hemolytic anemia induced in gnotobiotic neonatal small ruminants by parenteral administration of preparations containing exotoxin. *Am. J. Vet. Res.* **46**:1206-1211.
8. Kaczorek, M., F. Delpyroux, N. Chenciner, R. E. Streek, J. R. Murphy, P. Boquet, and P. Tiollais. 1983. Nucleotide sequence and expression of the diphtheria *tox228* gene in *Escherichia coli*. *Science* **221**:855-858.
9. Kendall, D. A., S. K. Bock, and E. T. Kaiser. 1986. Idealization of the hydrophobic segment of the alkaline phosphatase signal peptide. *Nature (London)* **321**:706-708.
10. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. McClure, W. R. 1985. Mechanism and control of transcription initiation in prokaryotes. *Annu. Rev. Biochem.* **54**:171-204.
12. Meinkoth, J., and G. Wahl. 1984. Hybridization of nucleic acids immobilised on solid supports. *Anal. Biochem.* **138**:267-284.
13. Ohtsuka, E., S. Matsuki, M. Ikehara, Y. Takahashi, and K. Matsubara. 1985. An alternative approach to deoxyoligonucleotides as hybridization probes by insertion of deoxyinosine at ambiguous codon positions. *J. Biol. Chem.* **260**:2605-2608.
14. Reznikoff, W., and L. Gold. 1986. Maximizing gene expression. Butterworths, Boston.
15. Souckova, A., and A. Soucek. 1972. Inhibition of the hemolytic action of alpha and beta lysins of *Staphylococcus pyogenes* by *Corynebacterium hemolyticum*, *C. ovis* and *C. ulcerans*. *Toxicol.* **10**:501-509.
16. Tabor, S., and C. C. Richardson. 1987. DNA sequencing with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA* **84**:4767-4771.
17. Zaki, M. M. 1965. Production of a soluble substance by *Corynebacterium ovis*. *Nature (London)* **205**:928-929.
18. Vieira, J., and J. Messing. 1987. Production of single-stranded plasmid DNA. *Methods Enzymol.* **154**:3-11.
19. von Heijne, G. 1985. Signal sequences. The limits of variation. *J. Mol. Biol.* **184**:99-105.
20. Waite, M. 1987. The phospholipases, p. 155. In D. J. Hanahan (ed.), Handbook of lipid research. Plenum Publishing Corp., New York.