Journal of Bacteriology, Sept. 1991, p. 5597–5603
0021-9193/91/185597-07$02.00/0
Copyright © 1991, American Society for Microbiology
Vol. 173, No. 18

# Cloning, Nucleotide Sequence, and Expression of the *Pasteurella haemolytica* A1 Glycoprotease Gene

KHALID M. ABDULLAH,[1] REGGIE Y. C. LO,[2] AND ALAN MELLORS[1]*

*Guelph-Waterloo Centre for Graduate Work in Chemistry[1] and Department of Microbiology,[2] University of Guelph, Guelph, Ontario, Canada N1G 2W1*

Received 15 March 1991/Accepted 3 July 1991

**Pasteurella haemolytica serotype A1 secretes a glycoprotease which is specific for O-sialoglycoproteins such as glycophorin A. The gene encoding the glycoprotease enzyme has been cloned in the recombinant plasmid pPH1, and its nucleotide sequence has been determined. The gene (designated *gcp*) codes for a protein of 35.2 kDa, and an active enzyme protein of this molecular mass can be observed in *Escherichia coli* clones carrying pPH1. In vivo labeling of plasmid-encoded proteins in *E. coli* maxicells demonstrated the expression of a 35-kDa protein from pPH1. The amino-terminal sequence of the heterologously expressed protein corresponds to that predicted from the nucleotide sequence. The glycoprotease is a neutral metalloprotease, and the predicted amino acid sequence of the glycoprotease contains a putative zinc-binding site. The gene shows no significant homology with the genes for other proteases of procaryotic or eucaryotic origin. However, there is substantial homology between *gcp* and an *E. coli* gene, *orfX*, whose product is believed to function in the regulation of macromolecule biosynthesis.**

*Pasteurella haemolytica* is the principal microorganism associated with bovine pneumonic pasteurellosis, a major cause of sickness and death in feedlot cattle in North America (15, 32). *P. haemolytica* has been divided into 16 serotypes based on soluble or extractable surface antigens (3). Among the 16 serotypes, serotype A1 is the predominant microorganism isolated from pneumonic lungs (27, 32). *P. haemolytica* A1 produces a number of antigens which are secreted into the culture supernatant during its growth. Some of these antigens include a heat-labile cytotoxin specific for ruminant leukocytes (24), a serotype-specific outer membrane protein (6, 12), a glycoprotease specific for sialoglycoproteins (1, 2, 21), and a neuraminidase (5). Vaccination of calves with a bacterium-free culture supernatant from logarithmic-phase cultures induces resistance to experimental challenge, and a vaccine based on the culture supernatant has been developed (Presponse) (23, 25).

The glycoprotease of *P. haemolytica* A1 is highly specific for O-glycosylated glycoproteins, and proteins which lack extensive O-sialoglycopeptide residues are not cleaved. The best-characterized glycoprotein substrate is glycophorin A from human erythrocytes. It is cleaved either in situ on the surface of erythrocyte plasma membrane or when the glycoprotein is in solution (2). The glycoprotease is a neutral metalloprotease and is nontoxic to cultured mammalian cells, including bovine pulmonary macrophages, bovine endothelial cells, and erythrocytes (2). The role of the glycoprotease in pathogenesis and in the induction of an immune response is unknown. A homogeneous enzyme preparation is difficult to isolate by conventional biochemical techniques. The use of molecular biology techniques to isolate the gene coding for the glycoprotease would facilitate further studies. The cloning and characterization of the genetic determinant encoding the *P. haemolytica* A1 leukotoxin has been described elsewhere (10, 11, 28). The gene encoding a serotype-specific outer membrane protein of *P. haemolytica* A1 has also been isolated (6, 12). Here, we report the cloning, nucleotide sequence, and expression of a gene coding for the *P. haemolytica* A1 glycoprotease.

* Corresponding author.

## MATERIALS AND METHODS

**Bacterial strains, plasmids, and culture conditions.** *Escherichia coli* HB101, TG-1, and CSR603 and *P. haemolytica* A1 have been described previously (9, 10, 22). The preparation of a clone bank which contains the *P. haemolytica* A1 genomic DNA carried in the vector pBR322 has also been described previously (10). The *E. coli* recombinant clones which encode soluble antigens of *P. haemolytica* A1 were isolated by using an antiserum directed against *P. haemolytica* A1 soluble antigens in a colony immunoblot assay (9). The M13 phage vectors mp18 and mp19 and the expression vectors pTTQ18 and pTTQ19 were from Pharmacia Chemicals Inc. (Dorval, Quebec, Quebec, Canada). The *E. coli* HB101 clones were cultured in LT medium supplemented with ampicillin to 100 μg/ml (LT + A) (9). *P. haemolytica* A1 cultures were grown in brain heart infusion broth. *E. coli* TG-1 was grown on Davis minimal medium (17).

**Enzymes and chemicals.** All restriction endonucleases and DNA-modifying enzymes were from Bethesda Research Laboratory (Burlington, Ontario, Canada) or Pharmacia Chemicals and were used as recommended by the supplier. [α-$^{32}$P]dATP (3,000 Ci/mmol) and Tran$^{35}$S-Label (1,130 Ci/mmol) were from ICN Biochemicals (St. Laurent, Quebec, Quebec, Canada). Immobilon polyvinylidene difluoride (PVDF) membranes were from Millipore (Mississauga, Ontario, Canada).

**Screening for glycoprotease activity and enzyme assay.** The *E. coli* recombinant clones were grown in LT + A broth overnight, subcultured (1/100), and grown for 4 h at 37°C to logarithmic phase. The cells were harvested by centrifugation (5,000 × *g*), washed in 50 mM *N*-2-hydroxyethylpiperazine-N′-ethanesulfonic acid (HEPES) buffer (pH 7.4), and lysed by passing them through a French press three times at a pressure of 17,000 lb/in$^2$. The lysates were centrifuged at 1,085 × *g* for 5 min to remove cellular debris, and the supernatants were assayed for glycoprotease activity. Su-

pernatant (0.1 ml) was incubated with 5 μg of $^{125}$I-glycophorin A (20 μCi/mg), prepared as described elsewhere (13, 14), in 0.1 ml of 50 mM HEPES (pH 7.4) for 16 h at 37°C. The unhydrolysed substrate and the products of hydrolysis were separated by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) (8) and detected by autoradiography using Cronex 4 X-ray film (Dupont, Wilmington, Del.). The enzyme activity of the glycoprotease was calculated from the percent disappearance of glycophorin A bands on an autoradiogram, measured by using a densitometer or by slicing the gel and counting the bands in a gamma counter (Beckman Instruments, Gamma 5500).

**In vivo labeling of cloned glycoprotease.** The recombinant plasmids were transformed in *E. coli* CSR603, and maxicells were prepared as described previously (22). Briefly, *E. coli* CSR603 carrying the recombinant plasmids was grown to mid-logarithmic phase at 37°C in Davis minimal medium supplemented with ampicillin (100 μg/ml) and 0.5% Casamino Acids. Ten milliliters of the culture was irradiated for 15 s at 400 μW/cm$^2$ with a germicidal lamp (General Electric). After 2 h of culture, 100 μl of D-cycloserine (2 mg/ml) was added, and the culture was grown overnight. Approximately 3 ml of the culture was centrifuged, and the washed pellet was suspended in 0.75 ml of Davis minimal medium supplemented with threonine (100 μg/ml), arginine (150 μg/ml), leucine (150 μg/ml), and proline (100 μg/ml). The cell suspension was incubated for 1 h at 37°C, after which 25 μCi of [$^{35}$S]methionine was added. For recombinants in the expression vectors pTTQ18 or pTTQ19, isopropyl-β-D-thiogalactoside (IPTG; 0.5 mM, final concentration) was included for induction of the *tac* promoter. After being labeled for 1 h, the cells were harvested in a microfuge and lysed by suspension in 150 μl of SDS-sample buffer (8). The labeled proteins were separated by SDS-PAGE (8) and identified by direct autoradiography of the dried gel.

**Subcloning and DNA sequencing.** The 3.3-kbp *Bam*HI-*Bgl*II fragment from pPH1 was subcloned into the *Bam*HI site of the expression vector pTTQ19. Selective fragments were removed from the insert DNA by digestion and religation with an endonuclease having one restriction site located on the insert DNA and one on the multicloning site in pTTQ19. A series of subclones were created, each containing only selected fragments of the original insert DNA from pPH1. Each subclone was then analyzed for the expression of the glycoprotease in the *E. coli* maxicell system. A 2.3-kbp *Bam*HI-*Hin*dIII fragment was found to contain the glycoprotease gene. This fragment was subcloned into the M13 phage vectors mp18 and mp19, which were propagated in *E. coli* TG-1. The nucleotide sequence of the insert DNA was determined by the dideoxy chain termination method as described previously (11). The method of Dale (4) was used to generate overlapping deletions of the insert DNA in either M13mp18 or M13mp19 vectors for sequencing of the entire insert DNA. The DNA sequence was analyzed by using the Pustell Sequence Analysis program (International Biotechnology Inc., Toronto, Ontario, Canada). The coding sequence was screened against the GenBank data base to search for sequence homology. The predicted amino acid sequence was also analyzed for its hydrophobicity and membrane-spanning regions by the SOAP program of PC Gene (Intelligenetics, Mountain View, Calif.).

**Expression of the glycoprotease gene.** The 2.3-kbp *Bam*HI-*Hin*dIII fragment from pPH1 was subcloned into expression vector pTTQ18 to form the plasmid pGP1. Expression was carried out by transformation of *E. coli* HB101 with pGP1. Cultures were grown overnight at 37°C in LT + A broth. The cultures were then subcultured (1/50) into LT + A broth and grown for 2 h. The cells were then harvested by centrifugation (5,000 × *g*), and the pellet was suspended in warm LT medium (without glucose) supplemented with ampicillin and IPTG (0.5 mM, final concentration) and grown for 90 min at 37°C. The culture was centrifuged (5,000 × *g*), and the pellet was suspended in 0.1 volume of 2× SDS-sample buffer (8). After being boiled for 5 min, the samples were separated by SDS-PAGE, and the proteins were visualized by staining with Coomassie brilliant blue R250. For the determination of glycoprotease activity, the cells were lysed by a French press and the lysate was assayed for activity as described above.

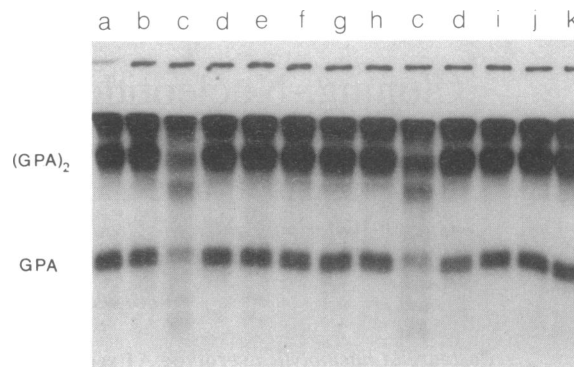


FIG. 1. Glycoprotease activity in *E. coli* recombinant clones. Lanes: a, $^{125}$I-glycophorin A only; b, *E. coli* carrying pBR322; c, *E. coli* carrying pPH1; d, *E. coli* carrying pPH8; e through k are the other *E. coli* clones of the genomic library. $(GPA)_2$, glycophorin A dimer; GPA, glycophorin A monomer.

**Preparation of periplasmic and cellular fractions.** *E. coli* cells carrying the plasmid pGP1 were induced with IPTG as described above and subjected to osmotic shock treatment by the method of Neu and Heppel (20). After osmotic shock, the cell suspension was stirred for 10 min on ice and centrifuged, and the supernatant containing the periplasmic fraction was dialyzed against distilled water and lyophilized. The cell pellet was resuspended in 2× SDS-sample buffer (8) and used as the cellular fraction. Fractions were assayed for cyclic phosphodiesterase and β-galactosidase as markers for periplasmic and cellular proteins, respectively, as described elsewhere (20).

**N-terminal amino acid analysis.** The glycoprotease expressed from pGP1 was recovered after SDS-PAGE and transferred by electroblotting onto a PVDF membrane (31). The PVDF membrane was stained with Coomassie brilliant blue R250 to locate the gene product and then destained, and the appropriate region carrying the glycoprotease was sliced out. The glycoprotease was then subjected to N-terminal amino acid analysis by the automated Edman procedure in a gas-phase peptide microsequencer.

**Nucleotide sequence accession number.** The nucleotide sequence reported here has been deposited in GenBank and given the accession number M62364.

## RESULTS

The lysates from 27 recombinant clones in *E. coli* were assayed for glycoprotease activity. The results in Fig. 1 show an autoradiogram of the hydrolysis of $^{125}$I-glycophorin A incubated with lysate from clone pPH1. A negative control of plasmid pBR322-transformed *E. coli* showed no hydroly-

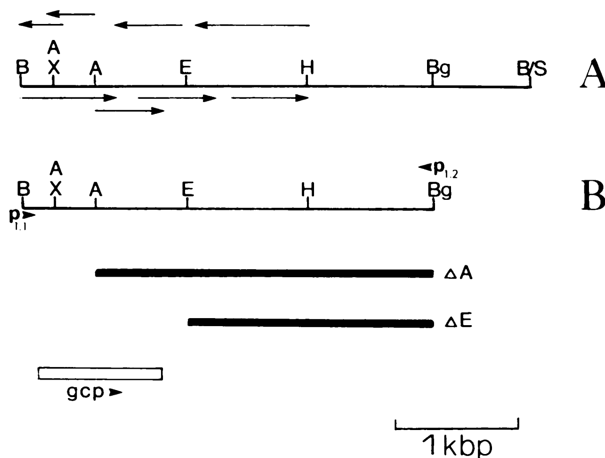


FIG. 2. Restriction maps of plasmids used in this study. (A) Restriction map of insert DNA in pPH1 and pPH8. The arrows indicate the direction and extent of DNA sequenced. (B) Subcloning of 3.3-kbp *Bam*HI-*Bgl*II fragment from pPH1 into pTTQ19 in either orientation. p, direction of expression from the *tac* promoter in pPH1.1 and pPH1.2; dark bars with ΔA and ΔE, internal fragments deleted in the plasmids pPH1.1A and pPH1.1E, respectively; open bar, open reading frame coding for *gcp* as deduced from the DNA sequence. Abbreviations: A, *Ava*I; B, *Bam*HI; Bg, *Bgl*II; B/S, *Bam*HI-*Sau*3A junction; E, *Eco*RI; H, *Hin*dIII; X, *Xba*I.

sis, and a positive control of *P. haemolytica* A1 culture supernatant showed complete hydrolysis (see Fig. 8). Prolonged autoradiography of the gel also showed weak hydrolysis of $^{125}$I-glycophorin A by the lysate from another clone, pPH8. Restriction endonuclease analysis of the recombinant plasmids pPH1 and pPH8 showed that the two plasmids contain identical inserts (Fig. 2), and pPH1 was chosen for further studies.

The plasmid pPH1 was transformed into *E. coli* CSR603, and the plasmid-encoded proteins were examined by in vivo $^{35}$S labeling. The results show that an additional protein of about 35 kDa was expressed from pPH1 (Fig. 3). This gives

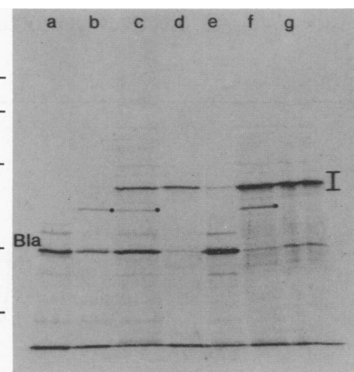


FIG. 3. Maxicell labeling of the plasmid-encoded proteins. Autoradiogram of 15% SDS-PAGE containing [$^{35}$S]methionine-labeled proteins expressed in *E. coli* carrying pBR322 (lane a), pPH1 (lane b), pPH1.1 (lane c), pPH1.2 (lane d), pPH1.1A (lane e), pPH1.1E (lane f), and pTTQ19 (lane g). The β-lactamase of pBR322 and the Lac repressor protein of pTTQ19 are indicated by Bla and I, respectively. Molecular size standards, starting from the top, represent 92.5, 66, 43, 31, and 21 kDa. The solid circles in the gel show the position of the 35-kDa protein corresponding to the glycoprotease.

an estimate of the size of the glycoprotease and is consistent with the estimated size of the enzyme in SDS-PAGE analysis of extracts of *P. haemolytica* A1 culture supernatant (2). To locate the coding region of the glycoprotease gene on pPH1, the 3.3-kbp *Bam*HI-*Bgl*II fragment was subcloned into expression vector pTTQ19 to produce subclones pPH1.1 and pPH1.2, which carry the insert DNA in opposite orientations (Fig. 2). The plasmid-encoded proteins from pPH1.1 and pPH1.2 were examined in the *E. coli* maxicell system, and the results are shown in Fig. 3. Only plasmid pPH1.1 expressed a plasmid-encoded protein identical in size to the cloned protein expressed from pPH1, suggesting that this plasmid encodes the correct orientation for expression. Two subclones were constructed from pPH1.1 by deleting selective internal fragments to yield the constructs pPH1.1A and pPH1.1E (Fig. 2). Subclone pPH1.1A (constructed by partial digestion of pPH1.1 with *Ava*I and religation) did not express the 35-kDa protein, whereas pPH1.1E still expressed a protein similar to that observed in pPH1 and pPH1.1 (Fig. 3). This shows that the *Ava*I site is within the glycoprotease gene and that the *Eco*RI site may be close to the end of the gene. On the basis of the size of the glycoprotease expressed in the maxicell analysis, a DNA fragment of about 1 kbp is required to encode the 35-kDa protein. This gene is therefore located within the 2.3-kbp *Bam*HI-*Hin*dIII fragment in the orientation depicted in pPH1.1.

The complete nucleotide sequence of the 2.3-kbp *Bam*HI-*Hin*dIII fragment was determined, and the 1.3-kbp sequence between the *Bam*HI and *Eco*RI sites is presented in Fig. 4. Both strands of the DNA were sequenced, either directly or by the use of overlapping deletions of the cloned DNA, in the phage vector M13mp18 or M13mp19. Most regions were sequenced at least three times independently. Analysis of the DNA sequence revealed one large open reading frame expressing in the direction anticipated from the maxicell labeling experiments (Fig. 2). The open reading frame covers 975 nucleotides and encodes 325 amino acids with a total predicted molecular weight of 35.2 kDa. These estimates are in agreement with the size of the protein expressed in the maxicell labeling experiments. The gene encoding the glycoprotease was designated *gcp*.

The glycoprotease was expressed from plasmid pGP1 (see below), separated by SDS-PAGE, and electroblotted onto PVDF membrane. The membrane region containing the glycoprotease was excised, and the N-terminal amino acid sequence was determined. The results from the first eight cycles are identical to the first eight amino acids predicted from the nucleotide sequence of *gcp* and confirm the assignment of the reading frame for the glycoprotease.

An examination of the nucleotide sequence upstream from *gcp* failed to reveal features similar to the promoter commonly found in *E. coli* (Fig. 4). Neither the consensus promoter sequence TATAAT nor the consensus RNA polymerase-binding site TTGACA was apparent, although a weak promoter could be observed at position −42 (TCGCC..TATTTT). Further, there is no apparent putative ribosome-binding site immediately preceding the ATG initiation codon of *gpc* (Fig. 4). It is possible that the *gcp* promoter is not readily recognized in *E. coli*, and this may explain the poor expression of glycoprotease activity in the initial clones pPH1 and pPH8. On the other hand, downstream from the termination codon of *gcp*, an mRNA structure consisting of a 14-bp stem-and-loop structure similar to the rho-independent transcriptional signal of *E. coli* could be identified (Fig. 4). In the two other loci sequenced from *P. haemolytica* A1, sequences similar to those of the *E. coli*

```
                                 -140      -130      -120      -110
                                 GGATCCAAGA ATATGAAAGC AAAGAGCTAC CGAATCCTGA
                                 B
-100       -90       -80       -70       -60       -50       -40
AAAACTGAAG TATGGCGAAC AATTCTAGTC GTACAGAGAA TAATGTGAGG GGCGTTCTTC GCCCCTTTTG

-30       -20       -10       1         10          20          30
GTTTTCTAAC TTATTTTGAC TTCTCCAACT ATG CGA ATT TTA GGT ATT GAA ACC TCT TGT GAT
                                 Met Arg Ile Leu Gly Ile Glu Thr Ser Cys Asp

    40          50          60          70          80          90
GAA ACC GGT GTT GCC ATT TAT GAT GAA GAC AAA GGC TTA GTG GCA AAC CAG CTT TAT
Glu Thr Gly Val Ala Ile Tyr Asp Glu Asp Lys Gly Leu Val Ala Asn Gln Leu Tyr

      100         110         120         130         140
AGC CAA ATT GAT ATG CAC GCC GAT TAC GGT GGC GTA GTC CCT GAA CTG GCT TCT CGA
Ser Gln Ile Asp Met His Ala Asp Tyr Gly Gly Val Val Pro Glu Leu Ala Ser Arg

150         160         170         180         190         200
GAC CAT ATC CGT AAA ACG TTG CCA CTA ATT CAA GAA GCC TTA AAA GAG GCC AAT CTG
Asp His Ile Arg Lys Thr Leu Pro Leu Ile Gln Glu Ala Leu Lys Glu Ala Asn Leu

  210         220         230         240         250         260
CAA CCC TCG GAT ATT GAC GGC ATT GCC TAT ACT GCC GGC CCA GGC TTG GTC GGG GCT
Gln Pro Ser Asp Ile Asp Gly Ile Ala Tyr Thr Ala Gly Pro Gly Leu Val Gly Ala

      270         280         290         300         310
TTA TTG GTC GGC TCA ACC ATT GCC CGT TCG CTG GCT TAT GCT TGG AAT GTT CCG GCA
Leu Leu Val Gly Ser Thr Ile Ala Arg Ser Leu Ala Tyr Ala Trp Asn Val Pro Ala

320         330         340         350         360         370
TTG GGC GTT CAC CAT ATG GAA GGG CAT TTA CTT GCC CCA ATG TTG GAA GAA AAT GCC
Leu Gly Val His His Met Glu Gly His Leu Leu Ala Pro Met Leu Glu Glu Asn Ala

  380         390         400         410         420         430
CCT GAA TTT CCG TTT GTG GCA TTA TTG ATT TCA GGT GGA CAC ACC CAA CTG GTA AAA
Pro Glu Phe Pro Phe Val Ala Leu Leu Ile Ser Gly Gly His Thr Gln Leu Val Lys

    440         450         460         470         480
GTT GAC GGC GTT GGG CAA TAC GAA CTA CTC GGG GAA TCA ATT GAT GAT GCT GCC GGT
Val Asp Gly Val Gly Gln Tyr Glu Leu Leu Gly Glu Ser Ile Asp Asp Ala Ala Gly

490         500         510         520         530         540
GAA GCC TTT GAC AAA ACA GGC AAA CTA CTC GGT TTG GAT TAC CCT GCC GGT GTA GCG
Glu Ala Phe Asp Lys Thr Gly Lys Leu Leu Gly Leu Asp Tyr Pro Ala Gly Val Ala

  550         560         570         580         590         600
ATG TCA AAA TTA GCC GAA TCC GGC ACG CCA AAT CGT TTT AAA TTC CCT CGT CCA ATG
Met Ser Lys Leu Ala Glu Ser Gly Thr Pro Asn Arg Phe Lys Phe Pro Arg Pro Met

      610         620         630         640         650         660
ACC GAC AGA CCG GGA CTG GAT TTC AGT TTC TCC GGT TTA AAA ACC TTT GCT GCG AAT
Thr Asp Arg Pro Gly Leu Asp Phe Ser Phe Ser Gly Leu Lys Thr Phe Ala Ala Asn

        670         680         690         700         710
ACG ATT AAA GCC AAT CTT AAT GAA AAT GGT GAA CTC GAT GAG CAA ACC AAA TGC GAT
Thr Ile Lys Ala Asn Leu Asn Glu Asn Gly Glu Leu Asp Glu Gln Thr Lys Cys Asp

720         730         740         750         760         770
ATT GCC CAC GCA TTC CAA CAA GCC GTG GTT GAT ACT ATT TTA ATT AAA TGC AAG CGA
Ile Ala His Ala Phe Gln Gln Ala Val Val Asp Thr Ile Leu Ile Lys Cys Lys Arg

    780         790         800         810         820         830
GCG TTA GAG CAA ACC GGC TAT AAA CGC TTA GTA ATG GCA GGC GGC GTA AGT GCC AAT
Ala Leu Glu Gln Thr Gly Tyr Lys Arg Leu Val Met Ala Gly Gly Val Ser Ala Asn

        840         850         860         870         880
AAA CAA TTA CGA GCA GAC CTT GCG GAA ATG ATG AAA AAA TTA AAA GGC GAA GTA TTC
Lys Gln Leu Arg Ala Asp Leu Ala Glu Met Met Lys Lys Leu Lys Gly Glu Val Phe

890         900         910         920         930         940
TAC CCT CGC CCA CAA TTT TGC ACT GAC AAC GGC GCA ATG ATT GCC TAC ACT GGC TTT
Tyr Pro Arg Pro Gln Phe Cys Thr Asp Asn Gly Ala Met Ile Ala Tyr Thr Gly Phe

  950         960         970         980         990       1000
CTT CGC TTA AAA ACG ATG AAC AAA CCG ACT TAA GC ATTAGCGTAA ACCCCGCTGG
Leu Arg Leu Lys Thr Met Asn Lys Pro Thr ---

      1010       1020       1030       1040       1050       1060       1070
CTATGACCGA ATTACCACCG ATTAATTAAC CTTTCAAGCG GTGAAATTTC TTGTTAATTT TGCAAAAATT

      1080       1090       1100       1110       1120       1130       1140
TAATCAAAAA TAACCGCTTG CTATATGATA GATTAAATTT ATGAATAATT ATGTAATTAG CCTACCTCCG

      1150       1160       1170
CACAGGAGCG TAGAAAACAT ATTCAAGCTG AATTC
                                 E
```

FIG. 4. Nucleotide sequence of insert DNA between the *Bam*HI and *Eco*RI sites in pPH1.1. The numbers above each line refer to nucleotide positions, which are arbitrarily numbered from −140 at the *Bam*HI site to 1170 at the *Eco*RI site. The predicted amino acid sequence for the glycoprotease is shown beneath the DNA sequence. The inverted repeats downstream from the *gcp* translational stop codon are underlined with arrows. Abbreviations: B, *Bam*HI; E, *Eco*RI.

promoters were identified (11, 12, 28, 29). It is possible that different types of promoters under different regulatory systems are utilized in *P. haemolytica* A1.

The predicted amino acid sequence of the glycoprotease was analyzed for its hydrophobicity and potential membrane spanning regions. Figure 5 shows a hydropathy plot of the glycoprotease analyzed by the SOAP program (7). The analysis classified the glycoprotease as a potential peripheral membrane protein but not as an integral membrane protein with transmembrane domains. This is consistent with the glycoprotease being found among the secreted products of *P. haemolytica* A1.

A search for the nucleotide sequence of *gcp* in data banks such as Genbank showed extensive homology with a DNA region upstream from the *E. coli rpsU-dnaG-rpoD* macromolecular synthesis operon. In particular, *gcp* is almost identical to an identified gene, designated *orfX*, in that region (19). A comparison of the predicted amino acids for the glycoprotease and the putative OrfX protein is shown in Fig. 6. Almost 76% of the amino acids of glycoprotease are identical to those of OrfX, suggesting that the two proteins probably have similar functions. On the other hand, the codons used in the two genes and their flanking nucleotide sequences are very dissimilar. A detailed comparison of the codon usage of the two genes from *E. coli* and *P. haemolytica* A1 will be presented elsewhere. Little is known about the function of the protein encoded by *orfX* except that it may be involved in regulation of expression of the *rpsU-dnaG-rpoD* operon (19). However, on the basis of current evidence, it would not be surprising to discover a proteolytic activity

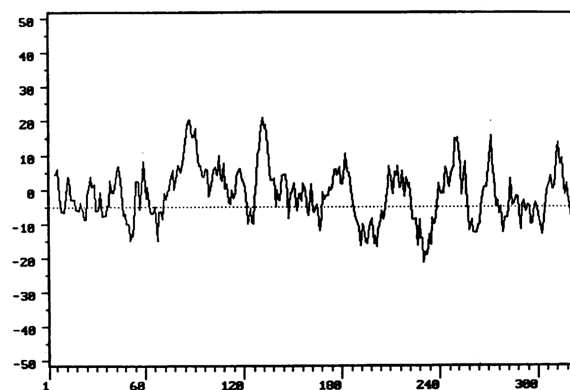


FIG. 5. Hydropathy profile of the glycoprotease protein. The method used was that of Klein et al. (7). The vertical axis gives the scale of the hydrophobic (positive) and hydrophilic (negative) values established for each window of nine residues. The horizontal axis gives the scale for the amino acids in protein.

```
              10         20         30         40         50
OrfX MRVLGIETSC DETGIAIYDD EKGLLANQLY SQVKLHADYG GVVPELASRD
Gcp  MRILGIETSC DETGVAIYDE DKGLVANQLY SQIDMHADYG GVVPELASRD

              60         70         80         90        100
     HVRKTVPLIQ AALKESGLTA KDIDAVAYTA GPGLVGALLV GATVGRSLAF
     HIRKTLPLIQ EALKEANLQP SDIDGIAYTA GPGLVGALLV GSTIARSLAY

             110        120        130        140        150
     AWDVPAIPVH HMEGHLLAPM LEDNPPEFPF VALLVSGGHT QLISVTGIGQ
     AWNVPALGVH HMEGHLLAPM LEENAPEFPF VALLISGGHT QLVKVDGVGQ

             160        170        180        190        200
     YELLGESIDD AAGEAFDKTA KLLGLDYPGG PLLSKMAAQG TAGRFVFPRP
     YELLGESIDD AAGEAFDKTG KLLGLDYPAG VAMSKLAESG TPNRFKFPRP

             210        220        230        240        250
     MTDRPGLDFS FSGLKTFAAN TIRDN----G -TDDQTRADI ARAFEDAVVD
     MTDRPGLDFS FSGLKTFAAN TIKANLNENG ELDEQTKCDI AHAFQQAVVD

             260        270        280        290        300
     TLMIKCKRAL DQTGFKRLVM AGGVSANRTL RAKLAEMMKK RRGEVFYARP
     TILIKCKRAL EQTGYKRLVM AGGVSANKQL RADLAEMMKK LKGEVFYPRP

             310        320        330        340
     EFCTDNGAMI AYAGMVRFKA GATADLGVSV RPRWPLAELP AA
     QFCTDNGAMI AYTGFLRLKT MNKPT
```

FIG. 6. Homology between the glycoprotease and OrfX. The segmented vertical lines show the locations of identical residues. The segmented horizontal lines represent breaks introduced to maximize homology. Potential zinc ligand of the glycoprotease and OrfX at His-111 and His-115 is underlined.

FIG. 7. High-level expression and secretion of the glycoprotease. (A) Total cellular proteins expressed from cultures of *E. coli* HB101 carrying pTTQ18 (lane 1), uninduced pGP1 (lane 2), pTTQ18 (lane 3), and pGP1 induced with IPTG (lane 4) separated by 15% SDS-PAGE and stained with Coomassie blue. The glycoprotease protein is indicated by an arrowhead. (B) Periplasmic fractions after osmotic shock treatments to *E. coli* HB101 carrying induced pGP1. Lane 1, Cytoplasmic fraction; lane 2, periplasmic fraction. Molecular size standards are in kilodaltons.

associated with the *orfX* protein, which may be part of the regulation of macromolecule biosynthesis.

The 2.3-kbp *Bam*HI-*Hin*dIII fragment was then subcloned into the *Bam*HI and *Hin*dIII site of the high-expression vector pTTQ18 to produce plasmid pGP1. This placed the initiation codon of *gcp* at an appropriate distance from the *tac* promoter to allow high-level expression of the glycoprotease. Figure 7A shows expression of the *P. haemolytica* A1 glycoprotease in *E. coli* carrying pGP1 upon induction with IPTG. No glycoprotease could be detected in *E. coli* carrying only pTTQ18 or in uninduced cultures (Fig. 7A). The same protein preparations were also assayed for glycoprotease activity, and the results in Fig. 8 showed that *E. coli* carrying pGP1 expressed glycoprotease activity similar to the enzyme activity present in culture supernatant of *P. haemolytica* A1. The overexpressed 35-kDa protein isolated by SDS-PAGE from *E. coli* lysates containing the glycoprotease activity was used to raise polyclonal antibody in rabbits. This antiserum was used to neutralize the glycoprotease activity in culture supernatants of *P. haemolytica* A1. Under conditions in which the *P. haemolytica* glycoprotease degraded 16.8 μg of glycophorin A in 30 min, when rabbit anti-glycoprotease antiserum was added at titers of 1/4, 1/8, 1/16, and 1/32, degradation of glycophorin A dropped to 4.0, 8.9, 8.4, and 11.5 μg of glycophorin A per 30 min, respectively. Control antisera against 35-kDa bands from lysates of plasmid-transformed *E. coli* which lacked the pGP1 gene showed no inhibition of the *P. haemolytica* glycoprotease.

The glycoprotease of *P. haemolytica* A1 is normally secreted in the culture medium, but its mechanism of secretion is not known. After IPTG induction of *E. coli* carrying pGP1, the cells were subjected to osmotic shock treatment to determine the location of the glycoprotease. Figure 7B shows localization of the glycoprotease in the periplasmic fraction of the *E. coli* clone; only traces of the glycoprotease protein were detected in the cellular fraction. An examination of the DNA-derived N-terminal amino acid sequence of the glycoprotease showed the absence of a conventional signal peptide sequence. Furthermore, N-terminal amino acid sequence analysis of the glycoprotease expressed in *E. coli* showed that there is no cleavage of amino acids from the N-terminal region during export of the glycoprotease to the periplasm in *E. coli*. This suggests that there may be a different mechanism involved in the secretion of the glycoprotease.

## DISCUSSION

The cloning and sequencing data show that the *P. haemolytica* A1 glycoprotease is a protein of 35.2 kDa. The enzyme has a predicted pI of 5.2, which is consistent with the finding that the enzyme activity is precipitated from culture supernatants by lowering the pH to 4.0 (2). In the initial stages of this work, only a low level of glycoprotease activity was expressed in lysates of *E. coli* transformed by pPH1 (or pPH8). This may be due to the inefficient activity of the *P. haemolytica* A1 *gcp* promoter in *E. coli*, to incomplete modification of the enzyme, or to instability of the enzyme in *E. coli*. Upon subcloning of the appropriate DNA fragment into pTTQ18, the level of glycoprotease activity in *E. coli* was increased, but the specific activity of the enzyme preparation was less than that observed in serum-free culture supernatants of *P. haemolytica* A1. The

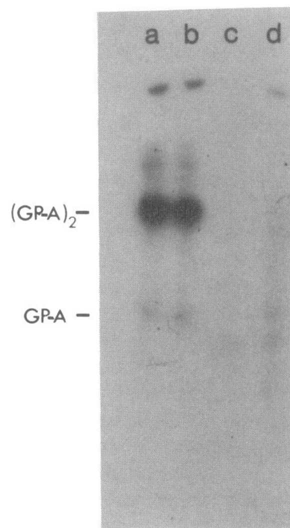


FIG. 8. Enzyme activity of the highly expressed glycoprotease in *E. coli* HB101. Lane a, $^{125}$I-glycophorin A only; lane b, negative control, protein extract from *E. coli* carrying pTTQ18; lane c, positive control, concentrated culture supernatant of *P. haemolytica* A1; lane d, protein extract from *E. coli* carrying pGP1 after IPTG induction. $(GPA)_2$, glycophorin A dimer; GPA, glycophorin A monomer.

lower specific activity of the recombinant product may be due to a lack of posttranslational processing of the cloned glycoprotease in *E. coli*.

Leukotoxin, for which the nucleotide sequence and the regulation of leukotoxin expression have been described elsewhere (11, 28, 29), is another secreted protein of *P. haemolytica* A1. The leukotoxin determinant is composed of four contiguous genes, *lktCABD*; the expression of *lktC* functions in the activation of leukotoxin (LktA), while proteins encoded by *lktB* and *lktD* are involved in the secretion of leukotoxin. It is not known whether the glycoprotease requires a similar activation mechanism, which might explain the lower specific activity of the enzyme expressed in *E. coli*. An examination of the amino terminus of pGP1 shows no conventional signal sequence such as those seen for a number of other secreted proteins characterized in bacteria (16, 26, 30). Since the glycoprotease is normally secreted from *P. haemolytica* A1, an alternative secretory mechanism not involving an amino terminus signal may be utilized, as has been reported for leukotoxin (28). Examination of the DNA sequence immediately downstream from the glycoprotease gene showed an open reading frame encoding a protein of about 9 kDa which might be involved in the secretion of the glycoprotease (data not shown). Experiments are under way to define the mechanism of secretion of the glycoprotease.

The glycoprotease of *P. haemolytica* A1, like other neutral metalloproteases of bacteria, such as thermolysin (18), can be inhibited by metal ion chelators, but there is no major sequence homology with these enzymes except for the presence of a potential zinc-binding site (Fig. 6). The unusual substrate specificity for the glycoprotease, namely, its specificity for O-sialoglycosylated proteins (2), and the lack of homology with other known proteases suggest that it may be a member of a distinct enzyme class.

The extensive homology between the *P. haemolytica* A1 glycoprotease and *E. coli* OrfX is an interesting finding. It has been shown, both by promoter analysis and Northern (RNA) blot analysis, that *orfX* is expressed in *E. coli* (19). Because of its possible role in the regulation of the *rpsU-dnaG-rpsD* macromolecular synthesis operon, OrfX is likely to be an intracellular molecule. On the other hand, the glycoprotease is a secreted enzyme of *P. haemolytica* A1 and could be one of the virulence factors in pathogenesis. It will be of interest to examine any possible evolutionary relationship between the respective genes.

Prior to this study, little was known about the relationship between the glycoprotease of *P. haemolytica* A1 and other bacterial proteases. High-level expression from pGP1 in *E. coli* upon induction with IPTG allows large-scale preparation of the glycoprotease free from other *P. haemolytica* A1 proteins. Such preparations can be used for detailed studies on the activity and immunological properties of the enzyme. Polyclonal antibody was prepared in rabbits against the 35-kDa protein overexpressed in *E. coli* transformed with plasmids carrying pGPQ1, and this antiserum neutralized the *P. haemolytica* glycoprotease activity. This supports the enzymatic evidence that the product of the cloned gene is the *P. haemolytica* glycoprotease. The antibody will be a useful reagent for further characterization of biochemical and biological properties of the glycoprotease. In addition, a DNA probe has been prepared from the *gcp* gene and has been used to screen serotypes of *P. haemolytica* for the gene, in comparison with the occurrence of the active glycoprotease in these strains (1).

## ACKNOWLEDGMENTS

We thank the HSC Biotechnology Centre for protein sequencing and P. E. Shewen, Ontario Veterinary College, for advice and assistance.

This work was supported by NSERC Canada grants to R.Y.C.L. and A.M.

## REFERENCES

1. **Abdullah, K. M., R. Y. C. Lo, and A. Mellors.** 1990. Distribution of glycoprotease activity and the glycoprotease gene among serotypes of *Pasteurella haemolytica*. Biochem. Soc. Trans. **18:**901–903.
2. **Abdullah, K. M., E. A. Udoh, P. E. Shewen, and A. Mellors.** Submitted for publication.
3. **Biberstein, E. L.** 1978. Biotyping and serotyping of *Pasteurella haemolytica*. Methods Microbiol. **10:**253–267.
4. **Dale, R. M. K.** 1985. A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondria 18s rDNA. Plasmid **13:**31–40.
5. **Frank, G. H., and L. B. Tabatabai.** 1981. Neuraminidase activity of *Pasteurella haemolytica* isolates. Infect. Immun. **32:**1119–1122.
6. **Gonzalez-Rayos, C., R. Y. C. Lo, P. E. Shewen, and T. J. Beveridge.** 1986. Cloning of a serotype-specific antigen from *Pasteurella haemolytica* A1. Infect. Immun. **53:**505–510.
7. **Klein, P., M. Kanehisa, and C. DeLisi.** 1985. The detection and classification of membrane-spanning proteins. Biochem. Biophys. Acta **815:**468–476.
8. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.
9. **Lo, R. Y. C., and L. A. Cameron.** 1986. A simple immunological detection method for the direct screening of genes from clone banks. Can. J. Biochem. Cell Biol. **64:**73–76.
10. **Lo, R. Y. C., P. E. Shewen, C. A. Strathdee, and C. N. Greer.** 1985. Cloning and expression of the leukotoxin gene of *Pasteurella haemolytica* A1 in *Escherichia coli* K-12. Infect. Immun. **50:**667–671.
11. **Lo, R. Y. C., C. A. Strathdee, and P. E. Shewen.** 1987. Nucleotide sequence of the leukotoxin genes of *Pasteurella*

*haemolytica* A1. Infect. Immun. **55**:1987–1996.

12. **Lo, R. Y. C., C. A. Strathdee, P. E. Shewen, and B. J. Cooney.** Molecular studies on Ssa1, a serotype-specific antigen of *Pasteurella haemolytica* A1. Infect. Immun., in press.
13. **Marchesi, V. T., and E. P. Andrews.** 1971. Glycoproteins: isolation from cell membranes with lithium diiodosalicylate. Science **174**:1247–1248.
14. **Markwell, M. A. K.** 1982. A new solid-state reagent to iodinate proteins. Anal. Biochem. **125**:427–432.
15. **Martin, S. W., A. H. Meek, D. G. Davis, R. G. Thomson, J. A. Johnson, A. Lopez, L. Stephens, R. A. Curtis, J. F. Prescott, S. Rosendal, M. Savan, A. J. Zubaidy, and M. R. Bolton.** 1980. Factors associated with mortality in feedlot cattle: the Bruce County beef cattle project. Can. J. Comp. Med. **44**:1–10.
16. **Michaelis, S., and J. Beckwith.** 1982. Mechanism of incorporation of cell envelope proteins in *Escherichia coli*. Annu. Rev. Microbiol. **36**:435–465.
17. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **Nakahama, K., K. Yoshimura, R. Marumoto, M. Kikuchi, I. S. Lee, T. Has, and H. Matsubara.** 1986. Cloning and sequencing of *Serratia* protease gene. Nucleic Acids Res. **14**:5843–5855.
19. **Nesin, M., J. R. Lupski, P. Svec, and G. N. Godson.** 1987. Possible new genes as revealed by molecular analysis of a 5-kb *Escherichia coli* chromosomal region 5′ to the *rpsU-dnaG-rpoD* macromolecular-synthesis operon. Gene **51**:149–161.
20. **Neu, H. C., and L. A. Heppel.** 1965. The release of enzymes from *Escherichia coli* by osmotic shock and during the formation of spheroplast. J. Biol. Chem. **240**:3685–3692.
21. **Otulakowski, G. L., P. E. Shewen, E. A. Udoh, A. Mellors, and B. N. Wilkie.** 1983. Proteolysis of sialoglycoprotein by *Pasteurella haemolytica* cytotoxic culture supernatant. Infect. Immun. **42**:64–70.
22. **Sancar, A., A. M. Hack, and W. D. Rupp.** 1979. A simple method for identification of plasmid-coded proteins. J. Bacteriol. **137**:692–693.
23. **Shewen, P. E., A. Sharp, and B. N. Wilkie.** 1988. Efficacy testing a *Pasteurella haemolytica* extract vaccine. Vet. Med. **83**:1078–1083.
24. **Shewen, P. E., and B. N. Wilkie.** 1982. Cytotoxin of *Pasteurella haemolytica* acting on bovine leukocytes. Infect. Immun. **35**:91–94.
25. **Shewen, P. E., and B. N. Wilkie.** 1988. Vaccination of calves with leukotoxic culture supernatant from *Pasteurella haemolytica*. Can. J. Vet. Med. **52**:30–36.
26. **Silhavy, T., S. Benson, and S. Emr.** 1983. Mechanism of protein localization. Microbiol. Rev. **47**:313–344.
27. **Smith, P. C.** 1983. Prevalence of *Pasteurella haemolytica* in transported calves. Am. J. Vet. Res. **44**:981–985.
28. **Strathdee, C. A., and R. Y. C. Lo.** 1989. Cloning, nucleotide sequence, and characterization of genes encoding the secretion function of the *Pasteurella haemolytica* leukotoxin determinant. J. Bacteriol. **171**:916–928.
29. **Strathdee, C. A., and R. Y. C. Lo.** 1989. Regulation of expression of the *Pasteurella haemolytica* A1 leukotoxin determinant. J. Bacteriol. **171**:5955–5962.
30. **Von Heijne, G.** 1983. Patterns of amino acids near signal sequence cleavage sites. Eur. J. Biochem. **133**:17–24.
31. **Walsh, M. J., J. McDougall, and B. Wittmann-Liebold.** 1988. Extended N-terminal sequencing of proteins of archaebacterial ribosomes blotted from two-dimensional gels onto glass fiber and PVDF membranes. Biochemistry **27**:6867–6876.
32. **Yates, W. D. G.** 1982. A review of infectious bovine rhinotracheitis, shipping fever pneumonia and viral-bacterial synergism in respiratory disease of cattle. Can. J. Comp. Med. **46**:225–263.